A mixed copula approach in computing the tranche price of CDS index

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

[a]Institute of Economics and Transport, 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

Copulae have been widely used in pricing of collateralized debt obligation (CDO). Among pricing models, the model with the Gaussian copula became the market standard. In the meanwhile other simple copulae were employed, however a lot of them bear different disadvantages in dependence modeling including destitution of heterogeneous sectorial dependence, asymmetrical dependence and tail dependence. For overcoming these disadvantages and also keeping the number of parameters small, we propose in this paper a mixed copula approach of pricing credit default swap index tranches. Six different component copulae are chosen from the elliptical and the Archimedean family to construct the mixed copula based models. The proposed approach introduced in this work has been further utilized to reproduce the spreads of the CDX NA IG index tranches. The empirical section discusses 43 different copula-based tranche pricing models, over which the mixed copula models show dominant performance against the other benchmark models.

Key words: Copula, Mixed Copula, Collateralized Debt Obligation, 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 including credit default swap (CDS), basket default swap, credit default swap index (CDX), collateralized debt obligation (CDO) have attracted more and more attention. From one perspective, CDO provides credit investors a new opportunity 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 CDO 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.

In the studies of the CDO pricing, the focus is in the dependence modeling of random default times. 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, including lacking the heterogeneity of dependence between sectors and the asymmetrical tail-dependence. This makes the exchangeable Gaussian copula based pricing not accurate.

In order to overcome drawbacks of the Gaussian copula model listed above and to understand causes of the crisis, many new methods have been proposed. Specifying the defaults dependence structure by choosing new copulae possess partly or whole features such as the heterogeneity of dependence in sectors and the asymmetrical tail-dependence. In choosing new copulae, literature is abundant, such as the Student-t copula model in Embrechts et al. [2001]; Demarta and McNeil [2005]; Schloegl and O'Kane [2005], the double-t

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

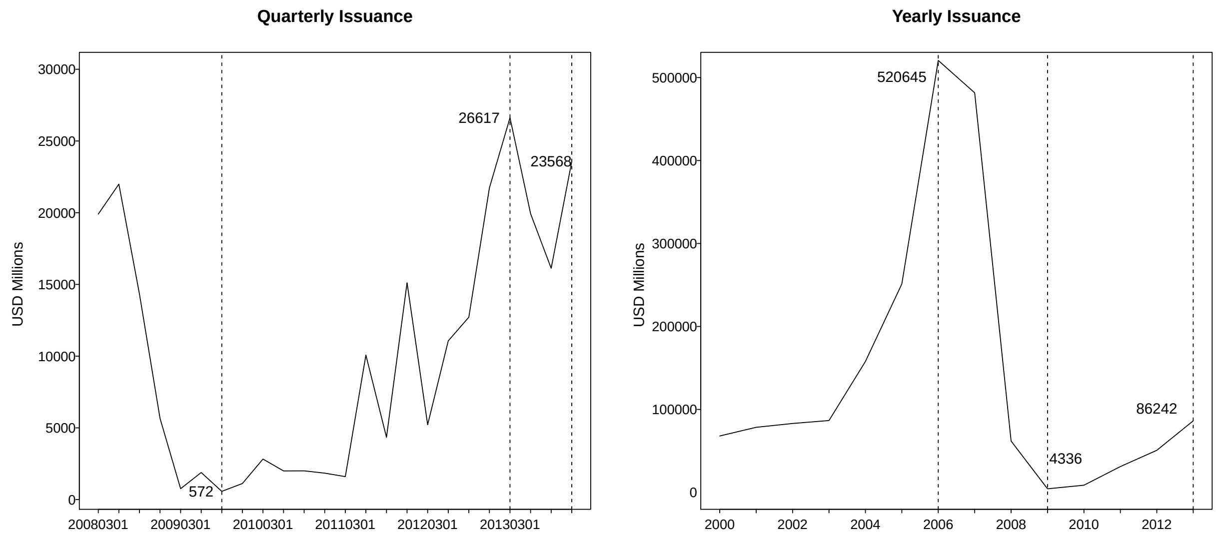

Figure 1: Issuance of CDOs. (left panel) Yearly issuance of CDOs with the data from the Securities Industry and Financial Markets Association. (right panel) Quarterly issuance of CDOs with the data from the Securities Industry and Financial Markets Association.

copula model in Hull and White [2004], the Clayton copula model in Schoenbucher and Schubert [2000]; Lindskog and McNeil [2001]; Schoenbucher [2002], the hierarchical Archimedean copula model in Hofert and Scherer [2008]; Hofert [2010]; Choros-Tomczyk et al. [2013].

This paper focuses on the CDO pricing approach based on the mixed copula, i.e. the convex combination of several copulae, which can be seen as a generalization of copula models employed in CDO pricing. In the mixed copula different copula families are joined together so that the virtues of different copulae can be utilized together for default dependence modeling. An example is to combine two Archimedean copulae together to grant the new model with feature of the asymmetrical tail-dependence, especially the upper tail-dependence such as the Joe and Gumbel copulae. With the Monte Carlo simulation at hand the mixed copula method can be conveniently incorporated in the evaluation of CDO tranche spreads.

Natural question arises whether it is reasonable to further investigate financial products, like CDO, that almost left the market after financial crisis 2008-2009. Figure 2 shows the yearly issuance (left panel) of CDOs and the quarterly issuance (right panel) of CDOs in the world. As can be seen from the left panel that after the financial crisis in 2008 the issuance of CDOs has already indeed sharply decreased and the issuance at the end of 2013 is almost the same as it in 2001. But from Childs [2013] and right panel of Figure 2 we observe that after the lowest reached volume happened in the 3rd quarter in 2009 CDOs' issuance booms again quarterly and reached a peak at the 1st quarter in 2013, which means that although CDOs have caused huge losses for credit investors in financial crisis, it is still an attractive vehicle for investment, especially for risk management in credit risk hedging. This implies that the study of more accurate CDO pricing is worthwhile.

The empirical study of this work uses the data set of the CDX NA IG (Credit Swap Index North America Investment Grade) tranches managed by Markit. Taken from `Bloomberg`, CDX NA IG is a CDS Series 19 index containing 125 names dispersed in 5 diverse sectors including the industrial sector (23 names), the consumer sector (44 names), the energy sector (16 names), the financial sector (19 names) and the technology, media and telecommunication (TMT) sector (23 names).

For the CDO pricing in this paper mixed copula based models are constructed and in every mixed copula model two component copulae from two families are incorporated, the elliptical family including the

Gaussian copula and the Student-t copula and the Archimedean family including Frank, Clayton, Gumbel and Joe copulae. Therefore totally 21 mixed copula models are employed, which can overcome drawbacks such as the heterogeneity of dependence between sectors and the asymmetrical tail-dependence existed in the exchangeable Gaussian copula model for the construction of random default times. In the modeling of a mixed copula with two constituents, the model is calibrated with up to three parameters. Although in this paper only the mixed copula model including two component copulae is used, yet with more powerful computation capacity this mixed copula model can be generalized to more than two components. It is also considered in this work to use another 22 copula models in the study, containing 14 elliptical copula models, 4 Archimedean copula models and 4 HAC models.

The main purpose of this paper is to employ the mixed copula models in reproduction of the spreads of CDO tranches. We calibrate the parameters in the mixed 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 this paper overall 4 tranches are evaluated consistent with the Bloomberg quoting convention.

This paper is structured as follows. Section 2 introduces copulae employed in the work. Section 3 dicusses the CDO structure and its pricing mechanism. The empirical study including the computation of tranche spread, the parameter calibration and analysis of the empirical results in Section 4. Section 5 concludes.

2. Copula Models

Copula is a function which joints marginal CDFs into a multivariate CDF. One can model the whole multivariate distribution separately by choosing the copula and the corresponding margins. Reader interested in the copula theory is referred to Joe [1997]; Nelsen [2006] and for copula application in finance to Cherubini et al. [2004].

According to the *Sklar's theorem* (Sklar [1959]), a joint CDF of d random variables $(X_1, X_2, \dots, X_d)^\top$ can be decomposed into a copula function C and d marginal CDFs $F_k,\ k = 1, \dots, d$, such that,

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

From (2) a copula C can be constructed by a d-dimensional joint CDF and inverse marginal CDFs,

$$C(u_1, \dots, u_d) \quad = \quad F\left\{F_1^{-1}(u_1), \dots, F_d^{-1}(u_d)\right\}. \tag{2}$$

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

$$C_{gs}(u; G) = \Phi_d\left\{\Phi^{-1}(u_1), \dots, \Phi^{-1}(u_d); R\right\},\ u_k \in [0,1],\ k = 1, \dots, d, \tag{3}$$

where R 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 symmetrical with zero tail dependence. Student-t copula on contrary has a non-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, $X = (X_1, \dots, X_d)^\top \in \mathbb{R}^d$. The Student-$t$ copula can be represented as follows,

$$\begin{aligned} C_t(u; \nu, \mu, R) \quad &= \quad \int_{-\infty}^{t^{-1}(u_1)} \cdots \int_{-\infty}^{t^{-1}(u_d)} \frac{\Gamma(\frac{\nu+d}{2})}{\Gamma(\frac{\nu}{2})\sqrt{(\pi\nu)^d \mid R \mid}} \left\{1 + \frac{(x-\mu)^T R^{-1}(x-\mu)}{\nu}\right\}^{-\frac{\nu+d}{2}} \mathrm{d}x, \\ &= \quad T_d\left\{t^{-1}(u_1; \nu), \dots, t^{-1}(u_d; \nu); \nu, \mu, R\right\}, \end{aligned} \tag{4}$$

where T_d is a d-dimensional Student-t CDF and t^{-1} is an inverse of a one dimensional Student-t distribution function.

Another important family is an Archimedean one, whose elements can be constructed as

$$C_A(u_1, \dots, u_d; \theta) = \begin{cases} \varphi^{-1}\left\{\varphi(u_1; \theta) + \cdots + \varphi(u_d; \theta); \theta\right\} & \text{if } \sum_{k=1}^d \varphi(u_k; \theta) \le \varphi(0; \theta), \\ 0 & \text{else,} \end{cases} \tag{5}$$

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, which own different properties of distribution: Frank, Clayton, Gumbel and Joe. Frank copulae is the only elliptically contoured Archimedean copula owning no tail dependence represented as follows,

$$\begin{aligned} C_f(u_1,\ldots,u_d;\theta) &= -\frac{1}{\theta}\log\left[1+\frac{\prod_{k=1}^{d}\{\exp(-\theta u_k)-1\}}{\{\exp(-\theta)-1\}^{d-1}}\right], && (6)\\ \varphi(t;\theta) &= -\log\left\{\frac{\exp(-\theta t)-1}{\exp(\theta)-1}\right\}, && (7)\end{aligned}$$

where $\theta \in (-\infty,+\infty)\backslash\{0\}$. Clayton copula has lower tail dependence and this is important for modeling losses. This copula is defined as

$$\begin{aligned} C_c(u_1,\ldots,u_d;\theta) &= \left(\sum_{k=1}^{d} u_k^{-\theta} - d + 1\right)^{-\theta^{-1}}, && (8)\\ \varphi(t;\theta) &= \frac{1}{\theta}(t^{-\theta}-1), && (9)\end{aligned}$$

where $\theta \in [-1/(d-1),\infty)\backslash\{0\}$. Gumbel copula is the only extreme value copula, and often used in modeling gains, which is represented as follows

$$\begin{aligned} C_g(u_1,\ldots,u_d;\theta) &= \exp\left\{-\sum_{k=1}^{d}(-\log u_k)^{\theta}\right\}^{\theta^{-1}}, && (10)\\ \varphi(t;\theta) &= \{-\log(t)\}^{\theta}, && (11)\end{aligned}$$

where $\theta \in [1,+\infty)$. Joe copula, having right tail dependence, can be represented as follows

$$\begin{aligned} C_j(u_1,\ldots,u_d;\theta) &= 1-\left\{\sum_{k=1}^{d}(1-u_k)^{\theta} - \prod_{k=1}^{d}(1-u_k)^{\theta}\right\}^{\theta^{-1}}, && (12)\\ \varphi(t;\theta) &= -\log\{1-(1-t)^{\theta}\}, && (13)\end{aligned}$$

where $\theta \in [1,+\infty)$.

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 $(U_1,\ldots,U_d)^\top$ is the same as that of $(U_{i_1},\ldots,U_{i_d})^\top$ for all $i_l \neq i_h$, which is restricted in the practice. A much more flexible model is the hierarchical Archimedean copula (HAC), $C(u_1,\ldots,u_d;\theta,s)$, where $s = (\ldots(i_1\ldots i_{j1})\ldots(\ldots)\ldots)$ stands for the HAC's structure and $i_l \in \{1,\ldots,d\}$ records the indices of the variables, and θ is the set of copula parameters. Details of HAC can be referred to Savu and Trede [2009], Okhrin et al. [2013] and Okhrin and Ristig [2014]. A special case of HAC, the d-dimensional fully nested HAC, is shown as follows,

$$\begin{aligned} C_{fHAC}(u_1,\ldots,u_d) &= C[C[\ldots C\{C(u_1,u_2;\varphi_1),u_3;\varphi_2\},\cdots,u_{d-1};\varphi_{d-2}],u_d;\varphi_{d-1}]\\ &= \varphi_{d-1}[\varphi_{d-1}^{-1}[\varphi_{d-2}[\ldots[\varphi_2^{-1}[\varphi_1\{\varphi_1^{-1}(u_1)+\varphi_1^{-1}(u_2)\}]+\varphi_2^{-1}(u_3)]\\ &\quad +\cdots+\varphi_{d-2}^{-1}(u_{d-1})]]+\varphi_{d-1}^{-1}(u_d)]. && (14)\end{aligned}$$

A 5-dimensional fully nested HAC together with a 5-dimensional partially nested HAC are illustrated in the left and right panels of Figure 2 respectively.

It is known that a convex combination of copulae is a copula itself, see Joe [1996], thus let

$$C_{mix}(u_1,\ldots,u_d;\theta_1,\ldots,\theta_I) = \sum_{i=1}^{I}\lambda_i C_i(u_1,\ldots,u_d;\theta_i),\ \sum_{i=1}^{I}\lambda_i = 1,\ u_k \in [0,1],\ k=1,\ldots,d, \qquad (15)$$

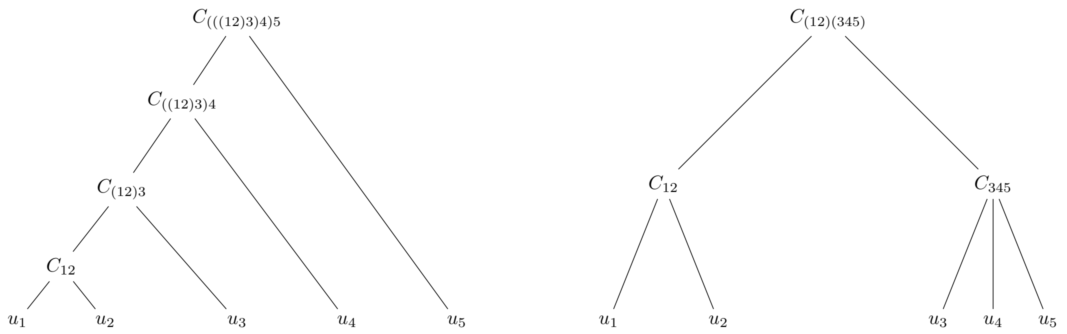

Figure 2: Examples of HAC with $d = 5$.

where λ_i is the i-th weight. And we call $C_i(u_1, \ldots, u_d; \theta_i)$ the i-th component copula with a parameter θ_i. $C_{mix}(u_1, \ldots, u_d; \theta_1, \ldots, \theta_I)$ can be seen as an unknown copula or a complicated copula composed by known $C_i(u_1, \ldots, u_d; \theta_i),\ i = 1, \ldots, I$, hence the mixed copula $C_{mix}(u_1, \ldots, u_d; \theta_1, \ldots, \theta_I)$ should inherit features from its component copulae, which is practical and reasonable in finance for capturing different joint behaviors such as the heterogeneity of dependence and the asymmetrical tail-dependence.

3. Collateralized debt obligations

The collateralized debt obligation 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 CDO issuer is the protection buyer which pays a fixed premium periodically and receives payment for the contingent loss of the credit portfolio. The CDO investor is the protection seller who receives the premium payments from the CDO issuer and takes responsibility to cover the issuer's contingent loss of the credit portfolio.

The CDO tranche technique uses attachment points and detachment points to define hierarchies of the product, which gives the loss percentages of the credit portfolio. In CDX NA IG product, four attachment points are $l_q = (0, 0.03, 0.07, 0.15)^\top$, thus the 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. CDO 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. Meanwhile let the maturity of the CDS index tranches be T, i.e. the length of the contract, 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 let the defaults 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 $\tau_k, k = 1, \ldots, d$, defined in the probability space $(\Omega, \mathcal{F}, \mathbb{P})$, be the default time of the kth entity standing for the survival length and r be the constant recovery rate.

Then one can specify the default of the kth entity by an indicator function $\mathbf{1}_{\{\cdot\}}$ as follows,

$$\Gamma_{k,t_j} = \mathbf{1}_{\{\tau_k \leq t_j\}},\ k = 1, \ldots, d. \tag{16}$$

Afterwards the portfolio loss process L_{t_j} is given through

$$L_{t_j} = \frac{1}{d} \sum_{k=1}^{d} (1 - r)\Gamma_{k,t_j},\ j = 0, \ldots, J. \tag{17}$$

Let $q = 1, \ldots, Q$ be the index of the qth tranche and L_{q,t_j} the tranche loss of the qth tranche at t_j. As the tranche loss is a function of the portfolio loss process, the qth tranche loss is given as follows,

$$L_{q,t_j} = \min\{\max\{L_{t_j} - l_q, 0\}, u_q - l_q\},\ j = 1, \ldots, J,\ q = 1, \ldots, Q. \tag{18}$$

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

$$P_{q,t_j} = u_q - l_q - L_{q,t_j}. \tag{19}$$

Under the non-arbitrage assumption the expectation of the accumulative payments generated by the protection buyer and seller should be equal. In the CDO 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,

$$DL_q \;=\; \mathbb{E}\left\{\sum_{j=1}^{J} \beta_{t_j} N(L_{q,t_j} - L_{q,t_{j-1}})\right\},\ q = 1, \ldots, Q, \tag{20}$$

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 not only a fixed payment, set as $B_q \in \{0.05, 0.01, 0.01, 0.0025\}$ from the first tranche to the fourth tranche based on the quotation convention of the CDX NA IG Series 19, but also an upfront payment, therefore the premium legs for diverse tranches equal to

$$PL_q \;=\; \mathbb{E}\left\{(u_q - l_q)NS_q^{CDO} - \sum_{j=1}^{J} B_q\beta_{t_j}(t_j - t_{j-1})N(P_{q,t_j} + P_{q,t_{j-1}})/2\right\}, \tag{21}$$

where in (21) S_q^{CDO} is the upfront payment rate.

According to (20) and (21) under the non-arbitrage assumption,

$$PL_q \;=\; DL_q, \tag{22}$$

then plugging (21) and (20) into (22) one obtains

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

Hence the qth CDS index tranche upfront payment rate S_q^{CDO} can be extracted as follows,

$$(u_q - l_q)S_q^{CDO}N + \mathbb{E}\left\{\sum_{j=1}^{J} 0.5B_q\beta_{t_j}(t_j - t_{j-1})N(P_{q,t_j} + P_{q,t_{j-1}})\right\} \;=\; DL_q, \tag{23}$$

therefore the S_q^{CDO} is obtained from (23) as follows,

$$\mathrm{S}_q^{CDO} \;=\; \mathbb{E}\left[\frac{\sum_{j=1}^{J} \beta_{t_j}\{(L_{q,t_j} - L_{q,t_{j-1}}) - 0.5B_q(t_j - t_{j-1})(P_{q,t_j} + P_{q,t_{j-1}})\}}{u_q - l_q}\right]. \tag{24}$$

3.2. Modeling of Marginal Default

As mentioned at the beginning, $\tau_k, k = 1, \dots, d$ is the random variable of survival length (or termed as the default time) of the kth entity in the reference pool, then F_k is denoted as the CDF of τ_k and $S_k(t)$ as a survival function such that

$$F_k(t) = \mathbb{P}(\tau_k \leq t),\ t \geq 0. \tag{25}$$

$$S_k(t) = 1 - \mathbb{P}(\tau_k \leq t) = 1 - F_k(t),\ t \geq 0. \tag{26}$$

In a bond market case the function $F_k(t)$ can be understood as the probability of a bond that can not survive over t, while the survival function can be understood as the probability of a bond which can survive over the length of t.

Meanwhile let τ_k have a PDF such that

$$f_k(t) = \frac{\partial F_k(t)}{\partial t} = -\frac{\partial S_k(t)}{\partial t} = \lim_{\Delta \to 0^+} \frac{\mathbb{P}(t \leq \tau_k \leq t + \Delta)}{\Delta}, \tag{27}$$

then hazard rate function $h_k(t)$ can be defined as

$$h_k(t) = \frac{f_k(t)}{1 - F_k(t)} = \frac{-S_k^{'}(s)}{S_k(t)}. \tag{28}$$

Then the $S_k(t)$ is obtained by integrating (28) in both sides over the interval $[0, t]$ such that

$$\int_0^t h_k(s)\mathrm{d}s = \int_0^t \frac{-S_k^{'}(s)}{S_k(s)}\mathrm{d}s.$$

After simple transformation the following representation of $F_k(t)$ is obtained,

$$F_k(t) = 1 - S_k(t) = 1 - \exp\left\{-\int_0^t h_k(s)\mathrm{d}s\right\}. \tag{29}$$

And for the simplicity the hazard rate function is set to be a constant h for all k, therefore

$$F_k(t) = 1 - \exp\left(-ht\right). \tag{30}$$

3.3. Modeling of Joint Defaults

Next copula function is used to model the joint behavior of these default times, $(\tau_1, \dots, \tau_d)^\top$.

As shown in (30) $\exp\left(-h\tau_k\right)$ is uniformly distributed over $[0, 1]$, thus let $U_k = \exp(-h\tau_k),\ k = 1, \dots, d$. The joint CDF of $(U_1, \dots, U_d)^\top$ is represented as

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

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

$$(\tau_1, \dots, \tau_d)^\top \quad = \quad \left(\frac{-\log U_1}{h}, \dots, \frac{-\log U_d}{h}\right)^\top. \tag{31}$$

By using (20), (21) and (22) the expectation of $\mathbb{E}[L_{q,t_j}],\ q = 1, \dots, Q$ and $j = 1, \dots, J$, is estimated through an average such that,

$$\hat{\mathbb{E}}[L_{q,t_j}] \quad = \quad \frac{1}{M}\sum_{m=1}^{M}\left(\min\left[\max\left\{\frac{1}{d}\sum_{k=1}^{d}(1-r)\mathbf{1}_{\{\mathfrak{z}_k^m \leq t_j\}} - l_q, 0\right\}, u_q - l_q\right]\right), \tag{32}$$

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

$$\widehat{S}_q^{CDO} = \widehat{\mathbb{E}}\left[\frac{\sum_{j=1}^{J}\beta_{t_j}\{(L_{q,t_j} - L_{q,t_{j-1}}) - 0.5B_q(t_j - t_{j-1})(P_{q,t_j} + P_{q,t_{j-1}})\}}{u_q - l_q}\right], \tag{33}$$

4. Empirical study

4.1. Data Set

The CDX NA IG index based tranche has four different maturity structures, 3 and 5 years and its underlying pool contains overall $d = 125$ CDS contracts. In this paper the maturity with 5 years of the CDX NA IG Series 19 is chosen, which is issued on 20120920 and ends on 20171220. And the pricing for all $Q = 4$ CDS index tranches is performed with 10 randomly chosen evaluation date points on 20140601, 20140703, 20140815, 20140923, 20141011, 20141117, 20141201, 20150107, 20150210, 20150315. In the pricing it is assumed that the risk-free rate as 0.0014 and recovery rate as 0.40 which is consistent with it used in Markit company which administrates the CDX NA IG Index. Next models for CDS index tranche pricing in the empirical study will be constructed .

4.2. Employed Models

Overall 43 copula models used in the study are listed below. In the following the notations are set as ga: Gaussian, t: Student-t, fr: Frank, cl: Clayton, gu: Gumbel, jo: Joe, $gai,\ i = 1, \ldots, 6$: 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. From the elliptical family of copulae an exchangeable Gaussian copula and an exchangeable Student-t copula are chosen such that

Model 1. Gaussian copula,

$$C(u_1, \ldots, u_d; \theta) \quad = \quad C_{ga}(u_1, \ldots, u_d; R_{ga}), \tag{34}$$

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

Model 2. Student-t copula,

$$C(u_1, \ldots, u_d; \theta) \quad = \quad C_t(u_1, \ldots, u_d; R_t, \nu). \tag{35}$$

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

For Gaussian copula diverse dependence structures are employed.

Model 3. Gaussian copula with sectorial dependence as in Figure 3 (a),

$$C(u_1, \ldots, u_d; \theta) \quad = \quad C_{ga1}(u_1, \ldots, u_d; R_{ga1}). \tag{36}$$

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

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

$$C(u_1, \ldots, u_d, u_{d+1}; \theta) \quad = \quad C_{ga2}(u_1, \ldots, u_d, u_{d+1}; R_{ga2}). \tag{37}$$

It is set that the random recovery U_{d+1} shown in (37) 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 sectorial dependence in Figure 3 (c),

$$C(u_1, \ldots, u_d, u_{d+1}; \theta) \quad = \quad C_{ga3}(u_1, \ldots, u_d, u_{d+1}; R_{ga3}). \tag{38}$$

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 sectorial dependence as in Figure 4 (a),

$$C(u_1, \ldots, u_d, u_{d+1}, \ldots, u_{d+5}; \theta) \quad = \quad C_{ga4}(u_1, \ldots, u_d, u_{d+1}, \ldots, u_{d+5}; R_{ga4}). \tag{39}$$

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

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

(a)

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

(b)

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

(c)

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.

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 sectorial dependence as in Figure 4 (b),

$$C(u_1,\ldots,u_d,u_{d+1},\ldots,u_{d+6};\theta) \quad = \quad C_{ga5}(u_1,\ldots,u_d,u_{d+1},\ldots,u_{d+5};R_{ga5}). \tag{40}$$

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 sectorial dependence as in Figure 4 (c),

$$C(u_1,\ldots,u_d,u_{d+1};\theta) \quad = \quad C_{ga6}(u_1,\ldots,u_d,u_{d+1};R_{ga6}). \tag{41}$$

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, 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 known. The mixed copula models with the Student-t copula as component copula use a fixed parameter of degree of freedom equal to 3.

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

(a)

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

(b)

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

(c)

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.

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,

$$
C(u_1, \ldots, u_d; \theta_a) \quad = \quad C_a(u_1, \ldots, u_d; \theta_a), \tag{42}
$$

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

Model 19. Gumbel HAC,

$$
C(u_1, \ldots, u_d, u_{d+1}; \theta) \quad = \quad C^1_{ng2}\{C^2_{ng2}(u_1, \ldots, u_d; \rho_{\mathcal{K}2}), u_{d+1}; \rho_{\mathcal{K}1}\}, \tag{43}
$$

where C^1_{ng2} is the root copula and C^2_{ng2} 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. The tree graph is illustrated in Figure 5 (a).

Model 20. Gumbel HAC,

$$
\begin{aligned}
C(u_1, \ldots, u_d; \theta) \quad = \quad & C^1_{ng3}\,\{ \\
& C^2_{ng3}\,(u_1, \ldots, u_{s_1}; \rho_{\mathcal{K}2})\,, \\
& C^2_{ng3}\,(u_{s_1+1}, \ldots, u_{s_1+s_2}; \rho_{\mathcal{K}2})\,, \ldots, \\
& C^2_{ng3}\,(u_{s_1+\cdots+s_5+1}, \ldots, u_d; \rho_{\mathcal{K}2})\,; \rho_{\mathcal{K}1}\},
\end{aligned} \tag{44}
$$

where $s_i,\ i = 1, \ldots, 5$ is the number of entities in ith sector, C^1_{ng3} means the root copula in the hierarchical Archimedean copula with a Gumbel generator function and C^2_{ng3} means the child copula in this model.

Model 20 is a hierarchical Archimedean copula 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. The model is given in (44) and the tree structure is illustrated in Figure 5 (b).

Model 21. Gumbel HAC,

$$\begin{aligned} C(u_1,\ldots,u_d,u_{d+1},\ldots,u_{d+5};\theta) &= C^1_{ng4}\{ \\ & \quad C^2_{ng4}(u_1,\ldots,u_{s_1},u_{d+1};\rho_{\mathcal{K}2}), \\ & \quad C^2_{ng4}(u_{s_1+1},\ldots,u_{s_1+s_2},u_{d+2};\rho_{\mathcal{K}2}),\ldots, \\ & \quad C^2_{ng4}(u_{s_1+\cdots+s_5+1},\ldots,u_d,u_{d+5};\rho_{\mathcal{K}2});\rho_{\mathcal{K}1}\}. \end{aligned} \tag{45}$$

This model is a generalization of the later one with five random recoveries being added in modeling, i.e. for each sector a single random recovery following uniform distribution. The tree graph is illustrated in Figure 5 (c).

Model 22. Gumbel HAC,

$$\begin{aligned} C(u_1,\ldots,u_d,u_{d+1},\ldots,u_{d+5};\theta) &= C^1_{ng5}[\\ & \quad C^2_{ng5}\left\{u_{d+1},C^3_{ng5}(u_1,\ldots,u_{s_1};\rho_{\mathcal{K}3});\rho_{\mathcal{K}2}\right\}, \\ & \quad C^2_{ng5}\left\{u_{d+2},C^3_{ng5}(u_{s_1+1},\ldots,u_{s_1+s_2};\rho_{\mathcal{K}3});\rho_{\mathcal{K}2}\right\},\ldots, \\ & \quad C^2_{ng5}\left\{u_{d+5},C^3_{ng5}(u_{s_1+\cdots+s_5+1},\ldots,u_d;\rho_{\mathcal{K}3});\rho_{\mathcal{K}2}\right\};\rho_{\mathcal{K}1}]. \end{aligned} \tag{46}$$

Model 22 is a HAC model with a Gumbel generator function using 5 random recoveries, $(U_{d+1},\ldots,U_{d+5})^\top$, and three 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 ith random recovery and the ith sector, where $i=1,\ldots,5$. The parameter $\rho_{\mathcal{K}1}$ control the dependence between the second layer child copulae, which can be shown in (46) and the structure in Figure 5 (d).

Next the mixed copula models from Model 23 to Model 43 are given. In a mixed copula model six copula models are employed such as the component copulae containing the exchangeable Gaussian copula, the Student-t copula with degree of freedom equal to 3, the Frank copula, the Clayton copula, the Gumbel copula and the Joe copula. It is set $w_l,\ l\in\{1,2\}$ as the weight for the lth component copula, then a general formula for mixed copula models to be used can be given as follows,

$$C_{comp1-comp2}(u_1,\ldots,u_d;\theta) = w_1C_{comp1}(u_1,\ldots,u_d;\theta_1)+w_2C_{comp2}(u_1,\ldots,u_d;\theta_2), \tag{47}$$

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 that Model 23 is a mixed Gaussian copula model with formula is given as follows,

Model 23. Mixed copula,

$$C_{ga-ga}(u_1,\ldots,u_d;\theta) = w_1C_{ga}(u_1,\ldots,u_d;\theta_1)+w_2C_{ga}(u_1,\ldots,u_d;\theta_2). \tag{48}$$

According to the convention in (47), C_{ga-ga} in Model 23 means that this model is constructed by mixture of two Gaussian (ga) copulae. All the 43 copula models used in this paper are listed in the Table 1.

4.3. Parameter Calibration

HAC, Archimedean copulae, mixed copulae and elliptical copulae have been introduced, which can be applied in CDS index tranche pricing by using the copula to construct the dependence structure of default times $(\tau_1,\ldots,\tau_d)^\top$. And (29) is used to model the marginal CDF of $\tau_k,\ k=1,\ldots,d$. 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 [2008].

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

Algorithm:

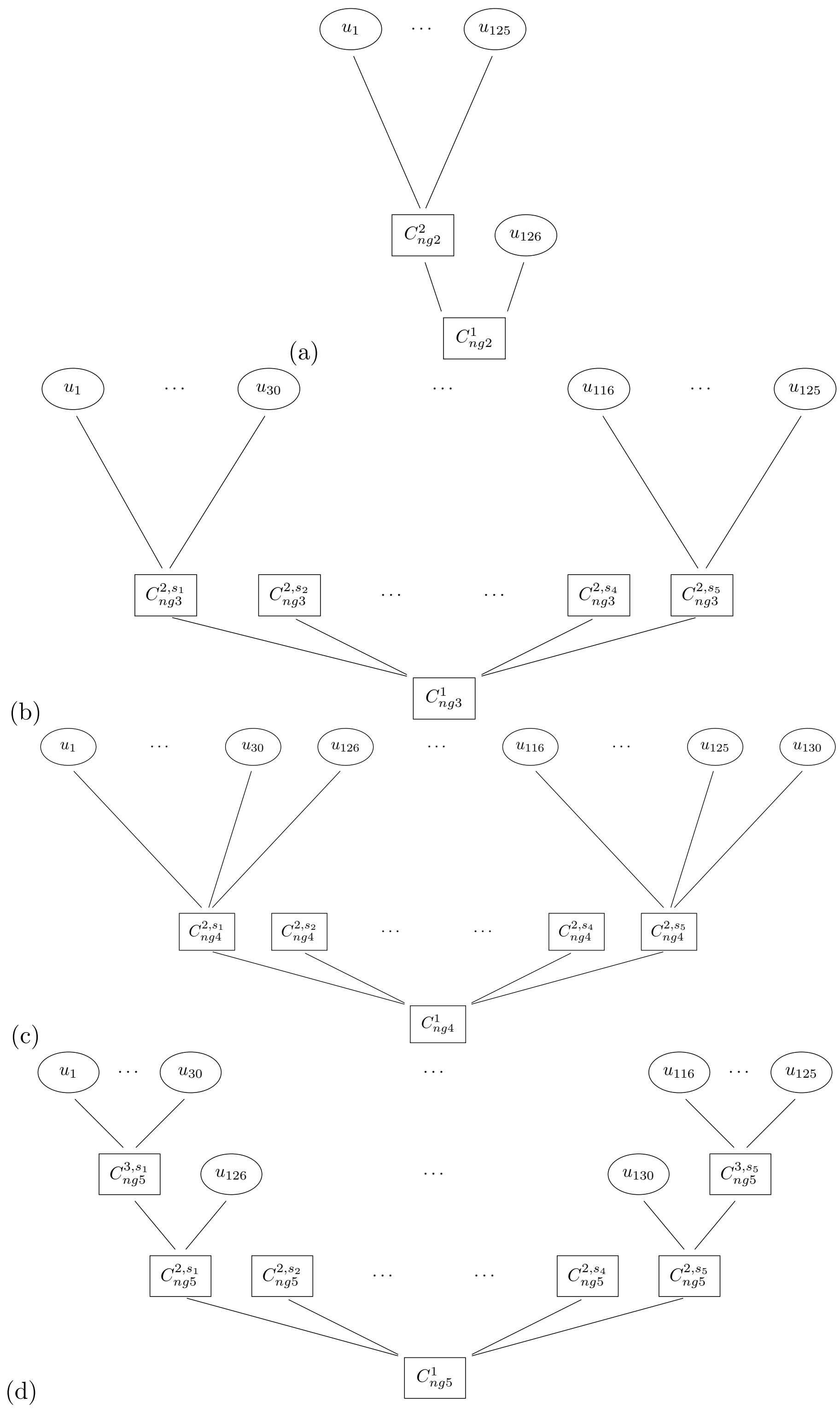

Figure 5: Tree structure plots for Model 19 in (a), Model 20 in (b), Model 21 in (c) and Model 22 in (d).

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 1: 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.

(1) Choose a copula model C listed in the Table 1.
(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 (20) and (21) using samples obtained from the last step.

For models embedded with one random recovery such as (37), (38), (41), (43), and with 5 random recoveries such as (39), (40), (45), (46) 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 (31) is used to obtain samples of default times $(\tau_1, \ldots, \tau_d)^\top$ which can be utilized to compute the portfolio loss in (17), qth tranche loss in (18) and the outstanding notional in (19). At last by (20) and (21) the qth default leg DL_q and the qth premium leg PL_q for CDS index tranche pricing can be obtained. Here it uses the notation $\widehat{S}_q^{CDO}$, defined under (33), as the tranche spreads (upfront rate version) by Monte Carlo simulations under models listed in Table 1 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,

$$RMSE = \sqrt{\frac{1}{Q}\sum_{q=1}^{Q}\left(\widehat{S}_q^{CDO} - S_q^{Market}\right)^2}. \tag{49}$$

According to the minimization of RMSE in (49) 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 4.4 the computation results according to the RMSE measure introduced in (49) are given. In Table 3 the mean of the RMSE measures based on 10 pricing points has been calculated and a ranking based on the mean of the RMSEs is given. For further comparison, Figure 6 shows the change of overall 43 models' RMSEs at 10 computation point with 10 lines. Also four plots at the three dimensional view for the pricing surface are given in Figure 7.

In regard to the ranking comparison based on RMSE and Figure 6 and Figure 7, one can interpret results as follows:

Model	2014-06-01	2014-07-03	2014-08-15	2014-09-23	2014-10-11	2014-11-17	2014-12-01	2015-01-07	2015-02-10	2015-03-15
C_{ga}	2.2922	2.2619	2.4178	2.4385	3.1289	2.5759	2.4286	3.8521	4.1498	4.0247
C_t	0.9311	0.9379	0.7822	1.0199	1.1113	0.8523	0.5980	0.9870	0.8766	0.6890
C_{ga1}	1.1925	1.3677	1.0352	1.1942	1.3885	1.0045	0.6820	1.3546	0.9738	0.8670
C_{ga2}	0.8950	1.1686	1.0378	1.1907	1.1609	1.0413	0.7312	1.2330	0.9976	0.8529
C_{ga3}	0.9967	1.1968	1.0673	1.1619	1.1388	0.8843	0.7487	1.1911	1.0011	0.8075
C_{ga4}	1.2470	1.2345	1.0258	1.2464	1.3178	0.9677	0.7460	1.3652	1.0481	0.8216
C_{ga5}	0.9564	0.9370	0.9717	1.0902	1.1562	0.8820	0.7495	1.3038	0.9638	0.7611
C_{ga6}	0.9823	1.0043	0.8785	0.9551	1.1900	0.8966	0.5408	1.1818	0.9146	0.8431
C_{t1}	1.0750	1.2578	1.0163	1.0913	1.3267	1.0149	0.8098	1.3154	1.0487	0.8948
C_{t2}	0.9606	1.1434	0.9875	0.9941	1.2812	0.8860	0.6890	1.2686	0.9294	0.8640
C_{t3}	0.9754	1.2018	0.8605	0.9522	1.2223	0.9980	0.7636	1.2843	1.0421	0.9218
C_{t4}	1.1004	1.1636	1.2325	1.0581	1.3691	1.0230	0.7231	1.5314	1.1052	0.8338
C_{t5}	0.8074	1.0869	0.9857	1.1088	1.1171	1.0824	0.7479	1.1168	0.9117	0.8889
C_{t6}	0.9936	1.0931	1.0113	1.0754	1.1161	0.8237	0.7067	1.1358	0.8964	0.8758
C_{fr}	1.1663	1.4403	1.2550	1.5306	1.4302	1.3227	0.9342	1.4879	1.2974	1.0513
C_{cl}	1.5779	1.5322	1.4523	1.6164	1.7540	1.4229	1.1476	1.7715	1.4000	1.0956
C_{gu}	0.4492	0.4035	0.4182	0.4762	0.5082	0.4139	0.4723	0.3481	0.2298	0.3327
C_{jo}	0.4333	0.4739	0.6395	0.5412	0.5889	0.5024	0.4598	0.3833	0.3188	0.3520
C_{ng2}	0.5519	0.4282	0.6490	0.5419	0.5706	0.4198	0.5595	0.3762	0.5008	0.3695
C_{ng3}	0.4434	0.4590	0.6689	0.6232	0.5552	0.4049	0.5278	0.3501	0.4068	0.3586
C_{ng4}	0.2298	0.3570	0.4553	0.5299	0.4961	0.4297	0.3318	0.3545	0.2908	0.3297
C_{ng5}	0.2269	0.4042	0.4120	0.4672	0.5337	0.3102	0.3557	0.4204	0.2793	0.2949
C_{cl-cl}	0.0780	0.0905	0.1590	0.2576	0.2041	0.1669	0.1740	0.2471	0.0662	0.2732
C_{cl-gu}	0.0578	0.1187	0.1204	0.0855	0.2984	0.1235	0.0577	0.0764	0.0288	0.0949
C_{cl-jo}	0.2128	0.1311	0.0623	0.0807	0.0397	0.1115	0.0811	0.0477	0.1110	0.1022
C_{fr-cl}	0.0997	0.0844	0.1458	0.2250	0.2336	0.1433	0.1843	0.1906	0.3231	0.1733
C_{fr-fr}	0.6093	0.8588	0.7040	0.8381	1.0122	0.8363	0.5136	0.8651	0.5894	0.4765
C_{fr-gu}	0.0240	0.0862	0.0835	0.1767	0.1867	0.1354	0.0772	0.1193	0.0869	0.0608
C_{fr-jo}	0.0565	0.0980	0.0821	0.1209	0.1291	0.1562	0.1660	0.0841	0.0961	0.1926
C_{gu-gu}	0.3165	0.2529	0.4252	0.4267	0.4621	0.3924	0.4073	0.3283	0.2566	0.3059
C_{gu-jo}	0.4130	0.3151	0.4464	0.4468	0.4898	0.4201	0.4171	0.3035	0.2720	0.3283
C_{ga-cl}	0.0916	0.1185	0.1219	0.2550	0.2230	0.0384	0.1236	0.1344	0.0849	0.1525
C_{ga-fr}	0.1204	0.1264	0.1398	0.2442	0.1865	0.1190	0.1263	0.1597	0.1474	0.1666
C_{ga-gu}	0.0716	0.0986	0.1177	0.2200	0.1755	0.1957	0.1728	0.1741	0.0665	0.1409
C_{ga-ga}	0.0706	0.0735	0.2194	0.2451	0.1898	0.1762	0.1655	0.1305	0.1463	0.1676
C_{ga-jo}	0.0727	0.0746	0.0241	0.2737	0.1283	0.1857	0.1647	0.1378	0.0592	0.0490
C_{ga-t}	0.1375	0.1061	0.1479	0.2153	0.1585	0.0918	0.1889	0.1386	0.0806	0.1336
C_{jo-jo}	0.4003	0.3284	0.5346	0.5060	0.5421	0.4024	0.4484	0.3979	0.3186	0.3283
C_{t-cl}	0.0737	0.1156	0.1132	0.1924	0.1143	0.0773	0.1687	0.1215	0.0625	0.1023
C_{t-fr}	0.0597	0.1240	0.0807	0.1493	0.1536	0.1101	0.1376	0.0959	0.1894	0.1272
C_{t-gu}	0.2184	0.2702	0.3866	0.4054	0.4484	0.4302	0.2998	0.4130	0.2731	0.2881
C_{t-jo}	0.2150	0.3050	0.4158	0.5076	0.5023	0.3826	0.2913	0.4119	0.2523	0.3376
C_{t-t}	1.0833	1.2192	1.1447	1.1150	1.3633	1.2651	1.1314	1.7237	2.0181	2.1420

Table 2: $RMSE$s after the calibration for copula models from Model 1 to Model 43 with the setting of runs of Monte Carlo simulation as $M = 10^4$.

1. In Table 3 it is found that according to the mean of RMSE 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 mixed copula models. Especially it can be seen that the top three models are not only the mixed 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.
2. In the ranking in Table 3 another result is that the elliptical copulae perform worst. 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 3 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.

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

Table 3: The ranking of 43 models under the mean RMSE. 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.

3. Hierarchical Archimedean copulae perform better than elliptical copulae. The best hierarchical Archimedean copula 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.

Therefore from the above analysis, some conclusions can be obtained. Firstly, the mixed copula model is superior against elliptical copula model, single parameter Archimedean copula model and hierarchical Archimedean copula model employed in Table 1 according to the mean RMSE ranking. Secondly, among the well performed mixed copula models the model employing a Gumbel or Joe or Clayton copula has better performance as the both 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.

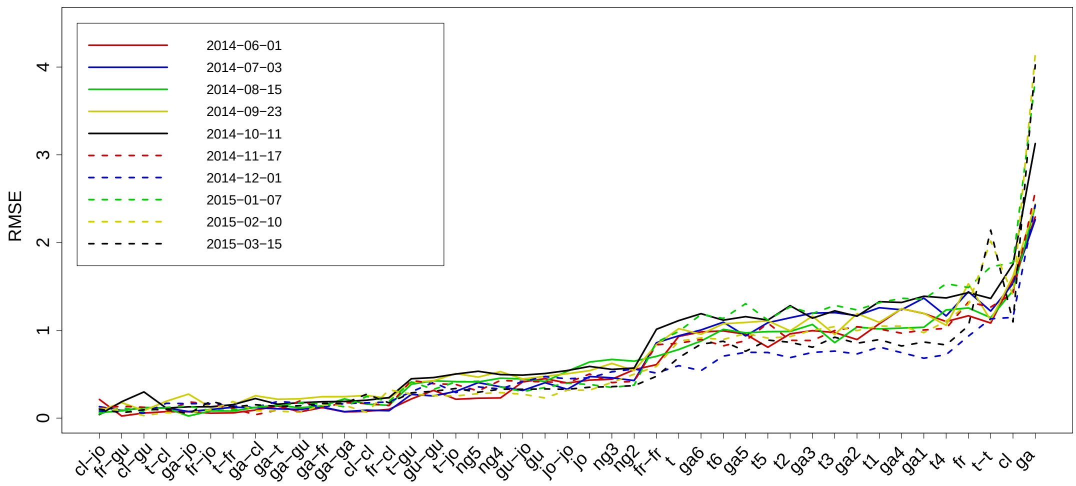

Figure 6: Comparison of RMSEs of 43 models at 10 computation points, 2014-06-01, 2014-07-03, 2014-08-15, 2014-09-23, 2014-10-11, 2014-11-17, 2014-12-01, 2015-01-07, 2015-02-10, 2015-03-15. Abbreviations: ga: Gaussian, t: Student-t, fr: Frank, cl: Clayton, gu: Gumbel, jo: Joe, gai, $i = 1, \ldots, 6$: 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.

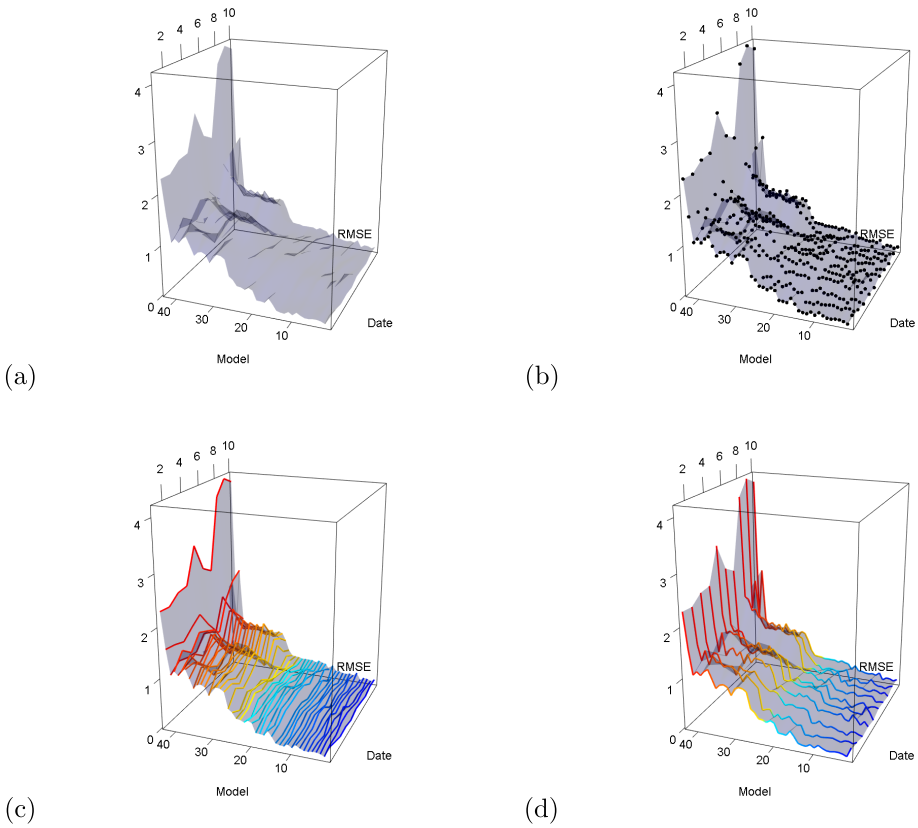

Figure 7: Three dimensional perspective illustrations of 43 Pricing models. (a) The RMSE surface of 43 models. (b) 430 RMSE points from 43 models. (c) 43 cross-sectional lines cutting from the axis of Model, from the best model C_{cl-jo} with the lowest mean of RMSE in deep blue to the worst model C_{ga} with the highest mean of RMSE in deep red. (d) 10 cross-sectional lines cutting from the axis of Date.

5. Conclusion

The goal of this paper is to construct defaults dependence structure mainly with mixed copulae for the CDS index tranche pricing. In this work totally 43 diverse copula models are employed, containing 21 mixed copulae with two component copulae coming from 2 elliptical copulae and 4 Archimedean copulae. At last all computation results are given out based on the RMSE measure. It is found that mixed copulae have superior performance than other copula models. Especially those mixed 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 CDS index tranche pricing. The rest two families, the Archimedean copula family and the nested Archimedean copula family, perform similarly and place in the middle of the ranking.

References

Cherubini, U., Luciano, E., Vecchiato, W., 2004. Copula Methods in Finance. WILEY.
Childs, M., 2013. Synthetic cdos making comeback as yields juiced. Bloomberg.

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., 2001. Modelling dependence with copulas and applications to risk management. Working Paper.

Hofert, J. M., 2010. Sampling nested archimedean copulas with applications to cdo pricing. Dissertation at the Ulm University.

Hofert, M., Scherer, M., 2008. Cdo pricing with nested archimedean copulas. Working Paper.

Hull, J., White, A., 2004. Valuation of a cdo and an nth 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. Distributions with Fixed Marginals and Related Topics.

Joe, H., 1997. Multivariate Models and Dependence Concepts. Chapman & Hall- London.

Li, D. X., 1999. The valuation of basket credit derivatives. CreditMetrics Monitor, 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 (2), 189–204.

Okhrin, O., Ristig, A., 2014. Hierarchical archimedean copulae: The hac package. Journal of Statistical Software 58 (4).

Savu, C., Trede, M., 2009. Hierarchies of archimedean copulas. Quantitative Finance 3 (3).

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

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

Schoenbucher, P., Schubert, D., 2000. Copula-dependent default risk in intensity models. Working Paper.

Sklar, A., 1959. Fonctions de repartition a n dimensions et leurs marges. Publications de l'Institut de statistique de l'Universit de Paris 8, 299–231.