Journal of Banking & Finance 36 (2012) 78–89

Contents lists available at ScienceDirect

Journal of Banking & Finance

journal homepage: www.elsevier.com/locate/jbf

# Credit rating dynamics in the presence of unknown structural breaks

Haipeng Xing [a,*], Ning Sun [a], Ying Chen [b,1]

[a] Department of Applied Mathematics and Statistics, State University of New York at Stony Brook, Stony Brook, NY 11794, United States
[b] MEAG New York Corporation, 540 Madison Avenue, New York, NY 10022, United States

ARTICLE INFO

*Article history:*
Received 22 May 2010
Accepted 9 June 2011
Available online 25 June 2011

*JEL classification:*
C13
C41
G12
G20

*Keywords:*
Credit risk
Hidden Markov model
Stochastic structural break

ABSTRACT

In many credit risk and pricing applications, credit transition matrix is modeled by a constant transition probability or generator matrix for Markov processes. Based on empirical evidence, we model rating transition processes as piecewise homogeneous Markov chains with unobserved structural breaks. The proposed model provides explicit formulas for the posterior distribution of the time-varying rating transition generator matrices, the probability of structural break at each period and prediction of transition matrices in the presence of possible structural breaks. Estimating the model by credit rating history, we show that the structural break in rating transitions can be captured by the proposed model. We also show that structural breaks in rating dynamics are different for different industries. We then compare the prediction performance of the proposed and time-homogeneous Markov chain models.



## 1. Introduction

In modern credit risk management, it is convenient to assume that the credit quality (or rating) of obligors (or firms) follows a time homogeneous Markov chain which is characterized by a credit rating transition matrix. Since the credit rating transition matrix summarizes historical changes in credit rating of obligors, it has wide applications in finance. In pricing models for bond and credit derivatives, the valuation of risky credit derivatives is based on the credit ratings of obligors (Jarrow et al., 1997, 1998; Lando, 2000; Acharya et al., 2006). In portfolio risk assessment, the risk is measured by the joint distribution of rating transitions for the loans and bonds which make up the portfolios of interest (Das et al., 2006; Frey and McNeil, 2007; Egloff et al., 2007; Duffie et al., 2009; Tsaig et al., 2011). In the Basel Accord for bank regulation, banks are required to construct credit rating transition matrices based on their own data to stress test their portfolios and evaluate evidence of rating transitions in external ratings (Treacy and Carey, 2000; Altman and Rijken, 2004; Gordy and Howells, 2006). In the credit rating industry, rating agencies such as Moody's, Standard and Poor, and Fitch publish reports on rating transitions for obligors or financial instruments, which are studied by credit risk managers, and some risk management tools such as Morgan's CreditMetrics are built on estimates of rating transition matrices.

The estimates of credit rating transition matrices published by rating agencies usually use a discrete-time setting. However, due to the availability of rating data and the well known advantages of using the continuous time Markov approach over the discrete one (Lando and Skødeberg, 2002), a continuous time homogeneous Markov framework is usually assumed for the rating process. In particular, suppose that there are $K$ rating categories and the rating migration process of a firm for the period $(0,t)$ is a continuous time homogeneous Markov chain with transition matrix $P(0,t)$, in which the $ij$th entry represents the probability of migrating from category $i$ to category $j$ during the period $(0,t)$. Similar to the discrete-time Markov process for which the rating transition matrix can be obtained by matrix multiplication from the one-period transition matrix, the matrix $P(0,t)$ can be represented, under the assumption of time homogeneity, through its generator matrix $\Lambda$, that is, for any time $t>0$,

$$P(0,t)=\exp(\Lambda t):=\sum_{k=0}^{\infty}\frac{\Lambda^k t^k}{k!}, \tag{1}$$

in which $\Lambda=(\lambda^{(i,j)})$ satisfies $\lambda^{(i,i)}=-\sum_{j\neq i}\lambda^{(i,j)}$ for $1\leqslant i\leqslant K$, and $\lambda^{(i,j)}\geqslant 0$ for $1\leqslant i\neq j\leqslant K$. The elements in $\Lambda$ can be estimated by their maximum likelihood estimators $\hat{\lambda}^{(i,j)}=\widetilde{N}_{ij}/\int_0^t Y_i(s)ds$, in which $Y_i(s)$ is the number of firms in rating class $i$ at time $s$ and $\widetilde{N}_{ij}$ is the total number of transitions from $i$ to $j(\neq i)$ over the period $(0,t)$ (Küchler and Sørensen, 1997, p. 26).

* Corresponding author. Tel.: +1 631 632 1892; fax: +1 631 632 8490.
*E-mail addresses:* xing@ams.sunysb.edu (H. Xing), nsun@ams.sunysb.edu (N. Sun), yingemma.chen@gmail.com (Y. Chen).
[1] This work has no connection with MEAG New York.

0378-4266/$ - see front matter © 2011 Elsevier B.V. All rights reserved.
doi:10.1016/j.jbankfin.2011.06.005

However, the assumptions of time homogeneity and Markovian behavior of the rating process in (1) have been challenged by recent studies on the presence of various non-Markovian behaviors such as industry heterogeneity, ratings drift, and time variations; see details in Altman (1998), Blume et al. (1998), Nickell et al. (2000), Bangia et al. (2002), Christensen et al. (2004) and Liao et al. (2009) and reference therein. Since the ratings are sometimes considered to ‘‘look through the cycle’’, firms’ ratings could be remarkably stable over credit cycles; see Fons (2002), Cantor and Mann (2003a,b), Altman and Rijken (2004) and Bruche and González-Aguado (2010). To address this issue without using firm-specific information, Frydman and Schuermann (2008) consider the behavior that firms with the same rating migrate at different speeds and propose a mixture model of two independent continuous time homogeneous Markov chains for the rating migration process. They apply the proposed model to analyze corporate credit rating history from Standard & Poor’s spanning 1981–2002.

This paper studies the variations in rating transition matrices and proposes a hidden Markov model which extends the continuous time homogeneous Markov chain and characterizes unobserved structural breaks in the credit rating dynamics. From a credit risk modeling perspective, it is important to study variations in rating transitions attributable to the state of the economy. Some credit risk modeling approaches classify the states of the economy as finite regimes and assume transition matrices change over time with the states of the economy. This ignores the fact that the states of the economy at different periods might be different even they are in the same regime. From this perspective, it would be better not to restrict the number of regimes. We hence assume the state of the economy to be continuous, and model the structural changes in the economy as the shifts of the state of the economy in a continuous space.

Motivated by the discussion above, we propose a stochastic structural break model for rating migration processes. Different from Crowder et al. (2005) who model the occurrence of defaults within a bond portfolio as hidden Markov process, we assume that, in our model, the generators of the rating transition matrices are constant between two adjacent structural changes in the economy. As the states of the economy may have infinite regimes, the generators are assumed to follow a continuous state and continuous time nonhomogeneous hidden Markov process, and be piecewise constant with the number, time and magnitude of the structural changes unobserved. We further model the structural breaks in the generators as a compounded Poisson process, in which the times of structural breaks follow a Poisson process with a constant rate and the entries of post-change generator matrices follow a Gamma distribution. These assumptions allow us to derive the distributions of the time-varying generators of rating migration matrices and the probability of structural breaks at each time period, given firms’ transition history. The derived distribution of generator matrices at a given period is a mixture of Gamma distributions, and the weights of mixture components can be computed explicitly using historical observations. As the number of mixture components changes over time, the model is allowed to incorporate various non-Markovian behaviors in empirical studies. From this perspective, our model extends the mixture model of two independent continuous time homogeneous Markov chains in Frydman and Schuermann (2008). The proposed model also implies a prediction formula for generator matrices with probable structural breaks in the future, for which the intensity or probability of structural breaks in generator matrices should be calibrated by historical observations.

We use the proposed model and developed inference procedure to study the monthly corporate credit ratings provided by Standard & Poor’s from January 1985 to September 2009. We show that the generator and transition matrices of rating transitions are indeed changing over time, and the estimated structural breaks are not only statistically significant but economically meaningful as well. The estimated times of structural breaks match the times of several significant structural changes in the economy. We also demonstrate that the generator or transition matrices in different industry categories have different behaviors, and specifically, industry sectors related to finance services are more susceptible to economic changes than other sectors. We further conduct out-of-sample forecast evaluations of our model against the time homogeneous Markov model and compare the performance of our model with that of a time-homogeneous Markov model. The comparison shows that our model provides a more accurate prediction for rating generator or probability matrices.

The remainder of the paper is organized as follows. Section 2 develops a stochastic structural break model for generator matrices and its inference procedure using continuous credit rating history. In Section 3, we study the in-sample and out-of-sample performance of our model on the data set, and discuss the estimation results and their economic implications. Section 4 provides some concluding remarks.

## 2. Stochastic structural break model

We assume that a rating transition process of an obligor follows a $K$-state non-homogeneous continuous time Markov process. This process is further characterized by a transition probability matrix $P(s,t)$ over the period $(s,t)$, in which the $ij$th element of $P(s,t)$ represents the probability that an obligor starting in state $i$ at time $s$ is in state $j$ at time $t$. Suppose that there are $n$ rating transitions observed over the period $(s,t)$. For a transition time $t_i$ in $(s,t)$, denote $\Delta N_{kj}(t_i)$ the number of transitions observed from state $k$ to state $j$ at time $t_i$, $\Delta N_k(t_i) = \sum_{1\leqslant j\leqslant K, j\neq k}\Delta N_{kj}(t_i)$, and $Y_k(t_i)$ the number of firms in state $k$ right before time $t_i$. The transition matrix $P(s,t)$ can be consistently estimated by the product-limit estimator

$$\widehat{P}(s,t) = \prod_{i=1}^{n}(I + \Delta\widehat{A}(t_i)),$$

in which

$$\Delta\widehat{A}(t_i) = \begin{pmatrix} -\frac{\Delta N_1(t_i)}{Y_1(t_i)} & \frac{\Delta N_{12}(t_i)}{Y_1(t_i)} & \frac{\Delta N_{13}(t_i)}{Y_1(t_i)} & \cdots & \frac{\Delta N_{1K}(t_i)}{Y_1(t_i)} \\ \frac{\Delta N_{21}(t_i)}{Y_2(t_i)} & -\frac{\Delta N_2(t_i)}{Y_2(t_i)} & \frac{\Delta N_{23}(t_i)}{Y_2(t_i)} & \cdots & \frac{\Delta N_{2K}(t_i)}{Y_2(t_i)} \\ \vdots & \vdots & & \cdots & \vdots \\ \frac{\Delta N_{K-1,1}(t_i)}{Y_{K-1}(t_i)} & \frac{\Delta N_{K-1,2}(t_i)}{Y_{K-1}(t_i)} & \cdots & -\frac{\Delta N_{K-1}(t_i)}{Y_{K-1}(t_i)} & \frac{\Delta N_{K-1,K}(t_i)}{Y_{K-1}(t_i)} \\ 0 & 0 & \cdots & \cdots & 0 \end{pmatrix},$$

see Andersen et al. (1995, Section IV.4). In the matrix above, the $k$th diagonal element counts the fraction of the exposed firms $Y_k(t_i)$ leaving the state at time $t_i$, and the $kj$th off-diagonal element count the fraction of transitions from the $k$th category to the $j$th category in the number of exposed firms at time $t_i$. Note that the variable $Y$ has incorporated the case of censoring for which there is no change in the estimator at the time of a censoring event. Furthermore, the last row in $\Delta\widehat{A}(t_i)$ is zero because the $k$th state (i.e., default state) is absorbent. Since the purpose of our model is to incorporate structural changes into credit rating dynamics, we assume from now on that the non-homogeneous continuous time Markov process can be decomposed as piecewise homogeneous continuous time Markov processes with unobserved structural breaks.

### 2.1. Model specification

Specifically, we assume that the structural breaks in credit rating generator matrices follow a Poisson process $\{N_{\Lambda}(t); t \geqslant 0\}$ with constant rate $\eta$, hence the duration between two adjacent structural breaks follows an exponential distribution with mean $1/\eta$. The generator matrices between two adjacent structural breaks are constant and the generator matrix at time $t$ is characterized as $\Lambda(t) = Q_{N_\Lambda(t)}$, where $Q_1, Q_2, \ldots$ are independent and identically distributed (i.i.d.) random generator matrices such that the off-diagonal elements $\lambda^{(i,j)}$ follow independently a Gamma $(\alpha_{ij}, \beta_i)$ prior distribution with the density function

$$g(\lambda^{(i,j)}) = \frac{\beta_i^{\alpha_{ij}}}{\Gamma(\alpha_{ij})}\left[\lambda^{(i,j)}\right]^{\alpha_{ij}-1}\exp(-\lambda^{(i,j)}\beta_i), \quad (i,j) \in \mathcal{K}, \tag{2}$$

in which $\mathcal{K} = \{(i,j) | i \neq j, 1 \leqslant i \leqslant K-1,\ 1 \leqslant j \leqslant K\}$. Note that the elements of the last row in the generator matrix, representing the rating migrations from the default category to others, are usually considered to be zero, our model also follows this convention. Furthermore, the above assumptions suggest that $\Lambda(t)$ follows a variant of compound Poisson process with rate $\eta$, and the time-dependent credit rating transition matrix $P(s,t)$ over the period $(s,t)$ can be characterized in the following way.

If there are $M$ structural breaks (or jumps) in the period $(s,t)$ and these $M$ change times, assuming as $s < \tau_1 < \ldots < \tau_M < t$, are observed, then the transition matrix in the period $(s,t)$ can be characterized as

$$\begin{aligned} P(s,t) &= \prod_{k=1}^{M+1} P(\tau_{k-1},\tau_k) = \prod_{k=1}^{M+1}\exp\left(\int_{\tau_{k-1}}^{\tau_k}\Lambda(u)du\right) \\ &= \prod_{k=1}^{M+1}\exp\left[(\tau_k - \tau_{k-1})\Lambda(\tau_k -)\right], \end{aligned}$$

in which $\tau_0 = s$, $\tau_{M+1} = t$. Note that the exponent in the above equation usually can not be simplified to $P(s,t) = \exp\left[\int_{\tau_0}^{\tau_{M+1}}\Lambda(u)du\right]$, because the components $P(\tau_0, \tau_1), \ldots, P(\tau_M, \tau_{M+1})$ may not commute.

If the generator $\Lambda(t)$ has no changes over the period $(s,t)$, then $P(s,t)$ becomes homogeneous and can be characterized by

$$P(s,t) = \exp\left(\int_s^t \Lambda(u)du\right) = \exp[(t-s)\Lambda(t-)].$$

In such a case, since we assume that the elements of $\Lambda(t)$ follow the conjugate Gamma priors (2), the posterior distribution of $\Lambda(t)$ given transition history over the period $(s,t)$ can be computed as follows. Suppose that there are $n$ realizations of a Markov chain with generator matrices $\Lambda(t)$. Denote, for the period $(s,t)$,$K_{s,t}^{(i,j)}$ the number of transitions from category $i$ to category $j$, $S_{s,t}^{(i)}$ the amount of time spent in category $i$, $\lambda_{s,t}^{(i,j)}$ the $ij$th entry in the generator $\Lambda(t)$, and $\mathcal{Y}_{s,t}$ the observed rating transitions over the period $(s,t)$, then the likelihood of $\mathcal{Y}_{s,t}$ given the constant $\Lambda(t)$ is expressed as

$$\begin{aligned} &\exp\left\{\sum_{i=1}^{K}\left[\sum_{j\neq i}K_{s,t}^{(i,j)}\log\lambda_{s,t}^{(i,j)} - \left(\sum_{j\neq i}\lambda_{s,t}^{(i,j)} + 1 - K\right)S_{s,t}^{(i)}\right]\right\} \\ &\quad\propto \prod_{i\neq j}(\lambda_{s,t}^{(i,j)})^{K_{s,t}^{(i,j)}} e^{-\lambda_{s,t}^{(i,j)}S_{s,t}^{(i)}}, \end{aligned} \tag{3}$$

see Küchler and Sørensen (1997, p. 26). Combining (3) with the assumed Gamma prior (2) yields the posterior distribution of $\lambda_{s,t}^{(i,j)}$ given $\mathcal{Y}_{s,t}$, which is Gamma $(K_{s,t}^{(i,j)} + \alpha_{ij}, S_{s,t}^{(i)} + \beta_i)$. Hence the element $\lambda_{s,t}^{(i,j)}$ can be estimated by the posterior mean of the Gamma distribution, i.e., $\hat{\lambda}_{s,t}^{(i,j)} = (K_{s,t}^{(i,j)} + \alpha_{ij})/(S_{s,t}^{(i)} + \beta_i)$.

Finally, we assume that firms' rating migrations from state $i$ to state $j$ at the period $(s,t)$ are conditional independent given the generator matrix at the period $(s,t)$. Furthermore, we assume that there are $n$ rating transitions observed over the sample period $(0,T)$. We also consider an evenly spaced partition for the period $(0,T)$, $0 = t_0 < t_1 < \ldots < t_L = T$, and assume that structural breaks can only happen at the times $t_1, \ldots, t_L$. We define the variables $I_1 = 1$ and $I_l = N_\Lambda(t_l-) - N_\Lambda(t_{l-1}-)$ for $l = 2, \ldots, L$ to indicate if $\Lambda(t)$ are same at the periods $(t_{l-2},\ t_{l-1})$ and $(t_{l-1},\ t_l)$, then $\{I_l\}$ is a sequence of i.i.d. Bernoulli random variables with success probability $p = 1 - \exp(-\eta T/L)$. We also assume that there is at most one structural break at time $t_l$. Note that these assumptions are reasonable to identify structural breaks in $\Lambda(t)$ as long as the partition of $(0,T)$ is fine enough.

### 2.2. Posterior distribution of $\Lambda(t)$ given $\mathcal{Y}_{(0,T)}$

We now derive the posterior distribution of $\Lambda(t_l) = (\lambda_{t_{l-1},t_l}^{(i,j)})_{(i,j)\in\mathcal{K}}$ given $\mathcal{Y}_{(0,T)}$. Let $R_l = \max\{t_{m-1} | I_m = 1,\ m \leqslant l\}$, i.e., $R_l$ represents the time of the most recent structural break up to time $t_{l-1}$. From Section 2.1, we know that the conditional distribution of $\lambda_{t_{m-1},t_l}^{(i,j)}$ given $R_l = t_{m-1}$ and $\mathcal{Y}_{t_{m-1},t_l}$ is Gamma $(K_{t_{m-1},t_l}^{(i,j)} + \alpha_{ij}, S_{t_{m-1},t_l}^{(i)} + \beta_i)$. Let $p_{m,l} = P(R_l = t_{m-1} | \mathcal{Y}_{t_{m-1},t_l})$. Then Appendix A shows that the posterior distribution of $\lambda_{t_{l-1},t_l}^{(i,j)}$ given $\mathcal{Y}_{(0,t_l)}$ can be expressed as a mixture of Gamma distributions,

$$\lambda_{t_{l-1},t_l}^{(i,j)} | \mathcal{Y}_{(0,t_l)} \sim \sum_{m=1}^{l} p_{m,l}\mathrm{Gamma}(K_{t_{m-1},t_l}^{(i,j)} + \alpha_{ij}, S_{t_{m-1},t_l}^{(i)} + \beta_i). \tag{4}$$

The mixture weight can be calculated recursively by $p_{m,l} = p_{m,l}^*/\sum_{m=1}^{l}p_{m,l}^*$, in which

$$p_{m,l}^* = \begin{cases} pf_{l,l}/f_{0,0} & m = l, \\ (1-p)p_{m,l-1}f_{m,l}/f_{m,l-1} & m < l. \end{cases} \tag{5}$$

The terms $f_{m,l}$ and $f_{0,0}$ in (5) are expressed as follows,

$$\begin{aligned} f_{m,l} &= \prod_{i,j\in\mathcal{K}}\Gamma(K_{t_{m-1},t_l}^{(i,j)} + \alpha_{ij})/(S_{t_{m-1},t_l}^{(i)} + \beta_i)^{(K_{t_{m-1},t_l}^{(i,j)} + \alpha_{ij})}, \\ f_{0,0} &= \prod_{i,j\in\mathcal{K}}\Gamma(\alpha_{ij})/\beta_i^{\alpha_{i,j}}. \end{aligned} \tag{6}$$

To find the posterior distribution of $\lambda_{t_{l-1},t_l}^{(i,j)}$ given $\mathcal{Y}_{(0,T)}$ for $1 \leqslant l < L$, we use the argument in Appendix A to show that the posterior distribution of $\lambda_{t_{l-1},t_l}^{(i,j)}$ given $\mathcal{Y}_{(t_l,T)}$ is given by

$$\begin{aligned} \lambda_{t_{l-1},t_l}^{(i,j)} | \mathcal{Y}_{(t_l,T)} &\sim p\ \mathrm{Gamma}(\alpha_{ij}, \beta_i) + (1-p) \\ &\quad\times \sum_{k=l+1}^{L}\tilde{p}_{k,l+1}\ \mathrm{Gamma}(K_{t_l,t_k}^{(i,j)} + \alpha_{ij}, S_{t_l,t_k}^{(i)} + \beta_i), \end{aligned} \tag{7}$$

in which the mixture weight $\tilde{p}_{k,l+1} = \tilde{p}_{k,l+1}^*/\sum_{k=l+1}^{L}\tilde{p}_{k,l+1}^*$ and

$$\tilde{p}_{k,l+1}^* = \begin{cases} pf_{l+1,l+1}/f_{00} & k = l+1, \\ (1-p)q_{l+2,k}f_{l+1,k}/f_{l+2,k} & k > l+1. \end{cases}$$

Following the proof in Appendix A, one can use the Bayes theorem to combine (4) and (7) to obtain the posterior distribution of $\lambda_{t_{l-1},t_l}^{(i,j)}$ given $\mathcal{Y}_{(0,T)}$, which is a mixture of Gamma distributions

$$\begin{aligned} \lambda_{t_{l-1},t_l}^{(i,j)} | \mathcal{Y}_{(0,T)} &\sim \sum_{1\leqslant m\leqslant l\leqslant k\leqslant L}\pi_{mlk}\mathrm{Gamma}(K_{t_{m-1},t_k}^{(i,j)} + \alpha_{ij}, S_{t_{m-1},t_k}^{(i)} + \beta_i), \\ &\quad 1 \leqslant l < L, \end{aligned} \tag{8}$$

in which $\pi_{mlk} = \pi_{mlk}^*/\sum_{1\leqslant m\leqslant l\leqslant k\leqslant L}\pi_{mlk}^*$ and

$$\pi_{mlk}^* = \begin{cases} pp_{m,l} & m \leqslant l = k, \\ (1-p)p_{m,l}\tilde{p}_{k,l+1}f_{m,k}f_{00}/(f_{m,l}f_{l+1,k}) & m \leqslant l < k. \end{cases} \tag{9}$$

### 2.3. Statistical inference and parameter estimation

#### 2.3.1. One- and multi-periods transition matrices

Given the posterior distribution (4) and (8), the $ij$th element of $\Lambda(t_l)$ can be estimated by its posterior mean, i.e.,

$$\hat{\lambda}^{(i,j)}_{t_{l-1},t_l} := E(\lambda^{(i,j)}_{t_{l-1},t_l}|\mathcal{Y}_{(0,T)}) = \begin{cases} \sum_{m=1}^{L} p_{m,l} \frac{K^{(i,j)}_{t_{m-1},t_l}+\alpha_{ij}}{S^{(i)}_{t_{m-1},t_l}+\beta_i} & l = L, \\ \sum_{1\leqslant m\leqslant l\leqslant k\leqslant L} \pi_{mlk} \frac{K^{(i,j)}_{t_{m-1},t_k}+\alpha_{ij}}{S^{(i)}_{t_{m-1},t_k}+\beta_i} & l < L. \end{cases} \quad (10)$$

The posterior variance of $\lambda^{(i,j)}_{t_{l-1},t_l}$ can be computed as follows,

$$\mathrm{Var}(\lambda^{(i,j)}_{t_{l-1},t_l}|\mathcal{Y}_{(0,T)}) = E\left[\left(\lambda^{(i,j)}_{t_{l-1},t_l}\right)^2|\mathcal{Y}_{(0,T)}\right] - \left[E(\lambda^{(i,j)}_{t_{l-1},t_l}|\mathcal{Y}_{(0,T)})\right]^2$$
$$= \begin{cases} \sum_{m=1}^{L}\left\{p_{m,l}\frac{\left[K^{(i,j)}_{t_{m-1},t_l}+\alpha_{ij}\right]\left[K^{(i,j)}_{t_{m-1},t_l}+\alpha_{ij}+1\right]}{\left[S^{(i)}_{t_{m-1},t_l}+\beta_i\right]^2}\right\} - \left[\hat{\lambda}^{(i,j)}_{t_{l-1},t_l}\right]^2 & l = L, \\ \sum_{1\leqslant m\leqslant l\leqslant k\leqslant L}\left\{\pi_{mlk}\frac{\left[K^{(i,j)}_{t_{m-1},t_k}+\alpha_{ij}\right]\cdot\left[K^{(i,j)}_{t_{m-1},t_k}+\alpha_{ij}+1\right]}{\left[S^{(i)}_{t_{m-1},t_k}+\beta_i\right]^2}\right\} - \left[\hat{\lambda}^{(i,j)}_{t_{l-1},t_l}\right]^2 & l < L. \end{cases} \quad (11)$$

Making use of (10) and $\widehat{\Lambda}(t_l) = (\hat{\lambda}^{(i,j)}_{t_{l-1},t_l})$, the one-period probability transition matrix $P(t_{l-1},t_l)$ can be estimated by

$$\widehat{P}(t_{l-1},t_l) = \exp\left[(t_l - t_{l-1})\widehat{\Lambda}(t_l)\right], \quad (12)$$

and the multi-period probability transition matrix $P(t_{m-1},t_k)$ can be estimated by

$$\widehat{P}(t_{m-1},t_k) = \prod_{l=m}^{k} P(t_{l-1},t_l) = \prod_{l=m}^{k} \exp\left[(t_l - t_{l-1})\widehat{\Lambda}(t_l)\right]. \quad (13)$$

Note that under certain regularity conditions, we could use (11) to compute asymptotic variances of the transition probability matrices (12) and (13), which provide an alternative way to compute the confidence sets discussed by Christensen et al. (2004).

*2.3.2. Probability of structural breaks*

Note that for $m \leqslant l \leqslant k$, the posterior distribution of $\lambda^{(i,j)}_{t_{l-1},t_l}$ given $\mathcal{Y}_{(0,T)}$ and the event $\{I_m = 1,\ I_{m+1} = \cdots = I_k = 0,\ I_{k+1} = 1\}$ is Gamma $(K^{(i,j)}_{t_{m-1},t_k} + \alpha_{ij}, S^{(i)}_{t_{m-1},t_k} + \beta_i)$, one can see from (8) that

$$\pi_{mlk} = P(I_m = 1, I_{m+1} = \cdots = I_k = 0, I_{k+1} = 1|\mathcal{Y}_{(0,T)})$$
$$= P\big(\Lambda(t_{m-1}-) \neq \Lambda(t_m-) = \ldots = \Lambda(t_k-) \neq \Lambda(t_{k+1}-)|\mathcal{Y}_{(0,T)}\big),$$

i.e., $\{\pi_{mlk}; 1 \leqslant m \leqslant l \leqslant k \leqslant L\}$ represents the conditional distributions of two most recent structural breaks before and after time $t_l$. Hence $P(I_{l+1} = 1|\mathcal{Y}_{(0,T)})$ can be computed by

$$P(I_{l+1} = 1|\mathcal{Y}_{(0,T)}) = \sum_{m=1}^{l} \pi_{mll} = p \Big/ \sum_{1\leqslant m\leqslant l\leqslant k\leqslant L} \pi^*_{mlk}. \quad (14)$$

The probability above can be used to infer whether structural breaks have occurred at time $t_l$. In fact, if there does exist a structural break at $t_l$ and the length $t_l - t_{l-1}$ of each period is fixed, one can show that, under the model assumption, the probability (14) converges to 1 if the number of observed transitions at $t_l$ goes to infinity. The key of the proof is to consider tests of structural break at $t_l$ and use the fact that the mixture weights in (4) and (8) are functions of those test statistics. (We skip the proof as it is tedious but technically not difficult.)

*2.3.3. Prediction of generator and transition matrices*

Since the model assume possible structural changes in the rating dynamics, it is interesting to see the out-of-sample performance of the model in the presence of possible structural breaks. Combining the assumption on the dynamics of structural breaks and the "on-line" estimate (4), we obtained the one-period ahead prediction for the distribution of the generator matrix $\Lambda(t_{L+1}) = \left(\lambda^{(i,j)}_{t_L,t_{L+1}}\right)$ in the period $(t_L, t_{L+1}) = (T, (1+\frac{1}{L})T)$ given the observation $\mathcal{Y}_{(0,T)}$

$$\lambda^{(i,j)}_{t_L,t_{L+1}}|\mathcal{Y}_{(0,t_L)} \sim p\ \mathrm{Gamma}(\alpha_{ij},\beta_i) + (1-p)$$
$$\times \sum_{m=1}^{L} p_{m,L}\ \mathrm{Gamma}(K^{(i,j)}_{t_{m-1},t_L} + \alpha_{ij}, S^{(i)}_{t_{m-1},t_L} + \beta_i). \quad (15)$$

Hence the one-period ahead prediction $\widetilde{\Lambda}_l(t_{L+1}) = \left(\tilde{\lambda}^{(i,j)}_{t_L,t_{L+1}}\right)$ given $\mathcal{Y}_{(0,t_L)}$ can be obtained via the posterior mean of (15),

$$\tilde{\lambda}^{(i,j)}_{t_L,t_{L+1}} = E(\lambda^{(i,j)}_{t_L,t_{L+1}}|\mathcal{Y}_{(0,t_L)}) = p\frac{\alpha_{ij}}{\beta_j} + (1-p)\sum_{m=1}^{L} p_{m,L}\frac{K^{(i,j)}_{t_{m-1},t_L} + \alpha_{ij}}{S^{(i)}_{t_{m-1},t_L} + \beta_i}, \quad (16)$$

and the one-period ahead prediction for the transition matrix can be obtained by the exponential transformation,

$$\widetilde{P}_L((t_L,t_{L+1}]) = \exp\left[(t_{L+1} - t_L)\widetilde{\Lambda}_l(t_{L+1})\right]. \quad (17)$$

Using the argument above, the multi-period ahead prediction of generator and probability transition matrices can be obtained similarly.

*2.3.4. Parameter estimation*

The inference procedure in the above sections involve the hyperparameters $p$ (or $\eta$), $\alpha_{ij}$ and $\beta_j$ for $(i,j) \in \mathcal{K}$, which can be replaced by their estimates in the empirical Bayes approach. Appendix B presents the maximum likelihood and expectation–maximization (EM) approaches to estimate these hyperparameters. Such empirical Bayes approach can be interpreted as the process of calibrating the probability of structural breaks and the prior distribution of rating transitions by historical observations.

## 3. Data analysis

The data are obtained from COMPUSTAT and consist of Standard & Poor's monthly credit ratings of firms over 25 years starting January 1985 and ending September 2009. There are a total of 2,160,809 rating records and 21,755 firms whose ratings were recorded at the end of each month. The data also contain the industry sector of each firm based on the Global Industry Classification Standard (GICS), but there is no other firm-specific information. The data contain ten rating categories, $\mathcal{AAA}, \mathcal{AA}, \mathcal{A}, \mathcal{BBB}, \mathcal{BB}, \mathcal{B}, \mathcal{CCC}, \mathcal{CC}, \mathcal{C}$ and $\mathcal{D}$ (default), and 25 rating subcategories. Subcategories are obtained by possibly adding "+" or "−" to the letter grade of categories, which shows relative standing within the major rating categories. We then clean the data as follows. We first group $\mathcal{C}$ and $\mathcal{CC}$ into $\mathcal{CCC}$ as the records in the former two rating categories are relatively few, and then remove rating records of two invalid ratings "N.M." and "Suspended". After the above data-cleaning process, we extract the initial rating and transition information from the rating records. Then we obtain 5185 initial rating and 6486 transition records covering 5185 firms, and eight rating categories, $\mathcal{AAA}, \mathcal{AA}, \mathcal{A}, \mathcal{BBB}, \mathcal{BB}, \mathcal{B}, \mathcal{CCC}$, and $\mathcal{D}$. In Fig. 1, the upper panel shows the percentage of the eight rating categories in each year during the period 1985–2009 (Note that records in 2009 only cover the first nine months), the bottom panel shows the number of initial rating and transition records in each year. Note that the first rating transition of a firm happens after it migrates away from its initial rating, and in 1985, there are 1286 initial ratings and only one rating transition. Hence the analysis of rating transition matrices in the following is based on the data from January 1986 to September 2009.

*3.1. Estimating the model*

We begin our analysis by estimating the stochastic structural break model at the level of $K = 8$ rating categories from January 1986 to September 2009. For convenience, we denote the beginning of 1986 as the time 0 and the end of September 2009 as the time $T$. Since the rating records are monthly, we partition the

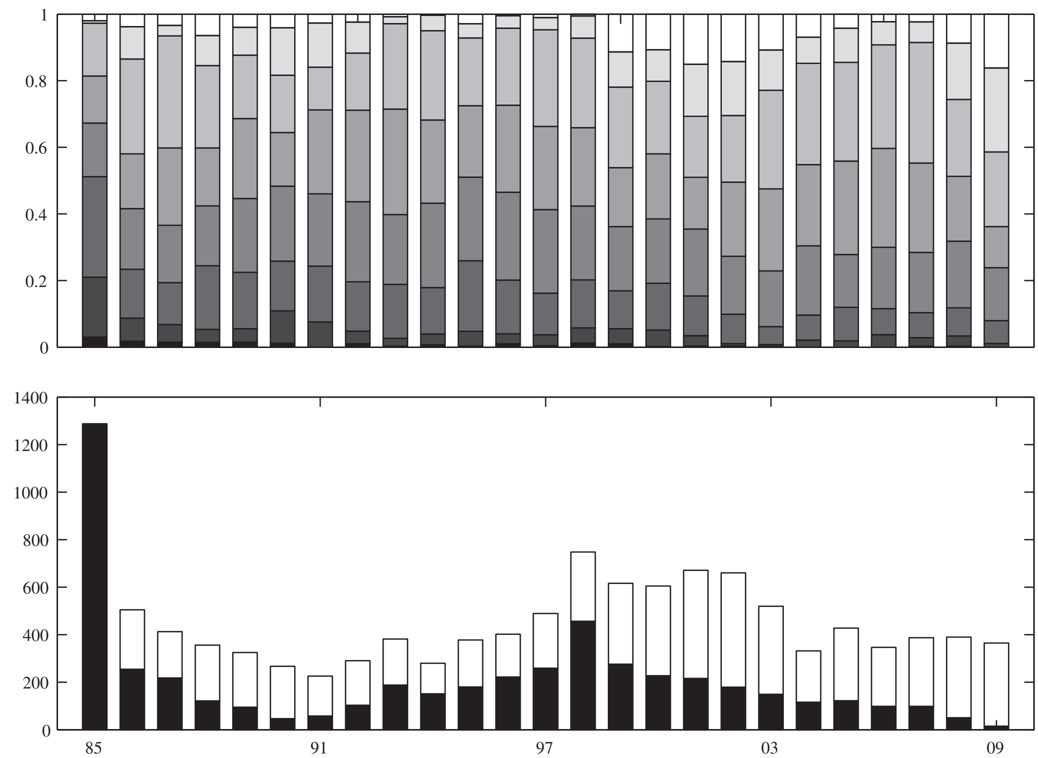


**Fig. 1.** Percentage of annual rating records (Upper panel) and the number of initial rating and transitions (Bottom panel) from January 1985 to September 2009. Upper panel: The bars in each year show the percentages of ratings in $\mathcal{AAA}, \mathcal{AA}, \mathcal{A}, \mathcal{BBB}, \mathcal{BB}, \mathcal{B}$ and $\mathcal{CCC}$ from top to bottom. Bottom panel: The number of initial ratings (black) and rating transitions (white) in each year.

sample period from January 1986 to September 2009 to $L = 285$ intervals and each interval corresponds to a calendar month. For example, the interval (0,1/12) represents the period of January 1986. Then we use the EM algorithm in Appendix B to obtain the following estimates of hyperparameters,

$$\hat{p} = 0.0143 \quad \text{or} \quad \hat{\eta} = 0.1732,$$

$$(\hat{\beta}_i) = (\,243.11 \quad 330.81 \quad 495.66 \quad 411.86 \quad 285.90 \quad 311.46 \quad 140.79 \quad \cdot\,),$$

$$(\hat{\alpha}_{ij}) = \begin{pmatrix} \cdot & .9067 & .9193 & .9193 & .9193 & .9193 & .9193 & .9193 \\ .9316 & \cdot & .9188 & .9316 & .9316 & .9316 & .9316 & .9316 \\ .9333 & .9333 & \cdot & .9333 & .9333 & .9333 & .9333 & .9333 \\ .9325 & .9325 & .9325 & \cdot & .9325 & .9325 & .9325 & .9325 \\ .9297 & .9297 & .9297 & .9297 & \cdot & .9297 & .9297 & .9297 \\ .9278 & .9278 & .9278 & .9278 & .9278 & \cdot & .9268 & .9277 \\ .9168 & .9168 & .9168 & .9168 & .9168 & .9125 & \cdot & .2655 \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \end{pmatrix}.$$

Using the above estimated hyperparameters, we first compute the estimates of time-varying generator matrices $\Lambda(t_l)$ by (10) and the estimates for transition probability matrices by $\widehat{P}((t_{l-1}, t_l]) = \exp\left[(t_l - t_{l-1})\widehat{\Lambda}(t_l)\right]$ for $l = 1, \ldots, L$. Fig. 2 shows some estimated one-year time-varying transition probabilities at each month. We find that these transition probabilities are piecewise constant at most periods, and their fluctuations seem to show some structural changes in the credit market. Furthermore, all transition probabilities change frequently in the period August 1990–May 1993, indicating that the credit market in that period was unstable and experiencing certain fluctuations. The sharp changes of transition probabilities in Fig. 2 are intrinsically linked with the structural changes in the credit market. For example, during the 2007–2009 financial crisis, all transition probabilities in Fig. 2 change at September 2007 ($l = 261$), then after seven month's fluctuation, most of these transition probabilities shift to another regime at April 2008 and to another level at October 2008. Such variations are significant for the transition probabilities of $\mathcal{AAA}-\mathcal{AAA}$, $\mathcal{AAA}-\mathcal{D}$, $\mathcal{A}-\mathcal{BBB}$ and $\mathcal{BB}-\mathcal{D}$. In particular, the default probabilities transited from the rating categories $\mathcal{AAA}$ increase from 0.004 to 0.007 at April 2008 and from 0.007 to 0.030 at October 2008, and the default probabilities transited from the rating categories $\mathcal{BB}$ increase from 0.0004 to 0.0008 at April 2008 and from 0.0008 to 0.0090 at October 2008. As these estimates are characterized by significant changes over different period, we next discuss the probability of structural breaks at each period.

To show the variation of the overall transition probability matrices over the sample period, we use the singular value decomposition (SVD) metric that is proposed in Jafry and Schuermann (2004). Such SVD metric for a transition matrix $P$ is defined as

$$M(P) = \frac{1}{K}\sum_{i=1}^{K}\sqrt{e_i[(P-I)'(P-I)]}, \tag{18}$$

in which $I$ is a $K \times K$ identity matrix and $e_i(\cdot)$ is the $i$th eigenvalue of the matrix. Note that $P - I$ represents the magnitude of the implied mobility of transition matrix. As shown in Jafry and Schuermann (2004), such metric has the capability of approximating the average probability of migrations across all rating categories. Fig. 3 shows the metric $M(\widehat{P}((t_{l-1}, t_l]))$ of estimated transition probability matrices for $l = 1, \ldots, L$. Note that the SVD metrics for the whole sample are piecewise constant during most periods. Furthermore, the SVD value (0.131) in August 1986–June 1990 is almost same as the value (0.135) in July 1999–May 2003. Similarly, the SVD value (0.082) in May 1993–February 1998 is close to the value (0.075) in June 2003–August 2007. As these four periods cover 203 months of the whole sample period (285 months), it seems to indicate that there exist regular states or regimes in the rating dynamics. We also notice that the SVD metric shows frequent changes during August 1990–April 1993, suggesting that the credit environment is unstable during

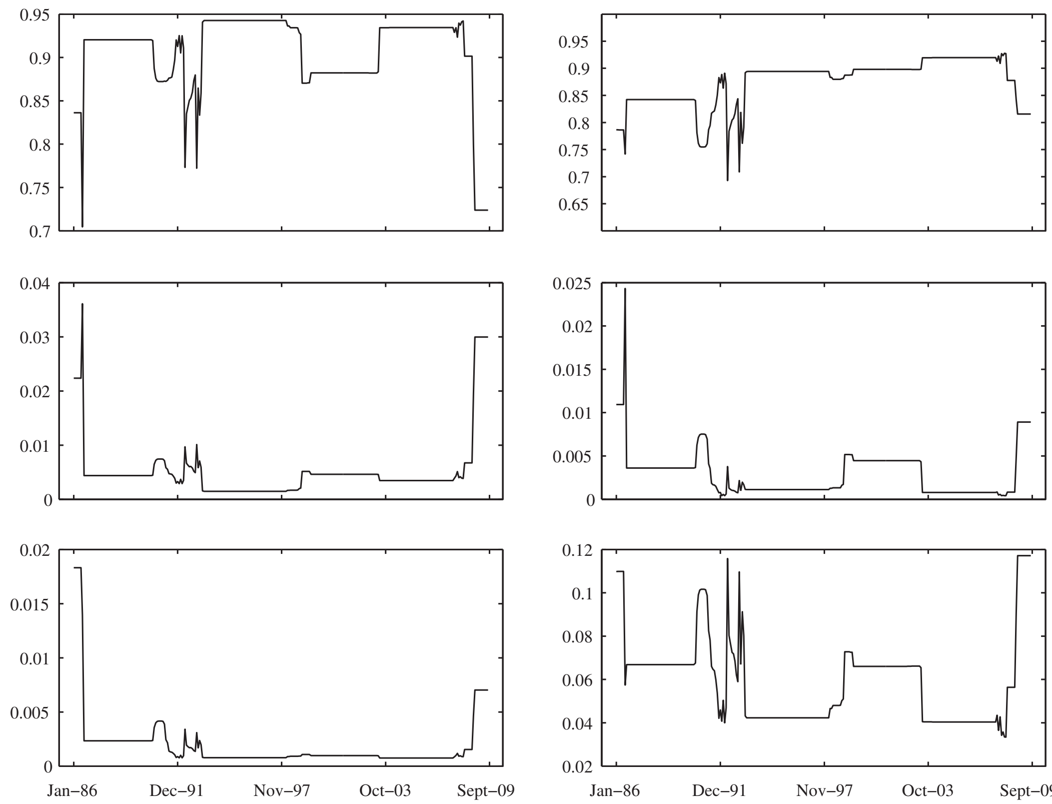


**Fig. 2.** Estimated transition probabilities $\mathcal{AAA}-\mathcal{AAA}, \mathcal{BB}-\mathcal{BB}, \mathcal{AAA}-\mathcal{D}, \mathcal{BB}-\mathcal{D}, \mathcal{AA}-\mathcal{B}$, and $\mathcal{A}-\mathcal{BBB}$ (from left to right and top to bottom).

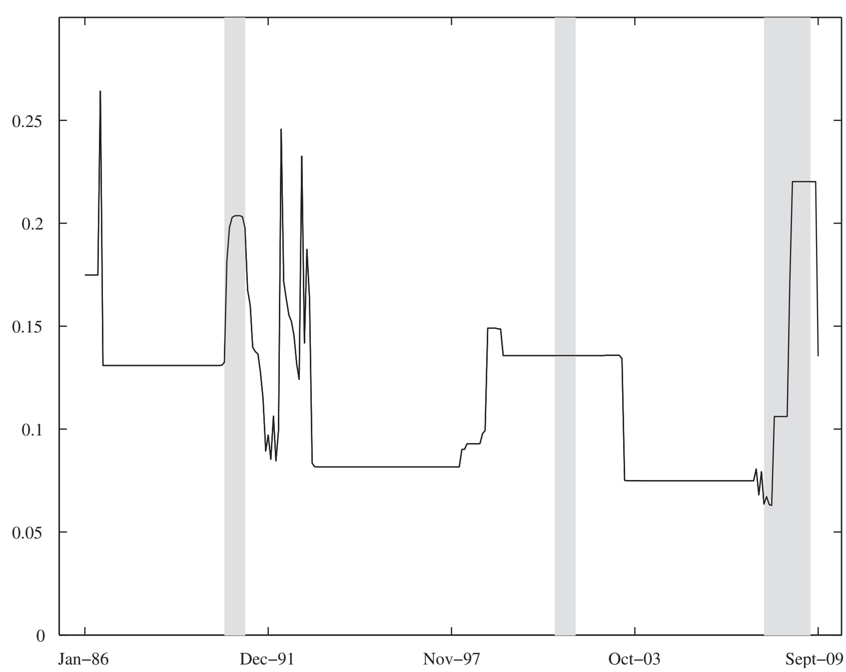


**Fig. 3.** The SVD metric of estimated transition probability matrices over time.

the period. Fig. 3 also shows three periods of economic recessions announced by NBER as shaded areas. During the first recession (July 1990–March 1991), the SVD metric of estimated transition probability matrices looks roughly constant (the values are between 0.18 and 0.20) and the period of such seemingly invariance overlaps with that of the first recession. The second recession (March 2001–November 2001) is in the middle of a much longer period (July 1999–May 2003) during which the SVD metric is constant, suggesting that business cycles do not necessarily have a significant impact on credit rating dynamics. The beginning month of the third recession (December 2007–June 2009) is three months after the period June 2003–August 2007, during which the SVD values are very stable, indicating that the credit environment began to change before the third recession. Furthermore, during the third recession, the SVD metric jumps from 0.06 in December 2007–March 2008 to 0.11 in April 2008–September 2008, then to 0.22 in October

2008–August 2009, and finally drops to 0.13 in September 2009, which is three months after the end of the recession. Note that such variations in the SVD metric during September 2007–September 2009 are consistent with the course of the financial crisis beginning in 2007, which will be further discussed in the next section.

### 3.2. Probability of structural break and segmentation

The proposed model assumes a Poisson process for structural breaks in continuous-time credit rating generator matrices (or a Bernoulli process for structural breaks in discrete time). We have derived the probability of a structural break at the $l$th month in Section 2.3.2, and now we use (14) to estimate the probability of a structural break in $\Lambda(t_l)$ at $t_l$ = January 1986, ..., September 2009. The result is shown in Fig. 4. We notice that most estimated probabilities of structural break are almost zero except at a few months. In particular, there are five months at which the probabilities of structural breaks are much larger than zero. These probabilities are 0.992 (July 1986), 0.527 (April 1991), 0.868 (January 1999), 0.969 (June 2003), and 0.572 (October 2008). It is intriguing to interpret these estimated probabilities and the corresponding times. The estimated probability of structural break at July 1986 might link with the fact that the US economy in 1986 began to change from a rapid recovery to a slower expansion, however, such interpretation may be undermined by the sample issue that the numbers of rating transitions before and after July 1986 are significantly different, as shown in Fig. 1. The estimated structural break at April 1991 is right after the ending month of a recession announced by NBER, which signifies the economic recovery from the 1990–1991 recession. We also notice that the estimated probability of structural break at June 1992 is 0.124, which is relatively small comparing with the magnitudes of other spikes. This small probability of structural break indicates that rating generator matrices are not stable around June 1992, which is consistent with the fluctuation of the SVD values around that period, as shown in Fig. 3. The estimated structural break at January 1999 follows a series of devastating events in the second half of 1998 such as Russia's default, Brazil's currency crisis and the severe disruption of the LTCM's (Long-Term Capital Management L.P.) crisis to the US commercial paper markets. The structural break at June 2003 is followed by the month when the Fed announced the lowest federal funds rate over the past 40 years, which might signify the turning point of US economy from recovery to healthy expansion. The most recent structural break in Fig. 4 is at October 2008, what is the exact period that the financial crisis, starting with the housing bubble burst in 2007, hits its peak.

To test if the structural breaks suggested by Fig. 4 are significant, we consider the test of $H_0$:$\Lambda(t)$ are constant during January 1986–September 2009 versus $H_1$:$\Lambda(t)$ are piecewise constant with break points at July 1986, April 1991, January 1999, June 2003, and October 2008. The likelihood ratio test yields a $p$-value less than $10^{-10}$, rejecting the null hypothesis. We then use the suggested break times to segment the period January 1986–September 2009 into six periods: January 1986–June 1986, July 1986–March 1991, April 1991–December 1998, January 1999–May 2003, June 2003–September 2008, and October 2008–September 2009. We show the maximum likelihood estimates of one-year transition probability matrices under $H_1$ and $H_0$ in Tables 1 and 2 respectively. Note that Table 1 skips the transition probability matrix in January 1986–June 1986 because of the small number of rating transitions in the period, as shown in Fig. 1. We find that the transition probability matrices in the segmented periods are much different from the one in Table 2. For instance, the default probabilities from all non-default rating categories during October 2008–September 2009 in Table 1 are much larger than those in Table 2. Such difference is due to the assumption of structural breaks in our model, as time-homogeneous Markov processes inevitably smoothes out the time variations of transition matrices, especially those before and after structural breaks. Note that the transition matrices in the segmented periods can also be obtained by combining the estimates $\widehat{P}((t_{l-1}, t_l])$ in Section 3.1 via (13), we did not present the results as they are similar to the ones in Table 1.

### 3.3. Industrial effects

The literature has shown the existence of industrial effect in rating transition probability matrices; see, for example, Nickell et al. (2000), Bangia et al. (2002) and reference therein. We now add

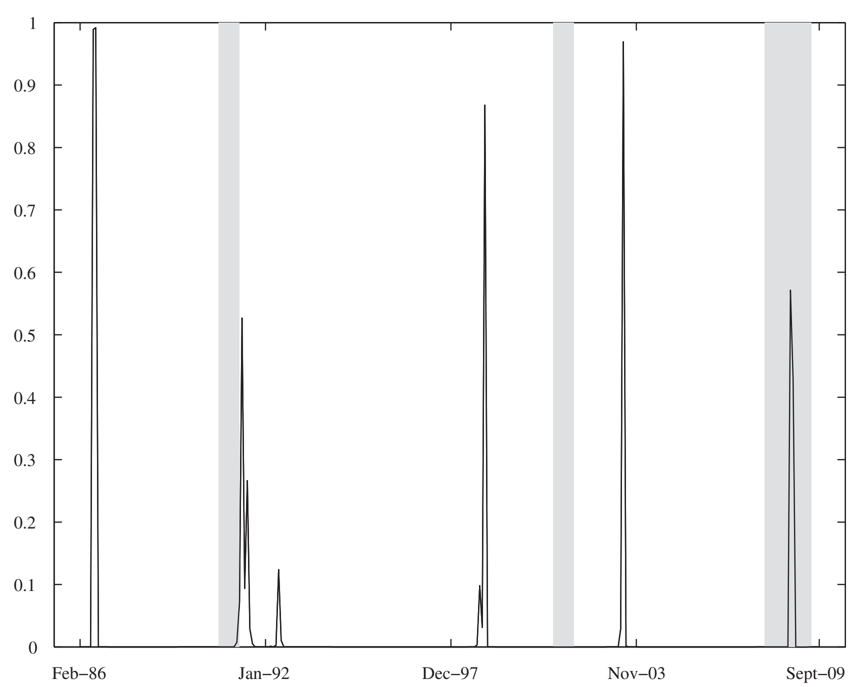


**Fig. 4.** Estimated probabilities of structural break at each month.

**Table 1**
Estimated one-year transition probability matrices over five periods.

| | $\mathcal{AAA}$ | $\mathcal{AA}$ | $\mathcal{A}$ | $\mathcal{BBB}$ | $\mathcal{BB}$ | $\mathcal{B}$ | $\mathcal{CCC}$ | $\mathcal{D}$ |
|---|---|---|---|---|---|---|---|---|
| July 1986–March 1991 | | | | | | | | |
| $\mathcal{AAA}$ | .9424 | .0390 | .0059 | .0044 | .0078 | .0004 | .0001 | .0000 |
| $\mathcal{AA}$ | .0103 | .8911 | .0865 | .0103 | .0007 | .0011 | .0001 | .0000 |
| $\mathcal{A}$ | .0010 | .0176 | .9063 | .0610 | .0080 | .0053 | .0007 | .0001 |
| $\mathcal{BBB}$ | .0000 | .0024 | .0535 | .8616 | .0615 | .0180 | .0026 | .0003 |
| $\mathcal{BB}$ | .0000 | .0007 | .0037 | .0541 | .8593 | .0689 | .0107 | .0025 |
| $\mathcal{B}$ | .0000 | .0005 | .0021 | .0031 | .0374 | .8872 | .0558 | .0138 |
| $\mathcal{CCC}$ | .0000 | .0000 | .0029 | .0005 | .0118 | .0471 | .8029 | .1347 |
| April 1991–December 1998 | | | | | | | | |
| $\mathcal{AAA}$ | .9515 | .0446 | .0038 | .0001 | .0000 | .0000 | .0000 | .0000 |
| $\mathcal{AA}$ | .0035 | .9379 | .0540 | .0044 | .0001 | .0000 | .0000 | .0000 |
| $\mathcal{A}$ | .0002 | .0091 | .9534 | .0335 | .0023 | .0012 | .0000 | .0002 |
| $\mathcal{BBB}$ | .0003 | .0014 | .0363 | .9300 | .0302 | .0018 | .0001 | .0000 |
| $\mathcal{BB}$ | .0005 | .0010 | .0039 | .0473 | .9062 | .0380 | .0025 | .0006 |
| $\mathcal{B}$ | .0002 | .0003 | .0018 | .0064 | .0473 | .9115 | .0299 | .0027 |
| $\mathcal{CCC}$ | .0000 | .0000 | .0011 | .0035 | .0200 | .0560 | .8961 | .0234 |
| January 1999–May 2003 | | | | | | | | |
| $\mathcal{AAA}$ | .9066 | .0756 | .0170 | .0008 | .0000 | .0000 | .0000 | .0000 |
| $\mathcal{AA}$ | .0009 | .8907 | .0982 | .0098 | .0003 | .0000 | .0000 | .0001 |
| $\mathcal{A}$ | .0003 | .0053 | .9273 | .0628 | .0027 | .0005 | .0003 | .0009 |
| $\mathcal{BBB}$ | .0005 | .0014 | .0169 | .9289 | .0458 | .0047 | .0009 | .0009 |
| $\mathcal{BB}$ | .0003 | .0003 | .0012 | .0199 | .9048 | .0621 | .0076 | .0038 |
| $\mathcal{B}$ | .0000 | .0002 | .0008 | .0024 | .0186 | .8873 | .0613 | .0293 |
| $\mathcal{CCC}$ | .0000 | .0000 | .0009 | .0019 | .0030 | .0234 | .7259 | .2448 |
| June 2003–September 2008 | | | | | | | | |
| $\mathcal{AAA}$ | .9647 | .0304 | .0005 | .0043 | .0001 | .0000 | .0000 | .0000 |
| $\mathcal{AA}$ | .0027 | .9688 | .0279 | .0005 | .0000 | .0000 | .0000 | .0000 |
| $\mathcal{A}$ | .0000 | .0081 | .9562 | .0334 | .0019 | .0001 | .0002 | .0000 |
| $\mathcal{BBB}$ | .0000 | .0009 | .0131 | .9542 | .0288 | .0025 | .0000 | .0003 |
| $\mathcal{BB}$ | .0000 | .0002 | .0009 | .0244 | .9281 | .0453 | .0008 | .0002 |
| $\mathcal{B}$ | .0000 | .0000 | .0012 | .0007 | .0347 | .9352 | .0261 | .0022 |
| $\mathcal{CCC}$ | .0000 | .0000 | .0008 | .0015 | .0020 | .0685 | .8607 | .0664 |
| October 2008–September 2009 | | | | | | | | |
| $\mathcal{AAA}$ | .8997 | .0945 | .0057 | .0001 | .0000 | .0000 | .0000 | .0000 |
| $\mathcal{AA}$ | .0000 | .8876 | .1081 | .0042 | .0001 | .0000 | .0000 | .0000 |
| $\mathcal{A}$ | .0000 | .0012 | .9242 | .0719 | .0026 | .0001 | .0000 | .0000 |
| $\mathcal{BBB}$ | .0000 | .0000 | .0099 | .9515 | .0350 | .0024 | .0009 | .0002 |
| $\mathcal{BB}$ | .0000 | .0000 | .0001 | .0146 | .8934 | .0801 | .0087 | .0031 |
| $\mathcal{B}$ | .0000 | .0000 | .0000 | .0001 | .0116 | .8876 | .0840 | .0167 |
| $\mathcal{CCC}$ | .0000 | .0000 | .0000 | .0000 | .0004 | .0535 | .7369 | .2092 |

**Table 2**
Estimated one-year time-homogeneous transition probability matrix (January 1986–September 2009).

| | $\mathcal{AAA}$ | $\mathcal{AA}$ | $\mathcal{A}$ | $\mathcal{BBB}$ | $\mathcal{BB}$ | $\mathcal{B}$ | $\mathcal{CCC}$ | $\mathcal{D}$ |
|---|---|---|---|---|---|---|---|---|
| $\mathcal{AAA}$ | .9416 | .0486 | .0060 | .0020 | .0017 | .0001 | .0000 | .0000 |
| $\mathcal{AA}$ | .0046 | .9222 | .0667 | .0057 | .0002 | .0006 | .0000 | .0000 |
| $\mathcal{A}$ | .0004 | .0094 | .9374 | .0477 | .0032 | .0013 | .0002 | .0004 |
| $\mathcal{BBB}$ | .0002 | .0012 | .0235 | .9329 | .0370 | .0043 | .0006 | .0004 |
| $\mathcal{BB}$ | .0002 | .0005 | .0018 | .0320 | .9075 | .0523 | .0043 | .0014 |
| $\mathcal{B}$ | .0001 | .0002 | .0012 | .0027 | .0328 | .9090 | .0434 | .0107 |
| $\mathcal{CCC}$ | .0000 | .0000 | .0010 | .0019 | .0075 | .0505 | .8205 | .1186 |

the industrial effect into our study and investigate how rating dynamics behave in different sectors. Among the 5185 firms used in our study, there are 44 firms containing no sector information and the rest 5144 firms are classified by the GICS system into 10 sectors (the numbers of firms and transition records in each sector are shown in parentheses): energy (396/852), materials (422/1039), manufacturing industrials (721/1612), consumer discretionary (1076/2519), consumer staples (304/622), health care (338/669), financials (823/1784), information technology (394/780), telecommunication (282/704), and utilities (385/998). Note that the number of transition records in each sector is too small to estimate a time-varying credit rating model, so we group these 10 sectors into two industry categories, so that the numbers of transition records in both categories are large enough for model estimation and comparable with each other. In particular, we group energy, materials, manufacturing, consumer discretionary and consumer staples as the first category, and the rest as the second. Besides the issue of sample size in each category, another reason of using such a rough way of grouping is that we try to put more traditional industries in the first category and procyclical industries in the second, so that the difference of rating dynamics between different categories can be shown more clearly. We then fit the stochastic structural break model to the rating records in each category. Similar to Section 3.1, we first estimate the hyperparameters and then compute the time-varying estimates for generator and rating transition matrices.

Let $\widehat{P}_1((t_{l-1},t_l])$ and $\widehat{P}_2((t_{l-1},t_l])$ denote the estimates of transition probability matrices in the first and second industry categories, respectively. Fig. 5 shows some estimated transition probabilities for each industry category and displays some interesting phenomena. Concerning the probability of staying at particular rating category, firms in the two industry categories show quite different behaviors. For instance, the probabilities of $\mathcal{AAA}$—$\mathcal{AAA}$ do not show much difference in both industry categories, but the probabilities of $\mathcal{BB}$—$\mathcal{BB}$ in the second category seem to have more fluctuations in the whole sample period. Furthermore, the second industry category shows larger default probabilities than the first. This phenomenon seems more significant after 2003. Furthermore, structural breaks (or jumps to larger values) of default probabilities happen earlier in the second industry category than the first.

To further evaluate these phenomena, we show in Fig. 6 the SVD metrics $M(\widehat{P}_1((t_{l-1},t_l]))$ and $M(\widehat{P}_2((t_{l-1},t_l]))$ defined by (18). The time-varying SVD metrics in Fig. 6 show some interesting results, compared with the metrics in Fig. 3. First, the metrics in Fig. 6 are stable in August 1986–June 1995, whereas during the same period the metrics in Fig. 3 show much fluctuation. This observation is difficult to interpret, and it seems to indicate that the stability of rating dynamics of all firms cannot be obtained by simple additive aggregation of different industries. Second, the SVD metrics of the first industry category are almost constant (0.209) in July 1999–May 2003, while in the same period, the SVD metrics of all firms are also almost constant (0.135). The SVD metric of the second category displays many fluctuations, indicating that the firms in the second industry category are more susceptible to economic shocks. This phenomenon is shown again during the most recent recession (December 2007–June 2009). The metrics of the second category jump from 0.099 to 0.144 right after the beginning of the recession, and then move to 0.192 at the peak of the recession in September and October 2008. While during the recession the metrics of the first industry category show only one jump (from 0.073 to 0.194) at October 2008. After the end of the recession (July 2009), the SVD metrics of the first category drop immediately to a much smaller value (0.06), while the values of the second category only drop to 0.168. This is consistent with the fact that the credit crisis hurts the finance and other service sectors before affecting the real economy. Furthermore, the SVD metrics of both categories in January 1986–February 1998 are close to those in June 2003–December 2007 before the financial crisis began, suggesting a normal regime exists for transition probabilities in both industry categories.

### 3.4. Out-of-sample forecasting

We now evaluate the forecast performance of the proposed model for each month $(t_L, t_{L+1})$ after the rating transition history $\mathcal{Y}_{(0,t_L)}$ is observed. In particular, at the end of month $t_L$, we first use the rating transition history $\mathcal{Y}_{(0,t_L)}$ as the training sample to estimate the hyperparameters by the EM algorithm in Appendix B, and then compute the one-month ahead prediction of the

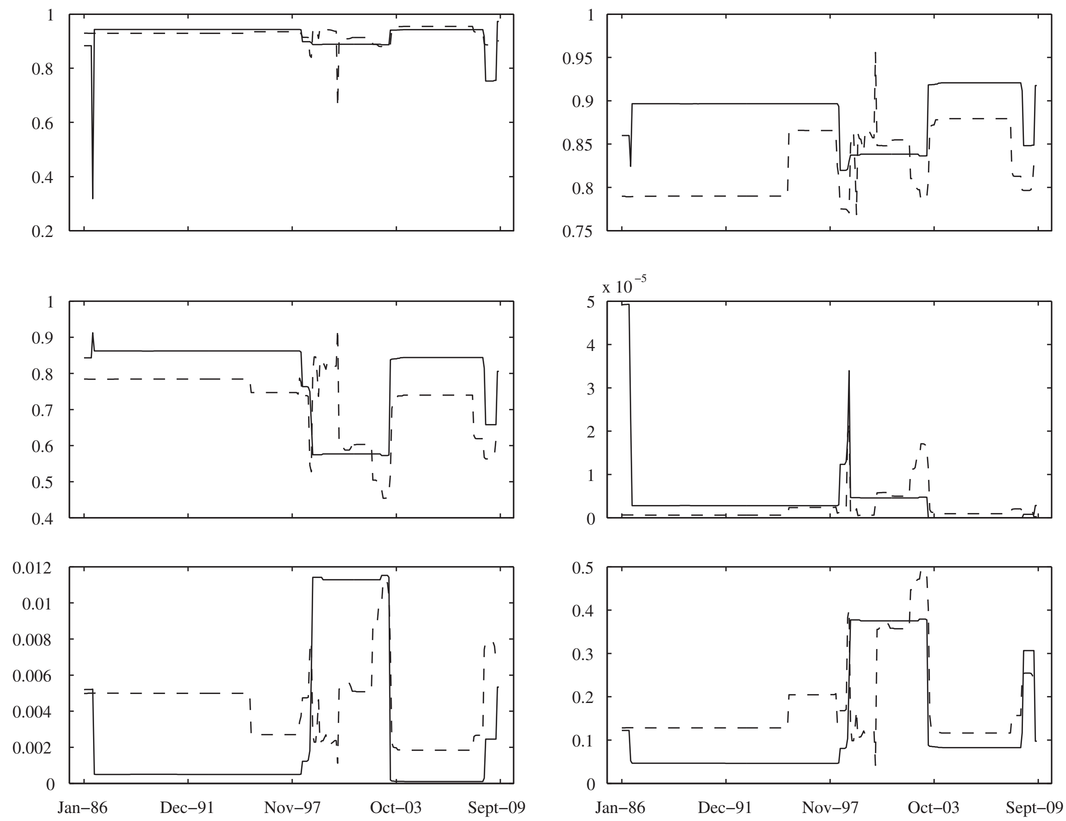


**Fig. 5.** Estimated transition probabilities $\mathcal{AAA}-\mathcal{AAA}, \mathcal{BB}-\mathcal{BB}, \mathcal{CCC}-\mathcal{CCC}, \mathcal{AAA}-\mathcal{D}, \mathcal{BB}-\mathcal{D}$ and $\mathcal{CCC}-\mathcal{D}$ (from left to right and top to bottom) for the first (solid line) and the second (dashed line) industry categories.

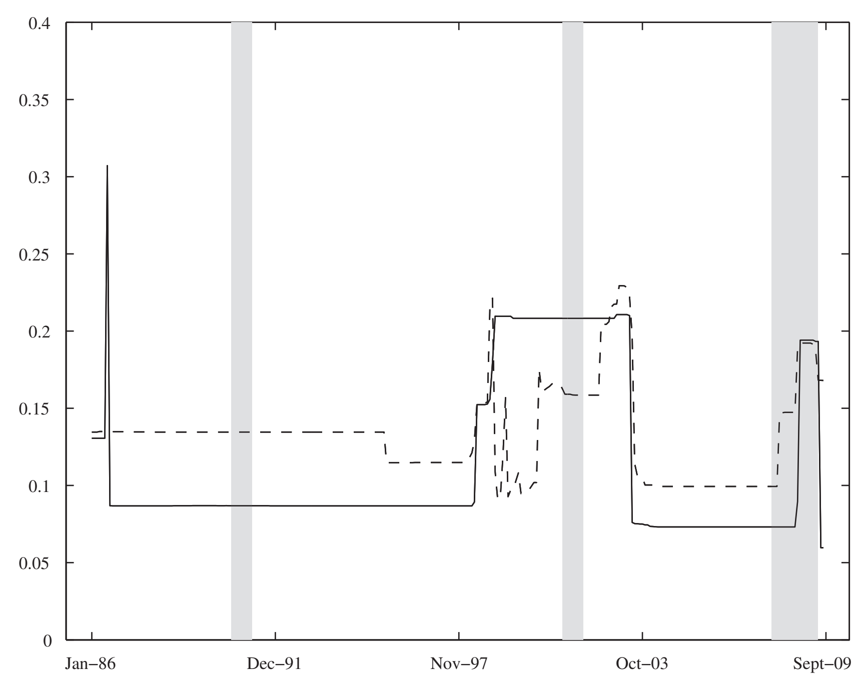


**Fig. 6.** The SVD metric of the estimated transition matrices for the first (solid line) and the second (dashed line) industry categories.

generator matrices $\widetilde{\Lambda}_L(t_{L+1})$ and the probability transition matrices $P((t_L, t_{L+1}))$ by (16) and (17), respectively. For comparison, we also use the time homogeneous Markov model to compute the one-month prediction of the generator matrices and the corresponding transition probability matrices. Specifically, under the assumption of the time homogeneous Markov model, we use one year historical transition records $\mathcal{Y}_{(t_{L-11}, t_L)}$ to compute the maximum likelihood estimate of the generator matrices, as stated in the second paragraph of Section 1. The reason of using samples with a moving window $(t_{L-11}, t_L)$ instead of the whole sample period $(0, t_L)$ is to avoid possible latent structural breaks in the long-run rating dynamics. Since the time-homogeneous Markov model

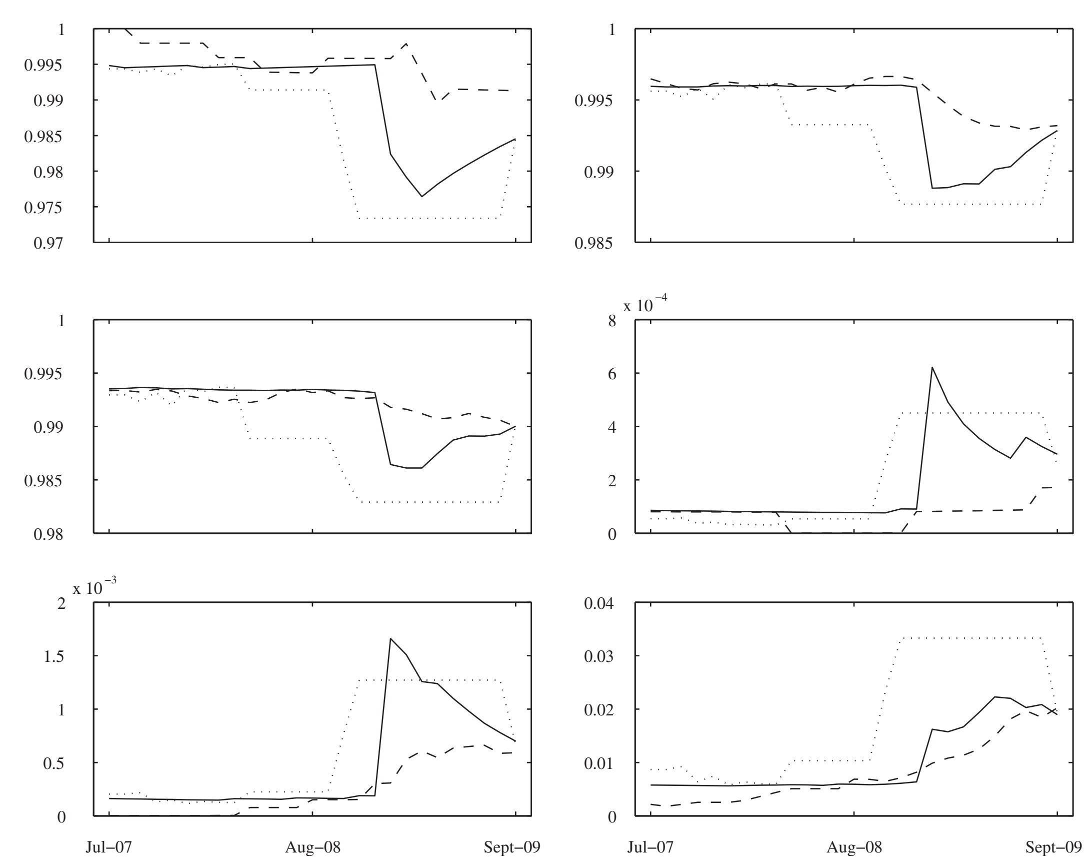


**Fig. 7.** One-month ahead predictions of transition probabilities $\mathcal{AAA}-\mathcal{AAA}$, $\mathcal{A}-\mathcal{A}$, $\mathcal{BB}-\mathcal{BB}$, $\mathcal{BB}-\mathcal{D}$, $\mathcal{B}-\mathcal{D}$ and $\mathcal{CCC}-\mathcal{D}$ (from left to right and top to bottom) using the stochastic structural break model (solid line) and the time-homogeneous Markov chain model (dashed line). Also shown are the estimated transition probabilities (dotted line) in Section 3.1.

assumes constant generator matrix, we then use the estimated generator as the generator matrix at the next period $(t_L, t_{L+1})$ and transform it to transition probability matrix.

We apply the above procedure to compute the one-month ahead prediction of transition matrices for $t_L$ = June 2007, ..., August 2009. Fig. 7 shows some one-month ahead predictions of transition probabilities obtained from our model and the time-homogeneous Markov model. Since we do not know the true dynamics of transition matrices during those periods, we use the in-sample estimates in Section 3.1 as a benchmark, which are shown as the dotted lines in Fig. 7. Comparing the out-of-sample predictions between our model and the time-homogeneous Markov model with moving windows, we find that the forecasts in both models show variations over time, but the predicted transition probabilities in the time-homogeneous Markov model with moving windows are much more smooth than those in our model. In particular, when the financial crisis starting from 2007 hit its peak in September and October 2008, the predicted transition probabilities in our model seem to catch this structural change much earlier than those in the time-homogeneous model.

## 4. Concluding remarks

We have developed a stochastic structural break model for rating transition matrices that generates piecewise constant rating generator matrices with unknown number, locations and magnitude of structural breaks following a compound Poisson process. The model allows exact computation, through explicit formulas given in Section 2, of the posterior distributions of the time-varying rating generator and transition probability matrices. From the posterior distribution of the generator matrices, given the rating history, one can make further inference on the rating transition probability matrices, the probability of structural break at each period and prediction of transition probability matrices in the presence of structural breaks. The hyperparameters in the model can be estimated by maximum likelihood or expectation–maximization algorithm in Appendix B.

The stochastic structural break model proposed here can be considered as an extension of the mixture model in Frydman and Schuermann (2008). Accordingly, the stochastic structural break model shares many statistical features with the mixture model. Furthermore, since our model assumes structural shifts among continuous states of generator matrices, the estimates for transition matrices become a mixture of continuous Markov chain with the number of mixture components much larger than two, which provides us further flexibility to address the concerns in Frydman and Schuermann (2008).

We find that the dynamics or estimated structural breaks in rating transitions are conditional on industry sectors, and the ratings of firms in service sectors are more vulnerable to economic shocks. We also show that the proposed model could incorporate possible structural breaks into the prediction of rating transition matrices, hence a better prediction can be obtained in the presence of structural changes or big economic shocks.

An interesting and challenging question following the proposed model is whether the structural break in rating transitions can be predicted or early detected based on historical information, which might provide a clue to prevent or mitigate the devastating impact of economic structural changes. Intuitively, the probability of structural breaks and historical information can be linked together via logistic regression. However, the identification of economic factors for structural breaks creates much difficulty in both statistical methodology and econometric theory, which should be investigated in further studies.

## Acknowledgement

The authors thank one anonymous referee for his/her helpful comment to improve the paper. This research is partially supported by the National Science Foundation.

## Appendix A. Proof of posterior distributions in Section 2.2

We first derive (4) or the posterior distribution of $\lambda^{(i,j)}_{t_{l-1},t_l}$ given $\mathcal{Y}_{(0,t_l]}$. Let $f(\cdot|\mathcal{Y}_{(0,t_l]})$ denote the density function of $\lambda^{(i,j)}_{t_{l-1},t_l}$ given $\mathcal{Y}_{(0,t_l]}$. Note that conditional on $I_l = 1$ or 0, we have

$$\begin{aligned} f(\lambda^{(i,j)}_{t_{l-1},t_l}|\mathcal{Y}_{(0,t_l]}) &\propto f(\lambda^{(i,j)}_{t_{l-1},t_l},\mathcal{Y}_{(t_{l-1},t_l]}|\mathcal{Y}_{(0,t_{l-1}]}) \\ &= pf(\lambda^{(i,j)}_{t_{l-1},t_l},\mathcal{Y}_{(t_{l-1},t_l]}|\mathcal{Y}_{(0,t_{l-1}]},I_l=1) \\ &\quad + (1-p)f(\lambda^{(i,j)}_{t_{l-1},t_l},\mathcal{Y}_{(t_{l-1},t_l]}|\mathcal{Y}_{(0,t_{l-1}]},I_l=0). \end{aligned} \tag{A.1}$$

Note that the first term in (A.1) can be written as

$$\begin{aligned} pf(\lambda^{(i,j)}_{t_{l-1},t_l},\mathcal{Y}_{(t_{l-1},t_l]}|\mathcal{Y}_{(0,t_{l-1}]},I_l=1) &= p^*_{l,l}f(\lambda^{(i,j)}_{t_{l-1},t_l}|\mathcal{Y}_{(0,t_l]},I_l=1) \\ &= p^*_{l,l}\mathrm{Gamma}(K^{(i,j)}_{t_{l-1},t_l}+\alpha_{ij},S^{(i)}_{t_{l-1},t_l}+\beta_i), \end{aligned} \tag{A.2}$$

in which

$$\begin{aligned} p^*_{l,l} &= pf(\mathcal{Y}_{(t_{l-1},t_l]}|\mathcal{Y}_{(0,t_{l-1}]},I_l=1) = p\int f(\mathcal{Y}_{(t_{l-1},t_l]}|\lambda^{(i,j)}_{t_{l-1},t_l})g(\lambda^{(i,j)}_{t_{l-1},t_l})d\lambda^{(i,j)}_{t_{l-1},t_l} \\ &= pf_{l,l}/f_{0,0}, \end{aligned}$$

and $f_{l,\,l}$ and $f_{0,0}$ are given in (6). The second term in (A.1) can be expanded as

$$\begin{aligned} &(1-p)f(\lambda^{(i,j)}_{t_{l-1},t_l},\mathcal{Y}_{(t_{l-1},t_l]}|\mathcal{Y}_{(0,t_{l-1}]},I_l=0) = (1-p)\sum_{m=1}^{l-1}P(R_{l-1} \\ &= t_{m-1}|\mathcal{Y}_{(0,t_{l-1}]},I_l=0)f(\lambda^{(i,j)}_{t_{l-1},t_l},\mathcal{Y}_{(t_{l-1},t_l]}|R_{l-1}=t_m,\mathcal{Y}_{(0,t_{l-1}]}, \\ &\quad I_l=0) = \sum_{m=1}^{l-1}p^*_{m,l-1}f(\lambda^{(i,j)}_{t_{l-1},t_l}|R_{l-1}=t_m,\mathcal{Y}_{(0,t_l]},I_l=0) \\ &= \sum_{m=1}^{l-1}p^*_{m,l-1}\mathrm{Gamma}(K^{(i,j)}_{t_{m-1},t_l}+\alpha_{ij},S^{(i,j)}_{t_{m-1},t_l}+\beta_i), \end{aligned} \tag{A.3}$$

in which

$$\begin{aligned} p^*_{m,l-1} &= (1-p)p_{m,l-1}f(\mathcal{Y}_{(t_{l-1},t_l]}|R_{l-1}=t_m,\mathcal{Y}_{(0,t_{l-1}]},I_l=0) \\ &= (1-p)p_{m,l-1}\frac{f(\mathcal{Y}_{t_{m-1},t_l},R_l=t_{m-1})}{f(\mathcal{Y}_{t_{m-1},t_{l-1}},R_{l-1}=t_{m-1})} \\ &= (1-p)p_{m,l-1}f_{m,l}/f_{m,l-1}. \end{aligned} \tag{A.4}$$

Hence combining (A.1)–(A.4) yields the posterior distribution (4).

We then derive (7), i.e., the posterior distribution of $\lambda^{(i,j)}_{t_{l-1},t_l}$ given $\mathcal{Y}_{(t_l,T)}$. We first reverse time and note that $\widetilde{I}_l = I_{L-l+1}$ are still i.i.d. Bernoulli and that the time-reversed Markov chain $\tilde{\lambda}^{(i,j)}_{t_{l-1},t_l} = \lambda^{(i,j)}_{t_{L-l},t_{L-l+1}}$ has the same transition probabilities as the Markov chain $\lambda^{(i,j)}_{t_{l-1},t_l}$. In other words, $\{\lambda^{(i,j)}_{t_{l-1},t_l}\}$ is a reversible Markov chain. Moreover, its stationary distribution is Gamma $(\alpha_{ij}, \beta_i)$. Using the similar argument for (4), we can prove the posterior distribution of $\lambda^{(i,j)}_{t_{l-1},t_l}$ given $\mathcal{Y}_{(t_l,T)}$ is given by (7).

We next apply the Bayes theorem to obtain (8), i.e., the posterior distribution of $\lambda^{(i,j)}_{t_{l-1},t_l}$ given $\mathcal{Y}_{(0,T)}$. Let $f(\cdot|\mathcal{Y}_{(t_l,T)})$ and $f(\cdot|\mathcal{Y}_{(0,T)})$ denote the density functions of $\lambda^{(i,j)}_{t_{l-1},t_l}$ given $\mathcal{Y}_{(t_l,T)}$ and $\mathcal{Y}_{(0,T)}$, respectively, and let $g$ denote the stationary density function of $\lambda^{(i,j)}_{t_{l-1},t_l}$ which is same as the prior Gamma distribution with parameters $\alpha_{ij}$ and $\beta_i$. Making use of the assumption that rating migrations of firms are conditionally independent in the period $(t_{l-1},t_l)$ given the generator matrix $\Lambda(t_l)$, we obtain

$$\begin{aligned} f(\lambda^{(i,j)}_{t_{l-1},t_l}|\mathcal{Y}_{(0,T)}) &\propto g(\lambda^{(i,j)}_{t_{l-1},t_l})f(\mathcal{Y}_{(0,T)}|\lambda^{(i,j)}_{t_{l-1},t_l}) \\ &\propto g(\lambda^{(i,j)}_{t_{l-1},t_l})f(\mathcal{Y}_{(0,t_l)}|\lambda^{(i,j)}_{t_{l-1},t_l})f(\mathcal{Y}_{(t_{l-1},t_l)}|\lambda^{(i,j)}_{t_{l-1},t_l}) \\ &\propto f(\mathcal{Y}_{(0,t_l)},\lambda^{(i,j)}_{t_{l-1},t_l})f(\mathcal{Y}_{(t_{l-1},t_l]},\lambda^{(i,j)}_{t_{l-1},t_l})/g(\lambda^{(i,j)}_{t_{l-1},t_l}) \\ &\propto f(\lambda^{(i,j)}_{t_{l-1},t_l}|\mathcal{Y}_{(0,t_l)})f(\lambda^{(i,j)}_{t_{l-1},t_l}|\mathcal{Y}_{(t_l,T)})/g(\lambda^{(i,j)}_{t_{l-1},t_l}), \end{aligned} \tag{A.5}$$

in which $f(\lambda^{(i,j)}_{t_{l-1},t_l}|\mathcal{Y}_{(0,t_l)})$ and $f(\lambda^{(i,j)}_{t_{l-1},t_l}|\mathcal{Y}_{(t_l,T)})$ are posterior distributions of $\lambda^{(i,j)}_{t_{l-1},t_l}$ given $\mathcal{Y}_{(0,t_l)}$ and $\mathcal{Y}_{(t_l,T)}$, respectively. Then we plug in (4), (7) and the Gamma prior (2) into (A.5), and notice the following fact

$$\begin{aligned} &\mathrm{Gamma}\Big(K^{(i,j)}_{t_{m-1},t_l}+\alpha_{ij},S^{(i,j)}_{t_{m-1},t_l}+\beta_i\Big) \\ &\quad\cdot\mathrm{Gamma}(K^{(i,j)}_{t_l,t_k}+\alpha_{ij},S^{(i,j)}_{t_l,t_k}+\beta_i)/\mathrm{Gamma}(\alpha_{ij},\beta_i) \\ &= \frac{f_{m,k}f_{0,0}}{f_{m,l}f_{l+1,k}}\mathrm{Gamma}\Big(K^{(i,j)}_{t_{m-1},t_k}+\alpha_{ij},S^{(i,j)}_{t_{m-1},t_k}+\beta_i\Big), \end{aligned}$$

we then obtain (8) or the posterior distribution of $\lambda^{(i,j)}_{t_{l-1},t_l}$ given $\mathcal{Y}_{(0,T)}$.

## Appendix B. Estimation of hyperparameters

From definition (5) of $p^*_{m,l}$, it follows that the conditional density function of $\mathcal{Y}_{(t_{l-1},t_l)}$ given $\mathcal{Y}_{(0,t_{l-1})}$ is

$$f(\mathcal{Y}_{t_{l-1},t_l}|\mathcal{Y}_{(0,t_{l-1})}) = \sum_{m=1}^{l}p^*_{m,l},$$

in which $p^*_{m,l}$ are functions of the hyperparameter vector $\Phi = \{p,\alpha_{ij},\beta_i|(i,j)\in\mathcal{K}\}$. Given $\Phi$ and the observed data $\mathcal{Y}_{(0,T)}$, the log-likelihood function is expressed as

$$l(\Phi) = \sum_{k=1}^{M}\log f(\mathcal{Y}_{t_{l-1},t_l}|\mathcal{Y}_{(0,t_{l-1})}) = \sum_{l=1}^{L}\log\left\{\sum_{m=1}^{l}p^*_{m,l}\right\}. \tag{B.1}$$

Maximizing (B.1) over $\Phi$ yields the maximum likelihood estimate $\widehat{\Phi}$.

Since $\Phi$ is a $(K(K-1)+1)$-dimensional vector and the functions $p^*_{m,l}$ have to be computed recursively for $1 \leqslant l \leqslant L$, direct maximization of (B.1) is computationally extensive. An alternative way is to use the EM algorithm which has the much simpler structure of the log likelihood $l_c(\Phi)$ of the complete data $\{K^{(i,j)}_{t_{l-1},t_l}, S^{(i)}_{t_{l-1},t_l}; (i,j)\in\mathcal{K}, 1\leqslant l\leqslant L\}$:

$$\begin{aligned} l_c(\Phi) &= \sum_{l=1}^{L}\sum_{i=1}^{K}\left\{\sum_{j\neq i}K^{(i,j)}_{t_{l-1},t_l}\log\lambda^{(i,j)}_{t_{l-1},t_l} - \left(\sum_{j\neq i}\lambda^{(i,j)}_{t_{l-1},t_l}+1-K\right)S^{(i)}_{t_{l-1},t_l}\right\} \\ &\quad + \sum_{l=1}^{L}\sum_{i=1}^{K}\left\{\sum_{j\neq i}(\alpha_{ij}-1)\log\lambda^{(i,j)}_{t_{l-1},t_l} - \left(\sum_{j\neq i}\lambda^{(i,j)}_{t_{l-1},t_l}\right)\beta_i \right. \\ &\quad \left. + \sum_{j\neq i}(\alpha_{ij}\log\beta_i - \log\Gamma(\alpha_{ij}))\right\}\mathbf{1}_{\{\Lambda(t_l)\neq\Lambda(t_{l-1})\}} \\ &\quad + \sum_{l=1}^{M}\{[\log(1-p)]\mathbf{1}_{\{\Lambda(t_l)=\Lambda(t_{l-1})\}} + (\log p)\mathbf{1}_{\{\Lambda(t_l)\neq\Lambda(t_{l-1})\}}\}. \end{aligned} \tag{B.2}$$

Since $l_c(\Phi)$ decomposes into gamma and binomial components, the E-step of the EM algorithm involves $E(\lambda^{(i,j)}_{t_{l-1},t_l}|\mathcal{Y}_{(0,T)}), E(\log\lambda^{(i,j)}_{t_{l-1},t_l}|\mathcal{Y}_{(0,T)})$ and the conditional probabilities $P(\Lambda(t_l)\neq\Lambda(t_{l-1})|\mathcal{Y}_{(0,T)}) = P(I_l = 1|\mathcal{Y}_{(0,T)})$. In view of (B.2), the M-step of the EM algorithm involves the updating formulas

$$\hat{p}_{\text{new}} = \sum_{l=1}^{L} P(\Lambda(t_l) \neq \Lambda(t_{l-1})|\mathcal{Y}_{(0,T)})/L,$$

$$\hat{\beta}_{i,\text{new}} = \frac{\sum_{l=1}^{L}\sum_{j\neq i}\hat{\alpha}_{ij,\text{old}}P(\Lambda(t_l) \neq \Lambda(t_{l-1})|\mathcal{Y}_{(0,T)})}{\sum_{l=1}^{L}E\left[\left(\sum_{j\neq i}\lambda^{(i,j)}_{t_{l-1},t_l}\right)\mathbf{1}_{\{\Lambda(t_l)\neq\Lambda(t_{l-1})\}}|\mathcal{Y}_{(0,T)}\right]}$$

$$\frac{\Gamma'(\hat{\alpha}_{ij,\text{new}})}{\Gamma(\hat{\alpha}_{ij,\text{new}})} = \Psi(\hat{\alpha}_{ij,\text{new}}) = \frac{\sum_{l=1}^{L}E\left(\log\lambda^{(i,j)}_{t_{l-1},t_l}\mathbf{1}_{\{\Lambda(t_l)\neq\Lambda(t_{l-1})\}}|\mathcal{Y}_{(0,T)}\right)}{\sum_{l=1}^{L}P(\Lambda(t_l)\neq\Lambda(t_{l-1})|\mathcal{Y}_{(0,T)}) + \log(\widehat{\beta}_{i,\text{old}})}, \tag{B.3}$$

in which $\Psi(\cdot)$ is the Digamma function, and the equation $\Psi(\cdot) = a$ can be solved numerically by grid search. Using the estimation procedure in Section 2.1, we can show that

$$E\left(\lambda^{(i,j)}_{t_{l-1},t_l}\mathbf{1}_{\{\Lambda(t_l)\neq\Lambda(t_{l-1})\}}|\mathcal{Y}_{(0,T)}\right) = \sum_{l\leqslant k\leqslant L}\pi_{lkl}\frac{K^{(i,j)}_{t_l,t_k}+\alpha_{ij}}{S^{(i)}_{t_l,t_k}+\beta_i},$$

which can be used to compute $\alpha_{ij,\text{new}}$ and $\beta_{i,\text{ new}}$ in (B.3). The iteration scheme (B.3) is carried out until convergence or until some prescribed upper bound on the number of iterations is reached.

## References


Acharya, V., Das, S., Sundaram, R., 2006. Arbitrage-free pricing of credit derivatives with rating transitions. Financial Analysts Journal 58, 28–44.

Altman, E., 1998. The importance and subtlety of credit rating migration. Journal of Banking and Finance 22, 1231–1247.

Altman, E., Rijken, H., 2004. How rating agencies achieve rating stability. Journal of Banking and Finance 28, 2679–2714.

Andersen, P., Borgan, O., Grill, R., Keiding, N., 1995. Statistical Models Based on Counting Processes. Springer, New York.

Bangia, A., Diebold, F., Kronimus, A., Schagen, C., Schuermann, T., 2002. Ratings migration and the business cycle, with application to credit portfolio stress testing. Journal of Banking and Finance 26, 445–474.

Blume, M., Lim, F., MacKinlay, A., 1998. The declining credit quality of us corporate debt: myth or reality? Journal of Finance 53, 1389–1413.

Bruche, M., González-Aguado, C., 2010. Recovery rates, default probabilities, and the credit cycle. Journal of Banking and Finance 34, 754–764.

Cantor, R., Mann, C., 2003a. Measuring the Performance of Corporate Bond Ratings. Moody's Special Comment.

Cantor, R., Mann, C., 2003b. Are Corporate Bond Ratings Procyclical. Moody's Special Comment.

Christensen, J., Hansen, E., Lando, D., 2004. Confidence sets for continuous-time rating transition probabilities. Journal of Banking and Finance 28, 2575–2602.

Crowder, M., Davis, M., Giampieri, G., Davis, M., 2005. Analysis of default data using hidden markov models. Quantitative Finance 5, 27–34.

Das, S., Freed, L., Geng, G., Kapadia, N., 2006. Correlated default risk. Journal of Fixed Income 16, 7–32.

Duffie, D., Eckner, A., Horel, G., Saita, L., 2009. Frailty correlated default. Journal of Finance 64, 2089–2123.

Egloff, D., Leippold, M., Vanini, P., 2007. A simple model of credit contagion. Journal of Banking and Finance 31, 2475–2492.

Fons, J., 2002. Understanding Moody's Corporate Bond Ratings and Rating Process. Moody's Special Comment.

Frey, R., McNeil, A., 2007. Dependence defaults in models of portfolio credit risk. Journal of Risk 6, 59–92.

Frydman, H., Schuermann, T., 2008. Credit rating dynamics and markov mixture models. Journal of Banking and Finance 32, 1062–1075.

Gordy, M., Howells, B., 2006. Procyclicality in basel II: can we treat the disease without killing the patient. Journal of Financial Intermediation 15, 395–417.

Jafry, Y., Schuermann, T., 2004. Measurement, estimation and comparison of credit migration matrices. Journal of Banking and Finance 28, 2603–2639.

Jarrow, R., Lando, D., Turnbull, S., 1997. A markov model for the term structure of credit risk spreads. Review of Financial Studies 10, 481–523.

Jarrow, R., Lando, D., Turnbull, S., 1998. A markov chain model for valuing credit risk derivatives. Journal of Derivatives 6, 97–108.

Küchler, U., Sørensen, M., 1997. Exponential Families of Stochastic Processes. Springer-Verlag, New York.

Lando, D., 2000. Some elements of rating-based credit risk modeling. In: Jegadeesh, N., Tuckman, B. (Eds.), Advanced Fixed-Income Valuation Tools. John Wiley, Chichester, pp. 193–215.

Lando, D., Skødeberg, T., 2002. Analyzing rating transitions and rating drift with continuous observations. Journal of Banking and Finance 26, 423–444.

Liao, H., Chen, T., Lu, C., 2009. Bank credit risk and structural credit models: agency and information asymmetry perspectives. Journal of Banking and Finance 33, 1520–1530.

Nickell, P., Perraudin, W., Varotto, S., 2000. Stability of rating transitions. Journal of Banking and Finance 24, 203–227.

Treacy, W., Carey, M., 2000. Credit risk rating systems at large us banks. Journal of Banking and Finance 24, 167–201.

Tsaig, Y., Levy, A., Wang, S., 2011. Analyzing the impact of credit migration in a portfolio setting. Journal of Banking and Finance 35, 3145–3157.