

A Parsimonious Multivariate Markov Chain Model for Credit Risk

Wai-Ki Ching ^{*} Tak-Kuen Siu [†] Li-min Li[‡] Tang Li[§] Wai-Keung Li [¶]

27 Mar. 2007

Abstract

In this paper, we develop a parsimonious and flexible multivariate Markov chain model for describing the dependency of transitions of ratings of credit risky entities. The model serves as a complement to copulas, which cannot be applied to describe the dependence of categorical time series, like credit ratings, effectively. The number of parameters involved here is only $O(sm^2 + s^2)$, where s is the dimension of the categorical time series of ratings and m is the number of possible ratings of a security. The new model has relatively small number of parameters and is more parsimonious and flexible when compared to the model of Siu et al. [18]. Our model also provides a great deal of flexibility in modelling both the positive and negative associations between the time series of credit ratings. Another distinctive feature of our model is its ability to handle the case when the dataset of ratings is short. This is an important feature since it is rather difficult or costly to obtain a long dataset of ratings in practice. We introduce an efficient method to estimate or calibrate the model. The estimation method can be formulated as a linear programming problem and it can be implemented using spreadsheets. We compare the estimation results and the computational efficiency of our model with those of Siu et al. [18] to illustrate the effect of incorporating both positive and negative associations on portfolio credit risks.

Key Words: Multivariate Markov chains, Parsimony, Ratings, Dependent Credit Risk, Positive and Negative Associations, Linear Programming.

^{*}Advanced Modeling and Applied Computing Laboratory, Department of Mathematics, The University of Hong Kong, Pokfulam Road, Hong Kong. E-mail: wching@hkusua.hku.hk. Research supported in part by RGC Grants, HKU CRCG Grants, Hung Hing Ying Physical Sciences Research Fund and HKU Strategic Research Theme Fund on Computational Physics and Numerical Methods.

[†]Department of Actuarial Mathematics and Statistics, School of Mathematical and Computer Sciences and the Maxwell Institute for Mathematical Sciences, Heriot-Watt University, Edinburgh EH14 4AS, UK.

[‡]Advanced Modeling and Applied Computing Laboratory, Department of Mathematics, The University of Hong Kong, Pokfulam Road, Hong Kong. E-mail: liminli@hkusua.hku.hk

[§]Advanced Modeling and Applied Computing Laboratory, Department of Mathematics, The University of Hong Kong, Pokfulam Road, Hong Kong. E-mail: litang@hkusua.hku.hk

[¶]Department of Statistics and Actuarial Science, The University of Hong Kong, Pokfulam Road, Hong Kong. Email: hrntlwk@hkucc.hku.hk

1 Introduction

Ratings data of credit risky entities provide us with an important piece of information for analyzing the credit quality of the entities. Some international rating agencies, like Standard & Poor, Moodys and Fitch, are the major providers of rating data. Investors can make use of the data provided by rating agencies to judge the credit qualities of corporate bonds, sovereign bonds and other securities and make their financial decisions. In practice, ratings change over time randomly according to the changes in the economic fundamentals of the entities, the economic judgments and forecasts made by analysts, macro-economic conditions, and others. Given the importance of ratings, academics and practitioners are looking for some appropriate quantitative methods to describe and analyze the transitions of ratings of credit risky entities. The discrete-time homogeneous Markov chain model provides a natural mean to model such transitions over time. It has widely been used by practitioners for modelling transitions of ratings.

Recently, there has been a considerable interest in modelling the dependency of individual credit risks in a portfolio. Knowledge in the dependency of credit risks is important for portfolio credit management. In particular, the model specification that explains and describes the dependency of the credit risks has significant implications for portfolio credit risk measurement and management. Different approaches have been introduced in the literature for this aspect. A natural approach to describe the inter-dependence of credit risks is the use of default correlations. It is a common approach for pricing multivariate credit derivatives, such as Collateral Debt Obligations (CDO). Li [15] proposes the use of copulas for modeling dependence of credit risks. The main advantage of the use of the copulas is that it can model asymmetric dependency and non-linear association of variables, which cannot be described by correlations. There are some technical issues on the practical implementation of the use of copulas for modelling dependent credit risks. One issue is on which parametric form of copulas function one should use in practice. There are different parametric forms of copulas functions, including Gaussian copulas, t -copulas and Clayton copulas. Each of them generates a particular type of asymmetric and non-linear dependent behavior. There are some statistical procedures that guide us how to choose a particular parametric form of copulas function. However, the theoretical justifications for these statistical procedures have yet been well-explored and understood. Moreover, some statistical procedures are rather ad hoc. Copulas can work well for describing asymmetric and non-linear dependent relationships when the underlying variables are continuous. However, when the underlying variables are discrete, like credit ratings, copulas cannot be applied to describe the dependence of the variables directly and effectively since according to Sklar's theorem, copulas cannot be specified uniquely when the underlying variables are discrete. Hence, there is a practical need to develop a method which can describe the dependence of credit ratings effectively. Some other models for dependent credit risks are the mixture model approach, in particular the Poisson mixture model in CreditRisk+ developed by Credit Suisse Financial Products in 1997 [8], the infectious default model

for Binomial Expansion Technique (BET) by Davis and Lo [9, 10] and the multivariate Markov chain model of Kijima et al. [14].

Modelling the inter-dependency and intra-dependency of ratings of individual entities in a portfolio is a crucial step for modelling dependency of credit risks. Kijima et al. [14] first introduce a multivariate Markov chain model to simulate the evolution of dependent ratings of several credit risks. Based on the multivariate Markov chain model in [2] Siu et al. [18] provide a different parameterization of multivariate Markov chain for modelling dependent credit risks and propose the use of credibility theory in actuarial science to calibrate the model. The model proposed by Siu et al. [18] can describe both the inter-dependent and intra-dependent of ratings. The number of parameters involved in the model is $O(s^2m^2)$, where s is the dimension of the categorical time series of ratings and m is the number of possible credit ratings of a security. To incorporate all possible pairwise dependencies between any two categorical time series for ratings, $C_2^s = s(s-1)/2$ parameters are involved. The model by Siu et al. [18] can only incorporate positive correlations (or associations), but not negative ones, between two ratings series. It can be a reasonable proxy if the two credit risks are from the same sector and the systematic effect in the sector dominates the idiosyncratic effect. However, when the idiosyncratic effect, or in particular, the competitive effect, within the sector dominates the systematic effect or when the two credit risks come from different sectors with negative associations, modelling positive associations only is not flexible enough.

In this paper, we shall develop a parsimonious and flexible multivariate Markov chain model for describing the dependency of transitions of ratings of credit risky entities. As in the model of Siu et al. [18], we suppose that the conditional probability distribution of the rating of a particular credit risk in the next period given the currently available information depends not only on its own current rating, but also on the current ratings of all other credit risks in the portfolio. In other word, the model can describe both the inter-dependency and intra-dependency of categorical time series of ratings. The model proposed in this paper is more parsimonious and flexible than the model introduced in [18]. The number of parameters involved here is only $O(sm^2 + s^2)$, where s is the dimension of the categorical time series of ratings and m is the number of possible ratings of a security. Our model also provides a great deal of flexibility in modelling both the positive and negative associations between the time series of credit ratings. We shall describe the precise meaning of positive and negative associations. In many cases, the stationary or the steady-state conditions are assumed in the estimation of model parameters. This means that a longer dataset is required for estimation. However, in practice, it might not be easy or it might be rather costly to get a long dataset for ratings. A distinctive feature of our model is its ability to accommodate the situation when the dataset is short. The estimation method can be formulated as a linear programming problem, which can be solved easily using spreadsheets. This provides market practitioners with a handful way to implement the estimation method. We compare the estimation results and the

computational efficiency of our proposed model with those of Siu et al. [18]. We shall also compare credit risk measures obtained from our model with those from Siu et al. [18] to illustrate the effect of incorporating both positive and negative associations on portfolio credit risks.

The rest of the paper is structured in the sequel. In Section 2, we introduce the multivariate Markov chain model for describing the dependency of credit risks and discuss some important properties of the model. Section 3 presents the estimation or calibration method of the model. In section 4, we provide some numerical examples to illustrate the implementation of the multivariate Markov chain model for credit risk measurement. We shall also present the estimation results of our model and compare them with those in Siu et al. [18] using real credit ratings data and the risk ratings data for bonds from four important industrial sectors in United States extracted by the interactive hidden Markov model (IHMM) in Ching et al. [5] using the quarterly bond defaults data of the four sectors. We shall compare the risk measures obtained from our model with those implied by the model of Siu et al. [18]. Concluding remarks are given in Section 5.

2 The New Multivariate Markov Chain Model

In this section, we shall first review the multivariate Markov chain model in Siu et al. [18]. Then, we shall present the main idea of the new multivariate Markov chain model. which has been successfully in sales demand forecasting [4]. The new model provides market practitioners with a more parsimonious and flexible way to model the dependency of ratings than the model of Siu et al. [18]. In particular, the new model can incorporate both positive and negative associations of ratings sequences while the model of Siu et al. [18] can only incorporate positive associations. The new model [4] can also accommodate the situation when the ratings dataset is short by adjusting a parameter which controls the rate of convergence of the sequence of the probability vectors of the process to the stationary distribution. This is an important feature of the new model since it is difficult and costly to obtain long ratings datasets and the accuracy of the estimation results depends on the rate of convergence. To investigate the rate of convergence to the stationary distribution, we shall represent the new model in the form of a vector-valued difference equation.

First, we fix a complete probability space $(\Omega, \mathcal{F}, \mathcal{P})$. Here, we consider a discrete-time economy and suppose that \mathcal{T} represents the time index set $\{0, 1, 2, \dots\}$ on which all economic activities take place. We consider a portfolio of n credit entities, for example, n corporate bonds, and deal with the situation when the ratings of these entities are dependent. Suppose $Y^{(1)}, Y^{(2)}, \dots, Y^{(s)}$ represent s ratings processes with common time index set \mathcal{T} . Let \mathcal{S} denote a set of unit vectors $\{e_1, e_2, \dots, e_m\}$ in R^m , where

$$e_i = (0, \dots, 0, \overbrace{1}^{i^{th} \text{entry}}, 0, \dots, 0)^T \in R^m .$$

Suppose, for each $j = 1, 2, \dots, s$, $Y^{(j)} := \{Y_t^{(j)}\}_{t \in \mathcal{T}}$ is a discrete-time, finite-state stochastic process on $(\Omega, \mathcal{F}, \mathcal{P})$ with state space \mathcal{S} , where $Y_t^{(j)}$ represents the ratings of the j^{th} entity at time t . In particular, “ $Y_t^{(j)} = e_i$ ” means that the j^{th} entity is in the i^{th} rating class at time t . Define the space $\hat{\mathcal{S}}$ as below:

$$\{s \in R^m | s = \sum_{i=1}^m \alpha_i e_i, \quad 0 \leq \alpha_i \leq 1, \quad \sum_{i=1}^m \alpha_i = 1\}.$$

Let $\{X_t^{(j)}\}_{t \in \mathcal{T}}$ denote the dynamics of the discrete probability distributions for $Y^{(j)}$, where $X_t^{(j)} \in \hat{\mathcal{S}}$ for each $t \in \mathcal{T}$. In particular, the i^{th} component $X_t^{(j)}(i)$ of the probability vector $X_t^{(j)}$ represents the probability that the j^{th} entity is in the i^{th} rating class at time t ; That is,

$$X_t^{(j)}(i) := \mathcal{P}(Y_t^{(j)} = e_i).$$

We also note that $\{X_t^{(j)}\}_{t \in \mathcal{T}}$ can be used to represent the ratings of the j^{th} entity at time t , for each $j = 1, 2, \dots, s$. Suppose the j^{th} entity is in the i^{th} rating class at time t ; that is, $Y_t^{(j)} = e_i$. In other words, the probability that the j^{th} entity is in the i^{th} rating class at time t is one, and,

$$X_t^{(j)} = e_i = (0, \dots, 0, \underbrace{1}_{i^{\text{th}} \text{ entry}}, 0, \dots, 0)^T.$$

Suppose $P^{(jk)}$ represents a transition matrix from the ratings of the k^{th} entity to the ratings of the j^{th} entity. In particular, $P^{(jj)}$ is the transition matrix for the ratings of the j^{th} entity. Then, under the multivariate Markov chain model for ratings in Siu et al. [18], the dynamics of the probability distributions of the ratings are governed by the following multivariate difference equations:

$$X_{t+1}^{(j)} = \sum_{k=1}^s \lambda_{jk} P^{(jk)} X_t^{(k)}, \quad \text{for } j = 1, 2, \dots, n, \quad (1)$$

where

$$\lambda_{jk} \geq 0, \quad 1 \leq j, k \leq n \quad \text{and} \quad \sum_{k=1}^n \lambda_{jk} = 1, \quad \text{for } j = 1, 2, \dots, s. \quad (2)$$

In Equation (1), the state probability distribution of the ratings of the j^{th} entity at time $(t + 1)$ depends on the weighted average of $P^{(jk)} X_t^{(k)}$ at time t . We note that

$$Y := (Y^{(1)}, Y^{(2)}, \dots, Y^{(s)})$$

forms a multivariate Markov chain under Equation (1), see for instance [18]. Equation

(1) can also be represented as the following matrix form:

$$X_{t+1} \equiv \begin{pmatrix} X_{t+1}^{(1)} \\ X_{t+1}^{(2)} \\ \vdots \\ X_{t+1}^{(n)} \end{pmatrix} = \begin{pmatrix} \lambda_{11}P^{(11)} & \lambda_{12}P^{(12)} & \cdots & \lambda_{1s}P^{(1s)} \\ \lambda_{21}P^{(21)} & \lambda_{22}P^{(22)} & \cdots & \lambda_{2s}P^{(2s)} \\ \vdots & \vdots & \vdots & \vdots \\ \lambda_{s1}P^{(s1)} & \lambda_{s2}P^{(s2)} & \cdots & \lambda_{ss}P^{(ss)} \end{pmatrix} \begin{pmatrix} X_t^{(1)} \\ X_t^{(2)} \\ \vdots \\ X_t^{(s)} \end{pmatrix} \equiv QX_t.$$

We note that the column sum of Q is not equal to one (the column sum of $P^{(jk)}$ is equal to one). In spite of this, there still exists a unique stationary distribution for the process Y . For more details, we refer readers to Ching et al. [2] and Siu et al. [18]. The number of parameters in the model of Siu et al. [18] is $O(s^2m^2)$, where s is the number of credit entities and m is the number of possible ratings of an entity. The precision of estimates of the parameters decreases as the number of parameters increases. In view of this, Ching et al. have proposed a simplified multivariate Markov chain model having fewer model parameters [6]. However, both models in [6] and [18] can only incorporate positive associations among the ratings of the entities since λ_{jk} 's and the components of the matrices $P^{(jk)}$'s are all non-negative. The estimation method in Siu et al. [18] is developed based on the existence of a unique stationary distribution X for the process Y such that $X = QX$ holds when the ratings dataset is sufficiently long. The matrices $P^{(jk)}$ can be estimated by counting the transition frequencies of the states among the entities. Then, the main idea of the estimation method for λ_{ij} is to minimize the ‘‘distance’’ between the \hat{X} and $\hat{Q}\hat{X}$, where \hat{X} is the estimate of X and \hat{Q} is the matrix Q having P_{ij} being estimated. We note that a long ratings dataset might be required to provide an accurate estimate for Q . However, in practice, it might be difficult and costly to obtain long ratings datasets. The accuracy of the estimate Q can be enhanced by speeding up the rate of convergence, and, this issue has not been discussed in Siu et al. [18].

In the sequel, we shall present a new multivariate Markov chain model for ratings of credit entities. This multivariate Markov chain model originates from the work of [2]. First, we let

$$Z_{t+1} = \frac{1}{m-1} (\mathbf{1} - X_t) ,$$

where $\mathbf{1}$ is the vector of all ones and the factor $(m-1)^{-1}$ is the normalization constant and the number of possible states $m \geq 2$. We note that there is a negative association between Z_{t+1} and X_t , i.e., if X_t increases (decreases), Z_{t+1} decreases (increases). To make the model more parsimonious (i.e., to reduce the number of parameters), we suppose that $P^{(ij)} = I$ when $i \neq j$. This idea has been adopted, justified and shown to be effective in modelling sales demand data sequences in Ching et al. [4]. Then, under the new multivariate Markov chain model, the dynamics of the probability vectors of the ratings

are governed by the following equations:

$$\begin{pmatrix} X_{t+1}^{(1)} \\ X_{t+1}^{(2)} \\ \vdots \\ X_{t+1}^{(s)} \end{pmatrix} = \underbrace{\Lambda^+ \begin{pmatrix} X_t^{(1)} \\ X_t^{(2)} \\ \vdots \\ X_t^{(s)} \end{pmatrix}}_{\text{Positive associated part}} + \frac{1}{m-1} \underbrace{\Lambda^- \begin{pmatrix} \mathbf{1} - X_t^{(1)} \\ \mathbf{1} - X_t^{(2)} \\ \vdots \\ \mathbf{1} - X_t^{(s)} \end{pmatrix}}_{\text{Negative associated part}} \quad (3)$$

where

$$\Lambda^+ = \begin{pmatrix} \lambda_{1,1}P^{(11)} & \lambda_{1,2}I & \cdots & \lambda_{1,s}I \\ \lambda_{2,1}I & \lambda_{2,2}P^{(22)} & \cdots & \lambda_{2,s}I \\ \cdots & \cdots & \cdots & \cdots \\ \lambda_{s,1}I & \cdots & \lambda_{s,s-1}I & \lambda_{s,s}P^{(ss)} \end{pmatrix}$$

and

$$\Lambda^- = \begin{pmatrix} \lambda_{1,-1}P^{(11)} & \lambda_{1,-2}I & \cdots & \lambda_{1,-s}I \\ \lambda_{2,-1}I & \lambda_{2,-2}P^{(22)} & \cdots & \lambda_{2,-s}I \\ \cdots & \cdots & \cdots & \cdots \\ \lambda_{s,-1}I & \cdots & \lambda_{s,-s+1}I & \lambda_{s,-s}P^{(ss)} \end{pmatrix}.$$

Here $\lambda_{i,j} \geq 0$ for $i = 1, 2, \dots, s$ and $j = \pm 1, \dots, \pm s$ and

$$\sum_{j=-s}^s \lambda_{i,j} = 1.$$

In Equation (3), Λ^+ and Λ^- represent the transition matrices for the positive and negative associations, respectively. From the second term of Equation (3), an increase (a decrease) in X_t leads to a decrease (or an increase) in X_{t+1} . In other words, there is a negative association between X_t and X_{t+1} . The term $\frac{1}{m-1}$ is a normalization constant, which makes $X_{t+1}^{(j)}$ on the left hand side of Equation (3) a probability vector, for each $j = 1, 2, \dots, s$.

We shall then investigate the existence of a stationary distribution under the new multivariate Markov chain model, the rate of convergence and how to speed up the rate of convergence. First, we represent the new multivariate Markov chain model as the

following vector-valued difference equation:

$$\begin{aligned}
X_{t+1} &= \begin{pmatrix} X_{t+1}^{(1)} \\ X_{t+1}^{(2)} \\ \vdots \\ X_{t+1}^{(s)} \end{pmatrix} \\
&= \begin{pmatrix} H_{1,1} & H_{1,2} & \cdots & H_{1,s} \\ H_{2,1} & H_{2,2} & \cdots & H_{2,s} \\ \cdots & \cdots & \cdots & \cdots \\ H_{s,1} & H_{s,2} & \cdots & H_{s,s} \end{pmatrix} \begin{pmatrix} X_t^{(1)} \\ X_t^{(2)} \\ \vdots \\ X_t^{(s)} \end{pmatrix} + \frac{1}{m-1} \begin{pmatrix} J_{1,-1} & J_{1,-2} & \cdots & J_{1,-s} \\ J_{2,-1} & J_{2,-2} & \cdots & J_{2,-s} \\ \cdots & \cdots & \cdots & \cdots \\ J_{s,-1} & \cdots & J_{s,-s+1} & J_{s,-s} \end{pmatrix} \begin{pmatrix} \mathbf{1} \\ \mathbf{1} \\ \vdots \\ \mathbf{1} \end{pmatrix} \\
&\equiv M_s X_t + \mathbf{b}.
\end{aligned}$$

Here

$$H_{ij} = \begin{cases} (\lambda_{i,j} - \frac{\lambda_{i,-j}}{m-1})P^{(ii)} & \text{if } i = j \\ (\lambda_{i,j} - \frac{\lambda_{i,-j}}{m-1})I & \text{otherwise} \end{cases}$$

and

$$J_{ij} = \begin{cases} \lambda_{i,-j}P^{(ii)} & \text{if } i = j \\ \lambda_{i,-j}I & \text{otherwise.} \end{cases}$$

We note that

$$\begin{aligned}
X_{t+1} &= M_s^2 X_{t-1} + (I + M_s)\mathbf{b} = M_s^3 X_{t-2} + (I + M_s + M_s^2)\mathbf{b} \\
&= M_s^{t+1} X_0 + \sum_{k=0}^t M_s^k \mathbf{b}
\end{aligned}$$

where $M_s^0 = I$.

Now the new multivariate Markov chain model has a stationary distribution if for a certain matrix norm $\|\cdot\|$ (let us choose $\|M_s\|_\infty$), $\|M_s\| < 1$. In this case, we have

$$\lim_{t \rightarrow \infty} X_t = \lim_{t \rightarrow \infty} \sum_{k=0}^t M_s^k \mathbf{b} = (I - M_s)^{-1} \mathbf{b}.$$

We also note that

$$\|M_s\|_\infty \leq \max_{1 \leq k \leq s} \left\{ m \left| \lambda_{k,k} - \frac{\lambda_{k,-k}}{m-1} \right| + \sum_{k \neq i} \left| \lambda_{k,i} - \frac{\lambda_{k,-i}}{m-1} \right| \right\}.$$

Hence, by controlling the the value $\|M_s\|_\infty$, the convergence rate of the process to the stationary distribution can be speeded up. This means one can control the rate of convergence by imposing an upper bound $\alpha < 1$ and introducing the following additional

constraints:

$$m \left| \lambda_{k,k} - \frac{\lambda_{k,-k}}{(m-1)} \right| + \sum_{k \neq i} \left| \lambda_{k,i} - \frac{\lambda_{k,-i}}{(m-1)} \right| \leq \alpha \quad \text{for } i = 1, 2, \dots, s.$$

Clearly if a smaller value of α is chosen, the rate of convergence to the stationary distribution becomes faster. Thus, reasonably accurate estimates for the unknown parameters can be obtained even when the ratings dataset is short. Thus, reasonably accurate estimates for the unknown parameters can be obtained even when the ratings dataset is short.

3 Estimation of Model Parameters

In this section, we present efficient methods for estimating $P^{(jj)}$ and λ_{jk} . For each ratings series of a credit entity, we can adopt the method in [2, 3, 4] to estimate the transition probability matrix. In particular, given a ratings series of an entity, we can count the transition frequencies from one arbitrary state to the others and construct the transition frequency matrix for the ratings series. Then, an estimate of the transition probability matrix of the ratings of the entity can be obtained by normalization. Here, we need to estimate $O(s \times m^2)$ transition frequency matrices for the new multivariate Markov chain model for ratings of s entities, each of which has m possible rating classes. The stationary probability vector X can be estimated from the proportion of the occurrence of each rating class in each of the ratings series.

Based on the idea at the end of the last section, if we take $\|\cdot\|$ to be $\|\cdot\|_1$, an estimate of λ_{jk} can be obtained by solving the following optimization problem (see [2, 3, 4]):

$$\min_{\lambda} \sum_i |[b_{j,k} - \hat{\mathbf{x}}^{(j)}]_i|$$

Subject to

$$\left\{ \begin{array}{l} b_{j,k} = \sum_{k=1}^s \left((\lambda_{j,k} - \frac{\lambda_{j,-k}}{m-1}) \Delta_{jk} \hat{\mathbf{x}}^{(k)} + \frac{1}{m-1} \lambda_{j,k} \Delta_{jk} \mathbf{1} \right) \\ \sum_{k=-s}^s \lambda_{jk} = 1, \quad \forall j = 1, 2, \dots, s. \\ \lambda_{jk} \geq 0, \quad \forall k = \pm 1, \dots, \pm s, j = 1, 2, \dots, s. \\ m \left| \lambda_{k,k} - \frac{\lambda_{k,-k}}{(m-1)} \right| + \sum_{k \neq i} \left| \lambda_{k,i} - \frac{\lambda_{k,-i}}{(m-1)} \right| \leq \alpha \quad \text{for } k = 1, 2, \dots, s. \end{array} \right. \quad (4)$$

Here,

$$\Delta_{jk} = \begin{cases} P^{(jj)} & \text{if } j = k \\ I & \text{if } j \neq k. \end{cases}$$

We can then formulate Problem (4) as s linear programming problems as below (see, for

example, [7, (p. 221)].

$$\min_{\lambda} \sum_i w_i$$

Subject to

$$\left\{ \begin{array}{l} w_i \geq [b_{j,k} - \hat{\mathbf{x}}^{(j)}]_i \\ w_i \geq -[b_{j,k} - \hat{\mathbf{x}}^{(j)}]_i \\ \sum_{k=-s}^s \lambda_{jk} = 1, \quad \forall j. \\ \lambda_{jk} \geq 0, \quad \forall k = \pm 1, \dots, \pm s, \forall j. \\ A\Lambda_j \leq \alpha, \forall j. \end{array} \right. \quad (5)$$

where

$$\Lambda_j = (\lambda_{j,1}, \lambda_{j,2}, \dots, \lambda_{j,s}, \lambda_{j,-1}, \dots, \lambda_{j,-s})^T, \quad A_i = [A_{1j}, A_{2j}],$$

$$A_{1j} = \begin{pmatrix} 1 & \cdots & m & \cdots & 1 & 1 \\ 1 & \cdots & m & \cdots & 1 & -1 \\ 1 & \cdots & m & \cdots & -1 & 1 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ -1 & \cdots & -m & \cdots & -1 & 1 \\ -1 & \cdots & \underbrace{-m}_{j \text{ th column}} & \cdots & -1 & -1 \end{pmatrix}$$

and

$$A_{2i} = -\frac{1}{m-1} A_{1i}.$$

Here, A_{1i} covers all of the rows in which each component takes one of the two possible values, 1 and -1 . Hence, A has $s \times 2^s$ rows.

We remark that the s linear programming problems can be solved using Excel spreadsheets. This provides market practitioners with a handy way to estimate the unknown parameters in the multivariate Markov chain model for ratings. The estimation procedure can be implemented with reasonably computational time and effort when there is a modest number of ratings series.

4 Practical Implementation and Estimation Results

In this section, we shall first provide a numerical example to demonstrate the practical implementation of the multivariate Markov chain model for hypothetical ratings. Then, we shall present the estimation results for the model using two sets of ratings data, namely, the ratings data of two assets extracted from Kijima et al. [14] and the risk ratings data of bonds from four important industrial sectors in United States, namely,

consumer, energy, transportation and media, extracted by the IHMM in Ching et al. [5] using the quarterly bond defaults data of the four sectors from Giampieri et al. [12]. As in [12] and [5], we suppose that bonds from each of the sectors are assigned one of the two risk ratings, namely, enhanced risk state and normal risk state. The values of the bonds depend on their risk ratings. In particular, the bonds with enhanced risk state potentially have lower values than those with normal risk state. We shall compare the empirical performance of the multivariate Markov chain model with that of the model in Siu et al. [18]. The Bayesian Information Criterion (BIC) will be adopted here as a criterion for model selection. We shall also compare the risk measures, namely, the Credit Value-at-Risk (CVaR) and the Credit Expected Shortfall (CES), obtained from the new model with different values of the parameter α with those obtained from the model of Siu et al. [18] and document the impact of incorporating different degrees of negative associations on the risk measures.

4.1 A Numerical Example

Here, we consider the following three hypothetical ratings series, each of which has three possible rating classes: 1, 2, 3.

$$\begin{array}{l} \hline A : 1, 2, 2, 1, 3, 3, 2, 3, 1, 3, 3, 2, 2, 3, 2, 1, 2, 1, 2, 2 \\ B : 2, 2, 1, 1, 2, 1, 2, 1, 2, 2, 2, 2, 3, 2, 2, 1, 2, 1, 2, 2 \\ C : 3, 1, 1, 2, 2, 1, 2, 3, 1, 1, 3, 3, 3, 2, 3, 2, 3, 3, 1, 2 \\ \hline \end{array}$$

First, we estimate all of the transition probability matrices $P^{(jj)}$ by counting the frequencies of transitions from one rating class to the others and the estimates are presented as below:

$$p^{(11)} = \begin{pmatrix} 0.0000 & 0.3750 & 0.1667 \\ 0.6000 & 0.3750 & 0.5000 \\ 0.4000 & 0.2500 & 0.3333 \end{pmatrix},$$

$$p^{(22)} = \begin{pmatrix} 0.1667 & 0.4167 & 0.0000 \\ 0.8333 & 0.5000 & 1.0000 \\ 0.0000 & 0.0833 & 0.0000 \end{pmatrix}$$

and

$$p^{(33)} = \begin{pmatrix} 0.3333 & 0.2000 & 0.3750 \\ 0.5000 & 0.2000 & 0.2500 \\ 0.1667 & 0.6000 & 0.3750 \end{pmatrix}.$$

Here, we suppose that the upper bound α ranges from 0.1 to 1.0, with an increment of 0.1. For each value of α , the estimates of the parameters λ_{ij} 's can be computed by solving the linear programming problems described in Section 3. Here, we need to solve three linear programming problems since we have three ratings series.

To compare the empirical performances of different models and perform model selec-

tion, we adopt the BIC as a criterion, which is defined as below:

$$BIC = -2L + q \log n ,$$

where

$$L = \sum_{j=1}^s \left(\sum_{i_0, k_1, \dots, k_s=1}^m n_{i_0, k_1, \dots, k_s}^{(j)} \log I \right) ,$$

$$I = \sum_{l=1}^m \sum_{k=1}^s \left(\lambda_{jk} - \frac{1}{m-1} \lambda_{j,-k} \right) p_{i_0, k_l}^{(jk)} + \frac{1}{m-1} \lambda_{j,-k}$$

is the log-likelihood of the model, and,

$$n_{i_0, k_1, k_2, \dots, k_s}^{(j)} = \sum x_{n+1}^{(j)}(i_0) x_n^1(k_1) x_n^2(k_2) \cdots x_n^s(k_s).$$

Here, q represents the number of independent parameters and n is the length of ratings series. The smaller the value of the BIC, the better the model is. In other words, the “best” model is given by the one with the smallest BIC.

For the sake of comparison, we also give the empirical results for the model of Siu et al. [18]. The BICs of the new model with different values of α and the model of Siu et al. [18] are reported in Table 4.1.

Table 4.1. The BIC for Different Models

Models	BIC
The model in [18]	607.4
The New Model ($\alpha = 0.1$)	247.0
The New Model ($\alpha = 0.2$)	246.2
The New Model ($\alpha = 0.3$)	245.3
The New Model ($\alpha = 0.4$)	245.5
The New Model ($\alpha = 0.5$)	247.2
The New Model ($\alpha = 0.6$)	246.5
The New Model ($\alpha = 0.7$)	245.9
The New Model ($\alpha = 0.8$)	245.5
The New Model ($\alpha = 0.9$)	245.2
The New Model ($\alpha^* = 1.0$)	244.9

From Table 4.1, the “best” model is the new model with $\alpha^* = 1.0$. This model is also better than the model of Siu et al. [18] in terms of fitting the hypothetical ratings data since the former has a smaller BIC than the latter. This might be attributed to the existence of negative associations among the three hypothetical ratings series. The optimal value $\alpha^* = 1.0$ is the most appropriate one to accommodate the given length of the ratings series according to the BIC. When the optimal value of $\alpha^* = 1.0$, the estimate

of $[\lambda_{ij}]$ is given by:

$$\begin{pmatrix} \lambda_{1,1} & \lambda_{1,2} & \lambda_{1,3} & \lambda_{1,-1} & \lambda_{1,-2} & \lambda_{1,-3} \\ \lambda_{2,1} & \lambda_{2,2} & \lambda_{2,3} & \lambda_{2,-1} & \lambda_{2,-2} & \lambda_{2,-3} \\ \lambda_{3,1} & \lambda_{3,2} & \lambda_{3,3} & \lambda_{3,-1} & \lambda_{3,-2} & \lambda_{3,-3} \end{pmatrix} = \begin{pmatrix} 0.2146 & 0.1599 & 0.1101 & 0.3094 & 0.0876 & 0.1184 \\ 0.0000 & 0.4707 & 0.0000 & 0.0000 & 0.3383 & 0.1910 \\ 0.2094 & 0.0000 & 0.0157 & 0.0000 & 0.3246 & 0.4503 \end{pmatrix}$$

and

$$M_1 = \begin{pmatrix} 0.0000 & 0.0225 & 0.01000 & 0.1161 & 0.0000 & 0.0000 & 0.0509 & 0.0000 & 0.0000 \\ 0.0359 & 0.0225 & 0.0299 & 0.0000 & 0.1161 & 0.0000 & 0.0000 & 0.0509 & 0.0000 \\ 0.0240 & 0.0150 & 0.0200 & 0.0000 & 0.0000 & 0.1161 & 0.0000 & 0.0000 & 0.0509 \\ \hline 0.0000 & 0.0000 & 0.0000 & 0.0503 & 0.1256 & 0.0000 & -0.0955 & 0.0000 & 0.0000 \\ 0.0000 & 0.0000 & 0.0000 & 0.2512 & 0.1507 & 0.3015 & 0.0000 & -0.0955 & 0.0000 \\ 0.0000 & 0.0000 & 0.0000 & 0.0000 & 0.0251 & 0.0000 & 0.0000 & 0.0000 & -0.0955 \\ \hline 0.2094 & 0.0000 & 0.0000 & -0.16230 & 0.0000 & 0.0000 & -0.0698 & -0.0419 & -0.0785 \\ 0.0000 & 0.2094 & 0.0000 & 0.0000 & -0.1623 & 0.0000 & -0.1047 & -0.0419 & -0.0524 \\ 0.0000 & 0.0000 & 0.2094 & 0.0000 & 0.0000 & -0.1623 & -0.0349 & -0.1257 & -0.0785 \end{pmatrix}$$

and

$$\mathbf{b}_1 = (0.2868, 0.3312, 0.2551, 0.1942, 0.4902, 0.1096, 0.3668, 0.3762, 0.4193)^T.$$

We remark that the computational time for the multivariate model [18] is 0.11 seconds, while the computation time for the new model with $\alpha = 1.0$ is 0.125 seconds in this numerical example. Hence, the computational times for the two models are similar.

4.2 Ratings Data of Two Assets

Now, we use the credit ratings data of two assets extracted from Kijima et al. [14]. This dataset has been adopted in Siu et al. [18] and is presented as follows:

Table 4.2. The Ratings of the Two Assets (Taken from Kijima et al. [14])

Asset/Year	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18
1	4	4	4	4	4	5	5	5	5	5	4	4	4	4	5	5	5	5
2	4	4	4	4	4	4	4	4	5	5	5	4	4	4	5	5	5	5

Here, the state “1” represents the highest rating class, say AAA. The estimates of the parameters λ_{jk} 's for different values of the upper bound α are reported in Table 4.3.

Table 4.3: $\Lambda^+ = (\lambda_{11}, \lambda_{12}, \lambda_{21}, \lambda_{22})$, $\Lambda^- = (\lambda_{1,-1}, \lambda_{1,-2}, \lambda_{2,-1}, \lambda_{2,-2})$

Models	Λ^+	Λ^-
$(\alpha = 0.1)$	(0.1333,0.0111,0.8556,0.0000)	(0.0111,0.1333,0.0000,0.8556)
$(\alpha = 0.2)$	(0.1417,0.0222,0.8361,0.0000)	(0.0222,0.1417,0.0000,0.8361)
$(\alpha = 0.3)$	(0.1500,0.0333,0.8167,0.0000)	(0.0333,0.1500,0.0000,0.8167)
$(\alpha = 0.4)$	(0.1583,0.0444,0.7972,0.0000)	(0.0444,0.1583,0.0000,0.7972)
$(\alpha = 0.5)$	(0.1667,0.0556,0.7778,0.0000)	(0.0556,0.1667,0.0000,0.7778)
$(\alpha = 0.6)$	(0.1750,0.0667,0.7583,0.0000)	(0.0667,0.1750,0.0000,0.7583)
$(\alpha = 0.7)$	(0.1833,0.0778,0.7389,0.0000)	(0.0778,0.1833,0.0000,0.7389)
$(\alpha = 0.8)$	(0.1917,0.0889,0.7194,0.0000)	(0.0889,0.1917,0.0000,0.7194)
$(\alpha = 0.9)$	(0.2000,0.1000,0.7000,0.0000)	(0.1000,0.2000,0.0000,0.7000)
$(\alpha^* = 1.0)$	(0.2083,0.1111,0.6806,0.0000)	(0.1111,0.2083,0.0000,0.6806)

Table 4.4. displays the BIC for the new model with different values of the upper bound α and the model of Siu et al. [18].

Table 4.4. The BIC

Models	BIC
The model in [18]	922.4
$(\alpha = 0.1)$	500.1
$(\alpha = 0.2)$	497.6
$(\alpha = 0.3)$	495.3
$(\alpha = 0.4)$	493.1
$(\alpha = 0.5)$	491.0
$(\alpha = 0.6)$	489.0
$(\alpha = 0.7)$	487.1
$(\alpha = 0.8)$	485.3
$(\alpha = 0.9)$	483.5
$(\alpha^* = 1.0)$	481.8

From Table 4.4, for each value of the upper bound α , the new model provides a better fit to the credit ratings data than the model of Siu et al. [18]. Moreover, the new model with a large value of the upper bound provides a better fit to the credit ratings data. In this case, the “best” model is the new model with $\alpha^* = 1$. When this optimal value of α is used, we have following model.

$$M_2 = \begin{pmatrix} M_{11} & M_{12} \\ M_{21} & M_{22} \end{pmatrix}$$

where

$$M_{11} = \left(\begin{array}{cc|cc|cc|cc} 0.0000 & 0.0000 & 0.0000 & 0.0000 & 0.0000 & 0.0000 & 0.0000 & 0.0000 \\ 0.0000 & 0.0000 & 0.0000 & 0.0000 & 0.0000 & 0.0000 & 0.0000 & 0.0000 \\ \hline 0.0000 & 0.0000 & 0.0000 & 0.0000 & 0.0000 & 0.0000 & 0.0000 & 0.0000 \\ 0.0000 & 0.0000 & 0.0000 & 0.0864 & 0.0139 & 0.0000 & 0.0000 & 0.0000 \\ \hline 0.0000 & 0.0000 & 0.0000 & 0.0247 & 0.0972 & 0.0000 & 0.0000 & 0.0000 \\ 0.0000 & 0.0000 & 0.0000 & 0.0000 & 0.0000 & 0.0000 & 0.0000 & 0.0000 \\ \hline 0.0000 & 0.0000 & 0.0000 & 0.0000 & 0.0000 & 0.0000 & 0.0000 & 0.0000 \\ 0.0000 & 0.0000 & 0.0000 & 0.0000 & 0.0000 & 0.0000 & 0.0000 & 0.0000 \end{array} \right),$$

$$M_{22} = \left(\begin{array}{cc|cc|cc|cc} 0.0000 & 0.0000 & 0.0000 & 0.0000 & 0.0000 & 0.0000 & 0.0000 & 0.0000 \\ 0.0000 & 0.0000 & 0.0000 & 0.0000 & 0.0000 & 0.0000 & 0.0000 & 0.0000 \\ \hline 0.0000 & 0.0000 & 0.0000 & 0.0000 & 0.0000 & 0.0000 & 0.0000 & 0.0000 \\ 0.0000 & 0.0000 & 0.0000 & 0.0909 & 0.0185 & 0.0000 & 0.0000 & 0.0000 \\ \hline 0.0000 & 0.0000 & 0.0000 & 0.0202 & 0.0926 & 0.0000 & 0.0000 & 0.0000 \\ 0.0000 & 0.0000 & 0.0000 & 0.0000 & 0.0000 & 0.0000 & 0.0000 & 0.0000 \\ \hline 0.0000 & 0.0000 & 0.0000 & 0.0000 & 0.0000 & 0.0000 & 0.0000 & 0.0000 \\ 0.0000 & 0.0000 & 0.0000 & 0.0000 & 0.0000 & 0.0000 & 0.0000 & 0.0000 \end{array} \right),$$

$$M_{21} = 0.1111I$$

and

$$M_{12} = 0.1111I$$

and

$$\mathbf{b}_2 = (0.0000, 0.0000, 0.0000, 0.0878, 0.1067, 0.0000, 0.0000, 0.0000, 0.0000, 0.0000, 0.0000, 0.0000, 0.0000, 0.0000, 0.0000)^T.$$

To evaluate the CVaR and the CES, we consider a portfolio with one unit of each of the two assets. As in Siu et al. [18] (see pp. 552-553 therein), we evaluate the loss from each of the two assets by first considering an unit interval and its uniform partitions $\cup_{i=1}^8 P_i$ with

$$P_i = \left[\frac{i-1}{8}, \frac{i}{8} \right),$$

for each $i = 1, 2, \dots, 8$.

We suppose that for each $j = 1, 2$ and each $i = 1, 2, \dots, m$, the loss from the j^{th} asset $\mathcal{L}_{t+1}(Y_{t+1}^{(j)})$ given $Y_{t+1}^{(j)} = e_i$ takes values from P_i and it follows an uniform distribution on the interval P_i . Given $Y_{t+1}^{(j)} = e_i$, we can then simulate the loss $\mathcal{L}_{t+1}(Y_{t+1}^{(j)})$ from the uniform distribution on P_i . The simulated losses for different assets and different rating classes are presented in Table 4.5.

Table 4.5. Simulated Losses for the Two Data Sequences

	1	2	3	4	5	6	7	8
1	0.1188	0.2009	0.3614	0.4321	0.6027	0.7019	0.8652	0.8970
2	0.0289	0.1857	0.3453	0.3773	0.5556	0.7240	0.8423	0.9257

The aggregate loss of the portfolio at time $t + 1$ is given by:

$$\mathcal{L}_{t+1}(Y_{t+1}) = \mathcal{L}_{t+1}^{(1)}(Y_{t+1}^{(1)}) + \mathcal{L}_{t+1}^{(2)}(Y_{t+1}^{(2)}) .$$

The CVaR and the CES are then evaluated by the formulae in Siu et al. [18] (see Formulae (11) - (13) for CVaR and Formula (14) for CES therein). The numerical results of the CVaR and the CES for the new model with different α and the model of Siu et al. [18] are presented in Table 4.6.

Table 4.6. The CVaR and the CES

Models	BIC	CVaR	CES	CVaR	CES
		($\beta = 0.05$)	($\beta = 0.05$)	($\beta = 0.01$)	($\beta = 0.01$)
The model in [18]	922.4	1.1593	1.1593	1.1593	1.1593
($\alpha = 0.1$)	500.1	1.1795	1.2030	1.2185	1.2290
($\alpha = 0.2$)	497.6	1.1830	1.2054	1.2201	1.2300
($\alpha = 0.3$)	495.3	1.1859	1.2073	1.2213	1.2309
($\alpha = 0.4$)	493.1	1.1882	1.2088	1.2224	1.2316
($\alpha = 0.5$)	491.0	1.1902	1.2101	1.2233	1.2322
($\alpha = 0.6$)	489.0	1.1919	1.2112	1.2240	1.2327
($\alpha = 0.7$)	487.1	1.1933	1.2122	1.2246	1.2331
($\alpha = 0.8$)	485.3	1.1946	1.2130	1.2252	1.2335
($\alpha = 0.9$)	483.5	1.1957	1.2138	1.2257	1.2338
($\alpha^* = 1.0$)	481.8	1.1966	1.2144	1.2261	1.2341

From Table 4.6, we see that the CES is always greater than or equal to the corresponding CVaR for a given probability level β and that the CES and the CVaR with $\beta = 0.01$ are greater than or equal to those with $\beta = 0.05$. These are consistent with the definition of the two risk measures. For the new model, both the CVaR and the CES with different values of the probability level β increase as the value of α increases, which indicates that the better the model is, the higher risk value the model has. This seems to be consistent with the intuition that a more strict measure of risk is obtained when the model becomes more restrictive. Our CES and CVaR with all different α are always greater than CES and CVaR obtained from model of Siu et al. [18] for both $\beta = 0.05$ and $\beta = 0.01$. In this case, the model incorporating both the positive and negative associations gives you the risk measures more strict than those obtained from the model that only incorporates the positive association. This reflects that without incorporating negative association presented in the ratings of the two assets leads to an under-estimation of the risk.

4.3 Risk Ratings Data for Bonds

In this section, we consider the risk ratings data for bonds from four industrial sectors in United States, namely, consumer, energy, media and transportation. These risk ratings data are extracted from the IHMM in Ching et al. [5] using the quarterly bonds defaults data for the four sectors from Giampieri et al. [12]. The bond defaults data set covers the period from January 1981 to December 2002 and the four sectors in United States and these data were originally taken from Standard & Poors' ProCredit6.2 database in Giampieri et al. [12]. Following Giampieri et al. [12] and Ching et al. [5], we suppose that the risk ratings of the bonds take one of the two possible values, namely, normal risk state (State 1) and enhanced risk state (State 2) and that bonds from each of the four sectors have the same ratings in each quarter. In other words, the bonds from each of the four sectors are homogeneous in their quarterly risk ratings. The length of quarterly risk ratings series of bonds for each of the four sectors is 88. Even through the bonds in the same sector have the same risk ratings. they may have different values. However, the risk ratings of the bonds do play an important role in their values. Here, we suppose that the values of the bonds depend on their risk ratings. In particular, the values of the bonds with enhanced risk state have lower values than bonds with normal risk state. In other words, the losses from trading bonds with enhanced risk state are higher than those from trading bonds with normal risk state. Table 4.7. reports the BIC of the new model with different values of α and the model of Siu et al. [18].

Table 4.7: The BIC

Models	BIC
The model in [18]	492.1
($\alpha = 0.1$)	445.4
($\alpha = 0.2$)	434.8
($\alpha = 0.3$)	424.9
($\alpha = 0.4$)	415.8
($\alpha = 0.5$)	407.6
($\alpha = 0.6$)	400.1
($\alpha = 0.7$)	393.7
($\alpha = 0.8$)	388.3
($\alpha = 0.9$)	384.2
($\alpha^* = 1.0$)	380.7

From Table 4.7, the optimal model is the new model with $\alpha^* = 1.0$. When this optimal value $\alpha^* = 1.0$ is applied, we have the following estimates of the unknown parameters of

the new model.

$$M_3 = \begin{pmatrix} 0.0504 & 0.0245 & 0.0032 & 0.0000 & 0.0611 & 0.0000 & 0.2255 & 0.0000 \\ 0.0055 & 0.0314 & 0.0000 & 0.0032 & 0.0000 & 0.0611 & 0.0000 & 0.2255 \\ \hline 0.0000 & 0.0000 & 0.1564 & 0.0952 & 0.0000 & 0.0000 & 0.1667 & 0.0000 \\ 0.0000 & 0.0000 & 0.0082 & 0.0714 & 0.0000 & 0.0000 & 0.0000 & 0.1667 \\ \hline -0.3750 & 0.0000 & 0.0000 & 0.0000 & 0.3545 & 0.1071 & 0.2500 & 0.0000 \\ 0.0000 & -0.3750 & 0.0000 & 0.0000 & 0.0205 & 0.2679 & 0.0000 & 0.2500 \\ \hline 0.0000 & 0.0000 & -0.1667 & 0.0000 & 0.0000 & 0.0000 & 0.4756 & 0.3000 \\ 0.0000 & 0.0000 & 0.0000 & -0.1667 & 0.0000 & 0.0000 & 0.0244 & 0.2000 \end{pmatrix}$$

and

$$\mathbf{b}_3 = (0.5726, 0.4606, 0.6699, 0.3260, 0.6250, 0.6250, 0.5000, 0.5000)^T.$$

To evaluate the CVaR and the CES, we consider a portfolio with one unit of a bond from each of the four sectors. In this case, we evaluate the loss from each of the four bonds by first considering an unit interval $[0, 1]$ and dividing it into two parts, namely, $P_1 := [0, 0.5)$ and $P_2 := [0.5, 1]$. Here, P_1 represents the state space of the losses from trading a bond with normal risk state and P_2 is the state space of losses from trading a bond with enhanced risk state.

Here, for each sector $j = 1, 2, 3, 4$ and each risk rating state $i = 1, 2$, the loss from trading a bond from the j^{th} sector $\mathcal{L}_{t+1}(Y_{t+1}^{(j)})$ given the risk rating state of the bond $Y_{t+1}^{(j)} = e_i$ takes values from P_i and it follows a uniform distribution on the interval P_i . Given $Y_{t+1}^{(j)} = e_i$, we simulate the loss $\mathcal{L}_{t+1}(Y_{t+1}^{(j)})$ from the uniform distribution on P_i . The simulated losses of the bonds with different sectors and different risk rating states are then displayed in Table 4.8.

Table 4.8: Simulated losses from trading the bonds

	1	2
1	0.1645	0.9147
2	0.2391	0.9781
3	0.2986	0.7978
4	0.0807	0.5144

The aggregate loss of the portfolio at time $t + 1$ is given by:

$$\mathcal{L}_{t+1}(Y_{t+1}) = \mathcal{L}_{t+1}^{(1)}(Y_{t+1}^{(1)}) + \mathcal{L}_{t+1}^{(2)}(Y_{t+1}^{(2)}) + \mathcal{L}_{t+1}^{(3)}(Y_{t+1}^{(3)}) + \mathcal{L}_{t+1}^{(4)}(Y_{t+1}^{(4)}).$$

The numerical results of the CVaR and the CES for the new model with different α and the model of Siu et al. [18] are presented in Table 4.9.

Table 4.9: The CVaR and the CES

Models	BIC	CVaR ($\beta = 0.05$)	CES ($\beta = 0.05$)	CVaR ($\beta = 0.01$)	CES ($\beta = 0.01$)
The model in [18]	492.1	2.0210	2.0210	2.2721	2.2721
($\alpha = 0.1$)	445.4	3.0273	3.2218	3.3495	3.4796
($\alpha = 0.2$)	434.8	3.1312	3.3049	3.4190	3.5352
($\alpha = 0.3$)	424.9	3.1901	3.3521	3.4584	3.5667
($\alpha = 0.4$)	415.8	3.2291	3.3833	3.4845	3.5876
($\alpha = 0.5$)	407.6	3.2572	3.4057	3.5032	3.6026
($\alpha = 0.6$)	400.1	3.2784	3.4227	3.5174	3.6139
($\alpha = 0.7$)	393.7	3.2950	3.4360	3.5285	3.6228
($\alpha = 0.8$)	388.3	3.3083	3.4466	3.5374	3.6299
($\alpha = 0.9$)	384.2	3.3219	3.4575	3.5465	3.6372
($\alpha^* = 1.0$)	380.7	3.3225	3.4580	3.5469	3.6375

In Table 4.9, the CES is always greater than or equal to the corresponding CVaR for a given probability level β and that the CES and the CVaR with $\beta = 0.01$ are greater than or equal to those with $\beta = 0.05$. This result is consistent with the definition of the two risk measures. As in Table 4.6, the numerical results in Table 4.9 are also consistent with the intuition that a more strict measure of risk is obtained when the model becomes more restrictive. The CVaR and the CES obtained from the optimal new model with all different values of α are larger than those obtained from the model of Siu et al. [18]. This reflects that without incorporating negative association presented in the quarterly risk ratings data of the bonds leads to an under-estimation of the risk from the portfolio. This has important implications for the practice of credit risk measurement and management.

5 Concluding Remarks

We developed a parsimonious and flexible multivariate Markov chain model for describing the dependency of transitions of ratings of credit risky entities. The number of parameters in the proposed model is only $O(sm^2 + s^2)$, which is relatively small compared with the number of parameters $O(s^2m^2)$ in the model of Siu et al. [18]. Our model can incorporate both the positive and negative associations between the time series of credit ratings. We also introduced an effective method to control the rate of convergence of the process to the stationary distribution and to handle the case when the dataset of ratings is short. This is an important feature since it is rather difficult or costly to obtain a long dataset of ratings in practice. We introduced an efficient method to estimate or calibrate the model and formulated the estimation problem as a linear programming problem, which can be implemented using spreadsheets. We compared the empirical performance and computational efficiency of the new model with that of the model of Siu et al. [18]. We found that the new model can provide a better fit to the two sets of ratings data and that the computational efficiencies of the two models are similar. We also compared credit

risk measures obtained from our model with those from Siu et al. [18] to illustrate the effect of incorporating both positive and negative associations on portfolio credit risks. We found that both the ratings of two assets taken from Kijima et al. [14] case and without incorporating negative association existed in the quarterly risk ratings data of the bonds extracted from the IHMM from Ching et al. [5] using the quarterly bonds defaults data from Giampieri et al. [12] leads to an underestimation of the risk from the portfolio of bonds from four industrial sectors, namely, consumer, media, transportation and energy. This finding has important implications for the practice of credit risk measurement and management. We also found that a more strict measure of risk is obtained when the model nested in the new model becomes more restrictive by controlling the upper bound of the rate of the convergence.

References

- [1] Artzner, P., Delbaen, F., 1995. *Default Risk Premium and Incomplete Markets*, Mathematical Finance **5** 187-195.
- [2] Ching, W., Fung, E. and Ng, M. 2002. *A Multivariate Markov Chain Model for Categorical Data Sequences and Its Applications in Demand Prediction*, IMA J. Manag. Math., **13** 187-199.
- [3] Ching, W. and Ng, M. 2006. *Markov Chains : Models, Algorithms and Applications*, International Series on Operations Research and Management Science, Springer, New York.
- [4] Ching, W., Li, L., Li, T. and Zhang S. 2007. *A New Multivariate Markov Chain Model with Applications to Sales Demand Forecasting*, Proceedings of the International Conference on Industrial Engineering and Systems Management 2007, Beijing, (in CD-ROM).
- [5] Ching, W., Siu, T., Li, L., Li, T. and Li, W. 2007. *Modeling Default Data via an Interactive Hidden Markov Model*, IMR Preprint Series 2007, No. 5., Department of Mathematics, the University of Hong Kong.
- [6] Ching W., Zhang S., and Ng M., 2007. On Multi-dimensional Markov Chain Models, to appear in Pacific Journal of Optimization.
- [7] V. Chvátal, 1983. *Linear Programming*, Freeman, New York.
- [8] Credit Suisse Financial Products 1997. Credit Risk+ a Credit Risk Management Framework. [http : //www.csfb.com/institutional/research/creditrisk.html](http://www.csfb.com/institutional/research/creditrisk.html).
- [9] Davis, M. and Lo, V., 2001. Infectious Defaults. Quantitative Finance, **1** 382-387.

- [10] Davis, M. and Lo, V., 2001. modeling Default Correlation in Bond Portfolio, In C. Alescander (ed.) *Mastering Risk Volume 2: Applications*. Financial Times, Prentice Hall, 141-151.
- [11] Duffie D. and Singleton K., 2003. *Credit Risk: Pricing, Measurement and Management*. Princeton University Press, Princeton.
- [12] Giampieri, G., Davis, M. Crowder M., 2005. Analysis of Default Data Using Hidden Markov Models. *Quantitative Finance*, **5** 27-34.
- [13] Jarrow, R., Lando, D., Turnbull, S. 1997. *A Markov Model for the Term Structure of Credit Spreads*, *Review of Financial Studies*, **10** 481-523.
- [14] Kijima, M., Komoribayashi, K., Suzuki, E., 2002. *A Multivariate Markov Model for Simulating Correlated Defaults*, *Journal of Risk*, **4** 1-32.
- [15] Li, D., 2000. *On Default Correlation: A Copula Function Approach*, *Journal of Fixed Income*, **9** (4) 43-54.
- [16] Madan, D., Unal, H., 1995. *Pricing the Risk of Default*, *Review of Derivatives Research* **2** 121-160.
- [17] Raftery, A., 1985. *A Model of Higher-Order Markov Chains*, *J. Royal Statist. Soc.*, **47** 528-539.
- [18] Siu, T., Ching, W., Ng, M. and Fung, E. 2005. *On Multivariate Credibility Approach for Portfolio Credit Risk Measurement*, *Quantitative Finance*, **5** 543-556.