

# Modeling of Contagious Rating Changes and Its Application to Multi-Downgrade Protection\*

Hidetoshi NAKAGAWA<sup>†</sup>

Graduate School of International Corporate Strategy, Hitotsubashi University  
2-1-2 Hitotsubashi, Chiyoda-ku, Tokyo 101-8439 Japan

February 17, 2009 (First version on Jan. 27, 2008)

## Abstract

This paper presents a new modeling of the intensities of typical rating change events with a multivariate affine jump process. The main feature of our model is that the intensity model has a mutually exciting structure in the sense that an event occurrence can jump not only the intensity of the same type of events but those of another event types. Also, such a modeling with the multivariate affine jump process enables us to execute the maximum likelihood estimation of the model parameters relatively easily and to compute the expectation of the number of events without Monte Carlo simulation. We actually try the maximum likelihood estimation for a specific model with some historical record of rating changes for the corporate bond issued by Japanese companies. In addition, we propose a new credit derivative named multi-downgrade protection (MDP) as an application of our mutually contagious intensity model. Some numerical illustrations on the expectation of the number of events and the fair value of MDP are also presented.

**Keywords:** Rating change, mutually exciting intensity model, affine jump process, maximum likelihood estimation, downgrade protection

## 1 Introduction

This paper presents a new modeling of the intensities of typical rating change events such as downgrade and upgrade within the top-down approach framework introduced by Giesecke and Goldberg [7]. More specifically, we apply a multivariate affine jump process to model the intensities. so that we can make the intensity model have a mutually exciting structure in the sense that an event occurrence can jump not only the intensity of the same type of events but those of another event types. Also, such a modeling with the multivariate affine jump process enables us to execute the maximum likelihood estimation of the model parameters relatively easily and to compute the expectation of the number of events without Monte Carlo simulation.

---

\*This research was supported by Grant-in-Aid for Scientific Research (A) No. 20241038 from Japan Society for the Promotion of Science (JSPS).

<sup>†</sup>E-mail: hnakagawa "at" ics.hit-u.ac.jp

Giesecke and Goldberg [7] call “bottom-up” the approach that firstly models each obligor’s default risk and then considers the dependence structure among the obligors to measure the default risk of the whole credit obligation portfolio. On the other hand, they call “top-down” the approach that sets aside the individual characteristics of each obligor and focuses on how the default events happen in the portfolio, in other words, on the counts of the defaults happened in the underlying portfolio.

Giesecke and Goldberg [7] indeed formulate, in the so-called reduced-form framework, the portfolio default intensity process with a self-exciting intensity process, which is originally studied in Hawkes [8]. The word “self-exciting” means that the intensity of next event can jump due to happening of the event. One can say the self-exciting intensity model is one of the models that reflect default contagion. Errais et al. [6] uses an affine jump-diffusion process type that is originally studied by Duffie et al. [5] to model the self-exciting intensity process for discussing a general procedure of obtaining the premium of index swaps and tranche swaps of credit default swap(CDS).

In this study we pay attention to typical rating change events such as downgrade and upgrade, rather than defaults. One reason is that consideration of contagious defaults sounds similar to remarking the clustering of downgrades or upgrades, which is actually observed in some Japanese actual data. Another reason is that empirical studies such as parameter estimation seem easier to do since rating change events happen much more frequently than defaults. Instead, we do not pay attention to modeling each specific transition probability from a particular grade to another particular grade (for example, from “AA” to “BBB”).

Dynamic rating transition is usually modeled by rating transition intensity matrix, but it is difficult to use the rating transition matrix framework so as to consider dynamic risk dependence in the portfolio. Though some non-Markov frameworks are used for modeling dependence like contagion, some unnatural strong assumptions are necessary to achieve the rating transition probability matrix via rating transition intensity matrix as is pointed out in Chapter 6 of Lando [10] and in Chapter 8 of Schönbucher [13]. Also, it seems quite hard to efficiently estimate all the specific rating transition intensities for any pair of pre and post rating grade.

The purpose of the paper is to utilize a multivariate affine jump process to model mutually exciting intensities of some typical rating change events (downgrades, upgrades and the other events) and to discuss possibility for some practical use of our model. Especially, we discuss parameter estimation from historical event observations and an application of our model to valuation of some financial product.

The maximum likelihood estimation for an affine jump process is argued in Azizpour and Giesecke [2] and Bowsher [4]. Azizpour and Giesecke [2] focus on the maximum likelihood estimation from the historical default data of U.S. corporate bonds with the portfolio default intensity given by a one dimensional affine jump (diffusion) process. The estimation method we actually try is largely based on their idea. Although Bowsher [4] is rather interested in high frequent events such as trades and mid-quote changes in the more liquid financial market, he treats the maximum likelihood estimation for some extensions of the mutually exciting intensity model introduced in Hawkes [8]. Also he mentions that some issues on the maximum likelihood estimators for multivariate non-stationary point processes are still unsolved.

Moreover, some useful mathematical properties of multivariate affine jump processes are studied in Errais et al. [6] and in Duffie et al. [5]. Their results help to compute the expectation

of the number of events without Monte Carlo simulation.

As an illustrative empirical study, we use the maximum likelihood method examined in Azizpour and Giesecke [2] to estimate the model parameters of a specific multivariate affine jump process from the data about issuer rating change events of Japanese enterprises from April 1998 to September 2008 by Rating and Investment Information, Inc.(R&I)<sup>1</sup> The estimation is still tentative, but consequently we partly observe if there exist some mutually exciting effects among downgrades, upgrade and the other events in the Japanese corporate bond market.

For the purpose of demonstrating some applications of our model to valuation of some financial product, we freshly propose a derivative named multi-downgrade protection (MDP). Simply put, MDP is supposed to be the contract that the protection seller pays to the buyer the amount according to the agreed rule over and over again whenever the target rating event (for example, from the investment grades to the speculative grades) happens in the underlying portfolio during the predetermined period.

In order to observe how the mutually exciting property of the event intensity model influences the value of some derivative, it is essential to consider such a multi payoff case as MDP. However, from a practical view, MDP seems more useful to manage the downgrade risk in the portfolio that consists of a wide variety and number of corporate bonds because it is thought that it is more important to consider how many downgrades will happen rather than which bond will be downgraded for portfolio management.

As numerical illustrations, we use the numerical approach presented in Errais et al. [6] to compute the expectation of the number of events as well as the theoretical value of MDP. Also, the influences of different values of some specific parameters to those quantities are discussed.

This paper is organized as follows. In section 2, we introduce our mutually exciting event intensity model of some point process with a multivariate affine jump process in the general setting. Additionally, we discuss the likelihood function for point processes and an explicit representation of the conditional expectation of affine jump processes. Section 3 presents the methodology and the results of the maximum likelihood estimation with the historical data of R& I's rating change events of Japanese enterprises. In section 4, we introduce multi-downgrade protection (MDP) and discuss the risk-neutral valuation of MDP. Also, demonstrated are some numerical illustrations about the expectation of the number of events as well as the theoretical value of MDP. Some concluding remarks are given in the last section.

## 2 Modeling of Contagious Rating Change

This section presents our rating change model that admits mutually contagious effects among downgrades, upgrades and the other events. In addition, we give some results on likelihood function of point processes as well as calculation of the conditional expectations of point processes with affine jump intensity.

---

<sup>1</sup>R&I is the largest rating agency to assign the rating to Japanese enterprises.

## 2.1 Rating change modeling via affine point processes

We will model contagious rating change events such as downgrade and upgrade via point processes with a multivariate affine jump process, which is a slight generalization of self-exciting intensity studied in Errais et al. [6].

Let  $(\Omega, \mathcal{F}, P)$  be a complete probability space and  $(\mathcal{F}_t)$  be the filtration that makes any processes appeared in this paper adapted.

For some  $m \in \mathbb{N}$ , let  $0(= \tau_0^i) < \tau_1^i < \tau_2^i < \dots$  ( $i = 1, \dots, m$ ) be  $(\mathcal{F}_t)$ -adapted point processes, that is, increasing sequences of  $(\mathcal{F}_t)$ -stopping times.  $\tau_k^i$  is regarded as the time when  $k$ -th event of type  $i$  happens. Also,  $N_t^1, \dots, N_t^m$  are counting processes associated with the point processes  $\{\tau_k^1\}_{k \in \mathbb{N}}, \dots, \{\tau_k^m\}_{k \in \mathbb{N}}$ , respectively.

Suppose that  $[N^i, N^j]_t = 0$  a.s. if  $i \neq j$ , that is, almost surely no simultaneous jump for different type events,.

In our rating change model, type 1 will be regarded as "downgrade", type 2 will be "upgrade" and type 3 will be "the others" such as withdrawal of rating.

Next,  $L_t^1, \dots, L_t^m$  are  $(\mathcal{F}_t)$ -adapted pure jump processes whose jump times coincide with those of  $N_t^1, \dots, N_t^m$ . More specifically, for each  $i$ ,  $L_t^i$  can be characterized by independently and identically distributed random variables  $\eta_1^i, \eta_2^i, \dots$  such as

$$L_t^i := \sum_{k=1}^{N_t^i} \eta_k^i.$$

Here we suppose that for any  $k \in \mathbb{N}$  and  $i = 1, \dots, m$ ,  $\eta_k^i$  is  $\mathcal{F}_{\tau_k^i}$ -measurable.

As for the rating change model, for example,  $\eta_k^1$  can be seen as variation range of downgrade occurred at time  $\tau_k^1$ . Then  $L_t^1$  stands for the cumulative sum of downgrade range up to time  $t$ . In the example later, we will suppose that  $\eta_k^i$  stands for the number of type  $i$  events occurred coincidentally at time  $\tau_k^i$ .

Then we need to specify the intensity process  $\lambda_t^i$  associated with  $N_t^i$  or equivalently  $L_t^i$ , namely, an  $(\mathcal{F}_t)$ -progressively measurable non-negative process so that the process  $M_t^i$  defined by

$$M_t^i := N_t^i - \int_0^t \lambda_s^i ds$$

is an  $(\mathcal{F}_t)$ -martingale.

In this study, we aim to model the intensities so that for any  $i$ ,  $\lambda_t^i$  can be influenced not only by occurrences of type  $i$  event itself (namely "self-exciting" effect), but also by the events other than type  $i$  (namely "mutually exciting").

For the purpose, for given  $d \in \mathbb{N}$ , we introduce some stochastic state vector  $\mathbf{X}_t := {}^t(X_t^1, \dots, X_t^d)$ , which is  $d$ -dimensional process driving the intensities of all types. We presume that  $\mathbf{X}_t$  is a strong solution to the following affine jump process<sup>2</sup>.

$$d\mathbf{X}_t = (K_0(t) + K_1(t) \cdot \mathbf{X}_t)dt + \sum_{i=1}^m \Xi^i d\mathbf{Z}_t^i, \quad \mathbf{X}_0 \in \mathbb{R}^d, \quad (1)$$

<sup>2</sup>As pointed out in Duffie et al. [5] and Errais et al. [6], it is possible to extend to the case of affine jump diffusion processes.

where  $K_0(t) \in \mathbb{R}^d$  and  $K_1(t) \in \mathbb{R}^{d \times d}$  are deterministic functions,  $\Xi^i$  is a  $d$ -dimensional diagonal matrix whose elements are non-negative, and  $\mathbf{Z}_t^i$  is a  $d$ -dimensional vector whose elements are either  $L_t^i$  or  $N_t^i$ .

This process can be regarded as a generalization of intensity process for the mutually exciting point processes studied in Hawkes [8]<sup>3</sup>.

Here we assume that each event-type intensity is represented as an affine transform of the multivariate affine jump process  $\mathbf{X}_t$ . In other words, for any  $i = 1, \dots, m$ ,  $\lambda_t^i = \Lambda_0^i(t) + \Lambda_1^i(t) \cdot \mathbf{X}_t$  for some deterministic functions  $\Lambda_0^i(t) \in \mathbb{R}$  and  $\Lambda_1^i(t) \in \mathbb{R}^d$ . ( $\mathbf{x} \cdot \mathbf{y}$  means the inner product of  $\mathbf{x}$  and  $\mathbf{y}$ .) Of course, we remark if  $\lambda_t^i$  becomes non-negative for any  $t \geq 0$ .

Such modeling has several remarkable characteristics. First, each event intensity moves deterministically in time between event times, while the intensity may jump at some event happening according to the components of the diagonal matrix  $\Xi^i$ . Second, the jump of one intensity can be caused by happenings of not only the target event itself but another events<sup>4</sup>. In this sense, the model reflects contagious rating transitions in a little more general form than the model of Errais et al. [6].

## 2.2 Likelihood function of parameters in point processes

We use the same measure change argument as Azizpour and Giesecke [2] and so on to obtain the likelihood function for the contagious event intensity model introduced in the previous subsection. From a similar proof of Proposition 3.1 and 3.2 in Azizpour and Giesecke [2], the next proposition follows.

**Proposition 2.1.** *Let  $\mathbf{X}_t$  be the  $d$ -dimensional process following (1).*

*Suppose that and that  $\Lambda^i(\mathbf{X}_t)$  for a deterministic function  $\Lambda^i : \mathbb{R}^d \rightarrow \mathbb{R}_+$  is the intensity process of the point process  $N_t^i$ , that is,  $N_t^i - \int_0^t \Lambda^i(\mathbf{X}_s) ds$  is an  $(\mathcal{F}_t)$ -martingale for any  $i = 1, \dots, m$ .*

*Define the process  $\rho_t^*$  by*

$$\rho_t^* = \exp\left(-\int_0^t \sum_{i=1}^m \log(\Lambda^i(\mathbf{X}_{s-})) dN_s^i - \int_0^t \sum_{i=1}^m (1 - \Lambda^i(\mathbf{X}_s)) ds\right), \quad (2)$$

*and assume that  $E[\rho_t^*] = 1$  for any  $t \in [0, T]$ . (In other words,  $\{\rho_t^*\}_{t \in [0, T]}$  is a true martingale under  $P$ .)*

*Let  $P^*$  be a probability measure equivalent to  $P$  such that the Radon-Nikodym density process of  $P^*$  with respect to  $P$  is given by  $\rho_t^*$ , that is,  $\rho_t^* = E\left[\frac{dP^*}{dP} \middle| \mathcal{F}_t\right]$ .*

*Then  $N_t^i - t$  ( $i = 1, \dots, m$ ) are  $P^*$ -martingales. This implies that each  $N_t^i$  is a  $P^*$ -standard Poisson process, that is, for any  $i$  and  $t < s$ ,  $P^*(N_s^i - N_t^i = 0 | \mathcal{F}_t) = e^{-(s-t)}$ .*

The proof is given in Appendix.

From now on, we discuss the specification of the likelihood function for our model. Let  $(\tilde{\tau}^i, \tilde{\eta}^i) := \{(\tilde{\tau}_k^i, \tilde{\eta}_k^i)\}_{k=1, \dots, \tilde{N}_T^i}$  be a set of observed samples during the period  $[0, T]$  of the marked

<sup>3</sup>Hawkes [8] assumes the stationarity on point processes.

<sup>4</sup>Conversely, the target event may not influence its own event intensity.

point process  $Z_t^i$  for  $i = 1, \dots, m$ . In a word,  $(\tilde{\tau}_k^i, \tilde{\eta}_k^i)$  means that the  $k$ -th type  $i$  event happened at time  $\tilde{\tau}_k^i$  and that the corresponding quantity  $\tilde{\eta}_k^i$  was observed at the same time.

Denote by  $\mathcal{L}(\Theta | \{\tilde{\tau}^i, \tilde{\eta}^i\}_{i=1, \dots, m})$  the likelihood function of the parameter set  $\Theta$  contained in  $\mathbf{X}_t$ .

By abuse of notation, let  $dP^* := \rho_T^*(\Theta) dP_\Theta$ . Then we have

$$\begin{aligned} \mathcal{L}(\Theta | \{\tilde{\tau}^i, \tilde{\eta}^i\}_{i=1, \dots, m}) &= P_\Theta(\{\tilde{\tau}^i, \tilde{\eta}^i\}_{i=1, \dots, m}) \\ &= E^* \left[ \frac{1}{\rho_T^*(\Theta)} \middle| \{\tilde{\tau}^i, \tilde{\eta}^i\}_{i=1, \dots, m} \right] P^*(\{\tilde{\tau}^i, \tilde{\eta}^i\}_{i=1, \dots, m}) \\ &= \frac{P^*(\{\tilde{\tau}^i, \tilde{\eta}^i\}_{i=1, \dots, m})}{\rho_T^*(\Theta; \{\tilde{\tau}^i, \tilde{\eta}^i\}_{i=1, \dots, m})} \end{aligned}$$

The last equality follows from the property that  $\rho_T^*(\Theta)$  becomes apparent if  $\{\tilde{\tau}^i, \tilde{\eta}^i\}_{i=1, \dots, m}$  are given.

Hence we have

$$\begin{aligned} \mathcal{L}(\Theta | \{\tilde{\tau}^i, \tilde{\eta}^i\}_{i=1, \dots, m}) &\propto 1/\rho_T^*(\Theta; \{\tilde{\tau}^i, \tilde{\eta}^i\}_{i=1, \dots, m}) \\ &\propto \exp \left( \int_0^T \sum_{i=1}^m \log(\Lambda^i(\tilde{\mathbf{X}}_{s-}^\Theta)) d\tilde{N}_s^i - \int_0^T \sum_{i=1}^m (1 - \Lambda^i(\tilde{\mathbf{X}}_s^\Theta)) ds \right) \\ &\propto \prod_{i=1}^m \exp \left( \int_0^T \log(\Lambda^i(\tilde{\mathbf{X}}_{s-}^\Theta)) d\tilde{N}_s^i - \int_0^T \Lambda^i(\tilde{\mathbf{X}}_s^\Theta) ds \right). \end{aligned} \quad (3)$$

Here  $\tilde{\mathbf{X}}_s^\Theta$  represents the sample path of  $\mathbf{X}$  constructed via (1) given  $\{\tilde{\tau}^i, \tilde{\eta}^i\}_{i=1, \dots, m}$ . The symbol  $A \propto B$  if and only if  $A$  is proportional to  $B$ .

The last expression implies that if  $\Lambda^i(\tilde{\mathbf{X}}_s^\Theta)$ , for each  $i$ , depends only upon a sub parameter set  $\Theta^i$  that differs from one event to another, each parameter set  $\Theta^i$  can be obtained by using maximum likelihood estimation separately on  $i$ .

### 2.3 Conditional expectation of point processes with multivariate affine jump intensity

As discussed in section 4, it is useful in some applications to achieve the explicit representation of  $E[Z_T^i | \mathcal{F}_t]$  ( $0 \leq t \leq T$ ), where either  $Z_T^i = L_T^i$  or  $Z_T^i = N_T^i$ . We note that in consideration of the pair  $(\mathbf{X}_t, \mathbf{Z}_t)$ , the model introduced in the last subsection fulfills the conditions of affine jump (diffusion) process studied in Duffie et al. [5].

For a rigorous argument about the distribution as well as the moments of the point processes, we usually need to discuss the characteristic functions of  $Z_T^i$  at the beginning. However, we need only the conditional expectations such as  $E[Z_T^i | \mathcal{F}_t]$  in this study, so we just mention the following proposition, a version of Corollary A.3 in Errais et al. [6].

**Proposition 2.2** (A version of Corollary A.3. in Errais et al. [6]). *Suppose that  $\mathbf{X}_t$  is a vector of affine jump processes satisfying (1) and that  $\mathbf{Z}_t$  is a pure jump process vector whose jump*

intensity is given by  $\Lambda_0^i(t) + \Lambda_1^i(t) \cdot \mathbf{X}_t$  for some deterministic functions  $\Lambda_0^i(t) \in \mathbb{R}$  and  $\Lambda_1^i(t) \in \mathbb{R}^d$ . In addition, let  $\mathbf{Y}_t := {}^t(\mathbf{X}_t, \mathbf{Z}_t)$ .

Moreover,  $\boldsymbol{\eta}_{\text{mean}}^i$  denotes a  $(d+m)$ -dimensional column vector whose elements are all the average of jump size of  $Z_t^i$ . For example,  $\boldsymbol{\eta}_{\text{mean}}^i = \sum_{k \geq 1} kP(\eta^i = k)$  if  $Z_t^i = L_t^i$  and  $\boldsymbol{\eta}_{\text{mean}}^i = 1$  if  $Z_t^i = N_t^i$ .

Then for any  $i \in \{1, \dots, m\}$  and  $T \geq t$ , we have

$$E[Z_T^i | \mathcal{F}_t] = A_{Z^i}(t, T) + B_{Z^i}(t, T) \cdot \mathbf{Y}_t.$$

Here  $A_{Z^i}(t, T) \in \mathbb{R}$  and  $B_{Z^i}(t, T) \in \mathbb{R}^{d+m}$  are given as follows:

$$B_{Z^i}(t, T) = \exp\left(\int_t^T \left[ \begin{pmatrix} {}^tK_1(s) & \mathbf{0}_{d \times m} \\ \mathbf{0}_{m \times d} & \mathbf{0}_{m \times m} \end{pmatrix} + \sum_{i=1}^m \left\{ \begin{pmatrix} \Lambda_1^i(s) \\ \mathbf{0}_m \end{pmatrix} {}^t\boldsymbol{\eta}_{\text{mean}}^i \right\} \begin{pmatrix} \Xi^i & \mathbf{0}_{d \times m} \\ \mathbf{0}_{m \times d} & U_{m \times m}^i \end{pmatrix} \right] ds\right) \mathbf{e}_{d+i}, \quad (4)$$

$$A_{Z^i}(t, T) = \int_t^T \left\{ \begin{pmatrix} K_0(s) \\ \mathbf{0}_{m'} \end{pmatrix} + \sum_{i=1}^m \Lambda_0^i(s) \begin{pmatrix} \Xi^i & \mathbf{0}_{d \times m} \\ \mathbf{0}_{m \times d} & U_{m \times m}^i \end{pmatrix} \boldsymbol{\eta}_{\text{mean}}^i \right\} \cdot B_{Z^i}(s, T) ds, \quad (5)$$

where  $U_{n \times n}^i$  is an  $n$ -dimensional diagonal matrix such that the element corresponding to  $Z^i$  is equal to one and the other elements are all zero,  $\mathbf{0}_n$  (resp.  $\mathbf{0}_{n \times n'}$ ) is an  $n$ -dimensional zero vector (resp.  $n \times n'$ -zero matrix), and  $\mathbf{e}_k$  is a  $(d+m)$ -dimensional vector such that  $k$ -th element is equal to one and the others are zero.

Refer to Proposition A.2 and Corollary A.3 in Errais et al. [6] for the proof and the related discussion.

### 3 Maximum Likelihood Estimation

In this section, we specify the event intensity model and tentatively try to estimate the model parameters from the actual data on rating change events of Japanese enterprises.

#### 3.1 The data

The data consists of issuer rating change events of Japanese enterprises from April 1998 to September 2008 by Rating and Investment Information, Inc.(R&I), which is the largest rating agency to assign the rating to Japanese enterprises<sup>5</sup>. Several events can occur at the same day, but it is impossible to identify the exact time when the events happen.

Figure 1 shows the trajectory of monthly numbers of downgrades, upgrades and the other events of R& I during April 1998 to September 2008. We can observe that there are more

<sup>5</sup>The data source is Bloomberg (the command "RATC" ). Strictly speaking, this study tries capturing the characteristics or the bias of the special agency of R&I rather than genuine (or objective) creditworthiness. However the agency seems getting responsible for the rating it has given, so we think the rating changes of R&I can pursue the time change of genuine creditworthiness to some extent.

In addition, it may be more suitable and useful to estimate the credit risk of each bond issuer according to the rating policy publicly announced by R&I. Nonetheless, it is difficult to capture the dependence and to measure the risk for at the portfolio level through the bottom-up approach.

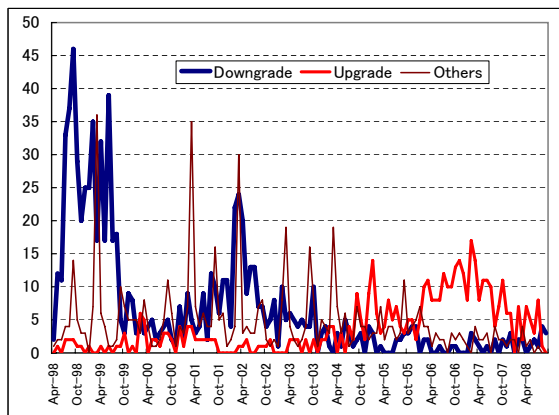


Figure 1: Trajectory of monthly numbers of downgrades, upgrades and the other events of R&I during April 1998 to September 2008.

downgrades from May 1998 to August 1999 and the first half in 2002 than other months while the number of upgrades remains relatively high since 2006.

During the above period, there are 848 downgrades, 467 upgrades, and 613 other events<sup>6</sup>. Among 848 downgrades, 159 is the number of downgrades from the rating over or equal to "A-" to the rating under or equal to "BBB+", which are regarded as the target downgrade for the payoff of MDP considered later.

Suppose that the initial date ( $t = 0$ ) of the data is April 1, 1998. In estimation, we adopt not the calendar time but a kind of business day-count convention, namely, the difference between any business days is set  $\Delta t := 1/250$ . The first downgrade happened on April 6, 1998 ( $3\Delta t$ ), the first upgrade on May 29, 1998 ( $39\Delta t$ ), and the first other event on April 28, 1998 ( $20\Delta t$ ). The final date of the data is September 30, 2008 and consequently  $T = 2584\Delta t$ .

### 3.2 Specification of the intensity model and the likelihood function

We try to estimate the parameters of the intensity model specified below from the data that consist of the sequence of event date as well as the number of simultaneous events. For the purpose, we apply the maximum likelihood method proposed in Azizpour and Giesecke [2].

Now, we specify the mutually exciting event intensity model introduced in Section 2. Let  $m = d = 3$ . Hereafter, we regard  $N_t^1, N_t^2$  and  $N_t^3$  as the counting process of the times when happen some downgrades, upgrades and the other events, respectively. Moreover we view  $L_t^i$  as the process of the cumulative number of type  $i$  events up to time  $t$ . This means that the jump size  $\eta_k^i$  is equal to the number of type  $i$  events which coincidentally happen at time  $\tau_k^i$ .

As for each event intensity, we suppose that each state process  $X_t^j$  ( $j = 1, 2, 3$ ) just becomes the intensity of the corresponding type of events, namely,  $\lambda_t^j = X_t^j$  ( $j = 1, 2, 3$ ).

<sup>6</sup>Almost all of 613 samples contains "NR(No Rating)" either in the item of Current Rating or Last Rating. Probably default or worsening of credit must be one of the main reasons for changes to NR, but we do not examine the reasons for NR. The data must be prepared for further analyses with more care.

We also assume that  $\mathbf{X}_t = {}^t(X_t^1, X_t^2, X_t^3)$  satisfies the following affine-jump type equation<sup>7</sup>.

$$\begin{pmatrix} dX_t^1 \\ dX_t^2 \\ dX_t^3 \end{pmatrix} = \begin{pmatrix} \kappa^1(c^1 - X_t^1) \\ \kappa^2(c^2 - X_t^2) \\ \kappa^3(c^3 - X_t^3) \end{pmatrix} dt + \sum_{i=1}^3 \begin{pmatrix} \xi^{1,i} \\ \xi^{2,i} \\ \xi^{3,i} \end{pmatrix} dL_t^i, \quad (6)$$

where  $\kappa^j, c^j, \{\xi^{j,i}\}_{i=1,2,3}$  ( $j = 1, 2, 3$ ) and the initial value  $X_0^j$  are all non-negative constant parameters.

Note that  $X_t^j$  ( $j = 1, 2, 3$ ) can be represented as

$$X_t^j = c^j + e^{-\kappa^j t}(X_0^j - c^j) + \int_0^t e^{-\kappa^j(t-s)} \sum_{i=1}^3 \xi^{j,i} dL_s^i. \quad (7)$$

Moreover we remark that  $X_t^j \geq \min\{c^j, X_0^j\} \geq 0$  for any  $t \geq 0$  provided that only positive jumps are allowed for every  $L_t^i$ , that is,  $P(\eta_k^i \geq 0) = 1$  for any  $k \in \mathbb{N}$ . The above immediately follows from (7) and the assumption that all the parameters are non negative.

As mentioned before, we apply the maximum likelihood method proposed in Azizpour and Giesecke [2] to estimate the model parameters denoted by  $\Theta^j := (X_0^j, \kappa^j, c^j, \xi^{j,i})_{i=1,2,3}$  ( $j = 1, 2, 3$ ). Let us remark what are different from the zero-factor model of Azizpour and Giesecke [2]. We pay attention to more than one events while Azizpour and Giesecke [2] see only the default events in the portfolio. Moreover, in our model the formulation of impact of events to the intensity is proportional to the number of events per day while the impact is represented in terms of the quadratic function of the number of events per day. However this is not an essential difference to estimate the parameters since we can indeed estimate the parameters separately for each  $j = 1, 2, 3$ . Furthermore, we estimate the initial value  $X_0^j$  and the long-term average  $c^j$  separately while Azizpour and Giesecke [2] assumed  $X_0^j = c^j$ .

Now we move to the likelihood function for our model. In subsection 2.2, we see that the likelihood function can be obtained like (3) in terms of the Radon-Nikodym density  $\rho_t^*$  defined in (2). Indeed, we need to check that  $E[\rho_t^*] = 1$  for the case where  $\mathbf{X}$  is specified by (6) and  $\mathbf{X}$  itself is the intensity vector, before applying (3) to our case. Some known sufficient conditions<sup>8</sup> for  $E[\rho_t^*] = 1$  are too strong to check it directly.

If both  $X_0^j$  and  $c^j$  are strictly positive, it follows from the above remark that  $X_t^j$  is bounded below away from zero. Also, if we implicitly suppose that the true intensity is given by  $\min\{X_t^j, K^j\}$  for some finite (but sufficiently large) constant  $K^j$ ,  $X_t^j$  is also bounded from above. Then we have  $E[\rho_t^*] = 1$  due to the bounded (true) intensity. The estimation from the actual data must not be influenced by the upper bound  $K^j$  that is sufficiently large, so we can ignore the upper bound at estimation.

<sup>7</sup>We may express the last term in the matrix form according to the general model (1).

<sup>8</sup>For example, see Lemma 19.6 in Liptser and Shiryaev [11]. In this case, one of the sufficient conditions is that there exists some constant  $K < \infty$  such that  $P$ -a.s.

$$\int_0^T \left( \frac{1}{X_s^j} - 1 \right)^2 X_s^j ds \leq K.$$

Once applying (3) is justified, we have that the likelihood function for (6) can be specified by

$$\mathcal{L}(\{\Theta^j\}_{j=1,2,3} | (\tilde{\tau}, \tilde{\eta})) \propto \prod_{j=1}^3 \exp\left(\int_0^T \log(\tilde{X}_{s-}^{j,\Theta^j}) d\tilde{N}_s^j - \int_0^T \tilde{X}_s^{j,\Theta^j} ds\right) \quad (8)$$

where  $\tilde{X}_t^{j,\Theta^j}$  is a sample path of  $X_t^j$  if the observations  $(\tilde{\tau}, \tilde{\eta})$  are given. This implies that each parameter set  $\Theta^j$  can be obtained separately on  $j = 1, 2, 3$ .

We know that maximizing the log-likelihood function  $l(\Theta^j | (\tilde{\tau}, \tilde{\eta})) := \log \mathcal{L}(\Theta^j | (\tilde{\tau}, \tilde{\eta}))$  is equivalent to maximizing

$$\int_0^T \log(\tilde{X}_{s-}^{j,\Theta^j}) d\tilde{N}_s^j - \int_0^T \tilde{X}_s^{j,\Theta^j} ds.$$

Using (7), we can represent the objective function to be maximized for achieving the maximum likelihood estimates of  $\Theta^j$  as follows.

$$\begin{aligned} & \sum_{k=1}^{\tilde{N}_t^j} \log \left\{ c^j + e^{-\kappa^j \tilde{\tau}_k^j} (X_0^j - c^j) + \sum_{i=1}^3 \xi^{j,i} \sum_{\tilde{\tau}_p^i < \tilde{\tau}_k^j} \tilde{\eta}_p^i e^{-\kappa^j (\tilde{\tau}_k^j - \tilde{\tau}_p^i)} \right\} \\ & - c^j T - \frac{X_0^j - c^j}{\kappa^j} (1 - e^{-\kappa^j T}) - \frac{1}{\kappa^j} \sum_{i=1}^3 \xi^{j,i} \sum_{k=1}^{\tilde{N}_t^i} \tilde{\eta}_k^i (1 - e^{-\kappa^j (T - \tilde{\tau}_k^i)}). \end{aligned} \quad (9)$$

### 3.3 Estimation result

We execute the maximum likelihood estimation of the parameters with the free statistical software package R. Specifically we use an intrinsic function `optim`<sup>9</sup> to maximize the objective function (9)<sup>10</sup>. We try the maximization for several different sets of initial values, and we finally choose the estimates that maximize the objective function among all the trials. Since we have no idea on any rigorous way to compute the standard error of the estimates in our non-stationary model, we regard as tentative standard errors the square roots of diagonal elements of the inverse of Hessian that is an output from the function `optim` of R.

The results of such estimation are given in Table 1. Tentative standard errors are given in parentheses.

(NA) stands for that it is impossible to obtain their square root since some diagonal elements of the inverse of Hessian turn out to be negative.

The optimization is considered dependable to some extent, since `optim` of R returns the success of convergence and the result is intuitively acceptable. We can observe some mutually exciting effects between downgrades and the other events in the sense that the estimates of  $\xi^{1,3}$  and  $\xi^{3,1}$  are positive. The estimate of  $X_0^1$  may seem too high, but it can be admissible in terms of the Japanese economic situation in 1998. Moreover, note that the estimate of  $c^3$  is zero.

<sup>9</sup>`optim(initial_values, obj_fun, method = "L-BFGS-B", lower = numeric(6), control=list(fnscale=-1), hessian=TRUE)`

<sup>10</sup>Azizpour and Giesecke [2] write “We perform a grid search over the discretized parameter space to solve the likelihood problem ...”.

Downgrade	$X_0^1$	$\kappa^1$	$c^1$	$\xi^{1,1}$	$\xi^{1,2}$	$\xi^{1,3}$
	57.44 (25.90)	4.74 (1.05)	7.48 (5.58)	1.87 (0.47)	0 (0.24)	0.31 (0.31)
Upgrade	$X_0^2$	$\kappa^2$	$c^2$	$\xi^{2,1}$	$\xi^{2,2}$	$\xi^{2,3}$
	4.80 (9.42)	3.38 (0.96)	6.79 (5.95)	0 (0.06)	2.03 (0.62)	0 (0.20)
The other events	$X_0^3$	$\kappa^3$	$c^3$	$\xi^{3,1}$	$\xi^{3,2}$	$\xi^{3,3}$
	26.98 (8.56)	0.30 (NA)	0 (7.32)	0.07 (0.04)	0 (NA)	0.06 (NA)

Table 1: The maximum likelihood estimates of each model parameter. The standard errors (the square roots of diagonal elements of the inverse of Hessian matrix that is numerically computed) are given in parentheses. NA means that the corresponding square root is impossible to compute. As the initial values  $X_0^j = 10$ ,  $\kappa^j = 6$ ,  $c^j = 2$ ,  $\xi^{j,i} = 0.3$  ( $\forall i, j$ )

This implies that only a happening of downgrade and the other event itself can cause successive irregular events like redemption, withdrawal of rating and new issue. It is seemingly strange, but it may be rational since such other events as redemption and withdrawal of rating are often caused by debasement of creditworthiness, that is, it looks natural to identify some other events with downgrades.

In Figure 2, we show the estimated paths of the intensities of downgrades, upgrades and the other events with the parameter estimates in Table 1. These paths seem quite consistent with the actual trajectories of each monthly event counts as seen in the left panel of Figure 1.

However we should strongly remark that the tentative standard errors are in whole large so that it is hard to be statistically confident about the significance of the most estimates. As goodness-of-fit tests such as Kolmogorov-Smirnov test and Prah’s test proposed in Azizpour and Giesecke [2], rejected is the hypothesis that event times after some time change follows independent and identically standard exponentially distributed. Anyway the model leaves much room for discussion in the procedure of estimation. More elaborate studies on the parameter estimation of the contagious rating change model will be described on another paper.

## 4 An Application: Valuation of Multi-Downgrade Protection

In this section, we formulate a new derivative named multi-downgrade protection (MDP) that pays off every time the target downgrade occurs in the underlying portfolio. Then we discuss a risk-neutral valuation of the protection leg of MDP under some tentative assumptions on the payoff form.

### 4.1 Multi-Downgrade Protection

In general, a regular downgrade protection is the contract that the protection seller pays to the buyer the amount according to the agreed rule when the rating of the particular reference

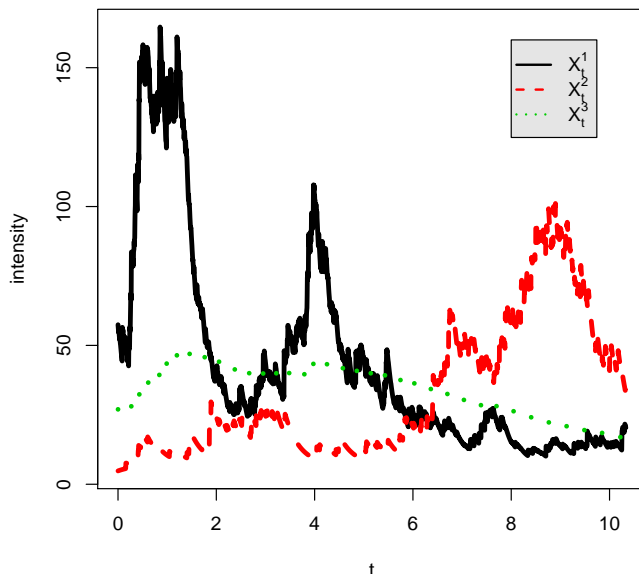


Figure 2: The estimated intensities of downgrades, upgrades and the other events: the estimates in Table 1 are used as the model parameters.

asset (usually corporate bond), assigned by some external rating agency such as Moody's, Standard & Poors(S&P), Rating and Investment Information(R&I) and so on, falls down below the predetermined threshold grade. In other words, the downgrade protection is a kind of credit derivatives in the sense that the payoff is dependent upon occurrence of some specific obligor's downgrade, which means that the obligor is estimated to be more likely to default than before. The main problem is how to compute the fair premium that the protection buyer should pay to the seller in return for the downgrade protection<sup>11</sup>.

The downgrade protection has not been so a popular financial instrument, but we guess that it has several *raison d'être*. Some Japanese institutional investors aim to track the benchmark index called NOMURA-BPI. The benchmark index is made up of only the corporate bonds with rating higher than or equal to "A". Therefore the passive investors that aim for accomplishment of the profit-loss performance close to NOMURA-BPI have to make the tracking portfolio with only the A-rated or the higher rated bonds.

If a bond in the tracking portfolio falls down below "A", many passive investors try to sell the downgraded bond at a time, and thus decline of the bond price may cause much loss on sale of the bond. The downgrade protection seems one of useful hedging tools for such a downgrade risk of the bond. The payoff trigger of most credit derivatives is the default event that directly makes an influence on the portfolio value, but the decrease of the bond value because of downgrade should be ignored for portfolio management.

As for the payoff of the downgrade protection, Aonuma [1] discusses a predetermined constant payoff independent of the time when the target downgrade happens, while Shimizu [14] studies

<sup>11</sup>The discussion below is easily extended to the swap-type protection that obliges the protection buyer to pay the periodic constant premium to the seller.

the payoff dependent on the difference of the bond price between just before and after downgrade in an interest rate tree of Hull-White model.

Now consider a new derivative named multi-downgrade protection (MDP). MDP is specified by the contract that pays off some amount according to the predetermined rule every time the target downgrade (for example, from the investment grades to the speculative grades) occurs in the underlying portfolio, independent of the individual name downgraded.

Although it is reasonable for reducing the hedging cost to protect only the particular corporate bonds that are expected to be downgraded, the protection will result in vain if the underlying bond keeps the same grade and another bonds without protection are downgraded. On the other hand, MDP costs may be rather costly, but it does not matter which bond is downgraded. For the purpose of managing the whole (large) portfolio, it looks more important to consider how many downgrades will happen rather than which bond will be downgraded. Therefore MDP can be a useful tool to manage the downgrade risk in the portfolio that consists of a wide variety and number of corporate bonds if the product structure is appropriately prepared.

## 4.2 Risk-neutral valuation of MDP

Hereafter, we use the risk-neutral valuation to evaluate a typical MDP.

Let  $N_t^*$  be the counting process of times when the target downgrade triggering the protection payoff happen, and  $L_t^*$  be the process of the cumulative number of the target downgrade up to time  $t$ . In short, we implicitly admit that more than one target downgrades can happen at the same time.

Denote by  $C_t^T$  a  $(\mathcal{F}_t)$ -predictable continuous process that stands for the protection payoff at time  $t$  if a target downgrade happens then. Let  $Q$  be a risk-neutral probability measure and fixed, and the price  $Z(t, T)$  at time  $t$  of the default-free zero-coupon bond with maturity  $T$  is given by

$$Z(t, T) = E^Q \left[ \exp \left( - \int_t^T r_u du \right) \middle| \mathcal{F}_t \right].$$

Then the fair value  $V_t^T$  at time  $t$  of a protection leg<sup>12</sup> of MDP with expiration  $T(\geq t)$  and payoff process  $C_t^T$  is specified as follows.

$$V_t^T = E^Q \left[ \int_t^T \exp \left( - \int_t^s r_u du \right) C_s^T dL_t^* \middle| \mathcal{F}_t \right].$$

---

<sup>12</sup>Some downgrade protection contracts have periodic premium convention like CDS. For such a case, we have to evaluate the premium leg of the protection. However, we think of only the protection leg with the lump-sum payment at the beginning of the contract period, since the premium leg of the protection for the case of paying periodic premiums does not have any essentially new issue for its valuation.

Using the integration-by-parts formula, we have

$$\begin{aligned}
V_t^T &= E^Q \left[ \exp\left(-\int_t^T r_u du\right) C_T L_T^* \middle| \mathcal{F}_t \right] - C_t^T L_t^* - E^Q \left[ \int_t^T L_s^* d\left\{ \exp\left(-\int_t^s r_u du\right) C_s^T \right\} \middle| \mathcal{F}_t \right] \\
&= E^Q \left[ \exp\left(-\int_t^T r_u du\right) C_T L_T^* \middle| \mathcal{F}_t \right] - C_t^T L_t^* \\
&\quad + E^Q \left[ \int_t^T L_s^* r_s \exp\left(-\int_t^s r_u du\right) C_s^T ds \middle| \mathcal{F}_t \right] - E^Q \left[ \int_t^T L_s^* \exp\left(-\int_t^s r_u du\right) dC_s^T \middle| \mathcal{F}_t \right] \quad (10)
\end{aligned}$$

For further calculation, we assume the followings.

**Assumption 4.1.** 1.  $\{r_t\}$  and  $\{L_t^*\}$  are independent.

2. For a given  $T$ ,  $\sigma_t^T$  is an  $(\mathcal{F}_t)$ -adapted positive process that represents the volatility of the default-free zero-coupon bond. Hence the price dynamics of  $Z(t, T)$  under  $Q$  is specified as

$$dZ(t, T) = Z(t, T)\{r_t dt + \sigma_t^T dW_t^Z\}, \quad Z(T, T) = 1,$$

where  $W_t^Z$  is a  $Q$ -standard Brownian motion that is independent of  $r_t$  and  $L_t^*$ .

3. For a fixed  $T$ ,  $C_t^T$  is given by  $Z(t, T)\varphi(t, T)$ , where  $\varphi(t, T)$  is an  $(\mathcal{F}_t)$ -adapted process defined by

$$\varphi(t, T) := \int_t^T E^Q[\bar{h}_u | \mathcal{F}_t] du,$$

and the process  $\bar{h}_t$  follows under  $Q$

$$d\bar{h}_t = \alpha(\beta - \bar{h}_t)dt + \sigma_h dW_t^h, \quad \bar{h}_0 > 0$$

where  $\alpha, \beta$  and  $\sigma_h$  are positive constants and  $W_t^h$  is a  $Q$ -standard Brownian motion that is independent of  $r_t, L_t^*$  and  $W_t^Z$ .

Under the third assumption, we can obtain the explicit representation of  $\varphi(t, T)$  as we see below.

As is well known as the Vasicek type interest rate model, it is easy to see that for  $s \geq t$

$$\bar{h}_s = \bar{h}_t e^{-\alpha(s-t)} + \beta(1 - e^{-\alpha(s-t)}) + \sigma_h e^{-\alpha(s-t)} \int_0^{s-t} e^{\alpha u} dW_u^h.$$

Therefore we can obtain

$$E^Q[\bar{h}_s | \mathcal{F}_t] = (\bar{h}_t - \beta)e^{-\alpha(s-t)} + \beta. \quad (11)$$

Hence

$$\varphi(t, T) = \int_t^T \{(\bar{h}_t - \beta)e^{-\alpha(u-t)} + \beta\} du = \frac{\bar{h}_t - \beta}{\alpha} (1 - e^{-\alpha(T-t)}) + \beta(T-t).$$

In addition, we can make sure the dynamics of  $\varphi(t, T)$  is given by  $\varphi(T, T) = 0$  and

$$d\varphi(t, T) = -\bar{h}_t dt + \frac{\sigma_h}{\alpha} dW_t^h.$$

At last, we can obtain

$$\begin{aligned} dC_t^T &= \varphi(t, T)dZ(t, T) + Z(t, T)d\varphi(t, T) \\ &= \{C_t^T r_t - Z(t, T)\bar{h}_t\}dt + C_t^T \sigma_t^Z dW_t^Z + Z(t, T)\frac{\sigma_h}{\alpha}dW_t^h. \end{aligned}$$

Now we can calculate (10) to a simpler representation.

$$\begin{aligned} V_t^T &= -C_t^T L_t^* + E^Q \left[ \int_t^T L_s^* r_s \exp\left(-\int_t^s r_u du\right) C_s^T ds \middle| \mathcal{F}_t \right] \\ &\quad - E^Q \left[ \int_t^T L_s^* \exp\left(-\int_t^s r_u du\right) \left[ \{C_s^T r_s - Z(s, T)\bar{h}_s\} ds + C_s^T \sigma_s^Z dW_s^Z + Z(s, T)\frac{\sigma_h}{\alpha} dW_s^h \right] \middle| \mathcal{F}_t \right] \\ &= -C_t^T L_t^* + E^Q \left[ \int_t^T L_s^* \exp\left(-\int_t^s r_u du\right) Z(s, T)\bar{h}_s ds \middle| \mathcal{F}_t \right] \\ &= -C_t^T L_t^* + Z(t, T) \int_t^T E^Q[L_s^* | \mathcal{F}_t] E^Q[\bar{h}_s | \mathcal{F}_t] ds \end{aligned} \quad (12)$$

The first equality follows from the observation  $\varphi(T, T) = 0$ , the second equality follows from the martingale property of stochastic integral with respect to Brownian motion and the cancellation of the second term and a part of the third term, and the last equality is due to the assumption of conditional independence among  $L_t^*, r_t$  and  $\bar{h}_t$  and the iterated conditioning of conditional expectations.

$E^Q[\bar{h}_s | \mathcal{F}_t]$  is given by (11), so the remaining issue to solve is how to compute  $E^Q[L_s^* | \mathcal{F}_t]$ , which is reduced to the argument in subsection 2.3.

**Remark 4.2.** The assumptions on the form of payoff in Assumption 4.1 are not essential. That is just a simple example for numerical computation, but we can mention that  $C_t^T = Z(t, T)\varphi(t, T)$  approximates the difference of the price of corporate zero-coupon bond between before and after downgrade, so such a payoff form has not a little meaning.

Let  $h_t^1$  and  $h_t^2$  be the credit spread processes for rating 1 (higher credit quality) and rating 2 (lower credit quality) respectively. Usually  $h_t^2 > h_t^1$  holds. We identify  $\bar{h}_t$  with  $h_t^2 - h_t^1$ .

Then, let us remark that  $C_t^T = Z(t, T)\varphi(t, T)$  can be regarded as a naive approximation of the difference of the price of corporate zero-coupon bond between before and after downgrade in the following sense. The assumption of independence between  $r_t$  and  $W_t^h$  enables us to proceed the calculation to the end.

$$\begin{aligned} &E^Q \left[ \exp\left(-\int_t^T \{r_u + h_u^1\} du\right) \middle| \mathcal{F}_t \right] - E^Q \left[ \exp\left(-\int_t^T \{r_u + h_u^2\} du\right) \middle| \mathcal{F}_t \right] \\ &= Z(t, T) E^Q \left[ \exp\left(-\int_t^T h_u^1 du\right) - \exp\left(-\int_t^T h_u^2 du\right) \middle| \mathcal{F}_t \right] \\ &\approx Z(t, T) E^Q \left[ \int_t^T \{h_u^2 - h_u^1\} du \middle| \mathcal{F}_t \right] \\ &= Z(t, T) \int_t^T E^Q[h_u^2 - h_u^1 | \mathcal{F}_t] du \\ &= Z(t, T)\varphi(t, T). \end{aligned}$$

Here the second approximation is due to the naive approximation  $e^{-x} \approx 1 - x$  if  $x$  is sufficiently close to zero.

### 4.3 Discussion on risk premium of MDP-target downgrade

In the previous section, we show a tentative parameter estimation from the historical data of rating change events by R&I and obtain the knowledge about the downgrade intensity under the physical probability measure. For the purpose of valuation of MDP, we need to know the MDP-target downgrade intensity under some risk-neutral probability measure.

Suppose that there exists a risk-neutral probability measure  $Q$  (equivalent to the physical measure  $P$ ) and that the MDP-target downgrade intensity is denoted by  $\lambda_t^{*,Q}$ .

In consideration of discussions in section 3 of Kusuoka [9] and in chapter 7 of Bielecki and Rutkowski [3], the Radon-Nikodim density process of  $Q$  with respect to  $P$  denoted by  $\rho_t = E\left[\frac{dQ}{dP}|\mathcal{F}_t\right]$  implies that there exist  $(-1, \infty)$ -valued predictable processes  $\beta_t^i$  ( $i = 1, 2, 3$ ) such that

$$\rho_t = 1 + \int_0^t \rho_{s-} \sum_{i=1}^3 \beta_s^i (dN_s^i - \lambda_s^i ds).$$

Immediately, it follows  $\lambda_t^{i,Q} = (1 + \beta_t^i)\lambda_t^i$ .

We have not empirically studied the quantification of risk premium of rating changes, so we have no established idea to measure  $\beta_t^i$ . As just a naive approach, it may be useful for risk premium measurement to examine the relation between the spread changes of corporate bonds before and after rating change in a portfolio.

In this paper, we assume that the risk premium of rating changes is ignored for numerical works in subsection 4.5 below. In other words, the event intensities are supposed to be unchanged under  $P$  and  $Q$ . Anyway, the quantification of risk premium of rating changes should be solved in near the future.

### 4.4 Model specification and preparation for numerical work

In this subsection, we prepare for the numerical works in the next subsection. From here to the end of this section, suppose that the intensities of downgrades, upgrades and the other events are given directly by (6).

To begin with, we specify the relation between the target downgrade and the whole downgrade. Figure 3 shows the trajectory of monthly numbers of the whole downgrades and the target downgrades (from the rating over or equal to "A-" to the rating under or equal to "BBB+") during April 1998 to September 2008. By naive observation of Figure 3, it can be permitted the assumption that the ratio of the number of target downgrades to the number of the whole downgrades is almost stable in time. Hence, let  $\theta \in (0, 1)$  is a given constant, and we suppose that  $E^Q[L_t^*] = \theta E^Q[L_t^1]$  for any  $t$ , ad hoc. Then we just interpret  $\theta$  as the conditional probability that a downgrade belongs to the target downgrade given the information about the total downgrades. In the below numerical example, we set  $\theta = 0.19$ , which is obtained as monthly average of the ratio of the target downgrade to the whole downgrade from R&I data.

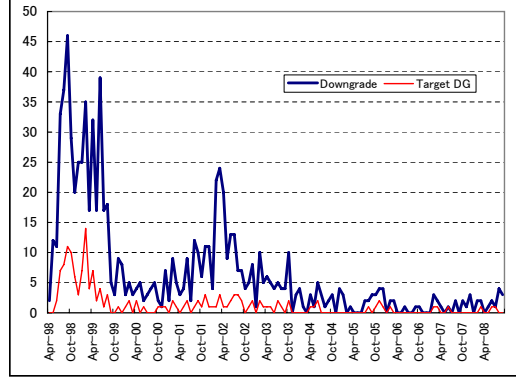


Figure 3: Trajectory of monthly numbers of the whole downgrades and the target downgrades of R& I (from the rating over or equal to "A-" to the rating under or equal to "BBB+") during April 1998 to September 2008.

Reminding us of the fair value formula (12) of MDP with the payoff  $C_t^T = Z(t, T)\varphi(t, T)$ , we can see

$$V_0^T = Z(0, T) \int_0^T E^Q[L_s^*] E^Q[\bar{h}_s] ds = \theta Z(0, T) \int_0^T E^Q[L_s^1] E^Q[\bar{h}_s] ds. \quad (13)$$

Therefore, in order to obtain the price  $V_0$ , we have to specify the price of the default-free zero-coupon bond  $Z(0, T)$ , the expectation of the substantial spread  $E^Q[\bar{h}_s]$  that determines the payoff at the target downgrades, and the expectation  $E^Q[L_s^1]$  of the cumulative number of downgrades.

Let us suppose that  $Z(0, T)$  is based on Cox-Ingersoll-Ross model, that is,

$$Z(0, T) = \left( \frac{2\gamma e^{\frac{(\gamma+\zeta)T}{2}}}{(\gamma+\zeta)(e^{\gamma T} - 1) + 2\gamma} \right)^{\frac{2\zeta r_\infty}{\sigma_r^2}} \exp\left( -\frac{2(e^{\gamma T} - 1)r_0}{(\gamma+\zeta)(e^{\gamma T} - 1) + 2\gamma} \right),$$

where  $\zeta$  is the mean-reverting speed,  $r_\infty$  is the long-term mean of  $\bar{h}$  and  $\sigma_r$  is the volatility of  $r_t$  and  $\gamma = \sqrt{\zeta^2 + 2\sigma_r^2}$ .

Also remember the expectation of the substantial spread  $E^Q[\bar{h}_s]$  that determines the payoff at the target downgrades is given by

$$E^Q[\bar{h}_s] = (\bar{h}_0 - \beta)e^{-\alpha s} + \beta,$$

where  $\alpha$  is the mean-reverting speed and  $\beta$  is the long-term mean of the spread  $\bar{h}$ .

Table 2 shows the fixed set of parameters on  $Z(0, T)$  and  $E^Q[\bar{h}_s]$  for the numerical illustration of the valuation of MDP below.

At last, we consider the method of computing the expected cumulative number of downgrades  $E[L_t^1]$  by applying Proposition 2.2 to the intensity model (6). Remark that the model (6) can

$\zeta$	$\sigma_r$	$r_0$	$r_\infty$	$\alpha$	$\beta$	$\bar{h}_0$
0.3	0.25	0.01	0.02	0.3	0.02	0.015

Table 2: The fixed set of parameters about  $Z(0, T)$  and  $E^Q[\bar{h}_s]$  for the numerical illustration below.

be obtained from the general affine jump process (1) by letting  $d = m = 3$  and

$$\begin{aligned} K_0(t) &\equiv K_0 = {}^t(\kappa^1 c^1, \kappa^2 c^2, \kappa^3 c^3), \quad K_1(t) \equiv K_1 = \text{diag}(-\kappa^1, -\kappa^2, -\kappa^3), \\ \Lambda^1 &= {}^t(1, 0, 0), \quad \Lambda^2 = {}^t(0, 1, 0), \quad \Lambda^3 = {}^t(0, 0, 1), \\ \Xi^1 &= \text{diag}(\xi^{1,1}, \xi^{2,1}, \xi^{3,1}), \quad \Xi^2 = \text{diag}(\xi^{1,2}, \xi^{2,2}, \xi^{3,2}), \quad \Xi^3 = \text{diag}(\xi^{1,3}, \xi^{2,3}, \xi^{3,3}). \end{aligned}$$

In addition, set  $\bar{\eta}^i$  as the elements of  $\boldsymbol{\eta}_{\text{mean}}^i$ .

Then Proposition 2.2 implies that

$$E[L_t^1] = A(0, t) + B(0, t) \cdot \mathbf{Y}_0 = A(0, t) + B(0, t) \cdot {}^t(X_0^1, X_0^2, X_0^3, 0, 0, 0), \quad (14)$$

where  $B(0, t)$  and  $A(0, t)$  are obtained as below.

$B(0, t)$  is specified by the product of the exponential mapping of the matrix  $H$  given by (15) below and  $\mathbf{e}_4 := {}^t(0, 0, 0, 1, 0, 0)$ .

$$H = \begin{pmatrix} -\kappa^1 + \bar{\eta}^1 \xi^{1,1} & \bar{\eta}^1 \xi^{2,1} & \bar{\eta}^1 \xi^{3,1} & 1 & 0 & 0 \\ \bar{\eta}^2 \xi^{1,2} & -\kappa^2 + \bar{\eta}^2 \xi^{2,2} & \bar{\eta}^2 \xi^{3,2} & 0 & 1 & 0 \\ \bar{\eta}^3 \xi^{1,3} & \bar{\eta}^3 \xi^{2,3} & -\kappa^3 + \bar{\eta}^3 \xi^{3,3} & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix}. \quad (15)$$

This exponential mapping  $\exp(tH)$  can be numerically calculated by using Runge-Kutta method. On the other hand, we have

$$A(0, t) = \int_0^t (\kappa^1 c^1, \kappa^2 c^2, \kappa^3 c^3, 0, 0, 0) \cdot \{\exp(uH)\mathbf{e}_4\} du. \quad (16)$$

#### 4.5 Numerical results

We give the estimates in Table 1 as the basic set of parameters for numerical illustration. As for the elements of  $\boldsymbol{\eta}_{\text{mean}}^i$ , set  $\bar{\eta}^1 = 1.84$ ,  $\bar{\eta}^2 = 1.37$ ,  $\bar{\eta}^3 = 1.74$ , which are estimated from the historical data of R& I.

Remark again that we assume that the risk premium of rating changes is ignored, namely,  $E^Q[L_t^i] = E[L_t^i]$  ( $i = 1, 2, 3$ ) although we should use the parameters estimated from not historical data but market data to evaluate the risk-neutral value of MDP, or at least we should consider the risk premium of rating changes for adjusting the intensities estimated from historical data.

Now we display some comparative results of numerical computation for different values of the parameters on the intensity of the whole downgrade; the initial value  $X_0^1$ , the long mean reverting speed  $\kappa^1$  and the mutually exciting components  $\xi^{k,l}$  ( $k \neq l$ ).

We numerically compute  $E[L_t^1]$  according to the expressions from (14) to (16) with Runge-Kutta method, and also calculate  $V_0^t$  based on (13) with  $E[L_t^1]$  obtained on ahead.

First, we change the value of the initial downgrade intensity  $X_0^1$  as the base  $X_0^1 = 57.4$  times 0.1, 0.5, 1, 2, 3, 5 to show the curves of  $E[L_t^1]$  and  $V_0^t$  for  $t \in [0, 5]$  in Figure 6. It is natural to observe that the larger the initial downgrade intensity is, the larger both the expectation of the downgrade counts and the value of MDP are. This implies that the initial target intensity is one of the most important parameters and that we should be careful about estimation of the initial downgrade intensity  $X_0^1$ . Moreover, we can see that all the curves of  $E[L_t^1]$  seem almost linear and in parallel after about one year. We conjecture that the intensity gets almost stationary as time passes under the estimated parameter set, so  $E[L_t^1]$  becomes independent of the initial intensity. As for  $V_0^t$ , all the curves looks strictly convex because, we guess, the protection payoffs at earlier stage becomes larger but less discounted as the time to maturity gets longer.

Second, we change the value of the recursion speed  $\kappa^1$  to the long-term downgrade intensity as the base  $\kappa^1 = 4.7$  times 0.6, 0.7, 0.8, 0.9, 1, 1.2 to show the curves of  $E[L_t^1]$  and  $V_0^t$  for  $t \in [0, 5]$  in Figure 5. It is natural to observe that the smaller the recursion speed is, the larger the expectation of the target downgrade counts, since the small recursion speed means that the downgrade intensity remains relatively high in our setting. Also, the curve shape of  $E[L_t^1]$  are turned from concave to convex as  $\kappa^1$  is decreasing. This implies that the level of  $\kappa^1$  determines whether the downgrade intensity is asymptotically stationary or not.

Third, we change the value of all the mutually exciting components  $\xi^{k,l}$  ( $k \neq l$ ) uniformly from 0 to 0.5 by 0.1 to show the curves of  $E[L_t^1]$  and  $V_0^t$  for  $t \in [0, 5]$  in Figure 6. As is also expected, we can observe that the larger  $\xi^{k,l}$  ( $k \neq l$ ) becomes, the more sharply the expectation of the target downgrade counts is increasing. We see the curve shape of  $E[L_t^1]$  are turned from concave to convex as  $\xi^{k,l}$  is increasing. We think that the level of  $\xi^{k,l}$  determines whether the downgrade intensity is asymptotically stationary or not.

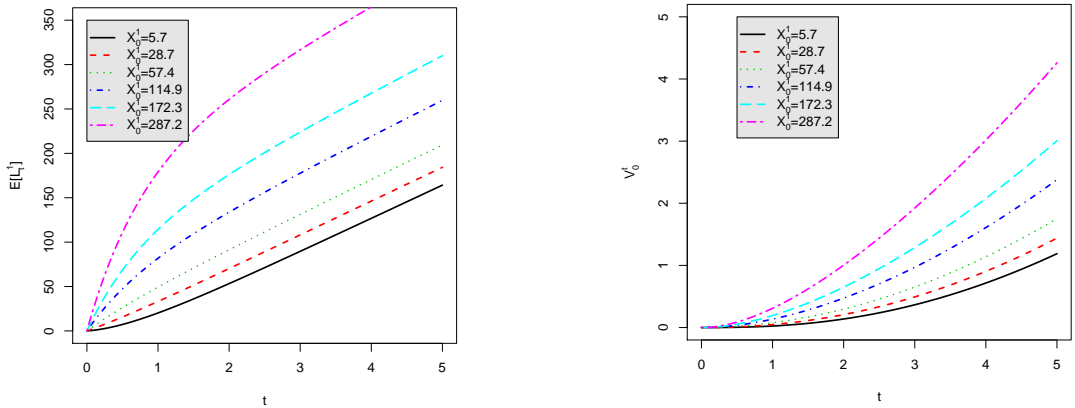


Figure 4:  $E[L_t^1]$  (on the left panel) and  $V_0^t$  (on the right panel) for different values of the initial downgrade intensity  $X_0^1$  as the base  $X_0^1 = 57.8$  times 0.1, 0.5, 1, 2, 3, 5.

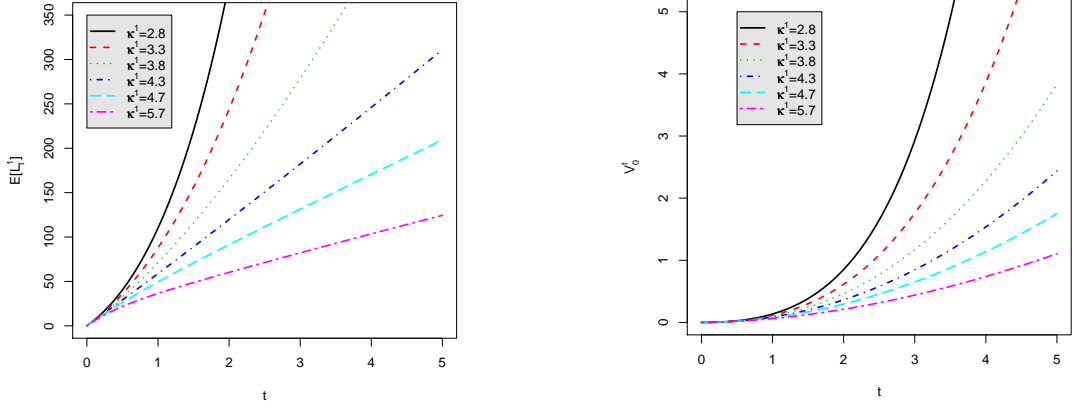


Figure 5:  $E[L_t^1]$  (on the left panel) and  $V_0^t$  (on the right panel) for different values of the reversion speed  $\kappa^1$ .

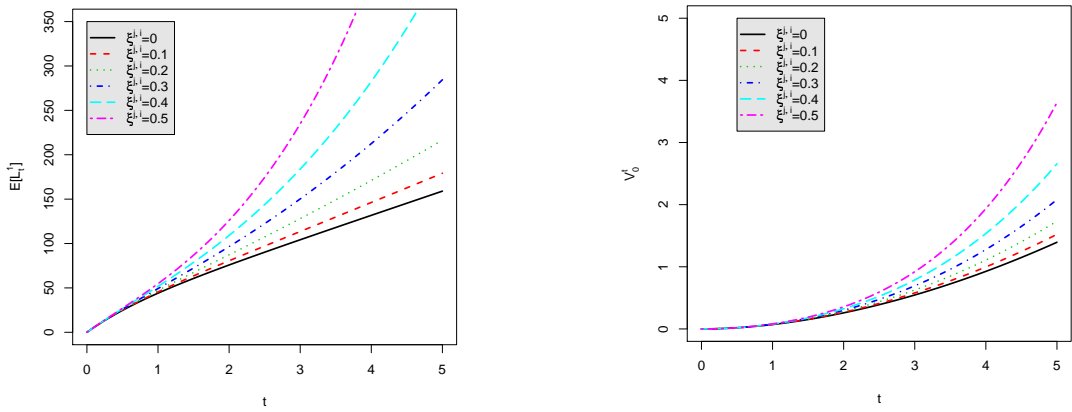


Figure 6:  $E[L_t^1]$  (on the left panel) and  $V_0^t$  (on the right panel) for different values of all the non-diagonal components  $\xi^{k,l}$  ( $k \neq l$ ) of the matrix  $\Xi^i$ .

## 5 Concluding Remarks

This paper especially focuses on construction of a mutually exciting intensity model of three typical rating change events of downgrades, upgrades and the other events. In view of feasibility of both the parameter estimation and the numerical computations on the expectation of the event counts, we utilize a multivariate affine jump process to formulate the mutually exciting intensity model of rating change events.

A specific model is empirically studied by the maximum likelihood estimation with the historical records on rating changes of Japanese enterprises. Some suggestive consequences are obtained from the estimation. For example, it is probable that there are some mutually exciting effects between downgrades and the other events. The optimization for the maximum likelihood estimates seems valid to some extent, but we have to confess that it is hard to be statistically confident about the significance of the most estimates.

In addition, we propose a new credit derivative named “multi-downgrade protection (MDP)” as an application of our mutually exciting intensity model. Some numerical illustrations on the expectation of the number of events and the fair value of MDP are also presented.

There still remain many assignments; revision of the method of processing historical data, further research about efficient parameter estimation and goodness-of-fit test, specification of downgrade risk premium, and so on.

## A The proof of Proposition 2.1

*Proof.* Note that we have

$$\rho_t^* = 1 + \int_0^t \rho_{s-}^* \sum_{i=1}^m \left( \frac{1}{\Lambda^i(\mathbf{X}_{s-})} - 1 \right) \{dN_s^i - \Lambda^i(\mathbf{X}_s)ds\}.$$

We also remark the  $P$ -conditional covariation  $\langle N^i, N^i \rangle_t = \int_0^t \Lambda^i(\mathbf{X}_s)ds$  for any  $i = 1, \dots, m$ .

As in the proof of Proposition 3.2 in the Appendix of Azizpour and Giesecke [2], the Girsanov-Meyer theorem (see Theorem 39 of section III in Protter [12]) implies that if  $M_t$  is a  $P$ -local martingale such that the  $P$ -conditional covariation  $\langle M, \rho^* \rangle$  exists, then

$$M_t - \int_0^t \frac{1}{\rho_{s-}^*} d\langle M, \rho^* \rangle_s$$

is a  $P^*$ -local martingale.

Under the assumptions that the quadratic covariation  $[N^i, N^k]_t = 0$  if  $i \neq k$  in section 2, we obtain  $\langle N^i, N^k \rangle_t = 0$  if  $i \neq k$ .

Thus we observe that letting  $M_t = N_t^i - \int_0^t \Lambda^i(\mathbf{X}_s)ds$  ( $i = 1, \dots, m$ ) leads to

$$\begin{aligned} M_t - \int_0^t \frac{1}{\rho_{s-}^*} d\langle M, \rho^* \rangle_s &= N_t^i - \int_0^t \Lambda^i(\mathbf{X}_s)ds - \int_0^t \frac{1}{\rho_{s-}^*} \cdot \rho_{s-}^* \sum_{k=1}^m \left( \frac{1}{\Lambda^i(\mathbf{X}_{s-})} - 1 \right) d\langle N^i, N^k \rangle_s \\ &= N_t^i - \int_0^t \Lambda^i(\mathbf{X}_s)ds - \int_0^t \left( \frac{1}{\Lambda^i(\mathbf{X}_{s-})} - 1 \right) \Lambda^i(\mathbf{X}_s)ds \\ &= N_t^i - t. \end{aligned}$$

□

## References

- [1] Aonuma, K.(2001), “An evaluation model for downgrade protection,” *Japan Journal of Industrial and Applied Mathematics*, **18**, 627-646
- [2] Azizpour, S. and K. Giesecke (2008), “Self-exciting corporate defaults: contagion vs. frailty,” *Working paper(Aug. 29, 2008), Stanford University*
- [3] Bielecki, T. R. and M. Rutkowski(2002), *Credit Risk: Modeling, Valuation and Hedging*, Springer
- [4] Bowsher, C. G. (2007), “Modelling security market events in continuous time: Intensity based, multivariate point process models,” *Journal of Econometrics*, **141**, 876-912
- [5] Duffie, D., J. Pan and K. Singleton(2000), “Transform analysis and asset pricing for affine jump-diffusions,” *Econometrica*, **68**(6), 1343-1376
- [6] Errais, E., K. Giesecke and L. R. Goldberg (2006), “Pricing credit from the top down with affine point processes,” *Working paper(Sep. 5, 2007), Stanford University*
- [7] Giesecke, K. and L. R. Goldberg (2005), “A top down approach to multi-name credit,” *Working paper(Jun. 6, 2008), Stanford University*
- [8] Hawkes, A. G.(1971), “Spectra of self-exciting and mutually exciting point processes,” *Biometrika*, **58**(1), 83-90
- [9] Kusuoka, S.(1999), “A Remark on default risk models,” *Advances in Mathematical Economics*, **1**, 69-82
- [10] Lando, D.(2004), *Credit risk modeling: theory and applications*, Princeton, Series in Finance
- [11] Liptser, R. S. and A. N. Shiryaev(2001), *Statistics of Random Processes I,II, 2nd. ed.*, Springer
- [12] Protter, P. E.(2004), *Stochastic integration and differential equations, 2nd. ed.*, Springer
- [13] Schönbucher, P.(2003), *Credit derivatives pricing models: models, pricing and implementation*, Wiley, UK
- [14] Shimizu, T.(2003), “Downgrade protection valuation model”, *Master thesis of Department of Industrial and Management Systems Engineering, Graduate School of Creative Science and Engineering, Waseda University (written in Japanese)*