Abstract

The multivariate modelling 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 dynamic behavior of credit ratings for several firms. Namely, individual credit ratings are modelled by univariate continuous time Markov chain, while their joint dynamics is modelled 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 modelling 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
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 the modelling of a multivariate process of risky securities, namely of their joint default processes. A typical example of such a product is an \( n^{th} \) to default swap which payout is contingent on the time and identity of the \( n^{th} \) default of a credit portfolio.

This joint modelling can be seen as 1) the modelling of each individual credit rating migration (where default is the worst rating and constitutes an absorbing state) and 2) the modelling of their joint evolution. This multivariate problem has been one of the most vexing technical challenge in the credit risk literature. CreditMetrics, e.g., technical document ([Greg et al., 1997]), one of the benchmarks in sophisticate default-risk pricing, uses one single Poisson based transition matrices for all the individual securities, and a Gaussian copula to model their joint behavior. Since all these individual Markov chains evolve according to the same transition matrix, the value of the portfolio is determined strictly by the number of bonds initially in each state, and does not allow for these bonds to have different characteristics. This extreme elimination of idiosyncratic default risk is a severe modelling restriction.

Clearly, to evaluate the price of a \( n^{th} \) to default swap, one needs the joint distribution of the vector of default times. These default times can be modelled with two general approaches, the structural approach initiated in Merton [1974]), and the reduced form, intensity based, approach, exemplified in Jarrow and Turnbull [1995]. Not however that these two pioneering papers only model a single default time. Recently, Schönbucher and Schubert [2001] propose a feasible model for a multivariate distribution of default times, based upon the reduce form approach. 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 parameters. We will show that this is in fact not the case.

An extension of the earlier intensity-based approach for the modelling of individual default is to use each bond’s credit rating information and model its evolution by a Markov chain, in discrete
and continuous time. This is proposed in Jarrow et al. [1997] for the first time. Typically, as is done by these authors, the law of motion must be studied under both the objective and martingale measures.

In this paper, we will follow this latter continuous time Markov chain approach for modelling the evolution of the credit ratings of each firm. To price credit default swaps or to calculate risk measures for portfolios, one needs to consider the joint behavior of the vector of credit ratings. To allow for obvious dependencies among these ratings, we will use the more recent copula discussed above. It 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., see Li [2000], they are most always of the Gaussian family or Student family. This limited menu of copulas leads to the belief in the literature that the choice of copula has a minor effect on risk modelling, e.g., Rosenberg and Schuermann [2004].

The remainder of the paper is structured as follows. Section 2, first gives the main assumptions required for modelling the joint behavior of multiple credit ratings. These hypotheses are motivated and discussed in further details in the Appendix. The section discusses the methodology for either pricing credit ratings derivatives or calculating measures of risk. Finally the section gives the algorithm for the simulation of multiple credit rating trajectories and outlines the pricing of an nth to default swap. Then, section 3 illustrates the implementation of the model. It shows 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 finally the computation of the default premia.

Finally, section 4 concludes after showing the properties of the model for 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 $X_t^{(1)}, \ldots, X_t^{(d)}$ at time t, where the credit ratings of a single firm take values in the ordered set $S = \{1, \ldots, m\}$,
where 1 is the best rating, and \( m \) is the default state. The goal is to model the joint behavior of the stochastic processes \( X^{(1)}, \ldots, X^{(d)} \) so as to price credit derivatives on the basket of bonds.

The assumptions made below are natural in a credit rating migration context. The first assumption concerns the dynamic behavior of each individual ratings, while the second relates to their joint distribution.

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

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

*Further: for any \( k \), the only absorbing state is state \( m \).*

This is a classic continuous-time modelling of each individual credit rating \( X^{(k)} \), whereby ratings can change at any time according to a Markov chain.

Clearly, the only absorbing state is default. 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 \( \tau_k \) 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 \( P_{ij}^{(k)}(t) = P \left( X_t^{(k)} = j \mid X_0^{(k)} = i \right) \), \( i, j \in S \), then it follows from Assumption 1 that \( P^{(k)}(t) = e^{t\Lambda^{(k)}} \), \( P^{(k)}(0) = I \). Hence \( P^{(k)}(t) = e^{t\Lambda^{(k)}} \), \( t \geq 0 \), and the distribution function \( F_i^{(k)} \) of \( \tau_k \), given \( X_0^{(k)} = i \), is given by

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

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 \( \Delta \) is a diagonal matrix, then \( P^{(k)}(t) = Me^{t\Delta}M^{-1} \) and
\( e^{t\Delta} \) is the diagonal matrix with \( (e^{t\Delta})_{ii} = e^{t\Delta_{ii}}, \) \( i = 1, \ldots, m. \) Therefore

\[
F_i^{(k)}(t) = \left( P^{(k)}(t) \right)_{im} = 1 + \sum_{j=1}^{m-1} M_{ij} e^{t\Delta_{jj}} \left( M^{-1} \right)_{jm}, \quad t \geq 0. \tag{2}
\]

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 modelling, is that the joint distribution of the infinitesimal generator can be represented by a copula as follows:

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

\[
\Lambda = \Phi \left( \Lambda^{(1)}, \ldots, \Lambda^{(d)}, C, \lambda \right) = \lambda (R - I), \tag{3}
\]

where

\[
R^{(k)} = I + \Lambda^{(k)}/\lambda, \quad G^{(k)}_{ij} = \sum_{l=1}^{j} R^{(k)}_{il}, \quad 1 \leq i, j \leq m, \ k = 1, \ldots, d, \quad \tag{4}
\]

\( G^{(k)}_{i0} = 0, \) and

\[
R_{\alpha\beta} = P \left( G^{(1)}_{\alpha_1, \beta_1 - 1} < U_1 \leq G^{(1)}_{\alpha_2, \beta_1}, \ldots, G^{(d)}_{\alpha_d, \beta_d - 1} < U_d \leq G^{(d)}_{\alpha_d, \beta_d} \right), \tag{5}
\]

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

It is important to note that (5) allows for multiple credit ratings changes at the same time. In particular, many firms could default at exactly the same time, a phenomenon rarely observed. This in fact arises only when the copula parameters imply very strong cross-sectional dependencies in the infinitesimal generator. If desired, this can be prevented by modifying the copula. We discuss how to do this in the Appendix, specifically, see Theorem 5 and the associated discussion.

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

\[
X_t = Y_{N_t}, \quad t \geq 0 \tag{6}
\]
where $N_t$ is a Poisson process with intensity $\lambda$. That link between the continuous and discrete-time modelling will be used for simulating continuous-time Markov chains. Further discussion of Assumption 2 is provided in Appendix.

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 distribution of each credit rating.

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

Here, the parameter $\sigma_k$ is the price of risk for the law of default time. As is standard practice for fixed-income derivative securities, it can be calibrated directly from market data, for example by fitting market prices of credit default swaps on the $k^{th}$ firm.

It follows from Assumption 3 that, if one sets

$$P_{ij}^{(k)}(t) = \mathcal{P}\left(X^{(k)}_t = j \mid X^{(k)}_0 = i\right), \quad i, j \in S,$$

then

$$\mathcal{P}^{(k)}(t) = e^{t\mathcal{L}^{(k)}} = e^{t\sigma_k \Lambda^{(k)}} = P^{(k)}(\sigma_k t), \quad t \geq 0. \quad (7)$$

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

$$\tilde{F}_i^{(k)}(t) = \mathcal{P}\left(\tau_k \leq t \mid X_0^{(k)} = i\right) = \left(\mathcal{P}^{(k)}(t)\right)_{im}^{(k)} = F_i^{(k)}(\sigma_k t), \quad i = 1, \ldots, m, \quad k = 1, \ldots, d. \quad (8)$$

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 $\tilde{\lambda} \geq \max \max_{1 \leq k \leq d, 1 \leq i \leq m} \tilde{\lambda}_i^{(k)}$, and $\tilde{\lambda}_i^{(k)} = \sigma_k \lambda_i^{(k)}$, $k = 1, \ldots, d$. 

6
That is, similarly to equation 3, we write for the risk-neutral infinitesimal generator that:

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

where

\[ \mathcal{R}_{\alpha\beta} = P \left( \mathcal{G}^{(1)}_{\alpha_1,\beta_1} \leq U_1 \leq \mathcal{G}^{(1)}_{\alpha_2,\beta_1}, \ldots, \mathcal{G}^{(d)}_{\alpha_d,\beta_d} \leq U_d \leq \mathcal{G}^{(d)}_{\alpha_d,\beta_d} \right), \]

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

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

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

\[ X_t = \tilde{Y}_{\tilde{N}_t}, \] (9)

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 calculate \( n^{th} \) to default swaps.

Given the parameters, it is easy to describe the technique that we propose to generate multiple credit ratings trajectories. Recall that 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. Namely, the inputs needed for simulation are, (1) the initial vector of credit ratings \( X_0 \) and (2) the discrete-time transition matrix \( \mathcal{R}. \)

Recall the relation (6), that is \( X_t = Y_{N_t}, \) for \( t \geq 0. \) Given an initial state, the transition to the next state is modelled by a Poisson process \( N, \) which intensity vector \( \lambda \) is a row of the transition matrix \( \mathcal{R}. \) 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 as follow:

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

2. Suppose that the constant $\lambda$, 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

$$X_{T_i} = k \text{ if } G_{j,k-1}^{(l)} < U_i^{(l)} \leq G_{j,k}^{(l)}, \quad j, k \in S. \quad (10)$$

Having generated multivariate trajectories, one can estimate the prices of the desired contracts 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 called a $n^{th}$ to default swap. Following Mashal and Naldi [2001], the contract can be described as follows:

- The contract starts at time $t = 0$ and its maturity is $T$.
- 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 $\tau_n$, provided $\tau_n \leq T$.
- At time $T$, if $\tau_{n,d} \leq T$, the insurance is triggered and the insurance payment is $N(1 \cdot RR_{n,d} \cdot a)$. 
• $r_t$ 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

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

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

$$p_f = f \frac{E_{\mathcal{P}} \left\{ \beta(\tau_{n,d}) (1 - RR_{n,d}) \mathbb{I}(\tau_{n,d} \leq T) \right\}}{E_{\mathcal{P}} \left\{ \sum_{i=1}^{\int T} \beta(i/f) \mathbb{I}\left( \tau_{n,d} > \frac{i}{f} \right) + \sum_{i=1}^{\int T} \beta(i/f) \mathbb{I}\left( \frac{i-1}{f} < \tau_{n,d} \leq \frac{i}{f} \right) \right\}}$$

$$= \frac{f E_{\mathcal{P}} \left\{ \beta(\tau_{n,d}) (1 - RR_{n,d}) \mathbb{I}(\tau_{n,d} \leq T) \right\}}{E_{\mathcal{P}} \left\{ \sum_{i=1}^{\int T} \beta(i/f) \mathbb{I}\left( \tau_{n,d} > \frac{i-1}{f} \right) \right\}}.$$ 

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

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

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

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

which reduces, under independence, to the simpler formula

$$q_f = f (1 - RR) \frac{B(T) F(T) - \int_0^T \partial_t B(t) F(t) dt}{\sum_{i=1}^{\int T} B\left( \frac{i}{f} \right) F\left( \frac{i-1}{f} \right)},$$
where, under $\mathcal{P}$, $\overline{RR}$ is the mean recovery rate, $F$ is the distribution function of the default time, and $\overline{F} = 1 - F$ is the associated survival function.

**Remark 1.** Since $B$ is non increasing, $\partial_t B(t) \leq 0$. So it is easy to check, from equations (11) and (13) 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 $F_{n,d}(t) \leq G_{n,d}(t)$ for all $0 \leq t \leq T$, then $p_f(F_{n,d}) \leq p_f(G_{n,d})$. Similarly, if $F(t) \leq G(t)$ for all $0 \leq t \leq T$, then $q_f(F) \leq q_f(G)$.

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

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

and

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

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, describing the data used along the way.

For estimating the infinitesimal generator of the simple Markov chains, we use the method proposed in Lando and Skødeberg [2002] and outline the procedure in Section 4.1. Credit default swaps (CDS) prices are used to estimate the risk neutral parameters $\sigma_k$, required to perform credit derivatives pricing. The corresponding returns are used to estimate the copula parameters, based on the pseudo-likelihood method of Genest, Ghoudi, and Rivest [1995], while the choice of an adequate copula family can be based on the methodology developed in Barbe, Genest, Ghoudi, and
Rémillard [1996] and Genest, Quessy, and Rémillard [2003].

In order to implement the proposed methodology, one first needs to estimate matrices $\Lambda^{(1)}, \ldots, \Lambda^{(d)}$. The data used to estimate these matrices consist of credit rating information from Moody’s 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, $\Lambda^{(1)} = \cdots = \Lambda^{(d)}$. Given enough date, we could of course relax this assumption. In practice, it is often needed so that there are 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 $k$, $1 \leq i \leq m$, $0 \leq j \leq m$, $j \neq i$, $1 \leq k \leq N$. Then $N_{ij} = \sum_{k=1}^{N} N_{ij}^{(k)}$ is the total number of transitions from state $i$ to state $j$, $1 \leq i \leq m$, $0 \leq j \leq m$, reported in the database, and $N_{i}^{(k)} = \sum_{j \neq i} N_{ij}^{(k)}$ 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, $1 \leq i \leq m$, $1 \leq k \leq N$. Set

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

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, the approach taken to estimate $\Lambda^{(1)}$ is the “cohort” method described in Lando and Skødeberg [2002]. Namely, the estimator of $\Lambda_{ij}^{(1)}$ is given by

$$\hat{\Lambda}_{ij}^{(1)} = \frac{N_{ij}}{L_i}, \quad 1 \leq i \leq m, 0 \leq j \leq m, j \neq i.$$ 

It follows that

$$\hat{\lambda}_i^{(1)} = \frac{N_i}{L_i}, \quad 1 \leq i \leq m.$$

Hence one can take

$$\hat{\lambda} = \max_{1 \leq i \leq m} \hat{\lambda}_i^{(1)} = \max_{1 \leq i \leq m} \frac{N_i}{L_i}.$$

Table 1 shows the resulting discrete-time generator of the transition matrix. Recall that $\lambda(t)$ is the transition probability. The negative values on the diagonal correspond to a very high probability of no migration in credit. The zero’s far from the diagonal represent the absence of a jump across many credit ratings in the data base.

### 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 $T$, we use $x_i$ and $r_i$ to denote the credit rating of the firm and the interest rate at period $i$, $1 \leq i \leq n$, 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 (11) to estimate $\sigma$. For example, one could choose $\sigma$ so as to minimize $\sum_{i=1}^{n} [Y_i - q_{f,x_i,r_i,T}(\sigma)]^2$.

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

To complete our numerical example, we chose a portfolio of 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 January 7, 2000, to February 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 $\sigma_1, \ldots, \sigma_7$ as well as the credit rating at the end of the period. The prices of risk range from 2.9 to 4.1, most of them being around 3.5. This is because these specific firms are all large high grade issuers whose risk characteristics do not vary greatly.

### 3.3 Estimation of the copula $C$

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 the well-known Gaussian copula as a first example. If $Y = (Y_1, \ldots, Y_d) \sim N_d(0, \Sigma)$, i.e. $Y$ is a $d$-dimensional Gaussian distribution with mean zero and covariance matrix $\Sigma$, with $\Sigma_{ii} = 1$ for all $1 \leq i \leq d$, then the Gaussian copula $C_\Sigma$ is defined as

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

where $N(\cdot)$ is the distribution function of the standard Gaussian distribution.

As a second example, consider the closely related Student copula. If $Y \sim N_d(0, \Sigma)$ and $V$ is an independent variable with a chi-square distribution with $\nu$ degrees of freedom, then $Y/\sqrt{V/\nu}$ has a multivariate Student distribution with parameters $\Sigma$ and $\nu$. 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 $\nu$ 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 in the form

$$C(u_1, \ldots, u_d) = \phi^{-1} \{\phi(u_1) + \cdots + \phi(u_d)\}, \quad (14)$$
where \( \phi : (0, 1] \rightarrow [0, \infty) \), is a bijection such that \( \phi(1) = 0 \) and

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

These three classes share the interesting property that \( \phi \), denoted the generator of the copula, is unique, up to a constant.

Table 3 gives the generators for three well-known Archimedean copulas: the Clayton, Frank, and Gumbel copulas. See Joe [1997] and Nelsen [1999] for further examples on copulas.

Due to a lack of data related to 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 \( Y = (Y^{(1)}, \ldots, Y^{(d)}) \) is invariant for monotone increasing transformations of its components \( Y^{(k)} \). That is, the copulas of \( Y \) and \( (\psi_1(Y^{(1)}), \ldots, \psi_1(Y^{(d)})) \) are the same whenever the \( \psi_k \)'s are strictly increasing transformations.

If 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 were 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 no empirical evidence to support or discourage this assumption.

Given returns \( Y_i = (Y^{(1)}_i, \ldots, Y^{(d)}_i) \), \( i = 1, \ldots, N \), one first calculates ranks \( \text{Rank}(Y^{(k)}_i) \) for each fixed \( k \). That is, \( \text{Rank}(Y^{(k)}_i) \) represents the rank of \( Y^{(k)}_i \) within \( Y^{(1)}_1, \ldots, Y^{(k)}_N \), with rank one assigned to the smallest number.

Next, one defines the pseudo-observations \( U_i \) as follows:

\[
U_i = \left( U^{(1)}_i, \ldots, U^{(d)}_i \right) = \frac{1}{N+1} \left( \text{Rank}(Y^{(1)}_i), \ldots, \text{Rank}(Y^{(d)}_i) \right), \quad i = 1, \ldots, N.
\]

If \( c_\sigma \) represents the density of the copula, the estimation of the parameter \( \theta \), possibly
multidimensional, is defined as the value $\hat{\theta}$ maximizing

$$\arg \max_{\theta} \sum_{i=1}^{N} \log \{c_{\theta}(U_i)\}. \quad (16)$$

Note that (16) 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 $\hat{\Sigma}$. 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 $\Sigma$ for the Gaussian copula model is given by

$$\hat{\Sigma} = \begin{pmatrix}
1.00 & 0.43 & 0.38 & 0.40 & 0.35 \\
0.43 & 1.00 & 0.59 & 0.48 & 0.34 \\
0.38 & 0.59 & 0.55 & 0.47 & 0.37 \\
0.40 & 0.48 & 0.47 & 0.54 & 0.33 \\
0.35 & 0.34 & 0.37 & 0.33 & 0.37
\end{pmatrix}. \quad (17)$$

Estimating the parameters $\Sigma$ and $\nu$ of the more general Student copula is a little bit trickier. It is recommended to follow a two-step procedure: first estimate $\hat{\Sigma}$, the estimate $\nu$, using the pseudo-maximum-likelihood method (16). It is known, e.g. B. Abdous and Rémillard [2004], that $\Sigma_{jk} = \sin(\tau_{jk} \pi/2)$, where $\tau_{jk}$ is the theoretical Kendall’s tau between components $Y^{(j)}$ and $Y^{(k)}$. If $\hat{\tau}_{jk}$ is an estimate of $\tau_{jk}$, then $\sin(\hat{\tau}_{jk} \pi/2)$ is an estimate of $\Sigma_{jk}$. 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 $\Delta$ 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 model given by

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

(18)

with an estimation of \( \hat{\nu} = 3.7 \) degrees of freedom.

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

### 3.3.1 Choice of the copula family

Since the choice of the copula family can be of paramount importance, one needs a way to choose the appropriate family. Some testing procedures have been studied so far. One method, suggested by Genest et al. [2003], is particularly well-suited for the Archimedean class. It is based on a one-dimensional characteristic of a copula, namely the probability transform \( K \), which is the distribution function of the random variable \( C(U) \). It is is estimated by an empirical function \( K_n \), and \( \sqrt{n}(K_n - K) \) is the so-called Kendall’s process. See Barbe et al. [1996] for additional details and important formulas. Test statistics can be defined as

\[
KS_N = \sup_{t \in [0,1]} \sqrt{n} \left| K_n(t) - \bar{K}(t) \right|,
\]

or

\[
CVM_N = \int_0^1 n \left\{ K_n(t) - \bar{K}(t) \right\}^2 d\bar{K}(t),
\]

where \( \bar{K} \) is the probability transform associated with the copula \( C_{\hat{\theta}} \). As shown in Genest et al. [2003], p-values can be calculated for the null hypothesis that the real copula belongs to a
given family. Graphical procedures based of $CVM_N$ can also be implemented to assist in the copula selection.

Other testing procedures have been proposed and proved to work well in Genest et al. [2005]. They are mainly based on the empirical copula

$$C_n(u_1, \ldots, u_d) = \# \left\{ i ; U_{i}^{(k)} \leq u_k, \text{ for all } k = 1, \ldots, d \right\} .$$

Formal statistical testing is besides the scope of this paper. However, we show below that the choice of copula has a major impact on the pricing of default risk.

3.3.2 Calculations of premia for $n^{th}$ to default swaps

Having estimated all parameters, one is a position to estimate the premiums $p_f$ for different copula families. The prices, in basis points, were then calculated by Monte Carlo methods using 100 000 repetitions. They are reported in Table 4.

Note that the choice of copula has a large impact on the prices. For the 1-st to default, the largest price is obtained for the Clayton family, while the smallest is obtained for the Student family. These roles are totally reverse 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. Further numerical examples are carried out 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 a portfolio composed of 10 firms having the following high-grade credit ratings: Aaa Aaa, Aa1, Aa1, Aa2, Aa2, Aa3, A1, A2, A3. We will assume here that the $\sigma_k$‘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 will assume that the prices of risk $\sigma_k$‘s are all equal to 5.
To study the effect of copulas beyond the commonly used Gaussian and Student, we select three additional families, namely the Clayton, Frank, and Gumbel families. For the purpose of 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 used by Hull and White [2005]. The five families are then indexed by one parameter, $\tau$, the Kendall tau. 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 the values $\tau = j/20$, for $j = 0, \ldots, 20$. We use the generator matrix estimated from the Moody's data set and shown in Table 1, as well as the values $f = 2$, $r = 2\%$ and $T = 5$ for all the simulations. For each copula and value of $\tau$, we conduct 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 the Kendall $\tau$ for the various copulas and orders of default. Consider Figure 1, $1^{st}$ to default. The different copulas price the swap on the risky securities in a very similar manner, as shown in the bottom plot. This is because, for these risky assets, the probability of even only 1 default out of 10, is very high. In contrast, the different copulas can result in very different prices for the low risk portfolio, as shown in the top plot. As a default is possible but unlikely, for these high grade securities, the choice of copula can greatly impact the default price (via the probability of default). Note that the impact of the copula on the premium is greatest when the Kendall $\tau$ is neither close to 0 nor to 1. This is because, when a high dependency, $\tau = 1$, is forced into the model, 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. Therefore their pricing will be nearly identical. Note how the Student and the Gaussian are very similar throughout the range of $\tau$, a result consistent with the existing literature.

Considering now only the top plots of figures 2 to 5, we notice the upward sloping curves as a function of $\tau$. This is because, for these high grade bonds, multiple defaults can really only result from a high level of dependence. That is, 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 swap 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 actual data. Then, we demonstrate its flexibility in modelling default-risk based instruments. Namely, for the most likely value of Kendall’s $\tau$ for actual data, that is 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.
<table>
<thead>
<tr>
<th></th>
<th>100</th>
<th>-15.66</th>
<th>11.36</th>
<th>3.53</th>
<th>0.23</th>
<th>0.38</th>
<th>0.08</th>
<th>0.08</th>
<th>0</th>
<th>0</th>
<th>0</th>
</tr>
</thead>
<tbody>
<tr>
<td>3.14</td>
<td>-30.84</td>
<td>13.59</td>
<td>11.64</td>
<td>1.44</td>
<td>0.68</td>
<td>0.17</td>
<td>0</td>
<td>0.08</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>1.02</td>
<td>4.22</td>
<td>-26.68</td>
<td>13.62</td>
<td>4.96</td>
<td>1.11</td>
<td>0.28</td>
<td>0.09</td>
<td>0.09</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>0.15</td>
<td>0.09</td>
<td>6.00</td>
<td>-28.02</td>
<td>15.50</td>
<td>4.64</td>
<td>0.63</td>
<td>0.09</td>
<td>0.15</td>
<td>0.03</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>0.04</td>
<td>0.13</td>
<td>0.58</td>
<td>7.17</td>
<td>-25.09</td>
<td>11.81</td>
<td>4.17</td>
<td>0.47</td>
<td>0.25</td>
<td>0.04</td>
<td>0.18</td>
<td>0.16</td>
</tr>
<tr>
<td>0.02</td>
<td>0.08</td>
<td>0.36</td>
<td>1.27</td>
<td>7.48</td>
<td>-28.11</td>
<td>15.45</td>
<td>3.09</td>
<td>0.87</td>
<td>0.23</td>
<td>0.08</td>
<td>0.11</td>
</tr>
<tr>
<td>0.10</td>
<td>0.08</td>
<td>0.10</td>
<td>0.30</td>
<td>1.50</td>
<td>10.90</td>
<td>-34.59</td>
<td>12.58</td>
<td>0.98</td>
<td>1.48</td>
<td>0.15</td>
<td>0.18</td>
</tr>
<tr>
<td>0.06</td>
<td>0.16</td>
<td>0.22</td>
<td>0.26</td>
<td>0.29</td>
<td>2.75</td>
<td>9.96</td>
<td>-36.39</td>
<td>15.42</td>
<td>4.98</td>
<td>1.12</td>
<td>0.48</td>
</tr>
<tr>
<td>0.03</td>
<td>0.18</td>
<td>0.09</td>
<td>0.25</td>
<td>1.01</td>
<td>4.94</td>
<td>8.58</td>
<td>-37.53</td>
<td>15.70</td>
<td>4.48</td>
<td>0.86</td>
<td>0.19</td>
</tr>
<tr>
<td>0.04</td>
<td>0.04</td>
<td>0.04</td>
<td>0.04</td>
<td>0.28</td>
<td>0.55</td>
<td>1.07</td>
<td>3.80</td>
<td>12.95</td>
<td>-43.69</td>
<td>13.86</td>
<td>6.69</td>
</tr>
<tr>
<td>0.04</td>
<td>0.04</td>
<td>0.04</td>
<td>0.23</td>
<td>0.23</td>
<td>0.23</td>
<td>0.13</td>
<td>0.58</td>
<td>0.75</td>
<td>2.75</td>
<td>11.88</td>
<td>-41.88</td>
</tr>
<tr>
<td>0</td>
<td>0.03</td>
<td>0.03</td>
<td>0</td>
<td>0.19</td>
<td>0.14</td>
<td>0.28</td>
<td>0.33</td>
<td>0.64</td>
<td>3.26</td>
<td>7.15</td>
<td>0</td>
</tr>
<tr>
<td>0.03</td>
<td>0</td>
<td>0.03</td>
<td>0.10</td>
<td>0.15</td>
<td>0.13</td>
<td>0.23</td>
<td>0.23</td>
<td>0.18</td>
<td>0.64</td>
<td>3.10</td>
<td>0</td>
</tr>
<tr>
<td>0</td>
<td>0.05</td>
<td>0.05</td>
<td>0.05</td>
<td>0</td>
<td>0.18</td>
<td>0.18</td>
<td>0.23</td>
<td>0.37</td>
<td>0.46</td>
<td>1.05</td>
<td>0</td>
</tr>
<tr>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0.30</td>
<td>0.60</td>
<td>0.30</td>
<td>0</td>
</tr>
<tr>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>0.15</td>
<td>0.10</td>
<td>0</td>
<td>0.05</td>
<td>0.05</td>
<td>0.10</td>
<td>0.10</td>
<td>0.10</td>
<td>0.30</td>
<td>0.25</td>
<td>0.45</td>
<td>0</td>
</tr>
<tr>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0.73</td>
<td>0</td>
</tr>
<tr>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0.59</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>0.08</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>0.05</td>
<td>0</td>
<td>0.05</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>0</td>
<td>0.09</td>
<td>0.03</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>0.04</td>
<td>0.04</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>0.06</td>
<td>0.04</td>
<td>0.02</td>
<td>0.02</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>0.13</td>
<td>0.10</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>0.41</td>
<td>0.22</td>
<td>0.06</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>0.46</td>
<td>0.25</td>
<td>0.25</td>
<td>0.06</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>2.38</td>
<td>1.11</td>
<td>0.40</td>
<td>0.08</td>
<td>0.08</td>
<td>0.20</td>
<td>0.04</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>7.92</td>
<td>2.33</td>
<td>1.72</td>
<td>0.23</td>
<td>0.08</td>
<td>0</td>
<td>0.11</td>
<td>0.04</td>
<td>0.04</td>
<td>0.04</td>
<td>0.46</td>
<td>0</td>
</tr>
<tr>
<td>13.43</td>
<td>4.87</td>
<td>5.10</td>
<td>1.60</td>
<td>0</td>
<td>0.04</td>
<td>0.09</td>
<td>0.04</td>
<td>0</td>
<td>0</td>
<td>0.27</td>
<td>0</td>
</tr>
<tr>
<td>-36.00</td>
<td>10.29</td>
<td>8.65</td>
<td>3.07</td>
<td>0.14</td>
<td>0.31</td>
<td>0.06</td>
<td>0.08</td>
<td>0.03</td>
<td>0</td>
<td>0.70</td>
<td>0</td>
</tr>
<tr>
<td>7.43</td>
<td>-38.29</td>
<td>15.07</td>
<td>7.25</td>
<td>0.56</td>
<td>1.23</td>
<td>0.46</td>
<td>0.08</td>
<td>0.13</td>
<td>0</td>
<td>1.26</td>
<td>0</td>
</tr>
<tr>
<td>1.12</td>
<td>2.41</td>
<td>6.03</td>
<td>-68.11</td>
<td>6.63</td>
<td>4.22</td>
<td>1.21</td>
<td>9.64</td>
<td>0.60</td>
<td>32.25</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>0.22</td>
<td>0.66</td>
<td>1.76</td>
<td>10.55</td>
<td>0</td>
<td>-91.89</td>
<td>19.79</td>
<td>24.18</td>
<td>14.51</td>
<td>1.54</td>
<td>18.69</td>
<td>0</td>
</tr>
<tr>
<td>1.32</td>
<td>2.21</td>
<td>3.09</td>
<td>0</td>
<td>7.06</td>
<td>-114.77</td>
<td>18.54</td>
<td>16.33</td>
<td>10.59</td>
<td>55.62</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>0.59</td>
<td>0</td>
<td>1.19</td>
<td>2.07</td>
<td>1.78</td>
<td>0</td>
<td>5.94</td>
<td>6.54</td>
<td>-129.00</td>
<td>12.48</td>
<td>96.90</td>
<td>0</td>
</tr>
<tr>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>7.62</td>
<td>-224.74</td>
<td>217.12</td>
<td>0</td>
</tr>
</tbody>
</table>

$\Lambda$ is the generator for the transition probability matrix $e^{\Lambda t}$
Table 2: Estimates of $\sigma$ for seven firms

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

$\sigma$ is the price of default risk expressed in basis points. Note that the other parameters are set to $RR = 50\%$, $r = 2\%$, $T = 5$, and $f = 2$.

Table 3: Multivariate Archimedean copulas and parameter estimates

<table>
<thead>
<tr>
<th>Model</th>
<th>$\phi_\theta(t)$</th>
<th>Range of $\theta$</th>
<th>$\hat{\theta}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Clayton</td>
<td>$(t^{-\theta} - 1)/\theta$</td>
<td>$(0, \infty)$</td>
<td>0.430</td>
</tr>
<tr>
<td>Frank</td>
<td>$\log \left( \frac{1 - e^{-\theta}}{1 - e^{-\theta}} \right)$</td>
<td>$(0, \infty)$</td>
<td>2.775</td>
</tr>
<tr>
<td>Gumbel-Hougaard</td>
<td>$</td>
<td>\log t</td>
<td>^{1/\theta}$</td>
</tr>
</tbody>
</table>

Table 4: Premia for the $n^{th}$ to default for different copula families

<table>
<thead>
<tr>
<th>Model</th>
<th>$1^{st}$</th>
<th>$2^{nd}$</th>
<th>$3^{rd}$</th>
<th>$4^{th}$</th>
<th>$5^{th}$</th>
<th>$6^{th}$</th>
<th>$7^{th}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Clayton</td>
<td>151.1</td>
<td>25.21</td>
<td>3.48</td>
<td>0.40</td>
<td>0.03</td>
<td>0.00</td>
<td>0.00</td>
</tr>
<tr>
<td>Frank</td>
<td>149.8</td>
<td>28.97</td>
<td>5.12</td>
<td>0.77</td>
<td>0.12</td>
<td>0.02</td>
<td>0.00</td>
</tr>
<tr>
<td>Gumbel</td>
<td>110.8</td>
<td>35.90</td>
<td>15.68</td>
<td>7.49</td>
<td>3.88</td>
<td>2.01</td>
<td>0.73</td>
</tr>
<tr>
<td>Gaussian</td>
<td>136.7</td>
<td>32.29</td>
<td>8.57</td>
<td>2.10</td>
<td>0.50</td>
<td>0.11</td>
<td>0.01</td>
</tr>
<tr>
<td>Student</td>
<td>109.8</td>
<td>39.26</td>
<td>17.01</td>
<td>7.53</td>
<td>3.15</td>
<td>1.06</td>
<td>0.24</td>
</tr>
</tbody>
</table>

The default premia are quoted in basis points.
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 $\Lambda$, with state space $S = S_1 \times \cdots \times S_d$. Further assume that the only absorbing state is the constant state $(m, \ldots, m)$.

That $\Lambda$ is the infinitesimal generator means that

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

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

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

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 $\alpha \in S$,

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

Therefore, $\alpha$ is an absorbing state if and only if $\lambda_\alpha = 0$. It follows that a Markov chain stays in a non absorbing state $\alpha$ a random time with exponential distribution with mean $1/\lambda_\alpha$, then goes to state $\beta \neq \alpha$ with probability $\frac{\lambda_{\alpha \beta}}{\lambda_\alpha}$.

Another way of interpreting the continuous time Markov chain is the following: if the chain is at a non absorbing state $\alpha \in S$, after a random period of time with exponential distribution with mean $1/\mu_\alpha$, $\mu_\alpha \geq \lambda_\alpha$, the states goes to state $\beta \neq \alpha$ with probability $R_{\alpha \beta} = \frac{\lambda_{\alpha \beta}}{\mu_\alpha}$ or stays at state $\alpha$ with probability $R_{\alpha \alpha} = 1 - \frac{\lambda_\alpha}{\mu_\alpha}$. For simplicity, assume that $\mu_\alpha = 0$ when $\lambda_\alpha = 0$. Note that any value $\mu_\alpha$ would also lead to the same generator.

Therefore let

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

22
To every $\mu \in M$ corresponds a unique transition probability matrix $R$ defined by

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

Note that $\mu_\alpha (R_{\alpha\beta} - I_{\alpha\beta}) = \Lambda_{\alpha\beta}$. The set of all $(\mu, R)$ 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 $\Lambda$ 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. It is the set $\mathcal{M}_0$ of all $(\mu, R)$ so that $R$ is a transition matrix of a discrete Markov chain $Y$ with the property that for each nonempty set $A = \{a_1, \ldots, a_k\} \subset \{1, \ldots, d\}$, $Y(A) = (Y(a_1), \ldots, Y(a_k))$ is a discrete time Markov chain.

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

$$\sum_{\beta \in S; \beta_j = \alpha_j, j \in A} R_{\alpha\beta}$$

depends only on $\alpha_A = \{\alpha_j, j \in A\}$.

For any $k = 1, \ldots, d$, $i, j = 1, \ldots, m$, set

$$G_{ij}^{(k)} = \sum_{l=1}^{j} R_{ij}^{(k)},$$

where $G_{i0}^{(k)} = 0$.

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 $\alpha, \beta \in S$,

$$R_{\alpha\beta} = P \left( G_{\alpha_1,\beta_1}^{(1)} < U_1 \leq G_{\alpha_2,\beta_1}^{(1)}, \ldots, G_{\alpha_d,\beta_d}^{(d)} < U_d \leq G_{\alpha_d,\beta_d}^{(d)} \right).$$

(20)

It follows that $R$, as defined by (20), satisfies (19). Of course $C$ is not uniquely defined by (20), but for a given copula $C$, and given marginal cumulative transition matrices $R^{(1)}(1), \ldots, R^{(d)}(d)$, $R$ is uniquely determined by (20).

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

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

It is now easy to characterize members $(\mu, R)$ 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 $\alpha \neq (m, m, \ldots, m)$.  

23
Proof. Take \((\mu, R) \in \mathcal{M}_0\). It follows from Proposition 3 and Corollary 4 that for any \(1 \leq i \neq j \leq m\), any \(1 \leq k \leq d\), and any \(\alpha \in \mathcal{S}\) with \(\alpha_k = i\),

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

Hence,

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

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

Finally, if \(\alpha = (m, m, \ldots, m)\), then \(\lambda_{\alpha} = 0\) because \(\alpha\) is an absorbing point. Therefore \(\mu_{\alpha} = 0\). \(\square\)

Based on Proposition 2 and relation (20), the following statement makes sense: The infinitesimal generator \(\Lambda\) is determined by a copula \(C\), and a constant \(\lambda \geq \max_{1 \leq k \leq d} \max_{1 \leq j < m} \lambda_i^{(k)}\).

Having fixed \(\lambda\), 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

\[
G_{ij}^{(k)} = \sum_{l=1}^{j} R_{il}^{(k)}, \quad 1 \leq i, j \leq m, \ k = 1, \ldots, d.
\]

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

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 \(j \in A\),

\[
(R_{A})_{\gamma, \delta} = \sum_{\beta \in \mathcal{S}; \beta_j = \delta_j, j \in A} R_{\alpha\beta}, \delta \in \mathcal{S}_A. \tag{21}
\]

**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

\[
(R^{(A)})_{\gamma, \delta} = P\left\{Y_1^{(A)} = \delta \mid Y_0^{(A)} = \gamma\right\}
\]

\[
= P\left\{Y_1^{(A)} = \delta \mid Y_0 = \alpha\right\}
\]

\[
= \sum_{\beta \in \mathcal{S}; \beta_j = \delta_j, j \in A} R_{\alpha\beta}.
\]

**Corollary 4.** Suppose that \(\Lambda\) is the infinitesimal generator of a continuous time Markov chain \(X_t\) with state space \(\mathcal{S} = S_1 \times \ldots S_d\), with the additional property that for each nonempty set \(A = \{a_1, \ldots, a_k\} \subset \{1, \ldots, d\}\), \(X^{(A)} = (X^{(a_1)}, \ldots, X^{(a_k)})\) is a continuous time Markov chain on \(S_A = \otimes_{j \in A} S_j\), with generator \(\Lambda^{(A)}\).
Then, a necessary and sufficient condition for this to happen is that for all nonempty set \( A \subset \{1, \ldots, d\} \), and all \( \alpha \in S \) so that \( \alpha_A = \gamma \in S_A \), i.e. \( \alpha_j = \gamma_j \) for all \( j \in A \),

\[
\Lambda^{(A)}_{\gamma, \delta} = \sum_{\beta \in S; \beta_{j=\delta_j}, j \in A} \Lambda_{\alpha \beta}, \delta \in S_A.
\tag{22}
\]

**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 \( X = (X^{(1)}, \ldots, X^{(d)}) \) on \( S = S_1 \times \cdots S_d \), with generator \( Q \) so that only one transition among the components is permitted at a time.

In that case, for any \( \alpha, \beta \in S \),

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

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 \( X = (X^{(1)}, \ldots, X^{(d)}) \) is clearly a Markov chain. It only remains to calculate the generator \( \bar{\Lambda} \).

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

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

Hence \( \bar{\Lambda} = \Lambda \).

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 \( \alpha, \beta \in S \), \( \Lambda \) 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 \( \tilde{P}(t) = \Lambda \tilde{P}(t) \), and \( \tilde{P}(0) = I \). Hence \( \tilde{P} = P \), by uniqueness of the solution. \( \square \)

**References**


