On a Multivariate Markov Chain Model for Credit Risk Measurement

Tak-Kuen Siu *, Wai-Ki Ching, †Eric S. Fung ‡and Michael K. Ng §

December 21, 2005

Abstract

In this paper, we use credibility theory to estimate credit transition matrices in a multivariate Markov chain model for credit rating. A transition matrix is estimated by a linear combination of the prior estimate of the transition matrix and the empirical transition matrix. These estimates can be easily computed by solving a set of Linear Programming (LP) problems. The estimation procedure can be implemented easily on Excel spreadsheets without requiring much computational effort and time. The number of parameters is $O(s^2m^2)$, where s is the dimension of the categorical time series for credit ratings and m is the number of possible credit ratings for a security. Numerical evaluations of credit risk measures based on our model are presented.

Keywords: Correlated Credit Migrations, Linear Programming, Transition Matrices, Credibility Theory

*The Corresponding Author: Tak-Kuen Siu, Department of Actuarial Mathematics and Statistics, School of Mathematical and Computer Sciences, Heriot-Watt University, Edinburgh EH14 4AS, UK. Tel: (+44) 131 451 3906; Fax: (+44) 131 451 3249; E-mail: T.K.Siu@ma.hw.ac.uk

†Wai-Ki Ching, Department of Mathematics, University of Hong Kong, Pokfulam Road, Hong Kong. Tel: (852) 2859-2256; E-mail: wkc@maths.hku.hk

‡Eric, S. Fung, Department of Mathematics, University of Hong Kong, Pokfulam Road, Hong Kong. Tel: (852) 2859-2789; E-mail: ericfung@hkusua.hku.hk

§Department of Mathematics, Hong Kong Baptist University, Kowloon Tong, Hong Kong; Email: mng@math.hkbu.edu.hk.

1 Introduction

Modelling credit risk is an important topic in quantitative finance and risk management. Recently, there has been much interest in modelling the dependency of credit risks of securities in portfolios due to the practical importance of risk analysis of portfolios of credit risky securities. The specification of the model that explains and describes the dependency of the credit risks can have significant implications in measuring and managing credit risky portfolios. There are two major approaches to modelling the dependency of credit risky securities in industries, namely copulas and Monte-Carlo simulation. Copulas have long been used in statistics, in particular survival analysis, and have also been widely applied in actuarial science. Li (2000) introduced the use of copula functions for credit risk measurement. Since then, copulas have become a popular tool for credit risk analysis in the finance and insurance industries. The main advantages of copulas are that they can capture the dependency of credit risks when the credit loss distributions do not belong to the elliptical class and that they can incorporate the dependency of more than two credit risks. Embrechts et al. (1999) documented the potential pitfalls of correlations and introduced the use of copulas for modelling dependent risks when the multivariate distribution is asymmetric. The Monte Carlo simulation technique is usually adopted together with copulas for modelling the dependency of credit risks. The monograph by Cherubini et al. (2004) provided a comprehensive discussion on the use of simulation methods with various copulas for modelling dependent risks.

Modelling the dynamics of transitions between credit ratings is vital for credit risk analysis in the finance industry. The discrete-time homogeneous Markov chain model has been used by academic researchers and market practitioners to model such transitions over time. The transition matrix represents the likelihood of the future evolution of the ratings. In practice, the transition matrix can be estimated from empirical data for credit ratings. Standard & Poor, Moodys and Fitch are the major providers of credit rating data. They provide and update from time to time historical data for individual companies and for countries. Credit Metrics (Gupton et al. (1997)) provides a very comprehensive account of practical implementation of transition matrices. For measuring and managing the risk of a credit portfolio, it is of practical importance to develop quantitative models that can describe the dependencies between the credit ratings of individual assets in the portfolio since the losses from the individual assets depend on their credit ratings. A multivariate Markov chain model provides a natural and convenient way to describe these dependencies.

Kijima et al. (2002) used a multivariate Markov chain model to simulate the evolution of correlated ratings of several credit risks. They applied their model to questions of pricing and risk measurement. They estimated the unknown parameters in their model by minimizing the squared error based on historic rating data only. Thomas et al. (2002) suggested that historic rating data alone is not adequate to reflect future movements in ratings when the future may not evolve smoothly from the past experience. The market view is a mixture of beliefs determined by both historic movements in ratings and a more extreme view of ratings. The latter can be specified by subjective views or other expert opinion. Incorporating expert opinion on future movements in ratings also plays a vital role in measuring and managing the risk of credit portfolios. From the perspectives of risk measurement and management, both regulators and risk managers may wish to use a more conservative view or pessimistic outlook than that predicted by using historic data only. It is of practical relevance to incorporate expert opinion into tools for credit risk measurement and management (see Alexander (2005) for an excellent discussion).

In this paper, we investigate the use of a discrete-time homogeneous multivariate Markov chain model for dependent credit risks introduced by Ching et al. (2002). 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. The multivariate Markov chain model can handle both temporal and cross-sectional dependencies of categorical time series for credit ratings. It is very useful for handling time series generated from similar sources or from the same source. In the model, the number of parameters involved is $O(s^2m^2)$, where s is the dimension of the categorical time series for credit ratings and m is the number of possible credit ratings of a security. We need $C_2^s = s(s-1)/2$ parameters to capture all possible pairwise correlations (or dependencies) between any two categorical time series for credit ratings and there are m possible credit ratings for each security. We have a $m \times m$ transition matrix. We cannot further reduce the factor of order s^2 unless it is given that some parameters are zero in advance. Here, we consider the simple case, namely the first-order multivariate Markov chain model. It is possible to extend this to a higher-order multivariate Markov chain model as discussed in Ching et al. (2004). The structure of our multivariate Markov chain model is different from that of Kijima et al. (2002). Kijima et al. (2002) assumed that the change in the credit ratings over a period called the 'credit increment' is specified by a single index model consisting of two components, namely the systematic part described by a single common factor and the firm-specific component. Their model is more suitable than our model for pricing credit risky securities. However, our model is more general and more analytically tractable than the model in Kijima et al. (2002). It requires fewer structural assumptions relating to the change in credit ratings. However, its applicability is limited by the number of parameters involved. When both the dimension of the categorical time series for credit ratings and the number of possible credit ratings are high, the simulation approach adopted in Kijima et al. (2002) is more applicable.

We employ actuarial credibility theory to combine two sources of information for estimating the transition matrix and other unknown model parameters in the multivariate Markov chain model. Our approach provides a consistent and convenient way to incorporate both historical rating data and another source of information, for instance, expert opinion or a subjective view. In contrast, the estimation method in Kijima et al. (2002) can only incorporate one source of information, namely historic rating data. We provide an estimate of the transition matrix as a linear combination of the empirical transition matrix and a prior transition matrix. The empirical transition matrix can be specified based on historic rating data while the prior transition matrix can be determined by expert opinion or some other subjective view. Hence, our model is more suitable for measuring and managing the risk of credit portfolios when the market view is a mixture of beliefs based on both historic data and expert opinion. In practice, it is difficult to obtain plentiful historic rating data, in which, the role of the prior transition matrix becomes more important. We can formulate our estimation problem as a set of Linear Programming (LP) problems, which is easy to implement and more computationally efficient compared with the minimization of the square error in Kijima et al. (2002). Our estimation procedure can be implemented easily using Excel spreadsheets. Our approach also aims to highlight the interplay between actuarial credibility theory and risk measurement. Some related works along this direction include Siu and Yang (1999), Siu et al. (2001) and Woo and Siu (2004).

The rest of the paper is organized as follows. In Section 2, we introduce the multivariate Markov chain model for credit risk measurement. Section 3 presents the credibility approach for estimating the model parameters. We discuss the portfolio credit risk measures under the multivariate Markov chain model in Section 4. Numerical examples are given to illustrate the effectiveness of the model for the evaluation of portfolio credit risk measures in Section 5. Finally, concluding remarks are given in Section 6.

2 The Model

Credit risk measurement and management are of practical importance and relevance in finance and banking. There are two important models for credit risk measurement and management, namely the structural approach and the reduced-form approach. The structural approach was originally proposed by Merton (1974) and assumes that the value of a corporation's asset is driven by a Geometric Brownian Motion. In this case, default is an endogenous event which occurs when the value of the corporation's asset triggers an *a priori* threshold level. The structural approach facilitates the rapid development of different models for pricing corporate debts and credit derivatives. It is also one of the most important models used to analyse the financial leverage and capital structure of a firm from the perspective of corporate finance. However, in practice, the value of a firm's asset is not observable. This can induce the technical difficulty of modelling the value of the firm's asset as an exogenously given variable. The reduced-form approach was considered by Artzner and Delbaen (1995), Duffie et al. (1996), Jarrow and Turnbull (1995) and Madan and Unal (1995). It assumes that default is an exogenous event and its occurrence is governed by a random point process. The reduced-form approach focuses on modelling the hazard rate of default. In our setting, we adopt the reduced-form approach and assume that the future transitions between ratings of correlated credit risks in a portfolio only depend on their current ratings. We consider modelling the dynamics of the ratings by the multivariate Markov chain model introduced by Ching et al. (2002). In the following, we present the main idea of the multivariate Markov chain model for credit risk measurement.

First, we fix a complete probability space $(\Omega, \mathcal{F}, \mathcal{P})$. Suppose $\mathcal{T}$ represents the time index set $\{0, 1, 2, \ldots, \infty\}$ on which all economic activities take place. In our setting, we consider a portfolio with n correlated credit risks, for instance, n corporate bonds with correlated credit ratings. On $(\Omega, \mathcal{F}, \mathcal{P})$, we define n categorical time series $Y^{(1)}, Y^{(2)}, \ldots, Y^{(n)}$ with common time index set $\mathcal{T}$. Let $\mathcal{S}$ denote a set of unit basis vectors $\{e_1, e_2, \ldots, e_m\}$ in R^m, where $e_i = (0, \ldots, 0, \overbrace{1}^{i^{th} entry}, 0, \ldots, 0)^T \in R^m$. For each $j = 1, 2, \ldots, n$, $Y^{(j)} := \{Y_t^{(j)}\}_{t \in \mathcal{T}}$ is a discrete-time and finite-state stochastic process with state space $\mathcal{S}$. Note that $\{Y_t^{(j)}\}_{t \in \mathcal{T}}$ represents the random transitions of the ratings of the j^{th} credit risk. In particular, the event $\{\omega \in \Omega | Y_t^{(j)}(\omega) = e_i\}$ means that the j^{th} credit risk is in the i^{th} rating class at time t. Define the space $\hat{\mathcal{S}}$ as

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

For each $j = 1, 2, \ldots, n$, we denote the dynamics of the discrete probability distributions $\{X_t^{(j)}\}_{t \in \mathcal{T}}$ for $Y^{(j)}$, where $X_t^{(j)} \in \hat{\mathcal{S}}$ for each $t \in \mathcal{T}$. In particular, for each $t \in \mathcal{T}$, the i^{th} entity of the probability vector $X_t^{(j)}$ represents the probability that the j^{th} credit risk is in the i^{th} rating class at time t. It is interesting to note that $\{X_t^{(j)}\}_{t \in \mathcal{T}}$ can also be used to represent the ratings of the j^{th} credit risk at time t, for each $j = 1, 2, \ldots, n$. Suppose that the j^{th} credit risk is in the i^{th} rating class at time t; that is, $Y_t^{(j)} = e_i$. This means that the probability of the j^{th} credit risk being in the i^{th} rating class at time t is equal to one. Hence, we have

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

Let $P^{(jk)}$ be a transition matrix from the states in the k^{th} sequence to the states in the j^{th} sequence. For each $j = 1, 2, \ldots, n$, $P^{(jj)}$ represents the transition matrix for the j^{th} categorical time series Y^j. Write $X_t^{(k)}$ for the state probability distribution of the k^{th} sequences at time t. Then, we assume that the dynamics of the probability distributions of the ratings for the j^{th} credit risk are governed by the following equation:

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

where

$$\lambda_{jk} \geq 0, \quad 1 \leq j, k \leq n \quad \text{and} \quad \sum_{k=1}^{n} \lambda_{jk} = 1, \quad \text{for} \quad j = 1, 2, \cdots, n. \tag{2}$$

The interpretation of Equation (1) is that the state probability distribution of the j^{th} categorical time series Y^j at time $(t+1)$ depends on the weighted average of

$P^{(jk)}X_t^{(k)}$ at time t. It is not difficult to check that the multivariate categorical time series $Y := (Y^{(1)}, Y^{(2)}, \ldots, Y^{(n)})$ forms a multivariate Markov chain under Equation (1). For each $t \in \mathcal{T}$, let Y_t be $(Y_t^{(1)}, Y_t^{(2)}, \ldots, Y_t^{(n)})$ and $\mathcal{F}_t$ the information set generated by the market observations $\{Y_0, Y_1, \ldots, Y_t\}$ of the ratings of the n credit risks up to and including time t. Since Y_t is known given $\mathcal{F}_t$, $Y_t = (e_{i_1}, e_{i_2}, \ldots, e_{i_n})$, for some $i_1, i_2, \ldots, i_n \in \{1, 2, \ldots, m\}$ given $\mathcal{F}_t$. In this case, $X_t^{(j)} = Y_t^{(j)} = e_{i_j}$ for each $t \in \mathcal{T}$ and $j = 1, 2, \ldots, n$. From Equation (1), we notice that the probability distribution $X_{t+1}^{(j)}$ of $Y_{t+1}^{(j)}$ given $\mathcal{F}_t$ depends only on $(Y_t^{(1)}, Y_t^{(2)}, \ldots, Y_t^{(n)})$. Hence, the multivariate Markov property follows. In this case, the conditional probability distribution for the ratings of the j^{th} credit risk at time $t+1$ depends on the ratings of all credit risks in the portfolio at time t. We can write Equation (1) in the following matrix form:

$$X_{t+1} \equiv \begin{pmatrix} X_{t+1}^{(1)} \\ X_{t+1}^{(2)} \\ \vdots \\ X_{t+1}^{(n)} \end{pmatrix} = \begin{pmatrix} \lambda_{11}P^{(11)} & \lambda_{12}P^{(12)} & \cdots & \lambda_{1n}P^{(1n)} \\ \lambda_{21}P^{(21)} & \lambda_{22}P^{(22)} & \cdots & \lambda_{2n}P^{(2n)} \\ \vdots & \vdots & \vdots & \vdots \\ \lambda_{n1}P^{(n1)} & \lambda_{n2}P^{(n2)} & \cdots & \lambda_{nn}P^{(nn)} \end{pmatrix} \begin{pmatrix} X_t^{(1)} \\ X_t^{(2)} \\ \vdots \\ X_t^{(n)} \end{pmatrix} \equiv QX_t$$

or

$$X_{t+1} = QX_t.$$

Although the column sum of Q is not equal to one (the column sum of $P^{(jk)}$ is equal to one), we still have the following proposition. The proof of the proposition can be found in Ching, Fung and Ng (2002).

Proposition 1 *Suppose that* $P^{(jk)}$ $(1 \leq j, k \leq n)$ *are irreducible and* $\lambda_{jk} > 0$ *for* $1 \leq j, k \leq n$. *Then, there is a vector*

$$X = [X^{(1)}, X^{(2)}, \cdots, X^{(n)}]^T$$

such that $X = QX$ *and*

$$\sum_{i=1}^{m} [X^{(j)}]^i = 1, \ \ 1 \leq j \leq n \ ,$$

where $[\cdot]^i$ *denote the* i^{th} *entry of the corresponding vector.*

The vector X in Proposition 1 contains the stationary probability distributions for the ratings of all credit risks in the portfolio. That is, for each j, $X^{(j)}$ represents the probability distribution for the ratings of the j^{th} credit risk in the long-run. For various mathematical and statistical properties of the multivariate Markov chain, refer to the recent paper by Ching et al. (2002).

Kijima et al. (2002) considered the use of the multivariate Markov model for simulating dependent credit risks. Their approach was based on the single-index model of modern portfolio theory and can be considered to be an extension of the credit risk model by Jarrow et al. (1997). They described the change in the credit ratings (or credit increment) as the single-index model, which consists of two components, namely the systematic component described by a single common factor and the firm-specific component. Compared with the model in Kijima et al. (2002), our model is more general and more analytical tractable. However, our model has more parameters than the model by Kijima et al. (2002). This limits its applicability. However, one cannot further reduce the order of the number of parameters given the generality of our model. Nevertheless, the estimation procedure of our model can be implemented easily on Excel spreadsheets without requiring much computational effort and time. Our model is intuitively appealing. It supposes that the transitions of the credit ratings are exogenous and describes the random behavior of transitions of the credit ratings directly.

The multivariate Markov chain model considered here is homogeneous. However, Das et al. (2004) provided a very comprehensive empirical investigation of correlated default risk and showed that the joint correlated default probabilities varied substantially over time. Their empirical result suggests that a homogeneous Markov chain is not appropriate for modelling the transitions of the ratings of a credit risk over time since it cannot describe the time-varying behavior of default risk. Although a non-homogeneous Markov chain may provide a more realistic way to describe the time-varying behavior of default risk, it makes the estimation procedure much more complicated and the model less analytically tractable. Hence, there is a tradeoff between the tractability of a model and a more realistic description of the empirical data.

In reality, the parameters in the matrix Q are unknown to market practitioners. In order to evaluate the risk of the credit portfolio, we have to estimate the unknown market parameters. We will discuss the use of credibility theory for the estimation in the next section.

3 Estimation of the Credit Transition Matrix

We relax the stringent assumption that the credit transition matrix is given in advance. In this section, we describe in some detail the case that the Q-matrix is unknown. We employ the idea of credibility theory to estimate the transition matrix and other unknown model parameters in the multivariate Markov chain model. Credibility theory has a long history in actuarial science. It has been widely applied in the actuarial discipline for calculating a policyholder's premium through experience rating of the policyholder's past claims. The main idea of credibility theory is to provide a consistent and convenient way to combine two different sources of information for premium calculation. It determines the weight that should be assigned to each source of information. Mowbray (1914), Bühlmann (1967) and Klugman, Panjer and Willmot (1997) together provide excellent accounts of actuarial credibility theory.

Bühlmann (1967) introduced a least squares approach for the estimation of credibility premiums without imposing stringent parametric assumptions for the claim models. Following an idea similar to the "Bühlmann least squares model", we provide market practitioners with an analytically tractable framework for estimating the Q-matrix without imposing any parametric assumptions. We assume that the estimate of each transition matrix in the Q-matrix can be represented as a linear combination of a prior transition matrix and the empirical transition matrix , where the empirical transition matrix is based on the frequencies of transitions between rating. Then, by noticing from Proposition 1 that there exists a vector X of stationary probability distributions such that $X = QX$, we can estimate the Q-matrix based on the vector X of the stationary distributions for the ratings.

Hu, Kiesel and Perraudin (2002) proposed an empirical Bayesian approach for the estimation of transition matrices for governments to evaluate and manage the risk of portfolios with credit exposures to emerging markets. Their estimator for transition matrices of ratings is also a linear combination of a prior matrix given by the empirical transition matrix (estimated directly from Standard & Poor's data) and a model-based updating matrix evaluated from the ordered probit model. They adopted empirical Bayesian techniques for estimating contingency tables and selected the weights of the linear combination of the prior and updating matrices by the goodness-of-fit $\mathcal{X}^2$-statistic. Our approach is different from their approach in that it is adapted to the multivariate Markov chain model and in that it adopts a different objective function associated with the method of optimization used for the estimation of the unknown model parameters. We present our approach as follows:

Suppose we are given a multivariate categorical time series of ratings of credit risks in the portfolio. We count the transition frequency $f_{rs}^{(jk)}$ from the rating r in the categorical time series $\{X_t^{(j)}\}$ of ratings of the j^{th} credit risk to the rating s in the categorical time series $\{X_t^{(k)}\}$ of the ratings of the k^{th} credit risk, where $r, s \in \{1, 2, \ldots, m\}$. Then, the transition frequency matrix can be constructed as follows:

$$F^{(jk)} = \begin{pmatrix} f_{11}^{(jk)} & \cdots & \cdots & f_{m1}^{(jk)} \\ f_{12}^{(jk)} & \cdots & \cdots & f_{m2}^{(jk)} \\ \vdots & \vdots & \vdots & \vdots \\ f_{1m}^{(jk)} & \cdots & \cdots & f_{mm}^{(jk)} \end{pmatrix}.$$

¿From $F^{(jk)}$, we get the estimate for the transition matrix $P^{(jk)}$ as follows:

$$\hat{P}^{(jk)} = \begin{pmatrix} \hat{p}_{11}^{(jk)} & \cdots & \cdots & \hat{p}_{m1}^{(jk)} \\ \hat{p}_{12}^{(jk)} & \cdots & \cdots & \hat{p}_{m2}^{(jk)} \\ \vdots & \vdots & \vdots & \vdots \\ \hat{p}_{1m}^{(jk)} & \cdots & \cdots & \hat{p}_{mm}^{(jk)} \end{pmatrix}$$

where

$$\hat{p}_{rs}^{(jk)} = \begin{cases} \dfrac{f_{rs}^{(jk)}}{\sum_{s=1}^{m} f_{rs}^{(jk)}} & \text{if } \sum_{s=1}^{m} f_{rs}^{(jk)} \neq 0 \\ \\ 0 & , \text{otherwise.} \end{cases}$$

For each $j = 1, 2, \ldots, n$, the empirical estimate $\hat{P}^{(jj)}$ of the transition matrix $P^{(jj)}$ for the j^{th} categorical time series Y^j is then given by:

$$\hat{P}^{(jj)} = \begin{pmatrix} \hat{p}_{11}^{(jj)} & \cdots & \cdots & \hat{p}_{m1}^{(jj)} \\ \hat{p}_{12}^{(jj)} & \cdots & \cdots & \hat{p}_{m2}^{(jj)} \\ \vdots & \vdots & \vdots & \vdots \\ \hat{p}_{1m}^{(jj)} & \cdots & \cdots & \hat{p}_{mm}^{(jj)} \end{pmatrix}$$

where

$$\hat{p}_{rs}^{(jj)} = \begin{cases} \dfrac{f_{rs}^{(jj)}}{\sum_{s=1}^{m} f_{rs}^{(jj)}} & \text{if } \sum_{s=1}^{m} f_{rs}^{(jj)} \neq 0 \\ \\ 0 & \text{, otherwise.} \end{cases}$$

Let $Q^{(jk)}$ denote the prior transition matrix for the estimation of $P^{(jk)}$. There are, in general, different approaches to determining the prior transition matrix. Market practitioners may determine the prior transition matrix based on prior knowledge of ratings of other firms or within the same industry. This resembles the determination of claim frequencies based on prior experience, in the experience rating of classical actuarial credibility. They may also base the prior transition matrix on their subjective beliefs about the different credit risks in the portfolio. The prior transition matrix can also be specified by the reference transition matrix produced by some well-renowned international credit rating agency such as Standard & Poor's, Moody's, Fitch and TRIS, etc. It may also be specified by the model-based estimate from the probit model in logistic regression analysis. Here, we specify the prior transition matrix by the transition matrix created by Standard & Poor's. It is well-known that the transition matrix produced by Standard & Poor's has been widely used as a benchmark for credit risk measurement and management in the finance and banking industries. For illustration, we assign a common prior transition matrix for two credit risky assets as the transition matrix created by Standard & Poor's shown in Table 2 to represent the belief that the transition matrices for the two credit risky assets are essentially the same based on the prior information. If more prior information about the credit rating of each asset is available, we can determine a more informative prior transition matrix for each credit risky asset. For a comprehensive overview and detailed discussion on the choice of prior distributions based on prior information, refer to some representative monographs on Bayesian Statistics, such as Lee (1997), Bernardo and Smith (2001) and Robert (2001), etc. Then, the estimate $P_e^{(jk)}$ of $P^{(jk)}$ is given by:

$$P_e^{(jk)} = w_{jk} Q^{(jk)} + (1 - w_{jk}) \hat{P}^{(jk)} \ , \ j, k = 1, 2, \ldots, n \ , \tag{3}$$

where $0 \leq w_{jk} \leq 1$, for each $j, k = 1, 2, \ldots, n$.

To estimate the Q-matrix, we also need to estimate the parameters λ_{jk}. From Proposition 1, we have seen that the multivariate Markov chain model has a stationary distribution X. In practice, the vector X can be estimated from the multivariate categorical time series of ratings by computing the proportion of the occurrence of each state in each of the categorical time series of ratings; let us denote the estimate of X by

$$\hat{X} = (\hat{X}^{(1)}, \hat{X}^{(2)}, \ldots, \hat{X}^{(n)})^T.$$

¿From Proposition 1, we have that

$$\begin{pmatrix} \lambda_{11}P_e^{(11)} & \lambda_{12}P_e^{(12)} & \cdots & \lambda_{1n}P_e^{(1n)} \\ \lambda_{21}P_e^{(21)} & \lambda_{22}P_e^{(22)} & \cdots & \lambda_{2n}P_e^{(2n)} \\ \vdots & \vdots & \vdots & \vdots \\ \lambda_{n1}P_e^{(n1)} & \lambda_{n2}P_e^{(n2)} & \cdots & \lambda_{nn}P_e^{(nn)} \end{pmatrix} \hat{\mathbf{X}} \approx \hat{\mathbf{X}} \ , \tag{4}$$

where the approximation sign means that the probability vector on the left hand side is approximated by the vector on the right hand side in the component-wise sense.

Let $\tilde{\lambda}^1_{jk} = \lambda_{jk}w_{jk}$ and $\tilde{\lambda}^2_{jk} = \lambda_{jk}(1-w_{jk})$. Then, it is easy to check that $\tilde{\lambda}^1_{jk}+\tilde{\lambda}^2_{jk} = \lambda_{jk}$, for each $j, k = 1, 2, \ldots, n$. We notice that the estimation of λ_{jk} and w_{jk} is equivalent to the estimation of $\tilde{\lambda}^1_{jk}$ and $\tilde{\lambda}^2_{jk}$. Then, (4) can be written in the following form:

$$\begin{pmatrix} \tilde{\lambda}^1_{11}Q^{(11)} + \tilde{\lambda}^2_{11}\hat{P}^{(11)} & \cdots & \tilde{\lambda}^1_{1n}Q^{(1n)} + \tilde{\lambda}^2_{1n}\hat{P}^{(1n)} \\ \tilde{\lambda}^1_{21}Q^{(21)} + \tilde{\lambda}^2_{21}\hat{P}^{(21)} & \cdots & \tilde{\lambda}^1_{2n}Q^{(2n)} + \tilde{\lambda}^2_{2n}\hat{P}^{(2n)} \\ \vdots & \vdots & \vdots \\ \tilde{\lambda}^1_{n1}Q^{(n1)} + \tilde{\lambda}^2_{n1}\hat{P}^{(n1)} & \cdots & \tilde{\lambda}^1_{nn}Q^{(nn)} + \tilde{\lambda}^2_{nn}\hat{P}^{(nn)} \end{pmatrix} \hat{\mathbf{X}} \approx \hat{\mathbf{X}} \ . \tag{5}$$

Now, we formulate our estimation problem as follows:

$$\begin{cases} \min_{\tilde{\lambda}^1, \tilde{\lambda}^2} \max_i \left| \left[\sum_{k=1}^m (\tilde{\lambda}^1_{jk} Q^{(jk)} + \tilde{\lambda}^2_{jk}\hat{P}^{(jk)}) \hat{X}^{(k)} - \hat{X}^{(j)} \right]_i \right| \\ \text{subject to} \quad \sum_{k=1}^n (\tilde{\lambda}^1_{jk} + \tilde{\lambda}^2_{jk}) = 1, \quad \tilde{\lambda}^1_{jk} \geq 0 \quad \text{and} \quad \tilde{\lambda}^2_{jk} \geq 0, \quad \forall j, k. \end{cases} \tag{6}$$

Let $O_j = \max_i \left| \left[\sum_{k=1}^m (\tilde{\lambda}^1_{jk} Q^{(jk)} + \tilde{\lambda}^2_{jk}\hat{P}^{(jk)}) \hat{X}^{(k)} - \hat{X}^{(j)} \right]_i \right|$. Then, Problem (6) can be reformulated as the following set of n Linear Programming problems (see, for instance, Ching et al. (2002)). For each j:

$$\left\{\begin{array}{ll} & \min_{\tilde{\lambda}^1,\tilde{\lambda}^2} O_j \\ \text{subject to} & \begin{pmatrix} O_j \\ O_j \\ O_j \\ O_j \\ \vdots \\ O_j \\ O_j \end{pmatrix} \geq \hat{X}^{(j)} - B_j \begin{pmatrix} \tilde{\lambda}_{j1}^1 \\ \tilde{\lambda}_{j1}^2 \\ \tilde{\lambda}_{j2}^1 \\ \tilde{\lambda}_{j2}^2 \\ \vdots \\ \tilde{\lambda}_{jn}^1 \\ \tilde{\lambda}_{jn}^2 \end{pmatrix}, \quad \begin{pmatrix} O_j \\ O_j \\ O_j \\ O_j \\ \vdots \\ O_j \\ O_j \end{pmatrix} \geq -\hat{X}^{(j)} + B_j \begin{pmatrix} \tilde{\lambda}_{j1}^1 \\ \tilde{\lambda}_{j1}^2 \\ \tilde{\lambda}_{j2}^1 \\ \tilde{\lambda}_{j2}^2 \\ \vdots \\ \tilde{\lambda}_{jn}^1 \\ \tilde{\lambda}_{jn}^2 \end{pmatrix}, \\ & O_j \geq 0, \\ & \sum_{k=1}^{n} (\tilde{\lambda}_{jk}^1 + \tilde{\lambda}_{jk}^2) = 1, \quad \tilde{\lambda}_{jk}^1 \geq 0 \quad \text{and} \quad \tilde{\lambda}_{jk}^2 \geq 0, \quad \forall j, k. \end{array}\right. \tag{7}$$

where

$$B_j = [Q^{(j1)}\hat{X}^{(1)} \mid \hat{P}^{(j1)}\hat{X}^{(1)} \mid Q^{(j2)}\hat{X}^{(2)} \mid \hat{P}^{(j2)}\hat{X}^{(2)} \mid \cdots \mid Q^{(jn)}\hat{X}^{(n)} \mid \hat{P}^{(jn)}\hat{X}^{(n)}].$$

Since there are n independent linear programming (LP) problems and each of them contains $O(m)$ constraints and $O(n)$ variables, the total computational complexity of solving such n LP problems is of $O(n^4 s)$ where s (dependent on n and m) is the number of binary bits needed to record all the data of the LP problems (see for example Fang and Puthenpura (1993)). In Section 5, we will demonstrate the effectiveness of our proposed method.

Kijima et al. (2002) estimated the unknown parameters in their model by the minimization of the squared error based on historic rating data only. Here, we adopt the idea of the Bühlmann credibility model to incorporate both prior information on the credit ratings and the historic rating data. The estimation procedure based on the Bühlmann credibility model can then be formulated as a set of LP problems. The Bühlmann credibility approach can incorporate additional prior information, such as expert opinion or rating data from other similar industries, when there is a lack of available historic rating data. In practice, one can implement our procedure by solving the LP problems on Excel spreadsheets. It is easy to implement without requiring much computational effort and time.

4 Credit Risk Measures

In recent years, there has been considerable interest in developing advanced techniques for credit risk measurement. The problem of credit risk measurement is more challenging than its market risk counterpart from both theoretical and practical perspectives. At the theoretical level, one has to apply various mathematical tools to model the possibility of defaults or rare events in a consistent way. From the practical viewpoint, it is not easy to obtain credit risk data, in comparison with its publicly available market risk counterpart. This makes the implementation and estimation of credit risk models even more difficult. In this section, we consider the problem of evaluating measures of risk, such as Value at risk (VaR) and Expected Shortfall (ES), for a portfolio of credit risks with correlated ratings.

4.1 Credit Value at Risk

In this subsection, we compute Credit VaR for a portfolio of correlated credit risks based on our multivariate Markov chain model. For illustration, we consider the one-step-ahead forecast Profit/Loss distribution for the portfolio. The method can also facilitate the evaluation of multi-step-ahead risk measures for the portfolio. Due to the fact that the predictive Profit/Loss distribution for the portfolio can be generated by the multivariate Markov chain model, we can obtain the Credit VaR for the portfolio easily.

We consider a simplified portfolio with n correlated credit risks, for instance, n corporate bonds with correlated transitions of ratings. For each $j = 1, 2, \ldots, n$ and $t \in \mathcal{T}$, we assume that the losses from the j^{th} credit risk at time t are exogenous and represented by an m-dimensional vector $L_t^j = (L_{t1}^j, L_{t2}^j, \ldots, L_{tm}^j) \in \mathcal{R}^m$, where L_{ti}^j represents the loss from the j^{th} credit risk at time t when $Y_t^{(j)} = e_i$. If L_{ti}^j is negative, the gain from the j^{th} credit risky entity at time t when $Y_t^{(j)} = e_i$ is given by $-L_{ti}^j$. The loss from the j^{th} credit risk at time t is a function of $Y_t^{(j)}$ and is denoted by $\mathcal{L}_t^{(j)}(Y_t^{(j)})$. It is easy to check that

$$\mathcal{L}_t^{(j)}(Y_t^{(j)}) = < L_t^j, Y_t^{(j)} > = \sum_{i=1}^{m} < L_t^j, e_i > I\{\omega \in \Omega | Y_t^{(j)}(\omega) = e_i\} ,$$

where $< x, y >$ denotes the inner product of two vectors $x, y \in \mathcal{R}^m$.

For each $t \in \mathcal{T}$, the aggregate loss $\mathcal{L}_t$ of the portfolio at time t is a function

$\mathcal{L}_t(Y_t)$ of Y_t and is given by:

$$\mathcal{L}_t(Y_t) := \sum_{j=1}^{n} < L_t^j, Y_t^{(j)} >= \sum_{j=1}^{n}\sum_{i=1}^{m} < L_t^j, e_i > I\{\omega \in \Omega | Y_t^{(j)}(\omega) = e_i\} \ .$$

One important statistic for credit risk measurement and management is the conditional expectation of the aggregate loss $\mathcal{L}_{t+1}$ of the portfolio at time $t+1$ given the information set $\mathcal{F}_t$. It can be evaluated as follows:

$$\begin{aligned} E_{\mathcal{P}}(\mathcal{L}_{t+1}(Y_{t+1})|\mathcal{F}_t) &= \sum_{j=1}^{n}\sum_{i=1}^{m} E_{\mathcal{P}}(< L_{t+1}^j, e_i > I\{\omega \in \Omega | Y_{t+1}^{(j)}(\omega) = e_i\}|\mathcal{F}_t) \\ &= \sum_{j=1}^{n}\sum_{i=1}^{m} < L_{t+1}^j, e_i > \mathcal{P}(\{Y_{t+1}^{(j)} = e_i\}|\mathcal{F}_t) \ . \end{aligned}$$

For each $j = 1, 2, \ldots, n$, we define the following conditional predictive distribution $P_{t+1|t}^{(j)}$ of $Y_{t+1}^{(j)}$ given $\mathcal{F}_t$:

$$P_{t+1|t}^{(j)} := (p_{t+1|t}^{(j1)}, p_{t+1|t}^{(j2)}, \ldots, p_{t+1|t}^{(jm)}) \ .$$

Following Elliott, Aggoun and Moore (1997), the conditional predictive probability $P_{t+1|t}^{(ji)}$ for each $i = 1, 2, \ldots, m$ can be evaluated as follows:

$$\begin{aligned} p_{t+1|t}^{(ji)} &:= P(\{Y_{t+1}^{(j)} = e_i\}|\mathcal{F}_t) = E_{\mathcal{P}}(< Y_{t+1}^{(j)}, e_i > |\mathcal{F}_t) \\ &= E_{\mathcal{P}}(< Y_{t+1}^{(j)}, e_i >)|_{Y_t=(e_{i_1}, e_{i_2}, \ldots, e_{i_n})} \ , \end{aligned}$$

where $f(Y_t)|_{Y_t=(e_{i_1}, e_{i_2}, \ldots, e_{i_n})}$ represents the value of the function f of the vector Y_t evaluated at $(e_{i_1}, e_{i_2}, \ldots, e_{i_n})$, for some $i_1, i_2, \ldots, i_n \in \{1, 2, \ldots, m\}$. Note that the last equality in the aforementioned expression follows from the multivariate Markov property of Y.

¿From Equation (1) in Section 2,

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

We can estimate the unknown parameters in the aforementioned equation as follows:

$$X_{t+1}^{(j)} = \sum_{k=1}^{n} \lambda_{jk} P^{(jk)} X_t^{(k)} \approx \sum_{k=1}^{n} (\tilde{\lambda}_{jk}^1 Q^{(jk)} + \tilde{\lambda}_{jk}^2 \hat{P}^{(jk)}) X_t^{(k)}, \quad \text{for} \quad j = 1, 2, \cdots, n \ .$$

Let $[V]^i$ denote the i^{th} element of the column vector V. Then, for each $j = 1, 2, \ldots, n$ and $i = 1, 2, \ldots, m$, we have

$$p_{t+1|t}^{(ji)} = E_{\mathcal{P}}(< Y_{t+1}^{(j)}, e_i >)|_{Y_t=(e_{i_1}, e_{i_2}, \ldots, e_{i_n})} = \mathcal{P}(\{Y_{t+1}^{(j)} = e_i\})|_{Y_t=(e_{i_1}, e_{i_2}, \ldots, e_{i_n})}$$

$$
\begin{aligned}
&= [X_{t+1}^{(j)}]^i|_{Y_t=(e_{i_1},e_{i_2},\ldots,e_{i_n})} = [X_{t+1}^{(j)}]^i|_{X_t=(e_{i_1},e_{i_2},\ldots,e_{i_n})} \\
&= [\sum_{k=1}^{n} \lambda_{jk} P^{(jk)} X_t^{(k)}]^i|_{X_t=(e_{i_1},e_{i_2},\ldots,e_{i_n})} \\
&\approx [\sum_{k=1}^{n} (\tilde{\lambda}_{jk}^1 Q^{(jk)} + \tilde{\lambda}_{jk}^2 \hat{P}^{(jk)}) X_t^{(k)}]^i|_{X_t=(e_{i_1},e_{i_2},\ldots,e_{i_n})} \ .
\end{aligned}
$$

This also implies that

$$
\begin{aligned}
E_{\mathcal{P}}(\mathcal{L}_{t+1}(Y_{t+1})|\mathcal{F}_t) &= \sum_{j=1}^{n}\sum_{i=1}^{m} < L_{t+1}^j, e_i > \mathcal{P}(\{Y_{t+1}^{(j)} = e_i\}|\mathcal{F}_t) \\
&= \sum_{j=1}^{n}\sum_{i=1}^{m} < L_{t+1}^j, e_i > p_{t+1|t}^{(ji)} \\
&\approx \sum_{j=1}^{n}\sum_{i=1}^{m} < L_{t+1}^j, e_i > [\sum_{k=1}^{n} (\tilde{\lambda}_{jk}^1 Q^{(jk)} + \tilde{\lambda}_{jk}^2 \hat{P}^{(jk)}) X_t^{(k)}]^i \\
&\quad |_{X_t=(e_{i_1},e_{i_2},\ldots,e_{i_n})} \ .
\end{aligned}
$$

For evaluating Credit VaR, we need to have complete knowledge of the joint conditional predictive distribution of Y_{t+1} given the information set $\mathcal{F}_t$. Based on the aforementioned result on the individual conditional predictive distribution $p_{t+1|t}^{(ji)}$, we obtain

$$
\begin{aligned}
P_{t+1|t}^{(j)} &:= (p_{t+1|t}^{(j1)}, p_{t+1|t}^{(j2)}, \ldots, p_{t+1|t}^{(jm)})^T \\
&\approx \sum_{k=1}^{n} (\tilde{\lambda}_{jk}^1 Q^{(jk)} + \tilde{\lambda}_{jk}^2 \hat{P}^{(jk)}) X_t^{(k)}|_{X_t=(e_{i_1},e_{i_2},\ldots,e_{i_n})} \ .
\end{aligned}
$$

Note that $Y_{t+1}^{(1)}, Y_{t+1}^{(2)}, \ldots, Y_{t+1}^{(n)}$ are conditionally independent given $\mathcal{F}_t$ or Y_t. Hence, the joint conditional predictive distribution $P_{t+1|t}$ of Y_{t+1} given the information $\mathcal{F}_t$ can be completely determined by:

$$
P_{t+1|t} := (P_{t+1|t}^{(1)}, P_{t+1|t}^{(2)}, \ldots, P_{t+1|t}^{(n)})^T \ , \tag{8}
$$

where $P_{t+1|t}$ is a $(n \times m)$-dimensional probability matrix.

Suppose the state space of $\mathcal{L}_{t+1|t}$ is $\{\mathcal{L}_{t+1}(1), \mathcal{L}_{t+1}(2), \ldots, \mathcal{L}_{t+1}(M)\}$, for some positive integer M, such that $\mathcal{L}_{t+1}(1) < \mathcal{L}_{t+1}(2) < \ldots < \mathcal{L}_{t+1}(M)$. For each $\tilde{k} = 1, 2, \ldots, M$, we let $I_{t+1,\tilde{k}}$ denote the set $\{(i_1, i_2, \ldots, i_n) \in \{1, 2, \ldots, m\}^n | \mathcal{L}_{t+1}(e_{i_1}, e_{i_2}, \ldots, e_{i_n}) = \mathcal{L}(\tilde{k})\}$. Then, the conditional predictive probability that the aggregate loss $\mathcal{L}_{t+1}$ equals $\mathcal{L}_{t+1}(\tilde{k})$ is given by:

$$
\mathcal{P}(\mathcal{L}_{t+1} = \mathcal{L}_{t+1}(\tilde{k})|\mathcal{F}_t) = \sum_{(i_1,i_2,\ldots,i_n)\in I_{t+1,\tilde{k}}} \prod_{j=1}^{n} p_{t+1|t}^{(ji_j)}
$$

$$\approx \sum_{(i_1,i_2,\ldots,i_n)\in I_{t+1,\tilde{k}}} \Big\{ \prod_{j=1}^{n} \Big[\sum_{k=1}^{n} (\tilde{\lambda}^1_{jk} Q^{(jk)} + \tilde{\lambda}^2_{jk} \hat{P}^{(jk)}) X_t^{(k)} \\ |_{X_t=(e_{r_1},e_{r_2},\ldots,e_{r_n})} \Big]^{i_j} \Big\} . \quad (9)$$

By definition, the VaR of the portfolio with probability level $\alpha \in (0,1)$ at time $t+1$ given the market information $\mathcal{F}_t$ is given by:

$$VaR_{\alpha,\mathcal{P}}(\mathcal{L}_{t+1}|\mathcal{F}_t) \;:=\; \inf\{L \in \mathcal{R} | \mathcal{P}(\mathcal{L}_{t+1} \geq L|\mathcal{F}_t) \leq \alpha\} \;, \quad (10)$$

where the probability level α is usually chosen to be 1% or 5% according to different purposes and practices of risk measurement and management.

Let $K^\star$ denote a positive integer in $\{1, 2, \ldots, M\}$ such that

$$\mathcal{P}(\mathcal{L}_{t+1} \geq \mathcal{L}_{t+1}(K^\star)|\mathcal{F}_t) = \sum_{\tilde{k}=K^\star}^{M} \mathcal{P}(\mathcal{L}_{t+1} = \mathcal{L}_{t+1}(\tilde{k})|\mathcal{F}_t) \leq \alpha \;, \quad (11)$$

and

$$\mathcal{P}(\mathcal{L}_{t+1} \geq \mathcal{L}_{t+1}(K^\star)+1|\mathcal{F}_t) = \sum_{\tilde{k}=K^\star+1}^{M} \mathcal{P}(\mathcal{L}_{t+1} = \mathcal{L}_{t+1}(\tilde{k})|\mathcal{F}_t) > \alpha \;. \quad (12)$$

Then, we have

$$VaR_{\alpha,\mathcal{P}}(\mathcal{L}_{t+1}|\mathcal{F}_t) = \mathcal{L}_{t+1}(K^\star) \;. \quad (13)$$

4.2 Credit Expected Shortfall

Acerbi and Tasche (2001) and Wirch and Hardy (1999) pointed out that one has to add an adjustment term to the Expected Shortfall in order to make it coherent when the loss distribution is discrete. Following these authors, the ES for the credit portfolio at time $t+1$ given the market information $\mathcal{F}_t$ up to and including time t with probability level α is defined as follows:

$$ES_\alpha(\mathcal{L}_{t+1}|\mathcal{F}_t) \;=\; \frac{1}{\alpha} E_{\mathcal{P}}(\mathcal{L}_{t+1} I_{\{\mathcal{L}_{t+1} \geq \mathcal{L}_{t+1}(K^\star)\}}|\mathcal{F}_t) + A(\alpha) \;, \quad (14)$$

where the adjustment term $A(\alpha)$ is given by:

$$A(\alpha) := \mathcal{L}_{t+1}(K^\star)\Big[1 - \frac{\mathcal{P}(\mathcal{L}_{t+1} \geq \mathcal{L}_{t+1}(K^\star)|\mathcal{F}_t)}{\alpha}\Big] \;. \quad (15)$$

To facilitate the computation, we can write the Expected Shortfall of the credit portfolio as follows:

$$
\begin{aligned}
ES_\alpha(\mathcal{L}_{t+1}|\mathcal{F}_t) \;=\;& \frac{1}{\alpha}\Big[\sum_{\tilde{k}=K^\star}^{M} \mathcal{L}_{t+1}(\tilde{k})\mathcal{P}(\mathcal{L}_{t+1} = \mathcal{L}_{t+1}(\tilde{k})|\mathcal{F}_t) - \mathcal{L}_{t+1}(K^\star) \\
& (\mathcal{P}(\mathcal{L}_{t+1} \geq \mathcal{L}_{t+1}(K^\star)|\mathcal{F}_t) - \alpha)\Big] \\
=\;& \frac{1}{\alpha}\Big\{\sum_{\tilde{k}=K^\star}^{M} \mathcal{L}_{t+1}(\tilde{k})\Big(\sum_{(i_1,i_2,\ldots,i_n)\in I_{t+1,\tilde{k}}} \prod_{j=1}^{n} p_{t+1|t}^{(ji_j)}\Big) - \mathcal{L}_{t+1}(K^\star) \\
& \Big[\sum_{\tilde{k}=K^\star}^{M}\Big(\sum_{(i_1,i_2,\ldots,i_n)\in I_{t+1,\tilde{k}}} \prod_{j=1}^{n} p_{t+1|t}^{(ji_j)}\Big) - \alpha\Big]\Big\} \\
\approx\;& \frac{1}{\alpha}\Big\{\sum_{\tilde{k}=K^\star}^{M} \mathcal{L}_{t+1}(\tilde{k})\Big\{\sum_{(i_1,i_2,\ldots,i_n)\in I_{t+1,\tilde{k}}} \Big\{\prod_{j=1}^{n} \\
& \Big[\sum_{k=1}^{n}(\tilde{\lambda}^1_{jk}Q^{(jk)} + \tilde{\lambda}^2_{jk}\hat{P}^{(jk)})X_t^{(k)}|_{X_t=(e_{r_1},e_{r_2},\ldots,e_{r_n})}\Big]^{i_j}\Big\}\Big\} - \\
& \mathcal{L}_{t+1}(K^\star)\Big\{\sum_{\tilde{k}=K^\star}^{M}\Big\{\sum_{(i_1,i_2,\ldots,i_n)\in I_{t+1,\tilde{k}}} \\
& \Big\{\prod_{j=1}^{n}\Big[\sum_{k=1}^{n}(\tilde{\lambda}^1_{jk}Q^{(jk)} + \tilde{\lambda}^2_{jk}\hat{P}^{(jk)})X_t^{(k)}|_{X_t=(e_{r_1},e_{r_2},\ldots,e_{r_n})}\Big]^{i_j}\Big\}\Big\} - \alpha\Big\}\Big\}
\end{aligned}
$$

5 A Numerical Experiment

In this section, we provide some numerical results to illustrate the use of our multivariate Markov chain model for evaluating credit VaR and ES for a portfolio of credit risks with dependent credit ratings.

5.1 An Example

In order to calculate the VaR and ES, we first have to estimate the unknown parameters in the multivariate Markov chain model. We use the data set extracted from that in Kijima et al. (2002) for the ratings of two assets, to estimate the multivariate Markov chain model. A total of eighteen years of past ratings of the two assets will be used for illustrating the numerical procedure for the estimation. In practice, it is more appropriate to use a data set with a higher sampling frequency over a shorter time interval, say two years of monthly data, instead of eighteen years of annual data, since the economic conditions may have changed dramatically over eighteen years. [1] Suppose AAA=1, AA=2, A=3, BBB=4, BB=5, B=6, CCC=7, Default=8. The historical ratings of the two assets are shown in Table 1. The following illustrates the numerical procedure of estimation.

Let S_1 and S_2 represent the ratings of the first and the second assets, respectively. By counting the transition frequencies of ratings,

$$S_1 : \begin{array}{l} 4 \rightarrow 4 \rightarrow 4 \rightarrow 4 \rightarrow 4 \rightarrow 5 \rightarrow 5 \rightarrow 5 \rightarrow 5 \rightarrow \\ 5 \rightarrow 4 \rightarrow 4 \rightarrow 4 \rightarrow 4 \rightarrow 5 \rightarrow 5 \rightarrow 5 \rightarrow 5 \end{array}$$

and

$$S_2 : \begin{array}{l} 4 \rightarrow 4 \rightarrow 4 \rightarrow 4 \rightarrow 4 \rightarrow 4 \rightarrow 4 \rightarrow 4 \rightarrow 5 \rightarrow \\ 5 \rightarrow 5 \rightarrow 4 \rightarrow 4 \rightarrow 4 \rightarrow 5 \rightarrow 5 \rightarrow 5 \rightarrow 5 \end{array}$$

we obtain the following results:

$$F^{(11)} = \begin{pmatrix} 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 7 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 2 & 7 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix} \quad \text{and} \quad F^{(22)} = \begin{pmatrix} 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 9 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 2 & 5 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix}.$$

By counting the inter-transition frequencies,

$$\begin{array}{lcccccccccccccccccccccc}
S_1 : 4 & & 4 & & 4 & & 4 & & 4 & & 5 & & 5 & & 5 & & 5 & & 5 & & 4 & & 4 \\
 & \nearrow & & \nearrow & & \nearrow & & \nearrow & & \nearrow & & \nearrow & & \nearrow & & \nearrow & & \nearrow & & \nearrow & & \nearrow & \\
S_2 : 4 & & 4 & & 4 & & 4 & & 4 & & 4 & & 4 & & 4 & & 5 & & 5 & & 5 & & 4 \\
S_1 : 4 & & 4 & & 5 & & 5 & & 5 & & 5 & & & & & & & & & & & & \\
 & \nearrow & & \nearrow & & \nearrow & & \nearrow & & \nearrow & & & & & & & & & & & & & \\
S_2 : 4 & & 4 & & 5 & & 5 & & 5 & & 5 & & & & & & & & & & & &
\end{array}$$

and

$$\begin{array}{lcccccccccccccccccccccc}
S_1 : 4 & & 4 & & 4 & & 4 & & 4 & & 5 & & 5 & & 5 & & 5 & & 5 & & 4 & & 4 \\
 & \searrow & & \searrow & & \searrow & & \searrow & & \searrow & & \searrow & & \searrow & & \searrow & & \searrow & & \searrow & & \searrow & \\
S_2 : 4 & & 4 & & 4 & & 4 & & 4 & & 4 & & 4 & & 4 & & 5 & & 5 & & 5 & & 4 \\
S_1 : 4 & & 4 & & 5 & & 5 & & 5 & & 5 & & & & & & & & & & & & \\
 & \searrow & & \searrow & & \searrow & & \searrow & & \searrow & & & & & & & & & & & & & \\
S_2 : 4 & & 4 & & 5 & & 5 & & 5 & & 5 & & & & & & & & & & & &
\end{array}$$

[1] The authors are grateful to the referee for pointing out this important issue.

we have:

$$F^{(12)} = \begin{pmatrix} 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 6 & 2 & 0 & 0 & 0 \\ 0 & 0 & 0 & 5 & 4 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix}, \quad F^{(21)} = \begin{pmatrix} 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 8 & 2 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 6 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix}.$$

Now, we normalize the columns. In practical situations, it is possible to get transition matrices with many columns of zeros. If there is a column of zeros, we follow the approach adopted in Ching et al. (2002) and replace it by a uniform distribution column vector. We then create the following transition matrix by filling in the columns with a discrete uniform distribution in the absence of information on the historical transition frequencies between other ratings for the two assets:

$$\hat{P}^{(11)} = \begin{pmatrix} \frac{1}{8} & \frac{1}{8} & \frac{1}{8} & 0 & 0 & \frac{1}{8} & \frac{1}{8} & \frac{1}{8} \\ \frac{1}{8} & \frac{1}{8} & \frac{1}{8} & 0 & 0 & \frac{1}{8} & \frac{1}{8} & \frac{1}{8} \\ \frac{1}{8} & \frac{1}{8} & \frac{1}{8} & 0 & 0 & \frac{1}{8} & \frac{1}{8} & \frac{1}{8} \\ \frac{1}{8} & \frac{1}{8} & \frac{1}{8} & \frac{7}{9} & \frac{1}{8} & \frac{1}{8} & \frac{1}{8} & \frac{1}{8} \\ \frac{1}{8} & \frac{1}{8} & \frac{1}{8} & \frac{2}{9} & \frac{7}{8} & \frac{1}{8} & \frac{1}{8} & \frac{1}{8} \\ \frac{1}{8} & \frac{1}{8} & \frac{1}{8} & 0 & 0 & \frac{1}{8} & \frac{1}{8} & \frac{1}{8} \\ \frac{1}{8} & \frac{1}{8} & \frac{1}{8} & 0 & 0 & \frac{1}{8} & \frac{1}{8} & \frac{1}{8} \\ \frac{1}{8} & \frac{1}{8} & \frac{1}{8} & 0 & 0 & \frac{1}{8} & \frac{1}{8} & \frac{1}{8} \end{pmatrix}, \quad \hat{P}^{(12)} = \begin{pmatrix} \frac{1}{8} & \frac{1}{8} & \frac{1}{8} & 0 & 0 & \frac{1}{8} & \frac{1}{8} & \frac{1}{8} \\ \frac{1}{8} & \frac{1}{8} & \frac{1}{8} & 0 & 0 & \frac{1}{8} & \frac{1}{8} & \frac{1}{8} \\ \frac{1}{8} & \frac{1}{8} & \frac{1}{8} & 0 & 0 & \frac{1}{8} & \frac{1}{8} & \frac{1}{8} \\ \frac{1}{8} & \frac{1}{8} & \frac{1}{8} & \frac{6}{11} & \frac{1}{3} & \frac{1}{8} & \frac{1}{8} & \frac{1}{8} \\ \frac{1}{8} & \frac{1}{8} & \frac{1}{8} & \frac{5}{11} & \frac{2}{3} & \frac{1}{8} & \frac{1}{8} & \frac{1}{8} \\ \frac{1}{8} & \frac{1}{8} & \frac{1}{8} & 0 & 0 & \frac{1}{8} & \frac{1}{8} & \frac{1}{8} \\ \frac{1}{8} & \frac{1}{8} & \frac{1}{8} & 0 & 0 & \frac{1}{8} & \frac{1}{8} & \frac{1}{8} \\ \frac{1}{8} & \frac{1}{8} & \frac{1}{8} & 0 & 0 & \frac{1}{8} & \frac{1}{8} & \frac{1}{8} \end{pmatrix},$$

$$\hat{P}^{(21)} = \begin{pmatrix} \frac{1}{8} & \frac{1}{8} & \frac{1}{8} & 0 & 0 & \frac{1}{8} & \frac{1}{8} & \frac{1}{8} \\ \frac{1}{8} & \frac{1}{8} & \frac{1}{8} & 0 & 0 & \frac{1}{8} & \frac{1}{8} & \frac{1}{8} \\ \frac{1}{8} & \frac{1}{8} & \frac{1}{8} & 0 & 0 & \frac{1}{8} & \frac{1}{8} & \frac{1}{8} \\ \frac{1}{8} & \frac{1}{8} & \frac{1}{8} & \frac{8}{9} & \frac{1}{4} & \frac{1}{8} & \frac{1}{8} & \frac{1}{8} \\ \frac{1}{8} & \frac{1}{8} & \frac{1}{8} & \frac{1}{9} & \frac{3}{4} & \frac{1}{8} & \frac{1}{8} & \frac{1}{8} \\ \frac{1}{8} & \frac{1}{8} & \frac{1}{8} & 0 & 0 & \frac{1}{8} & \frac{1}{8} & \frac{1}{8} \\ \frac{1}{8} & \frac{1}{8} & \frac{1}{8} & 0 & 0 & \frac{1}{8} & \frac{1}{8} & \frac{1}{8} \\ \frac{1}{8} & \frac{1}{8} & \frac{1}{8} & 0 & 0 & \frac{1}{8} & \frac{1}{8} & \frac{1}{8} \end{pmatrix}, \quad \hat{P}^{(22)} = \begin{pmatrix} \frac{1}{8} & \frac{1}{8} & \frac{1}{8} & 0 & 0 & \frac{1}{8} & \frac{1}{8} & \frac{1}{8} \\ \frac{1}{8} & \frac{1}{8} & \frac{1}{8} & 0 & 0 & \frac{1}{8} & \frac{1}{8} & \frac{1}{8} \\ \frac{1}{8} & \frac{1}{8} & \frac{1}{8} & 0 & 0 & \frac{1}{8} & \frac{1}{8} & \frac{1}{8} \\ \frac{1}{8} & \frac{1}{8} & \frac{1}{8} & \frac{9}{11} & \frac{1}{6} & \frac{1}{8} & \frac{1}{8} & \frac{1}{8} \\ \frac{1}{8} & \frac{1}{8} & \frac{1}{8} & \frac{2}{11} & \frac{5}{6} & \frac{1}{8} & \frac{1}{8} & \frac{1}{8} \\ \frac{1}{8} & \frac{1}{8} & \frac{1}{8} & 0 & 0 & \frac{1}{8} & \frac{1}{8} & \frac{1}{8} \\ \frac{1}{8} & \frac{1}{8} & \frac{1}{8} & 0 & 0 & \frac{1}{8} & \frac{1}{8} & \frac{1}{8} \\ \frac{1}{8} & \frac{1}{8} & \frac{1}{8} & 0 & 0 & \frac{1}{8} & \frac{1}{8} & \frac{1}{8} \end{pmatrix}.$$

Based on the historical ratings of the two assets, we obtain the following estimates for the stationary probability distributions of the ratings of the first and the second credit risks:

$$\hat{X}_1 = (0, 0, 0, \frac{1}{2}, \frac{1}{2}, 0, 0, 0)^T \quad \text{and} \quad \hat{X}_2 = (0, 0, 0, \frac{11}{18}, \frac{7}{18}, 0, 0, 0)^T$$

Finally, by solving a set of Linear Programming (LP) problems, the estimated multivariate Markov chain model for S_1 and S_2 is given by:

$$\begin{cases} X_{n+1}^{(1)} & = & 0.614Q^{(11)}X_n^{(1)} + 0.000\hat{P}^{(11)}X_n^{(1)} + 0.000Q^{(12)}X_n^{(2)} + 0.384\hat{P}^{(12)}X_n^{(2)} \\ X_{n+1}^{(2)} & = & 0.000Q^{(21)}X_n^{(1)} + 1\hat{P}^{(21)}X_n^{(1)} + 0.000Q^{(22)}X_n^{(2)} + 0.000\hat{P}^{(22)}X_n^{(2)}. \end{cases}$$

It is natural to use the average probability to replace the missing information in the estimated transition matrix P since the data only appear in two different states, 4 and 5, in P. Although most of the entries in P are equal to zero, the missing information can be given by the prior transition matrix once the weighting of the corresponding prior transition matrix, $Q^{(kk)}$, is non-zero. The aforementioned estimated model can be used to generate the predictive probability distributions for evaluating credit risk measures.

The traditional reduced-form approach assumes that the losses from a credit risk at each particular rating class are exogenous. It is supposed that the losses from credit risks in different rating classes can be evaluated based on some accounting information and principles. Following the traditional reduced-form approach, we also assume that the losses from credit risks at different rating classes are given in advance. For illustration, we consider a portfolio of the two assets with correlated ratings. We suppose that the aggregate loss from the portfolio is given by:

$$\mathcal{L}_{t+1}(Y_{t+1}^{(1)}, Y_{t+1}^{(2)}) = \mathcal{L}_{t+1}^{(1)}(Y_{t+1}^{(1)}) + \mathcal{L}_{t+1}^{(2)}(Y_{t+1}^{(2)}) \ . \tag{16}$$

For each $j = 1, 2$, the rating of the j^{th} asset $Y_{t+1}^{(j)}$ at time $t+1$ can take values in the set of unit basis vectors $\{e_1, e_2, \ldots, e_8\} \in \mathcal{R}^8$. Now, we consider a unit interval $[0, 1]$ and its uniform partition $\cup_{i=1}^{8} P_i$ with

$$P_i = [\frac{i-1}{8}, \frac{i}{8}) \ , \tag{17}$$

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

We further assume that for each $j = 1, 2$ and $i = 1, 2, \ldots, 8$, the loss from the j^{th} asset $\mathcal{L}_{t+1}^{(j)}(Y_{t+1}^{(j)})$ given that $Y_{t+1}^{(j)} = e_i$ can take values in the interval P_i. More precisely, the conditional distribution of the loss $\mathcal{L}_{t+1}^{(j)}(Y_{t+1}^{(j)})$ given that $Y_{t+1}^{(j)} = e_i$ is a uniform distribution on the interval P_i, for each $i = 1, 2, \ldots, 8$. Suppose the j^{th} asset is in the highest rating class at time $t+1$, say the class AAA. This implies that $Y_{t+1}^{(j)} = e_1$. In this case, the loss from the j^{th} asset $\mathcal{L}_{t+1}^{(j)}(Y_{t+1}^{(j)})$ at time $t+1$ can take values in $P_1 = [0, \frac{1}{8})$. For each $j = 1, 2$ and $i = 1, 2, \ldots, 8$, we assign a value to the loss $\mathcal{L}_{t+1}^{(j)}(Y_{t+1}^{(j)})$ from the j^{th} asset at time $t+1$ given that $Y_{t+1}^{(j)} = e_i$ by a pseudo-random number drawn from a uniform distribution on the interval P_i. The simulated results are shown in Table 3. We summarize the future possible values of the aggregate loss from the portfolio, their corresponding ratings of the underlying assets and their corresponding conditional predictive probabilities given the current information in Table 4. The set $I_{t+1,\tilde{k}}$ is also shown in Table 4. The conditional predictive probability of the aggregate loss $\mathcal{L}$ given the current market information for each combination of ratings of the assets in the portfolio is computed by Equation (9).

Based on the numerical results in Table 4, the ordered aggregate losses from the credit portfolio at time $t+1$ based on the simulated results in Table 3 are given by { 0.1477 0.2297 0.3045 0.3866 0.3903 0.4610 0.4640 0.4961 0.5461 0.5472 0.5782 0.6178 0.6316 0.6744 0.7067 0.7308 0.7387 0.7564 0.7773 0.7884 0.8094 0.8428 0.8877 0.8941 0.9170 0.9248 0.9259 0.9479 0.9610 0.9800 0.9876 1.0431 1.0445 1.0472 1.0510 1.0792 1.0828 1.0854 1.1266 1.1561 1.1583 1.2037 1.2105 1.2423 1.2425 1.2575 1.2743 1.2743 1.2871 1.3267 1.3578 1.4208 1.4259 1.4450 1.4526 1.5284 1.5442 1.5892 1.6210 1.6276 1.7075 1.7393 1.7909 1.8227 } .

For evaluating Credit VaR for the portfolio, we choose the value $K^\star$ such that it satisfies Equation (11) and Equation (12) for a given probability level α. In practice, the probability level α is usually chosen to be either 0.05 or 0.01. Hence, we adopt these two commonly used values of α for illustration.

First, we set $\alpha = 0.05$. From Table 4, we obtain:

$$\sum_{\tilde{k}=41+1}^{64} \mathcal{P}(\mathcal{L}_{t+1} = \mathcal{L}_{t+1}(\tilde{k})|\mathcal{F}_t) \leq 0.05$$

and

$$\sum_{\tilde{k}=41}^{64} \mathcal{P}(\mathcal{L}_{t+1} = \mathcal{L}_{t+1}(\tilde{k})|\mathcal{F}_t) > 0.05 \ .$$

This implies that

$$K^\star = 41 \ .$$

Hence, the numerical value of the VaR is given by:

$$VaR_{\alpha,\mathcal{P}}(\mathcal{L}_{t+1}|\mathcal{F}_t) = \mathcal{L}_{t+1}(K^\star) = 1.1583 \ .$$

¿From the numerical value of the VaR and the numerical values in Table 4, we can evaluate the Expected shortfall of the portfolio using Equation (15). Note that

$$\begin{aligned}
& \frac{1}{\alpha}\Big\{\sum_{\tilde{k}=K^\star}^{M}\mathcal{L}_{t+1}(\tilde{k})\Big\{\sum_{(i_1,\ldots,i_n)\in I_{t+1,\tilde{k}}}\Big\{\prod_{j=1}^{n}\Big[\sum_{k=1}^{n}(\tilde{\lambda}^1_{jk}Q^{(jk)}+\tilde{\lambda}^2_{jk}\hat{P}^{(jk)})X_t^{(k)} \\
& |_{X_t=(e_{r_1},\ldots,e_{r_n})}\Big]^{i_j}\Big\}\Big\} \\
= \; & 14.44287 \ ,
\end{aligned}$$

and

$$\begin{aligned}
& \frac{1}{\alpha}\Big\{\mathcal{L}_{t+1}(K^\star)\Big\{\sum_{\tilde{k}=K^\star}^{M}\Big\{\sum_{(i_1,\ldots,i_n)\in I_{t+1,\tilde{k}}}\Big\{\prod_{j=1}^{n}\Big[\sum_{k=1}^{n}(\tilde{\lambda}^1_{jk}Q^{(jk)}+\tilde{\lambda}^2_{jk}\hat{P}^{(jk)})X_t^{(k)} \\
& |_{X_t=(e_{r_1},\ldots,e_{r_n})}\Big]^{i_j}\Big\}\Big\}-\alpha\Big\}\Big\} \\
= \; & 13.15134 \ .
\end{aligned}$$

Therefore, we obtain:

$$ES_\alpha(\mathcal{L}_{t+1}|\mathcal{F}_t) = 1.291532 \ .$$

Now, we set $\alpha = 0.01$ and follow the same procedure as before. ¿From Table 4, we obtain:

$$\sum_{\tilde{k}=48+1}^{64}\mathcal{P}(\mathcal{L}_{t+1}=\mathcal{L}_{t+1}(\tilde{k})|\mathcal{F}_t) \leq 0.01$$

and

$$\sum_{\tilde{k}=48}^{64}\mathcal{P}(\mathcal{L}_{t+1}=\mathcal{L}_{t+1}(\tilde{k})|\mathcal{F}_t) > 0.01 \ .$$

This implies that

$$K^\star = 48 \ .$$

Hence, we have

$$VaR_{\alpha,\mathcal{P}}(\mathcal{L}_{t+1}|\mathcal{F}_t) = \mathcal{L}_{t+1}(K^\star) = 1.2743 \ .$$

In this case, we have:

$$\begin{aligned}
& \frac{1}{\alpha}\Big\{\sum_{\tilde{k}=K^\star}^{M}\mathcal{L}_{t+1}(k)\Big\{\sum_{(i_1,\ldots,i_n)\in I_{t+1,\tilde{k}}}\Big\{\prod_{j=1}^{n}\Big[\sum_{k=1}^{n}(\tilde{\lambda}^1_{jk}Q^{(jk)}+\tilde{\lambda}^2_{jk}\hat{P}^{(jk)})X_t^{(k)} \\
& |_{X_t=(e_{r_1},\ldots,e_{r_n})}\Big]^{i_j}\Big\}\Big\}\Big\} \\
= \; & 1.611218 \ ,
\end{aligned}$$

and

$$
\begin{aligned}
& \frac{1}{\alpha}\Big\{\mathcal{L}_{t+1}(K^\star)\Big\{\sum_{\tilde{k}=K^\star}^{M}\Big\{\sum_{(i_1,\ldots,i_n)\in I_{t+1,\tilde{k}}}\Big\{\prod_{j=1}^{n}\Big[\sum_{k=1}^{n}(\tilde{\lambda}_{jk}^{1}Q^{(jk)}+\tilde{\lambda}_{jk}^{2}\hat{P}^{(jk)})X_t^{(k)} \\
& |_{X_t=(e_{r_1},\ldots,e_{r_n})}\Big]^{i_j}\Big\}\Big\}-\alpha\Big\}\Big\} \\
= \ & 0.178402 \ .
\end{aligned}
$$

Therefore, the Expected Shortfall of the portfolio is given by:

$$ES_\alpha(\mathcal{L}_{t+1}|\mathcal{F}_t) = 1.432816 \ .$$

The numerical results for VaR and ES are consistent with intuition. The VaR (ES) increases from 1.1583 (1.291532) to 1.2743 (1.432816) as the probability level α decreases from 5% to 1%. For each fixed probability level α, the ES is always greater than the VaR.

5.2 Efficiency of Parameter Estimation

Finally, we report the efficiency of our proposed method for parameter estimation. Figure 1 shows the total computational times for model parameters estimation. The computational times include the construction of the transition matrices and the linear programming solver for (6). All the calculations were made with CPU=AMD1800+ with RAM=512. Here, we set the number of states to be eight, which is the same as in our previous numerical example. Credit rating sequences are simulated and the computational times required for the construction of our proposed multivariate Markov models are recorded. In Figure 1, we show the computational times for different numbers of credit rating sequences. We see from the figure that the computation times increase quadratically with respect to the number n of sequences. As we mentioned in Section 3, there are n linear programming problems to be solved, so these numerical results show that our proposed estimation method is quite efficient. Indeed, the construction time of our multivariate Markov models in the previous example only requires a few seconds.

The implementation of the estimation procedure in the model by Kijima, Komoribayashi and Suzuki (2002) involves the computation of a two-stage minimization iteratively (see Algorithm 2 therein). Our estimation procedure involves solving a set of LP problems, which can be implemented easily on Microsoft Excel spreadsheets. Most of the calculations in our estimation procedures only involve elementary mathematical operations.

6 Conclusion and Further Research

We have considered one of the most important topics in risk management, namely the modelling of dependent risks. In particular, we applied the multivariate Markov chain model to incorporate the dependency of the ratings of credit risks in a portfolio. Our model allows the next period's rating of a particular credit risk to depend not only on its current rating but also on the current ratings of other credit risks in the portfolio. We proposed an estimate for the transition matrix based on a linear combination of the empirical transition matrix and a prior estimate of the transition matrix. The estimation method can incorporate both the historic rating data and another source of information about rating data, for instance, expert opinion or subjective views, which is also important for credit risk measurement and management. The estimates of the unknown parameters and the transition matrices can be obtained by solving a set of LP problems. Our estimation method is analytically tractable, easy to implement and computationally efficient. It can be implemented on Excel spreadsheets easily. However, its applicability is limited by the number of parameters involved, which depends on the dimension of the multivariate categorical time series and the number of possible credit ratings.

One may also explore applications of the multivariate Markov chain model to portfolio management and financial econometrics. For portfolio management, it is important to model the correlation of risky securities. It has been documented in the literature that Markov chain models are feasible approximations to continuous-state time series models. Hence, it would be interesting to explore the use of the multivariate Markov chain model as a model of the dependency of multivariate sources of risk in portfolio management, and to investigate the efficiency of the approximation. It would also be interesting to explore the use of the multivariate Markov chain model to approximate the dependency between multivariate financial time series, which is a very important topic in financial econometrics (see Diebold (2003) and Patton (2004)). The advantage of using the multivariate Markov chain model in these two areas is that the model is easy to implement and computationally efficient. However, the applicability of the multivariate Markov chain model in these two areas is also limited by the number of parameters involved.

There are also some possible directions for further research from the technical perspective. First, it is interesting to explore the possibility of using the risk-neutral transition matrix as a specification of the prior transition matrix. In this case, one can combine the information obtained from both the risk-neutral probability and the real-world probability for the estimation of the unknown parameters and the transition matrix. It may also be interesting to investigate whether it is possible to explain the choice of the risk-neutral transition matrix as the prior probability matrix under a certain combination of preference and probability structures. Finally, the investigation of the use of probit models in logistic regression analysis for the specification of the prior transition matrix is also interesting.

Acknowledgments

The authors would like to thank the referees for many valuable and helpful comments and suggestions. Their valuable comments and suggestions improved the paper. This work was supported in part by RGC Grant No. HKU 7126/02P, 7130/02P, 7046/03P, 7035/04P and 7035/05P, and HKU CRCG Grant Nos. 10204436, 10205105 and 10205775.

References

Alexander, C., 2005. The Present and Future of Risk Management. Journal of Financial Econometrics 3 (1), 3-25.

Artzner, P., Delbaen, F., 1995. Default Risk Premium and Incomplete Markets. Mathematical Finance 5, 187-195.

Acerbi, C., Tasche, D., 2001. On the Coherence of Expected Shortfall. Journal of Banking and Finance 26 (7), 1487-1503.

Bernardo, J. M., Smith, A. M. F., 2001. Bayesian Theory. John Wiley & Sons, New York.

Bühlmann, H., 1967. Experience Rating and Credibility Theory. ASTIN Bulletin 4, 199-207.

Cherubini, U., Luciano, E. 2002. Bivariate Option Pricing with Copulas. Applied Mathematical Finance 8, 69-85.

Cherubini, U., Luciano, E., Vecchiato, W. 2004. Copulas Methods in Finance, John Wiley & Sons , New York.

Ching, W., Fung, E., Ng, M., 2002. A Multivariate Markov Chain Model for Categorical Data Sequences and Its Applications in Demand Predictions. IMA Journal of Management Mathematics 13, 187-199.

Ching, W., Fung, E., Ng, M., 2004. Higher-order Multivariate Markov Chains: Models, Algorithms and Applications, revised and re-submitted to INFORMS Journal of Computing.

Chvatal, V., 1983. Linear Programming, Freeman, New York.

Das, S., Freed, L., Geng, G. and Kapadia, N. 2004. Correlated Default Risk. Working Paper.

Diebold, F.X. 2003. The ET Interview: Professor Robert F. Engle. Econometric Theory, 19, 1159-1193.

Duffie, D., Schroder, M., Skiadas, C., 1996. Recursive Valuation of Defaultable Securities and the Timing of the Resolution of Uncertainty. Annal of Applied Probability 6, 1075-1090.

Elliott, R. J., Aggoun, L., Moore, J. B., 1997. Hidden Markov models: estimation and control, Springer-Verlag, New York.

Embrechts, P., Mcneil, A., Straumann, D., 1999. Correlation and Dependence in Risk Management: Properties and Pitfalls. Risk May, 69-71.

Fang, S and Puthenpura, S.,1993. Linear Optimization and Extensions. Prentice-Hall, Englewood Cliffs, NJ.

Gupton, G. M., Finger, C. C., Bhatia, M., 1997. $CreditMetrics^{TM}$, J.P. Morgan, New York.

Hennessy, D. A., Lapan, H. E. 2002. The Use of Archimedean Copulas To Model Portfolio Allocations. Mathematical Finance 12 (2), 143-154.

Hu, Y. T., Kiesel, R., Perraudin, W., 2002. The Estimation of Transition Matrices for Sovereign Ratings. Journal of Banking and Finance 26 (7), 1383-1406.

Jarrow, R., Turnbull, S., 1995. Pricing Options on Financial Derivatives Subject to Default Risk. Journal of Finance 50, 53-86.

Jarrow, R., Lando, D., Turnbull, S. (1997) A Markov Model for the Term Structure of Credit Spreads. Review of Financial Studies 10, 481-523.

Kijima, M., Komoribayashi, K., Suzuki, E., 2002. A Multivariate Markov Model for Simulating Correlated Defaults. Journal of Risk 4(4), 1-32.

Klugman, S., Panjer, H. H., Willmot, G. E., 1997. Loss Models: From Data to Decisions. John Wiley & Sons, New York.

Lando, D. 1998. On Cox Processes and Credit Risky Securities. The Review of Derivatives Research 2, 99-120.

Lee, P. M., 1997. Bayesian Statistics: An Introduction. Edward Arnold, London.

Li, D., 2000. On Default Correlation: A Copula Function Approach. Journal of Fixed Income 9 (4), 43-54.

Madan, D., Unal, H., 1995. Pricing the Risk of Default. Review of Derivatives Research 2, 121-160.

Merton, R., 1974. On the Pricing of Corporate Debt: The Risk Structure of Interest Rates. The Journal of Finance 29, 449-470.

Mowbray, A. N., 1914. How Extensive a Payroll Exposure is Necessary to give a Dependent Pure Premium. Proceedings of the Causality Actuarial Society 1, 24-30.

Nahmias, S., Schmidt, C., 1984. An Efficient Heuristic for Multi-item Newsboy Problem with a Single Constraint. Naval Research Logistics Quarterly 21, 463-474.

Patton, A. 2004. Modelling Asymmetric Exchange Rate Dependence. Working Paper. London School of Economics, United Kingdom.

Pliska, S. R. (2003) “Introduction to Mathematical Finance: Discrete Time Models”, Blackwell Publishers, Oxford.

Robert, C. P., 2001. The Bayesian Choice, Springer-Verlag, New York.

Ross, S., 2000. Introduction to Probability Models. Academic Press, San Diego.
Siu, T. K., Yang, H. L., 1999. Subjective Risk Measures: Bayesian Predictive Scenarios Analysis. Insurance: Mathematics and Economics 25, 157-169.
Siu, T. K., Tong, H., Yang, H. L., 2001. Bayesian Risk Measures for Derivatives via Random Esscher Transform. North American Actuarial Journal, 5(3), 78-91.
Thomas, L.C., Allen, D.E. and Morkel-Kingsbury, N., 2002. A Hidden Markov Chain Model for the Term Structure of Credit Risk Spreads. International Review of Financial Analysis, 11, 311-329.
Wirch, J. L., Hardy, M. R., 1999. A Synthesis of Risk Measures for Capital Adequacy. Insurance: Mathematics and Economics 25, 337-347.
Woo, W. H., Siu, T. K., 2004. A Dynamic Binomial Expansion Technique for Credit Risk Measurement: A Bayesian Filtering Approach. Applied Mathematical Finance, 11(2), 165-186.

7 Tables and Figures

Asset/Year	1	2	4	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

Table 1. The Ratings of the Two Assets.

Asset/Year	AAA	AA	A	BBB	BB	B	CCC	D
AAA	0.9193	0.0746	0.0048	0.0008	0.0004	0.0000	0.0000	0.0000
AA	0.6400	0.9181	0.0676	0.0060	0.0006	0.0012	0.0003	0.0000
A	0.0700	0.0227	0.9169	0.0512	0.0056	0.0025	0.0001	0.0004
BBB	0.0400	0.0270	0.0556	0.8788	0.0483	0.0102	0.0017	0.0024
BB	0.0400	0.0010	0.0061	0.0775	0.8148	0.0790	0.0111	0.0101
B	0.0000	0.0010	0.0028	0.0046	0.0695	0.8280	0.0396	0.0545
CCC	0.1900	0.0000	0.0037	0.0075	0.0243	0.1213	0.6045	0.2369
D	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	1.0000

Table 2. The Transition Probability Table from Standard & Poor's (1999)

j/i	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

Table 3. The Simulated Losses (i.e. $\mathcal{L}_{t+1}^{(j)}(e_i)$)

$\tilde{k}$	The Value of Aggregate Loss at State $\mathcal{L}(\tilde{k})$	$\mathcal{P}(\mathcal{L}_{t+1} = \mathcal{L}_{t+1}(\tilde{k})\|\mathcal{F}_t)$	$I_{t+1,\tilde{k}}$
1	0.1477	0	{ 1 , 1 }
2	0.2297	0	{ 2 , 1 }
3	0.3045	0	{ 1 , 2 }
4	0.3866	0	{ 2 , 2 }
5	0.3903	0	{ 3 , 1 }
6	0.461	0	{ 4 , 1 }
7	0.464	0	{ 1 , 3 }
8	0.4961	0.0001	{ 1 , 4 }
9	0.5461	0	{ 2 , 3 }
10	0.5472	0	{ 3 , 2 }
11	0.5782	0.0002	{ 2 , 4 }
12	0.6178	0	{ 4 , 2 }
13	0.6316	0	{ 5 , 1 }
14	0.6744	0.0002	{ 1 , 5 }
15	0.7067	0	{ 3 , 3 }
16	0.7308	0	{ 6 , 1 }
17	0.7387	0.0009	{ 3 , 4 }
18	0.7564	0.0005	{ 2 , 5 }
19	0.7773	0	{ 4 , 3 }
20	0.7884	0	{ 5 , 2 }
21	0.8094	0.0441	{ 4 , 4 }
22	0.8428	0	{ 1 , 6 }
23	0.8877	0	{ 6 , 2 }
24	0.8941	0	{ 7 , 1 }
25	0.917	0.0028	{ 3 , 5 }
26	0.9248	0	{ 2 , 6 }
27	0.9259	0	{ 8 , 1 }
28	0.9479	0	{ 5 , 3 }
29	0.961	0	{ 1 , 7 }
30	0.98	0.1894	{ 5 , 4 }
31	0.9876	0.1322	{ 4 , 5 }
32	1.0431	0	{ 2 , 7 }

$\tilde{k}$	The Value of Aggregate Loss at State $\mathcal{L}(\tilde{k})$	$\mathcal{P}(\mathcal{L}_{t+1} = \mathcal{L}_{t+1}(\tilde{k})\vert\mathcal{F}_t)$	$I_{t+1,\tilde{k}}$
33	1.0445	0	{ 1 , 8 }
34	1.0472	0	{ 6 , 3 }
35	1.051	0	{ 7 , 2 }
36	1.0792	0.0121	{ 6 , 4 }
37	1.0828	0	{ 8 , 2 }
38	1.0854	0	{ 3 , 6 }
39	1.1266	0	{ 2 , 8 }
40	1.1561	0	{ 4 , 6 }
41	1.1583	0.5682	{ 5 , 5 }
42	1.2037	0	{ 3 , 7 }
43	1.2105	0	{ 7 , 3 }
44	1.2423	0	{ 8 , 3 }
45	1.2425	0.0017	{ 7 , 4 }
46	1.2575	0.0364	{ 6 , 5 }
47	1.2743	0	{ 4 , 7 }
48	1.2743	0.0016	{ 8 , 4 }
49	1.2871	0	{ 3 , 8 }
50	1.3267	0	{ 5 , 6 }
51	1.3578	0	{ 4 , 8 }
52	1.4208	0.0051	{ 7 , 5 }
53	1.4259	0	{ 6 , 6 }
54	1.445	0	{ 5 , 7 }
55	1.4526	0.0047	{ 8 , 5 }
56	1.5284	0	{ 5 , 8 }
57	1.5442	0	{ 6 , 7 }
58	1.5892	0	{ 7 , 6 }
59	1.621	0	{ 8 , 6 }
60	1.6276	0	{ 6 , 8 }
61	1.7075	0	{ 7 , 7 }
62	1.7393	0	{ 8 , 7 }
63	1.7909	0	{ 7 , 8 }
64	1.8227	0	{ 8 , 8 }

Table 4. The Aggregate Loss

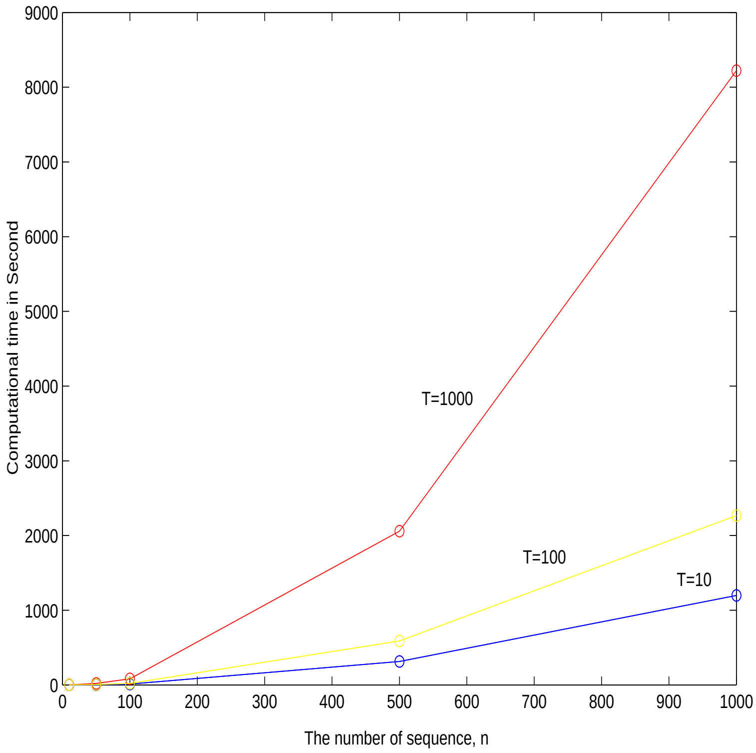

Figure 1: The total computational times for model parameters estimation.