Credit migration and basket derivatives pricing with copulas

Tony Berrada, Debbie Dupuis, Eric Jacquier, Nicolas Papageorgiou, Bruno Rémillard *

Forthcoming Journal of Credit Risk
First version: January 2005
This version: September 2006

Abstract

The multivariate modeling of default risk is a crucial aspect of the pricing of credit derivative products referencing a portfolio of underlying assets, and the evaluation of Value at Risk of such portfolios. This paper proposes a model for the joint dynamics of credit ratings of several firms. Namely, individual credit ratings are modeled by univariate continuous time Markov chain, while their joint dynamic is modeled using copulas. A by-product of the method is the joint laws of the default times of all the firms in the portfolio. The use of copulas allows us to incorporate our knowledge of the modeling of univariate processes, into a multivariate framework. The Normal and Student copulas commonly used in the literature as well as by practitioners do not produce very different estimates of default risk prices. We show that this result is restricted to these two two basic copulas. That is, for any other family of copula, the choice of the copula greatly affects the pricing of default risk.

Key Words: Copula, Markov chain, credit risk, credit rating migration

J.E.L. classification: G10, G20, G28, C16

*Papageorgiou is the corresponding author, nicolas.papageorgiou@hec.ca. The authors are at HEC Montreal and can be reached at firstname.lastname@hec.ca. Funding in partial support of this work was provided by the NSERC, the FQRNT, the FQRSC, and the Institut de Finance Mathématique de Montréal. The paper has benefitted from the comments of a referee and the Finance seminar participants at U. of Liège and Mastricht. We thank Hyung-Seob Kim for computational assistance

1 Introduction

The last few years have witnessed an unprecedented growth in the use of credit derivatives. In addition to single-issuer credit derivatives, e.g., total-return swaps, credit-spread options, and credit-default swaps, there is now significant demand for credit derivative products based on a portfolio of underlying assets. This poses the far more complicated technical challenge of modeling a multivariate process of risky securities, namely their joint default process. A typical example of such a product is an n^{th} to default swap where payout is contingent on the time and identity of the n^{th} default of a credit portfolio.

One needs to model each individual credit rating migration, where default is the worst rating and constitutes an absorbing state, but also their joint evolution. This multivariate problem has been one of the most vexing technical challenges in the credit risk literature. For example [Greg et al., 1997] (CreditMetrics, a sophisticated benchmark for default-risk pricing), use one single Poisson based transition matrix for all the individual securities, and a Gaussian copula to model their joint behavior. Since the individual Markov chains all evolve according to the same transition matrix, the value of the portfolio is determined strictly by the number of bonds initially in each state, without allowing allow these bonds to have different characteristics. This extreme elimination of idiosyncratic default risk is a severe modeling restriction.

One needs the joint distribution of the vector of default times to price n^{th}-to-default swaps. Default times can be modeled with the structural approach as initiated in Merton [1974] or with the reduced form, intensity based, approach, such as in Jarrow and Turnbull [1995]. These pioneering papers model only a single default time. Schönbucher and Schubert [2001] propose a feasible model based upon the reduced form approach, for the multivariate distribution of default times. Hull and White [2005] document the behavior of a stylized copula based models, e.g., with equal pair-wise correlations. Their results could be seen to imply that the copula parameters have little impact on the pricing of n^{th} to default instruments. We will show that this is in fact not the case. An extension of the earlier intensity-based approach for the modeling of individual default is to use each bond's credit rating information and model its evolution by a Markov chain. This is first proposed in Jarrow et al. [1997]. Typically, as is done by these authors, one must study the law of motion under both the objective and martingale measures.

In this paper, we follow the continuous time Markov chain approach initially used to model individual credit ratings. However, our objective is to price credit default swaps or to calculate risk measures for portfolios. So we need to model the joint behavior of multiple credit ratings. We use the recent copula approach to model obvious dependencies among these ratings. This will allow us to model the joint evolution of the credit ratings as a continuous time Markov chain. While copulas have already been proposed, e.g., Li [2000], they are most always of the Gaussian family or Student family. This limited menu of copulas has lead to the belief in the literature that the choice of copula has a minor effect on risk modeling, see e.g., Rosenberg and Schuermann [2004].

The remainder of the paper is structured as follows. Section 2 first gives the main assumptions required to model the joint behavior of multiple credit ratings. The assumptions are motivated and discussed in further details in the Appendix. We then discuss the methodology for pricing credit rating derivatives and calculating measures of risk. We also give the algorithm for the simulation of multiple credit rating trajectories and outline the pricing of an nth to default swap. The model is implemented in Section 3. We show how to estimate 1) the infinitesimal generator of each univariate Markov chain, 2) the price of default risk which links the objective and the risk-neutral measures, 3) the chosen copula and its parameters, and 4) the default premia. Section 4 compares the different copulas with numerical examples and simulations.

2 The Model

2.1 Assumptions on the process of default

Suppose that a portfolio is composed of d risky bonds with respective ratings $X_t^{(1)}, \ldots, X_t^{(d)}$, taking values in the ordered set $\mathcal{S} = \{1, \ldots, m\}$, where 1 is the best rating, and m is the default state. The goal is to model the joint behavior of the stochastic processes $X^{(1)}$, …, $X^{(d)}$ so as to price credit derivatives on a basket of bonds. The assumptions made below are natural in a credit rating migration context. The first assumption concerns the dynamics of the individual ratings while the second relates to their joint distribution.

Assumption 1. *For any $1 \leq k \leq d$, $X^{(k)} = \left(X_t^{(k)}\right)_{t \geq 0}$ is a continuous time Markov chain on S, with infinitesimal generator $\Lambda^{(k)}$, that is*

$$P\left(X_{t+h}^{(k)} = j \middle| X_t^{(k)} = i \right) = \begin{cases} h\Lambda_{ij}^{(k)} + o(h), & j \neq i, \\ 1 + h\Lambda_{ii} + o(h), & j = i, \end{cases} \quad , \quad i, j \in S.$$

Furthermore, for any k, the only absorbing state is state m.

This is standard continuous-time modeling of the individual credit ratings $X^{(k)}$, whereby ratings can change at any time according to a Markov chain. The only absorbing state is default and therefore if one sets $\lambda_i^{(k)} = -\Lambda_{ii}^{(k)}$, Assumption 1 implies that $\lambda_i^{(k)} > 0$ if $i < m$ and $\lambda_m^{(k)} = 0$, for all $1 \leq k \leq d$.

Now denote by τ_k the default time of the k^{th} firm, that is,

$$\tau_k = \inf\left\{ t > 0;\ X_t^{(k)} = m \right\}, \quad k = 1, \ldots, d.$$

By definition of the infinitesimal generator, if $P_{ij}^{(k)}(t) = P\left(X_t^{(k)} = j \middle| X_0^{(k)} = i \right)$, $\quad i, j \in S$, then it follows from Assumption 1 that $\dot{P}^{(k)}(t) = \Lambda^{(k)} P^{(k)}(t)$, $\quad P^{(k)}(0) = I$. Hence $P^{(k)}(t) = e^{t\Lambda^{(k)}}$, $t \geq 0$, and the distribution function $F_i^{(k)}$ of τ_k, given $X_0^{(k)} = i$, is given by

$$F_i^{(k)}(t) = P\left(\tau_k \leq t \,\middle|\, X_0^{(k)} = i \right) = \left(P^{(k)}(t) \right)_{im}, \quad i = 1, \ldots, m, \quad k = 1, \ldots, d.$$

Further remark that $P^{(k)}$ can be written explicitly. In the particular case that $\Lambda^{(k)}$ is diagonalizable, i.e. $\Lambda^{(k)} = M \Delta M^{-1}$ where Δ is a diagonal matrix, $P^{(k)}(t) = M e^{t\Delta} M^{-1}$ and $e^{t\Delta}$ is the diagonal matrix with $\left(e^{t\Delta}\right)_{ii} = e^{t\Delta_{ii}}$, $i = 1, \ldots, m$. Consequently,

$$F_i^{(k)}(t) = \left(P^{(k)}(t)\right)_{im} = 1 + \sum_{j=1}^{m-1} M_{ij} e^{t\Delta_{jj}} \left(M^{-1}\right)_{jm}, \quad t \geq 0.$$

Note that one could also consider non-homogeneous Markov chains, but we believe that the added value is not worth the complications, mainly for estimation.

The second assumption, central to our modeling, is that the joint distribution of the infinitesimal generator can be represented by a copula as follows:

Assumption 2. *The stochastic process $X = \left(X^{(1)}, \ldots, X^{(d)}\right)$ is a Markov chain with state space S^d and infinitesimal generator Λ. Furthermore, Λ is determined by the generators $\Lambda^{(1)}, \ldots, \Lambda^{(d)}$, by a copula C, and a constant $\lambda \geq \max_{1 \leq k \leq d} \max_{1 \leq i \leq m} \lambda_i^{(k)}$, through the following relations:*

$$\Lambda = \Phi\left(\Lambda^{(1)}, \ldots, \Lambda^{(d)}, C, \lambda\right) = \lambda(R - I), \tag{1}$$

where

$$R^{(k)} = I + \Lambda^{(k)}/\lambda, \qquad G_{ij}^{(k)} = \sum_{l=1}^{j} R_{il}^{(k)}, \quad 1 \leq i, j \leq m, \ k = 1, \ldots, d,$$

$G_{i0}^{(k)} = 0$, and

$$R_{\alpha\beta} = P\left(G_{\alpha_1,\beta_1-1}^{(1)} < U_1 \leq G_{\alpha_a,\beta_1}^{(1)}, \ldots, G_{\alpha_d,\beta_d-1}^{(d)} < U_d \leq G_{\alpha_d,\beta_d}^{(d)}\right), \tag{2}$$

for any $\alpha, \beta \in S^d$,where $U = (U_1, \ldots, U_d)$ has distribution function C.

It is important to note that (2) allows for multiple simultaneous credit rating changes. One may worry that many firms could default at exactly the same time, a phenomenon rarely observed. We would require the independence of the Markov chains to allow only one default at one time, see Theorem 5 in the appendix. However this is not a concern, rather a flexibility of the modeling. The application to typical data in Section 3.3 will show that the clustering effect is quite mitigated.

Finally, note that if Y is a discrete time Markov chain on S^d with transition matrix R, one can write

$$X_t = Y_{N_t}, \ t \geq 0 \tag{3}$$

where N_t is a Poisson process with intensity λ. That link between continuous and discrete-time modeling will be used for simulating continuous-time Markov chains. The Appendix provides further discussion of Assumption 2.

The next two assumptions characterize the law of the credit rating processes $X^{(1)}, \ldots, X^{(d)}$, under the risk neutral measure $\mathcal{P}$. They essentially state that the distribution of ratings is preserved in the passage to the risk-neutral measure. First, consider the univariate marginal distributions of credit ratings.

Assumption 3. *Under the risk neutral measure $\mathcal{P}$, the process $X^{(k)}$ is a homogeneous continuous time Markov chain with infinitesimal generator $\mathcal{L}^{(k)} = \sigma_k \Lambda^{(k)}$.*

Here, the parameter σ_k is the price of risk for the law of default time. Following standard practice for fixed-income derivative securities, e.g., Jarrow et al. [1997], the risk premium σ_k can be calibrated directly from market data, for example by fitting market prices of credit default swaps.

It follows from Assumption 3 that, if one sets

$$\mathcal{P}_{ij}^{(k)}(t) = \mathcal{P}\left(X_t^{(k)} = j \middle| X_0^{(k)} = i\right), \quad i, j \in S,$$

then

$$\mathcal{P}^{(k)}(t) = e^{t\mathcal{L}^{(k)}} = e^{t\sigma_k \Lambda^{(k)}} = P^{(k)}(\sigma_k t), \quad t \geq 0.$$

This is because the equation $\dot{\mathcal{P}}^{(k)}(t) = \mathcal{L}^{(k)}\mathcal{P}^{(k)}(t), \quad \mathcal{P}^{(k)}(0) = I$ has a unique solution. In particular, under the risk neutral measure $\mathcal{P}$, the distribution function $\tilde{F}_i^{(k)}$ of the default time τ_k, given $X_0^{(k)} = i$, is given by

$$\tilde{F}_i^{(k)}(t) = \mathcal{P}\left(\tau_k \leq t \middle| X_0^{(k)} = i\right) = \left(\mathcal{P}^{(k)}(t)\right)_{im} = F_i^{(k)}(\sigma_k t), \quad i = 1, \ldots, m, \quad k = 1, \ldots, d.$$

Finally, the fourth Assumption states that the cross-sectional dependence structure of the infinitesimal generator is also preserved under the change of measure.

Assumption 4. *Under the risk neutral measure $\mathcal{P}$, the process X is a continuous time Markov chain with infinitesimal generator $\mathcal{L} = \Phi\left(\mathcal{L}^{(1)}, \ldots, \mathcal{L}^{(d)}, C, \tilde{\lambda}\right)$, where $\tilde{\lambda} \geq \max_{1\leq k\leq d} \max_{1\leq i\leq m} \tilde{\lambda}_i^{(k)}$, and $\tilde{\lambda}_i^{(k)} = \sigma_k \lambda_i^{(k)}$, $k = 1, \ldots, d$.*

That is, similarly to equation (1), we write for the risk-neutral infinitesimal generator that:

$$\mathcal{L} = \tilde{\lambda}(\mathcal{R} - I),$$

where

$$\mathcal{R}_{\alpha\beta} = P\left(\mathcal{G}_{\alpha_1,\beta_1-1}^{(1)} < U_1 \leq \mathcal{G}_{\alpha_a,\beta_1}^{(1)}, \ldots, \mathcal{G}_{\alpha_d,\beta_d-1}^{(d)} < U_d \leq \mathcal{G}_{\alpha_d,\beta_d}^{(d)}\right),$$

$\alpha, \beta \in S^d$, and $U = (U_1, \ldots, U_d)$ has distribution function C, where $\mathcal{R}^{(k)} = I + \mathcal{L}^{(k)}/\tilde{\lambda}$, and

$$\mathcal{G}_{ij}^{(k)} = \sum_{l=1}^{j} \mathcal{R}_{il}^{(k)}, \quad 1 \leq i, j \leq m.$$

Note that, as stated under the historical measure, the Markov chain X can be written under the risk neutral measure $\mathcal{P}$ as

$$X_t = \tilde{Y}_{\tilde{N}_t},$$

where $\tilde{N}_t$ is a Poisson process with intensity $\tilde{\lambda}$ and $\tilde{Y}$ is a discrete Markov chain on S^d with transition matrix $\mathcal{R}$.

2.2 Simulation of multiple credit ratings trajectories

Given a model for the joint behavior of the credit rating processes $X^{(1)}, \ldots, X^{(d)}$, it is straightforward to use Monte Carlo methods to compute prices of basket credit derivatives. We now show how to simulate multiple credit ratings trajectories, and in Section 2.3 how to price n^{th} to default swaps.

Given the parameters, one can easily describe the technique proposed to generate multiple credit ratings trajectories. Simulations under the risk-neutral and the objective measures are similar as a result of Assumptions 3 and 4. Hence, we only describe simulation under the objective measure. The inputs needed for simulation are: (1) the initial vector of credit ratings X_0 and (2) the discrete-time transition matrix R.

Recall relation (3): $X_t = Y_{N_t}$, for $t \geq 0$. Given an initial state, the transition to the next state is modeled by a Poisson process N where the intensity vector λ is a row of the transition matrix R. Since the Poisson process N only has a finite number of jumps up to time T, we can simulate the whole trajectory of X, after X_0 and up to X_T. The steps are:

1. Set $T_0 = 0$ and generate $n = N_T$ arrival times $T_1, \ldots, T_n$ of the Poisson process N of intensity λ, up to time T.

2. Suppose that the constant λ, the matrices $G^{(1)}, \ldots, G^{(d)}$, and the copula C are given as in Assumption 2. Then, for each $i = 1, \ldots, n$, generate $U_i = \left(U_i^{(1)}, \ldots, U_i^{(m)}\right) \sim C$. For a fixed $1 \leq l \leq m$, if $X_{T_{i-1}}^{(l)} = j$, then

$$X_{T_i} = k \qquad \text{if } G_{j,k-1}^{(l)} < U_i^{(l)} \leq G_{j,k}^{(l)}, \qquad j, k \in S.$$

Having generated multivariate trajectories, one can now price contracts on baskets of securities and VAR of portfolios, based on the expectations of trajectories.

2.3 Pricing of credit default swaps

The main class of contract we want to price is the nth to default swap. Following Mashal and Naldi [2001], the contract can be described as follows:

- The contract starts at time $t = 0$ and matures at time $t = T$.
- The notional value of the contract is N.
- $\tau_{(i,d)}$, $i = 1, \ldots, d$ is the i-th shortest default time and $\tau_{(1,d)} \leq \tau_{(2,d)} \leq \cdots \leq \tau_{(d,d)}$.
- p_f is the percentage yearly premium and it is paid f times a year (at the end of each period), so the net amount due each period is $p_f N/f$, until default.
- RR_j is the recovery rate for name j, $j = 1, \ldots, d$, and $RR_{n,d}$ stands for the recovery rate of the n-th to default. These rates are assume to be predictable.

- a is the accrued premium (in percentage), i.e. Na is the amount the insurance holder owes the insurance seller since the last payment, until τ_n, provided $\tau_n \le T$.
- At time T, if $\tau_{n,d} \le T$, the insurance is triggered and the insurance payment is $N(1-RR_{n,d}-a)$.
- r_t is the risk free instantaneous interest rate at time t and $\beta(t) = e^{-\int_0^t r_s ds}$.

Under the risk neutral probability measure $\mathcal{P}$, the value of the yearly premium p_f verifies

$$E_{\mathcal{P}}\left[N\frac{p_f}{f}\sum_{i=1}^{fT}\beta(i/f)\mathbb{I}\left(\tau_{n,d} > \frac{i}{f}\right) - \beta(\tau_{n,d})N(1-RR_{(n)}-a)\mathbb{I}\left(\tau_{n,d}\le T\right)\right] = 0.$$

Note that if $\frac{i-1}{f} < \tau_{n,d} \le \frac{i}{f}$, then $\beta(\tau_{n,d})a = \beta(i/f)\frac{p_f}{f}$, so

$$\begin{aligned} p_f &= f\,\frac{E_{\mathcal{P}}\left\{\beta(\tau_{n,d})(1-RR_{n,d})\mathbb{I}(\tau_{n,d}\le T)\right\}}{E_{\mathcal{P}}\left\{\sum_{i=1}^{fT}\beta(i/f)\mathbb{I}\left(\tau_{n,d} > \frac{i}{f}\right) + \sum_{i=1}^{fT}\beta(i/f)\mathbb{I}\left(\frac{i-1}{f} < \tau_{n,d} \le \frac{i}{f}\right)\right\}} \\ &= f\,\frac{E_{\mathcal{P}}\left\{\beta(\tau_{n,d})(1-RR_{n,d})\mathbb{I}(\tau_{n,d}\le T)\right\}}{E_{\mathcal{P}}\left\{\sum_{i=1}^{fT}\beta(i/f)\mathbb{I}\left(\tau_{n,d} > \frac{i-1}{f}\right)\right\}}. \end{aligned}$$

In particular, if the interest rate, the default times and the recovery rates are independent, the above formula reduces to

$$p_f = f\,(1-\overline{RR})\,\frac{B(T)F_{n,d}(T) - \int_0^T \partial_t B(t)F_{n,d}(t)dt}{\sum_{i=1}^{fT} B\left(\frac{i}{f}\right)\bar{F}_{n,d}\left(\frac{i-1}{f}\right)}, \tag{4}$$

where, under $\mathcal{P}$, $B(t)$ is the value of a zero coupon bond with maturity t, $\overline{RR}$ is the mean recovery rate, $F_{n,d}$ is the distribution function of the n-th default time, and $\bar{F}_{n,d} = 1 - F_{n,d}$ is the associated survival function.

In the special case of a single debt, i.e. $m = 1$, the contract is simply called a credit default swap (CDS). Its value q_f is thus given by

$$q_f = f\,\frac{E_{\mathcal{P}}\left\{\beta(\tau)(1-RR)\mathbb{I}(\tau\le T)\right\}}{E_{\mathcal{P}}\left\{\sum_{i=1}^{fT}\beta(i/f)\mathbb{I}\left(\tau > \frac{i-1}{f}\right)\right\}},$$

which simplifies, under independence, to:

$$q_f = f\,(1-\overline{RR})\,\frac{B(T)F(T) - \int_0^T \partial_t B(t)F(t)dt}{\sum_{i=1}^{fT} B\left(\frac{i}{f}\right)\bar{F}\left(\frac{i-1}{f}\right)}, \tag{5}$$

where under $\mathcal{P}$, $\overline{RR}$ is the mean recovery rate, F is the distribution function of the default time, and $\bar{F} = 1 - F$ is the associated survival function.

Remark 1. Since B is non increasing, $\partial_t B(t) \le 0$. It is thus easy to check from equations (4) and (5) that p_f, as of function of $F_{n,d}$, and q_f, as a function of F, are non decreasing. This means that if $F_{n,d}(t) \le G_{n,d}(t)$ for all $0 \le t \le T$, then $p_f(F_{n,d}) \le p_f(G_{n,d})$. Similarly, if $F(t) \le G(t)$ for all $0 \le t \le T$, then $q_f(F) \le q_f(G)$.

In particular, if $G(t) = F(\sigma t)$, then $q_f(\sigma)$ is monotone increasing in σ, so given a premium q and F, there is a unique " implicit" σ so that $q_f(\sigma) = q$.

Note that the continuous time limit $f \to \infty$ also exists. This means that premium would be paid continuously. It follows that

$$p_\infty = \lim_{f\to\infty} p_f = \frac{E_{\mathcal{P}}\left\{\beta(\tau_{n,d})(1 - RR_{n,d})\mathbb{I}(\tau_{n,d} \le T)\right\}}{E_{\mathcal{P}}\left\{\int_0^{T\wedge\tau_{n,d}} \beta(t)dt\right\}}$$

and

$$q_\infty = \lim_{f\to\infty} q_f = \frac{E_{\mathcal{P}}\left\{\beta(\tau)(1 - RR)\mathbb{I}(\tau \le T)\right\}}{E_{\mathcal{P}}\left\{\int_0^{T\wedge\tau} \beta(t)dt\right\}},$$

where $x \wedge y = \min(x, y)$. Numerical results for n-th to default swaps are provided in Section 4.

3 Implementation

We now show how to estimate the parameters necessary to implement the model and describe the data used.

To estimate the infinitesimal generator of the simple Markov chains, we use the method proposed in Lando and Skødeberg [2002]. We outline the procedure in Section 4.1. Credit default swaps (CDS) prices are used to estimate the risk neutral parameters σ_k, required to perform credit derivatives pricing. The corresponding returns are used to estimate the copula parameters, based on the pseudo-likelihood method of Genest et al. [1995].

To implement the proposed methodology, one first needs to estimate the matrices $\Lambda^{(1)}, \ldots \Lambda^{(d)}$. To do this we use the Moody's credit rating for all corporate risky bonds from April 26, 1982 to January 7, 2004. For sake of completeness and to remove possible biases, we include all the 23 credit classes: Aaa (=1), Aa1, Aa2, Aa3, A1, A2, A3, Baa1, Baa2, Baa3, Ba1, Ba2, Ba3, B1, B2, B3, Caa, Caa1, Caa2, Caa3, Ca, C, and D (=23). Withdrawals (WR) corresponding to recall are treated as all withdrawals corresponding to maturity. The observations thus consist of all transitions from one class to another, and the length of time between these transitions.

3.1 Estimation of the historical transition matrices

For simplicity, we assume that the d firms all have the same historical transition probabilities, that is, $\Lambda^{(1)} = \cdots = \Lambda^{(d)}$. With enough data, one could of course relax this assumption. In practice, it is often needed so as to have enough transitions from state to state in the data.

If N is the total number of bonds in the database, let $N_{ij}^{(k)}$ denotes the total number of transitions from state i to state j for bond k, $1 \le i \le m$, $0 \le j \le m, j \ne i$, $1 \le k \le N$. Then

$N_{ij} = \sum_{k=1}^{N} N_{ij}^{(k)}$ is the total number of transitions from state i to state j, $1 \le i \le m$, $0 \le j \le m$, reported in the database, and $N_i^{(k)} = \sum_{j \neq i} N_{ij}$ is the total number of transitions starting from state i, in the database.

Further let $L_i^{(k)}$ be the total occupation time of state i for bond k, that is, the total time the bond k spent in state i for the period considered, $1 \le i \le m$, $1 \le k \le N$. Set

$$L_i = \sum_{k=1}^{N} L_i^{(k)}, \quad 1 \le i \le m.$$

Note that when constructing the dataset, a transition from any state (say i) to the state WR (corresponding to maturity of the bond or its recall) is not taken into account, but the time spent in state i prior to moving to such a WR must be taken into account, when calculating L_i.

Because a full maximum likelihood estimation would be quite complex, even impossibly difficult due to the dependence structure and the small number of transitions for a given portfolio, we use the "cohort" method described in Lando and Skødeberg [2002] to estimate $\Lambda^{(1)}$. Namely, the estimator of $\Lambda_{ij}^{(1)}$ is given by

$$\widehat{\Lambda_{ij}^{(1)}} = \frac{N_{ij}}{L_i}, \qquad 1 \le i \le m, 0 \le j \le m, j \neq i.$$

It follows that

$$\widehat{\lambda_i^{(1)}} = \frac{N_i}{L_i}, \; 1 \le i \le m.$$

Hence one can take

$$\hat{\lambda} = \max_{1 \le i \le m} \widehat{\lambda_i^{(1)}} = \max_{1 \le i \le m} \frac{N_i}{L_i}.$$

Table 1 shows the resulting discrete-time generator of the transition matrix. Recall that $e^{-\Lambda t}$ is the transition probability. The negative values on the diagonal correspond to a very high probability of no migration in credit. The zeros far from the diagonal represent the absence of a jump across many credit ratings in the data base.

Table 1: $100\hat{\Lambda}$ based on Moody's ratings, all corporate risky bonds Apr. 26, 1982 to Jan. 7, 2004.

-15.66	11.36	3.53	0.23	0.38	0.08	0.08	0	0	0	0	0
3.14	-30.84	13.59	11.64	1.44	0.68	0	0.17	0	0	0.08	0
1.02	4.22	-25.68	13.82	4.96	1.11	0.28	0.09	0.09	0	0	0
0.15	0.69	6.00	-28.02	15.50	4.64	0.63	0.09	0.15	0.03	0	0
0.04	0.13	0.58	7.17	-25.09	11.81	4.17	0.47	0.25	0.04	0.18	0.16
0.02	0.08	0.36	1.27	7.48	-28.11	13.45	3.99	0.87	0.23	0.08	0.11
0.10	0.08	0.10	0.30	1.50	10.90	-34.59	12.58	6.98	1.48	0.15	0.18
0.06	0.16	0.22	0.26	0.29	2.75	9.96	-36.39	15.42	4.98	1.12	0.48
0.03	0.18	0.09	0.25	0.37	1.01	4.94	8.58	-37.53	15.70	4.48	0.86
0.04	0.04	0.04	0.04	0.28	0.55	1.07	3.80	12.95	-43.69	13.86	6.69
0.04	0.04	0.04	0.04	0.23	0.23	0.46	0.96	3.87	11.87	-39.96	9.26
0	0	0.04	0.04	0.04	0.22	0.13	0.58	0.75	2.75	11.88	-41.88
0	0.03	0.03	0	0.03	0.19	0.14	0.28	0.33	0.64	3.26	7.15
0.03	0	0.03	0	0.10	0.15	0.13	0.23	0.23	0.18	0.64	3.10
0	0	0.05	0.09	0.05	0	0.18	0.18	0.23	0.37	0.46	1.05
0	0.15	0.10	0	0.05	0.05	0.10	0.10	0.10	0.30	0.25	0.45
0	0	0	0	0	0	0	0	0.30	0.60	0.30	0.30
0	0	0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0	0.73	0
0	0	0	0	0	0	0.59	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0	0	0

0	0	0	0	0	0	0	0	0	0	0
0.08	0	0	0	0	0	0	0	0	0	0
0.05	0	0.05	0	0	0	0	0	0	0	0
0	0.09	0.03	0	0	0	0	0	0	0	0
0.04	0.04	0	0	0	0	0	0	0	0	0
0.06	0.04	0.02	0.02	0	0	0	0	0	0	0
0.13	0.10	0	0	0	0	0	0	0	0	0.03
0.41	0.22	0.06	0	0	0	0	0	0	0	0
0.46	0.25	0.25	0.06	0	0	0	0	0	0	0.03
2.38	1.11	0.40	0.08	0.08	0	0.20	0.04	0	0	0.04
7.92	2.33	1.72	0.23	0.08	0	0	0.11	0.04	0.04	0.46
13.43	4.87	5.10	1.60	0	0	0.04	0.09	0.04	0	0.27
-36.00	10.29	8.65	3.67	0.14	0.31	0.06	0.08	0.03	0	0.70
7.43	-38.29	15.07	7.25	0.56	1.23	0.46	0.08	0.13	0	1.26
2.89	11.14	-53.87	20.86	2.25	7.61	2.29	0.78	0.96	0.05	2.38
0.85	4.03	7.06	-53.89	6.02	9.74	5.52	3.53	2.54	0.15	12.83
1.21	2.41	2.41	6.03	-68.11	6.63	4.22	1.21	9.64	0.60	32.25
0.22	0.66	1.76	10.55	0	-91.89	19.79	24.18	14.51	1.54	18.69
1.32	0	2.21	3.09	0	7.06	-114.77	18.54	16.33	10.59	55.62
0	2.18	0.73	2.18	0	2.91	8.72	-150.47	50.16	13.08	69.78
0.59	0	1.19	2.97	1.78	0	5.94	6.54	-129.00	12.48	96.90
0	0	0	0	0	0	0	0	7.62	-224.74	217.12
0	0	0	0	0	0	0	0	0	0	0

Λ is the generator for the transition probability matrix $e^{\Lambda t}$

3.2 Estimation of the price of risk $\sigma^{(k)}$

Suppose one has prices $Y_1 = q_{f,x_1,r_1,T}(\sigma), \ldots, Y_n = q_{f,x_n,r_n,T}(\sigma)$ of credit defaults swaps of a firm for a fixed T where x_i and r_i denote the credit rating of the firm and the interest rate at time i, $1 \le i \le n$, respectively. Since $\Lambda^{(1)}$ is now assumed to be known, the distribution functions $F_1, \ldots, F_m$ are also known. Hence one can use formula (4) to estimate σ. For example, one could choose σ so as to minimize $\sum_{i=1}^{n} \left[Y_i - q_{f,x_i,r_i,T}(\sigma)\right]^2$.

Note that if the credit rating and the interest rate do not change over the period considered, then $q_j(\sigma) = \bar{Y}$, so σ can be found explicitly since $q_f(\sigma)$ is monotone increasing and has range $(0, \infty)$, see Remark 1 in section 2.3.

We use seven firms : Bank One Corp., Bear Stearns Companies Inc,. Goldman Sachs Group Inc., Lehman Brothers Holdings Inc., Merrill Lynch & Co., American Express Co., and Countrywide Home Loans Inc. We collected the weekly prices (in basis points) of credit default swaps for a maturity of $T = 5$ years, from Jan. 7, 2000, to Feb. 2, 2004, where payments were made twice a year, i.e. $f = 2$. For simplicity, we assume a constant interest rate $r = 2\%$ and a mean recovery rate $\overline{RR} = 50\%$, since the worst rating of the group is A3. Note that all firms remained in the same credit rating, with the exception of Lehman Brothers Holdings which switched from A3 to A2 after 43 weeks.

Table 2 shows the estimates of the coefficients $\sigma_1, \ldots, \sigma_7$ and the credit ratings at the end of the period. The prices of risk range from 2.9 to 4.1, most of them being around 3.5, which is expected since these specific firms are large high grade issuers whose risk characteristics do not vary much.

Table 2: Ratings and estimates of σ for seven stocks

Firm	Bank One	Bear Stearns	Goldman Sachs	Lehman Bros	Merrill Lynch	American Express	Countrywide Home Loans
Rating	Aa3	A2	Aa3	A2	Aa3	A1	A3
State	4	6	4	6	4	5	7
$\hat{\sigma}$	3.62	3.22	4.16	3.26	4.36	3.67	2.93

Ratings on Feb. 2, 2004. σ is the price of default risk expressed basis points. Estimates based on data from Jan. 7, 2000 to Feb. 2, 2004. Other parameters are set to $\overline{RR} = 50\%$, $r = 2\%$, $T = 5$, and $f = 2$

3.3 Estimation of the copula

Recall that a d-dimensional copula $C(u_1, \ldots, u_d)$ is a joint distribution function of uniformly distributed random variables $U_1, \ldots, U_d$. Consider first the well-known Gaussian copula. Let $Y = (Y_1, \ldots, Y_d) \sim N_d(0, \Sigma)$, i.e. Y is a d-dimensional Gaussian distribution with mean zero and covariance matrix Σ, where $\Sigma_{ii} = 1$ for all $1 \leq i \leq d$, then the Gaussian copula C_Σ is defined as

$$C_\Sigma(u_1, \ldots, u_d) = P\left\{N(Y_1) \leq u_1, \ldots, N(Y_d) \leq u_d\right\}, \quad u_1, \ldots, u_d \in [0, 1],$$

where $N(\cdot)$ is the distribution function of the standard Gaussian distribution. Note that as $\Sigma_{ii} = 1$ for all $1 \leq i \leq d$, Σ is also a correlation matrix.

As a second example, consider the closely related Student copula. If $Y \sim N_d(0, \Sigma)$ as above and V is an independent variable with a chi-square distribution with ν degrees of freedom, then $Y/\sqrt{V/\nu}$ has a multivariate Student distribution with parameters Σ and ν. The Student copula $C_{\Sigma,\,\nu}$ is then defined as the joint distribution function $U_i = T_\nu\left(Y_i/\sqrt{V/\nu}\right)$, $i = 1, \ldots, d$, where $T_\nu(\cdot)$ is the distribution function of a Student distribution with ν degrees of freedom.

Another popular family of copulas is the Archimedean family defined by Genest and MacKay [1986]. A copula is said to be Archimedean when it can be expressed as

$$C(u_1, \ldots, u_d) = \phi^{-1}\left\{\phi(u_1) + \cdots + \phi(u_d)\right\},$$

where $\phi : (0, 1] \to [0, \infty)$, is a bijection such that $\phi(1) = 0$ and

$$(-1)^i \frac{d^i}{dx^i}\, \phi^{-1}(x) > 0, \quad 1 \leq i \leq d.$$

Table 3 gives the generators for three well-known Archimedean copulas: the Clayton, Frank, and Gumbel copulas. These three classes share the interesting property that ϕ, denoted the generator of the copula, is unique, up to a constant. See Joe [1997] and Nelsen [1999] for further examples on copulas.

Table 3: Archimedean copulas, parameter estimates for seven firms

Model	$\phi_\theta(t)$	Range of θ	$\hat{\theta}$
Clayton	$\frac{t^{-\theta}-1}{\theta}$	$(0,\infty)$	0.43
Frank	$\log\left(\frac{1-e^{-\theta}}{1-e^{-\theta t}}\right)$	$(0,\infty)$	2.78
Gumbel-Hougaard	$\lvert\log t\rvert^{1/\theta}$	$(0,1)$	0.77

Due to a lack of data on credit migration, the estimation of copula parameters must rely on proxies. This approach is validated by the property that the copula corresponding to a random vector $Y = \left(Y^{(1)},\dots,Y^{(d)}\right)$ is invariant for monotone increasing transformations of its components $Y^{(k)}$. That is, the copulas of Y and $\left(\psi_1\left(Y^{(1)}\right),\dots,\psi_1\left(Y^{(d)}\right)\right)$ are the same whenever the ψ_k's are strictly increasing transformations.

If Credit default Swaps (CDS) prices are available, the associated returns could serve as proxies to estimate the parameters of the copula and to choose the "right" copula family. If no reasonably large data set of CDS prices is available, one could use the daily log-returns of the corresponding companies as proxies, under the assumption that the higher the return, the higher the default time, and the higher the credit rating. There is presently no empirical evidence to discourage this assumption, and in fact, Mashal et al. [2003] favor the practice.

We proceed as follows. Given returns $Y_i = \left(Y_i^{(1)},\dots,Y_i^{(d)}\right)$, $i = 1,\dots,N$, one first calculates ranks $\mathrm{Rank}\left(Y_i^{(k)}\right)$ for each fixed k. That is, $\mathrm{Rank}\left(Y_i^{(k)}\right)$ represents the rank of $Y_i^{(k)}$ within $Y_1^{(k)},\dots,Y_N^{(k)}$, with rank one assigned to the smallest number. Next, one defines the pseudo-observations U_i as follows:

$$U_i = \left(U_i^{(1)},\dots,U_i^{(d)}\right) = \frac{1}{N+1}\left(\mathrm{Rank}\left(Y_i^{(1)}\right),\dots,\mathrm{Rank}\left(Y_i^{(d)}\right)\right),\ i = 1,\dots,N.$$

If c_θ represents the density of the copula, the estimation of the parameter θ, possibly multi-dimensional, is defined as the value $\hat{\theta}$ maximizing

$$\arg\max_\theta \sum_{i=1}^{N} \log\left\{c_\theta(U_i)\right\}. \tag{6}$$

Note that (6) still makes sense even if the time independence of returns is not satisfied. Serial independence is used mainly to calculate the estimation error, see Genest et al. [1995] for additional details.

For the Gaussian copula, there is an explicit expression for the correlation matrix $\hat{\Sigma}$. Each component $\hat{\Sigma}_{jk}$ is the so-called van der Waerden correlation coefficient, defined as the Pearson correlation coefficient of the pairs $\left\{N^{-1}\left(U_i^{(j)}\right), N^{-1}\left(U_i^{(k)}\right)\right\}$, $i = 1, \ldots, n$.

For our data set, the estimation of Σ for the Gaussian copula model is given by

$$\hat{\Sigma} = \begin{pmatrix} 1.00 & & & & & & \\ 0.43 & 1.00 & & & & & \\ 0.42 & 0.48 & 1.00 & & & & \\ 0.38 & 0.59 & 0.55 & 1.00 & & & \\ 0.40 & 0.50 & 0.47 & 0.54 & 1.00 & & \\ 0.31 & 0.29 & 0.25 & 0.28 & 0.25 & 1.00 & \\ 0.35 & 0.34 & 0.37 & 0.33 & 0.37 & 0.35 & 1.00 \end{pmatrix}.$$

Estimating the parameters Σ and ν of the more general Student copula is a little bit trickier. It is recommended to follow a two-step procedure: first estimate Σ, then estimate ν, using the pseudo-maximum-likelihood method (6). It is known, e.g., Abdous et al. [2004], that $\Sigma_{jk} = \sin(\tau_{jk}\pi/2)$, where τ_{jk} is the theoretical Kendall's tau between components $Y^{(j)}$ and $Y^{(k)}$. If τ_{jk} is an estimate of τ_{jk}, then $\sin(\hat{\tau}_{jk}\pi/2)$ is an estimate of Σ_{jk}. However, the new estimated matrix $\hat{\Sigma}_0$ is not necessarily a correlation matrix. It can be transformed into a correlation matrix $\hat{\Sigma}$ in the following way: take the symmetric square root $\hat{\Sigma}_1$ of the matrix $\hat{\Sigma}_0\hat{\Sigma}_0^\top$, then form the diagonal matrix Δ with components $\Delta_{jj} = \sqrt{\left(\hat{\Sigma}_1\right)_{jj}}$. Finally set, $\hat{\Sigma} = \Delta^{-1}\hat{\Sigma}_1\Delta^{-1}$.

Using that method with our data set yields an estimated correlation matrix for the Student copula given by

$$\hat{\Sigma} = \begin{pmatrix} 1.00 & & & & & & \\ 0.50 & 1.00 & & & & & \\ 0.49 & 0.56 & 1.00 & & & & \\ 0.43 & 0.66 & 0.62 & 1.00 & & & \\ 0.46 & 0.56 & 0.53 & 0.58 & 1.00 & & \\ 0.42 & 0.36 & 0.33 & 0.33 & 0.31 & 1.00 & \\ 0.44 & 0.39 & 0.44 & 0.38 & 0.43 & 0.49 & 1.00 \end{pmatrix},$$

with an estimate of $\hat{\nu} = 3.7$ degrees of freedom.

The parameter estimates for the three Archimedean copulas are listed in Table 3. Note that, while θ varies greatly with the specific copula, it does not have the same interpretation, as even the ranges of possible values differ.

3.3.1 Differences between copulas

To illustrate the contagion effect (the clustering of defaults), we simulated 100,000 five-year periods of joint ratings of the seven firms. Simulation is under the risk neutral dynamics, with Clayton and Gumbel copulas. Parameter values were set to previously fitted values, i.e., in the calculation of n-th to defaults swaps. There were 33,600 periods with at least one default (33,910 defaults total) for the Clayton copula, and 30,702 default times (33,106 defaults total) for the Gumbel. Note that 33.9% or 33.1% do not represent probability of defaults since the simulation is done under the risk neutral dynamics. The effect of the coefficients $\sigma_1, \ldots, \sigma_7$ is to accelerate time. For example, if they were all equal to 5, then the percentages would reflect the probability of defaults in a 25-year period instead of a 5-year period. In the present case, they are all of order 3 so 33.91% represents the probability of at least one default in the seven firms over a 15-year period for the Clayton model. Table 4 summarizes the results.

Table 4: Risk Neutral Probability of simultaneous defaults for two copulas

Copula	1	2	3	4	5	6	7
Clayton	99.1	0.91	0.01	0	0	0	0
Gumbel	93.8	5.05	0.844	0.26	0.1	0.01	0

100,000 periods of 5 years simulated for the 7 firms, with parameters in Table 2 and 3. Probabilities of n (1 to 7) defaults over one 5-year period in %.

Table 4 confirms that these copulas do not produce massive clustering of default. However it also shows that the Gumbel copula produces significantly more clustering than the Clayton. Clearly, once one departs from the Normal and Student copulas, very different patterns of dependencies can be simulated. The clustering of the defaults is tied to the upper-tail dependence of the fitted copula. We see for example in Table 4 that the Gumbel's larger upper-tail dependence than the Clayton leads to higher clustering of default. More on upper-tail dependence can be found in Joe [1997].

This shows that the choice of the appropriate family of copula can be very important. While formal statistical testing is beyond the scope of this paper, the above results show that it is an important issue for future research. Indeed, formal tests have only appeared recently in the literature. See Genest et al. [2006] for discussions of formal tests, and Dupuis et al. [2006] for an application to equity and CDS prices.

3.3.2 Pricing nth to default swaps

After estimating the parameters, one can now estimate the premiums p_f for different copula families. We computed the prices, in basis points, by Monte Carlo methods using 100,000 repetitions. They are reported in Table 5.

Table 5: Premia for the n^{th} to default, basket of seven firms.

	Order of default						
Model	1^{st}	2^{nd}	3^{rd}	4^{th}	5^{th}	6^{th}	7^{th}
Clayton	151	25.2	3.5	0.4	0.03	0.00	0.00
Frank	150	29.0	5.1	0.8	0.12	0.00	0.00
Gumbel	111	35.9	15.7	7.5	3.88	2.01	0.73
Gaussian	137	32.3	8.6	2.1	0.50	0.11	0.01
Student	110	39.3	17.0	7.5	3.15	1.06	0.24

Default premia in basis points.

Note that the choice of copula also has a large impact on the prices. For the 1^{st} to default, the largest price is obtained for the Clayton family, and the smallest for the Student family. These roles are reversed for the 2^{nd} to default price. It shows that a copula may dominate for a given n^{th} to default, but not necessarily for all. We discuss further numerical examples in the next section.

4 Properties of the Model and Concluding Remarks

To illustrate the resulting effect of the copula and its dependence parameters on the prices of n^{th} to default swaps, we consider two portfolios. The first one is made of 10 firms with the following high-grade credit ratings: Aaa Aaa, Aa1, Aa1, Aa2, Aa2, Aa3, A1, A2, A3. We assume here that the σ_k's are all equal to 3 basis points, a reasonable value given the estimates of Table 2. The second portfolio is made of 10 low-grade risky bonds with the following ratings: Caa1, Caa1, Caa2, Caa2,Caa3, Caa3, Ca, Ca, C, C. We assume σ_k's of 5.

To study the effect of copulas beyond the relatively commonly Gaussian and Student, we add the Clayton, Frank, and Gumbel families described earlier. For this simulation we will use the equi-correlated Gaussian and Student families (i.e. $\Sigma_{jk} = \rho$ for all $j \neq k$), similar to the restriction in Hull and White [2005]. We set $\nu = 3.7$ for the Student, the fitted value in our real data example in Section 3.2, taken as plausible behavior. As the value of ν does not affect the value of Kendall's τ for the Student, the five families are then indexed by one parameter, τ. Note that $\tau = 0$ corresponds to the independence copula $C(u_1, \ldots, u_d) = u_1 \cdots u_d$, while $\tau = 1$ corresponds to the Fréchet copula $C(u_1, \ldots, u_d) = \min(u_1, \ldots, u_d)$. We used $\tau = j/20$, for $j = 0, \ldots, 20$. We use the generator matrix estimated from the Moody's data set, shown in Table 1, and the values $f = 2$, $r = 2\%$ and $T = 5$ for all the simulations. For each copula and value of τ, we make 100,000 draws of the vector of survival times, to estimate the premia for 1^{st} to 10^{th} to default credit swap instruments.

Figures 1 to 5 show plots of the premia in basis points versus Kendall's τ for the various copulas and orders of default. Figure 1 considers 1^{st} to default. The bottom plot shows that the different copulas produce very similar prices for the swap on the risky securities. For these risky assets, the probability of even only 1 default out of 10 is very high and the effect of the dependence structure is minimal. In contrast, the top plot shows that the copulas produce very different prices for the low risk portfolio. As default unlikely for these high grade securities, the choice of copula can greatly impact the default price, via the probability of default. Note that the copulas produce the most different premia when Kendall's τ is neither close to 0 nor to 1. This is because, when a high dependency is forced into the model ($\tau = 1$), the copulas will all be consistent with a very low probability of all these high-grade securities defaulting. On the other end, when we model very low dependency ($\tau = 0$), all the copulas will agree that the possibility of one of these securities defaulting is at its highest and their pricing will be nearly identical. Unlike the other copulas, the Student and the Gaussian are very similar throughout the range of τ, a result consistent with the existing literature.

Considering only the top plots of Figures 2 to 5, we notice the upward sloping curves as a function of τ. Again, for these high grade bonds, multiple defaults can really only result from a high level of dependence. Conversely, for independent, high-grade securities, the probability of multiple defaults is extremely low. Note also that the prices are lower for the higher order default swaps.

In summary, we propose to price n^{th} to default swaps using continuous time Markov chains and copulas to model the joint dynamics of default. The method is easy to implement and prices can be calculated rapidly. We implement the model on real data and demonstrate the model's flexibility for default-risk based instruments. For the most likely values of Kendall's τ (i.e. for neither independent nor perfectly correlated vectors of risky securities) the choice of the copula for credit-default pricing can have a very large impact on prices. Also, while the Gaussian and the Student exhibit very similar behavior, this is not the case for other, possibly more general copula. This suggest further research to investigate which copula fit credit instruments best.

Appendix on Continuous-time Markov Chains

Suppose that $X_t = \left(X_t^{(1)}, \ldots, X_t^{(d)}\right)$ is a homogeneous time Markov chain with infinitesimal generator Λ and state space $\mathcal{S} = S_1 \times \cdots S_d$. Further assume that the only absorbing state is the constant state $(m, \ldots, m)$. If Λ is the infinitesimal generator then

$$P\left(X_{t+h} = \beta \middle| X_t = \alpha\right) = \begin{cases} h\Lambda_{\alpha\beta} + o(h), & \beta \neq \alpha, \\ 1 + h\Lambda_{\alpha\alpha} + o(h), & \beta = \alpha, \end{cases}, \quad \alpha, \beta \in \mathcal{S}.$$

It follows that $\Lambda_{\alpha\beta} \geq 0$, $\beta \neq \alpha$, and

$$\lambda_\alpha = -\Lambda_{\alpha\alpha} = \sum_{\beta \neq \alpha} \Lambda_{\alpha\beta}.$$

Furthermore, if

$$P_{\alpha\beta}(t) = P\left(X_t = \beta \middle| X_0 = \alpha\right), \quad \alpha, \beta \in \mathcal{S},$$

then $\dot{P}(t) = \Lambda P(t)$, $P(0) = I$. Hence, using the usual convention $\Lambda^0 = I$, P can be written as

$$P(t) = e^{t\Lambda} = \sum_{n=0}^{\infty} \frac{t^n}{n!} \Lambda^n, \quad t \geq 0.$$

For example, if $d = 1$, $S = \{1, 2\}$, and $\Lambda = \begin{pmatrix} -\lambda & \lambda \\ 0 & 0 \end{pmatrix}$, then

$$P(t) = \begin{pmatrix} e^{-\lambda t} & 1 - e^{-\lambda t} \\ 0 & 1 \end{pmatrix}, \quad t \geq 0.$$

It is known that for all $\alpha \in \mathcal{S}$,

$$P(X_{s+u} = \alpha \text{ for all } u < t \,|\, X_s = \alpha) = e^{-\lambda_\alpha t}.$$

Therefore, α is an absorbing state if and only if $\lambda_\alpha = 0$. It follows that a Markov chain stays in a non absorbing state α an exponentially distributed (mean $= 1/\lambda_\alpha$) time and then goes to state $\beta \neq \alpha$ with probability $\Lambda_{\alpha\beta}/\lambda_\alpha$.

Another way of interpreting the continuous time Markov chain is the following: if the chain is at a non absorbing state $\alpha \in \mathcal{S}$, then after an exponentially distributed (mean $= 1/\lambda_\alpha$) time, $\mu_\alpha \geq \lambda_\alpha$, the chain goes to state $\beta \neq \alpha$ with probability $R_{\alpha\beta} = \Lambda_{\alpha\beta}/\mu_\alpha$ or stays at state α with probability $R_{\alpha\alpha} = 1 - \lambda_\alpha/\mu_\alpha$. For simplicity, assume that $\mu_\alpha = 0$ when $\lambda_\alpha = 0$. Note that any value μ_α would also lead to the same generator.

Let

$$M = \{\mu : \mathcal{S} \mapsto [0,\infty); \mu_\alpha \geq \lambda_\alpha,\ \mu_\alpha = 0 \text{ when } \lambda_\alpha = 0\},$$

and then every $\mu \in M$ corresponds a unique transition probability matrix R defined by

$$R_{\alpha\beta} = \begin{cases} \frac{\Lambda_{\alpha\beta}}{\mu_\alpha}, & \lambda_\alpha > 0, \beta \neq \alpha, \\ 1 - \frac{\lambda_\alpha}{\mu_\alpha}, & \lambda_\alpha > 0, \beta = \alpha, \\ 0, & \lambda_\alpha = 0, \beta \neq \alpha, \\ 1, & \lambda_\alpha = 0, \beta = \alpha. \end{cases}$$

Note that $\mu_\alpha(R_{\alpha\beta} - I_{\alpha\beta}) = \Lambda_{\alpha\beta}$. The set of all (μ, R) with $\mu \in M$ is denoted by $\mathcal{M}$. For every such R, one can define a discrete time Markov chain $(Y_n)_{n\geq 0}$ with transition matrix R.

It follows that an equivalent way of describing the continuous time Markov chain with generator Λ is to prescribe the jumping rates $\mu \in M$, together with the transition matrix R.

There is a subset of $\mathcal{M}$ which is most interesting. Consider the set $\mathcal{M}_0$ of all (μ, R) so that R is a transition matrix of a discrete Markov chain Y with the property that for each nonempty set $A = \{a_1, \ldots, a_k\} \subset \{1, \ldots, d\}$, $Y^{(A)} = \left(Y^{(a_1)}, \ldots, Y^{(a_k)}\right)$ is a discrete time Markov chain.

It follows from Proposition 3 that this is equivalent to the following condition: For each nonempty set $A \subset \{1, \ldots, d\}$,

$$\sum_{\beta\in\mathcal{S};\ \beta_j=\gamma_j,\ j\in A} R_{\alpha\beta} \tag{7}$$

depends only on $\alpha_A = \{\alpha_j,\ j \in A\}$. For any $k = 1, \cdots, d$, $i, j = 1, \ldots, m$, set

$$G^{(k)}_{ij} = \sum_{l=1}^{j} R^{(k)}_{ij},$$

where $G^{(k)}_{i0} = 0$. Next, given a copula C on $[0,1]^d$, if $U = (U_1, \ldots, U_d)$ is a random vector with distribution function C, define, for any $\alpha, \beta \in \mathcal{S}$,

$$R_{\alpha\beta} = P\left(G^{(1)}_{\alpha_1,\beta_1-1} < U_1 \leq G^{(1)}_{\alpha_a,\beta_1}, \ldots, G^{(d)}_{\alpha_d,\beta_d-1} < U_d \leq G^{(d)}_{\alpha_d,\beta_d}\right). \tag{8}$$

It follows that R, as defined by (8), satisfies (7). Of course C is not uniquely defined by (8), but for a given copula C, and given marginal cumulative transition matrices $R^{(1)}, \ldots, R^{(d)}$, $\mathcal{R}$ is uniquely determined by (8).

It follows that for any $A \subset D$,

$$R^{(A)}_{\alpha\beta} \leq \min_{k\in A} R^{(k)}_{\alpha_k\beta_k}.$$

It is now easy to characterize members (μ, R) of $\mathcal{M}_0$ satisfying Assumption 2.

Proposition 2. *Under Assumption 2, for any $(\mu, R) \in \mathcal{M}_0$, there is a constant $\lambda > 0$ such that $\mu\alpha = \lambda$ for all $\alpha \neq (m, m, \ldots, m)$.*

Proof. Take $(\mu, R) \in \mathcal{M}_0$. It follows from Proposition 3 and Corollary 4 below that for any $1 \le i \neq j \le m$, any $1 \le k \le d$, and any $\alpha \in \mathcal{S}$ with $\alpha_k = i$,

$$\Lambda_{ij}^{(k)} = \sum_{\beta \in \mathcal{S};\, \beta_k = j} \Lambda_{\alpha\beta} = \sum_{\beta \in \mathcal{S};\, \beta_k = j} \mu_\alpha R_{\alpha\beta} = \mu_\alpha R_{ij}^{(k)}.$$

Hence,

$$\lambda_i^{(k)} = \mu_\alpha \left(1 - R_{ii}^{(k)}\right).$$

If $i < m$, then $\lambda_i^{(k)} > 0$, and μ_α is constant for any $\alpha \neq (m, m, \ldots, m)$.

Finally, if $\alpha = (m, m, \ldots, m)$, then $\lambda_\alpha = 0$ since α is an absorbing point. Therefore $\mu_\alpha = 0$. ◻

Based on Proposition 2 and relation (8), the following statement makes sense: The infinitesimal generator Λ is determined by a copula C, and a constant $\lambda \ge \max_{1 \le k \le d} \max_{1 \le j < m} \lambda_i^{(k)}$. Having fixed λ, the marginal transition matrices $R^{(k)}$ corresponding to $\Lambda^{(k)}$ are $R^{(k)} = I + \Lambda^{(k)}/\lambda$, while the cumulative transition matrix $G^{(k)}$ is defined by

$$G_{ij}^{(k)} = \sum_{l=1}^{j} R_{il}^{(k)}, \quad 1 \le i, j \le m, \; k = 1, \ldots, d.$$

Proposition 3 and Corollary 4 used above are as follows.

Proposition 3. *Assume R is a transition matrix of a discrete Markov chain $Y \in \mathcal{S} = S_1 \times \cdots S_d$, with the property that for each nonempty set $A = \{a_1, \ldots, a_k\} \subset \{1, \ldots, d\}$, $Y^{(A)} = \left(Y^{(a_1)}, \ldots, Y^{(a_k)}\right)$ is a discrete time Markov chain on $\mathcal{S}_A = \otimes_{j \in A} S_j$, with transition matrix $\mathcal{R}^{(A)}$.*

Then, a necessary and sufficient condition for this is that for all nonempty set $A \subset \{1, \ldots, d\}$, and all $\alpha \in \mathcal{S}$ so that $\alpha_A = \gamma \in \mathcal{S}_A$, i.e. $\alpha_j = \gamma_j$ for all $j \in A$,

$$(R_A)_{\gamma,\delta} = \sum_{\beta \in \mathcal{S};\, \beta_j = \delta_j,\, j \in A} R_{\alpha\beta}, \delta \in \mathcal{S}_A.$$

Proof. The condition is clearly sufficient. To prove that it is also necessary, remark that for any fixed $\alpha \in \mathcal{S}$ with $\alpha_A = \gamma$, it follows from the Markov hypothesis that

$$\begin{aligned} \left(R^{(A)}\right)_{\gamma\delta} &= P\left\{ Y_1^{(A)} = \delta \middle| Y_0^{(A)} = \gamma \right\} \\ &= P\left\{ Y_1^{(A)} = \delta \middle| Y_0 = \alpha \right\} \\ &= \sum_{\beta \in \mathcal{S};\, \beta_j = \delta_j,\, j \in A} R_{\alpha\beta}. \end{aligned}$$

Corollary 4. *Suppose that Λ is the infinitesimal generator of a continuous time Markov chain X_t with state space $\mathcal{S} = S_1 \times \cdots S_d$, with the additional property that for each nonempty set $A = \{a_1, \ldots, a_k\} \subset \{1, \ldots, d\}$, $X^{(A)} = \left(X^{(a_1)}, \ldots, X^{(a_k)}\right)$ is a continuous time Markov chain on $\mathcal{S}_A = \otimes_{j \in A} S_j$, with generator $\Lambda^{(A)}$.*

Then, a necessary and sufficient condition for this is that for all nonempty set $A \subset \{1,\ldots,d\}$, and all $\alpha \in \mathcal{S}$ so that $\alpha_A = \gamma \in \mathcal{S}_A$, i.e. $\alpha_j = \gamma_j$ for all $j \in A$,

$$\left(\Lambda^{(A)}\right)_{\gamma,\delta} = \sum_{\beta\in\mathcal{S};\ \beta_j=\delta_j,\ j\in A} \Lambda_{\alpha\beta}, \delta \in \mathcal{S}_A.$$

Theorem 5. *Suppose $X^{(1)},\ldots,X^{(d)}$ are Markov chains with infinitesimal generators $\Lambda^{(1)},\ldots,\Lambda^{(d)}$ on the state spaces $S_1,\ldots,S_d$. Then there is a unique way to define a Markov chain $X = \left(X^{(1)},\ldots,X^{(d)}\right)$ on $\mathcal{S} = S_1 \times \cdots S_d$, with generator Q so that only one transition among the components is permitted at a time.*

In that case, for any $\alpha,\beta \in \mathcal{S}$,

$$\Lambda_{\alpha\beta} = \sum_{k=1}^{d} \left(\Lambda^{(k)}\right)_{\alpha_k\beta_k} \prod_{j\neq k} I_{\alpha_j\beta_j},$$

and it follows that the Markov chains $X^{(1)},\ldots,X^{(d)}$ are all independent.

Proof. If $X^{(1)},\ldots,X^{(d)}$ are independent, then $X = \left(X^{(1)},\ldots,X^{(d)}\right)$ is clearly a Markov chain. It only remains to calculate the generator $\tilde{\Lambda}$.

By hypothesis, $P_{\alpha\beta}(t) = \prod_{k=1}^{d}\left(P^{(k)}(t)\right)_{\alpha_k\beta_k}$, if $\beta \neq \alpha$, so

$$\tilde{\Lambda}_{\alpha\beta} = \frac{d}{dt}P_{\alpha\beta}(t)\Big|_{t=0} = \sum_{k=1}^{d} \left(\Lambda^{(k)}\right)_{\alpha_k\beta_k} \prod_{j\neq k} I_{\alpha_j\beta_j} = \Lambda_{\alpha\beta}.$$

Hence $\tilde{\Lambda} = \Lambda$.

On the other hand, if more than one simultaneous transition is prohibited among the components, and if all components are Markov chains, it follows from Corollary 4 that for any $\alpha,\beta \in \mathcal{S}$, Λ is given by

$$\Lambda_{\alpha\beta} = \sum_{k=1}^{d} \left(\Lambda^{(k)}\right)_{\alpha_k\beta_k} \prod_{j\neq k} I_{\alpha_j\beta_j}.$$

Setting

$$\tilde{P}_{\alpha\beta}(t) = \prod_{k=1}^{d} \left(P^{(k)}(t)\right)_{\alpha_k\beta_k},$$

one can check that $\dot{\tilde{P}}(t) = \Lambda\tilde{P}(t)$, and $\tilde{P}(0) = I$. Hence $\tilde{P} = P$, by uniqueness of the solution. □

References

B. Abdous, C. Genest, and B. Rémillard. Dependence properties of meta-elliptical distributions. In P. Duchesne and B. Rémillard, editors, *Statistical Modeling and Analysis for Complex Data Problem*, pages 1–15. Kluwer Academic Publishing, 2004.

D. Dupuis, E. Jacquier, N. Papageorgiou, and B. Remillard. An empirical study of dependence of credit default data and of equity prices. Technical report, HEC Montreal, 2006.

C. Genest and J. MacKay. The joy of copulas: bivariate distributions with uniform marginals. *Amer. Statist.*, 40:280–283, 1986.

C. Genest, K. Ghoudi, and L.-P. Rivest. A semiparametric estimation procedure of dependence parameters in multivariate families of distributions. *Biometrika*, 82(3):543–552, 1995.

C. Genest, B. Rémillard, and D. Beaudoin. Goodness-of-fit tests for copula models: A critical review and a power study. Technical report, GERAD, 2006.

G. Greg, C. Finger, and M. Bhatia. *CreditMetrics Technical Document*. RiskMetrics Group, New York, 1997.

J. Hull and A. White. Valuation of a CDO and an n^{th} to default cds without Monte Carlo simulation. Technical report, Forthcoming, Journal of Derivatives, 2005.

R. A. Jarrow and S. M. Turnbull. Pricing derivatives on financial securities subject to credit risk. *Journal of Finance*, L:53–85, 1995.

R. A. Jarrow, D. Lando, and S. M. Turnbull. A Markov model for the term structure of credit risk spreads. *The Review of Financial Studies*, 10:481–523, 1997.

H. Joe. *Multivariate Models and Dependance Concepts*, volume 73 of *Monographs on Statistics and Applied Probability*. Chapmann and Hall, London, 1997.

D. Lando and T.M. Skødeberg. Analyzing rating transition and rating drift with continuous observations. *Journal of Banking & Finance*, 26:423–444, 2002.

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

R. Mashal and M. Naldi. Pricing multiname credit derivatives: Heavy tailed hybrid approach. Technical report, Columbia University, 2001.

R. Mashal, M. Naldi, and A. Zeevi. On the dependence of equity and asset returns. *Risk*, 16 (October):83–87, 2003.

R. Merton. On the pricing of corporate debt: The risk structure of interest rates. *Journal of Finance*, 29:449–470, 1974.

R.B. Nelsen. *An Introduction to copulas*, volume 139 of *Lectures Notes in Statistics*. Springer-Verlag , New York, 1999.

J. V. Rosenberg and T. Schuermann. A general approach to integrated risk management with skewed, fat-tailed risks. Technical report, New York Federal Reserve Staff Report 185, 2004.

P.J. Schönbucher and D. Schubert. Copula-dependent default risk in intensity models. Technical report, Department of Statistics, Bonn University, 2001.

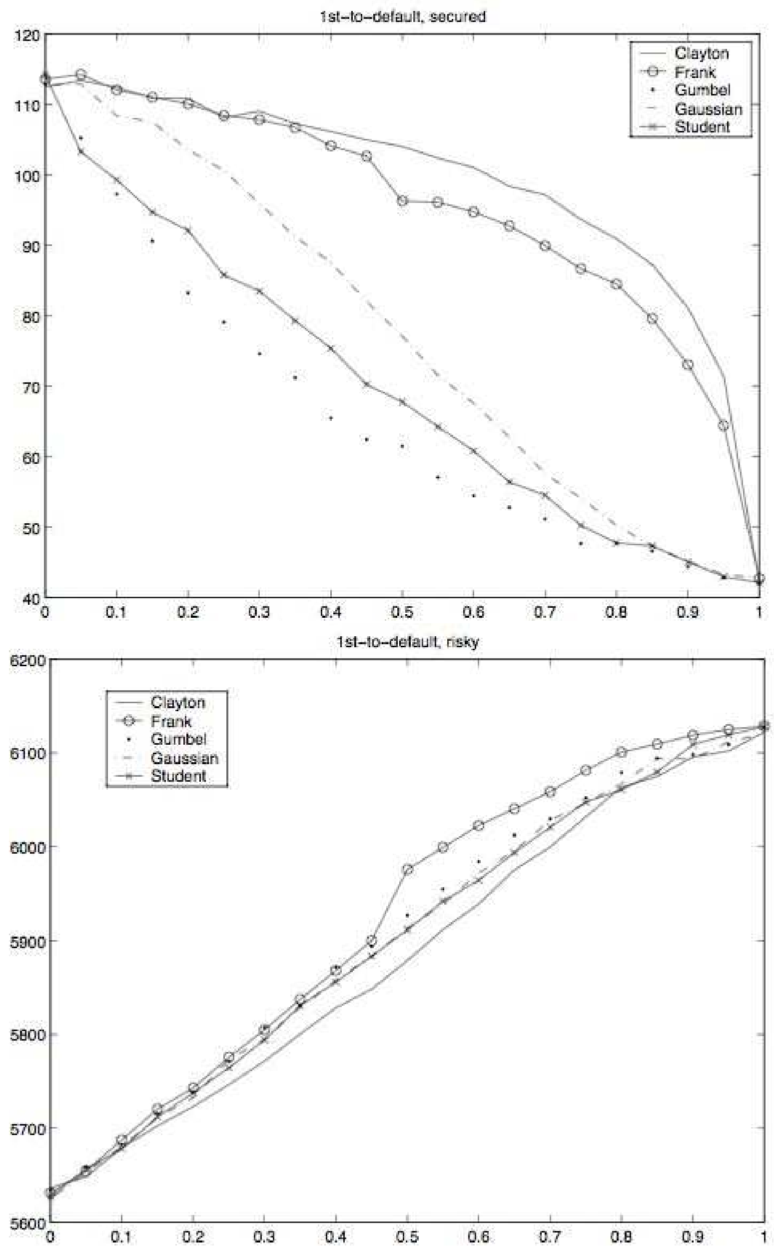

Figure 1: Premium (basis points) for 1^{st} to default for five-year contracts with bi-annual payment vs Kendall's τ for different copulas. Top plot: high grade; bottom plot: low grade.

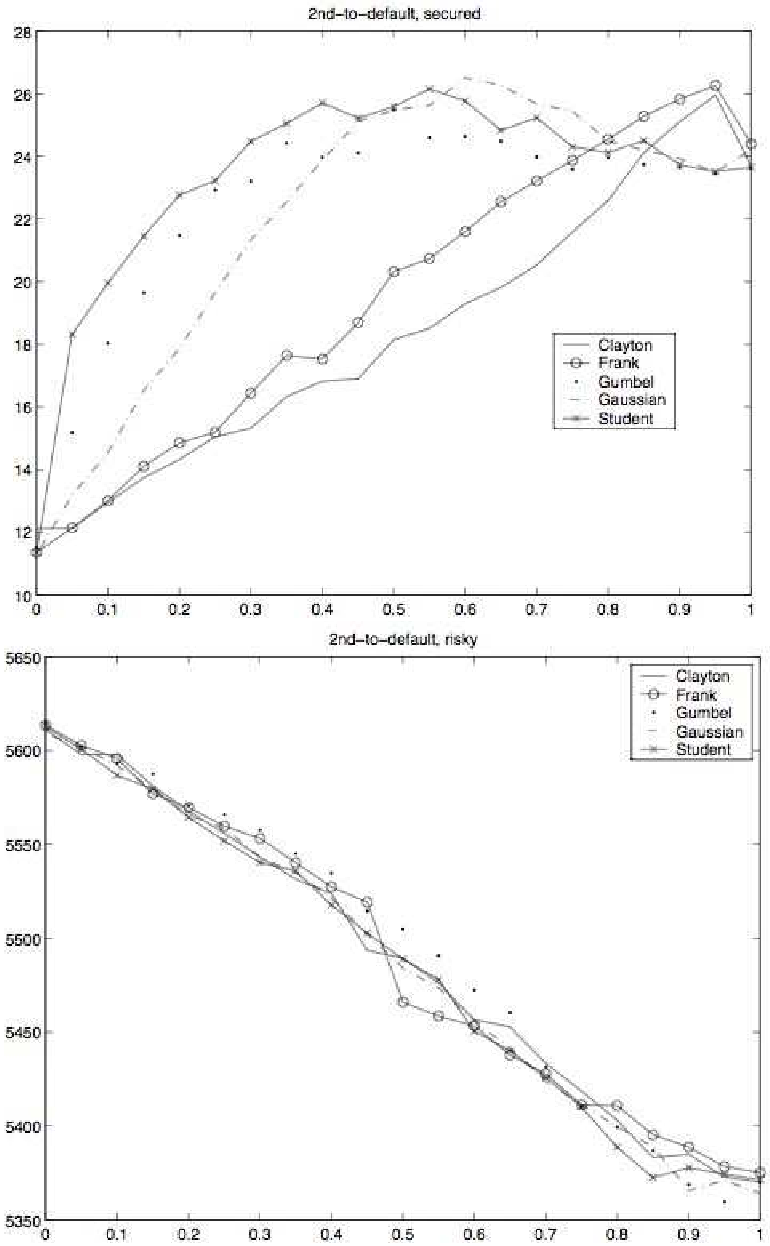

Figure 2: Premium (basis points) for 2^{nd} to default for five-year contracts with bi-annual payment vs Kendall's τ for different copulas. Top plot: high grade; bottom plot: low grade.

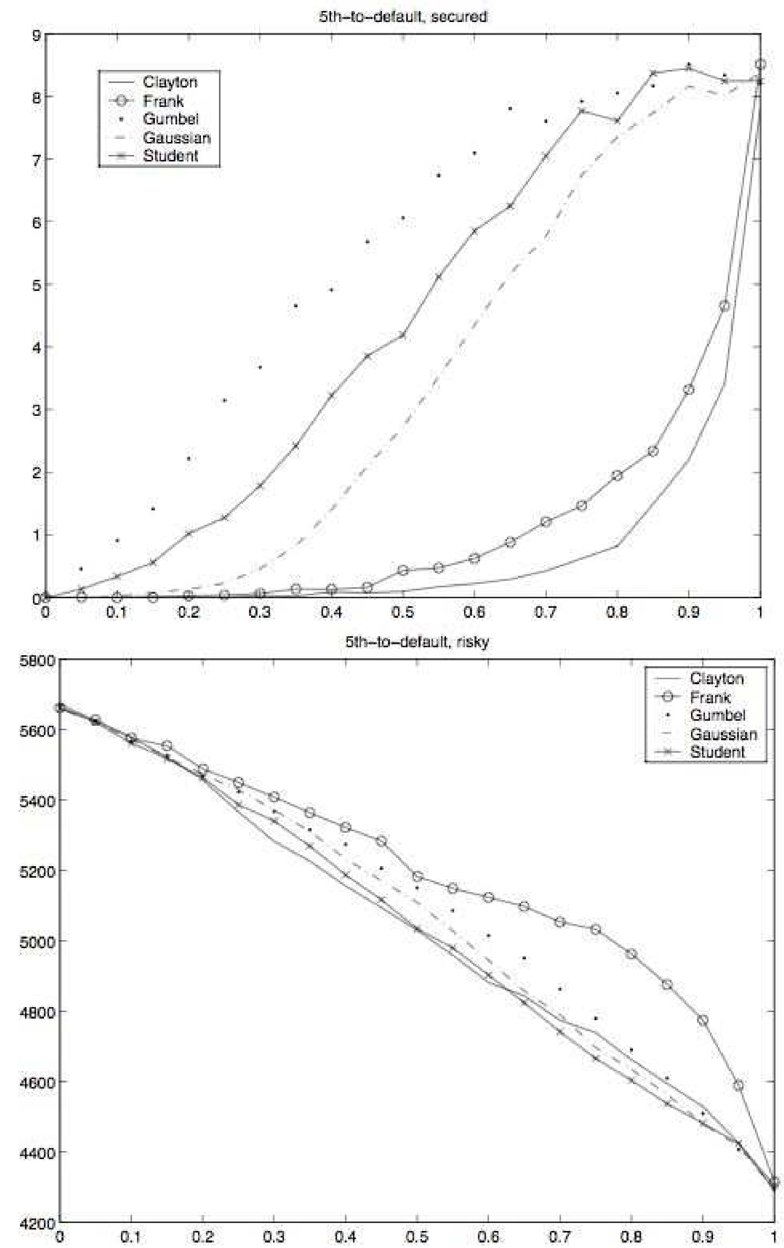

Figure 3: Premium (basis points) for 5^{th} to default for five-year contracts with bi-annual payment vs Kendall's τ for different copulas. Top plot: high grade; bottom plot: low grade.

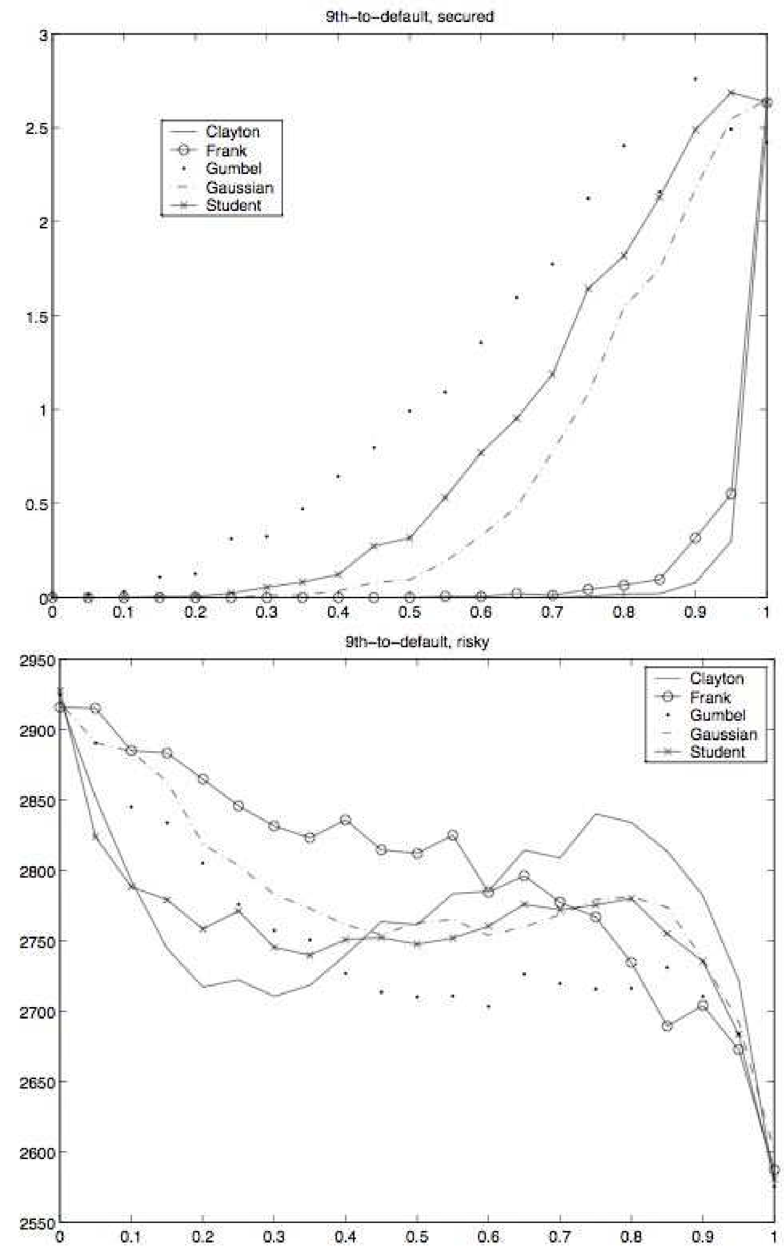

Figure 4: Premium (basis points) for 9^{th} to default for five-year contracts with bi-annual payment vs Kendall's τ for different copulas. Top plot: high grade; bottom plot: low grade.

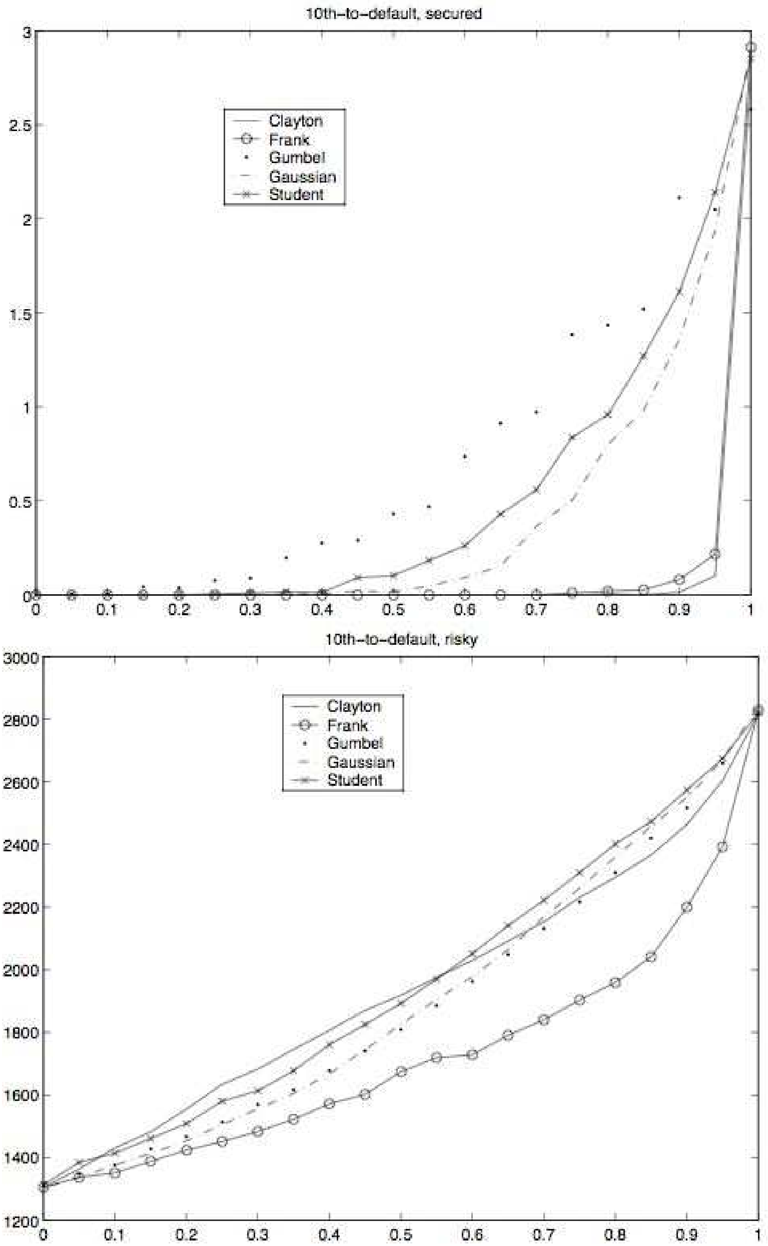

Figure 5: Premium (basis points) for 10^{th} to default for five-year contracts with bi-annual payment vs Kendall's τ for different copulas. Top plot: high grade; bottom plot: low grade.