Hitotsubashi ICS-FS Working Paper Series

FS-2014-E-001

Numerical analysis of rating transition matrix depending on latent macro factor via nonlinear particle filter method

Hidetoshi Nakagawa

Graduate School of International Corporate Strategy, Hitotsubashi University

Hideyuki Takada

Department of Information Science, Toho University

First version: June 27, 2014

Numerical analysis of rating transition matrix depending on latent macro factor via nonlinear particle filter method

Hidetoshi Nakagawa[†*]and Hideyuki Takada[†††]

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

†† *Department of Information Science, Toho University Miyama 2-2-1, Funabashi, Chiba 274-8510, Japan*

June 27, 2014

Abstract

We propose a new nonlinear filtering model for a better estimation of credit rating transition matrix consistent with the hypothesis that rating transition intensities as well as dynamics of financial asset prices depend on some unobservable macroeconomic factor. We attempt a branching particle filter method to numerically obtain the conditional distribution of the latent factor. For an illustration, we analyze a rating transition history of Japanese enterprises. As a result, we realize that our model can capture some contagion effect of credit events and an interpolative role of financial market information on the rating transition intensities.

2010 Mathematics Subject Classification codes: 91G40, 91G60.
Keywords: Credit risk, credit rating transition, nonlinear filtering, branching particle filter

1 Introduction

Credit ratings, whether they are assigned by public rating agencies or by some internal rating procedures, play a significant role in both credit risk management and defaultable debt valuation. Accordingly, it is important for risk management to assess the possibility of rating changes as well as defaults in the future as accurately as possible.

In general, a so-called cohort method, namely the maximum likelihood estimation with counting data of credit events over a long period, has been often used for estimation of credit rating transition probability matrix. With this method, various applications are possible by obtaining an infinitesimal generator matrix of the credit rating transitions. The idea behind this method is the presumption that stable and reliable rating transition probabilities can be estimated by observing actual credit events for a long period.

However, as some empirical studies imply, it seems appropriate to suppose that the rating transition probabilities fluctuate as time passes because of irregular economic cycles and so forth.

[*]hnakagawa@ics.hit-u.ac.jp
[†]Corresponding author: hideyuki.takada@is.sci.toho-u.ac.jp

Thus, it is rational to consider an estimation problem of rating transition probabilities under the assumption that such probabilities are not constant in time but dependent on some dynamic economic factors. In addition, we should notice that such economic factors would not necessarily be specified. Indeed some literature suggests that it is better to suppose that some latent factors can exist for analyzing historical data of credit events.

For example, Duffie et al. [2009] suggest a strong evidence for the presence of frailty, that is, unobservable common factors to explain default clustering observed in U.S. corporate defaults. We adhere fundamentally to the concept that the uncertainty about frailty may imply an additional correlation beyond so-called doubly stochastic framework. Therefore we start with the hypothesis that the rating transition probabilities are dependent upon some unobservable variable, which may be interpreted as some macroeconomic factor.

We construct an intensity-based rating transition model that enables us to apply the non-linear filtering methodology[1] introduced by Frey and Runggaldier [2010]. Frey and Runggaldier [2010] study the pricing of credit derivatives in reduced-form portfolio credit risk model under incomplete information and then succeed in representing information-driven default contagion via filtering, that is, successive updating of the conditional distribution of an unobservable factor in reaction to incoming default observations.

Different from Frey and Runggaldier [2010], we consider multiple rating classes including default, and then it is necessary to construct a dynamic transition probability matrix. Since it is much complicated to do with the dynamic transition probability matrix, instead we directly specify appropriate transition intensity processes so that we can avoid the embedding problem for stochastic matrices. Specifically, in order to keep tractability, we assume that transition intensities are driven by a one-dimensional Ornstein-Uhlenbeck process.

Moreover we note that we stay under the real-world measure to see how both credit events and some financial asset prices are related to dynamics of the unobservable factor, while Frey and Runggaldier [2010] consider the pricing of the credit risky products under a risk-neutral measure. Here we assume that the dynamics of some observable asset price process, such as market indexes, may depend on the same unobservable factor.

As another model of credit rating transitions with latent dynamic factors, Koopman et al. [2006] proposed the Multi-state Latent Factor Intensity (MLFI) model, in other words, an intensity-based credit rating transition model with multiple states which are driven by exogenous covariates and latent dynamic factors (one of which can be interpreted as the credit cycle). It is remarkable that their parameter estimation is based on Monte Carlo maximum likelihood method for high-dimensional integrations. Our approach is an alternative to such a maximum likelihood method. We introduce a new model that facilitates computational tractability by assuming that a single unobservable factor could grasp not only the dynamics of rating transitions but also the credit contagion effects combined with Bayesian updating via filtering.

As for the previous researches analyzing the external rating transition history of Japanese enterprises, see Nakagawa [2010], Yamanaka et al. [2011a], [2011b], [2012] for example. Different from our main purpose of estimating the rating transition intensities, they use a top-down approach for modeling upgrade and downgrade intensities to examine if there exist self-exciting effects of

[1] In another area of finance, a nonlinear filtering with point process observations is, for example, studied by Ceci and Geradi [2006] and Zeng [2003] for analysis of (ultra) high frequency trading.

upgrade or downgrade in the same industry group and mutually exciting effects among different industry groups.

Next we turn our eyes towards how to solve our nonlinear filtering problem numerically. For the purpose, we rely upon a kind of branching particle filtering algorithm, which is mentioned by Frey and Runggaldier [2010] for default event analysis, and we extend it into the case of multiple rating classes. Specifically, in our particle filtering algorithm, we need to trace the rating transition dynamics, or the variation of the distribution of the credit ratings over target firms in the sample pool.

The rest of the paper is organized as follows. Section 2 introduces our filtering model to specify a signal process and observation processes. Section 3 is devoted to derive some filtering equations and a numerical algorithm relying upon the branching particle filter. Section 4 provides preliminaries for empirical analysis based on the rating transition history of Japanese enterprises and the time series data of the Tokyo Stock Price Index (TOPIX). Numerical results and discussions are presented in section 5. Finally, section 6 mentions some concluding remarks. Appendix contains the proofs of the theorems.

2 The filtering model

This section provides the filtering model for estimating rating transition matrix depending on latent macroeconomic factor. Our filtering model is similar to (partly simpler than) that of Frey and Runggaldier [2010], but has some extension.

2.1 General setting and notation

To begin with, let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space and suppose that $\mathbb{P}$ is a real-world probability measure. We construct a credit rating transition model on this filtered probability space. Hereafter we often call an "event" for a credit rating change event including a default.

Let $M \in \mathbb{N}$ be a number of firms rated by some credit rating agency and let $\mathcal{C} = \{1, 2, \dots, M\}$ be a set of the rated firms. Also, let D be a natural number more than 2 and let $\mathcal{R} = \{1, 2, \dots, D\}$ be a set of credit ratings. Suppose that 1 is the highest credit rating such as "AAA", 2 is the second highest rating such as "AA", and so forth and that D is the default state. When we write $(\alpha, \beta) \in \mathcal{R}^2$ for an event, the first component α stands for the previous rating just before the event and the second component β stands for the current rating just after the event.

For each $i \in \mathcal{C}$, denote by $\{Y_t^i\}_{t\geq 0}$ a time-inhomogeneous Markov chain on state space $\mathcal{R}$ such that Y_t^i indicates the credit rating at time t of firm i. If a rating change with $(\alpha, \beta) \in \mathcal{R}^2$ happens to firm i at time t, it follows that $Y_{t-}^i = \alpha$ and $Y_t^i = \beta$. Thus the $\{Y_t^i\}_{t\geq 0}$ is a càdlàg process. We write $\mathbf{Y}_t$ for the vector $(Y_t^1, \dots, Y_t^M)$. Moreover we denote by Λ_t the common transition rate matrix of the Markov chain $\mathbf{Y}_t$. We also denote by T_n the n-th event time of some firm in $\mathcal{C}$ and by $\xi_n \in \mathcal{C}$ the identity of the firm corresponding to the n-th event time. In addition, the default time of firm $i \in \mathcal{C}$ is given by $\tau_i = \inf\{t > 0 : Y_t^i = D\}$. Next we define $\{X_t\}_{t\geq 0}$ by a real-valued process and will regard it as some macroeconomic factor that reflects the real-time business condition. We implicitly presume that the larger (resp. smaller) X_t is, the worse (resp. better) the business condition is. Then we assume that the transition rate matrix Λ_t of the Markov chain $\mathbf{Y}_t$ is driven by X_t, that is, specified as $\Lambda_t = (\lambda_{\alpha,\beta}(X_t))_{1\leq\alpha,\beta\leq D}$ with nonnegative functions $\{\lambda_{\alpha,\beta}(x)\}_{1\leq\alpha,\beta\leq D}$.

Moreover, we introduce another real-valued process $\{S_t\}_{t\geq 0}$ regarded as the value process of some financial asset or index such as stock price index observed in the market.

Before specifying the dynamics of the processes $\{X_t\}_{t\geq 0}$, $\{Y_t^i\}_{t\geq 0}$ $(i \in \mathcal{C})$ and $\{S_t\}_{t\geq 0}$, we define some filtrations[2]. Let $\mathcal{H}_t^i = \sigma(Y_s^i : 0 \leq s \leq t)$ for $i \in \mathcal{C}$ and $\mathcal{H}_t = \mathcal{H}_t^1 \vee \mathcal{H}_t^1 \vee \cdots \vee \mathcal{H}_t^M$. In other words, the filtration $(\mathcal{H}_t^i)_{t\geq 0}$ stands for the information about the firm i's history of credit ratings (including default) until time t and $(\mathcal{H}_t)_{t\geq 0}$ stands for the information about the history of credit ratings of the whole economy. Let $\mathcal{F}_t^X = \sigma(X_s : 0 \leq s \leq t)$ and $\mathcal{F}_t^S = \sigma(S_s : 0 \leq s \leq t)$ be the σ-algebra generated by the process $\{X_t\}_{t\geq 0}$ and $\{S_t\}_{t\geq 0}$ respectively. Also, define $\mathcal{G}_t = \mathcal{F}_t^S \vee \mathcal{H}_t$ and $\mathcal{F}_t = \mathcal{G}_t \vee \mathcal{F}_t^X$. Hence we can consider that the filtration $\mathbb{G} = (\mathcal{G}_t)_{t\geq 0}$ stands for the information available for market participants while $\mathbb{F} = (\mathcal{F}_t)_{t\geq 0}$ stands for the complete information that they are unable to utilize. We should remark that the process $\{X_t\}_{t\geq 0}$ is not $\mathbb{G}$-adapted (but $\mathbb{F}$-adapted) while $\{Y_t^i\}_{t\geq 0}$ $(i \in \mathcal{C})$ and $\{S_t\}_{t\geq 0}$ are both $\mathbb{G}$-adapted.

2.2 Signal process and observation processes

Now we specify the dynamics of the processes $\{X_t\}_{t\geq 0}$, $\{Y_t^i\}_{t\geq 0}$ $(i \in \mathcal{C})$ and $\{S_t\}_{t\geq 0}$ as follows. First, we assume that the latent macroeconomic factor process ("signal" in terms of the general filtering theory) X_t is a strong solution of the following SDE[3]:

$$dX_t = -\kappa X_t dt + c dW_t, \tag{1}$$

where $\kappa \in \mathbb{R}$, $c > 0$ and $\{W_t\}_{t\geq 0}$ is a one-dimensional $(\mathbb{P}, \mathbb{F})$-standard Brownian motion.

Second, we assume that rating transitions and the default occurred for firm i in the universe $\mathcal{C}$ is specified by[4].

$$Y_t^i = Y_0^i + \int_0^t \int_U \delta_i(X_{s-}, Y_{s-}^i, u) K_i(X_{s-}, Y_{s-}^i, u) \mathcal{N}(ds, du), \quad i \in \mathcal{C}$$

where $\delta_i(x, y, u)$ denotes a function taking values in $\{1 - y, 2 - y, \ldots, D - y\}\backslash\{0\}$, $K_i(x, y, u)$ denotes a function taking values in $\{0, 1\}$ and $\mathcal{N}(ds, du)$ denotes a $(\mathbb{P}, \mathbb{F})$- standard Poisson random measure on $(\mathbb{R}_+ \times U)$ for some Euclidean space U. Let $\nu(du)ds$ be a compensator measure of $\mathcal{N}(ds, du)$. Suppose that $\mathcal{N}(ds, du)$ is independent of the Brownian motion W_t. Let $U_i(x, y) = \{u \in U | K_i(x, y, u) \neq 0\}$. Then we also assume that $\nu(U_i(x, y) \cap U_j(x, y)) = 0$ for $i \neq j$ $(i, j \in \mathcal{C})$. This assumption implies that Y_t^i and Y_t^j has no simultaneous jumps if $i \neq j$. We remark that $\delta_i(x, y, u)$ stands for the difference between a new rating of firm i and its last rating y at the time of its rating transition and that $K_i(x, y, u)$ represents whether some event occurs for firm i. Moreover, the transition law of Y_t^i does not depend on the other firms' credit ratings.

Let $N_t^i = \sum_{n=1}^{\infty} \mathbf{1}_{\{T_n \leq t, \xi_n = i\}}$ be the counting process of firm i's events. One then has

$$N_t^i - \int_0^t \mathbf{1}_{\{Y_{s-}^i \neq D\}} \nu\left(U_i(X_{s-}, Y_{s-}^i)\right) ds$$

[2] All the filtration are supposed to satisfy the usual conditions, namely, they are supposed to be right-continuous and complete.

[3] Frey and Runggaldier [2010] consider a multi-dimensional (signal) factor that may jump according to a Poisson random measure.

[4] Frey and Runggaldier [2010] consider only default events while they suppose that one firm's default intensity can depend on other firms' defaults

$$= N_t^i - \int_0^t \int_U \mathbf{1}_{\{Y_{s-}^i \neq D\}} K_i(X_{s-}, Y_{s-}^i, u)\nu(du)ds$$
$$= \int_0^t \int_U \mathbf{1}_{\{Y_{s-}^i \neq D\}} K_i(X_{s-}, Y_{s-}^i, u)\mathcal{N}(ds, du) - \int_0^t \int_U \mathbf{1}_{\{Y_{s-}^i \neq D\}} K_i(X_{s-}, Y_{s-}^i, u)\nu(du)ds$$
$$= \int_0^t \int_U \mathbf{1}_{\{Y_{s-}^i \neq D\}} K_i(X_{s-}, Y_{s-}^i, u) \left[\mathcal{N}(ds, du) - \nu(du)ds\right].$$

Since the last term is a $\mathbb{F}$-martingale, we can identify $\nu\left(U_i(X_{t-}, Y_{t-}^i)\right)$ as the firm i's event intensity process $\lambda_{(Y_{t-}^i, \bullet)}(X_t)$.

Finally, we assume that the asset value process S_t is a strong solution of the following SDE

$$\frac{dS_t}{S_t} = \mu(X_t)dt + \sigma dB_t, \tag{2}$$

where $\mu(x)$ is a measurable function, $\sigma > 0$ and $\{B_t\}_{t \geq 0}$ is a $(\mathbb{P}, \mathbb{F})$-standard Brownian motion, independent of W_t and $\mathcal{N}(ds, du)$.

3 Filtering

We suppose that the process X_t is not directly observable for market participants and that the available information is specified by the filtration $\mathbb{G} = (\mathcal{G}_t)_{t \geq 0}$ generated by the events $\mathbf{Y}_t$ and the value process S_t of some financial asset or index. The purpose of this section is, for a given bounded function $f : \mathbb{R} \to \mathbb{R}$, to find some useful expression of the following filter

$$\pi_t(f) := \mathbb{E}_{\mathbb{P}}\left[f(X_t)|\mathcal{G}_t\right]$$

in a recursive form.

3.1 Preliminary

First we use a measure change argument from the reference measure approach in order to reduce the filter $\pi_t(f)$ with respect to $\mathcal{G}_t$ to that with respect to $\mathcal{H}_t$. For this purpose, we redefine a probability space as the product space of a space supporting $(X_t, \mathbf{Y}_t)$ and the other supporting S_t as follows. This argument is analogous to that of Frey and Runggaldier [2010]. Define a probability space $(\Omega, \mathcal{F}, \mathbb{Q})$ such that $(X_t, \mathbf{Y}_t)$ and S_t are independent under the probability measure $\mathbb{Q}$. More explicitly, $(X_t, \mathbf{Y}_t)$ are defined on the probability space $(\Omega^1, \mathcal{F}^1, \mathbb{Q}^1)$ while S_t are defined on the probability space $(\Omega^2, \mathcal{F}^2, \mathbb{Q}^2)$, so we define

$$\Omega = \Omega^1 \times \Omega^2,\ \mathcal{F} = \mathcal{F}^1 \otimes \mathcal{F}^2,\ \mathbb{Q} = \mathbb{Q}^1 \times \mathbb{Q}^2.$$

In addition, for $\omega = (\omega_1, \omega_2) \in \Omega$, we set $X_t(\omega) = X_t(\omega_1)$, $\mathbf{Y}_t(\omega) = \mathbf{Y}_t(\omega_1)$ and $S_t(\omega) = S_t(\omega_2)$. The filtrations $(\mathcal{H}_t)$ and $(\mathcal{F}_t^S)$ are redefined on $(\Omega^1, \mathcal{F}^1, \mathbb{Q}^1)$ and $(\Omega^2, \mathcal{F}^2, \mathbb{Q}^2)$ respectively. Therefore we can view the investors' filtration $\mathbb{G} = (\mathcal{G}_t) = (\mathcal{H}_t \vee \mathcal{F}_t^S)$ as the product σ-algebra $(\mathcal{H}_t \times \{\emptyset, \Omega_2\}) \vee (\{\emptyset, \Omega_1\} \times \mathcal{F}_t^S) \equiv \sigma\{A \times B \mid A \in \mathcal{H}_t,\ B \in \mathcal{F}_t^S\}$.

Then, we can define the original probability measure $\mathbb{P}$ on the space $(\Omega, \mathcal{F})$ by the following Radon-Nikodym density process

$$\left.\frac{d\mathbb{P}}{d\mathbb{Q}}\right|_{\mathcal{F}_t} = L_t(X, Z) := \exp\left(\int_0^t \frac{a(X_s)}{\sigma^2} dZ_s - \frac{1}{2}\int_0^t \left(\frac{a(X_s)}{\sigma}\right)^2 ds\right), \tag{3}$$

where $a(x) = \mu(x) - \dfrac{\sigma^2}{2}$ and $Z_t = \log \dfrac{S_t}{S_0}$ is the process such that $\dfrac{Z_t}{\sigma}$ is a standard Brownian motion under $\mathbb{Q}$. Indeed, it follows from Girsanov-Maruyama theorem that the process given by

$$B_t(\omega) = \frac{Z_t(\omega_2)}{\sigma} - \int_0^t \frac{a(X_s(\omega_1))}{\sigma} ds$$

is a standard Brownian motion under the measure $\mathbb{P}$. Hence we can easily see that S_t follows the original SDE (2) under $\mathbb{P}$. Now we have the following lemma.

Lemma 3.1. *It follows*

$$\pi_t(f) = \frac{\mathbb{E}_{\mathbb{Q}^1}[f(X_t)L_t(X,Z)|\mathcal{H}_t]}{\mathbb{E}_{\mathbb{Q}^1}[L_t(X,Z)|\mathcal{H}_t]},$$

where $L_t(X,Z)$ is given by (3)[5].

This lemma implies that calculation of the filter $\pi_t(f)$ can be reduced to the conditional expectation with respect to $\mathcal{H}_t$ under $\mathbb{Q}^1$ instead of $\mathcal{G}_t$ under $\mathbb{P}$. The readers can see the proof in Appendix A.1.

3.2 Formulas for the filter between events/at events

In this subsection we present a couple of formulas for the filter $\pi_t(f) := \mathbb{E}_{\mathbb{P}}[f(X_t)|\mathcal{G}_t]$. One is the filter at time between events, that is, at time $t \in (T_{n-1}, T_n)$ and the other is at events $\{T_n\}$. For this purpose, throughout in this subsection we focus on a fixed period $[T_{n-1}, T_n)$ from $(n-1)$-th event until n-th event. Accordingly, we denote by $\bar{X}_s := X_{s+T_{n-1}}$ $(s \in [0, T_n - T_{n-1})$, namely, the solution of the next SDE

$$\bar{X}_s = \bar{X}_0 - \int_0^s \kappa \bar{X}_v dv + c\bar{W}_s, \qquad \bar{X}_0 = X_{T_{n-1}}, \tag{4}$$

where we set $\bar{W}_s := W_{s+T_{n-1}} - W_{T_{n-1}}$. Similarly, let $\bar{Y}_s := Y_{s+T_{n-1}}$ $(s \in [0, T_n - T_{n-1})$. Each component $\bar{Y}^i_s$ $(i \in \mathcal{C})$ can be regarded as the solution of

$$\bar{Y}^i_s = \bar{Y}^i_0 + \int_0^s \int_U \delta_i(\bar{X}_{v-}, \bar{Y}^i_{v-}, u) K_i(\bar{X}_{v-}, \bar{Y}^i_{v-}, u) \bar{\mathcal{N}}(dv, du), \qquad \bar{Y}^i_0 = Y^i_{T_{n-1}},$$

where $\bar{\mathcal{N}}([0,s), U) := \mathcal{N}([0, T_{n-1}+s), U) - \mathcal{N}([0, T_{n-1}), U)$.

Thus the stopping time $\bar{T}_1 := T_n - T_{n-1}$ can be defined as the first event time of the above "bar-model". Denote by $\bar{\xi}_1 \in \{1, 2, \cdots, m\}$ the identity of the name at time $\bar{T}_1$. We also remark that the total intensity $\lambda^{\text{all}}(X_t)$ at time t of the next event for all surviving firms can be given by

$$\lambda^{\text{all}}(X_t) = \sum_{i=1}^M \mathbf{1}_{\{\tau_i > t\}} \sum_{\beta \neq Y^i_t} \lambda_{(Y^i_t, \beta)}(X_t). \tag{5}$$

[5] More exactly, since Z is regarded as the non-random path obtained in terms of the observed market value S of the asset, we should see $L_t(X,Z)$ as $L_t(X,z)|_{z=Z}$ where

$$L_t(X,z) = \exp\left(\int_0^t \frac{a(X_s)}{\sigma^2} dz_s - \frac{1}{2} \int_0^t \left(\frac{a(X_s)}{\sigma} \right)^2 ds \right).$$

As a consequence, the $\mathcal{F}_\infty^{\bar{X}}$-conditional law of $\bar{T}_1$ is obtained as below.

$$\mathbb{Q}^1_{(\bar{X}_0,\bar{\mathbf{Y}}_0)}\left(\bar{T}_1 > t | \mathcal{F}_\infty^{\bar{X}} \right) = \exp\left(-\int_0^t \lambda^{\text{all}}(\bar{X}_s) ds \right), \tag{6}$$

where $\mathbb{Q}^1_{(\bar{X}_0,\bar{\mathbf{Y}}_0)}$ stands for the joint law of $\{\bar{X}_t\}$ and $\{\bar{\mathbf{Y}}_t\}$ with initial values $(\bar{X}_0, \bar{\mathbf{Y}}_0)$ under $\mathbb{Q}^1$. Also, we mention that the density of the first event of "bar-model" is given by

$$\lim_{h\to 0} \frac{\mathbb{Q}^1_{(\bar{X}_0,\bar{\mathbf{Y}}_0)}\left(\bar{T}_1 \in (t, t+h], \bar{\xi}_1 = i, \bar{Y}_0^i = \alpha, \bar{Y}^i_{\bar{T}_1} = \beta \mid \mathcal{F}_\infty^{\bar{X}} \right)}{h} = \lambda_{(\alpha,\beta)}(\bar{X}_t) \exp\left(-\int_0^t \lambda^{\text{all}}(\bar{X}_s) ds \right). \tag{7}$$

3.2.1 The filter between events

First, we present the formula at any time between events of the filter $\pi_t(f) := \mathbb{E}_{\mathbb{P}}[f(X_t)|\mathcal{G}_t]$. As we see below, the filter at time $t \in (T_{n-1}, T_n)$ can be achieved from the distribution $\pi_{T_{n-1}}(dx)$ of the latent factor at the last event and the observed path $\{S_u\}_{T_{n-1} \le u \le t}$ of the market value of some financial asset or index.

Theorem 3.2. *Let $\mathbb{E}^{\mathbb{Q}^1}_{(x,\mathbf{y})}[\cdot]$ be the expectation operator under the law $\mathbb{Q}^1_{(x,\mathbf{y})}$. For $t \in (T_{n-1}, T_n)$, we have*

$$\pi_t(f) = \frac{\int_{\mathbb{R}} \mathbb{E}^{\mathbb{Q}^1}_{(x,\mathbf{Y}_{T_{n-1}})}\left[f(\bar{X}_{t-T_{n-1}}) L_{t-T_{n-1}}(\bar{X},\bar{Z}) \exp\left(-\int_0^{t-T_{n-1}} \lambda^{\text{all}}(\bar{X}_s) ds \right) \right] \pi_{T_{n-1}}(dx)}{\int_{\mathbb{R}} \mathbb{E}^{\mathbb{Q}^1}_{(x,\mathbf{Y}_{n-1})}\left[L_{t-T_{n-1}}(\bar{X},\bar{Z}) \exp\left(-\int_0^{t-T_{n-1}} \lambda^{\text{all}}(\bar{X}_s) ds \right) \right] \pi_{T_{n-1}}(dx)}, \tag{8}$$

where $\pi_{T_{n-1}}(dx)$ is given below in (10) in the next Theorem and the process $\{L_s(\bar{X},\bar{Z})\}_{s\ge 0}$ is defined by

$$L_s(\bar{X},\bar{Z}) \overset{def}{=} \exp\left(\int_0^s \frac{a(\bar{X}_v)}{\sigma^2} d\bar{Z}_v - \frac{1}{2}\int_0^s \left(\frac{a(\bar{X}_v)}{\sigma} \right)^2 dv \right), \quad \bar{Z}_v = \log \frac{S_{v+T_{n-1}}}{S_{T_{n-1}}}. \tag{9}$$

The proof is given in Appendix A.2.

3.2.2 The filter at some credit event

Secondly, we display the formula of the filter $\pi_{T_n}(f) := \mathbb{E}_{\mathbb{P}}[f(X_{T_n})|\mathcal{G}_{T_n}]$ at an event time T_n.

Theorem 3.3. *Let $\bar{T}_1 := T_n - T_{n-1}$. At the event time T_n, we have*

$$\pi_{T_n}(f) = \frac{\int_{\mathbb{R}} \mathbb{E}^{\mathbb{Q}^1}_{(x,\mathbf{Y}_{T_{n-1}})}\left[f\left(\bar{X}_{\bar{T}_1}\right) L_{\bar{T}_1}(\bar{X},\bar{Z}) \lambda_{(Y^{\xi_n}_{T_{n-1}}, Y^{\xi_n}_{T_n})}(\bar{X}_{\bar{T}_1}) \exp\left(-\int_0^{\bar{T}_1} \lambda^{\text{all}}(\bar{X}_s) ds \right) \right] \pi_{T_{n-1}}(dx)}{\int_{\mathbb{R}} \mathbb{E}^{\mathbb{Q}^1}_{(x,\mathbf{Y}_{T_{n-1}})}\left[L_{\bar{T}_1}(\bar{X},\bar{Z}) \lambda_{(Y^{\xi_n}_{T_{n-1}}, Y^{\xi_n}_{T_n})}(\bar{X}_{\bar{T}_1}) \exp\left(-\int_0^{\bar{T}_1} \lambda^{\text{all}}(\bar{X}_s) ds \right) \right] \pi_{T_{n-1}}(dx)}, \tag{10}$$

where $L_{\bar{T}_1}(\bar{X},\bar{Z})$ is given by (9) *with $t = \bar{T}_1$.*

The proof is given in Appendix A.3.

3.3 Algorithm with particle filter

In reality it is necessary to approximate the expected values appeared in Theorem 3.2 and 3.3 so as to numerically compute the filter $\pi_t(f)$. Broadly, particle filter is a method to approximate the conditional distribution $\mathbb{P}(X_t \in \bullet|\mathcal{G}_t)$ with some suitable discrete random measures of the form

$$\mathbb{P}(X_t \in \bullet|\mathcal{G}_t) \approx \sum_p \eta_t^p \mathbf{1}_{x_t^p}$$

with some sample points $\{x_t^p\}$ and their consistent stochastic "weights" $\{\eta_t^p\}$.

In this subsection, we summarize the numerical algorithm for computing filters $\pi_t(f)$ in case of $f(x) = x$ via a particle system as in Frey and Runggaldier [2010] and Del Moral and Miclo [2000].

Originally the particle filter algorithm presented below is introduced by Crisan and Lyons [1999] as "the minimal variance branching method[6]."

The branching particle system is constructed over equally discretized time steps $t_k = k\Delta, k \in \mathbb{N}$ with the sequence of occupation measures $\{\tilde{\pi}_{t_k}\}_{k=1,2,\cdots}$ approximating the conditional distributions $\pi_{X_{t_k}|\mathcal{G}_{t_k}}$ for each time step. As the filtering equations (8) and (10) are represented in recursive form, the occupation measure $\tilde{\pi}_{t_k}$ is computed from $\tilde{\pi}_{t_{k-1}}$ and similar procedures are repeated for subsequent time steps. Let $\boldsymbol{x}_k = (x_k^1, x_k^2, \cdots, x_k^p, \cdots, x_k^{n_k})$ denotes the set of n_k particles at time t_k living in the state space of X. Roughly speaking, the branching particle filter is constructed by a two-stage procedure. In the *prediction stage*, for each particle x_k^p, one generates a trajectory $(X_s^p)_{t_k \le s \le t_{k+1}}$ governed by the SDE (4). In the *updating stage*, which is further divided into two kinds of procedures depending on whether an event occurred or not, the new particle system is constructed by letting each particle branch into a random number of offspring. The specific particle filter algorithm is described below.

The convergence results are discussed in Bain and Crisan [2008].

In order to describe the variation of the credit rating distribution, we denote by $M_\alpha(t), \alpha \in \mathcal{R}\backslash\{D\}$ the number of firms belonging at time t to the rating α so $\sum_{\alpha=1}^{D-1} M_\alpha(0) = M$ holds.

Algorithm 3.4 (Branching Particle System)**.** *In order to derive the discrete filter distribution* $\{\tilde{\pi}_{t_k}\}_{k=0,1,\ldots}$*, we have to follow the several steps:*

Step 0. *Initialization*

Set the number n_0 of initial particles, a discretized time step size Δ and the number $\{M_\alpha(0)\}_{\alpha\in\mathcal{R}\backslash\{D\}}$ of the firms for each credit rating at initial time.

Step 1. *Initial discrete distribution*

The initial discrete distribution $\tilde{\pi}_0$ is given by the occupation measure of n_0 particles of mass $1/n_0$, that is, $\tilde{\pi}_0 = \frac{1}{n_0}\sum_{i=1}^{n_0} \delta_{x^i(0)}$. Here $\boldsymbol{x}_0 = (x^1(0), x^2(0), \cdots, x^{n_0}(0))$ represents independent draws from the initial distribution $\pi_0 := \mathbb{P}(X_0 \in \bullet)$.

Step 2. *Prediction stage*

[6] In Budhiraja et al. [2007], the algorithm is called "variance reduction scheme: a branching particle filter."

Given the particles $\boldsymbol{x}_k = (x_k^1, x_k^2, \cdots, x_k^{n_k})$ *at time* t_k*, generate* n_k *independent trajectories* $\big(\bar{\mathbf{X}}_s\big)_{0\le s\le\Delta} = \Big(\big(\bar{X}_s^1\big)_{0\le s\le\Delta}, \big(\bar{X}_s^2\big)_{0\le s\le\Delta}, \cdots, \big(\bar{X}_s^{n_k}\big)_{0\le s\le\Delta}\Big)$ *such that, for each* $p \in \{1, \cdots, n_k\}$*,* $\bar{X}_s^p$ *starts from* x_k^p *at time 0 and then follows the SDE of* X_t *given by* (1).
Thus the estimated total intensity $\{\lambda^{\mathrm{all}}(\bar{X}_s^p)\}_{0\le s\le\Delta}$*, defined by* (5)*, can be obtained of the form*

$$\lambda^{\mathrm{all}}(\bar{X}_s^p) = \sum_{\alpha\in\mathcal{R}\setminus\{D\}} M_\alpha(t_k) \sum_{\beta\neq\alpha} \lambda_{(\alpha,\beta)}(\bar{X}_s^p).$$

Step 3. <u>*Updating stage*</u>

Given the new observation $(S_s)_{t_k\le s\le t_{k+1}}$*, or* $\bar{Z}_s = \log\frac{S_{s+t_k}}{S_{t_k}}$*, compute* $\tilde{w}_{k+1}^p$ *for each* $\bar{X}_s^p$ *obtained in the last prediction stage as*

$$\begin{aligned}\tilde{w}_{k+1}^p &= L_{t_{k+1}}(\bar{X}^p, \bar{Z}) \exp\left(-\int_0^\Delta \lambda^{\mathrm{all}}(\bar{X}_s^p) ds\right)\\ &= \exp\left(\int_0^\Delta \frac{a(\bar{X}_s^p)}{\sigma^2} d\bar{Z}_s - \int_0^\Delta \left\{\frac{1}{2}\Big(\frac{a(\bar{X}_s^p)}{\sigma}\Big)^2 + \lambda^{\mathrm{all}}(\bar{X}_s^p)\right\} ds\right).\end{aligned}$$

(In fact, the stochastic integral of the first term is calculated by Euler-Maruyama approximation scheme with the sample trajectory obtained in the last prediction stage and the observed path of logarithm return of the asset.) Whether some event occurred during $(t_k, t_{k+1}]$ *or not, we redefine* w_{k+1}^p *as follows:*

$$w_{k+1}^p := \begin{cases} \tilde{w}_{k+1}^p & \text{if there is no event during } (t_k, t_{k+1}],\\ \lambda_{(Y_{t_k}^\xi, Y_{t_{k+1}}^\xi)}(\bar{X}_\Delta^p) \times \tilde{w}_{k+1}^p & \text{if an event occurs at } t_{k+1},\end{cases}$$

where ξ *indicates the identity of the firm corresponding the event.*
Now we compute $v_{k+1}^p = \frac{n_k w_{k+1}^p}{\sum_{q=1}^{n_k} w_{k+1}^q}$ *for every* $p = 1, \cdots, n_k$ *and obtain* $\{o_{k+1}^p\}_{p=1,\cdots,n_k}$ *as below:*

$$o_{k+1}^p = \begin{cases} [v_{k+1}^p] & \text{with probability } 1 + [v_{k+1}^p] - v_{k+1}^p\\ [v_{k+1}^p] + 1 & \text{with probability } v_{k+1}^p - [v_{k+1}^p]\end{cases}$$

where $[v]$ *stands for the integer part of* v*.*
Denote by $n_{k+1} = \sum_{p=1}^{n_k} o_{k+1}^p$ *the total number of particles at time* t_{k+1}*.*
Then each particle x_k^p *at time* t_k *independently generates* o_{k+1}^p *offspring of* $\big(\bar{X}_s^p\big)_{0\le s\le\Delta}$ *starting at* x_k^p *for every* $p = 1, \cdots, n_k$ *and denote by* $\boldsymbol{x}_{k+1} = (x_{k+1}^1, x_{k+1}^2, \cdots, x_{k+1}^{n_{k+1}})$ *all the realized random samples of* $\bar{\mathbf{X}}_\Delta$*.*
Thus one can achieve the approximated discrete distribution at t_{k+1} *as*

$$\tilde{\pi}_{t_{k+1}} = \frac{1}{n_{k+1}} \sum_{p=1}^{n_{k+1}} \delta_{x_{k+1}^p}.$$

Furthermore, in order to calculate the total intensity $\lambda^{\mathrm{all}}(X_{t_{k+1}})$ *at the next time, we obtain the cardinality at the next time as* $M_\alpha(t_{k+1}) := M_\alpha(t_k) - 1$ *and* $M_\beta(t_{k+1}) := M_\beta(t_k) + 1$ *if one transition from the rating* α *to* β *occurred during* $(t_k, t_{k+1}]$*.*

Step 4.

Proceed from k to $k+1$ and go to **Step 2** *until some time horizon.*

In Step 3, the updating stage, each particle is replaced by the particles of which the number is randomly given by o^p. This procedure is worked in a consistent manner; particles with small weights w^p have almost zero offspring while those with large weights are replaced by several offspring. We mention that most of the calculation time with our algorithm is caused by sampling of the random number o^p in the updating stage.

4 Preliminaries for empirical analysis with rating transition history of Japanese enterprises

In this section we prepare the model parameters necessary for filtering. First, based on the one year transition probability matrix, we compute the constant generator matrix which can be considered as the rating transition rate matrix for the case of standard macroeconomic condition. Next, in order to introduce the dynamics, specify the function (functional) form of the intensity matrix and then estimate the parameters to fit with the credit cycle. Finally stochastic dynamics of the macroeconomic factor X_t and market index S_t are specified. Here we refer the Tokyo Stock Price Index (TOPIX) for the estimation of the model parameters of S_t.

4.1 Intensity matrix of rating transitions

Most Japanese portfolio managers keep track of the rating transitions announced by the rating agency R&I. Therefore, in our empirical study, we focus on the history of credit rating transitions of Japanese companies given by R&I, which is available with Bloomberg. Table 1 displays one year rating transition probability matrix $P(1)$ released by R&I in 2012.

Table 1: One year rating transition probability matrix $P(1)$ released by R&I in 2012.

	AAA	AA	A	BBB	BB	B	CCC	Default
AAA	0.909686	0.090314	0	0	0	0	0	0
AA	0.008859	0.939353	0.051107	0.000681	0	0	0	0
A	0.000145	0.018317	0.940544	0.038959	0.001454	0	0	0.000581
BBB	0	0.000286	0.036582	0.934267	0.027151	0.000429	0	0.001286
BB	0	0	0.002558	0.079284	0.865729	0.025575	0.001279	0.025575
B	0	0	0	0.007634	0.099237	0.770992	0.007634	0.114504
CCC	0	0	0	0	0	0.047619	0.880952	0.071429
Default	0	0	0	0	0	0	0	1

We suppose that the transition rates are governed by the latent macro factor X_t. In order to peruse the filtering algorithm described in the Section 3.3 numerically, we need to determine the transition rate functions. For a given $P(t)$, we want to find transition rate matrix $Q = (q_{i,j})_{1\leq i,j\leq 8}$ which satisfies

$$P(t) = \exp(tQ).$$

As we can see in Section 8.3 in [2003], one of the most numerically robust estimation methods is the fitting of the transition rate matrix approximately, i.e., $\hat{Q}$ can be obtained by solving the following optimization problem.

$$\begin{aligned} \text{minimize} \quad & \| P(1) - e^{Q} \|_2 \\ \text{subject to} \quad & q_{i,j} \geq 0 \ \forall i,j \in \{1,2,\cdots,8\}, i \neq j, \\ & q_{i,i} = -\sum_{j\neq i} q_{i,j} \ \forall i \in \{1,2,\cdots,7\}, \\ & q_{8,j} = 0 \ \forall j \in \{1,2,\cdots,8\}. \end{aligned}$$

Here the $\| \cdot \|_2$ - norm is defined as the square root of the sum of the squared elements of matrix. An initial value Q_0 for optimization algorithm is set as $Q_0 = U\Gamma U^{-1}$ with $U = [u_1, u_2, \cdots, u_8]$ and $\Gamma = diag(\gamma_1, \cdots, \gamma_8)$, where γ_i is an eigenvalue of $P(1)$ and u_i is the corresponding eigenvector. This depends on the fact that if there exists a non-singular matrix U and a diagonal matrix Γ satisfying $P = U\Gamma U^{-1}$, then a matrix Q that satisfies $P = \exp Q$ can be also diagonalised with U. As a result, we achieve the approximated matrix $\hat{Q}$ as follows

$$\hat{Q} = \begin{pmatrix} -0.0958 & 0.096 & 6.92e{-9} & 7.55e{-5} & 2.14e{-6} & -6.77e{-8} & 1.60e{-9} & 1.69e{-6} \\ 0.009 & -0.067 & 0.054 & 0.0005 & 4.44e{-5} & 5.46e{-7} & 3.40e{-8} & 1.85e{-5} \\ 6.48e{-5} & 0.019 & -0.066 & 0.041 & 0.001 & 2.7e{-5} & 4.08e{-7} & 0.0008 \\ 2.53e{-6} & 9.35e{-5} & 0.039 & -0.073 & 0.030 & 2.11e{-5} & 4.36e{-6} & 0.001 \\ 6.04e{-8} & 2.04e{-5} & 0.001 & 0.088 & -0.078 & 0.031 & 0.002 & 0.025 \\ 4.68e{-9} & 2.47e{-6} & 0.0001 & 0.003 & 0.122 & -0.336 & 0.009 & 0.128 \\ 6.71e{-11} & 9.11e{-8} & 6.51e{-6} & 6.92e{-5} & 3.21e{-8} & 0.056 & -0.127 & 0.071 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix}.$$

Here the approximated transition probability $\exp \hat{Q}$ is given by

$$\exp \hat{Q} = \begin{pmatrix} 0.909075 & 0.088396 & 0.002395 & 0.000126 & 0.000006 & 0.000000 & 0.000000 & 0.000003 \\ 0.008509 & 0.939250 & 0.050584 & 0.001530 & 0.000085 & 0.000000 & 0.000000 & 0.000040 \\ 0.000143 & 0.018155 & 0.940466 & 0.038797 & 0.001569 & 0.000044 & 0.000002 & 0.000826 \\ 0.000000 & 0.000440 & 0.036379 & 0.934240 & 0.027090 & 0.000419 & 0.000028 & 0.001400 \\ 0.000000 & 0.000006 & 0.002825 & 0.079083 & 0.865642 & 0.025391 & 0.001594 & 0.025420 \\ 0.000000 & 0.000000 & 0.000325 & 0.007303 & 0.099262 & 0.770997 & 0.007646 & 0.114460 \\ 0.000000 & 0.000000 & 0.000014 & 0.000226 & 0.002850 & 0.046138 & 0.880572 & 0.070199 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \end{pmatrix}.$$

The approximation error is $\| P(1) - e^{\hat{Q}} \|_2 = 0.004866$, seems sufficiently accurate.

In view of the low sample of default events occurred in both AAA and AA, we merged them into one credit rating class. The new credit rating class is denoted by $AAA\&AA$. In addition, given the low sample of firms granted the rating B and CCC, we merged these rating classes into one rating class and write as $B\&CCC$. Therefore the new set of ratings is redefined as $\tilde{\mathcal{R}} = \{AAA\&AA, A, BBB, BB, B\&CCC, D\} = \{1,2,3,4,5,6\}$ and the new transition rate matrix $\tilde{Q} = (\tilde{q}_{i,j})_{1\le i,j\le 6}$ can be derived with the relations;

$\tilde{q}_{1,1} = \hat{q}_{1,1} + \hat{q}_{1,2} + \hat{q}_{2,1} + \hat{q}_{2,2}$, $\tilde{q}_{1,\beta} = \hat{q}_{1,\beta} + \hat{q}_{2,\beta}$ $(3 \le \beta \le 5)$, $\tilde{q}_{1,5} = \hat{q}_{1,6} + \hat{q}_{1,7} + \hat{q}_{2,6} + \hat{q}_{2,7}$,
$\tilde{q}_{\alpha,1} = \hat{q}_{\alpha,1} + \hat{q}_{\alpha,2}$ $(3 \le \alpha \le 5)$, $\tilde{q}_{5,1} = \hat{q}_{6,1} + \hat{q}_{6,2} + \hat{q}_{7,1} + \hat{q}_{7,2}$, $\tilde{q}_{5,\beta} = \hat{q}_{6,\beta} + \hat{q}_{7,\beta}$, $(3 \le \beta \le 5)$,
$\tilde{q}_{5,5} = \hat{q}_{6,6} + \hat{q}_{6,7} + \hat{q}_{7,6} + \hat{q}_{7,7}$.

We conclude that the resulting matrix is

$$\tilde{Q} = \begin{pmatrix} -0.0545 & 0.0538 & 6e-04 & 5e-05 & 6.5e-07 & 0.0000202 \\ 0.0194 & -0.0627 & 0.0414 & 1e-03 & 3e-05 & 0.00081 \\ 1e-04 & 0.0388 & -0.0701 & 0.0301 & 3e-05 & 0.00104 \\ 2e-05 & 0.0014 & 0.0880 & -0.1475 & 0.0328 & 0.02533 \\ 2.6e-06 & 0.0002 & 0.0033 & 0.1216 & -0.3246 & 0.19955 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix}.$$

4.2 Specification of event intensities and preliminary analysis

According to R&I, since $P(1)$ is estimated from the data recorded during 34 years between 1978 and 2011, we can consider this intensity matrix $\tilde{Q}$ as the transition rate matrix conditional to standard macroeconomic state $X = 0$. The next step is to construct $\tilde{Q}(x)$ as a continuous function of $x \in \mathcal{S}^X$, where $\mathcal{S}^X$ denotes the state space of X, preserving some conditions such as described in Section 2.2.

4.2.1 Dynamic event intensities

We specify the event intensities $\lambda_{\alpha,\beta}(X_t))_{1\le\alpha,\beta\le D}$, where the intensities are classified into three types of events; upgrade, downgrade and default. First, for an upgrade $(\alpha,\beta) \in (\mathcal{R}\backslash\{D\})^2$ with $\alpha > \beta$, we assume that the function $\lambda_{(\alpha,\beta)}(x)$ is decreasing in x and $\alpha - \beta$. Second, for a downgrade $(\alpha,\beta) \in (\mathcal{R}\backslash\{D\})^2$ with $\alpha < \beta$, we assume that the function $\lambda_{(\alpha,\beta)}(x)$ is increasing in x and is decreasing in $\beta - \alpha$. And last, for a default $(\alpha, D) \in (\mathcal{R}\backslash\{D\}) \times \{D\}$, we assume that the function $\lambda_{(\alpha,D)}(x)$ is increasing in x and is increasing in α.

We first formulate the default intensity processes $\lambda_{Y_t^i,D}(X_t)$ depending on the macroeconomic state X_t at time t. Suppose that the firm i, classified into the rating α at time t, has the default intensity $\lambda_{\alpha,D}(X_t)$ expressed as

$$\lambda_{\alpha,D}(x) = \lambda_{\alpha,D} \cdot \exp(C_{\alpha,D} \cdot x),$$

for $\lambda_{\alpha,D}, C_{\alpha,D} \in \mathbb{R}_+$ to be estimated. Similarly, suppose that α rated firm i at time t has the rating transition intensities $\lambda_{\alpha,\beta}(x)$ expressed as

$$\lambda_{\alpha,\beta}(x) = \begin{cases} \lambda_{\alpha,\beta} \cdot \exp(C_{\downarrow} \cdot x) & \text{if } \alpha < \beta \ , \\ \lambda_{\alpha,\beta} \cdot \exp(-C_{\uparrow} \cdot x) & \text{if } \alpha > \beta \ , \end{cases}$$

for $\lambda_{\alpha,\beta}, C_{\downharpoonleft}, C_{\upharpoonleft} \in \mathbb{R}_+$ to be estimated. Then the transition rate matrix can be expressed as

$$\begin{pmatrix} -\lambda_{1,1}(X_t) & \lambda_{1,2}(X_t) & \lambda_{1,3}(X_t) & \cdots & \lambda_{1,6}(X_t) \\ \lambda_{2,1}(X_t) & -\lambda_{2,2}(X_t) & \lambda_{2,3}(X_t) & \cdots & \lambda_{2,6}(X_t) \\ \vdots & \vdots & \vdots & & \vdots \\ \lambda_{5,1}(X_t) & \lambda_{5,2}(X_t) & \lambda_{5,3}(X_t) & \cdots & \lambda_{5,6}(X_t) \\ 0 & 0 & 0 & \cdots & 0 \end{pmatrix}$$

$$= \begin{pmatrix} \lambda_{1,1}(X_t) & \lambda_{1,2}e^{C_{\downharpoonleft}X_t} & \lambda_{1,3}e^{C_{\downharpoonleft}X_t} & \lambda_{1,4}e^{C_{\downharpoonleft}X_t} & \lambda_{1,5}e^{C_{\downharpoonleft}X_t} & \lambda_{1,6}e^{C_{1,6}X_t} \\ \lambda_{2,1}e^{-C_{\upharpoonleft}X_t} & \lambda_{2,2}(X_t) & \lambda_{2,3}e^{C_{\downharpoonleft}X_t} & \lambda_{2,4}e^{C_{\downharpoonleft}X_t} & \lambda_{2,5}e^{C_{\downharpoonleft}X_t} & \lambda_{2,6}e^{C_{2,6}X_t} \\ \lambda_{3,1}e^{-C_{\upharpoonleft}X_t} & \lambda_{3,2}e^{-C_{\upharpoonleft}X_t} & \lambda_{3,3}(X_t) & \lambda_{3,4}e^{C_{\downharpoonleft}X_t} & \lambda_{3,5}e^{C_{\downharpoonleft}X_t} & \lambda_{3,6}e^{C_{3,6}X_t} \\ \lambda_{4,1}e^{-C_{\upharpoonleft}X_t} & \lambda_{4,2}e^{-C_{\upharpoonleft}X_t} & \lambda_{4,3}e^{-C_{\upharpoonleft}X_t} & \lambda_{4,4}(X_t) & \lambda_{4,5}e^{C_{\downharpoonleft}X_t} & \lambda_{4,6}e^{C_{4,6}X_t} \\ \lambda_{5,1}e^{-C_{\upharpoonleft}X_t} & \lambda_{5,2}e^{-C_{\upharpoonleft}X_t} & \lambda_{5,3}e^{-C_{\upharpoonleft}X_t} & \lambda_{5,4}e^{-C_{\upharpoonleft}X_t} & \lambda_{5,5}(X_t) & \lambda_{5,6}e^{C_{5,6}X_t} \\ 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix}$$

where the diagonal elements are given by

$$\lambda_{\alpha,\alpha}(X_t) = -\left(\sum_{\alpha<\beta} \lambda_{\alpha,\beta}e^{C_{\downharpoonleft}X_t} + \sum_{\alpha>\beta} \lambda_{\alpha,\beta}e^{-C_{\upharpoonleft}X_t} + \lambda_{\alpha,D}e^{C_{\alpha,D}X_t}\right).$$

4.2.2 Joint calibration of $C_{\alpha D}, C_{\downharpoonleft}, C_{\upharpoonleft}$

We split the historical rating transitions data recorded by R&I from April 1998 to March 2012 into two portions of data sets. The first one, starting from April 1998 and ending in March 2008, is devoted to the estimations of the parameters employing transition intensities as described in previous subsection. The second one, starting from April 2008 and ending in March 2012, is used for filtering. Based on the rating history available with Bloomberg, the number of transition change events between the states in $\tilde{\mathcal{R}} = \{1, 2, 3, 4, 5, 6\}$ are summarized in Table 2.

Table 2: The number of transition change events recorded by R&I. (Source: Bloomberg)

Year	1998	1999	2000	2001	2002	2003	2004	2005	2006	2007	2008
Up grade	4	10	9	7	3	5	18	24	51	39	10
Down grade	170	76	17	55	52	20	15	13	8	6	40
population	812	756	737	687	674	637	608	611	623	650	673
rate(Up)	0.005	0.013	0.012	0.010	0.005	0.008	0.03	0.039	0.082	0.060	0.015
rate(Down)	0.21	0.101	0.023	0.080	0.078	0.031	0.025	0.021	0.013	0.009	0.059

Let $\mathcal{T} \in \{1998, 1999, \cdots, 2008\}$ denotes the one year period and $P_D(\alpha, \mathcal{T})$, $P_{DG}(\mathcal{T})$ and $P_{UG}(\mathcal{T})$ denote the historical default probability of rating α, historical downgrade event probability of all the rating and historical upgrade event probability of all the rating during the period $\mathcal{T}$ respectively. We want to find parameters set $\Theta = \{\{C_{\alpha,D}\}_{\alpha=1,\cdots,5}, C_{\downharpoonleft}, C_{\upharpoonleft}\}$ which control functions $\lambda_{\alpha,\beta}(x)$ so as to explain historical rating transition data suitably. For this purpose, we consider 18 parameters $X_{1998}, X_{1999}, \cdots, X_{2008}, C_{\alpha,D}(\alpha = 1, \cdots, 5), C_{\downharpoonleft}, C_{\upharpoonleft}$ jointly and minimize the value evaluated by the following function $A : \mathbb{R}^{18} \to \mathbb{R}$.

$$A(\{X_{\mathcal{T}}\}_{\mathcal{T}=1999,\cdots,2008}, \Theta)$$

$$
\begin{aligned}
&= \sum_{\mathcal{T}=1998}^{2008} \sum_{\alpha=1}^{5} \left(\left[1 - \exp(-\lambda_{\alpha,D} e^{C_{\alpha,D} X_t})\right] - P_D(\alpha, \mathcal{T}) \right)^2 \\
&\quad + \sum_{\mathcal{T}=1998}^{2008} \left(\sum_{\alpha<\beta} \left[1 - \exp(-\lambda_{\alpha,\beta} e^{C_{\downharpoonleft} X_t})\right] - P_{DG}(\mathcal{T}) \right)^2 \\
&\quad + \sum_{\mathcal{T}=1998}^{2008} \left(\sum_{\alpha>\beta} \left[1 - \exp(-\lambda_{\alpha,\beta} e^{-C_{\upharpoonleft} X_t})\right] - P_{UG}(\mathcal{T}) \right)^2 .
\end{aligned}
$$

Thus Θ can be estimated by solving the following optimization problem.

$$
\begin{aligned}
\min_{X_{\mathcal{T}}, \Theta} \quad & A(\{X_{\mathcal{T}}\}_{\mathcal{T}=1999,\cdots,2008}, \Theta) \\
\text{subject to} \quad & X_{\mathcal{T}} \in \mathcal{S}^X, \ \{C_{\alpha D}\}_{\alpha=1,\cdots,5}, C_{\downharpoonleft}, C_{\upharpoonleft} \geq 0
\end{aligned}
$$

The results are shown in Table 3 and we also mention that the above minimization is almost independent of the choice of the initial values.

Table 3: Estimated parameter set Θ

C_{1D}	C_{2D}	C_{3D}	C_{4D}	C_{5D}	$C_{\downharpoonleft}$	$C_{\upharpoonleft}$
1.7102	4.6344	5.4332	3.3367	2.9750	2.3673	4.0739

4.3 Specification of the dynamics of S_t and X_t

As mentioned above, we refer to TOPIX as the asset price process $\{S_t\}$ governed by 2. The volatility σ of TOPIX return can be determined by the sample standard deviation of historical daily return $\sqrt{\mathbf{Var}[\log(S_{t+1}/S_t)]}$. However, determination of κ, c and $\mu(x)$ is not straightforward. Assume

$$
\mu(x) = \tanh(-\mu x) \tag{11}
$$

to achieve $\mu(x) < 0$ when $x > 0$, and $\mu(x) > 0$ when $x < 0$. And then take six month moving average of the TOPIX denoted by $\mathbf{MA}(t)$ and their daily return $\{\hat{\mu}_t\}_{t=1,2,\cdots}$ as follows.

$$
\begin{aligned}
\mathbf{MA}(t) &:= \frac{1}{21 \times 6} \sum_{s=t-21\times 6}^{t} TOPIX(s), \\
\hat{\mu}_t &:= \log \frac{\mathbf{MA}(t+1)}{\mathbf{MA}(t)}.
\end{aligned}
$$

In order to estimate the parameter μ which controls the trend of the TOPIX return, we assume that the series $\{\hat{\mu}_t\}_{t=1,2,\cdots}$ are realizations of $\{\mu(X_t)\}_{t=1,2,\cdots}$, where the random variables $\{X_t\}_{t=1,2,\cdots}$ has the density

$$
f(x) = \sqrt{\frac{\kappa}{\pi c^2}} \exp\left(-\frac{\kappa}{c^2} x^2\right),
$$

which represents the steady state of the OU process X_t. Let D_i be the empirical percentile points defined by $D_i = \mathbb{P}(\hat{\mu} > x_i)$ for arbitrary selected real values $x_i \in (-\infty, \infty)$ and M_i be theoretical percentile points defined by $M_i = \mathbb{P}(\mu(X) > x_i),\ X \sim N\left(0, \frac{c^2}{2\kappa}\right)$, for $i = 1, 2, \cdots, N$. Then the parameters κ, c and μ can be estimated by solving the following minimization problem

$$\begin{aligned} \min_{\kappa,c,\mu} \quad & \sum_{i=1}^{N} (D_i - M_i)^2 \\ \text{subject to} \quad & \kappa > 0, c > 0, \mu > 0, \end{aligned}$$

and the estimated parameters are displayed in Table 4.

Table 4: Estimated parameters

κ	c	μ	σ
1.8548	0.1814	2.7142	0.2291

5 Results and considerations

5.1 Transitions of filtered X_t

In this section we illustrate the numerical results based on the credit rating history of Japanese companies and the historical TOPIX data. In order to pursue the particle filter algorithm, we take $n_0 = 10000$ and $\Delta = 1/365$. All the other parameters are set to the values specified in Section 4. Figure 1 shows the historical data of logarithm of daily TOPIX and the timing of events such as upgrade, downgrade and default captured by R&I.

In Figure 1, the dates when defaults, upgrades and downgrades occurred are indicated with the mark " $\circ$ ", " $*$ " and " $+$ " respectively. For example, the first upgrade event marked with " $*$ " at the far left in the Figure 1 shows that one upgrade event occurred on April 4, 2009.

For reader's convenience, based on not only R&I data but also information available such as in Japanese news website, we compiled past defaults occurred during the second term (from April 2008 to March 2012) in Table 5. Since the credit ratings granted just before default are not easy to fix, we duplicated them so as to be consistent with the cohort data of NEWS RELEASE published by R&I. For example, Aiful Corporation went into default on September 9, 2009 and its credit rating given by R&I just before the default was BB. This is also displayed with the mark " $\circ$ " in Figure 1. We mention that from April 2010 to March 2012, there was no default within the firms that R&I had granted a rating.

Table 5: The past defaults occurred from April 2008 to March 2012.

Date	May 29, 2009	Sep. 18, 2009	Sep. 19,2009	Dec. 25, 2009	Jan. 19, 2010
Defaulted firm	Joint Corp.	Aiful Corp.	Willcom	Takefuji	Japan Air Line
Credit rating	BBB	BB	BBB	BBB	BB

Figure 2 illustrates the transitions of the percentile points of the occupation measure $\tilde{\pi}_t$ calculated every day from April 1, 2009 to March 31, 2012. Quick overview of both figures demonstrate that the down grade events and default events tend macroeconomic factor process X_t under the

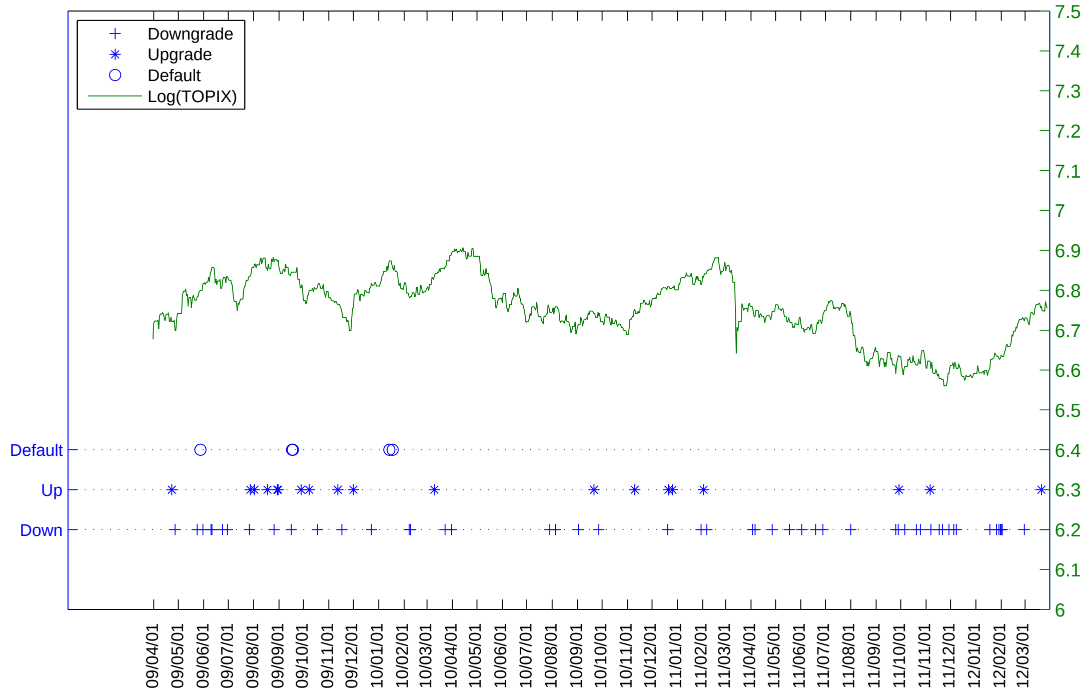


Figure 1: Observables

investor's filtration $\mathcal{G}_t$ to increase, while the upgrade events tend X_t to decrease. When no events are observed, we can see that X_t simply obey the SDE (1) and then X_t is forced to revert to the long-run average level 0.

The default contagion, directly induced by the jump up of X_t, can be captured by this model. How the default intensity of the surviving firms may change from before and after a default can be calculated at each time of the default as follows. Default intensity for rating class α, evaluated just before default τ is given by $\hat{\lambda}_{\alpha,D}(\tau-) = \mathbb{E}[\lambda_{\alpha,D}(X_{\tau-})|\mathcal{G}_{\tau-}]$ and evaluated just at τ is $\hat{\lambda}_{\alpha,D}(\tau) = \mathbb{E}[\lambda_{\alpha,D}(X_{\tau})|\mathcal{G}_{\tau}]$. Table 6 shows the default contagion effect, i.e., the changes of filtered default intensity for each rating class just before and after the default of the firm listed in the left side of the table. For instance, the default intensity of the rating class 5 (B and CCC) in our model, jumps upward from 9.859% to 11.063% due to the default of the Aiful Corporation.

5.2 Does TOPIX give additional information for the filtering?

We next examine if daily observation of TOPIX brings about additional information for the filtering. Figure 3 illustrates the transitions of some percentile points of the occupation measure $\tilde{\pi}_t$ without TOPIX for the same period as Figure 2 with TOPIX. Different from Figure 2, it is worth noting that the filtered X_t without TOPIX moves upwards or downwards monotonically during no credit events.

For a comparison between Figure 2 and 3, let us focus on the period from April 1, 2010 to July 30, 2010. During the three months sandwiched between a couple of adjacent downgrades, no credit event was observed. As is seen from Figure 1, TOPIX displays the downward trend during this

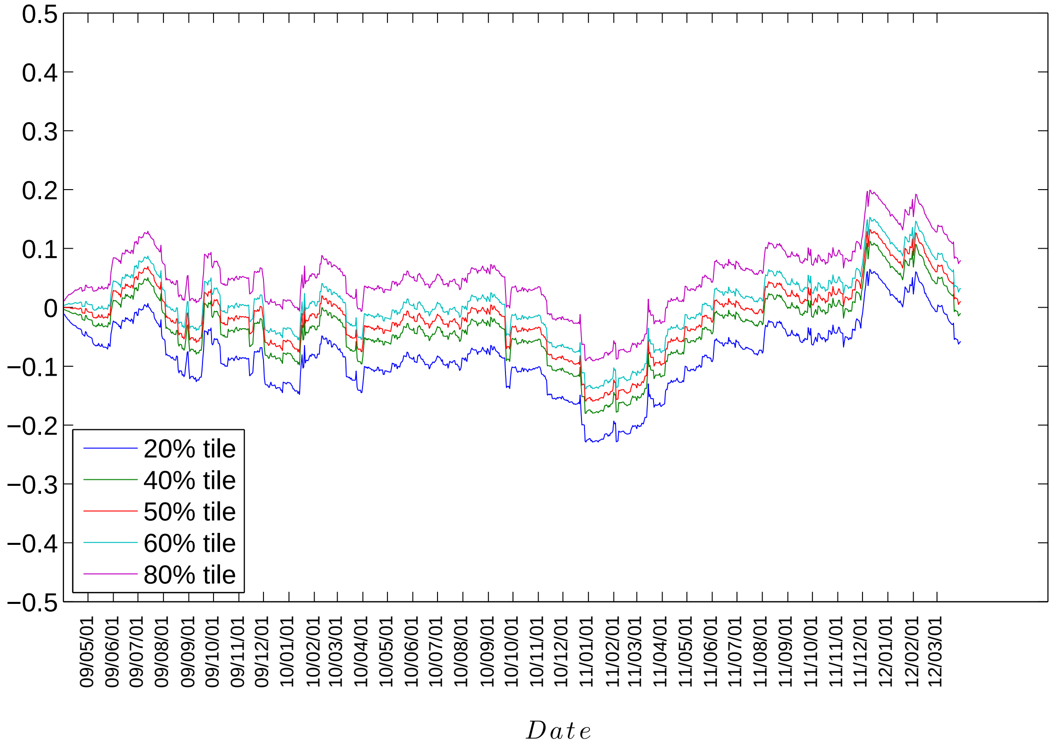


Figure 2: Filterd X_t

Table 6: Default contagion effects

Before/After the default of		Rating 1	Rating 2	Rating 3	Rating 4	Rating 5
Joint Corporation (BBB)	Before	0.0020%	0.061%	0.160%	4.576%	10.575%
	After	0.0021%	0.067%	0.177%	4.872%	11.183%
Aiful Corporation (BB)	Before	0.0019%	0.056%	0.143%	4.235%	9.859%
	After	0.0020%	0.066%	0.176%	4.819%	11.063%
Willcom (BBB)	Before	0.0020%	0.066%	0.176%	4.819%	11.063%
	After	0.0022%	0.079%	0.217%	5.479%	12.406%
Takefuji (BBB)	Before	0.0020%	0.066%	0.175%	4.786%	10.992%
	After	0.0021%	0.175%	0.205%	5.273%	11.983%
Japan Air Line (BB)	Befire	0.0019%	0.058%	0.151%	4.362%	10.118%
	After	0.0021%	0.069%	0.184%	4.944%	11.315%

period, implying that it would be appropriate to suppose that the macroeconomic state actually deteriorated in this period. These facts would be consistent with the equation (11), which assumed that TOPIX has downward trend when the macroeconomic factor stays in a bad state and vice versa.

Therefore we expect that the observation of some market index such as TOPIX could help to supplement some information about the latent macro factor X_t while no event information is updated. In Figure 4 and 5, we compare the estimated distribution functions of the $\tilde{\pi}$ for both with and without TOPIX observations, respectively at the beginning date and the ending one of the focused period. It appears from Figure 4 that on April 1, 2010, $\tilde{\pi}_t$ filtered with TOPIX observation is upper than that without TOPIX. To the contrary, $\tilde{\pi}_t$ with TOPIX observation is much lower than that without TOPIX on July 30, 2010. This result implies that the filtering with TOPIX may capture a deterioration trend of the latent macro factor during the period more precisely. Thus we can mention that TOPIX is to some extent worth observing for a better filtering of X_t. If we can find a more informative and observable index than TOPIX, it is likely that the accuracy of the filter will be more improved. [7]

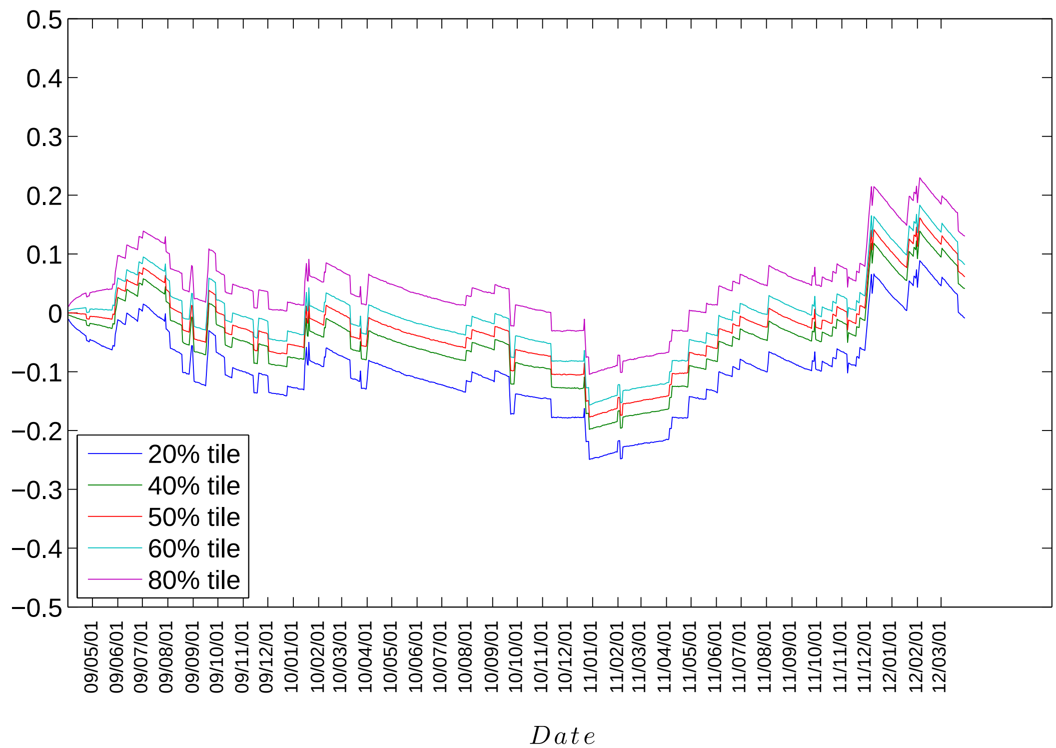
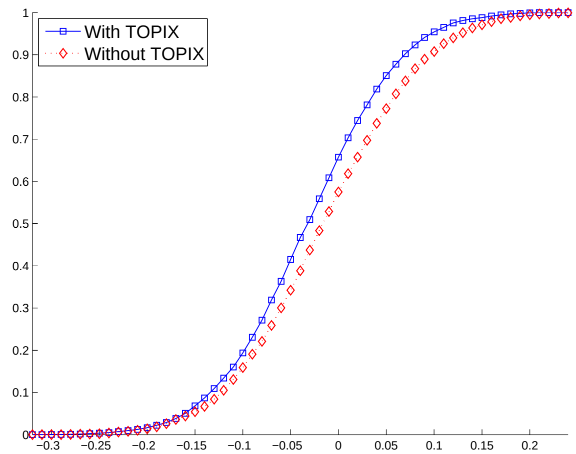


Figure 3: Filtered X_t: Without observations of TOPIX

5.3 Sensitivity analysis

In this subsection, we investigate a kind of robustness with respect to each of the model parameters via some numerical illustrations. More specifically, we see the impact on the filtering of overestimat-

[7] In order to quantitatively judge which model with or without TOPIX should be selected, we tentatively applied the Bayes factor criterion (refer to Jeffreys [1961] for example), that is recently applied to a nonlinear filtering with point process observation by Scott and Zeng [2011] due to its computational tractability. For two comparative models, the Bayes factor is defined by the ratio of integrated likelihood of one model (the model with TOPIX in our case) to that of the other (the model without TOPIX). Kass and Raftery [1995] suggested that if the Bayes factor takes a value more than three, it can be concluded that the former is better than the latter. In our case, we indeed found that the approximated Bayes factor almost always takes values around one. Therefore we cannot conclude from the Bayes factor that the model with TOPIX is more significant than that without TOPIX.



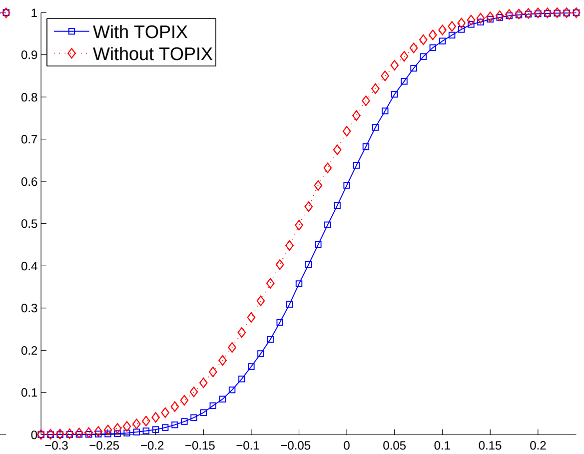


Figure 4: CDF of $\tilde{\pi}_t$ as of April 1, 2010

Figure 5: CDF of $\tilde{\pi}_t$ as of July 30, 2010

ing or underestimating the model parameters σ, μ, κ and c by changing the value of each parameter upwards or downwards from its base value summarized in Table 4. First, we examine the sensitivity with respect to the volatility σ and the drift μ of the TOPIX process $\{S_t\}$. It appears from Figure 6 and 7 that the approximated filter $\tilde{\pi}_t$ is more sensitive to σ than μ since the 40%-60% range of the approximated filters of the case with $\sigma = 0.1$ and $\sigma = 0.3$ are not overlapped in some periods. In addition, it seems that the larger σ, the more uncertain but the less sharply approximated filter $\tilde{\pi}_t$ fluctuates. Thus, the result indicates that the estimation of the asset price volatility is more important for the filtering.

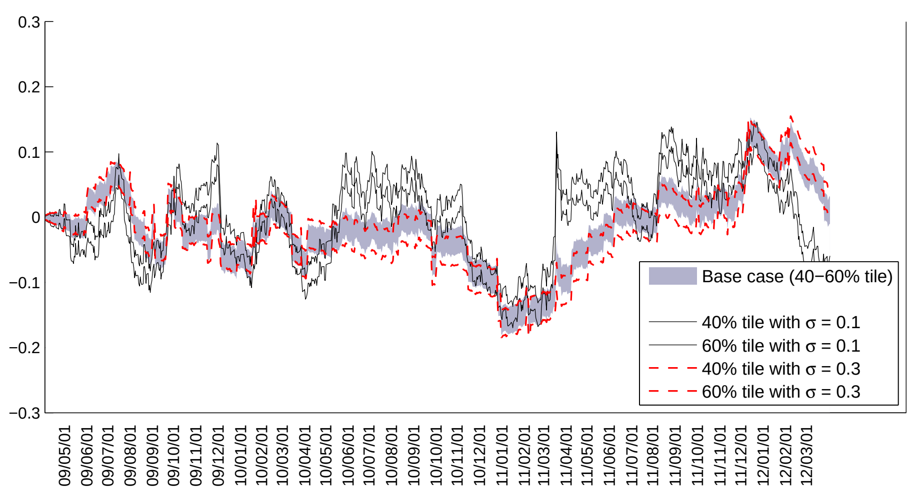


Figure 6: σ=0.1 or 0.3 with fixed μ, κ, c

Second, we see if the approximated filter $\tilde{\pi}_t$ is sensitive to the rate κ of mean reversion and the volatility c of the latent macro factor $\{X_t\}$. Figure 8 displays that $\tilde{\pi}_t$ fluctuates more sharply as κ gets smaller. Similarly, Figure 9 shows that $\tilde{\pi}_t$ fluctuates more sharply as c gets larger. It goes without saying that the results are consistent with the characteristics of the mean-reverting model

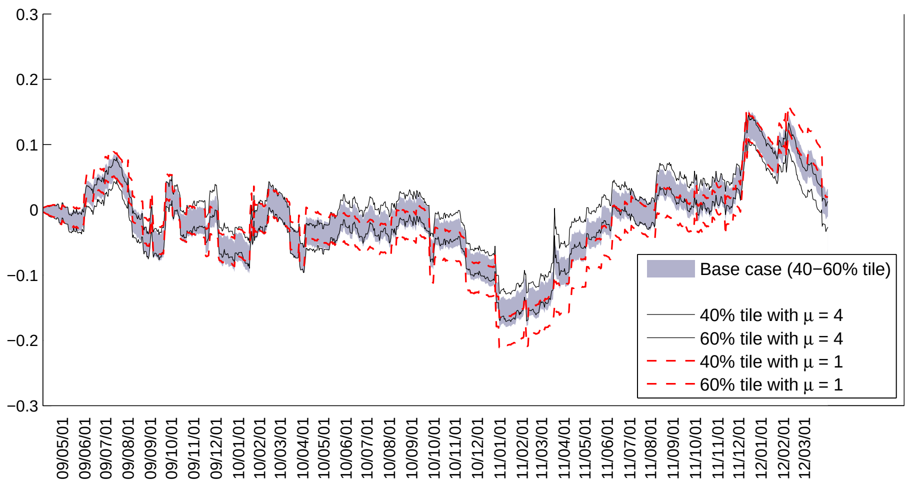


Figure 7: μ=4 or 1 with fixed σ, κ, c

because the larger κ and the smaller c imply that X_t is unlikely to be away from zero for a long time.

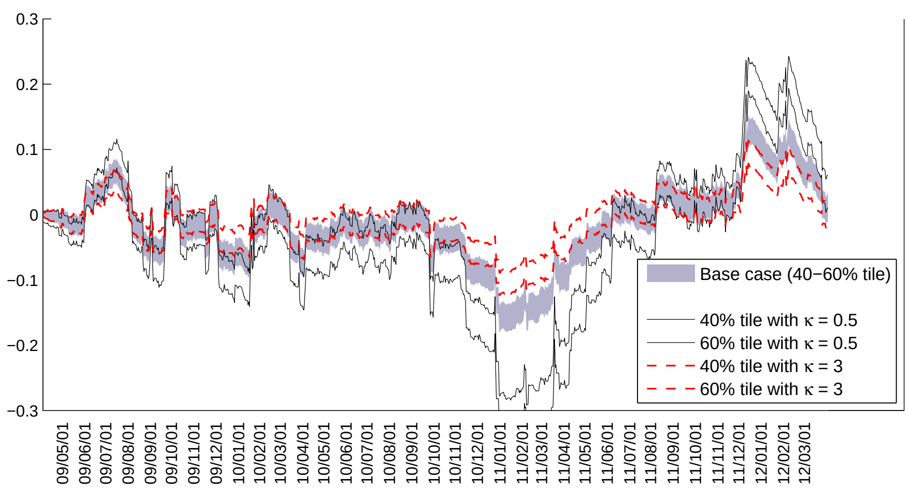


Figure 8: κ=0.5 or 3 with fixed σ, μ, c

6 Concluding remarks

Our contribution of this study to the field of credit risk research is summarized as follows. First, we introduce a new filtering model that assumes some latent macroeconomic factor has some influence on the frequency of credit rating transition events as well as default events. Since rating transition events are likely to be observed more often than default, the conditional distribution of the latent factor is expected to be estimated more accurately with observations on rating transition. Also, this model can contain some covariates like stock indices continuously observed in the market so as

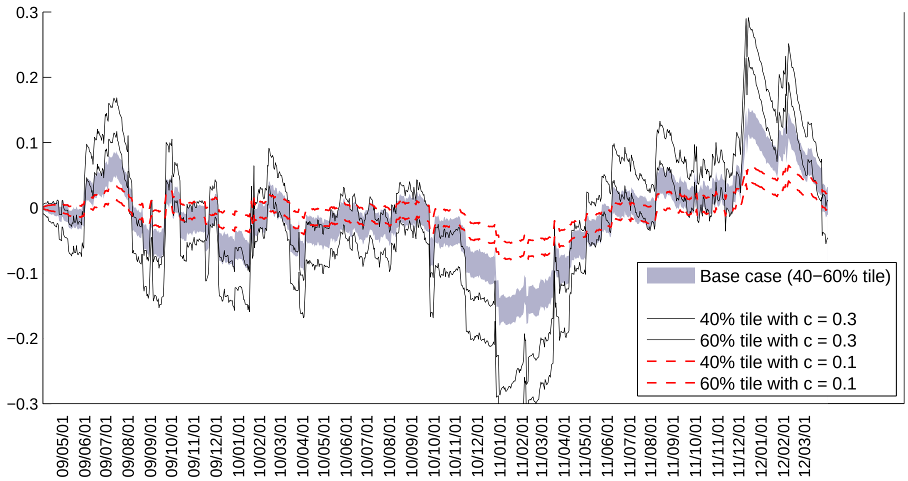


Figure 9: c=0.3 or 0.1 with fixed σ, μ, κ

to supplement information of the latent factor while no credit event happens. This implies that if we select a suitable continuous process related to the credit cycles, it is probable that such a more frequent observation from the market can improve the accuracy of filtering of the latent factor.

Second, we illustrate via some empirical analyses that our model is practically useful and tractable to examine what is related to occurrence of credit events for each rating category. Specifically, we naively apply a branching particle filter technique with historical data of credit rating transition and defaults in Japan to achieve the best estimate of the unobservable macroeconomic factor. Hence we can analyze how the credit events, the stock index, and the latent factor are related. As a consequence of the empirical analyses, we realize that a contagion effect among credit events as well as a relation between credit events and some financial market index are not negligible. At last, this enables us to prompt credit risk revaluation in terms of the estimated transition probability matrix under the real-world probability measure, without referring to any credit risky instruments traded in the market.

In addition, we have to mention a few of our future challenges. The convergence of the branching particle filter has not been argued in this paper, although it will be resolved by a discussion similar to the previous studies (for instance, Bain and Crisan [2008] and Xiong and Zeng [2011]). It is too technical to mention in this paper, so we will be ready for another paper for discussing such a theoretical issue.

From an empirical viewpoint, we try to look for a more suitable observable factor than TOPIX for improving the predictive accuracy of credit rating changes. This seems relevant to investigating what the latent factor really is.

Moreover, we may need to extend the model so that several observable and/or unobservable factors can be contained. In fact, the particle filter method has an advantage for high-dimensional problem, so an extension to multi-dimensional case cannot be an obstacle to numerical analysis.

Acknowledgements The authors are grateful to participants of the CSFI (Center for the Study of Finance and Insurance) Workshop on March 2014 at Osaka University and participants of the

Bachelier Finance Society 8th World Congress (BFS 2014) in Brussels for valuable comments which help to improve this paper.

A Proofs

Hereafter we will write $L_t(\omega_1, \omega_2)$ for $L_t(X(\omega_1), Z((\omega_2))$ in spite of abuse of notation.

A.1 Proof of Lemma 3.1

It follows from Kallianpur-Striebel formula that

$$\pi_t(f)(\omega) := \mathbb{E}_{\mathbb{P}}\left[f(X_t)\middle|\mathcal{G}_t\right](\omega) = \frac{\mathbb{E}_{\mathbb{Q}}[f(X_t)L_t \mid \mathcal{G}_t](\omega)}{\mathbb{E}_{\mathbb{Q}}[L_t \mid \mathcal{G}_t](\omega)}.$$

As we remark that $\mathbb{Q} = \mathbb{Q}^1 \times \mathbb{Q}^2$, all we have to show is

$$\mathbb{E}_{\mathbb{Q}^1\times\mathbb{Q}^2}\left[f(X_t)L_t(\cdot,\cdot)\middle|\mathcal{H}_t \vee \mathcal{F}_t^S\right](\omega_1,\omega_2) = \mathbb{E}_{\mathbb{Q}^1}\left[f(X_t)L_t(\cdot,\omega_2)\middle|\mathcal{H}_t\right](\omega_1). \tag{12}$$

Fix $A \in \mathcal{H}_t$ and $B \in \mathcal{F}_t^S$ arbitrarily. We have

$$\begin{aligned}
&\int_{A\times B}\mathbb{E}_{\mathbb{Q}^1\times\mathbb{Q}^2}\left[f(X_t)L_t(\cdot,\cdot)\middle|\mathcal{H}_t\vee\mathcal{G}_t\right](\omega_1,\omega_2)d(\mathbb{Q}^1\times\mathbb{Q}^2)(\omega_1,\omega_2)\\
&\quad=\int_{A\times B} f(X_t(\omega_1))L_t(\omega_1,\omega_2)d(\mathbb{Q}^1\times\mathbb{Q}^2)(\omega_1,\omega_2)\\
&\quad=\int_{\Omega^1\times B}\mathbf{1}_A(\omega_1)f(X_t(\omega_1))L_t(\omega_1,\omega_2)d(\mathbb{Q}^1\times\mathbb{Q}^2)(\omega_1,\omega_2)\\
&\quad=\int_{\Omega^1\times B}\mathbb{E}_{\mathbb{Q}^1\times\mathbb{Q}^2}\left[\mathbf{1}_A f(X_t)L_t(\cdot,\cdot)\middle|\mathcal{F}_t^S\right](\omega_1,\omega_2)d(\mathbb{Q}^1\times\mathbb{Q}^2)(\omega_1,\omega_2).
\end{aligned}$$

Here we define[8] for $w \in \Omega^2$

$$\varphi_A(w) := \mathbb{E}_{\mathbb{Q}^1}\left[\mathbf{1}_A(\omega_1)f(X_t(\omega_1))L_t(\omega_1,w)\right].$$

We should note that

$$\begin{aligned}
\varphi_A(w) &= \mathbb{E}_{\mathbb{Q}^1}\left[\mathbf{1}_A(\omega_1)f(X_t(\omega_1))L_t(\omega_1,w)\right]\\
&= \int_A f(X_t(\omega_1))L_t(\omega_1,w)d\mathbb{Q}^1(\omega_1)\\
&= \int_A \mathbb{E}_{\mathbb{Q}^1}\left[f(X_t)L_t(\cdot,w)\middle|\mathcal{H}_t\right](\omega_1)d\mathbb{Q}^1(\omega_1).
\end{aligned}$$

Then it follows from Fubini's theorem and the last note

$$\int_{\Omega^1\times B}\mathbb{E}_{\mathbb{Q}^1\times\mathbb{Q}^2}\left[\mathbf{1}_A f(X_t)L_t(\cdot,\cdot)\middle|\mathcal{F}_t^S\right](\omega_1,\omega_2)d(\mathbb{Q}^1\times\mathbb{Q}^2)(\omega_1,\omega_2)$$

[8] In general, for any $\mathcal{B}$-measurable random variable X and any random variable Y independent of $\mathcal{B}$ and a Borel function Φ with $\mathbb{E}[|\Phi(X,Y)|] < \infty$,

$$\mathbb{E}[\Phi(X,Y)|\mathcal{B}] = \varphi(X) \text{ a.s.} \quad \varphi(x) := \mathbb{E}[\Phi(x,Y)]$$

(See the appendix of Lamberton and Lapayre [1996].)

$$
\begin{aligned}
&= \int_{\Omega^1\times B} \varphi_A(\omega_2) d(\mathbb{Q}^1\times\mathbb{Q}^2)(\omega_1,\omega_2)\\
&\overset{\text{Fubini}}{=} \int_{\Omega^1} d\mathbb{Q}^1(\omega_1)\int_B \varphi_A(\omega_2) d\mathbb{Q}^2(\omega_2)\\
&= \int_B \varphi_A(\omega_2) d\mathbb{Q}^2(\omega_2)\\
&= \int_B \left\{\int_A \mathbb{E}_{\mathbb{Q}^1}\left[f(X_t)L_t(\cdot,w)\middle|\mathcal{H}_t\right](\omega_1)\right\}\Big|_{w=\omega_2} d\mathbb{Q}^2(\omega_2)\\
&\overset{\text{Fubini}}{=} \int_{A\times B} \mathbb{E}_{\mathbb{Q}^1}\left[f(X_t)L_t(\cdot,w)\middle|\mathcal{H}_t\right]\Big|_{w=\omega_2}(\omega_1) d(\mathbb{Q}^1\times\mathbb{Q}^2)(\omega_1,\omega_2).
\end{aligned}
$$

Due to the arbitrariness of A and B, equation (12) holds. □

A.2 Proof of Theorem 3.2

It follows from Lemma 3.1 that

$$\pi_t(f) = \frac{\mathbb{E}_{\mathbb{Q}^1}[f(X_t)L_t|\mathcal{H}_t]}{\mathbb{E}_{\mathbb{Q}^1}[L_t|\mathcal{H}_t]}.$$

Since we notice

$$\pi_t(f) = \mathbb{E}_{\mathbb{P}}[f(X_t)|\mathcal{G}_t] = \int_{\mathbb{R}} f(x)\pi_t(dx),$$

we can see that the following equality holds at time T_{n-1}:

$$\mathbb{E}_{\mathbb{Q}^1}[L_{T_{n-1}}f(X_{T_{n-1}})|\mathcal{H}_{T_{n-1}}] \;=\; \mathbb{E}_{\mathbb{Q}^1}[L_{T_{n-1}}|\mathcal{H}_{T_{n-1}}]\int_{\mathbb{R}} f(x)\pi_{T_{n-1}}(dx). \tag{13}$$

Moreover we can see $\mathcal{H}_t = \mathcal{H}_{T_{n-1}} \vee \sigma\{T_n\wedge s \mid s\le t\}$ for $t\in[T_{n-1},T_n)$. In general, it follows from Dellacherie formula that for any $\mathcal{F}_\infty$-integrable random variable U,

$$\mathbb{E}_{\mathbb{Q}^1}[U\mathbf{1}_{\{T_n>t\}}|\mathcal{H}_t] = \mathbf{1}_{\{T_n>t\}}\frac{\mathbb{E}_{\mathbb{Q}^1}[U\mathbf{1}_{\{T_n>t\}}|\mathcal{H}_{T_{n-1}}]}{\mathbb{E}_{\mathbb{Q}^1}[\mathbf{1}_{\{T_n>t\}}|\mathcal{H}_{T_{n-1}}]}.$$

Now we can apply the above result to the numerator $\mathbb{E}_{\mathbb{Q}^1}[f(X_t)L_t|\mathcal{H}_t]$ of (A.2) by setting $U = f(X_t)L_t$ for $t\in[T_{n-1},T_n)$. Since $\mathcal{H}_t\subset\mathcal{F}_t$, L_t is $\mathcal{F}_t$-measurable, we have for $t\in[T_{n-1},T_n)$,

$$
\begin{aligned}
&\mathbb{E}_{\mathbb{Q}^1}[f(X_t)L_t|\mathcal{H}_t]\\
&= \mathbb{E}_{\mathbb{Q}^1}[f(X_t)L_t\mathbf{1}_{\{T_n>t\}}|\mathcal{H}_t] = \frac{\mathbb{E}_{\mathbb{Q}^1}\left[f(X_t)L_t\mathbf{1}_{\{T_n>t\}}|\mathcal{H}_{T_{n-1}}\right]}{\mathbb{E}_{\mathbb{Q}^1}\left[\mathbf{1}_{\{T_n>t\}}|\mathcal{H}_{T_{n-1}}\right]}\\
&= \frac{\mathbb{E}_{\mathbb{Q}^1}\left[\ \mathbb{E}_{\mathbb{Q}^1}\left[f(X_t)L_t\mathbf{1}_{\{T_n>t\}}|\mathcal{F}_{T_{n-1}}\right]\ \middle|\ \mathcal{H}_{T_{n-1}}\right]}{\mathbb{E}_{\mathbb{Q}^1}\left[\mathbf{1}_{\{T_n>t\}}|\mathcal{H}_{T_{n-1}}\right]}\\
&= \frac{\mathbb{E}_{\mathbb{Q}^1}\left[\ L_{T_{n-1}}\mathbb{E}^{\mathbb{Q}^1}_{(X_{T_{n-1}},\mathbf{Y}_{T_{n-1}})}\left[f(\bar{X}_{t-T_{n-1}})\bar{L}_{t-T_{n-1}}\mathbf{1}_{\{\bar{T}_1>t-T_{n-1}\}}\right]\ \middle|\ \mathcal{H}_{T_{n-1}}\right]}{\mathbb{E}_{\mathbb{Q}^1}\left[\mathbf{1}_{\{T_n>t\}}|\mathcal{H}_{T_{n-1}}\right]}\\
&= \frac{\mathbb{E}_{\mathbb{Q}^1}\left[\ L_{T_{n-1}}\mathbb{E}^{\mathbb{Q}^1}_{(X_{T_{n-1}},\mathbf{Y}_{T_{n-1}})}\left[f(\bar{X}_{t-T_{n-1}})\bar{L}_{t-T_{n-1}}\mathbb{E}^{\mathbb{Q}^1}_{(X_{T_{n-1}},\mathbf{Y}_{T_{n-1}})}\left[\mathbf{1}_{\{\bar{T}_1>t-T_{n-1}\}}|\mathcal{F}^{\bar{X}}_\infty\right]\right]\ \middle|\ \mathcal{H}_{T_{n-1}}\right]}{\mathbb{E}_{\mathbb{Q}^1}\left[\mathbf{1}_{\{T_n>t\}}|\mathcal{H}_{T_{n-1}}\right]}
\end{aligned}
$$

$$= \frac{\mathbb{E}_{\mathbb{Q}^1}\left[L_{T_{n-1}} \mathbb{E}^{\mathbb{Q}^1}_{(X_{T_{n-1}}, \mathbf{Y}_{T_{n-1}})} \left[f(\bar{X}_{t-T_{n-1}}) \bar{L}_{t-T_{n-1}} \exp\left(- \int_0^{t-T_{n-1}} \lambda^{\mathrm{all}}(\bar{X}_s) ds \right) \right] \;\middle|\; \mathcal{H}_{T_{n-1}} \right]}{\mathbb{E}_{\mathbb{Q}^1}\left[\mathbf{1}_{\{T_n > t\}} | \mathcal{H}_{T_{n-1}} \right]}.$$

The equality (6) in the beginning of subsection 3.2 justifies that the conditional expectation under $\mathbb{Q}^1$ is changed to the expectation under $\mathbb{Q}^1_{(X_{T_{n-1}}, \mathbf{Y}_{T_{n-1}})}$ in the last third equality. Finally applying the relation (13) to the numerator of the last term implies

$$\begin{aligned}
&\mathbb{E}_{\mathbb{Q}^1}\left[L_{T_{n-1}} \mathbb{E}^{\mathbb{Q}^1}_{(\bar{X}_{T_{n-1}}, \mathbf{Y}_{T_{n-1}})} \left[f(\bar{X}_{t-T_{n-1}}) \bar{L}_{t-T_{n-1}} \exp\left(- \int_0^{t-T_{n-1}} \lambda^{\mathrm{all}}(\bar{X}_s) ds \right) \right] \;\middle|\; \mathcal{H}_{T_{n-1}} \right] \\
&\quad = \mathbb{E}_{\mathbb{Q}^1}[L_{T_{n-1}} | \mathcal{H}_{T_{n-1}}] \int_{\mathbb{R}} \mathbb{E}^{\mathbb{Q}^1}_{(x, \mathbf{Y}_{T_{n-1}})} \left[f(\bar{X}_{t-T_{n-1}}) \bar{L}_{t-T_{n-1}} \exp\left(- \int_0^{t-T_{n-1}} \lambda^{\mathrm{all}}(\bar{X}_s) ds \right) \right] \pi_{T_{n-1}}(dx).
\end{aligned}$$

As for the denominator of (A.2), due to the same argument, we have

$$\begin{aligned}
&\mathbb{E}_{\mathbb{Q}^1}[L_t | \mathcal{H}_t] \\
&\quad = \frac{1}{\mathbb{E}_{\mathbb{Q}^1}\left[\mathbf{1}_{\{T_n > t\}} | \mathcal{H}_{T_{n-1}} \right]} \mathbb{E}_{\mathbb{Q}^1}\left[L_{T_{n-1}} \mathbb{E}^{\mathbb{Q}^1}_{(\bar{X}_{T_{n-1}}, \mathbf{Y}_{T_{n-1}})} \left[\bar{L}_{t-T_{n-1}} \exp\left(- \int_0^{t-T_{n-1}} \lambda^{\mathrm{all}}(\bar{X}_s) ds \right) \right] \;\middle|\; \mathcal{H}_{T_{n-1}} \right] \\
&\quad = \frac{\mathbb{E}_{\mathbb{Q}^1}[L_{T_{n-1}} | \mathcal{H}_{T_{n-1}}]}{\mathbb{E}_{\mathbb{Q}^1}\left[\mathbf{1}_{\{T_n > t\}} | \mathcal{H}_{T_{n-1}} \right]} \int_{\mathbb{R}} \mathbb{E}^{\mathbb{Q}^1}_{(x, \mathbf{Y}_{T_{n-1}})} \left[\bar{L}_{t-T_{n-1}} \exp\left(- \int_0^{t-T_{n-1}} \lambda^{\mathrm{all}}(\bar{X}_s) ds \right) \right] \pi_{T_{n-1}}(dx).
\end{aligned}$$

And then $\frac{\mathbb{E}_{\mathbb{Q}^1}[L_{T_{n-1}} | \mathcal{H}_{T_{n-1}}]}{\mathbb{E}_{\mathbb{Q}^1}\left[\mathbf{1}_{\{T_n > t\}} | \mathcal{H}_{T_{n-1}} \right]}$ is cancelled between the numerator and the denominator of (A.2), so we can obtain Equation (8). □

A.3 Proof of Theorem 3.3

The argument is almost similar to the proof of Theorem 3.2, so we just mention the crucial part below. We should remark that $\mathcal{H}_{T_n} = \mathcal{H}_{T_{n-1}} \vee \sigma(T_n, \xi_n, Y^{\xi_n}_{T_n})$. Because of the Equality (7), we can see

$$\begin{aligned}
&\mathbb{E}_{\mathbb{Q}^1}\left[f(X_{T_n}) L_{T_n} | \mathcal{H}_{T_n} \right] \\
&\quad = \mathbb{E}_{\mathbb{Q}^1}\left[f(X_{T_n}) L_{T_n} | \mathcal{H}_{T_{n-1}} \vee \sigma(T_n, \xi_n, Y^{\xi_n}_{T_n}) \right] \\
&\quad = \mathbb{E}_{\mathbb{Q}^1}\left[L_{T_{n-1}} \mathbb{E}_{\mathbb{Q}^1}\left[f(X_{T_n}) \frac{L_{T_n}}{L_{T_{n-1}}} \;\middle|\; \mathcal{F}_{T_{n-1}}, T_n, \xi_n, Y^{\xi_n}_{T_n} \right] \;\middle|\; \mathcal{H}_{T_{n-1}}, T_n, \xi_n, Y^{\xi_n}_{T_n} \right] \\
&\quad = \mathbb{E}_{\mathbb{Q}^1}\left[L_{T_{n-1}} \mathbb{E}^{\mathbb{Q}^1}_{(X_{T_{n-1}}, \mathbf{Y}_{T_{n-1}})} \left[f(\bar{X}_{T_1}) \bar{L}_{T_n - T_{n-1}} \;\middle|\; \bar{T}_1 = T_n - T_{n-1}, \bar{\xi}_1, \bar{Y}^{\bar{\xi}_1}_{\bar{T}_1} \right] \;\middle|\; \mathcal{H}_{T_{n-1}}, T_n, \xi_n, Y^{\xi_n}_{T_n} \right] \\
&\quad \propto \mathbb{E}_{\mathbb{Q}^1}\left[L_{T_{n-1}} \mathbb{E}^{\mathbb{Q}^1}_{(X_{T_{n-1}}, \mathbf{Y}_{T_{n-1}})} \left[f(\bar{X}_{T_1}) \bar{L}_{T_n - T_{n-1}} \lambda_{(\bar{Y}^{\bar{\xi}_1}_0, \bar{Y}^{\bar{\xi}_1}_{\bar{T}_1})} (\bar{X}_{T_n - T_{n-1}}) \exp\left(- \int_0^{T_n - T_{n-1}} \lambda^{\mathrm{all}}(\bar{X}_s) ds \right) \right] \right. \\
&\qquad \left. \;\middle|\; \mathcal{H}_{T_{n-1}}, T_n, \xi_n, Y^{\xi_n}_{T_n} \right] \\
&\quad \propto \int_{\mathbb{R}} \mathbb{E}^{\mathbb{Q}^1}_{(x, \mathbf{Y}_{T_{n-1}})} \left[f\left(\bar{X}_{T_n - T_{n-1}} \right) \bar{L}_{T_n - T_{n-1}} \lambda_{(\bar{Y}^{\bar{\xi}_1}_0, \bar{Y}^{\bar{\xi}_1}_{\bar{T}_1})} (\bar{X}_{T_n - T_{n-1}}) \exp\left(- \int_0^{T_n - T_{n-1}} \lambda^{\mathrm{all}}(\bar{X}_s) ds \right) \right] \pi_{T_{n-1}}(dx).
\end{aligned}$$

This immediately concludes the proof.

□

References

[2008] Bain, A. and D. Crisan, 2008, *Fundamentals of Stochastic Filtering*, Stochastic Modelling & Applied Probability **60**, Springer.

[2007] Budhiraja, A., L. Chen and C. Lee, 2007, "A survey of numerical methods for nonlinear filtering problems", *Physica* D **230**, 27-36.

[2006] Ceci, C. and A. Gerardi, 2006, "A Model for High Frequency Data under Partial Information: A Filtering Approach", *International Journal of Theoretical and Applied Finance*, **9**(4), 555-576.

[1999] Crisan, D. and T. Lyons, 1999, "A Particle Approximation of the Solution of the Kushner-Stratonovitch Equation", *Probability Theory and Related Fields*, **115**, 549-578.

[2000] Del Moral, P. and L. Miclo, 2000, "Branching and Interacting Particle Systems. Approximations of Feynman-Kac Formulae with Applications to Non-linear Filtering", *Seminaire de probabilities (Strasbourg)*, tome 34, 1-145.

[2009] Duffie, D., A. Eckner, G. Horel, and L. Saita, 2009, "Frailty correlated default", *Journal of Finance*, **64**, 2089-2123.

[2010] Frey, R. and W. Runggaldier, 2010, "Pricing Credit Derivatives under Incomplete Information: a Nonlinear-Filtering Approach", *Finance and Stochastics*, **14**(4), 495-526.

[2012] Frey, R. and T. Schmit, 2012, "Pricing and Hedging of Credit Derivatives via the Innovations Approach to Nonlinear Filtering", *Finance and Stochastics*, **16**(1), 105-133.

[1961] Jeffreys, H., 1961, Theory of Probability. Oxford University Press, London.

[1995] Kass, R. E. and Raftery, A. E., 1995,Bayes factors and model uncertainty. *Journal of the American Statistical Association*, **90**, 773-795.

[1990] Kliemann, W. H., G. Kock, and F. Marchetti, 1990, "On the Unnormalized Solution of the Filtering Problem with Counting Process Observations", *IEEE Transactions and Information Theory*, **36**(6), 1415 - 1425.

[2006] Koopman, S. J., A. Lucas, and A. Monteiro, 2006, "The Multi-State Latent Factor Intensity Model for Credit Rating Transitions", *Tinbergen Institute Discussion Paper No. TI 05-071/4*

[2005] McNeil, A., R. Frey, and P. Embrechts, 2005, *Quantitative Risk Management: Concepts, Techniques and Tools*. Princeton University Press.

[2010] Nakagawa, H., 2010, "Modeling of Contagious Downgrades and Its Application to Multi-Downgrade Protection", *JSIAM Letters*, **2**, 65-68.

[1996] Lamberton, D. and Lapayre, B., 1996, *Introduction to Stochastic Calculus Applied to Finance*, Chapman and Hall.

[2004] Schönbucher, P., "Information Driven Default Contagion", 2004, *Preprint, Department of Mathematics, ETH Zürich.*.

[2003] Schönbucher, P., 2003, *Credit Derivatives Pricing Models: Models, Pricing and Implementation*, Wiley Finance.

[2011] Xiong, J. and Y. Zeng, 2011, "A Branching Particle Approximation to a Filtering Micro-movement Model of Asset Price", *Statistical Inference for Stochastic Processes*, **14**, 111-140.

[2011a] Yamanaka, S., M. Sugihara, and H. Nakagawa, 2011, "Analysis of Credit Event Impact with Self-Exciting Intensity Model", *JSIAM Letters*, **3**, 49-52.

[2011b] Yamanaka, S., M. Sugihara, and H. Nakagawa, 2011, "Analysis of Downgrade Risk in Credit Portfolios with Self-Exciting Intensity Model", *JSIAM Letters*, **3**, 93-96.

[2012] Yamanaka, S., M. Sugihara, and H. Nakagawa, 2012, "Modeling of Contagious Credit Events and Risk Analysis of Credit Portfolios", *Asia-Pacific Financial Markets*, **19**, 43-62.

[2003] Zeng, Y., 2003, "A Partially-Observed Model for Micro-movement of Asset Prices with Bayes Estimation via Filtering", *Mathematical Finance*, **13**, 411-444.

[2011] Scott, L. C. and Zeng, Y., 2011, "Filtering with Counting Process Observations: Application to the Statistical Analysis of the Micromovement of Asset Price", *The Oxford Handbook of Nonlinear Filtering, 1019-1046* .