Entropy 2020, 22, 1072; doi:10.3390/e22101072 www.mdpi.com/journal/entropy

Article
Statistical Surveillance of Structural Breaks in Credit Rating Dynamics

Haipeng Xing 1,*, Ke Wang 2, Zhi Li 1 and Ying Chen 3,†

1 Department of Applied Mathematics and Statistics, State University of New York, Stony Brook, NY 11794, USA; zhi.li.3@stonybrook.edu
2 Quantitative Research, J.P. Morgan Chase, 383 Madison Ave, NY, 10179, USA; wkkate6@gmail.com
3 Point72 Asset Management, 55 Hudson Yards, New York, NY 10001, USA; yingemma.chen@gmail.com
* Correspondence: haipeng.xing@stonybrook.edu
† Ying Chen contributed to this article in her personal capacity. The information, views, and opinions expressed herein are solely her own and do not necessarily represent the views of Point72. Point72 is not responsible for, and did not verify for accuracy, any of the information contained herein.

Received: 22 August 2020; Accepted: 22 September 2020; Published: 24 September 2020

Abstract: The 2007–2008 financial crisis had severe consequences on the global economy and an intriguing question related to the crisis is whether structural breaks in the credit market can be detected. To address this issue, we chose firms’ credit rating transition dynamics as a proxy of the credit market and discuss how statistical process control tools can be used to surveil structural breaks in firms’ rating transition dynamics. After reviewing some commonly used Markovian models for firms’ rating transition dynamics, we present several surveillance rules for detecting changes in generators of firms’ rating migration matrices, including the likelihood ratio rule, the generalized likelihood ratio rule, the extended Shiryaev’s detection rule, and a Bayesian detection rule for piecewise homogeneous Markovian models. The effectiveness of these rules was analyzed on the basis of Monte Carlo simulations. We also provide a real example that used the surveillance rules to analyze and detect structural breaks in the monthly credit rating migration of U.S. firms from January 1986 to February 2017.

Keywords: credit rating; financial surveillance; rating transition dynamics; statistical process control; structural break

1. Introduction

The 2007–2008 global financial crisis, originally triggered by the U.S. subprime mortgage crisis in 2007 and quickly spread over the U.S. and the rest of the world via the U.S. and international banking systems, caused severe economic, political, and social consequences over the world. The crisis not only triggered the European Debt Crisis that began in the European Union at the end of 2009, but also led to a global economic recession which was considered the most severe one since the Great Depression in 1930s. In addition to this, the COVID-19 pandemic declared by the World Health Organization on March 2020 placed the global financial system under further strain and triggered a global economic downturn [1]. These events raise some important questions for economists, financial practitioners, and government regulators, and one of them is whether quantitative tools can be developed to surveil the occurrence of crises or structural breaks in the credit market or its sub-markets.

The question is difficult to answer directly because, in contrast to most quickest change detection problems in which the system dynamics are known, dynamical models have not been found which can extract and aggregate credit information at the microeconomic or macroeconomic level, and in the meanwhile, describe the general movement of the credit market over time. Hence, in order to provide
possible insights into the question of detecting structural changes in the credit market, this paper proposes to use firms’ rating transition dynamics as a proxy of the credit market and presents several surveillance rules for structural breaks in firms’ credit rating dynamics.

The reason for choosing firms’ credit ratings as a proxy of the credit market is the following. First, as an opinion on the relative ability of an entity to meet a financial commitment provided by credit rating agencies (CRAs), credit ratings are an information good at the intersection between the demand and supply for capital, which mitigates the information asymmetry in the credit market and hence enhances capital market efficiency and transparency [2]. Second, besides several types of information revealed by firms’ credit rating transitions, such as the conflicts of interest between investors and CRAs [3,4], the effect of a CRA’s reputational concern on rating quality [5,6], the interaction between the business cycle and firms’ incentives [7], and so forth, it has been shown that another type of information, structural breaks in the credit market, can be extracted from firms’ credit rating transitions [8–10]. Third, credit rating models play an important role in modern credit risk management. For example, in the pricing of bonds and credit, derivative, structural, and reduced-form models are used to model firms’ credit risk and credit ratings [10–14]; in bank regulation, banks are required to construct credit rating models to stress test their portfolios and evaluate evidence of rating transitions in external ratings [15,16].

By choosing firms’ rating transition dynamics as a proxy of the credit market, this paper makes the following contributions to the literature of financial surveillance. First, to the best of our knowledge, this is the first study that extends statistical process control methods to study the problem of monitoring financial stability at the regulatory level. Specifically, we combine rules of change-point detection with unknown pre- and after-change distributions with the state-of-the-art models of credit rating dynamics and propose different types of surveillance rules for structural breaks in rating transition dynamics. Second, instead of using a single or aggregated time series that is summarized from the credit market, the surveillance methods proposed in the paper are based on credit behavior of a collection of firms, which not only provides each firm’s credit quality at the microeconomic level but also aggregates the structural break information over firms for the surveillance purpose. Third, the empirical study in the paper compares the proposed surveillance rules for structural breaks in firms’ rating transition dynamics and demonstrates the effectiveness of these rules in the surveillance of the credit market.

We next elaborate the proposed methods and empirical study as follows. Since firms’ rating transition dynamics have been chosen as a proxy of the credit market, one can link the variability of rating transition dynamics with the stability of the credit market. Before presenting surveillance rules for firms’ credit rating dynamics, we review models that characterize firms’ credit rating migrations as discrete- or continuous-time Markov chains in Section 2. A typical model of these assumes that firms’ rating migrations follow discrete-time homogeneous Markov chain, which is characterized by a probability transition matrix. Though it is usually used by CRAs to publish firms’ yearly rating transition matrices, the discrete-time model is criticized due to the availability of rating data and the need for forecasting rating transition matrices over any time horizon. Therefore, firms’ rating transitions are also modeled in continuous time and all surveillance rules discussed in this paper are also based on continuous-time settings. We consider two types of continuous-time models in the paper. The first one assumes that firms’ rating transitions follow continuous-time homogeneous Markov chains that are characterized by a generator matrix of probability transition matrices. The second model is the stochastic structural break (SSB) [8], which assumes that firms’ rating transitions follow a continuous-time piecewise homogeneous Markov chain characterized by piecewise constant generator matrices. The SSB model assumes the existence of multiple unknown structural breaks in firms’ rating transition dynamics and provides attractive online estimates of generator matrices and probabilities of structural breaks.

We then present several surveillance rules based on continuous-time models of firms’ rating transition dynamics. Since both the pre- and after-change distributions of firms’ rating generator matrices are unknown, rules of detecting structural breaks in firms’ rating dynamics involve steps
of estimating pre- and after-change distributions. In particular, we consider the following rules. 

The first rule is based on the likelihood ratio tests using historical and current rating records, which can be further specified as maximized likelihood ratios and mixture likelihood ratios. The second rule involves the generalized likelihood ratio test when firms’ rating transitions follow continuous-time Markov chains with different generators before and after the break point. The third rule applies the extended Shiryaev’s detection rule [17] to continuous-time Markov chains with different pre- and after-change generators. The fourth rule is a Bayesian detection rule for exponential families with unknown multiple change-points [18] and is based on estimated probabilities of the most recent change-point. We apply the fourth rule to the SSB model for firms’ rating transition dynamics and obtain a surveillance rule for structural breaks.

As general properties of these surveillance rules have been discussed in the literature, we focus on the effectiveness of these rules for detecting changes in firms’ rating transition dynamics. Section 4 describes our simulation studies for detecting structural breaks in four-state continuous-time Markov chains with break points in generator matrices. In particular, we discuss the false alarm probabilities and conditional detection delay of the four surveillance rules in different scenarios. To further demonstrate the performance of these surveillance rules in detecting structural breaks in firms’ rating transition dynamics, Section 5 walks through a real example in which structural breaks in the U.S. firms’ monthly rating transitions are detected. In particular, after a preliminary analysis of the U.S. firms’ rating transitions during 1986–2016, we applied the surveillance rules discussed in Sections 3 and 4 and detect sequentially structural breaks in firms’ rating transition dynamics from January 1997 to February 2017.

The result in Section 5 shows that surveillance rules can help us detect structural breaks in firms’ rating transition dynamics, and some of detected break points can be traced back to economic events that might cause structural changes of the credit market. However, structural breaks detected by different surveillance rules can be very different, indicating that further studies need to be carried out. Section 6 concludes the paper with some remarks.

2. Models of Credit Rating Dynamics

In the study of credit rating dynamics, it is convenient to assume that a firm’s credit ratings follow a time-homogeneous Markov chain. Suppose that there are $K$ rating categories which are ordered as $\{1, \ldots, K\}$ from excellent to very bad credit quality. In particular, we denote category $K$ a “default” state of the firm which is an absorbing state.

2.1. Discrete-Time Homogeneous Rating Transition Matrices

Consider first a discrete-time setting for firms’ rating transitions. Denote by $\eta_t \in \{1, \ldots, K\}$ the rating of firm $l$ at time $t$. Suppose that $\{\eta_t\}_{t \geq 0}$ follows a discrete-time $K$-state Markov chain, which is characterized by a probability transition matrix $P = (p_{ij})_{1 \leq i, j \leq K}$ and $p_{ij} = \mathbb{P}(\eta_t = j | \eta_{t-1} = i)$ is the probability that firm $l$ migrates from rating category $i$ at time $t-1$ to rating category $j$ at time $t$. Then for $i = 1, \ldots, K, \sum_{j=1}^{K} p_{ij} = 1$, and $p_{KK} = 1$ since $K$ is a default state. The probability transition matrix $P$ is a one-period (usually, one-year) transition matrix, for rating transitions over multiple periods; the transition probability matrix can be obtained by matrix multiplication. In particular, the firm $l$’s $(t-s)$-period ($t > s \geq 0$) probability transition matrix is expressed as $P^{t-s}$.

Suppose that all firms’ rating transitions follow a discrete-time $K$-state Markov chain with one-year probability transition matrix $P = (p_{ij})_{1 \leq i, j \leq K}$. Assume that there are $N_i$ firms in a given rating category $i$ at the time $t-1$ and $N_{ij}$ firms migrating from $i$ at time $t-1$ to $j$ ($j \neq i$) at time $t$; the log-likelihood of observed ratings from $t-1$ to $t$ is

$$
\log L(P | Y_{t-1:t}) = \sum_{i=1}^{K} \left[ \sum_{j \neq i} N_{ij} \log p_{ij} + (N_i - \sum_{j \neq i} N_{ij}) \log(1 - \sum_{j \neq i} p_{ij}) \right].
$$

(1)

Then the maximum likelihood estimate for $P$ is given by $\hat{p}_{ij} = \frac{N_{ij}}{N_i}$ for $j \neq i$ and $\hat{p}_{ii} = 1 - \sum_{j \neq i} \hat{p}_{ij}$ for $i = 1, \ldots, K$.

In the literature of measuring changes of probability transition matrices for general Markov chains, mobility indices of probability transition matrices are usually used. Reference [19] proposed indices for Markov matrices using eigenvalues and determinants. Reference [20] presented a set of criteria by which the performance of a proposed metric for arbitrary transition matrices should be judged. In particular, Reference [20] summarized criteria in three distinct aspects. The persistence criteria require that a metric is consistent with some simple, intuitively appealing interpretations of the transition matrix $P$; the convergence criteria specify that a metric needs to establish an ordering among transition matrices $P$ that is consistent with the rate at which the multiperiod transition matrices $P^s$ ($s > 0$ is an integer) converge to a limiting transition matrix; and the temporal aggregation criteria remove the influence of the length of the basic time period on comparisons of mobility. Reference [21] argued that, since credit migration matrices are generally diagonally dominant and then the decay rates towards its steady-state are generally slow, the temporal aggregation criteria can be neglected given that credit migration matrices are usually evaluated for fixed periods considerably shorter than the decay time of the system. Instead, Reference [21] highlighted the importance of an additional persistence, distribution discriminatory criterion, which can discriminate matrices with the same row-wise probabilities of change but different distributions across each row, and then proposed a scalar SVD metric defined as the average of the singular values of the mobility matrix for a credit migration matrix $P$:

$$M(P) = \frac{1}{K} \sum_{i=1}^{K} \sqrt{\lambda_i(P-I')(P-I)},$$

in which $I$ is a $K \times K$ identity matrix and $\lambda_i(\cdot)$ is the $i$th eigenvalue of the matrix. Note that $P-I$ represents the magnitude of the implied mobility of migration matrix. Reference [8] later used the SVD metric to show time variations in estimated credit migration matrices.

One concern regarding the discrete-time method of a estimating credit rating transition matrix is that the estimated probability of rare events can be zero due to the observability of the rare events. Besides, the method does not provide transition probabilities in fractional periods ([22], Chapter 6). Since the continuous-time history of rating transitions is usually accessed by the rating agencies or banks using an internal rating system, a continuous-time homogeneous Markov framework is usually assumed for the rating process.

2.2. Continuous-Time Homogeneous Markov Transition Matrices

Suppose that 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 $(0,t)$. Similarly 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 obtained through its generator matrix $\Lambda$ under the assumption of time homogeneity. That is, for any time $t > 0$,

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

in which $\Lambda = (\lambda^{(ij)})$ satisfies $\lambda^{(ij)} = -\sum_{j \neq i} \lambda^{(ij)}$ for $1 \leq i \leq K$, and $\lambda^{(ij)} \geq 0$ for $1 \leq i \neq j \leq K$. Note that the elements of the last row in the generator matrix $\Lambda$ represent the rating transitions from the default category to others and are usually considered as zeros.

Suppose that there are $n$ realizations of a Markov chain with the generator matrix $\Lambda$. Denote by, for the period $(0,t)$, $K_{ij}^{(i)}$ the number of transitions from category $i$ to category $j$, $S_{ij}^{(i)}$ the amount of time
spent in category $i$, and $Y_{0,t}$ the observed rating transitions over the period $(0,t)$. The log-likelihood of $Y_{0,t}$ given the generator $\Lambda$ is expressed as

$$
\log L(\Lambda|Y_{0,t}) = \sum_{i=1}^{K} \left[ \sum_{j\neq i}^{K} K_{0,t}^{(i,j)} \log \lambda^{(i,j)} - \left( \sum_{j\neq i}^{K} \lambda^{(i,j)} + 1 - K \right) S_{0,t}^{(i,j)} \right];
$$

(4)

see [23]. Then the maximum likelihood estimates of elements of $\Lambda$ are given by $\tilde{\lambda}^{(i,j)} = K_{0,t}^{(i,j)}/S_{0,t}^{(i,j)}$.

One may also assume a prior distribution for the off-diagonal elements $\lambda^{(i,j)}$ and obtain a Bayesian estimate of the generator matrix. In particular, suppose that the off-diagonal elements $\lambda^{(i,j)}$ follow independently a gamma($a_{ij}, \beta_{i}$) prior distribution with the density function

$$
g(\lambda^{(i,j)}) = \frac{\beta_{i}^{a_{ij}}}{\Gamma(a_{ij})} [\lambda^{(i,j)}]^{a_{ij}-1} \exp(-\lambda^{(i,j)}\beta_{i}), \quad (i, j) \in K,
$$

(5)

in which $K = \{(i, j)| i \neq j, 1 \leq i \leq K - 1, 1 \leq j \leq K\}$. Combining the conjugate gamma prior (5) with the likelihood function (4) yields the posterior distribution of $\lambda^{(i,j)}$ given $Y_{0,t}$, which is gamma($K_{0,t}^{(i,j)} + a_{ij}, S_{0,t}^{(i,j)} + \beta_{i}$). Then the element $\lambda^{(i,j)}$ can be estimated by the posterior mean of the gamma distribution, i.e., $\tilde{\lambda}^{(i,j)} = (K_{0,t}^{(i,j)} + a_{ij})/(S_{0,t}^{(i,j)} + \beta_{i})$.

2.3. Continuous-Time Piecewise-Homogeneous Transition Matrices

Some empirical studies show that the observed firms’ rating transitions are not stationary over time, so we may move forward and assume that a firm’s rating transition process follows a $K$-state non-homogeneous continuous time Markov process. In general, such a process can be 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 a firm starting in state $i$ at time $s$ is in state $j$ at time $t$. Suppose that there are $m$ rating transitions observed over the period $(s,t)$. For a transition time $t_k$ in $(s,t)$, denote by $\Delta N_{ij}(t_k)$ the number of transitions observed from state $i$ to state $j$ at time $t_k$, and $Y_{i}(t_k)$ the number of firms in state $i$ right before time $t_k$. The non-homogeneous transition matrix $P(s,t)$ can be consistently estimated by the product-limit estimator

$$
\hat{P}(s,t) = \prod_{k=1}^{m} (I + \Delta \hat{A}(t_k)),
$$

(6)

in which

$$
\Delta \hat{A}(t_k) = \begin{pmatrix}
-\Delta N_{11}(t_k) & \Delta N_{21}(t_k) & \Delta N_{31}(t_k) & \cdots & \Delta N_{K1}(t_k) \\
\Delta N_{12}(t_k) & -\Delta N_{22}(t_k) & \Delta N_{32}(t_k) & \cdots & \Delta N_{K2}(t_k) \\
\Delta N_{13}(t_k) & \Delta N_{23}(t_k) & -\Delta N_{33}(t_k) & \cdots & \Delta N_{K3}(t_k) \\
\vdots & \vdots & \vdots & \ddots & \vdots \\
\Delta N_{1K}(t_k) & \Delta N_{2K}(t_k) & \Delta N_{3K}(t_k) & \cdots & -\Delta N_{KK}(t_k)
\end{pmatrix};
$$

(7)

see [24]. In the matrix above, the $k$th diagonal element counts the fraction of the exposed firms $Y_{i}(t_k)$ leaving the state at time $t_k$, and the $(ij)$’th off-diagonal element counts the fraction of transitions from the category $i$ to the category $j$ in the number of exposed firms at time $t_k$. 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 $\Delta \hat{A}(t_k)$ is zero because the $K$th state (i.e., default state) is absorbent.
The mixture weight can be calculated recursively by

\[ p_\Lambda = \frac{\lambda_{ij}^{(s)}}{\sum_{m=1}^{k} p_{m,l}^{*(s)}} \]

The second assumption of the model is that the piecewise constant generator matrices \( \Lambda \) are independent and identically distributed (i.i.d.) random matrices and their off-diagonal element \( \lambda_{ij} \) follows a gamma \((\alpha_{ij}, \beta_{ij})\) prior distribution whose density function is given by (5).

For the period \((s, t)\), denote by \( k^{(s,j)} \) the number of transitions from category \( i \) to category \( j \), \( s_{ij} \) the amount of time spent in category \( i \), \( \lambda_{ij} \) the \( ij \)th entry in the generator \( \Lambda \), and \( Y_{s,t} \) the observed rating transitions over the period \((s, t)\). The SSB model suggests a filtering approach to estimate the piecewise constant generator matrices \( \Lambda(t) \). In particular, they partitioned the sample period \((0, T)\) as \( 0 = t_0 < t_1 < \cdots < t_L = T \) and assumed that structural breaks can only occur at the times \( t_1, \ldots, t_L \). Let \( l = 1 \) and \( l = N_{\Lambda}(t_{l-}) - N_{\Lambda}(t_{l-1}) \) for \( l = 2, \ldots, L \); then \( l_1 \) indicates if \( \Lambda(t) \) are the same during the periods \((t_{l-2}, t_{l-1})\) and \((t_{l-1}, t_l)\) and \( \{l_i\} \) are a sequence of i.i.d. Bernoulli random variables with success probability \( p = 1 - \exp(-\eta T/L) \). To derive the posterior distribution of \( \Lambda(t) = (\lambda_{ij})_{i,j} \) given \( Y_{(0,T)} \) they further computed the probability that the most recent break time prior to \( t_{l-1} \) is \( t_{l-1} \), i.e., \( p_{m,l} = P(R_l = t_{m-1} | Y_{(0,t_{l-1})}) \), where \( R_l = \max \{t_{m-1} | t_{m-1} = 1, m \leq l \} \) is the most recent break time up to time \( t_{l-1} \), and showed that the posterior distribution of \( \lambda_{ij} \) given \( Y_{(0,t_{l-1})} \) is a mixture of gamma distributions,

\[
\lambda_{ij} \mid Y_{(0,t_{l-1})} \sim \sum_{m=1}^{l} \Gamma(K_{m,l-1}^{(s,i)} + \alpha_{ij}, S_{m-1,l}^{(s)} + \beta_{ij})
\]

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

\[
p^{*(s)}_{m,l} = \begin{cases} p f_{i,j} / f_{0,0} & m = l, \\ (1-p) p_{m,l-1} f_{m,l} / f_{m,l-1} & m < l, \end{cases}
\]

and the terms \( f_{m,l} \) and \( f_{0,0} \) are expressed as

\[
f_{m,l} = \prod_{i,l \in \mathcal{K}} \Gamma(K_{m-1,l}^{(s,i)} + \alpha_{ij}) / (S_{m-1,l}^{(s)} + \beta_{ij}) \Gamma(K_{m-1,l}^{(s,i)} + \alpha_{ij}), \\ f_{0,0} = \prod_{i,j \in \mathcal{K}} \Gamma(\alpha_{ij}) / \beta_{ij}^{\alpha_{ij}}.
\]
We use (11) to construct a surveillance rule for structural breaks in the next section.

3. Methods for Surveillance

We now consider several surveillance rules for structural breaks in the generator of rating migration matrices, using the continuous-time models of rating transition dynamics in Sections 2.2 and 2.3. Suppose that firms’ rating migrations follow a K-state continuous-time Markov chain whose probability transition matrices are characterized by their generator matrices $\Lambda(t)$. Denote by $\mathcal{Y}_{s:t}$ the observed ratings during the period $(s, t)$; then its sufficient statistic is given by $\{K_{s,t}^{(ij)}, S_{s,t}^{(ij)} \mid 1 \leq i,j \leq K\}$, where $K_{s,t}^{(ij)}$ and $S_{s,t}^{(ij)}$ are defined in Section 2.3. Assume that the generator matrix $\Lambda(t)$ is $\Lambda_0$ for $t < v$ and $\Lambda_1$ for $t \geq v$.

If both $\Lambda_0$ and $\Lambda_1$ are known, one may consider the cumulative sum (CUSUM) rule [25] to detect the change-point $v$,

$$N_{\text{CUSUM}} = \inf\{n : \max_{0 \leq k < n} \log[L(\Lambda_1; \mathcal{Y}_{k+1,n})/f(\Lambda_0; \mathcal{Y}_{k+1,n})] \geq c\}$$

with $c$ satisfying $E_0(N_{\text{CUSUM}}) = \gamma$. Reference [26] showed the asymptotic minimax property of the CUSUM rule, and it suggests that as $\gamma \to \infty$, the CUSUM rule (12) asymptotically minimizes the worst-case expected delay $\bar{E}_1(T) = \sup_{v \geq 1} \text{ess sup} E[(T - v + 1)^+ | \mathcal{Y}_{0,v}]$ over the class $\mathcal{F}_\gamma$ of all rules $T$ satisfying the constraint $E_0(T) \geq \gamma$, and furthermore, as $\gamma \to \infty$, $\bar{E}_1(N_{\text{CUSUM}}) \sim \inf_{T \in \mathcal{F}_\gamma} \bar{E}_1(T) \sim (\log \gamma)/I(L_1, L_0)$, where $I(L_1, L_0) = E_1\{\log(L(\Lambda_1; \mathcal{Y}_{1,t})/L(\Lambda_0; \mathcal{Y}_{1,t}))\}$ is the Kullback–Leiber information number.

The CUSUM rule cannot be used to detect changes in rating transition dynamics since the pre- and after-changes of generator matrices are unknown by definition of the problem. Hence we consider the following surveillance rules that assume either $\Lambda_1$ is unknown or both $\Lambda_0$ and $\Lambda_1$ are unknown. To better present the idea, we ignore the assumption that state $K$ is an absorbing state in this section and will make necessary modifications on discussed rules later for real data implementation.

3.1. Likelihood Ratio (LR) Tests

To study sequentially whether there is a change in the generator matrices of firms’ rating transition process, we first consider a hypothesis testing problem for observed ratings transitions $\mathcal{Y}_{t-1,t}$ at the period $(t - 1, t)$, $H_0 : \Lambda = \Lambda_0$ versus $H_1 : \Lambda \neq \Lambda_0$, where the generator matrix $\Lambda_0 = (\lambda(t))_{1 \leq i,j \leq K}$.

Since the log-likelihood function of the data for generator matrix $\Lambda$ is given by (4), the logarithm of the likelihood ratio statistic for the hypothesis testing problem is

$$Z_{LR}(\Lambda_0; \mathcal{Y}_{t-1,t}) := 2 \left( \max_{\Lambda \neq \Lambda_0} \log L(\Lambda; \mathcal{Y}_{t-1,t}) - \log L(\Lambda_0; \mathcal{Y}_{t-1,t}) \right)$$

$$= 2 \sum_{i=1}^{K} \left[ \sum_{j \neq i} K_{t-1,t}^{(ij)} \log \frac{\hat{\lambda}_{(ij)}^{(i)}}{\lambda_0^{(ij)}} - \sum_{j \neq i} (\hat{\lambda}_{(ij)} - \lambda_0^{(ij)}) S_{t-1,t}^{(ij)} \right],$$

where $\hat{\lambda}_{(ij)}$ is the forward estimate of $\lambda_{(ij)}$, and $S_{t-1,t}^{(ij)}$ is the worst-case expected delay.
where \( \hat{\lambda}^{(ij)} = K^{(ij)}_{t-1,K}/S^{(i)}_{t-1,j} \) is the maximum likelihood estimate (MLE) of \( \lambda^{(ij)} \) (1 \( i \neq j \leq K \)). Besides using the MLE of \( \Lambda \), one may specify gamma prior distributions for elements of \( \Lambda \) and obtain a mixture likelihood ratio (MLR). Suppose elements of the generator matrix \( \Lambda \) follow the gamma distribution (3); the MLR of the data is expressed as

\[
Z_{\text{MLR}}(\Lambda_0; Y_{1:t-1}) := 2 \log \left( \int L(\Lambda | Y_{1:t-1}) g(\Lambda) d\Lambda \right) - 2 \log L(\Lambda_0 | Y_{1:t-1})
\]

\[
= 2 \sum_{i=1}^{K} \left[ \sum_{j \neq i} \alpha_{ij} \log \beta_i - \log(\alpha_{ij}) - (K^{(ij)}_{t-1,i} + \alpha_{ij}) \log(S^{(i)}_{t-1,i}) + \log(\beta_i) \right]
\]

\[+ \log(K^{(ij)}_{t-1,i} + \alpha_{ij}) - K^{(ij)}_{t-1,i} \log(\hat{\lambda}^{(ij)}_{t-1}) + \left( \sum_{j \neq i} \lambda^{(ij)}_{t-1} + 1 - K \right) S^{(i)}_{t-1,i} \].

Based on the LR and MLR statistics, (13) and (14), we obtain the following rules that detect whether a change-point occurs at the period \( Y_{t+1} \) in the generator of rating migration matrices,

\[
N_{\text{LR}} = \inf\{ t \geq t_0 : Z_{\text{LR}}(\Lambda_0; Y_{1:t-1}) \geq c \}, \quad (15)
\]

and

\[
N_{\text{MLR}} = \inf\{ t \geq t_0 : Z_{\text{MLR}}(\Lambda_0; Y_{1:t-1}) \geq c \}. \quad (16)
\]

The above two rules assume the pre-change distribution of firms’ rating transitions is known. Since \( \Lambda_0 \) is unknown in reality, one can replace it by its MLE using historical observations. In particular, given the observed rating transitions \( Y_{0:t-1} \) during the period \((0, t-1]\), let \( \hat{\Lambda}_0(Y_{0:t-1}) \) be the MLE of \( \Lambda_0 \) during the period \((0, t-1]\). Then rules (15) and (16) can be modified as follows:

\[
\hat{N}_{\text{LR}} = \inf\{ t \geq t_0 : Z_{\text{LR}}(\hat{\Lambda}_0(Y_{0:t-1}); Y_{1:t-1}) \geq c \}, \quad (17)
\]

\[
\hat{N}_{\text{MLR}} = \inf\{ t \geq t_0 : Z_{\text{MLR}}(\hat{\Lambda}_0(Y_{0:t-1}); Y_{1:t-1}) \geq c \}. \quad (18)
\]

Besides constructing detection rules directly, the LR and MLR statistics (13) and (14) can also be used to construct exponentially weighted moving average (EWMA) control charts. Control charts were first introduced by [27,28] to monitor production processes and were later extended to different disciplines. The EWMA control chart, based on the EWMA statistic, was introduced by [29]. Its multivariate version was implemented by [30] and extended by [31] to monitor multivariate time series. The EWMA control chart has been applied in financial surveillance. For example, Reference [32] constructed multivariate EWMA control charts based on the Mahalanobis distance to sequentially detect a change in the parameters of the Cox–Ingersoll–Ross term structure model for interest rates. Using the LR and MLR statistics (13) and (14), one can easily construct the following EWMA control charts for generator matrices of rating transition matrices:

\[
W_{t,\text{LR}} = (1 - \alpha)W_{t-1,\text{LR}} + \alpha Z_{\text{LR}}(\hat{\Lambda}_0(Y_{0:t-1}); Y_{1:t-1}), \quad t \geq 1,
\]

\[
W_{t,\text{MLR}} = (1 - \alpha)W_{t-1,\text{MLR}} + \alpha Z_{\text{MLR}}(\hat{\Lambda}_0(Y_{0:t-1}); Y_{1:t-1}), \quad t \geq 1.
\]

As it is usually difficult to find the distributions of the above two EWMA statistics, one may use bootstrap methods to estimate the distribution of the EWMA statistics for real data implementation.

3.2. Generalized Likelihood Ratio (GLR) Rule

The LR and MLR statistics (or their modified versions) can be considered as likelihood ratio statistics for the problem of testing whether a change point occurs at \( t - 1 \) given observations \( Y_{0:t} \). In general, since both \( \Lambda_0 \) and \( \Lambda_1 \) are unknown in practice, one may consider the GLR statistic for testing
the null hypothesis of no change-point based on $Y_{0,1}, \ldots, Y_{t-1,t}$, versus the alternative hypothesis of a single change-point prior to $t$ but not before $t_0$:

$$Z_{GLR}(Y_{0,t}) := \max_{n_0 \leq k \leq n} \left\{ \sup_{\Lambda_1} \log L(A_1; Y_{0,k}) + \sup_{\Lambda_2} \log L(A_2; Y_{t,k}) - \sup_{\Lambda_0} \log L(A_0; Y_{0,t}) \right\}$$

$$= \max_{n_0 \leq k \leq n} \sum_{1 \leq i \leq K} \left\{ K_{0,k}^{(i,j)} \log \frac{K_{0,k}^{(i,j)}}{s_{0,k}^{(i,j)}} + K_{i,j}^{(i,j)} \log \frac{K_{i,j}^{(i,j)}}{s_{i,j}^{(i,j)}} - K_{0,0}^{(i,j)} \log \frac{s_{0,0}^{(i,j)}}{s_{0,0}^{(i,j)}} \right\}$$

where $\sup_{\Lambda_0}$ is the maximum likelihood under the null hypothesis, and $\sup_{\Lambda_1}$ and $\sup_{\Lambda_2}$ are obtained by maximizing the likelihood under the hypothesis of a single change-point occurring at $k + 1$. Then the GLR rule with unknown pre- and post-change parameters is

$$N_{GLR} = \inf \left\{ t > t_0 : Z_{GLR}(Y_{0,t}) \geq c \right\}. \quad (19)$$

The GLR rule in the normal case has been discussed by [33,34]. For the GLR rule in the continuous-time Markov chain, we will study its performance via simulation in Section 4.

### 3.3. An Extended Shiryaev’s Detection Rule

Instead of the non-Bayesian approach, Bayesian rules can be used for sequential surveillance of the rating migration matrix. Reference [35,36] assumed a geometric prior distribution on $v$ and a loss function that takes value $c$ for each observation taken at or after $v$ and 1 for a false alarm before $v$, provided an optimal sequential detection of the change-time $v$ by the Bayesian approach. He showed that an alarm would be triggered as soon as the posterior probability exceeds some fixed level when a change has occurred using optimal stopping theory. By applying Shiryaev’s result to rating transition observations, we have

$$P\{v \leq n | Y_{0,n}\} = R_{p,n} / (R_{p,n} + p^{-1})$$

where $p$ is the parameter of the geometric distribution $P\{v = n\} = p(1 - p)^{n-1}$ and $R_{p,n} = \sum_{k=1}^{n} (1 - p)^{-(n-k+1)}$, $L(Y_{k,n}| A_1)$, $L(Y_{k,n}| A_0)$, the Bayes rule declares at time

$$N_{Shi} = \inf \{ n \geq 1 : R_{p,n} \geq c \},$$

when a change has occurred. Following [37,38], we can set $p = 0$ and obtain the Shiryaev–Roberts rule for detecting a change-point in firms’ rating transition dynamics:

$$N_{Shi,Rob} = \inf \left\{ n \geq 1 : \sum_{k=1}^{n} L(Y_{k,n}| A_1) \geq c \right\}.$$  

According to [37], when $p \to 0$, $N_{Shi,Rob}$ is asymptotically Bayes-risk efficient.

Reference [17] extended Shiryaev’s Bayesian detection rule and described asymptotically optimal Bayesian and frequentist solutions to the problem of sequential change-point detection in multiparameter exponential families when the pre- and post-change parameters are unknown. The extended Shiryaev’s rule has the following form:

$$N_{exShi} = \inf\{ n > n_p : P\{v \leq n | v \geq n - k_p, Y_{0,n}\} \geq c \}.$$  

(20)

Following [17], we may construct the extended Shiryaev’s Bayesian detection rule for change-points in rating migration matrices as follows. Assume that $v$ follows a geometric prior distribution with parameter $p$ but is constrained to be larger than $n_0$, and $A_0, A_1$ are generator matrices in which the $ij$th
we obtain the extended Shiryaev’s rule for detecting change-points in generators of continuous-time
presents the forward estimates of generator matrices, the probability of the most recent change-points,
period, because the quickest detection in most engineering applications assumes that the system
Entropy 2020, 22, 1072
3.4. Bayesian Rules Based on the Stochastic Structural Break Model
4. Monte Carlo Results
In this section we go through numerical experiments to discuss the performance of surveillance
rules in Section 3. To fix the idea, we assume the rating system has $K = 4$ rating categories
in which there is no default category (or absorbing state), and all firms’ rating transitions follow
four-state continuous-time Markov chains. For convenience, we further assume that there are always
$n = 1000$ firms with their observed ratings at each time $t$.
We first consider the issue of simulating a generator matrix for firms’ rating transitions. There are
many ways to simulate generator matrices of continuous-time homogeneous Markov chains. Instead of
using estimated generator matrices from real data, we try to simulate a $K$-state generator matrix
for rating transitions. We note that firms’ one-year rating transition probability matrix, or probabilities of staying at the previous rating categories, are usually larger than 0.9.
Second, according to CRA’s rating procedures, a firm’s rating is usually upgraded or downgraded sequentially based on the superiority order of rating categories. This suggests that the off-diagonal transition probabilities $p_{ij}$ decrease with the value of $|i - j|$. Due to these two features, we simulated generator matrix $\Lambda = (\lambda^{(i,j)})_{1 \leq i,j \leq K}$ in the following way. For $1 \leq i \neq j \leq K$, we simulated $\lambda^{(i,j)}$ from a uniform distribution $\text{Unif}(a^{-|i-j|} - 1, a^{-|i-j|})$ ($a$ is a positive constant), and for $1 \leq i \leq K$, we set $\lambda^{(i,i)} = - \sum_{j \neq i} \lambda^{(i,j)}$. The positive constant $a$ can be chosen to adjust the difference between off-diagonal elements, and in our study, we chose $a = 16$. Table 1 shows three sets of simulated generators and their one-period transition probability matrices. The SVD metrics of three one-period rating transition matrices are given below.

$$M(P_A) = 0.07850, \quad M(P_B) = 0.06716, \quad M(P_C) = 0.08793.$$ 

![Table 1. Simulated generators and their transition probability matrices.](image)

We then carried out the following experiment. Suppose that firm $l$’s $(1 \leq l \leq n)$ rating transitions follow a continuous-time homogeneous Markov chain with generator matrix $\Lambda_0$ for $t < v$ and with generator matrix $\Lambda_1$ for $t \geq v$, respectively. The following two scenarios are considered.

(S1) $\Lambda_0 = \Lambda_A$ and $\Lambda_1 = \Lambda_B$.
(S2) $\Lambda_0 = \Lambda_A$ and $\Lambda_1 = \Lambda_C$.

Note that in (S1), $|M(P_1) - M(P_0)| = |M(P_A) - M(P_B)| = 0.01134$, and in (S2), $|M(P_1) - M(P_0)| = |M(P_A) - M(P_C)| = 0.00943$. This indicates that the sizes of changes in transition probability matrices in (S1) and (S2) are very small (For the case $|M(P_1) - M(P_0)| > 0.02$, all surveillance rules have no detection delay with the settings of our simulation. Those results are not presented in the paper as all surveillance rules provided very good performances).

We considered four surveillance rules presented in Section 3 and determined the threshold $c$ in the surveillance rule as follows. Assume that there is no change in generators of firms’ rating transitions, that is, all rating transitions follow the continuous-time homogeneous Markov chain with generator $\Lambda_A$ at each period. We then computed the statistics $Z$ in the surveillance rule. We ran said simulation $m = 1000$ times and obtained $\{Z_1, \ldots, Z_m\}$. We then chose the 99% quantile of $\{Z_1, \ldots, Z_m\}$ as the threshold $c$ in the surveillance rule.

For scenarios (S1) and (S2), we considered three locations of structural breaks $v = 50, 100, 200$; hence, in total we had six scenarios. We ran 500 simulations in each scenario. In each simulation, we generated $n$ Markov chains at time $t = 1, \ldots, T$ ($T = 1000$) and applied surveillance rules. Using the 500 detected locations of change-point, we estimated the conditional detection delay (CDD), defined as $CDD_N = E(N - v | N > v)$, and the false alarm probability (FAP) defined as $P(N < v)$ when there was a change at $v$. Table 2 shows the estimated CDDs and FAPs and their standard errors in parentheses. The estimated FAPs for all surveillance rules increase with the location $v$ of the
change-point. Among the four surveillance rules in the study, the FAP of the LR rule increased fastest and more significantly than the other rules. The FAP of the GLR rule increased less than that of the LR rule, but a little more than the ex.Shi, SSB.prob, and SSB.lr rules. Compared to the FAP, all surveillance rules’ CDDs did not change significantly. In particular, the CDDs of the ex.Shi and SSB rules were larger than those of the LR and GLR rules, and the CDDs of the GLR rule were larger than that of the LR rule. Table 2 seems to suggest the following. Compared with the other three rules, the LR rule had relatively larger FAPs but smaller CDDs. The two Bayesian rules, ex.Shi and SSB, had lower FAPs but slightly larger CDDs than the LR and the GLR rules. The GLR rule seems to have had relatively balanced FAPs and CDDs compared to other rules.

Table 2. False alarm probability and conditional detection delay.

| Rules | $v = 50$ | $v = 100$ | $v = 200$ | $v = 50$ | $v = 100$ | $v = 200$ |
|-------|----------|----------|----------|----------|----------|----------|
| (S1)  |          |          |          |          |          |          |
| LR    | 0.248 (0.019) | 0.354 (0.021) | 0.810 (0.017) | 2.420 (0.823) | 1.659 (0.130) | 1.916 (0.269) |
| GLR   | 0.062 (0.011) | 0.120 (0.014) | 0.342 (0.021) | 2.746 (0.041) | 2.607 (0.041) | 2.596 (0.045) |
| ex.Shi| 0.036 (0.008) | 0.076 (0.012) | 0.260 (0.019) | 4.340 (0.052) | 4.284 (0.051) | 4.300 (0.056) |
| SSB   | 0.036 (0.008) | 0.064 (0.011) | 0.238 (0.019) | 4.315 (0.053) | 3.667 (0.049) | 3.766 (0.052) |
| (S2)  |          |          |          |          |          |          |
| LR    | 0.278 (0.020) | 0.342 (0.021) | 0.782 (0.018) | 0.343 (0.040) | 0.571 (0.062) | 0.358 (0.063) |
| GLR   | 0.064 (0.011) | 0.120 (0.015) | 0.376 (0.022) | 2.068 (0.033) | 1.932 (0.036) | 1.987 (0.040) |
| ex.Shi| 0.070 (0.011) | 0.092 (0.013) | 0.274 (0.020) | 3.357 (0.041) | 3.236 (0.043) | 3.154 (0.046) |
| SSB   | 0.064 (0.011) | 0.074 (0.012) | 0.248 (0.019) | 3.340 (0.041) | 2.659 (0.041) | 2.697 (0.043) |

5. Data Analysis

In this section we describe an empirical study of using the surveillance rules in Section 4 to analyze structural breaks in the U.S. firms’ rating transition dynamics during 1986–2016 and discuss the implication to structural breaks in the credit market during the period.

5.1. Data Description

The data were obtained from COMPUSTAT and consist of Standard and Poor’s monthly credit ratings of firms starting January 1986 and ending February 2017 (The proposed surveillance rules can also be applied to firms’ rating data provided by other credit rating agencies, such as Moody’s and Fitch Group). There were a total of 703,085 rating records and 5802 firms whose ratings were recorded at the end of each month. The raw data contained ten rating categories, A A A, A A, A, BBB, BB, B, CCC, CC, C, and D (default), and 25 rating subcategories which were obtained by adding “+” or “-” to the letter grades of categories and showing relative standing within the major rating categories. For convenience, we grouped C and CC into CCC, as the records in the former two rating categories were few. Hence we had eight rating categories in total, which are denoted by A A A, A A, A, BBB, BB, B, CCC, and D.

The top and middle panels of Figure 1 display the number of rated firms and proportion of firms’ ratings in each rating category from 1986 to 2016 and show the following facts. First, the total number of rated firms increased from 1738 in 1986, peaked at 2833 in 2002, and decreased to 2156 in 2016. Second, the numbers and proportions of firms rated A A A, A A, and A decreased significantly during 1986–2016, while the number and proportion of firms rated BBB almost doubled up from 291 to 719 and from 16.7% to 33.3%, respectively, during the whole period. This means the number of firms rated as “investment grade ratings” (i.e., A A A, A A, A, BBB, BB, B, CCC, CC, C, and D) peaked around 4% during 2002–2003 and at 2.8% at 2009.
As it is not obvious to see whether structural breaks occur in the number and proportion of rated firms, the bottom panel shows the SVD metric of estimated one-year transition probability matrices during 1986–2016. Each year’s transition probability matrix is computed based on a continuous-time homogeneous Markov chain model. The SVD curve ranges from 0.0810 to 0.2540 during the sample period and has three local maximums, which are 0.2326 in 2002, 0.2540 in 2009, and 0.2351 in 2016. The variation of the SVD metric in one-year rating transition probability matrices indicates the non-stationarity of firms’ rating transition dynamics and the instability of the credit market during 1986–2016.

5.2. Surveillance of Structural Breaks

Next we carried out surveillance procedures for structural breaks in firms’ rating transition dynamics. Since the pre- and post-change generator matrices of rating transition dynamics are not known, the surveillance procedures include steps of determining threshold values of surveillance rules, computing the surveillance rule statistic, and deciding whether a structural break should be claimed each month, whose details are given below.
Let $t_0 = \text{January 1997}, T = \text{February 2017}$, and $n$ be the length of the sliding window in which observed ratings are used for surveillance. At each month $t = t_0, t_0 + 1, \ldots, T$, we used rating data $Y_{t-n+1,t}$ at months $\{t - n + 1, t - n + 2, \ldots, t\}$ to compute the surveillance statistic and compare it with the corresponding threshold in the surveillance rule. If the statistic exceeded the threshold, a structural break $\tau_k = t$ was claimed; otherwise, it was concluded that ratings dynamics in month $t$ had no structural break and we proceeded to the next month. Then we obtained a series of structural breaks at $\tau_1, \tau_2, \ldots$. For convenience, we also denote $\tau_0 = t_0$.

The threshold $c$ in the surveillance rule between two adjacent months of structural breaks $\tau_k$ and $\tau_{k+1}$ ($k = 0, 1, \ldots$) is determined by Monte Carlo or bootstrap methods using the rating records $Y_{\tau_k - n, \tau_k}$ at months $\{\tau_k - n, \tau_k - n + 1, \ldots, \tau_k\}$. Specifically, for the LR rule (15), we estimate the pre-change generator matrix $\Lambda_0$ using data $Y_{\tau_k - n, \tau_k}$ and then use the estimated generator $\hat{\Lambda}_0$ to simulate $m = 1000$ groups of rating transitions. Each group of simulated rating transition contains a set of Markov chains in which each Markov chain follows a continuous-time homogeneous Markov chain with generator $\hat{\Lambda}_0$ and the number of Markov chains is the same as that of firms at period $[\tau_k, \tau_k + 1]$. The LR statistic is computed for each group of simulated rating transitions and the $99\%$ quantile of these $m$ LR statistics is chosen as the threshold value $c_{LR}$ for the LR surveillance rule. For the GLR rule, we used rating records $Y_{\tau_k - n, \tau_k}$ to generate $m$ bootstrap samples and computed the GLR statistic for each bootstrap sample. Then the threshold value $c_{GLR}$ was chosen as the $99\%$ quantile of the $m$ bootstrap GLR statistics. The threshold values $c_{\text{ex.Shi}}$ and $c_{\text{SSB}}$ were determined similarly to those in the GLR rule via the bootstrap method. Note that all thresholds need to be determined after the month of structural break $\tau_k$ is claimed, and the months of structural breaks detected by different surveillance rules are not necessarily the same.

We implemented four surveillance rules in the study, which were the LR rule $N_{\text{LR}}$, the GLR rule $N_{\text{GLR}}$, the extended Shiryaev’s rule $N_{\text{ex.Shi}}$, and the change-point probability rule $N_{\text{SSB}}$ in Section 3 for $n = 60$ months. Figure 2 plots the statistics of four surveillance rules in the moving window from January 1997 to February 2017. The piecewise horizontal lines in each panel of Figure 2 are the thresholds used in each rule for each segmented series. Note that those thresholds were estimated via Monte Carlo or bootstrap methods whenever a structural break was claimed. Besides, depending on the user’s definition of structural breaks, the significance level of those thresholds can also be adjusted in practice. Table 3 lists all structural breaks detected by four surveillance rules. Note that the detected months with structural break are quite different, except that all rules detected changes in 2009, which seems to agree with the fact that the 2007–2008 financial crisis significantly deteriorated the credit market. This is consistent with Figure 2 in which the SVD metric peaks in 2002 and 2009. However, exact months suggested by the four rules have an up to six-month difference, suggesting that some of these rules have relatively large detection delays though the month with the true structural break being unknown. Besides, it seems that the LR and GLR rules are more sensitive to structural changes in generator matrices, as they suggest more structural breaks than other rules.

Table 3. Detected structural breaks in firms’ rating transition dynamics.

| Rules   | Structural Breaks                           |
|---------|--------------------------------------------|
| LR      | August 2000, March 2002, August 2009, August 2010 |
| GLR     | November 1999, May 2005, September 2005, June 2009, March 2013 |
| ex.Shi  | July 2003, March 2005, March 2009, September 2013 |
| SSB     | August 2004, October 2009, October 2012     |
6. Conclusions

This paper studied the problem of the sequential surveillance of structural breaks in firms’ rating transition dynamics. After reviewing some discrete- and continuous-time Markov chain models of firms’ rating transition dynamics, we presented several surveillance rules in statistical process control and applied them to continuous-time Markovian models of rating transition dynamics. Our simulation study showed that the GLR rule performs better than other rules in terms of the trade-off between the false alarm probability and the conditional detection delay. As an example, we applied four surveillance rules to detect structural breaks in the rating transition dynamics of the U.S. firms during 1997–2017. The detected structural breaks show that large structural changes in firms’ rating transitions can be captured, but different surveillance rules tend to provide different structural breaks when they are not significant. The real example also showed that the GLR rule is more sensitive to structural changes in generator matrices.

Besides the specific application of surveillance rules in firms’ rating transition dynamics, this paper shows the application of statistical process control in monitoring financial stability at the regulatory level. Actually, as statistical process control has been discussed and used in various areas of finance, including portfolio optimization, financial time series analysis, and technical analysis of trading, its application to monitoring the stability of the market and the economy does not catch the eye of the statistician or the econometrician. Our study fills in this gap, and the analysis in the paper also suggests that more effective surveillance rules should be studied in subsequent research.

Figure 2. Sequential surveillance of firms’ rating transition dynamics using the LR, GLR, ex.Shi, and SSB rules during January 1997 — February 2017.
Author Contributions: Conceptualization, H.X.; methodology, H.X. and K.W.; software, H.X., K.W., and Y.C.; validation, Z.L. and Y.C.; formal analysis, H.X. and K.W.; investigation, H.X.; resources, H.X. and Z.L.; data curation, X.H. and Z.L.; writing—original draft preparation, H.X. and K.W.; writing—review and editing, H.X., Z.L., and Y.C.; visualization, H.X.; supervision, H.X.; project administration, H.X.; funding acquisition, H.X. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Conflicts of Interest: The authors declare no conflict of interest.

Abbreviations

The following abbreviations are used in this manuscript:

CUSUM cumulative sum
GLR Generalized likelihood ratio
SSB Stochastic structural break
SVD Singular value decomposition
LR Likelihood ratio

References

1. Financial Stability Board. COVID-19 Pandemic: Financial Stability Implications and Policy Measures Take—Report Submitted to the G20 Financial Ministers and Governors; Financial Stability Board: Basel, Switzerland, 15 July 2020.
2. Langohr, H.M.; Langohr, P.T. The Rating Agencies and their Credit Ratings: What They Are, How They Work and Why They Are Relevant; John Wiley & Sons Ltd.: Chichester, UK, 2010.
3. Opp, C.; Opp, M.; Harris, M. Rating agencies in the face of regulation. J. Financ. Econ. 2013, 108, 46–61. [CrossRef]
4. Skreta, V.; Veldkamp, L. Rating shopping and asset complexity: A theory of rating inflation. J. Monet. Econ. 2009, 56, 678–695. [CrossRef]
5. Mariano, B. Market power and reputational concerns in the ratings industry. J. Monet. Econ. 2012, 36, 1616–1626. [CrossRef]
6. Mathis, J.; McAndrews, J.; Rochet, J.C. Rating the raters: Are reputation concerns powerful enough to discipline rating agencies? J. Monet. Econ. 2009, 56, 657–674. [CrossRef]
7. Povel, P.; Singh, R.; Winton, A. Booms, busts, and fraud. Rev. Financ. Stud. 2007, 20, 1219–1254. [CrossRef]
8. Xing, H.; Sun, N.; Chen, Y. Rating agencies achieve rating stability. J. Bank. Financ. 2012, 36, 78–89. [CrossRef]
9. Xing, H.; Chen, Y. Dependence of structural breaks in rating transition dynamics on economic and market variations. Rev. Econ. Financ. 2018, 11, 1–18.
10. Xing, H.; Yu, Y. Firm’s credit risk in the presence of market structural breaks. Risks 2018, 6, 136, doi:10.3390/risks6040136. [CrossRef]
11. Merton, R.C. On the pricing of corporate debt: The risk structure of interest rates. J. Financ. 1974, 29, 449–470.
12. Leland, H.E. Corporate debt value, bond covenants, and optimal capital structure. J. Financ. 1994, 49, 1213–1252. [CrossRef]
13. Duffie, D.; Saita, L.; Wang, K. Multi-period corporate default prediction with stochastic covariates. J. Financ. Econ. 2007, 83, 635–665. [CrossRef]
14. Duffie, D.; Eckner, A.; Horel, G.; Saita, L. Frailty correlated default. J. Financ. 2009, 64, 2089–2123. [CrossRef]
15. Altman, E.; Rijken, H. How rating agencies achieve rating stability. J. Bank. Financ. 2004, 28, 2679–2714. [CrossRef]
16. Gordy, M.; Howells, B. Procyclicality in Basel II: Can we treat the disease without killing the patient. J. Financ. Intermediat. 2006, 15, 395–417. [CrossRef]
17. Lai, T.L.; Xing, H. Sequential change-point detection when the pre- and post-change parameters are unknown. Seq. Anal. 2010, 29, 162–175. [CrossRef]
18. Lai, T.L.; Liu, T.; Xing, H. A Bayesian approach to sequential surveillance in exponential families. Commun. Stat. Theory Methods 2009, 38, 2958–2968. [CrossRef]
19. Shorrocks, A.F. The measurement of mobility. Economietica 1978, 46, 1013–1024. [CrossRef]
20. Geweke, J.; Marshall, R.C.; Zarkin, G.A. Mobility indices in continuous time Markov chains. *Econometrica* **1986**, *54*, 1407–1423. [CrossRef]
21. Jafry, Y.; Schuermenn, T. Measurement, estimation and comparison of credit migration matrices. *J. Bank. Financ.* **2004**, *28*, 2603–2639. [CrossRef]
22. Lando, D. *Credit Risk Modeling: Theory and Applications*; Princeton University Press: Princeton, NJ, USA, 2009.
23. Küchler, U.; Sørensen, M. *Exponential Families of Stochastic Processes*; Springer: New York, NY, USA.
24. Andersen, P.; Borgan, O.; Gill, R.; Keiding, N. *Statistical Models Based on Counting Processes*; Springer: New York, NY, USA, 2012.
25. Page, E.S. Continuous inspection schemes. *Biometrika* **1954**, *41*, 100–114. [CrossRef]
26. Lorden, G. Procedures for reacting to a change in distribution. *Ann. Math. Stat.* **1971**, *42*, 1897–1908. [CrossRef]
27. Shewhart, W. Quality control charts. *Bell Syst. Tech. J.* **1926**, *5*, 593–603. [CrossRef]
28. Shewhart, W. *Economic Control of Quality of Manufactured Product*; D. Van Nostrand Company: Toronto, ON, Canada, 1931.
29. Roberts, S. Control charts tests based on geometric moving averages. *Technometrics* **1959**, *1*, 239–250. [CrossRef]
30. Lowry, C.; Woodall, W.; Champ, C.; Rigdon, S. A multivariate exponentially weighted moving average chart. *Technometrics* **1992**, *34*, 46–53. [CrossRef]
31. Kramer, H.; Schmid, W. EWMA charts for multivariate time series. *Seq. Anal.* **1997**, *16*, 131–154. [CrossRef]
32. Schmid, W.; Tzotchev, D. Statistical surveillance of the parameters of a one-factor Cox-Ingersoll-Ross model. *Seq. Anal.* **2004**, *23*, 379–412. [CrossRef]
33. Pollak, M.; Siegmund, D. Sequential detection of a change in a Normal mean when the initial value is unknown. *Ann. Stat.* **1991**, *19*, 394–416. [CrossRef]
34. Siegmund, D.; Venkatraman, E.S. Using the generalized likelihood ratio statistic for sequential detection of a change-point. *Ann. Stat.* **1995**, *23*, 255–271. [CrossRef]
35. Shiryaev, A.N. On Optimum Methods in quickest detection problems. *Theor. Probab. Appl.* **1963**, *8*, 22–46. [CrossRef]
36. Shiryaev, A.N. *Optimal Stopping Rules*; Springer: New York, NY, USA, 1978.
37. Pollak, M. Optimal detection of a change in distribution. *Ann. Stat.* **1985**, *13*, 206–227. [CrossRef]
38. Roberts, S. A comparison of some control chart procedures. *Technometrics* **1966**, *8*, 411–430. [CrossRef]
39. Frisén, M. Optimal sequential surveillance for finance, public health, and other areas. *Seq. Anal.* **2009**, *28*, 310–337. [CrossRef]