A Comparison Study of Pricing Credit Default Swap Index Tranches with Convex Combination of Copulae

Ostap Okhrin[a], Yafei Xu[a,b,*]

[a] Chair of Econometrics and Statistics, esp. in Transportation Science, Technische Universität Dresden, Würzburger Str. 35, 01187 Dresden, Germany.
[b] School of Business and Economics, Humboldt-Universität zu Berlin, Spandauer Strasse 1, 10178 Berlin, Germany.

Abstract

Copula as a tool for dependence modeling have been widely utilized in pricing portfolio-like financial derivatives, e.g. Credit Default Swap Index (CDX) tranches. Among the pricing models, the model equipped with the Gaussian copula has become the market benchmark. Albeit thereafter some other copulae were employed to improve Gaussian model, yet a lot of them suffer weak points, especially in destitution of heterogeneous sectoral dependence, asymmetric dependence and fat tail dependence. For increasing the pricing accuracy and also keeping the model parsimonious, we propose in this paper an approach of convex combination of copulae (cc-copula) in pricing CDX tranches. Copulae from elliptical and the Archimedean families are chosen as components to construct the cc-copula models. In the a posteriori study of this paper the proposed cc-copula approach is applied to reproduce the spreads of the CDX NA IG tranches. The results evince that the cc-copula based pricing models have dominant performance compared to the benchmarks.

Keywords: Copula, Convex Combination of Copulae, CDS Index, Credit Risk

1. Introduction

In recent years, the financial innovation has accelerated significantly with introduction of a lot of new types of contracts and derivative products. In credit derivative market new vehicles, for instance, credit default swap index (CDX) has attracted more and more attention. From one perspective, CDX provides credit investors possibility to diversify their credit portfolio's risk in contrary to a single CDS contract. It has a multi-name protection for the credit portfolios by employing a slicing technique termed as *tranche* under a large pool of debtors. From another perspective, the complex mechanism of the CDX contract brings investors challenges in the accurate pricing of the product, where one of the core questions is in modeling of the dependence of random default times.

*Corresponding author
Email addresses: ostap.okhrin@tu-dresden.de (Ostap Okhrin), yafei.xu@hu-berlin.de (Yafei Xu)

Preprint submitted to Elsevier *29th August 2016*

In the studies of the CDX pricing, the cynosure is in the dependence modeling of random default times. Since CDX and CDO (collateralize debt obligation) have analogous pricing philosophy, therefore literature for CDX pricing can be referred to those for CDO pricing. Firstly proposed in Li (1999) and Li (2000), the Gaussian factor copula model in CDO pricing focuses on modeling the multi-name default times with a high dimensional exchangeable Gaussian copula combined with a transformation of the single-name survival function. Although being simple in dependence modeling, there are a lot of drawbacks in the Gaussian copula thoroughly discussed in the literature over the last decade. These drawbacks include the destitution of the heterogeneity of dependence between sectors and the asymmetric tail-dependence. This makes the exchangeable Gaussian copula based pricing not accurate.

In order to overcome drawbacks listed above, various new methods have been proposed. Specifying the defaults dependence structure by choosing new copulae possessing partly or whole features such as the heterogeneity of dependence in different sectors and the asymmetrical tail-dependence. In choosing new copulae, literature is abundant, such as the Student-t copula model in Embrechts et al. (2003); Demarta and McNeil (2005); Schloegl and O'Kane (2005), the double-t copula model in Hull and White (2004), the Clayton copula model in Schönbucher and Schubert (2000); Lindskog and McNeil (2001); Schönbucher (2002), the hierarchical Archimedean copula model in Hofert and Scherer (2011); Hofert (2010); Choros-Tomczyk et al. (2013), just to name a few.

This paper focuses on the CDX pricing approach based on the convex combination of copulae (cc-copula). Within this project we intend to convexly combine different copulae in order to acquire advantageous properties from component copulae. In the cc-copula models different copula families are convexly combined together so that the merits of different copulae can be utilized together for default dependence modeling. The proposed empirical study uses the data set of the CDX NA IG (Credit Swap Index North America Investment Grade) Series 19 tranches managed by Markit. The CDX NA IG Series 19 containing 125 names dispersed in 5 diverse sectors, issued on 20120920 and ends on 20171220. The main purpose of this paper is to employ the cc-copula models in reproduction of the spreads of CDX tranches to achieve higher accuracy of CDX tranche pricing. We calibrate the parameters in the cc-copula models with numerical optimizations, whose objective function is root-mean-square error (RMSE) based on the theoretical spreads and the real market spreads. In the paper overall 4 tranches are evaluated consistent with the Bloomberg quoting convention.

This paper is structured as follows. Section 2 introduces the fundamental of copula. Section 3 discusses the CDX structure and the pricing mechanism. The empirical study in section 4 includes the computation of tranche spread, the parameter calibration and the performance comparison of models. Section 5 concludes.

2. Copula Models

2.1. Basics of Copula

Copula is a function which joints marginal distributions into a multivariate distribution and is in essence a multivariate cumulative distribution function with all marginals being uniformly distributed. To construct a multivariate cumulative distribution function separately by choosing the copula function and the corresponding margins, which is derived from the Sklar's Theorem.

Theorem 1. *Sklar's Theorem, c.f. Sklar (1959)*
Every multivariate cumulative distribution function $H(x_1, \ldots, x_d) = \mathbb{P}[X_1 \leq x_1, \ldots, X_d \leq x_d]$ of a random vector $(X_1, X_2, \ldots, X_d)$ can be expressed in terms of its marginals $F_i(x) = \mathbb{P}[X_i \leq x]$ and a copula C, such that

$$H(x_1, \ldots, x_d) \quad = \quad C\left\{F_1(x_1), \ldots, F_d(x_d)\right\}. \tag{1}$$

If $F_i(\cdot)$ are continuous, then C is unique.

Reader interested in the copula theory is referred to Nelsen (2006) and Joe (2014) and for copula application in finance to Cherubini et al. (2004).

Two elliptical copulae used in this work are Gaussian copula and Student-t copula. The first one is given by,

$$C_{gs}(u_1, \ldots, u_d; G) = \Phi_d\left\{\Phi^{-1}(u_1), \ldots, \Phi^{-1}(u_d); G\right\}, \ u_k \in [0,1], \ k = 1, \ldots, d, \tag{2}$$

where G is a $(d \times d)$ correlation matrix, Φ_d is a d-dimensional standard Gaussian CDF and Φ is a one dimensional standard Gaussian CDF. Gaussian copula is symmetric with zero tail dependence.

Let $\nu \in (1, +\infty)$ be the degree of freedom and $R = (1 - \frac{2}{\nu})\mathrm{Var}(X)$ the $(d \times d)$ correlation matrix, $X = (X_1, \ldots, X_d)^\top \in \mathbb{R}^d$. The Student-$t$ copula can be represented as follows,

$$\begin{aligned} C_t(u_1, \ldots, u_d; \nu, \mu, R) &= \int_{-\infty}^{t^{-1}(u_1)} \cdots \int_{-\infty}^{t^{-1}(u_d)} \frac{\Gamma(\frac{\nu+d}{2})}{\Gamma(\frac{\nu}{2})\sqrt{(\pi\nu)^d \mid R \mid}} \left\{1 + \frac{(x-\mu)^T R^{-1}(x-\mu)}{\nu}\right\}^{-\frac{\nu+d}{2}} \mathrm{d}x, \\ &= T_d\left\{t^{-1}(u_1; \nu), \ldots, t^{-1}(u_d; \nu); \nu, \mu, R\right\}, \end{aligned} \tag{3}$$

where T_d is a d-dimensional Student-t CDF and t^{-1} is an inverse of a one dimensional Student-t distribution function. Student-t copula has a non-zero tail dependence.

Another important family is the Archimedean copula family, whose elements can be constructed as

$$C_A(u_1, \ldots, u_d; \theta) = \begin{cases} \varphi^{-1}\left\{\varphi(u_1; \theta) + \cdots + \varphi(u_d; \theta); \theta\right\} & \text{if } \sum_{k=1}^d \varphi(u_k; \theta) \leq \varphi(0; \theta), \\ 0 & \text{else}, \end{cases} \tag{4}$$

where a decreasing function $\varphi\colon [0,1] \to [0, +\infty)$ is the generator function with $\varphi(1) = 0$ and $\varphi(+\infty) = 1$. Here four most well-known Archimedean copulae are considered, i.e. Frank, Clayton, Gumbel and Joe. Table

Archimedean Copula	Representation $C(u_1,\dots,u_d;\theta)$	Generator Function $\varphi(t;\theta)$	Parameter θ
Frank	$-\frac{1}{\theta}\log\left[1+\frac{\prod_{k=1}^{d}\{\exp(-\theta u_k)-1\}}{\{\exp(-\theta)-1\}^{d-1}}\right]$	$-\log\left\{\frac{\exp(-\theta t)-1}{\exp(\theta)-1}\right\}$	$(-\infty,+\infty)\backslash\{0\}$
Clayton	$\left(\sum_{k=1}^{d}u_k^{-\theta}-d+1\right)^{-\theta^{-1}}$	$\frac{1}{\theta}(t^{-\theta}-1)$	$[-1/(d-1),\infty)\backslash\{0\}$
Gumbel	$\exp\left\{-\sum_{k=1}^{d}(-\log u_k)^{\theta}\right\}^{\theta^{-1}}$	$\{-\log(t)\}^{\theta}$	$[1,+\infty)$
Joe	$1-\left\{\sum_{k=1}^{d}(1-uk)^{\theta}-\prod_{k=1}^{d}(1-u_k)^{\theta}\right\}^{\frac{1}{\theta}}$	$-\log\left\{1-(1-t)^{\theta}\right\}$	$[1,+\infty)$

Table 1: Structures of common Archimedean copulae.

1 lists the representations, generator functions and parameter spaces of these four common Archimedean copulae.

Frank copula is the only elliptically contoured Archimedean copula owning no tail dependence. Clayton copula has lower tail dependence but no upper tail dependence and this is important for modeling losses. Gumbel copula is the only extreme value copula, and often used in modeling gains. Joe copula, having upper tail dependence. Therefore those copulae with such a specific attribute are used in our study.

As is given before, a simple multivariate Archimedean copula has two weak points. Firstly, it typically uses a single parameter of the generator function $\varphi(\cdot)$ to specify the dependence structure. Secondly, Archimedean copula implies that the distribution of $(U_1,\dots,U_d)^\top$ is the same as that of $(U_{i_1},\dots,U_{i_d})^\top$ for all $i_l \neq i_h$, $l,h \in \{1,\dots,d\}$, which is not common in the practice. A much more flexible model is the hierarchical Archimedean copula (HAC), $C(u_1,\dots,u_d;\theta,s)$, where s stands for the HAC's structure, and θ is the set of copula parameters. Details of HAC can be referred to Savu and Trede (2010), Okhrin et al. (2013) and Okhrin and Ristig (2014). A special case of HAC, the d-dimensional fully nested HAC, is shown as follows,

$$\begin{aligned} C_{fHAC}(u_1,\dots,u_d) &= C[C[\dots C\{C(u_1,u_2;\varphi_1),u_3;\varphi_2\},\cdots,u_{d-1};\varphi_{d-2}],u_d;\varphi_{d-1}] \\ &= \varphi_{d-1}[\varphi_{d-1}^{-1}[\varphi_{d-2}[\dots[\varphi_2^{-1}[\varphi_1\{\varphi_1^{-1}(u_1)+\varphi_1^{-1}(u_2)\}]+\varphi_2^{-1}(u_3)] \\ &\quad +\cdots+\varphi_{d-2}^{-1}(u_{d-1})]]+\varphi_{d-1}^{-1}(u_d)]. \end{aligned} \tag{5}$$

2.2. Convex Combination of Copulae

It is known that a convex combination of distribution functions is again a distribution function, same holds for copulae, see Joe (1996), thus let

$$C(u_1,\dots,u_d;\theta_1,\dots,\theta_I)=\sum_{i=1}^{I}\lambda_i C_i(u_1,\dots,u_d;\theta_i),\ \sum_{i=1}^{I}\lambda_i=1,\ u_k\in[0,1],\ k=1,\dots,d, \tag{6}$$

where λ_i is the weight parameter of the i-th component copula and I stands for the number of the component copulae in cc-copula. $C_i(u_1,\dots,u_d;\theta_i)$ is the i-th component copula with the parameter θ_i. And

$C(u_1, \ldots, u_d; \theta_1, \ldots, \theta_I)$ can be thought as a complicated but flexible joint distribution composed of known copula functions of $C_i(u_1, \ldots, u_d; \theta_i),\ i = 1, \ldots, I$, hence the convex combined copula $C(u_1, \ldots, u_d; \theta_1, \ldots, \theta_I)$ will inherit features from its component copulae of $C_i(u_1, \ldots, u_d; \theta_i)$, which is practical and reasonable in finance for capturing different joint behaviors such as the heterogeneity of dependence and the asymmetrical tail-dependence.

Example 1. *A cc-copula with Clayton and Joe component copulae such that*

$$C(u_1, u_2; \theta_1, \theta_2) = \lambda C_{Clayton}(u_1, u_2; \theta_1) + (1 - \lambda) C_{Joe}(u_1, u_2; \theta_2). \tag{7}$$

Example 1 gives a cc-copula with Clayton copula and Joe copula as components. Figure 1 illustrates this example. In this copula there are three parameters, i.e. θ_1, θ_2 used for the dependence structure in Clayton copula and Joe copula separately. The third parameter, λ, is used for the combination of the two components, which can control the attributes inheriting from the both component copulae. For instance in Figure 1, both copula structure parameters, θ_1, θ_2, are given as known constants. It is set that $\theta_1 = \theta_2 = 0.7$. And the ten weight parameter are set that $\lambda \in \{0.1, 0.2, \ldots, 0.9, 1.0\}$. It is clear that when λ is small, say 0.1, then the Joe copula owns a large weight in the cc-copula. This implies that the upper triangular panel contain figures with more observations accumulated in the upper tail area. This means that this cc-copula is an upper tail dependence characterized copula. Analogously, when λ is large then, say $\lambda = 0.9$, then the Clayton copula will own larger weight, hence the cc-copula will have the lower tail dependence structure, which can be advocated by the contour plot in the first upper triangle.

Therefore, competing against the classical elliptical copula (zero-tail dependence, see Gaussian copula) and the common Archimdean copula (only upper-tail dependence or lower-tail dependence, see Gumbel copula, Joe copula, Clayton copula), the cc-copula model with its adaptivity and flexibility in inheriting of assets from different component copulae, has a promising application future.

3. Credit Default Swap Index

The CDX is a structured credit derivative which can be used to protect against default of the multi-name credit. The portfolio's default risk is divided into slices using the tranche technique, which slices the risk into different hierarchies with a ranking. The CDX issuer is the protection buyer which pays a fixed premium periodically and receives payment for the contingent loss of the credit portfolio. The CDX investor is the protection seller who receives the premium payments from the CDX issuer and takes responsibility to cover the issuer's contingent loss of the credit portfolio.

The CDX tranche technique uses attachment points and detachment points to define hierarchies of the product, which gives the loss percentages of the credit portfolio. The sliced hierarchy is also termed as the tranche. In CDX NA IG product, four attachment points are $l_q = (0, 0.03, 0.07, 0.15)^\top$, thus the

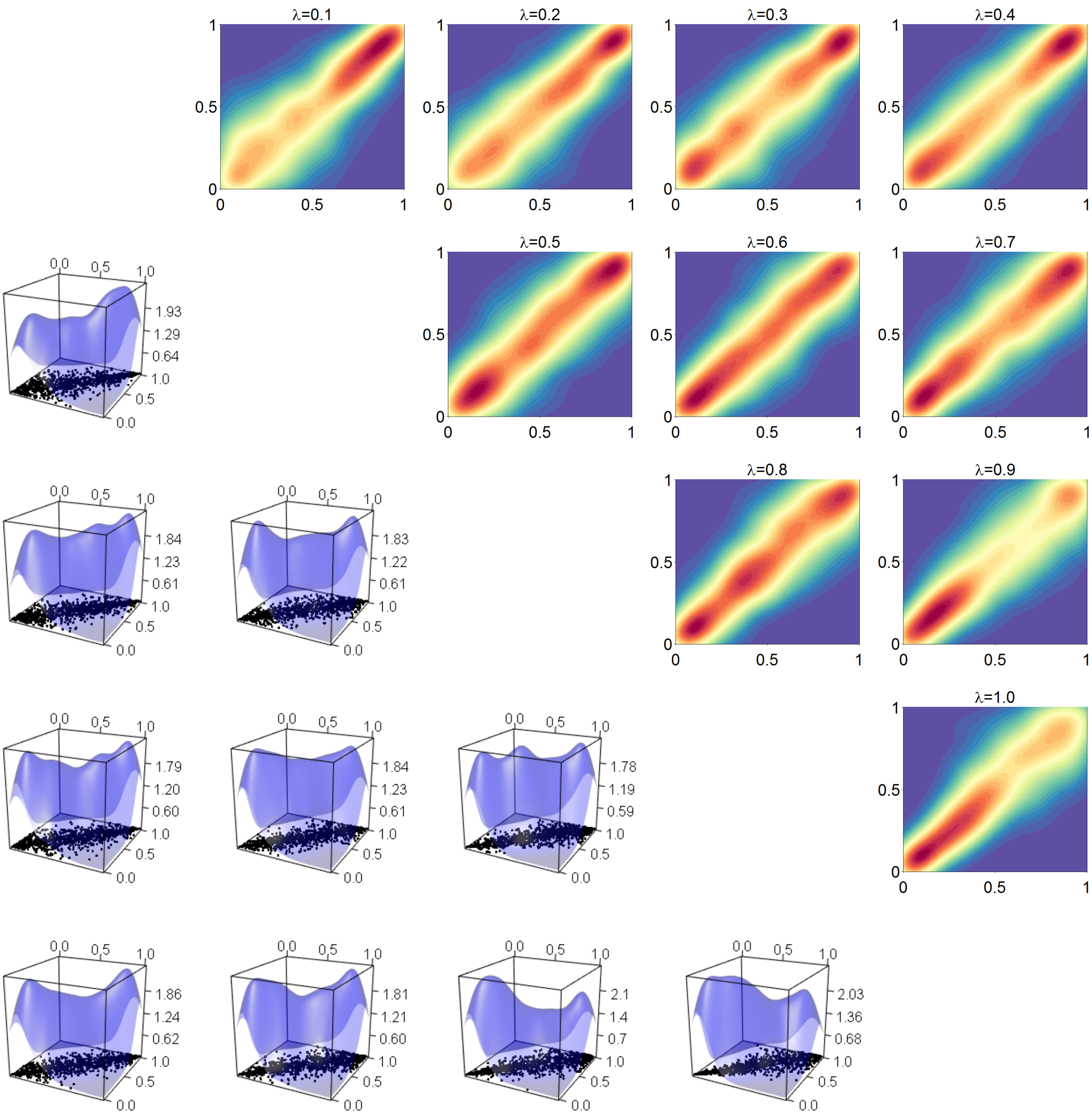

Figure 1: The lower triangular plots illustrate 2-dimensional kernel density estimations containing scatter plots of (U_1, U_2). The scatter points are obtained from 1000 simulations of the cc-copula of Clayton-Joe with Kendall's $\tau = 0.7$ for the both component copulae and $\lambda \in \{0.1, 0.2, \ldots, 0.9, 1\}$, i.e. $C(u_1, u_2; \theta_1, \theta_2) = \lambda C_{Clayton}(u_1, u_2; \theta_1) + (1-\lambda) C_{Joe}(u_1, u_2; \theta_2)$. The upper triangular plots introduce the corresponding contours of the scatter points under ten λs.

corresponding detachment points are $u_q = (0.03, 0.07, 0.15, 1)^\top$. When contingent loss happens between an attachment point and a detachment point of a hierarchy then the notional will be decreased and the periodic payments for the portfolio protection buyer will be reduced either. When contingent loss increases over the detachment point of a hierarchy, then the protection seller pays no premium any more and the protection buyer covers the corresponding losses.

3.1. CDX Pricing

Firstly, let a credit portfolio containing d reference entities with overall N notional principal being equally distributed on entities, i.e. every entity shares $1/d$ of the overall investment. In the meanwhile let the maturity of the CDS index tranches be T, i.e. the length of the contract duration, and premiums are payed at points $t_j,\ j = 1, \ldots, J$ and it is set $t_0 = 0$. In the practice, credit events can occur at any point of the interval $[0, t_J],\ t_J = T$. For simplicity assume the default occur in the midpoint of the two premium payment dates, i.e. $(t_j + t_{j+1})/2$, see Choros-Tomczyk et al. (2013). Then let the random variable $\tau_k, k = 1, \ldots, d$, be the default time of the k-th entity standing for the survival length and r be the constant recovery rate.

The portfolio loss process L_{t_j} is given through

$$L_{t_j} = \frac{1}{d} \sum_{k=1}^{d} (1 - r) \mathbf{1}_{\{\tau_k \le t_j\}},\ j = 0, \ldots, J, \tag{8}$$

where the indicator function $\mathbf{1}_{\{\cdot\}}$ stands for the default indication of the k-th entity. Let $q = 1, \ldots, Q$ be the index of the q-th tranche and L_{q,t_j} the tranche loss of the q-th tranche at t_j. As the tranche loss is a function of the portfolio loss process, the q-th tranche loss is given as follows,

$$L_{q,t_j} = \min\{\max\{L_{t_j} - l_q, 0\}, u_q - l_q\},\ j = 1, \ldots, J,\ q = 1, \ldots, Q. \tag{9}$$

In the run of a CDX tranche, if credit events of underlying entities occur then the premium to be paid in the next period needs to be adjusted according to the outstanding notional P_{q,t_j}

$$P_{q,t_j} = u_q - l_q - L_{q,t_j}. \tag{10}$$

Under the non-arbitrage assumption the expectation of the accumulative payments generated by the protection buyer and seller should be equal. In the CDX pricing study two terminologies for these two expectations are used, the default leg DL_q which represents the expectation of the aggregated compensation payments from the protection seller side, and the premium leg PL_q which stands for the expectation of the aggregated premium payments from the protection buyer side. The default leg DL_q is thus formalized as follows,

$$DL_q \;=\; \mathbb{E}\left\{ \sum_{j=1}^{J} \beta_{t_j} N (L_{q,t_j} - L_{q,t_{j-1}}) \right\},\ q = 1, \ldots, Q, \tag{11}$$

where β_{t_j} is the discount function dependent on the survival length at each payment point.

As in the market practice the protection buyer of a tranche needs to pay upfront payment for every tranche based on the quotation convention of the CDX NA IG Series 19, therefore the premium legs for diverse tranches equal to

$$PL_q = \mathbb{E}\left\{(u_q - l_q)NS_q^{CDX} - \sum_{j=1}^{J} B_q\beta_{t_j}(t_j - t_{j-1})N(P_{q,t_j} + P_{q,t_{j-1}})/2\right\}, \tag{12}$$

where in (12) S_q^{CDX} is the upfront payment rate.

According to the non-arbitrage assumption, the default leg (11) should equals the premium leg (12), what leads to

$$PL_q = DL_q, \tag{13}$$

then plugging (12) and (11) into (13) one obtains

$$\mathbb{E}\left\{(u_q - l_q)NS_q^{CDX} - \sum_{j=1}^{J} B_q\beta_{t_j}(t_j - t_{j-1})N(P_{q,t_j} + P_{q,t_{j-1}})/2\right\} = \mathbb{E}\left\{\sum_{j=1}^{J}\beta_{t_j}N(L_{q,t_j} - L_{q,t_{j-1}})\right\}.$$

Hence the q-th CDS index tranche upfront payment rate S_q^{CDX} can be extracted as follows,

$$(u_q - l_q)S_q^{CDX}N + \mathbb{E}\left\{\sum_{j=1}^{J} 0.5B_q\beta_{t_j}(t_j - t_{j-1})N(P_{q,t_j} + P_{q,t_{j-1}})\right\} = DL_q, \tag{14}$$

therefore the S_q^{CDX} is obtained from (14) as follows,

$$S_q^{CDX} = \mathbb{E}\left[\frac{\sum_{j=1}^{J}\beta_{t_j}\{(L_{q,t_j} - L_{q,t_{j-1}}) - 0.5B_q(t_j - t_{j-1})(P_{q,t_j} + P_{q,t_{j-1}})\}}{u_q - l_q}\right]. \tag{15}$$

3.2. Modeling of Joint Defaults

As mentioned at the beginning, $\tau_k, k = 1, \ldots, d$ is the random variable of survival length (or termed as the default time) of the k-th entity in the reference pool, then let F_k be denoted as the CDF of τ_k and $S_k(t)$ as a survival function. The marginal defaults are assumed to follow homogeneous Poisson process with intensity h, therefore survival times till default has a distribution function of the form

$$F_k(t) = 1 - \exp(-ht). \tag{16}$$

Next the copula function is employed for modeling the joint behavior of default times, $(\tau_1, \ldots, \tau_d)^\top$.

As in (16), $\exp(-h\tau_k)$ is uniformly distributed over $[0, 1]$, thus let $U_k = \exp(-h\tau_k),\ k = 1, \ldots, d$. The joint CDF of $(U_1, \ldots, U_d)^\top$ is represented as

$$\mathbb{P}(U_1 \leq u_1, \ldots, U_d \leq u_d) = C(u_1, \ldots, u_d).$$

Samples of $(U_1, \dots, U_d)^\top$ are obtained from the copula function $C(u_1, \dots, u_d)$, and using the fact that $U_k = \exp(-h\tau_k),\ k = 1, \dots, d$ one can obtain

$$(\tau_1, \dots, \tau_d)^\top \quad = \quad \left(\frac{-\log U_1}{h}, \dots, \frac{-\log U_d}{h} \right)^\top . \tag{17}$$

By using (11), (12) and (13) the expectation of $\mathbb{E}[L_{q,t_j}],\ q = 1, \dots, Q$ and $j = 1, \dots, J$, is estimated through

$$\widehat{\mathbb{E}}[L_{q,t_j}] \quad = \quad \frac{1}{M} \sum_{m=1}^{M} \left(\min \left[\max \left\{ \frac{1}{d} \sum_{k=1}^{d} (1-r) \mathbf{1}_{\{\mathfrak{z}_k^m \leq t_j\}} - l_q, 0 \right\}, u_q - l_q \right] \right), \tag{18}$$

where $(\mathfrak{z}_1^m, \dots, \mathfrak{z}_d^m)^\top$ is the m-th Monte Carlo sample of the default times $(\tau_1, \dots, \tau_d)^\top$. Therefore at last the empirical representations for spreads of CDS index tranches (upfront rate version) is obtained with the following formula.

$$\widehat{S}_q^{CDX} = \widehat{\mathbb{E}} \left[\frac{\sum_{j=1}^{J} \beta_{t_j} \{ (L_{q,t_j} - L_{q,t_{j-1}}) - 0.5 B_q (t_j - t_{j-1})(P_{q,t_j} + P_{q,t_{j-1}}) \}}{u_q - l_q} \right]. \tag{19}$$

4. Empirical study

4.1. Data Set

In the empirical study, the data of CDX NA IG index is employed. The CDX NA IG index based tranche has four different maturity structures (3, 5, 7 and 10 years) and its underlying entity pool contains overall $d = 125$ CDS contracts. In this paper the maturity with 5 years of the CDX NA IG Series 19 is used, which is issued on 20120920 and ends on 20171220. And the pricing for all $Q = 4$ CDS index tranches is computed with 10 randomly chosen evaluation date points (20140601, 20140703, 20140815, 20140923, 20141011, 20141117, 20141201, 20150107, 20150210, 20150315). In the pricing it is assumed that the risk-free rate as 0.0014 (consistent with the mean of LIBOR of the ten dates) and recovery rate as 0.40 being consistent with it used in Markit company which administrates the CDX NA IG index, see Markit™ (2008). The illustration of the spreads of the four tranches and the corresponding CDS is given in Figure 2 and the data set is given in Table 2.

4.2. Employed Models

Overall 43 copula models used in the study are described below. In the following the notations are set as ga: Gaussian; t: Student-t; fr: Frank; cl: Clayton; gu: Gumbel; jo: Joe; $gai,\ i = 1, \dots, 6$: Gaussian with the correlation matrix $R_{gai},\ i = 1, \dots, 6$; $tj,\ j = 1, \dots, 6$: Student-t with the same correlation matrix structure as R_{gai}; ng: HAC with the Gumbel generator function.

From the elliptical family of copulae an exchangeable Gaussian copula and an exchangeable Student-t copula are chosen in Model 1 and 2.

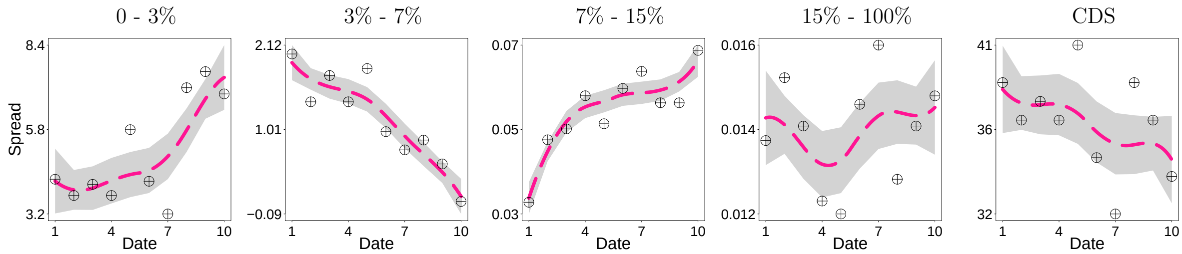

Figure 2: Spreads of four tranches of the CDX NA IG Series 19 and the corresponding CDS spreads illustrated with scatter points, at ten dates 20140601, 20140703, 20140815, 20140923, 20141011, 20141117, 20141201, 20150107, 20150210, 20150315. The dashed line gives a local polynomial regression with its confidence boundaries constraining the shading area.

Date	0-3%	3-7%	7-15%	15-100%	CDS
2014/06/01	4.250	2.000	0.036	0.014	39
2014/070/3	3.750	1.375	0.048	0.015	37
2014/08/15	4.094	1.719	0.050	0.014	38
2014/09/23	3.750	1.375	0.056	0.012	37
2014/10/11	5.775	1.810	0.050	0.012	41
2014/11/17	4.188	0.985	0.057	0.015	35
2014/12/01	3.183	0.747	0.060	0.016	32
2015/01/07	7.065	0.875	0.055	0.013	39
2015/02/10	7.559	0.563	0.055	0.014	37
2015/03/15	6.874	0.073	0.064	0.015	34

Table 2: Spreads of four tranches of the CDX NA IG Series 19 and the corresponding CDS spreads.

(a)
$$
R_{ga1} = \begin{pmatrix}
1 & \cdots & \rho_2 & \rho_1 & \cdots & \cdots & \cdots & \cdots & \rho_1 \\
 & \ddots & & \vdots & & & & & \vdots \\
\rho_2 & \cdots & 1 & \rho_1 & & & & & \vdots \\
\rho_1 & \cdots & \rho_1 & 1 & & & & & \\
\vdots & & & & & & & & \\
\vdots & & & & & & & & \vdots \\
 & & & & \ddots & & & & \vdots \\
 & & & & & 1 & \rho_1 & \cdots & \rho_1 \\
\vdots & & & & & \rho_1 & 1 & \cdots & \rho_2 \\
\vdots & & & & & \vdots & & \ddots & \\
\rho_1 & \cdots & \cdots & & \cdots & \rho_1 & \rho_2 & \cdots & 1
\end{pmatrix}_{d \times d}
$$

(b)
$$
R_{ga2} = \begin{pmatrix}
1 & \cdots & \rho_1 & \rho_1 & \cdots & \cdots & \cdots & \cdots & \cdots & \rho_1 \\
 & \ddots & & \vdots & & & & & & \vdots \\
\rho_1 & \cdots & 1 & \rho_1 & & & & & & \vdots \\
\rho_1 & \cdots & \rho_1 & 1 & \cdot & & & & & \\
\vdots & & & & \cdot & & & & & \\
\vdots & & & & & & & & \cdot & \vdots \\
 & & & & \ddots & & & & & \vdots \\
 & & & & & 1 & \rho_1 & \cdots & \cdots & \rho_1 \\
 & & & & & \rho_1 & 1 & \cdots & \rho_1 & \vdots \\
\vdots & & & & & \vdots & & \ddots & & \vdots \\
\vdots & & & & & \vdots & \rho_1 & \cdots & 1 & \rho_1 \\
\rho_1 & \cdots & \cdots & & \cdots & \rho_1 & \cdots & \cdots & \rho_1 & 1
\end{pmatrix}_{(d+1)\times(d+1)}
$$

(c)
$$
R_{ga3} = \begin{pmatrix}
1 & \cdots & \rho_2 & \rho_2 & \cdots & \cdots & \cdots & \cdots & \rho_2 & \rho_1 \\
 & \ddots & & \vdots & & & & & \vdots & \vdots \\
\rho_2 & \cdots & 1 & \rho_2 & & & & & \vdots & \vdots \\
\rho_2 & \cdots & \rho_2 & 1 & & & & & & \\
\vdots & & & & & & & & & \\
\vdots & & & & & & & & \vdots & \vdots \\
 & & & & \ddots & & & & \vdots & \vdots \\
 & & & & & 1 & \rho_2 & \cdots & \rho_2 & \rho_1 \\
\vdots & & & & & \rho_2 & 1 & \cdots & \rho_2 & \rho_1 \\
\vdots & & & & & \vdots & & \ddots & & \vdots \\
\rho_2 & \cdots & \cdots & & \cdots & \rho_2 & \rho_2 & \cdots & 1 & \rho_1 \\
\rho_1 & \cdots & \cdots & & \cdots & \rho_1 & \rho_1 & \cdots & \rho_1 & 1
\end{pmatrix}_{(d+1)\times(d+1)}
$$

Figure 3: The structure of the correlation matrix (a) R_{ga1} is utilized in Model 3 and Model 9. And the structure of the correlation matrix (b) R_{ga2} is utilized in Model 4 and Model 10. The structure of the correlation matrix (c) R_{ga3} is utilized in Model 5 and Model 11.

Model 1. Gaussian copula,

$$C(u_1, \ldots, u_d; \theta) \quad = \quad C_{ga}(u_1, \ldots, u_d; R_{ga}), \tag{20}$$

where R_{ga} is the correlation matrix with equal correlation in off-diagonal elements.

Model 2. Student-t copula,

$$C(u_1, \ldots, u_d; \theta) \quad = \quad C_t(u_1, \ldots, u_d; R_t, \nu). \tag{21}$$

where R_t is the correlation matrix with equal correlation in off-diagonal elements.

For Gaussian copulae with diverse dependence structures are given in Model 3 to Model 8.

Model 3. Gaussian copula with sectoral dependence illustrated in Figure 3 (a),

$$C(u_1, \ldots, u_d; \theta) \quad = \quad C_{ga1}(u_1, \ldots, u_d; R_{ga1}). \tag{22}$$

Here two parameters are used, ρ_2 for controlling the dependence within a sector and ρ_1 to specify the dependence between sectors. The correlation matrix of Model 3 is given in Figure 3 (a).

Model 4. Gaussian copula with sectoral dependence as in Figure 3 (b),

$$C(u_1, \ldots, u_d, u_{d+1}; \theta) \quad = \quad C_{ga2}(u_1, \ldots, u_d, u_{d+1}; R_{ga2}). \tag{23}$$

It is set that the random recovery U_{d+1} shown in (23) is uniformly distributed. The parameter ρ_1 is the unique parameter for the dependence structure as given in Figure 3 (b).

Model 5. Gaussian copula with sectoral dependence in Figure 3 (c),

$$C(u_1, \ldots, u_d, u_{d+1}; \theta) \quad = \quad C_{ga3}(u_1, \ldots, u_d, u_{d+1}; R_{ga3}). \tag{24}$$

This model is a generalization of Model 4 that let the parameter ρ_2 specify the dependence within and between sectors. U_{d+1} is a random recovery as in latter model. Parameter ρ_1 controls the dependence structure between U_{d+1} and $(U_1, \ldots, U_d)^\top$. The corresponding correlation matrix is illustrated in Figure 3 (c).

Model 6. Gaussian copula with sectoral dependence as in Figure 4 (a),

$$C(u_1, \ldots, u_d, u_{d+1}, \ldots, u_{d+5}; \theta) \quad = \quad C_{ga4}(u_1, \ldots, u_d, u_{d+1}, \ldots, u_{d+5}; R_{ga4}). \tag{25}$$

As diverse sectors may have heterogeneous recovery rates, therefore Model 6 let $(U_{d+1}, \ldots, U_{d+5})$ be six different uniformly distributed random recovery rates for each vector separately. Figure 4 (a) presents the correlation matrix for Model 6 where the parameter ρ_2 is responsible for within sector dependence and the parameter ρ_1 for between sectors dependence.

Model 7. Gaussian copula with sectoral dependence as in Figure 4 (b),

$$C(u_1, \ldots, u_d, u_{d+1}, \ldots, u_{d+6}; \theta) \quad = \quad C_{ga5}(u_1, \ldots, u_d, u_{d+1}, \ldots, u_{d+5}; R_{ga5}). \tag{26}$$

Model 7 still keeps the six heterogeneous recovery rates setting but it is modified that the parameter ρ_3 is used to specify the dependence structure within sectors and the parameter ρ_2 to control the dependence between $U_s,\ s = d+1, \ldots, d+5$ and 5 different sectors. At last the parameter ρ_1 is used to specify the dependence between blocks as described in Figure 4 (b).

Model 8. Gaussian copula with sectoral dependence as in Figure 4 (c),

$$C(u_1, \ldots, u_d, u_{d+1}; \theta) \quad = \quad C_{ga6}(u_1, \ldots, u_d, u_{d+1}; R_{ga6}). \tag{27}$$

This model still uses 3 parameters to specify the dependence structure of $(U_1, \ldots,\ U_d, U_{d+1})^\top$. For the within-sector dependence, the parameter ρ_3 and the parameter ρ_2 are used to control the between-sector dependence. At last the parameter ρ_1 is for the dependence between U_{d+1}, which stands for the single random recovery rate, and $(U_1, \ldots, U_d)^\top$.

As the Gaussian copula has zero tail-dependence, therefore another member of elliptical copula with the fat tail-dependence feature, the Student-t copula, is considered. Models 9-14 are the Student-t copulae,

(a)
$$R_{ga4} = \begin{pmatrix}
1 & \cdots & \rho_2 & \rho_2 & \rho_1 & \cdots & \cdots & & \cdots & \cdots & \rho_1 \\
 & \ddots & & \vdots & \vdots & & & & & & \vdots \\
\rho_2 & \cdots & 1 & \rho_2 & \vdots & & & & & & \vdots \\
\rho_2 & \cdots & \rho_2 & 1 & \rho_1 & & & & & & \\
\rho_1 & \cdots & \cdots & \rho_1 & 1 & & & & & & \\
\vdots & & & & & & & & & & \\
\vdots & & & & & & & & & & \vdots \\
 & & & & & \ddots & & & & & \vdots \\
 & & & & & & 1 & \rho_1 & \cdots & \cdots & \rho_1 \\
 & & & & & & \rho_1 & 1 & \cdots & \rho_2 & \rho_2 \\
\vdots & & & & & & \vdots & & \ddots & & \vdots \\
\vdots & & & & & & \vdots & \rho_2 & \cdots & 1 & \rho_2 \\
\rho_1 & \cdots & \cdots & & & \cdots & \rho_1 & \rho_2 & \cdots & \rho_2 & 1
\end{pmatrix}_{(d+5)\times(d+5)}$$

(b)
$$R_{ga5} = \begin{pmatrix}
1 & \cdots & \rho_3 & \rho_2 & \rho_1 & \cdots & \cdots & & \cdots & \cdots & \rho_1 \\
 & \ddots & & \vdots & \vdots & & & & & & \vdots \\
\rho_3 & \cdots & 1 & \rho_2 & \vdots & & & & & & \vdots \\
\rho_2 & \cdots & \rho_2 & 1 & \rho_1 & & & & & & \\
\rho_1 & \cdots & \cdots & \rho_1 & 1 & & & & & & \\
\vdots & & & & & & & & & & \\
\vdots & & & & & & & & & & \vdots \\
 & & & & & \ddots & & & & & \vdots \\
 & & & & & & 1 & \rho_1 & \cdots & \cdots & \rho_1 \\
 & & & & & & \rho_1 & 1 & \cdots & \rho_3 & \rho_2 \\
\vdots & & & & & & \vdots & & \ddots & & \vdots \\
\vdots & & & & & & \vdots & \rho_3 & \cdots & 1 & \rho_2 \\
\rho_1 & \cdots & \cdots & & & \cdots & \rho_1 & \rho_2 & \cdots & \rho_2 & 1
\end{pmatrix}_{(d+5)\times(d+5)}$$

(c)
$$R_{ga6} = \begin{pmatrix}
1 & \cdots & \rho_3 & \rho_2 & \cdots & \cdots & & \cdots & \cdots & \rho_2 & \rho_1 \\
 & \ddots & & \vdots & & & & & & \vdots & \vdots \\
\rho_3 & \cdots & 1 & \rho_2 & & & & & & \vdots & \vdots \\
\rho_2 & \cdots & \rho_2 & 1 & & & & & & & \\
\vdots & & & & & & & & & & \\
\vdots & & & & & & & & & \vdots & \\
 & & & & & \ddots & & & & \vdots & \\
 & & & & & & 1 & \rho_2 & \cdots & \rho_2 & \\
\vdots & & & & & & \rho_2 & 1 & \cdots & \rho_3 & \vdots \\
\vdots & & & & & & \vdots & & \ddots & & \vdots \\
\rho_2 & \cdots & \cdots & & & \cdots & \rho_2 & \rho_3 & \cdots & 1 & \rho_1 \\
\rho_1 & \cdots & \cdots & & & & & \cdots & \cdots & \rho_1 & 1
\end{pmatrix}_{(d+1)\times(d+1)}$$

Figure 4: the structure of the correlation matrix (a) R_{ga4} is utilized in Model 6 and Model 12. The structure of the correlation matrix (b) R_{ga5} is utilized in Model 7 and Model 13. And the structure of the correlation matrix (c) R_{ga6} is utilized in Model 8 and Model 14.

denoted by C_{t1}, C_{t2}, C_{t3}, C_{t4}, C_{t5}, C_{t6}, with the same correlation matrix structures shown in Figures 3 and 4. Ten different degrees of freedom are obtained by calibration for the Student-t copula of Model 2. Then these ten calibrated parameters are plugged into Models 9-14 as the known parameters. The mixed copula models with the Student-t copula as component copula use a fixed parameter of degree of freedom equal to 3.

After models constructed by elliptical family of copula, in the following the Archimedean copula based models are given. As introduced before the Archimedean copula members share different tail-dependence structures. Model 15 to Model 18 are four diverse Archimedean copula models represented as follows,

$$C(u_1, \ldots, u_d; \theta_a) \quad = \quad C_a(u_1, \ldots, u_d; \theta_a), \tag{28}$$

where $a = cl, jo, gu, fr$, standing separately for Clayton, Joe, Gumbel and Frank copula.

Model 19 to Model 22 are HAC copulae used in the empirical study.

Model 19. Gumbel HAC,

$$C(u_1, \ldots, u_d, u_{d+1}; \theta) \quad = \quad C^1_{ng2}\{C^2_{ng2}(u_1, \ldots, u_d; \rho_{\mathcal{K}2}), u_{d+1}; \rho_{\mathcal{K}1}\}, \tag{29}$$

where C_{ng2}^1 is the root copula and C_{ng2}^2 is the child copula. Model 19 is a Gumbel HAC copula with one parameter $\rho_{\mathcal{K}1}$ for dependence between sectors and random recovery rate U_{d+1}. And $\rho_{\mathcal{K}2}$ is used for dependence of d entities.

Model 20. Gumbel HAC,

$$
\begin{aligned}
C(u_1, \dots, u_d; \theta) \quad = \quad & C_{ng3}^1 \{ \\
& C_{ng3}^2 (u_1, \dots, u_{s_1}; \rho_{\mathcal{K}2}), \\
& C_{ng3}^2 (u_{s_1+1}, \dots, u_{s_1+s_2}; \rho_{\mathcal{K}2}), \dots, \\
& C_{ng3}^2 (u_{s_1+\dots+s_5+1}, \dots, u_d; \rho_{\mathcal{K}2}); \rho_{\mathcal{K}1}\},
\end{aligned}
\tag{30}
$$

where $s_i,\ i = 1, \dots, 5$ is the number of entities in the i-th sector, C_{ng3}^1 means the root copula in the HAC with a Gumbel generator function and C_{ng3}^2 means the child copula in this model. Model 20 is a HAC without random recovery using a root copula and 5 child copulae. The parameter $\rho_{\mathcal{K}2}$ is for dependence within a sector and $\rho_{\mathcal{K}1}$ for dependence between sectors.

Model 21. Gumbel HAC,

$$
\begin{aligned}
C(u_1, \dots, u_d, u_{d+1}, \dots, u_{d+5}; \theta) \quad = \quad & C_{ng4}^1 \{ \\
& C_{ng4}^2 (u_1, \dots, u_{s_1}, u_{d+1}; \rho_{\mathcal{K}2}), \\
& C_{ng4}^2 (u_{s_1+1}, \dots, u_{s_1+s_2}, u_{d+2}; \rho_{\mathcal{K}2}), \dots, \\
& C_{ng4}^2 (u_{s_1+\dots+s_5+1}, \dots, u_d, u_{d+5}; \rho_{\mathcal{K}2}); \rho_{\mathcal{K}1}\}.
\end{aligned}
\tag{31}
$$

This model has five random recoveries, i.e. for each sector a single random recovery following uniform distribution.

Model 22. Gumbel HAC,

$$
\begin{aligned}
C(u_1, \dots, u_d, u_{d+1}, \dots, u_{d+5}; \theta) \quad = \quad & C_{ng5}^1 [\\
& C_{ng5}^2 \left\{ u_{d+1}, C_{ng5}^3 (u_1, \dots, u_{s_1}; \rho_{\mathcal{K}3}); \rho_{\mathcal{K}2} \right\}, \\
& C_{ng5}^2 \left\{ u_{d+2}, C_{ng5}^3 (u_{s_1+1}, \dots, u_{s_1+s_2}; \rho_{\mathcal{K}3}); \rho_{\mathcal{K}2} \right\}, \dots, \\
& C_{ng5}^2 \left\{ u_{d+5}, C_{ng5}^3 (u_{s_1+\dots+s_5+1}, \dots, u_d; \rho_{\mathcal{K}3}); \rho_{\mathcal{K}2} \right\}; \rho_{\mathcal{K}1}].
\end{aligned}
\tag{32}
$$

Model 22 is a HAC model with a Gumbel generator function using 5 random recoveries, $(U_{d+1}, \dots, U_{d+5})^\top$, and three dependence parameters. $\rho_{\mathcal{K}3}$ is utilized for within-sector dependence, i.e. all 5 sectors share the same dependence parameter in every sector. $\rho_{\mathcal{K}2}$ is employed for dependence between the i-th random recovery and the i-th sector, where $i = 1, \dots, 5$. The parameter $\rho_{\mathcal{K}1}$ control the dependence between the second layer child copulae.

Next the cc-copula models from Model 23 to Model 43 are given. In a cc-copula, six copulae are employed as the component copulae containing the exchangeable Gaussian copula, the Student-t copula with degree

Model	Notation	Model	Notation	Model	Notation	Model	Notation
1	C_{ga}	12	C_{t4}	23	C_{ga-ga}	34	C_{fr-fr}
2	C_t	13	C_{t5}	24	C_{ga-t}	35	C_{fr-cl}
3	C_{ga1}	14	C_{t6}	25	C_{ga-fr}	36	C_{fr-gu}
4	C_{ga2}	15	C_{fr}	26	C_{ga-cl}	37	C_{fr-jo}
5	C_{ga3}	16	C_{cl}	27	C_{ga-gu}	38	C_{cl-cl}
6	C_{ga4}	17	C_{gu}	28	C_{ga-jo}	39	C_{cl-gu}
7	C_{ga5}	18	C_{jo}	29	C_{t-t}	40	C_{cl-jo}
8	C_{ga6}	19	C_{ng2}	30	C_{t-fr}	41	C_{gu-gu}
9	C_{t1}	20	C_{ng3}	31	C_{t-cl}	42	C_{gu-jo}
10	C_{t2}	21	C_{ng4}	32	C_{t-gu}	43	C_{jo-jo}
11	C_{t3}	22	C_{ng5}	33	C_{t-jo}		

Table 3: Abbreviations: ga: Gaussian, t: Student-t, fr: Frank, cl: Clayton, gu: Gumbel, jo: Joe, $gai,\ i = 1, \ldots, 6$: Gaussian with the correlation matrix $R_{gai},\ i = 1, \ldots, 6$, $tj,\ j = 1, \ldots, 6$: Student-t with the same correlation matrix structure as $R_{gai},\ i = 1, \ldots, 6$, ng: HAC with the Gumbel generator function.

of freedom equal to 3, the Frank copula, the Clayton copula, the Gumbel copula and the Joe copula. It is set $\lambda,\ \lambda \in [0, 1]$ as the weight for the component copulae, then a general formula for cc-copula models with two components to be used can be given as follows,

$$C_{comp1-comp2}(u_1, \ldots, u_d; \theta) \quad = \quad \lambda C_{comp1}(u_1, \ldots, u_d; \theta_1) + (1 - \lambda) C_{comp2}(u_1, \ldots, u_d; \theta_2), \tag{33}$$

where the $comp1$, $comp2 \in \{ga, t, fr, cl, gu, jo\}$ and parameters θ_1 and θ_2 belong correspondingly to the component copula 1 and 2. An example of a cc-copula is given as follows,

Model 23. cc-copula with two Gaussian components,

$$C_{ga-ga}(u_1, \ldots, u_d; \theta) \quad = \quad \lambda C_{ga}(u_1, \ldots, u_d; \theta_1) + (1 - \lambda) C_{ga}(u_1, \ldots, u_d; \theta_2). \tag{34}$$

According to the convention in (33), C_{ga-ga} in Model 23 means that this model is constructed by two Gaussian (ga) copulae. All the 43 copula models used in this paper are listed in the Table 3.

4.3. Parameter Calibration

HAC, Archimedean copulae, elliptical copulae and cc-copula have been introduced, which can be applied in CDS index tranche pricing by using the copula to construct the dependence structure of default times $(\tau_1, \ldots, \tau_d)^\top$. In this work it is assumed the hazard function as a constant scalar h and this quantity is implied from the market spreads of the CDX NA IG Index Series 19. For a detailed method of implication of h it is referred to Hofert and Scherer (2011).

The exact computation of tranche prices can be performed by the following algorithm.

Algorithm:

(1) Choose a copula model C listed in the Table 3.

(2) Sample by $M = 10^4$ runs of Monte Carlo simulation according to $(U_1, \ldots, U_d)^\top \sim C$.

(3) Obtain samples of $(u_{m,1}, \ldots, u_{m,d})^\top$, $m = 1, \ldots, M$.

(4) Compute (11) to (12) using samples obtained from the step 3.

For models embedded with one random recovery such as (23), (24), (27), (29), and with 5 random recoveries such as (25), (26), (31), (32) one needs to obtain samples respectively according to $(U_1, \ldots, U_d, U_{d+1})^\top \sim C$ and $(U_1, \ldots, U_d, U_{d+1}, \ldots, U_{d+5})^\top \sim C$ in step (2) of algorithm.

After $(U_1, \ldots, U_d)^\top \sim C$ is sampled from copulae, then (17) is used to obtain samples of default times $(\tau_1, \ldots, \tau_d)^\top$ which can be utilized to compute the portfolio loss in (8), q-th tranche loss in (9) and the outstanding notional in (10). At last by (11) and (12) the q-th default leg DL_q and the q-th premium leg PL_q for CDS index tranche pricing can be obtained. Here it uses the notation $\widehat{S}_q^{CDX}$, defined under (19), as the tranche spreads (upfront rate version) by Monte Carlo simulation under models listed in Table 3 and S_q^{Market} as the real market tranche spread (upfront rate version). And for the parameter calibration, the following measure is utilized, which is a root-mean-square error (RMSE) such that,

$$RMSE = \sqrt{\frac{1}{Q}\sum_{q=1}^{Q}\left(\widehat{S}_q^{CDX} - S_q^{Market}\right)^2}. \tag{35}$$

According to the minimization of RMSE in (35) the calibration is performed.

As it is given that RMSE is an argument representation, therefore it is needed to perform numerical optimization to calibrate parameters. For all these models the grid search with the multi-core parallel computation in the optimization is employed.

4.4. Results and Analysis

In Table 4 the mean of RMSEs (MRMSE) based on ten pricing points has been calculated and a ranking based on the mean of the RMSEs is given. In Table 5 and 6 the parameter calibration of the 21 cc-copula models is given. According to the ranking of MRMSE and the parameter calibration, we interpret results as follows:

1. The cc-copula with Archimedean components obtains advantage in CDX pricing. In Table 4 it is found that according to the mean of RMSEs the top three best performed models are correspondingly C_{cl-jo}, C_{fr-gu} and C_{cl-gu}. And it is shown that the top 17 models are all cc-copula models. Especially it can

Rank	Model	MRMSE	Rank	Model	MRMSE	Rank	Model	MRMSE	Rank	Model	MRMSE
1	C_{cl-jo}	0.0980	12	C_{ga-ga}	0.1585	23	C_{jo}	0.4693	34	C_{t4}	1.0222
2	C_{fr-gu}	0.1037	13	C_{cl-cl}	0.1717	24	C_{ng3}	0.4798	35	C_{ga2}	1.0309
3	C_{cl-gu}	0.1062	14	C_{fr-cl}	0.1803	25	C_{ng2}	0.4967	36	C_{t1}	1.0851
4	C_{t-cl}	0.1142	15	C_{t-gu}	0.3433	26	C_{fr-fr}	0.7303	37	C_{ga4}	1.1020
5	C_{ga-jo}	0.1170	16	C_{gu-gu}	0.3574	27	C_{t}	0.8785	38	C_{ga1}	1.1060
6	C_{fr-jo}	0.1182	17	C_{t-jo}	0.3621	28	C_{ga6}	0.9387	39	C_{t4}	1.1140
7	C_{t-fr}	0.1228	18	C_{ng5}	0.3705	29	C_{t6}	0.9728	40	C_{fr}	1.2916
8	C_{ga-cl}	0.1344	19	C_{ng4}	0.3805	30	C_{ga5}	0.9772	41	C_{t-t}	1.4206
9	C_{ga-t}	0.1399	20	C_{gu-jo}	0.3852	31	C_{t5}	0.9854	42	C_{cl}	1.4770
10	C_{ga-gu}	0.1433	21	C_{gu}	0.4052	32	C_{t2}	1.0004	43	C_{ga}	2.9570
11	C_{ga-fr}	0.1536	22	C_{jo-jo}	0.4207	33	C_{ga3}	1.0194			

Table 4: The ranking of 43 copula based models under the mean RMSE (MRMSE). Abbreviations: ga: Gaussian, t: Student-t, fr: Frank, cl: Clayton, gu: Gumbel, jo: Joe, gai, $i = 1, \dots, 6$: Gaussian with the correlation matrix R_{gai}, $i = 1, \dots, 6$, tj, $j = 1, \dots, 6$: Student-t with the same correlation matrix structure as R_{gai}, $i = 1, \dots, 6$, ng: HAC with the Gumbel generator function.

be seen that the top three models are not only the cc-copula models but also their components are all from the Archimedean family and it is quite clear that if a model belongs to a member in the top five rank then there must be at least one component copula coming from a Gumbel copula or a Joe copula or a Clayton copula, the copulae with lower or upper tail-dependence. The comparison in RMSE measure of the best three models and the worst three models is shown in the first two rows of Figure 5. It is clear the gap between the best and the worst in RMSE gauge is quite large.

2. Dearth of asymmetric tail dependence leads to the failure for the elliptical family in the MRMSE ranking. In the ranking in Table 4 another result is that the group of elliptical copulae perform the worst, see last two rows of Figure 5. One can see that the worst ten models are almost all elliptical copulae. And under the same structures, the Gaussian copula models and the Student-t copula models are compared pair by pair, and it is found that in every structure introduced by Figure 3 and 4 the Student-t copula models perform similarly to the Gaussian copula models. The last column in Table 4 shows that the elliptical copulae are not appropriate for modeling the defaults dependence under the context of CDX NA IG Series 19 index tranche. It can be seen clearly that the Gaussian copulae and the Student-t copulae rank in quite low place.

3. The cc-copula models as a group outperform the competing models. Hierarchical Archimedean copulae perform better than elliptical copulae, see Figure 5. The best HAC model is C_{ng5} ranked at the 18th

place being better than the best single parameter Archimedean copula C_{gu} ranked in place of 21. And the best performed elliptical copula is C_{ga6} ranking at the 28th place. Elliptical family performs the worst. Single parameter Archimedean copula models show bifurcating performance. The copulae with upper tail dependence structure (Gumbel and Joe copulae) show fair performance, while the lower tail-dependent model (Clayton copula) and the zero tail-dependent model (Frank copula) belong to the tail group.

4. The cc-copula models have adaptivity and flexibility to CDX pricing. The Figure 7 shows the relationship between the weight parameter and the date in upper triangular. We can observe that the cc-copula model can reduce the RMSE by adjusting the weight parameter in order to be more flexible in different environment. The elliptical copula, the Archimedean copula and the HAC cannot own such properties, which could be a reason of obtaining higher MRMSE in the ranking. Last but not the least, cc-copula models perform stably through ten pricing dates. According to the Figure 6 it can be observed that the cc-copula models' RMSEs vary more stable than the other models. The elliptical models vary stronger through ten pricing dates.

Therefore from the above analysis, some conclusions can be obtained. Firstly, the cc-copula model is superior against elliptical copula model, single parameter Archimedean copula model and HAC model, according to the mean RMSE ranking. Secondly, among the well performed cc-copula models the model employing a Gumbel or Joe or Clayton copula has better performance as the both components share the asymmetrical tail-dependence. At last it is concluded that the elliptical copula model are not appropriate for the CDS index tranche pricing as its elliptical distribution and symmetrical tail-dependence.

5. Conclusion

The goal of this paper is to construct defaults dependence structure mainly with cc-copulae for the CDX tranche pricing. In this work totally 43 diverse copula models are employed, containing 21 cc-copulae with two component copulae coming from two elliptical copulae and four Archimedean copulae. At last all computation results are given out based on the RMSE measure. It is found that cc-copula models have dominant performance compared with other copula models. Especially those cc-copulae which own at least one asymmetrical component copula coming from the Gumbel, Joe or Clayton copula, show top performance. It is a clear evidence that joint defaults are asymmetrically tail-dependent. And in the other three families, the elliptical family performs the worst, which means that those copulae without tail-dependence feature and asymmetrical distribution are not suitable for CDX tranche pricing. The rest two families, the Archimedean copula family and the HAC family, perform similarly and place in the middle of the ranking.

Reference

Cherubini, U., Luciano, E., Vecchiato, W., 2004. Copula Methods in Finance. WILEY.

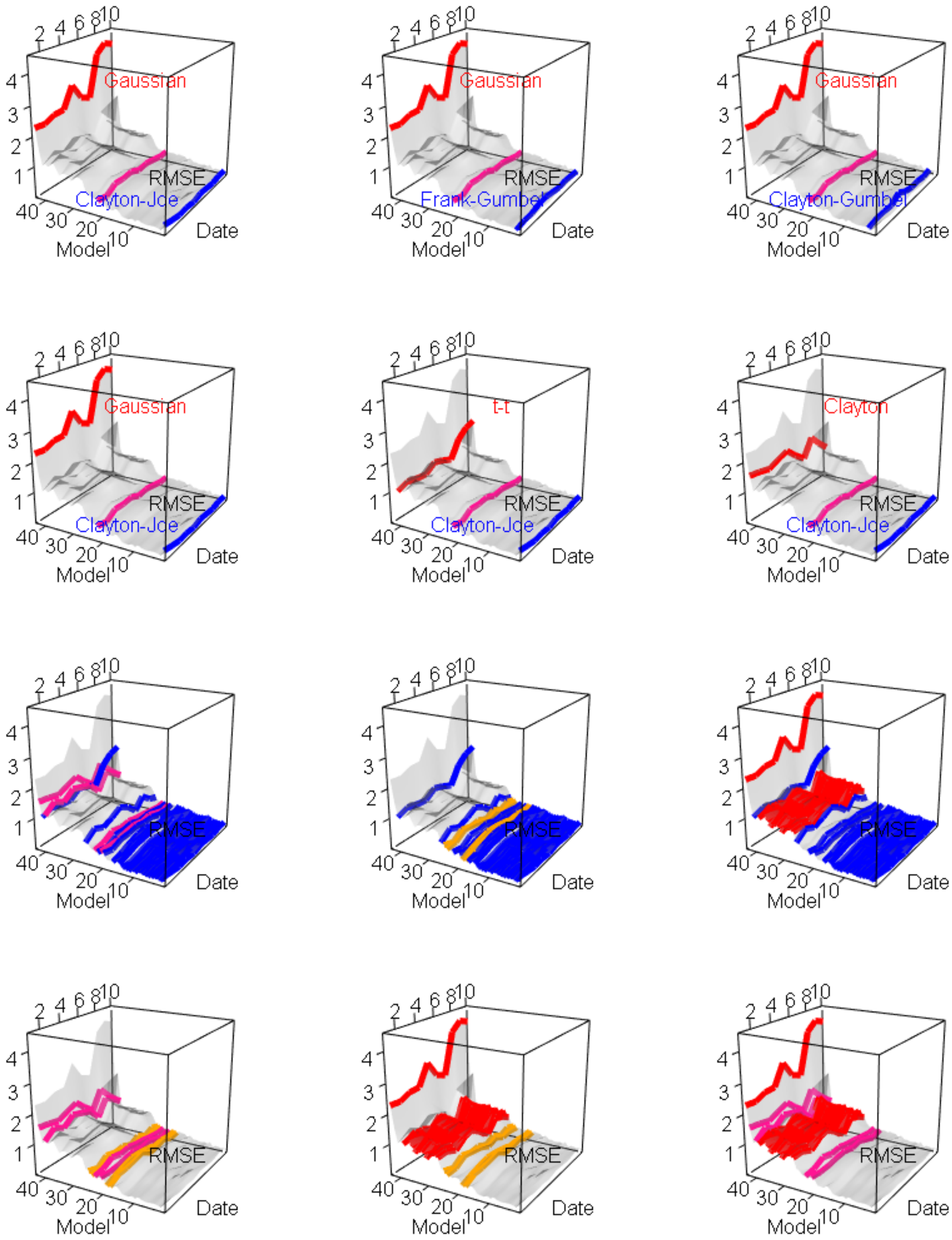

Figure 5: Comparison of RMSEs between the three best performed models (C_{cl-jo}, C_{fr-gu}, C_{cl-gu}) and the worst performed model (C_{ga}) at ten date points (see the first row). Comparison of RMSEs between the three worst performed models (C_{ga}, C_{t-t}, C_{cl}) and the best performed model (C_{cl-jo}) at ten date points (see the second row). The pink lines in the first and second row stands for the median performed model, C_{jo-jo}. Comparison of RMSEs of models between four different families (see the third and the fourth row). The red lines stands for the models from the elliptical family, the orange for the HAC family, the pink for the Archimedean family and the blue for the cc-copula family. The transparent gray terrain in every plot stands for the RMSE surface of all 43 models.

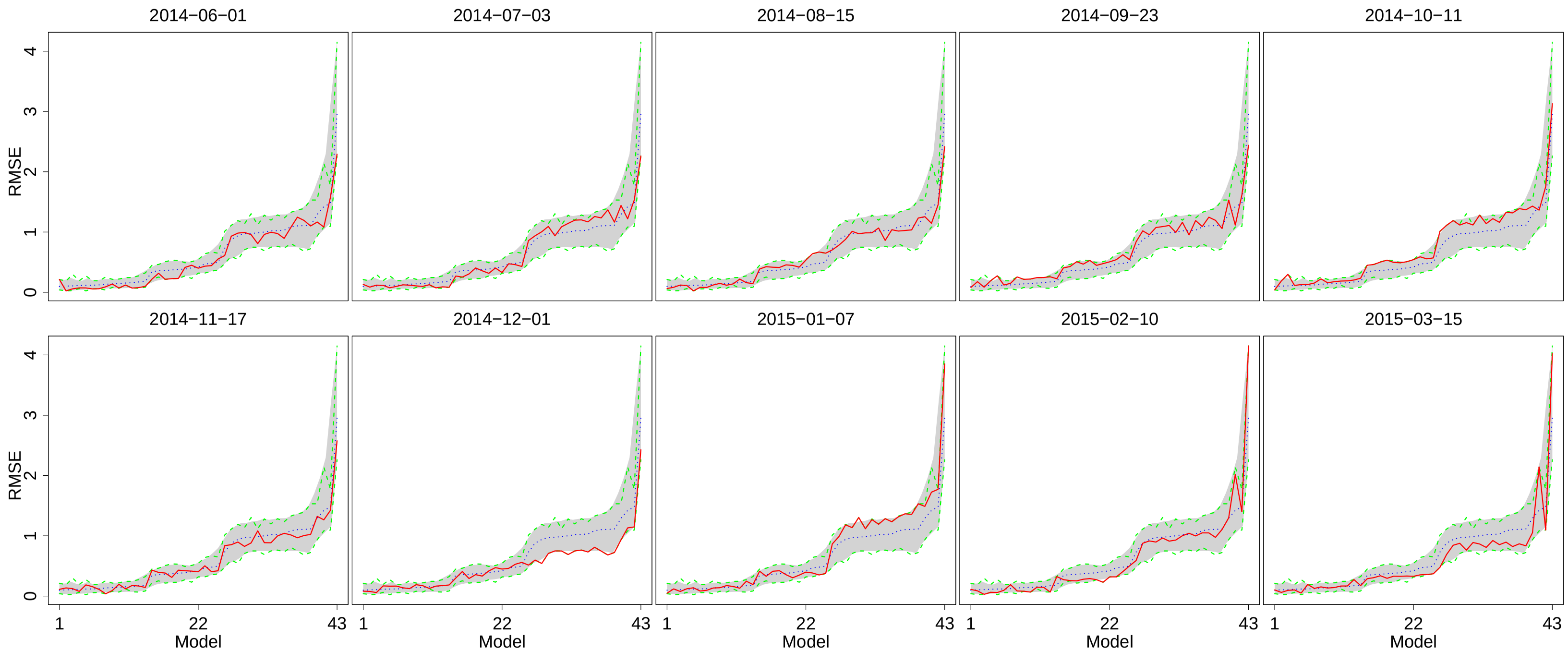

Figure 6: RMSEs comparison of 43 models at ten pricing dates. The red line stands for the RMSE of the corresponding pricing date. The green line stands for the RMSE bounds of ten dates. The black dashed line shows the mean RMSE through ten pricing date points. The shading area is limited by 0.05 and 0.95 nonlinear local quantile regressions.

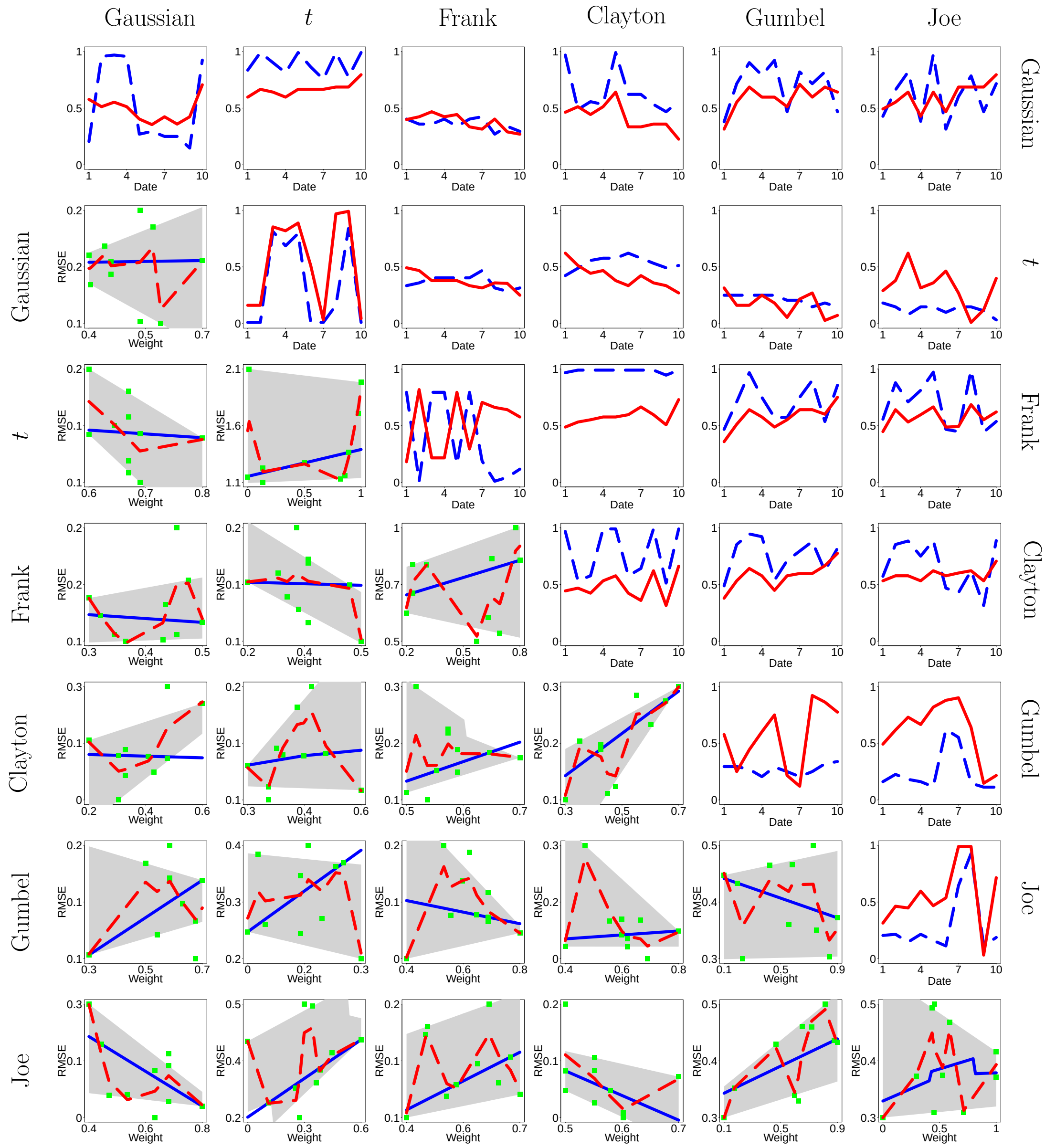

Figure 7: The lower triangle shows the relationship between the weight (λ) and the RMSE for 21 cc-copula models. The blue solid line stands for the 0.5 quantile regression with 0.15 and 0.85 quantile regressions as the shading boundaries. The dashed red line stands is a local polynomial linear regression. The upper triangle illustrates the two parameters' series. The red solid line stands for λ and the blue dashed for θ_1.

Model		20140601	20140703	20140815	20140923	20141011	20141117	20141201	20150107	20150210	20150315
C_{cl-cl}	θ_1	0.443	0.990	0.990	0.532	0.532	0.990	0.990	0.512	0.990	0.443
	θ_2	0.968	0.532	0.577	0.990	0.990	0.577	0.641	0.990	0.512	0.990
	λ	0.443	0.468	0.423	0.532	0.577	0.423	0.359	0.621	0.314	0.661
C_{cl-gu}	θ_1	0.532	0.532	0.641	0.577	0.463	0.577	0.597	0.488	0.512	0.557
	θ_2	0.488	0.853	0.946	0.921	0.537	0.710	0.794	0.879	0.621	0.819
	λ	0.379	0.532	0.641	0.577	0.448	0.577	0.597	0.597	0.666	0.774
C_{cl-jo}	θ_1	0.577	0.577	0.577	0.532	0.577	0.577	0.597	0.512	0.468	0.488
	θ_2	0.572	0.853	0.883	0.750	0.887	0.468	0.428	0.621	0.314	0.887
	λ	0.532	0.577	0.577	0.532	0.621	0.577	0.601	0.621	0.532	0.706
C_{fr-cl}	θ_1	0.314	0.359	0.379	0.403	0.359	0.379	0.448	0.270	0.206	0.294
	θ_2	0.968	0.990	0.990	0.990	0.990	0.990	0.990	0.990	0.946	0.990
	λ	0.488	0.532	0.552	0.577	0.577	0.597	0.666	0.597	0.508	0.730
C_{fr-fr}	θ_1	0.166	0.794	0.188	0.143	0.794	0.226	0.794	0.794	0.794	0.794
	θ_2	0.794	0.010	0.794	0.794	0.161	0.794	0.188	0.010	0.044	0.117
	λ	0.182	0.818	0.216	0.216	0.794	0.294	0.706	0.661	0.641	0.577
C_{fr-gu}	θ_1	0.314	0.359	0.423	0.403	0.314	0.379	0.423	0.314	0.294	0.359
	θ_2	0.468	0.710	0.968	0.750	0.572	0.572	0.750	0.901	0.537	0.857
	λ	0.359	0.512	0.641	0.577	0.488	0.552	0.641	0.641	0.601	0.750
C_{fr-jo}	θ_1	0.359	0.423	0.379	0.423	0.403	0.379	0.359	0.359	0.270	0.270
	θ_2	0.557	0.879	0.710	0.812	0.972	0.468	0.448	0.990	0.443	0.537
	λ	0.448	0.641	0.532	0.597	0.666	0.488	0.492	0.686	0.552	0.621
C_{gu-gu}	θ_1	0.270	0.246	0.294	0.314	0.250	0.181	0.339	0.206	0.161	0.113
	θ_2	0.294	0.294	0.270	0.206	0.290	0.250	0.206	0.250	0.319	0.339
	λ	0.577	0.250	0.443	0.601	0.750	0.216	0.121	0.921	0.863	0.774
C_{gu-jo}	θ_1	0.339	0.250	0.270	0.294	0.294	0.206	0.206	0.206	0.250	0.077
	θ_2	0.161	0.226	0.182	0.161	0.117	0.621	0.552	0.147	0.113	0.113
	λ	0.492	0.617	0.730	0.666	0.819	0.879	0.901	0.641	0.147	0.216
C_{ga-cl}	θ_1	0.226	0.946	0.990	0.946	0.314	0.990	0.990	0.990	0.990	0.990
	θ_2	0.968	0.488	0.557	0.532	0.990	0.621	0.621	0.532	0.468	0.557
	λ	0.463	0.512	0.443	0.512	0.641	0.334	0.334	0.359	0.359	0.226
C_{ga-fr}	θ_1	0.990	0.990	0.923	0.990	0.990	0.990	0.946	0.968	0.990	0.968
	θ_2	0.403	0.359	0.359	0.403	0.339	0.403	0.423	0.270	0.339	0.294
	λ	0.399	0.423	0.468	0.423	0.443	0.334	0.314	0.403	0.290	0.270

Table 5: Calibration of cc-copulas' parameters, i.e. θ_1, θ_2, λ. Abbreviations: ga: Gaussian, t: Student-t, fr: Frank, cl: Clayton, gu: Gumbel, jo: Joe.

Model		20140601	20140703	20140815	20140923	20141011	20141117	20141201	20150107	20150210	20150315
C_{ga-gu}	θ_1	0.294	0.270	0.359	0.314	0.270	0.250	0.339	0.206	0.206	0.206
	θ_2	0.379	0.715	0.901	0.790	0.921	0.468	0.819	0.715	0.819	0.468
	λ	0.314	0.552	0.686	0.597	0.597	0.512	0.710	0.597	0.686	0.641
C_{ga-ga}	θ_1	0.887	0.250	0.270	0.250	0.990	0.972	0.921	0.990	0.887	0.206
	θ_2	0.206	0.956	0.968	0.956	0.270	0.294	0.250	0.250	0.147	0.923
	λ	0.577	0.512	0.552	0.512	0.403	0.354	0.423	0.359	0.423	0.706
C_{ga-jo}	θ_1	0.290	0.270	0.314	0.226	0.314	0.250	0.314	0.250	0.250	0.250
	θ_2	0.428	0.661	0.818	0.383	0.972	0.314	0.601	0.784	0.468	0.715
	λ	0.492	0.552	0.641	0.428	0.641	0.468	0.686	0.686	0.686	0.794
C_{ga-t}	θ_1	0.314	0.339	0.314	0.314	0.339	0.294	0.294	0.250	0.206	0.250
	θ_2	0.834	0.990	0.901	0.812	0.990	0.863	0.754	0.990	0.774	0.990
	λ	0.597	0.666	0.641	0.597	0.666	0.666	0.666	0.686	0.686	0.794
C_{jo-jo}	θ_1	0.147	0.161	0.246	0.161	0.216	0.206	0.147	0.147	0.853	0.079
	θ_2	0.206	0.216	0.147	0.216	0.161	0.113	0.646	0.928	0.113	0.188
	λ	0.314	0.463	0.448	0.597	0.468	0.537	0.990	0.990	0.032	0.715
C_{t-cl}	θ_1	0.463	0.715	0.774	0.617	0.928	0.730	0.532	0.887	0.754	0.819
	θ_2	0.423	0.488	0.557	0.577	0.577	0.621	0.577	0.532	0.492	0.512
	λ	0.621	0.512	0.443	0.468	0.379	0.334	0.423	0.359	0.334	0.270
C_{t-fr}	θ_1	0.681	0.784	0.887	0.990	0.956	0.972	0.883	0.818	0.537	0.853
	θ_2	0.334	0.359	0.403	0.403	0.403	0.403	0.468	0.314	0.285	0.314
	λ	0.492	0.468	0.379	0.379	0.379	0.334	0.314	0.359	0.354	0.250
C_{t-gu}	θ_1	0.137	0.379	0.314	0.226	0.117	0.887	0.054	0.443	0.468	0.079
	θ_2	0.250	0.250	0.250	0.250	0.250	0.206	0.206	0.147	0.182	0.147
	λ	0.314	0.161	0.161	0.250	0.181	0.054	0.216	0.270	0.028	0.072
C_{t-jo}	θ_1	0.137	0.270	0.294	0.290	0.181	0.226	0.028	0.863	0.010	0.250
	θ_2	0.182	0.147	0.079	0.147	0.147	0.099	0.147	0.147	0.113	0.032
	λ	0.290	0.379	0.621	0.314	0.359	0.463	0.270	0.010	0.121	0.399
C_{t-t}	θ_1	0.863	0.794	0.010	0.010	0.010	0.077	0.839	0.010	0.010	0.077
	θ_2	0.010	0.010	0.812	0.686	0.794	0.010	0.010	0.161	0.839	0.010
	λ	0.161	0.161	0.853	0.818	0.887	0.512	0.032	0.968	0.990	0.044

Table 6: Calibration of cc-copula's parameters, i.e. θ_1, θ_2, λ. Abbreviations: ga: Gaussian, t: Student-t, fr: Frank, cl: Clayton, gu: Gumbel, jo: Joe.

Choros-Tomczyk, B., Haerdle, W.K., Okhrin, O., 2013. Valuation of collateralized debt obligations with hierarchical Archimedean copulae. Journal of Empirical Finance 24, 42–62.

Demarta, S., McNeil, A.J., 2005. The t copula and related copulas. International Statistical Review 73, 111–129.

Embrechts, P., Lindskog, F., McNeil, A., 2003. Modelling dependence with copulas and applications to risk management. Handbook of Heavy Tailed Distributions in Finance: Handbooks in Finance .

Hofert, J.M., 2010. Sampling nested Archimedean copulas with applications to CDO pricing. Dissertation of Ulm University .

Hofert, M., Scherer, M., 2011. CDO pricing with nested Archimedean copulas. Quantitative Finance 11, 775–87.

Hull, J., White, A., 2004. Valuation of a CDO and an n-th to default CDS without Monte Carlo simulation. Journal of Derivatives 12, 8–23.

Joe, H., 1996. Families of m-variate distributions with given margins and m(m-1)/2 bivariate dependence parameters. Lecture Notes-Monograph Series , 120–141.

Joe, H., 2014. Dependence Modeling with Copulas. Chapman and Hall/CRC, Boca Raton.

Li, D.X., 1999. The valuation of basket credit derivatives. CreditMetrics Monitor 4, 34–50.

Li, D.X., 2000. On default correlation: A copula function approach. Journal of Fixed Income 9, 43–54.

Lindskog, F., McNeil, A., 2001. Common poisson shock models: Applications to insurance and credit risk modelling. ASTIN BULLETIN 33, 209–238.

Nelsen, R., 2006. An Introduction to Copulas. Springer- New York.

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

Okhrin, O., Ristig, A., 2014. Hierarchical archimedean copulae: The HAC package. Journal of Statistical Software 58.

Savu, C., Trede, M., 2010. Hierarchies of Archimedean copulas. Quantitative Finance 10, 295–304.

Schloegl, L., O'Kane, D., 2005. A note on the large homogeneous portfolio approximation with the student-t copula. Finance Stochastics 9, 577–584.

Schönbucher, P., 2002. Taken to the limit: Simple and not-so-simple loan loss distribution. Working Paper .

Schönbucher, P., Schubert, D., 2000. Copula-dependent default risk in intensity models. Working paper .

Sklar, A., 1959. Fonctions de répartition à n dimension et leurs marges. Publications de l'Institut de Statistique de l'Université de Paris 8, 299–231.

Markit$^{\mathrm{TM}}$, 2008. Markit credit indices: A primer. Technical Report, URL: `https://www.markit.com/news/Credit%20Indices%20Primer.pdf`.