HITOTSUBASHI UNIVERSITY

Center for Financial Engineering Education (CFEE)

CFEE Discussion Paper Series No. 2011-1

Estimating Default Probabilities Using Corporate Bond Price Data

Reiko Tobe

Center for Financial Engineering Education (CFEE)
Department of Economics, Hitotsubashi University
2-1 Naka, Kunitachi Tokyo, 186-8601 Japan
http://www.econ.hit-u.ac.jp/~finmodel

Estimating Default Probabilities Using Corporate Bond Price Data

Reiko Tobe [*]

March 13, 2012

Abstract

This paper discusses estimating method of implied default probability from the interest spread. Instead of using Duffie & Singleton(1999) model directly, we use statistical estimating procedure of Takahashi(2011). We shall provide a complete proof of the consistency of the estimator in the Takahashi(2011) model employing the method of Wald(1949) which gives the proof of consistency of maximum likelihood estimate. We also verify regularity conditions for asymptotic properties of the estimator in the Weibull survival function case. Furthermore, we provide an empirical analysis of implied survival probability estimating. In the empirical analysis, we discuss asymptotic properties using bootstrap method. An interesting finding is that the number of the data and the existence of the outlier data have a major effect on the bootstrap estimating results. When we analyze the impact of the bankruptcy of Lehman Brothers, we observe that the influences differ with the 5 individual companies.

Keywords: Implied default probability; Statistical model; Parametric model; Pseudo maximum likelihood estimator; Consistency; Asymptotic normality; Bootstrap;

1 Introduction

In this paper we study estimating method of default probabilities of individual companies and the representative default probabilities of rating classes. Many rating agencies publish bond issuer's rating using their own methodology in measuring creditworthiness and a specific rating scale. Typically, ratings are expressed as letter grades that range, for example, from 'AAA' to 'D' to communicate the agency's opinion of relative level of credit risk. For example, a corporate bond that is rated 'AA' is viewed by some rating agency's as having a higher credit quality than a corporate bond with a 'BBB' rating. But the 'AA' rating isn't a guarantee that it will not default, only that, in their opinion, it is less likely to default than the 'BBB' bond. Some rating agencies also announce annual default probabilities of each rating class.

For modeling credit risk, two classes of models exists; structural model and reduced-form model. Reduced-form model was originally introduced by Jarrow & Turnbull (1992), and subsequently studied by Jarrow and Turnbull (1995), Duffie and Singleton (1999) among others. By using Duffie and Singleton(1999) model, we may calculate "implied default probability" by calibration. "Implied default probability" means default

[*]Graduate School of Economics, Hitotsubashi University

probability which is incorporated by the bond prices and it gives us view and expectations in the market. Takahashi(2011) introduced statistical model for estimating implied default probability. They assumed observed corporate bond price contains some error. As a result of containing error term, we can discuss asymptotic properties of the estimator. In Takahashi(2011), they gave a sketchy proof for consistency and a complete proof for asymptotic normality. They also recommended using bootstrap method for examining asymptotic properties when we apply their method to real data.

The first objective of this paper is to give the complete proof of the consistency of the pseudo maximum likelihood estimator which is proposed by Takahashi(2011). We employ the method of Wald(1949) which gives the proof of consistency of maximum likelihood estimate. His proof assumes the compact property and uses the strong law of large numbers. Similarly, we prove consistency of our estimator under some regularly conditions using the strong law of large numbers.

Furthermore we shall verify regularity conditions for consistency and asymptotic normality of our estimator. When we estimate the statistical model of Takahashi(2011), we may use parametric model for estimating survival function. The parametric models we consider are exponential model and Weibull model. Needless to say, exponential distribution is a special case of Weibull distribution, we need only to verify the conditions in the Weibull model. That is the second objective of this paper.

The last objective is to apply our model to real data to estimate default probabilities(survival probabilities). We present several numerical results in this paper. First, in order to set known constants of our model, we examine comparative study varying the recovery rate and the coefficient of error term. Second, we compare the results on two parametric survival functions; exponential survival function and Weibull survival function. The sum of squared residuals from the exponential survival function turned out much larger than that of the Weibull survival function. Third, we compare different recovery models; RM recovery model proposed by Duffie & Singleton(1999) and RT recovery model proposed by Jarrow & Turnbull(1995). The examination showed that there is no great difference between the two. Fourth, we show estimating results and bootstrap results using 8 individual companies' data and 5 rating classes' data on April 16, 2004. There is a close relation between the distribution of bootstrap estimates and the number of data. For example, the number of Toyota's data was only 3 and the estimated bootstrap distribution of Toyota was quite different from normal distribution. It seems the larger amount of data(bonds) being used, the more like the bootstrap distribution is normal distribution. The existence of outlier data also influence upon the distribution of bootstrap estimates. For example, when we excluded 3 outlier data from the original "AA" class data, the warps are reduced and the distributions of bootstrap parameter estimates look more like normal distribution. Kolmogorov-Smirnov test also supported the influence of outlier data on the estimating results. Finally we apply our statistical estimating method to time series data. Since we are interested in the impact of the bankruptcy of Lehman Brothers(September 15, 2008), we use the data from January 4, 2008 to March 31, 2009 and select 5 companies, which might be affected by the bankruptcy. In fact, all 5 companies' survival probabilities were declining after the bankruptcy, but the influences differ with the individual company in some points.

This paper is organized as follows. Section 2 provides a review of statistical estimating method of implied survival probabilities proposed by Takahashi(2011). The proof of asymptotic property of our estimator is discussed in Section 3. Some empirical results are presented in Section 4, and the last Section is devoted to the summary and concluding remarks.

2 Review of the model and some development

In this section, we provide a brief explanation of Duffie and Singleton(1999), which derive the relationship between the default probabilities and the interest yield. We review Takahashi(2011) model which is the main model of this paper. We also show the application of the model to the different recovery type case.

2.1 Setup and results on implied default probabilities

Throughout this chapter, $\mathscr{W} = \{W(t), \ \ t \leq 0\}$ denotes standard Brownian motion, on a probability space $(\Pi, \mathscr{F}, P_r)$. The filtration $\{\mathscr{F}_t\}_{t\geq 0}$ is generated by $\mathscr{W}$. We suppose there are no arbitrage opportunities in the market and the market is assumed to be complete.

Let $P(0,t)$ be time 0 price of the defaultable zero coupon bond (corporate bond) maturing at t and $P^*(0,t)$ be time 0 price of the corresponding default free bond (government bond). The spread (difference of the yield to maturity between corporate bond and government bond) Υ is thus defined by

$$\Upsilon(0) = -\frac{1}{t}\log\left[\frac{P(0,t)}{P^*(0,t)}\right]$$

Let Q be the unique equivalent Martingale measure, then,

$$P^*(0,t) = E^Q\left\{\exp\left[-\int_0^t r(v)dv\right]\right\}$$

where $r(v)$ is an instantaneous risk free spot rate and E^Q denotes the expectation under the measure Q (Harrison and Pliska[1981]). Let τ be the time of default, then the survival function $G(t)$ is given by

$$\begin{aligned} G(t) &= Q\{\tau > t\} \\ &= \exp\left\{-\int_0^t \lambda(u)du\right\} \end{aligned}$$

where $\lambda(t)$ is the hazard function. We also let δ_t be the recovery rate in the event of default at t.

The formula between the survival function and the interest spread varies with the definition of recovery model. We consider the following two recovery models which are in common use.

RM Type Recovery (cf. Duffie & Singleton(1999)) By assuming that $\delta_t \times P(\tau-,t)$ is recovered at the time of default, we have

$$P(t,T) = E^Q\left\{\exp\left[-\int_t^T (r(v) + (1-\delta_v)\lambda(v))\,dv\right] |\mathscr{F}_t\right\}$$

where $P(\tau-,t)$ is the price of corporate bond just before default. If λ is assumed to be non random function and δ to be constant, we have

$$P(0,t) = E^Q\left\{\exp\left[-\int_0^t (r(v) + (1-\delta)\lambda(v))\,dv\right]\right\}$$

$$
\begin{aligned}
&= E^Q\left\{\exp\left[-\int_0^t r(v)dv\right]\exp\left[-(1-\delta)\int_0^t \lambda(v)dv\right]\right\}\\
&= P^*(0,t)Q\{\tau>t\}^{1-\delta}\\
&= P^*(0,t)G(t)^{1-\delta}
\end{aligned}
$$

Under RM type recovery, spread becomes

$$
\begin{aligned}
\Upsilon(t) &= -\frac{1}{t}\log\left\{\frac{P(0,t)}{P^*(0,t)}\right\}\\
&= -\frac{1}{t}\log\left\{G(t)^{1-\delta}\right\},
\end{aligned}
$$

and implied survival function[1] is

$$
\begin{aligned}
G(t) &= \left(e^{-t\Upsilon}\right)^{\frac{1}{1-\delta}}\\
&= \left(\frac{P(0,t)}{P^*(0,t)}\right)^{\frac{1}{1-\delta}}
\end{aligned}
$$

RT Type Recovery (cf. Jarrow & Turnbull(1995)) Under RT type recovery, recovery rate δ_t is formulated as

$$
\delta_t = (1-L_t)P^*(t,T)
$$

for a given fractional recovery process $(1-Lt)$. We assume default hazard rate process λ be independent (under Q) of short rate r and recovery process be constant. With constant δ and non-random λ, the payoff at maturity in the event of default is $\delta\times 1$. Then we have

$$
\begin{aligned}
P(u,s) &= E_u^Q\left\{\exp\left[-\int_u^s r(v)dv\right]\delta\cdot I_{\{\tau\le s\}}+\exp\left[-\int_u^s r(v)dv\right]\cdot 1\cdot I_{\{\tau>s\}}|\tau>u\right\}\\
&= E_u^Q\left\{\exp\left[-\int_u^s r(v)dv\right]\right\}E_u^Q\left\{\delta\cdot I_{\{\tau\le s\}}+1\cdot I_{\{\tau>s\}}|\tau>u\right\}\\
&= P^*(u,s)\left(\delta+(1-\delta)E_u^Q\left\{I_{\{\tau>s\}}|\tau>u\right\}\right)\\
&= \delta P^*(u,s)+(1-\delta)P^*(u,s)Q_u\{\tau>s\}
\end{aligned}
$$

By setting $u=0,\ s=t$, it follows that

$$
\begin{aligned}
P(0,t) &= \delta P^*(0,t)+(1-\delta)P^*(0,t)Q_u\{\tau>s\}\\
&= \delta P^*(0,t)+(1-\delta)P^*(0,t)G(t)
\end{aligned}
$$

Under RT type recovery, spread becomes

$$
\begin{aligned}
\Upsilon^\sharp(t) &= -\frac{1}{t}\log\left\{\frac{P(0,t)}{P^*(0,t)}\right\}\\
&= -\frac{1}{t}\log\{\delta+(1-\delta)G(t)\},
\end{aligned}
$$

[1] In this paper, implied survival function means survival function which is incorporated by the bond price.

and implied survival function is

$$
\begin{aligned}
G^{\sharp}(t) &= \frac{1}{1-\delta}\left(e^{-t\Upsilon}-\delta\right) \\
&= \frac{1}{1-\delta}\left(\frac{P(0,t)}{P^{*}(0,t)}\right)
\end{aligned}
$$

where we define $\Upsilon^{\sharp}(t)$ is the spread under RT type recovery and $G^{\sharp}$ is survival function under RT type recovery.

In both type of recovery model, the formula is derived which presents the relationship of the default probabilities and the interest spread. Takahashi(2011) proposed the statistical model of estimating procedure of implied default probabilities under this relationship. We review the results in the next subsection.

2.2 Review of the Statistical Model by Takahashi(2011)

We have an exact estimate of the ”true” survival function or each t and the distribution function of default time τ under the equivalent martingale measure Q.

$$
\tau \sim 1-G(t:\theta) \qquad \theta \in \Theta \subset R^{d} \qquad d \geq 1
$$

Since there may be many sources of errors such as a lack of our information and a noise in the market, we suppose observed $P(0,t_i)$ contains some error.

$$
\hat{P}(0,t_i) = P(0,t_i) + \sigma h_{t_i}\varepsilon_{t_i} \qquad i = 1,2,\ldots,n
$$

where σ is unknown constant and h_t is a known function with $h_0 = 0$ and h_t is an increasing in t. We consider

$$
h_t = [1-P^{*}(0,t)]^{\alpha}
$$

for some α. On the other hand, $P^{*}(0,t_i)$ is supposed to be observed without any error. After having observed $\hat{P}(0,t_i),\ i = 1,\ldots,n$, we calculate ”observed spread”

$$
\hat{\Upsilon}(t_i) = -\frac{1}{t}\log\left\{\frac{p(0,t_i)}{p^{*}(0,t_i)} + \sigma\rho_{t_i}\varepsilon_{t_i}\right\}
$$

where ρ_{t_i} is $\frac{h_t}{p^{*}(0,t_i)}$. Using the observed spread $\hat{\Upsilon}(t_i)$, implied survial function $S(t)$ in stead of $G(t)$ becomes

$$
\begin{aligned}
S(t) &= \left(e^{-t\hat{\Upsilon}}\right)^{\frac{1}{1-\delta}} \\
&= \left(e^{-t\cdot-\frac{1}{t}\log\left\{\frac{p(0,t_i)}{p^{*}(0,t_i)}+\sigma\rho_{t_i}\varepsilon_{t_i}\right\}}\right)^{\frac{1}{1-\delta}} \\
&= \left(e^{\log\left\{\frac{p(0,t_i)}{p^{*}(0,t_i)}+\sigma\rho_{t_i}\varepsilon_{t_i}\right\}}\right)^{\frac{1}{1-\delta}}
\end{aligned}
$$

It is convenient to rewrite $S(t)$ as

$$
S^{1-\delta}(t) = G_t^{1-\delta} + \sigma\rho_t\varepsilon_t
$$

Then we assume the following parametric model. Let Θ be a subset of R^d, where d denote the dimension of the parameter space. We also let

$$G(t:\theta) = Q\{\tau > t:\theta\}$$

be a survival function parameterized by d-dimensional vector $\theta = (\theta_1,\ldots,\theta_d)' \in \Theta$. After having observed $\{s^\sharp_{t_i}, i = 1,\ldots,n\}$, the pseudo maximum likelihood estimate $\hat{\vartheta}$ of θ is defined by

$$\hat{\vartheta} = \arg_{\theta\in\Theta}\left[\min\sum_{i=1}^{n}\left\{\frac{G(t_i,\theta)^{1-\delta} - s_{t_i}^{1-\delta}}{\rho_{t_i}}\right\}^2\right].$$

Let

$$\Xi(t,\theta) = \frac{G(t,\theta)^{1-\delta}}{\rho_t}$$

and

$$X(t) = \frac{S_t^{1-\delta}}{\rho_t} \quad and \quad x(t) = \frac{s_t^{1-\delta}}{\rho_t},$$

then, $\hat{\vartheta}$ may be rewritten as

$$\hat{\vartheta} = \arg_{\theta\in\Theta}\left[\min\sum_{i=1}^{n}\{\Xi(t_i,\theta) - x_{t_i}\}^2\right].$$

and $X(t)$ as

$$X(t) = \Xi(t,\theta) + \sigma\varepsilon_t.$$

2.3 Statistical Model under RT Type Recovery

We shall next derive statistical model under RT type recovery. We set the same assumptions to $P(0,t_i)$, $P^*(0,t_i)$ and $\hat{\Upsilon}(t_i)$ as those under RM type recovery. Using the observed spread $\hat{\Upsilon}(t_i)$, implied survival function $S^\sharp(t)$ under RT type recovery in stead of $G^\sharp(t)$ becomes

$$\begin{aligned}
S^\sharp(t) &= \frac{1}{1-\delta}\left(e^{-t\hat{\Upsilon}} - \delta\right) \\
&= \frac{1}{1-\delta}\left(e^{-t\cdot-\frac{1}{t}\log\left\{\frac{p(0,t_i)}{p^*(0,t_i)}+\sigma\rho_{t_i}\varepsilon_{t_i}\right\}} - \delta\right) \\
&= \frac{1}{1-\delta}\left(\frac{p(0,t)}{p^*(0,t)} - \delta\right) + \frac{1}{1-\delta}\sigma\rho_{t_i}\varepsilon_{t_i} \\
&= G^\sharp(t) + \frac{1}{1-\delta}\sigma\rho_{t_i}\varepsilon_{t_i}
\end{aligned}$$

After having observed $\{s^\sharp_{t_i}, i = 1,\ldots,n\}$, the pseudo maximum likelihood estimate $\hat{\vartheta}$ of θ is defined by

$$\hat{\vartheta} = \arg_{\theta\in\Theta}\left[\min\sum_{i=1}^{n}\left\{\frac{(1-\delta)\left(G^\sharp(t_i,\theta) - s^\sharp_{t_i}\right)}{\rho_{t_i}}\right\}^2\right].$$

Let

$$\Xi^{\sharp}(t,\theta) = \frac{(1-\delta)G^{\sharp}(t,\theta)}{\rho_t}$$

and

$$X^{\sharp}(t) = \frac{(1-\delta)S_t^{\sharp}}{\rho_t} \quad and \quad x^{\sharp}(t) = \frac{(1-\delta)s_t^{\sharp}}{\rho_t},$$

then, $\hat{\vartheta}$ may be rewritten as

$$\hat{\vartheta} = \arg_{\theta\in\Theta}\left[\min\sum_{i=1}^{n}\left\{\Xi^{\sharp}(t_i,\theta) - x_{t_i}^{\sharp}\right\}^2\right].$$

and $X^{\sharp}(t)$ as

$$X^{\sharp}(t) = \Xi^{\sharp}(t,\theta) + \sigma\varepsilon_t.$$

2.4 RM Type Recovery and RT Type Recovery

The pseudo maximum likelihood estimate under RM type recovery $\hat{\vartheta}_{RM}$ of θ is defined by

$$\begin{aligned}
\hat{\vartheta}_{RM} &= \arg_{\theta\in\Theta}\left[\min\sum_{i=1}^{n}\left\{\frac{G(t_i,\theta)^{1-\delta} - s_{t_i}^{1-\delta}}{\rho_{t_i}}\right\}^2\right] \\
&= \arg_{\theta\in\Theta}\left[\min\sum_{i=1}^{n}\left\{\frac{G(t_i,\theta)^{1-\delta} - \left(\left(e^{-t_i\hat{\Upsilon}(t_i)}\right)^{\frac{1}{1-\delta}}\right)^{1-\delta}}{\rho_{t_i}}\right\}^2\right] \\
&= \arg_{\theta\in\Theta}\left[\min\sum_{i=1}^{n}\left\{\frac{G(t_i,\theta)^{1-\delta} - e^{-t_i\hat{\Upsilon}(t_i)}}{\rho_{t_i}}\right\}^2\right]
\end{aligned}$$

The pseudo maximum likelihood estimate $\hat{\vartheta}_{RT}$ of θ is defined by

$$\begin{aligned}
\hat{\vartheta}_{RT} &= \arg_{\theta\in\Theta}\left[\min\sum_{i=1}^{n}\left\{\frac{(1-\delta)\left(G^{\sharp}(t_i,\theta) - s_{t_i}^{\sharp}\right)}{\rho_{t_i}}\right\}^2\right] \\
&= \arg_{\theta\in\Theta}\left[\min\sum_{i=1}^{n}\left\{\frac{(1-\delta)\left(G^{\sharp}(t_i,\theta) - \frac{1}{1-\delta}\left(e^{-t_i\hat{\Upsilon}(t_i)} - \delta\right)\right)}{\rho_{t_i}}\right\}^2\right] \\
&= \arg_{\theta\in\Theta}\left[\min\sum_{i=1}^{n}\left\{\frac{(1-\delta)G^{\sharp}(t_i,\theta) - \left(e^{-t_i\hat{\Upsilon}(t_i)} - \delta\right)}{\rho_{t_i}}\right\}^2\right] \\
&= \arg_{\theta\in\Theta}\left[\min\sum_{i=1}^{n}\left\{\frac{(1-\delta)G^{\sharp}(t_i,\theta) + \delta - e^{-t_i\hat{\Upsilon}(t_i)}}{\rho_{t_i}}\right\}^2\right]
\end{aligned}$$

G	Statistical RM $G^{1-\delta}$ $(\delta = 0.3)$	Statistical RT $(1-\delta)G+\delta$ $(\delta = 0.3)$	RM-RT $G^{1-\delta}$ $-(1-\delta)G+\delta$
1	1	1	0
0.99	0.992989	0.993	-0.000011
0.98	0.985958	0.986	-0.000042
0.97	0.978904	0.979	-0.000096
0.96	0.971829	0.972	-0.000171
0.95	0.964732	0.965	-0.000268
0.94	0.957612	0.958	-0.000388
0.93	0.950469	0.951	-0.000531
0.92	0.943304	0.944	-0.000696
0.91	0.936115	0.937	-0.000885
0.9	0.928902	0.93	-0.001098

Table 1: Simulation of RM model and RT model

Difference between $\hat{\vartheta}_{RM}$ and $\hat{\vartheta}_{RT}$ is the term

$$\begin{aligned} RM &: \quad G(t_i,\theta)^{1-\delta} \\ RT &: \quad (1-\delta)G^{\sharp}(t_i,\theta)+\delta \end{aligned}$$

We consider Weibull distribution for the survival function $G(t_i,\theta)$ and $G^{\sharp}(t_i,\theta)$. Table 1 is a simulation of RM model and RT model for $G = 1, 0.99, \dots, 0.9$.

From Table 1 we can see the difference $G^{1-\delta} - \{(1-\delta)G+\delta\}$ is negative in all cases. The smaller the survival probability G gets, the larger the difference $G^{1-\delta} - \{(1-\delta)G+\delta\}$ becomes.

2.5 Calibration

We may also suppose a model in which error term is not contained to the price of defaultable bond $P(0,t)$ (and of course $P(0,t)^*$). It means we assume that there is no noise in the market and $P(0,t)$ is determined with no error in the market. Under RM type recovery and the assumption of no error term, we can estimate the survival function by the following equation;

$$G(t) = G(t,\theta) \qquad \theta \in \Theta \subset R^d \quad d \geq 1.$$

Observe $\Upsilon(t_i)$ $i = 1,\dots,n$ and then estimate θ from

$$\hat{\vartheta} = \min_{\theta\in\Theta}\left[\sum_{i=1}^{n}\left\{G(t_i,\theta) - \left(e^{-t_i\Upsilon(t_i)}\right)^{\frac{1}{1-\delta}}\right\}^2\right].$$

We can estimate the survival function under RT type recovery in the similar way;

$$G(t) = G(t,\theta) \qquad \theta \in \Theta \subset R^d \quad d \geq 1.$$

Observe $\Upsilon(t_i)$ $i = 1,\dots,n$ and then estimate θ from

$$\hat{\vartheta} = \min_{\theta\in\Theta}\left[\sum_{i=1}^{n}\left\{G(t_i,\theta) - \frac{1}{1-\delta}\left(e^{-t_i\Upsilon(t_i)} - \delta\right)\right\}^2\right].$$

3 Asymptotic Properties

The aim of this section is to justify the estimating method described in the previous section. We shall first give a complete proof of the consistency of the pseudo maximum likelihood estimator. We employ the method of Wald(1949) which gives the proof of consistency of maximum likelihood estimator. Furthermore, we give a verification of the conditions for the consistency and the asymptotic normality of our estimator in the case of Weibull survival function.

3.1 Proof of the Consistency

In this subsection, we prove that under certain regularity conditions, our estimator $\hat{\vartheta}$ is strongly consistent.

Let,

$$Q(t, X_t, \theta) = \{\Xi(t, \theta) - X_t\}^2.$$

For any $\rho > 0$, let

$$Q(t, X_t, \theta, \rho) = \inf_{|\theta' - \theta| < \rho} Q(t, X_t, \theta'),$$

and, for any $r > 0$, let,

$$\varphi(t, X_t, r) = \inf_{|\theta| > r} Q(t, X_t, \theta).$$

We shall now impose some regularity conditions for the consistency:

[Regularity Condition C2] (conditions for the consistency)

1. If $\lim_{i\to\infty} \theta_i = \theta$,
$$\lim_{i\to\infty} Q(t, X_t, \theta_i) = Q(t, X_t, \theta)$$
for all (t, X_t) except on a set of measure zero under the true parameter vector θ^0. The null set may depend on θ, but not on the sequence $\{\theta_i, i = 1, 2, \ldots\}$.

2. For any pairs $\theta \neq \theta'$,
$$\Xi(t, \theta) \neq \Xi(t, \theta'), \quad \textit{at least one } t$$

3. For sufficiently small ρ, the expected values
$$\int_{-\infty}^{\infty} Q(t, X_t, \theta, \rho) dF(X_t, \theta^0) < \infty,$$
where θ^0 denotes the true parameter.

4. For the true parameter vector θ^0, we have
$$\int_{-\infty}^{\infty} Q(t, X_t, \theta^0) dF(X_t, \theta^0) < \infty.$$

5. If $\lim_{i\to\infty} |\theta_i| = \infty$, then $\lim_{i\to\infty} Q(t, X_t, \theta_i) = \infty$ for any (t, X_t) except perhaps on a fixed set of measure zero according to the true parameter vector θ^0.

Let, θ^0 be the true parameter vector, then it follows that,

$$E_{\theta^0}\{Q(t,X_t,\theta^0)\} = E_{\theta^0}\left\{(\Xi(t,\theta^0)-X_t)^2\right\} = \sigma^2 \tag{1}$$

and,

$$\begin{aligned} E_{\theta^0}\{Q(t,X_t,\theta)\} &= E_{\theta^0}\left\{(\Xi(t,\theta)-X_t)^2\right\} \\ &= E_{\theta^0}\left\{\left(\Xi(t,\theta^0)-X_t+\Xi(t,\theta)-\Xi(t,\theta^0)\right)^2\right\} \\ &= E_{\theta^0}\left\{\left(\Xi(t,\theta^0)-X_t\right)^2+\left(\Xi(t,\theta)-\Xi(t,\theta^0)\right)^2\right\} \\ &= \sigma^2+\left(\Xi(t,\theta)-\Xi(t,\theta^0)\right)^2 \end{aligned} \tag{2}$$

We consider the case where Θ is compact. Let $\tilde{\Theta}$ be a compact subset of Θ, which does not contain the true parameter vector θ^0.

[Lemma 1] For any $\theta \in \tilde{\Theta}$,

$$E_{\theta^0}\{Q(t,X_t,\theta)\} > E_{\theta^0}\{Q(t,X_t,\theta^0)\}. \tag{3}$$

Proof. It follows from **[C2]**-3 and **[C2]**-4 that the expectated values in (3) exsist. Lemma1 follows easily from (2) and **[C2]**-2.

[Lemma 2]

$$\lim_{\rho\to 0} E_{\theta^0}\left\{Q(t,X_t,\theta,\rho)\right\} = E_{\theta^0}\left\{Q(t,X_t,\theta)\right\} \tag{4}$$

Proof. It follows from **[C2]**-1 that,

$$\lim_{\rho\to 0} Q(t,X_t,\theta,\rho) = Q(t,X_t,\theta). \tag{5}$$

except on a set whose probability measure is zero. Since $Q(t,X_t,\theta,\rho)$ is a decreasing function of ρ, and $Q(t,X_t,\theta,\rho) > 0$, there exists an integrable function ψ such that

$$|Q(t,X_t,\theta,\rho)| \le \psi(t,X_t,\theta) \tag{6}$$

for all ρ. It follows from (5) and Lebesgue's Convergence Thorem[2] , that

$$\lim_{\rho\to 0}\int_{-\infty}^{\infty} Q(t,X_t,\theta,\rho)dF(X_t,\theta^0) = \int_{-\infty}^{\infty} Q(t,X_t,\theta)dF(X_t,\theta^0). \tag{7}$$

[2] To show $\lim_{\rho\to 0}\int_{-\infty}^{\infty} Q(t,X_t,\theta,\rho)dF(X_t,\theta^0) = \int_{-\infty}^{\infty} Q(t,X_t,\theta)dF(X_t,\theta^0)$ is equivalent to show, for any sequence $\{\rho_n;\rho_n\to 0\}$, $\lim_{n\to\infty}\int_{-\infty}^{\infty} Q(t,X_t,\theta,\rho_n)dF(X_t,\theta^0) = \int_{-\infty}^{\infty} Q(t,X_t,\theta)dF(X_t,\theta^0)$. Since $Q(t,X_t,\theta,\rho_1)$ is integrable from **[C2]**-3, then,(6) means for any sequense $\{\rho_n;\rho_n\to 0\}$, $|Q(t,X_t,\theta,\rho_n)| \le Q(t,X_t,\theta,\rho_1)$.

Thus, Lemma2 is proved.

[Lemma 3]

$$\lim_{r\to\infty} E_{\theta^0}\left\{\varphi(t,X_t,r)\right\} = \infty \tag{8}$$

Proof. According to **[C2]**-5,

$$\lim_{r\to\infty}\varphi(t,X_t,r) = \lim_{r\to\infty}\inf_{|\theta|>r} Q(t,X_t,\theta) = \infty \tag{9}$$

for any (t,X_t) except on a set of measure zero under the true parameter vector θ^0. $\varphi(t,X_t,r)$ is an increasing function of r. If there exists $r>0$ such that the measure F of a set $\{(t,X_t)\in\mathscr{X};\varphi(t,X_t,r)=\infty\}$ is positive, then, $\lim_{r\to\infty}\varphi(t,X_t,r)=\infty$ on a set $\{(t,X_t)\in\mathscr{X};\varphi(t,X_t,r)=\infty\}$. So, Lemma3

$$\lim_{r\to\infty}\int\varphi(t,X_t,r)dF(t,X_t,\theta^0) = \infty$$

obviously holds. Thus, we shall merely consider the case when $\varphi(t,X_t,r)<\infty$ on $\{(t,X_t)\in\mathscr{X}$. Since $\varphi(t,X_t,r)$ is an increasing function of r, we define, for any sequence $r_i; r_1 < r_2 < \ldots < r_i < \ldots$

$$\begin{aligned}
\varphi'(t,X_t,r_1) &= \varphi(t,X_t,r_1)\\
\varphi'(t,X_t,r_i) &= \varphi(t,X_t,r_i)-\varphi(t,X_t,r_{i-1}) \quad i=2,3,\ldots
\end{aligned}$$

Since

$$\varphi(t,X_t,r_i) = \sum_{j=1}^{i}\varphi'(t,X_t,r_j)$$

and

$$\varphi'(t,X_t,r_j)\geq 0,$$

$$\begin{aligned}
\lim_{r_i\to\infty}\int\varphi(t,X_t,r_i)dF(t,X_t,\theta^0) &= \lim_{r_i\to\infty}\int\sum_{j=1}^{i}\varphi'(t,X_t,r_j)dF(t,X_t,\theta^0) \qquad (10)\\
&= \lim_{r_i\to\infty}\sum_{j=1}^{i}\int\varphi'(t,X_t,r_j)dF(t,X_t,\theta^0)\\
&= \sum_{j=1}^{\infty}\int\varphi'(t,X_t,r_j)dF(t,X_t,\theta^0)\\
&= \int\sum_{j=1}^{\infty}\varphi'(t,X_t,r_j)dF(t,X_t,\theta^0)\\
&= \int\lim_{r_i\to\infty}\varphi(t,X_t,r_i)dF(t,X_t,\theta^0)\\
&= \infty
\end{aligned}$$

The second and fifth equality are from the change the order of infinite summation and integration and the 6th equal is from (9). Thus, Lemma3 is proved.

[Theorem 1] (The consistency theorem)

$$P_{\theta^0}\left\{\lim_{n\to\infty}\hat{\vartheta}_n = \theta^0\right\} = 1. \tag{11}$$

Proof. Let r_0 be a positive number chosen such that

$$E_{\theta^0}\{\varphi(t, X_t, r_0)\} > E_{\theta^0}\{Q(t, X_t, \theta^0)\}. \tag{12}$$

The existence of such a positive number follows from Lemma3.

Let $\breve{\Theta}$ be the subset of $\tilde{\Theta}$ consisting of all parameter vector θ of $\tilde{\Theta}$ for which $|\theta| \le r_0$. With each parameter vector $\theta \in \breve{\Theta}$, we associate a positive value ρ_θ such that

$$E_{\theta^0}\left\{Q(t, X_t, \theta, \rho_\theta)\right\} > E_{\theta^0}\left\{Q(t, X_t, \theta^0)\right\}. \tag{13}$$

The existence of such ρ_θ follows from Lemma1 and Lemma2.

To prove that $\hat{\theta}_n$ is strongly consistent we have to show that

$$P_{\theta^0}\left\{\frac{\inf_{\theta\in\tilde{\Theta}}\sum_{i=1}^{n} Q(t_i, X_{ti}, \theta)}{\sum_{i=1}^{n} Q(t_i, X_{ti}, \theta^0)} > 1\right\} = 1 - \varepsilon \tag{14}$$

for all n sufficiently large. This means that for all n sufficiently large the values of θ^0 which minimizes $\sum_{i=1}^{n} Q(t_i, X_{ti}, \theta^0)$, $\hat{\theta}_n$, belongs to neighborhood of θ^0 with probability one.

Since $\breve{\Theta}$ is compact, there exists a finite open covering of $\breve{\Theta}$

$$\{\mathscr{O}(\theta^{(l)}, \rho_\theta^{(l)}), l = 1, \ldots, L\},$$

where, $\mathscr{O}(\theta, \rho_\theta)$ is a sphere with center θ and radius ρ_θ ; $\breve{\Theta} \subset \mathscr{O}(\theta^{(1)}, \rho_\theta^{(1)}) + \ldots + \mathscr{O}(\theta^{(L)}, \rho_\theta^{(L)})$.

Clearly,

$$\inf_{\theta\in\tilde{\Theta}}\sum_{i=1}^{n} Q(t_i, X_{t_i}, \theta) \ge \min_{l=1,\ldots,L}\sum_{i=1}^{n} Q(t_i, X_{t_i}, \theta^{(l)}, \rho_\theta^{(l)}).$$

Hence

$$P_{\theta^0}\left\{\frac{\inf_{\theta\in\tilde{\Theta}}\sum_{i=1}^{n} Q(t_i, X_{ti}, \theta)}{\sum_{i=1}^{n} Q(t_i, X_{ti}, \theta^0)} > 1\right\} \ge P_{\theta^0}\left\{\frac{\min_{l=1,\ldots,L}\sum_{i=1}^{n} Q(t_i, X_{t_i}, \theta^{(l)}, \rho_\theta^{(l)})}{\sum_{i=1}^{n} Q(t_i, X_{ti}, \theta^0)} > 1\right\} \tag{15}$$

Thus from (15) we deduce

$$\begin{aligned} 1 - P_{\theta^0}&\left\{\frac{\inf_{\theta\in\tilde{\Theta}}\sum_{i=1}^{n} Q(t_i, X_{ti}, \theta)}{\sum_{i=1}^{n} Q(t_i, X_{ti}, \theta^0)} > 1\right\} \\ \le\ & 1 - P_{\theta^0}\left\{\frac{\min_{l=1,\ldots,L}\sum_{i=1}^{n} Q(t_i, X_{t_i}, \theta^{(l)}, \rho_\theta^{(l)})}{\sum_{i=1}^{n} Q(t_i, X_{ti}, \theta^0)} > 1\right\} \\ =\ & P_{\theta^0}\left\{\frac{\min_{l=1,\ldots,L}\sum_{i=1}^{n} Q(t_i, X_{t_i}, \theta^{(l)}, \rho_\theta^{(l)})}{\sum_{i=1}^{n} Q(t_i, X_{ti}, \theta^0)} \le 1\right\} \end{aligned} \tag{16}$$

By the strong law of large numbers, (12) and (13), the right-hand side of (16) is smaller than ε for all n sufficiently large. This completes the proof of Theorem1.

3.2 Verification of the Regularity Conditions for consistency

We shall now verify the regularity conditions for consistency and the asymptotic normality of our estimator when we apply Weibull distribution to the survival function. When we estimate the statistical model of Takahashi(2011), we consider exponential model and Weibull model for the parametric survival function. Needless to say, exponential distribution is the special case of Weibull distribution, we need only to verify the regularity conditions in the Weibull model. Weibull survival function is defined in the form of

$$G_{Wei}(\theta_1, \theta_2, t) = \exp\left\{-(t\theta_1)^{\theta_2}\right\}, \qquad t \geq 0.$$

We now verify the conditions of [C2] which is proved in the previous subsection. **Verification of the conditions of [C2].** We shall verify the conditions for consistency [C2] in the Weibull distribution case. Condition [C2]-1 holds clearly. To show [C2]-2, let $(\theta_1, \theta_2) \neq (\theta_1', \theta_2')$ and define $\Delta\theta_2 = \theta_2 - \theta_2'$. Then,

$$\begin{aligned}(t\theta_1)^{\theta_2} - (t\theta_1')^{\theta_2'} &= (t\theta_1)^{\theta_2+\Delta\theta_2} - (t\theta_1')^{\theta_2'} \\ &= t\theta_2'(t^{\Delta\theta_2}\theta_1^{\theta_2'+\Delta\theta_2} - \theta_1^{'\theta_2'}) \\ &\neq 0 \ \ at \ \ least \ \ one \ \ t,\end{aligned}$$

so that $(t\theta_1)^{\theta_2} \neq (t\theta_1')^{\theta_2'}$ at least one t. Since $G(t,\theta) = e^{-(t\theta 1)^{\theta 2}}$ is strictly monotonic, $(t\theta_1)^{\theta_2} \neq (t\theta_1')^{\theta_2'}$ implies $G(t,\theta) \neq G(t,\theta')$, and hence that $\Xi(t,\theta) \neq \Xi(t,\theta')$ at least one t.

We next show the existence of $E[Q(t,X_t,\theta^0)]$ ([C2]-4). The integrand of the expectation

$$\begin{aligned}E[Q(t,X_t,\theta^0)] &= \int_{-\infty}^{\infty} Q(t,x_t,\theta^0)dF(x_t,\theta^0) \\ &= \int_{-\infty}^{\infty} \{\Xi(t,\theta^0) - x_t\}^2 dF(x_t,\theta^0).\end{aligned}$$

is written as follows.

$$\begin{aligned}\{\Xi(t,\theta^0) - x_t\}^2 &= \left\{\frac{G(t,\theta^0)^{1-\delta}}{\rho_t} - x_t\right\}^2 \\ &= \left\{\frac{(e^{-(t\theta_1)^{\theta_2}})^{1-\delta}}{\rho_t} - x_t\right\}^2 \\ &= \left\{\frac{(e^{-\theta_1^{\theta_2}})^{1-\delta}(e^{-t^{\theta_2}})^{1-\delta}}{\rho_t} - \frac{s_t^{1-\delta}}{\rho_t}\right\}^2 \\ &= \left\{\frac{(e^{-\theta_1^{\theta_2}})^{1-\delta}(e^{-t^{\theta_2}})^{1-\delta}P^*(0,t)}{h_t} - \frac{s_t^{1-\delta}P^*(0,t)}{h_t}\right\}^2\end{aligned}$$

Note that $(e^{-\theta_1^{\theta_2}})^{1-\delta}$ is positive constant. From $0 \leq t$, $\theta_2 > 0$ and $0 < \delta < 1$, $(e^{-t^{\theta_2}})^{1-\delta} \leq 1$. It follows that $s_t^{1-\delta} \leq 1$ and $0 < P^*(0,t) \leq 1$. Therefore $\{\Xi(t,\theta^0) - x_t\}^2 \to \infty$ as $h_t \to 0$; $h_t \neq 0$ implies $E[Q(t,x_t,\theta^0)] < \infty$. So [C2]-4 holds. [C2]-3 can be verified in the same way. For [C2]-5, if $lim_{i\to\infty}|\theta_i| = \infty$, $\Xi(t,\theta_i) \to 0$. [C2]-5 holds clearly.

3.3 Verification of the Regularity Conditions for Asymptotic Normality

The asymptotic normality is proved completely by Takahashi(2011), so we describe the conditions for the asymptotic normality without proof.

[Regularity Condition C3] (conditions for the asymptotic normality)

1. The pseudo data will be sampled in a way that the time sequence $\{t_i, i = 1,2,\dots\}$ are selected so that there is a positive definite matrix $\Sigma(\theta^0)$ and $B(\theta^0)$ for which

$$\lim_{n\to\infty} \frac{1}{n} \sum_{i=1}^{n} Var\{\psi(\theta^0, t_i, X_{t_i})\} = \lim_{n\to\infty} \sigma^2 \Sigma(\theta^0, n) = \sigma^2 \Sigma(\theta^0) \tag{17}$$

where $(l,m)^{th}$ element of $\Sigma(\theta^0,n) = \frac{1}{n\sigma^2}\sum_{i=1}^{n} Var\{\psi(\theta^0,t_i,X_{t_i})\}$ is defined by

$$\frac{1}{n}\sum_{i=1}^{n} \frac{\partial}{\partial\theta_l}\Xi(t_i,\theta^0)\frac{\partial}{\partial\theta_m}\Xi(t_i,\theta^0)$$

and,

$$\lim_{n\to\infty} \frac{1}{n} \sum_{i=1}^{n} \left[\frac{\partial^2}{\partial\theta^2}\psi^{(l)}(\theta^0, t_i, X_{t_i})\right] = \lim_{n\to\infty} B(\theta^0, n) = B(\theta^0) \tag{18}$$

where $(k,m)^{th}$ element of $\frac{1}{n}\sum_{i=1}^{n}\left[\frac{\partial^2}{\partial\theta^2}\psi^{(l)}(\theta^0,t_i,X_{t_i})\right]$ is given by

$$\begin{aligned}
& \frac{1}{n}\sum_{i=1}^{n}\left[\frac{\partial^3}{\partial\theta_k\partial\theta_m\partial\theta_l}\Xi(t_i,\theta^0)\right]\left\{\Xi(t_i,\theta^0) - X_t\right\} \\
+ \quad & \frac{1}{n}\sum_{i=1}^{n}\left[\frac{\partial^2}{\partial\theta_m\partial\theta_l}\Xi(t_i,\theta^0)\right]\left[\frac{\partial}{\partial\theta_k}\Xi(t_i,\theta^0)\right] \\
+ \quad & \frac{1}{n}\sum_{i=1}^{n}\left[\frac{\partial^2}{\partial\theta_l\partial\theta_k}\Xi(t_i,\theta^0)\right]\left[\frac{\partial}{\partial\theta_m}\Xi(t_i,\theta^0)\right] \\
+ \quad & \frac{1}{n}\sum_{i=1}^{n}\left[\frac{\partial^2}{\partial\theta_k\partial\theta_m}\Xi(t_i,\theta^0)\right]\left[\frac{\partial}{\partial\theta_l}\Xi(t_i,\theta^0)\right]
\end{aligned}$$

and the $(k,m)^{th}$ element of $B(\theta^0,n)$ is defined by

$$\begin{aligned}
& \frac{1}{n}\sum_{i=1}^{n}\left[\frac{\partial^2}{\partial\theta_m\partial\theta_l}\Xi(t_i,\theta^0)\right]\left[\frac{\partial}{\partial\theta_k}\Xi(t_i,\theta^0)\right] \\
+ \quad & \frac{1}{n}\sum_{i=1}^{n}\left[\frac{\partial^2}{\partial\theta_l\partial\theta_k}\Xi(t_i,\theta^0)\right]\left[\frac{\partial}{\partial\theta_m}\Xi(t_i,\theta^0)\right] \\
+ \quad & \frac{1}{n}\sum_{i=1}^{n}\left[\frac{\partial^2}{\partial\theta_k\partial\theta_m}\Xi(t_i,\theta^0)\right]\left[\frac{\partial}{\partial\theta_l}\Xi(t_i,\theta^0)\right].
\end{aligned}$$

2. For every $\theta \in \Theta$, the time sequence $\{t_i, i = 1, 2, \dots\}$ are selected so that

$$\frac{1}{n}\sum_{i=1}^{n}\left[\frac{\partial^2}{\partial\theta_l\partial\theta_m}\Xi(t_i,\theta)\right]^2 \quad \textit{converges as } n \to \infty, \quad \textit{for all } l \textit{ and } m \tag{19}$$

$$\frac{1}{n}\sum_{i=1}^{n}\left[\frac{\partial^3}{\partial\theta_k\partial\theta_l\partial\theta_m}\Xi(t_i,\theta)\right]^2 \quad \textit{converges as } n \to \infty, \quad \textit{for all } k,\ l \textit{ and } m\ . \tag{20}$$

3. The sequence

$$\left\{\psi(\theta^0,t_i,X_{t_i}) = \left[\psi^{(1)}(\theta^0,t_i,X_{t_i}),\dots,\psi^{(d)}(\theta^0,t_i,X_{t_i})\right]', i = 1,2,\dots\right\}$$

obeys the multivariate central limit theorem;

$$\left[\sum_{i=1}^{n} Var\left\{\psi(\theta^0,t_i,X_{t_i})\right\}\right]^{-\frac{1}{2}}\sum_{i=1}^{n}\psi(\theta^0,t_i,X_{t_i}) \to N(0,I_d) \quad \textit{as } n \to \infty \tag{21}$$

where $0 = (0,\dots,0)$ is the d dimensional zero vector and I_d is the $d \times d$ identity matrix.

Next we shall verify the conditions of [C3] in the case of Weibull survival function.

Verification of the conditions [C3]. We next show the conditions of the asymptotic normality in the Weibull distribution case. To show condition[C3]-1, consider $\alpha = (\alpha 1, \alpha 2)' \neq 0$.

$$\begin{aligned}
&(\alpha 1,\alpha 2)\cdot\begin{pmatrix}\frac{\partial}{\partial\theta_1}\Xi\\ \frac{\partial}{\partial\theta_2}\Xi\end{pmatrix}\\
= \ & (\alpha 1,\alpha 2)\cdot\begin{pmatrix}\dfrac{\left(e^{(-t\theta_1)^{\theta_2}}\right)^{1-\delta}(1-\delta)(-t\theta_1)^{\theta_2}\frac{\theta_2}{\theta_1}}{\rho_t}\\ \dfrac{\left(e^{(-t\theta_1)^{\theta_2}}\right)^{1-\delta}(1-\delta)(-t\theta_1)^{\theta_2}\ln[-t\theta_1]}{\rho_t}\end{pmatrix}\\
= \ & \alpha 1\frac{\left(e^{(-t\theta_1)^{\theta_2}}\right)^{1-\delta}(1-\delta)(-t\theta_1)^{\theta_2}\frac{\theta_2}{\theta_1}}{\rho_t} + \alpha 2\frac{\left(e^{(-t\theta_1)^{\theta_2}}\right)^{1-\delta}(1-\delta)(-t\theta_1)^{\theta_2}\ln[-t\theta_1]}{\rho_t}\\
= \ & \frac{\left(e^{(-t\theta_1)^{\theta_2}}\right)^{1-\delta}(1-\delta)(-t\theta_1)^{\theta_2}}{\rho_t}\left(\alpha 1\frac{\theta_2}{\theta_1} + \alpha 2\ln[-t\theta_1]\right)
\end{aligned}$$

Since $\left(e^{(-t\theta_1)^{\theta_2}}\right)^{1-\delta} > 0$, $1-\delta > 0$, $(-t\theta_1)^{\theta_2} < 0$, $\theta_2/\theta_1 > 0$, $\ln[-t\theta_1] < 0$ and $\rho_t > 0$ by definition, $(\alpha 1,\alpha 2)\cdot\begin{pmatrix}\frac{\partial}{\partial\theta_1}\Xi\\ \frac{\partial}{\partial\theta_2}\Xi\end{pmatrix}$ is nonzero on an interval. Therefore,

$$\left\{(\alpha 1,\alpha 2)\cdot\begin{pmatrix}\frac{\partial}{\partial\theta_1}\Xi\\ \frac{\partial}{\partial\theta_2}\Xi\end{pmatrix}\right\}^2 = \alpha'\Sigma(\theta^0)\alpha > 0$$

Since this conclusion is true for any $\alpha \neq 0$, Σ must be positive definite matrix. The existence of a positive matrix $B(\theta^0)$ can be showed in the similar way .

Condition [C3]-2 holds by inspection of $\frac{1}{n}\sum_{i=1}^n \left(\frac{\partial^2}{\partial\theta_l\partial\theta_m}\Xi(t_i,\theta)\right)^2$ and $\frac{1}{n}\sum_{i=1}^n \left(\frac{\partial^3}{\partial\theta_k\partial\theta_l\partial\theta_m}\Xi(t_i,\theta)\right)^2$, $(l,m,k=1,2)$. In the case of $k=1$, $l=1$ and $m=1$, it follows that

$$\begin{aligned}
&\frac{1}{n}\sum_{i=1}^n\left(\frac{\partial^2}{\partial\theta_1\partial\theta_1}\Xi(t_i,\theta)\right)^2\\
&= \frac{1}{n}\sum_{i=1}^n\left(\frac{\left(e^{(-t\theta_1)^{\theta_2}}\right)^{1-\delta} t_i^2\,(1-\delta)\,(-t_i\theta_1)^{-2+\theta_2}\,(-1+\theta_2)\theta_2}{\rho_t}\right.\\
&\quad+\left.\frac{\left(e^{(-t\theta_1)^{\theta_2}}\right)^{1-\delta} t_i^2\,(1-\delta)^2\,(-t_i\theta_1)^{-2+2\theta_2}\,\theta_2^2}{\rho_t}\right)^2,
\end{aligned}$$

and

$$\begin{aligned}
&\frac{1}{n}\sum_{i=1}^n\left(\frac{\partial^3}{\partial\theta_1\partial\theta_1\partial\theta_1}\Xi(t_i,\theta)\right)^2\\
&= \frac{1}{n}\sum_{i=1}^n\left(-\frac{\left(e^{(-t\theta_1)^{\theta_2}}\right)^{1-\delta} t_i^3\,(1-\delta)\,(-t_i\theta_1)^{-3+\theta_2}\,(-2+\theta_2)(-1+\theta_2)\theta_2}{\rho_t}\right.\\
&\quad+\frac{\left(e^{(-t\theta_1)^{\theta_2}}\right)^{1-\delta} t_i^3\,(1-\delta)^3\,(-t_i\theta_1)^{-3+2\theta_2}\,(-1+\theta_2)\theta_2^2}{\rho_t}\\
&\quad-\frac{\left(e^{(-t\theta_1)^{\theta_2}}\right)^{1-\delta} t_i^3\,(1-\delta)^3\,(-t_i\theta_1)^{-3+3\theta_2}\,\theta_2^3}{\rho_t}\\
&\quad+\left.\frac{\left(e^{(-t\theta_1)^{\theta_2}}\right)^{1-\delta} t_i^3\,(1-\delta)^2\,(-t_i\theta_1)^{-3+2\theta_2}\,\theta_2^2(-2+2\theta_2)}{\rho_t}\right)^2
\end{aligned}$$

Since $0<t_i<T<\infty$, $\rho_{t_i}=\frac{(1-P^*(t_i))^\alpha}{P^*(t_i)}$, $0<P^*(t_i)<1$ and $\alpha>0$, $\frac{1}{n}\sum_{i=1}^n\left(\frac{\partial^2}{\partial\theta_1\partial\theta_1}\Xi(t_i,\theta)\right)^2$ and $\frac{1}{n}\sum_{i=1}^n\left(\frac{\partial^3}{\partial\theta_1\partial\theta_1\partial\theta_1}\Xi(t_i,\theta)\right)^2$ converges as $n\to\infty$. The rest case can be verified in the similar way.

By the equation (17), condition [C3]-3

$$\begin{aligned}
&\left[\sum_{i=1}^n Var\left\{\psi(\theta^0,t_i,X_{t_i})\right\}\right]^{-\frac{1}{2}}\sum_{i=1}^n\psi(\theta^0,t_i,X_{t_i})\\
&= \frac{1}{\sqrt{n}}\left[\sum_{i=1}^n Var\left\{\psi(\theta^0,t_i,X_{t_i})\right\}\right]^{-\frac{1}{2}}\sum_{i=1}^n\psi(\theta^0,t_i,X_{t_i})\to N(0,I_d)\quad as\ n\to\infty
\end{aligned}$$

is rearranged to

$$\frac{1}{\sqrt{n}}\sum_{i=1}^n\psi(\theta^0,t_i,X_{t_i})\to N(0,\sigma^2\Sigma(\theta^0))\quad as\ n\to\infty \tag{22}$$

We have already shown that

$$\frac{1}{n}\sum_{i=1}^{n} Var\{\psi(\theta^0, t_i, X_{t_i})\} \to \sigma^2\Sigma(\theta^0)$$

is non zero, so that only condition

$$\lim_{n\to\infty}\frac{1}{n}\sum_{i=1}^{n}\int_{\|\psi\|>\varepsilon\sqrt{n}} \|\psi\|^2 dF_i(\psi) = 0 \tag{23}$$

for any $\varepsilon > 0$ is needed, where $\|\cdot\|$ is the Euclidean norm. Since

$$\begin{aligned}
\|\psi\| &= \sqrt{|\psi^{(1)}|^2 + |\psi^{(2)}|^2} \\
&= \sqrt{|\frac{\partial}{\partial\theta_1}\Xi(t_i,\theta^0)\,(\Xi(t_i,\theta^0)-x)\,|^2 + |\frac{\partial}{\partial\theta_2}\Xi(t_i,\theta^0)\,(\Xi(t_i,\theta^0)-x)\,|^2},
\end{aligned}$$

it follows that for any $\varepsilon > 0$,

$$\frac{1}{n}\sum_{i=1}^{n}\int_{\|\psi\|>\varepsilon\sqrt{n}} \|\psi\|^2 dF_i(\psi) \to 0$$

as $n \to \infty$, from which the condition (23) follows in the Weibull case.

4 Numerical Results

In this section, we present some numerical results. We apply our model to real data to estimate default probabilities(survival probabilities). We describe data sources at first, and then show the results of comparative study varying α and δ. We also compare results on different recovery models and on two parametric distributions. Then we show estimating results and bootstrap results using 8 individual companies' data and 5rating classes' data on April 16, 2004. Finally we apply our statistical estimating method to time series data(January 2008 to March 2009).

4.1 Data Description

Corporate bond prices and government bond prices were obtained from Bloomberg Financial Markets and Japan Securities Dealers Association covering the period April 2003 to March 2009. Specifically, we collected daily price data for OTC Bond Transactions. We also obtained bond issuer's rating information from Bloomberg. Rating and Investment Information, Inc. (R & I) is one of the leading rating companies in Japan and it grade for bond issuers by the rating categories from AAA to CCC. The data we used in the comparative study is bond price data of Mitsui & Co., LTD. also known as Mitsui Bussan on April 16, 2004. We use 1520 individual companies' data for estimating survival probabilities on one day, and 5 individual companies' data for time series analysis.

4.2 Selection of a constant α the recovery rate δ

In this subsection, we show the results of comparative study varying α and δ. For the purpose of equalizing the other conditions, we use only Weibull distribution for the survival function $G(\theta,t)$ in this subsection.

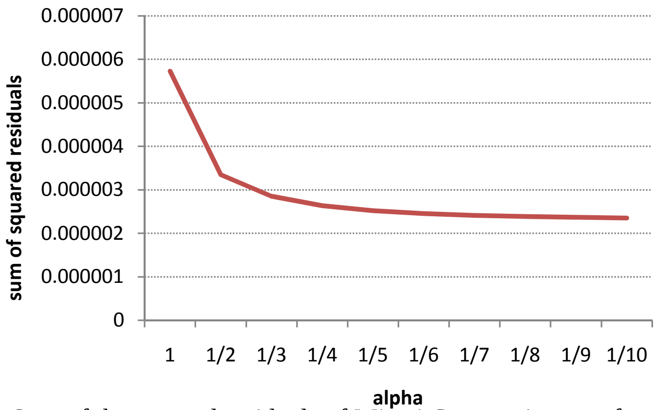

Figure 1: Sum of the squared residuals of Mitsui Corporation as a function of α

A constant α is defined by

$$\rho_t = \frac{[1 - P^*(0,t)]^\alpha}{P^*(0,t)}.$$

The restriction at the maturity $\rho_0 = 0$ suggests that $\alpha > 0$. We first consider $\alpha = 1$. (It means $\rho_t = \frac{1-P^*(0,t)}{P^*(0,t)}$.) Since $P^*(0,t)$ is near to 1 with small t and then ρ_t takes a small value, the weight of the data becomes very large with small t. We get biased estimating results and larger values of sum of squared residuals. Next, we calculate sum of the squared residuals with different α ; $\alpha = 1, \frac{1}{2}, \ldots, \frac{1}{10}$ Figure 1 shows the sum of the squared residuals of Mitsui Corporation as a function of α. In this case, we see that the sum of squared residuals is decreasing as $\alpha \to 0$. We get similar results in most individual company cases. But as mentioned above, α is supposed to take positive value. Since the sums of the squared residuals with $\alpha = \frac{1}{3}$ are enough small in most cases, we set α to be $\frac{1}{3}$ throughout the rest of this section.

We next consider about recovery rate. In our model, recovery rate, δ, is assumed to be non-random. We use δ when we calculate the survival function

$$\begin{aligned} G(t) &= \exp\{t\Upsilon(t : P*, P)\}^{\frac{1}{1-\delta}} \\ &= \left(\frac{P}{P*}\right)^{\frac{1}{1-\delta}}, \end{aligned}$$

and of course, when calculate pseudo random sample(data), s_{t_i} in the statistical model;

$$s_t = \left(\exp\left\{\log\left[\frac{P}{P*} + \sigma\rho_t\varepsilon_t\right]\right\}\right)^{\frac{1}{1-\delta}}$$

Figure1 shows the variation of the survival probability of Mitsui Corporation with the time to maturity for $\delta = 0.15$, $\delta = 0.3$, $\delta = 0.45$. The plot "Observed" is the values of survival probabilities calculated from the bond prices. The line "Model" is the values of estimated survival probabilities. Table 2 shows the values of the estimated parameters of Weibull model for $\delta = 0.15$, $\delta = 0.3$, $\delta = 0.45$. The recovery rate δ becomes larger, survival probability tends to smaller. An interesting results is that the values of θ_2 are the same with different δ. Since θ_2 is the shape parameter of Weibull distribution, we

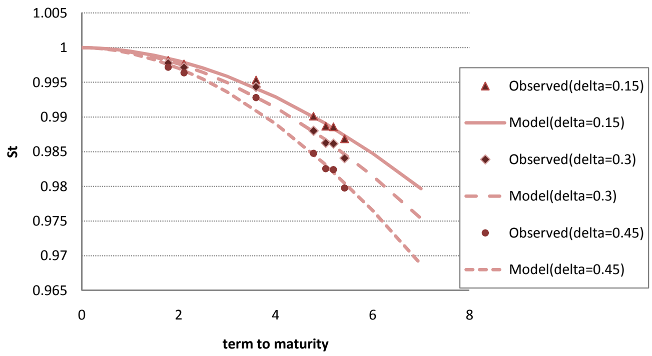

Figure 2: Variation of the survival probability with the time to maturity for different δ

δ	θ_1	θ_2
0.15	0.01825	1.88807
0.3	0.02023	1.88807
0.45	0.02298	1.88807

Table 2: Estimated parameters of Weibull model for different δ

see that the shape of the survival function does not vary with the recovery rate δ in our model.

The similar results are obtained for the statistical model under RT type recovery (Figure2, Table 3). Throughout the rest of this section, we set the recovery rate δ to be 0.3.

4.3 Selection of the parametric model

In this subsection, we consider the parametric model for the survival function of our estimating model. We use two parametric models for the survival function: exponential distribution and Weibull distribution. We set the survival function with exponential distribution G_{exp} and the survival function with Weibull distribution G_{Wei}. $G_{exp}(\theta,t)$ and $G_{Wei}(\theta_1,\theta_2,t)$ is defined by

$$G_{exp}(\theta,t) \quad = \quad \exp\{-t\theta\}, \qquad t \geq 0$$

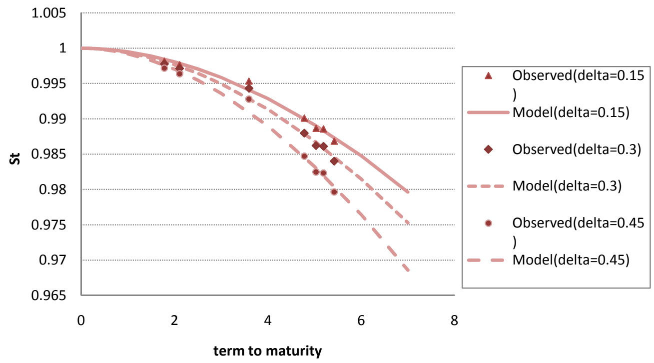

Figure 3: Variation of the survival probability with the time to maturity for different δ

δ	θ_1	θ_2
0.15	0.01825	1.88807
0.3	0.02023	1.88807
0.45	0.02298	1.88807

Table 3: Estimated parameters of Weibull model for different δ Mitsui Corp

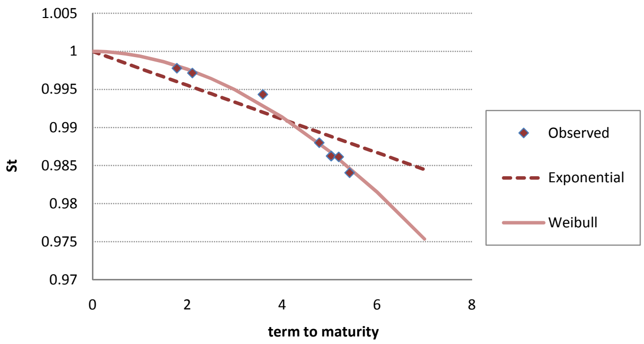

Figure 4: Estimated survival functions of Mitsui & Co. for the exponential and the Weibull model

$$G_{Wei}(\theta_1,\theta_2,t) \quad = \quad \exp\left\{-(t\theta_1)^{\theta_2}\right\}, \qquad t \geq 0$$

respectively. Figure 4 shows the estimated survival functions of Mitsui & Co. for the exponential and the Weibull model. Figure 5 shows the cumulative value of the sum of squared residuals for the exponential(dotted line) and the Weibull model(staggered line).

It is clear that the exponential survival function is estimated far from the ”Observed” plot.(See Figure 4. ”Observed” plots indicate the survival probabilities calculated from the interest yields.) And the sum of squared residuals from the exponential survival function is much larger than that of the Weibull survival function. The result is of course from the fact that the exponential distribution is allowed less 1 parameter than Weibull distribution. We shall thus apply Weibull distribution for the rest of our analysis.

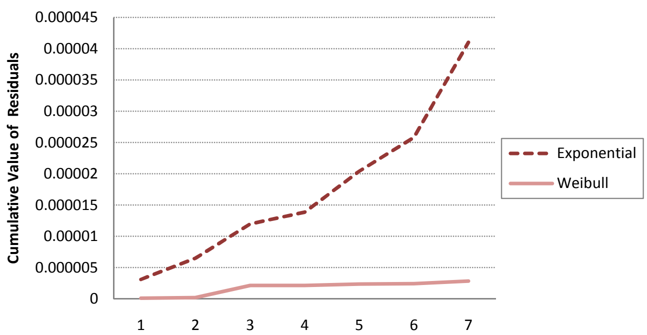

Figure 5: Cumulative value of the sum of squared residuals for the exponential and the Weibull model

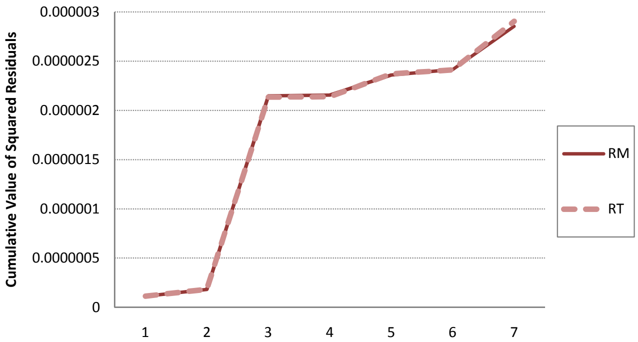

Figure 6: Cumulative value of the sum of squared residuals under RM type recovery and under RT type recovery

4.4 Two recovery models; RM type recovery & RT type recovery

As mentioned in chapter 2, the pseudo maximum likelihood estimate of statistical model under RM type recovery and RT type recovery are defined by

$$\hat{\vartheta} = \arg_{\theta\in\Theta}\left[\min\sum_{i=1}^{n}\left\{\frac{G(t_i,\theta)^{1-\delta}-s_{t_i}^{1-\delta}}{\rho_{t_i}}\right\}^2\right]$$

$$\hat{\vartheta} = \arg_{\theta\in\Theta}\left[\min\sum_{i=1}^{n}\left\{\frac{(1-\delta)\left(G^{\sharp}(t_i,\theta)-s_{t_i}^{\sharp}\right)}{\rho_{t_i}}\right\}^2\right]$$

respectively, where,

$$s(t) = \left(e^{-t\Upsilon}\right)^{\frac{1}{1-\delta}}$$

$$s^{\sharp}(t) = \frac{1}{1-\delta}\left(e^{-t\Upsilon}-\delta\right).$$

We also defined the pseudo maximum likelihood estimate of calibration model under RM type recovery and RT type recovery as follows:

$$\hat{\vartheta} = \min_{\theta\in\Theta}\left[\sum_{i=1}^{n}\left\{G(t_i,\theta)-\left(e^{-t_i\Upsilon(t_i)}\right)^{\frac{1}{1-\delta}}\right\}^2\right]$$

$$\hat{\vartheta} = \min_{\theta\in\Theta}\left[\sum_{i=1}^{n}\left\{G(t_i,\theta)-\frac{1}{1-\delta}\left(e^{-t_i\Upsilon(t_i)}-\delta\right)\right\}^2\right].$$

Figure 6 shows the cumulative value of the sum of squared residuals of Mitsui & Co. for the statistical model under RM type recovery(solid line) and that for the statistical model under RT type recovery(dotted line).

We see from Figure 6 that the lines are lying on top each other and that the difference between the sum of squared residuals from RM type recovery and RT type recovery is very small. Table 4 indicates the absolute value of residuals and the sum of the squared residuals when we apply statistical and calibration model to 10 individual companies. The above table is for the statistical model and the bottom is for the calibration model.

Company	Statistical Model RM	RT	RM/RT
Toyota	0.0000127	0.0000127	RT
NTT Data	0.0016643	0.0015528	RT
Mitsui & Co.	0.0004898	0.0004917	RM
Asahi Breweries	0.0007438	0.0006811	RT
Kajima	0.0008121	0.0008242	RM
NEC	0.0009871	0.0009972	RM
Kintetsu	0.0015136	0.0015379	RM
Sumitomo Realty	0.0006707	0.0006997	RT
Cosmo Oil	0.0000993	0.0001037	RM
JAL	0.0015214	0.0015431	RM
Company	**Calibration RM**	**RT**	**RM/RT**
Toyota	0.0000120	0.0000118	RM
NTT Data	0.0015446	0.0015477	RM
Mitsui & Co.	0.0004964	0.0004982	RM
Asahi Breweries	0.0005521	0.0005585	RM
Kajima	0.0006030	0.0006375	RM
NEC	0.0008421	0.0007921	RM
Kintetsu	0.0014051	0.0013742	RM
Sumitomo Realty	0.0006942	0.0007389	RM
Cosmo Oil	0.0000794	0.0000984	RM
JAL	0.0011855	0.0012178	RM

Table 4: The mean of the absolute value of residuals for the statistical model and calibration under RM type recovery and RT type recovery. Residuals are calculated by $\hat{G}(t) - G(t)$. The rightmost item is the model name which take the smaller value of the sum of squared residuals.

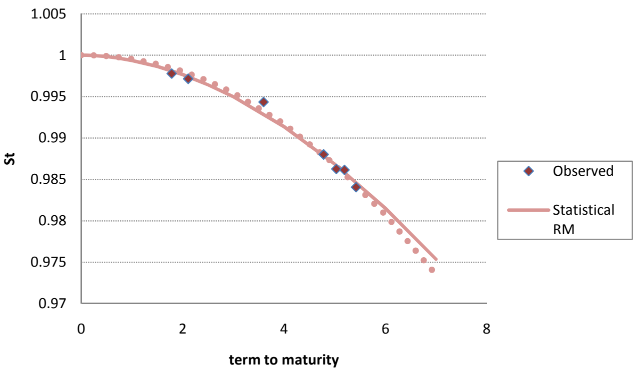

Figure 7: Estimated survival function of Mitsui Corporation for the statistical model and the calibration model

The rightmost item is the model name which takes the smaller value of the sum of squared residuals.

We see that RM model and RT model have little difference in residuals, but the following results are interesting. When we apply calibration estimating model, residuals of the model under RM type recovery are calculated smaller than that under RT type recovery for all 10 companies, though when applying statistical model, we can't say which model has smaller residuals.

4.5 Statistical Model and Calibration

Our estimators of statistical model under RM type recovery and calibration model under RM type recovery are written as follows

$$\hat{\vartheta} = \arg_{\theta\in\Theta}\left[\min\sum_{i=1}^{n}\left\{\frac{G(t_i,\theta)^{1-\delta}-s_{t_i}^{1-\delta}}{\rho_{t_i}}\right\}^2\right]$$

$$\hat{\vartheta} = \min_{\theta\in\Theta}\left[\sum_{i=1}^{n}\{G(t_i,\theta)-s_{t_i}\}^2\right]$$

respectively, where

$$s_{t_i} = \left(e^{-t_i\Upsilon(t_i)}\right)^{\frac{1}{1-\delta}}.$$

Figure 7 shows the estimated survival function of Mitsui Corporation for the two models; the statistical model (solid line) and the calibration model(dotted line) and Figure 8 shows the cumulative values of the sum of squared residuals for the two models, where we take RM type recovery model.

We find that the residuals calculated by the calibration model are smaller than that by the statistical model for all 10 companies(See Table 4). We also find that the cumulative values of the sum of squared residuals for the statistical model are take smaller values than for the calibration model in the first and second times of cumulation, and the larger values after the third cumulation(See Figure7). It is because we defined h_t (or ρ_t) as a function of the price of default free bond $P^*(0,t)$ and a constant α for our statistical model. When the data is close to the maturity, $P^*(0,t)$ becomes nearly 1 and so ρ_t gets near to 0. That is the reason of the behavior of the cumulative values of the sum of squared residuals.

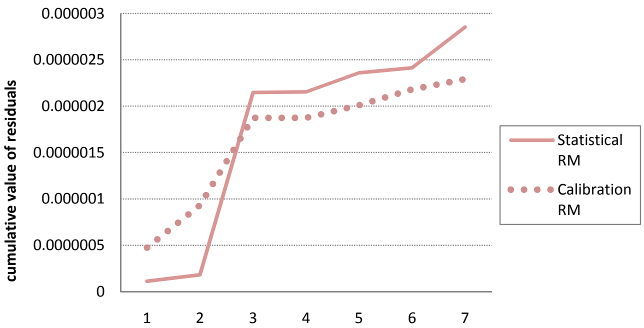

Figure 8: Cumulative values of the sum of squared residuals for the statistical model and the calibration model

	RM		RT	
	θ_1	θ_2	θ_1	θ_2
$\tilde{\theta}^{(Boots)}$	0.02040	1.89376	0.02051	1.89701
$\hat{\theta}$	0.02023	1.88807	0.02030	1.89005
$\widetilde{Bias}(\hat{\theta})^{(Boots)}$	0.000174	0.00569	0.000211	0.00696
$\widetilde{Var}(\hat{\theta})^{(Boots)}$	0.000010	0.01594	0.000010	0.01600
$\widetilde{SD}_{\hat{\theta}}(t)^{(Boots)}$	0.00319	0.12627	0.00321	0.12651

Table 5: The parameter estimate of the bootstrap estimation for Mitsui & Co. using statistical model. The notations are given in Appendix.

The bootstrap analysis showed that the estimate of statistical model has some warp. Table 5 and Table 6 indicate the parameter estimate of the bootstrap estimation for Mitsui & Co.. Table 5 is for the statistical model and Table 6 is for calibration.

Figure 9 shows the distribution of bootstrap parameter estimate of Mitsui & Co. for statistical model and Figure 10 shows the distribution of bootstrap parameter estimate for calibration. The left graph in Figure 9 and Figure 10 shows the histogram of the bootstrap estimate $\hat{\vartheta}_1{}^{(m)}$ and the right graph shows that of $\hat{\vartheta}_2{}^{(m)}$.

As seen from a comparison between Figure 9 and Figure 10, the distribution for statistical model has some warp, although that for calibration doesn't.

	RM		RT	
	θ_1	θ_2	θ_1	θ_2
$\tilde{\theta}^{(Boots)}$	0.02567	2.10442	0.02573	2.10599
$\hat{\theta}$	0.02543	2.0954	0.02554	2.0988
$\widetilde{Bias}(\hat{\theta})^{(Boots)}$	0.000242	0.00907	0.000186	0.00724
$\widetilde{Var}(\hat{\theta})^{(Boots)}$	0.000018	0.02878	0.000018	0.02901
$\widetilde{SD}_{\hat{\theta}}(t)^{(Boots)}$	0.00423	0.16966	0.00425	0.17033

Table 6: The parameter estimate of the bootstrap estimation for Mitsui & Co. using calibration. The notations are given in Appendix.

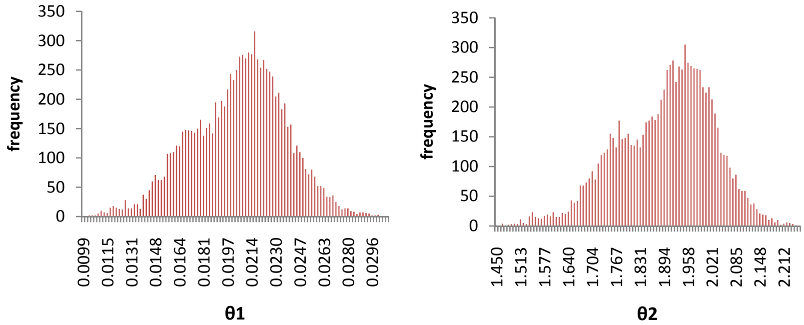

Figure 9: Distribution of bootstrap parameter estimate for statistical model

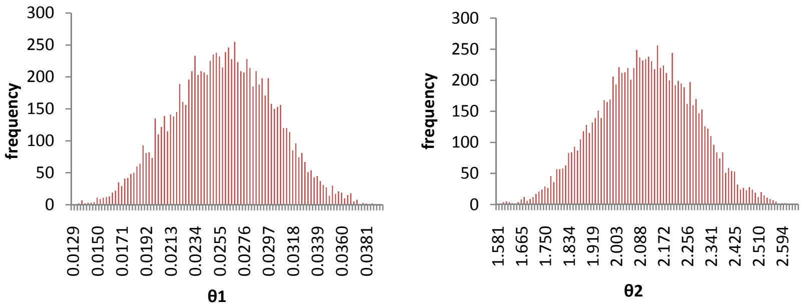

Figure 10: Distribution of bootstrap parameter estimate for calibration

4.6 Estimating results for 8 individual companies

We now present estimating results for individual companies. We set α to be $\frac{1}{3}$ and the recovery rate δ to be 0.3, and use statistical model of Weibull survival function under RM type recovery. We analyze the following 8 individual companies; Toyota Motor Corporation, NTT Data Corporation, Mitsui & Co., Ltd., Asahi Breweries, Ltd., Kajima Corporation, Kintetsu Corporation, Japan Airlines Co., Ltd. and Cosmo Oil Co., Ltd.. Table 7 indicates the data information and the parameter estimates of 8 individual companies. Figure 11 shows the estimated survival function of 8 individual companies.

Toyota's estimated Weibull shape parameter $\hat{\theta}_2$ is the smallest of the 8 companies and it is nearly 1. That means straight survival function and the almost constant hazard rate(constant instantaneous default probabilities). When the shape parameter $\hat{\theta}_2$ is larger than 1, the instantaneous default probabilities decline as $t \longrightarrow 0$.

Table 8 indicates the bootstrap parameter estimates of 8 individual companies. Figure 12 shows the distribution of bootstrap parameter estimate of 8 individual companies. The upper graph shows the histogram of the bootstrap estimate $\hat{\vartheta}_1{}^{(m)}$ and the lower graph shows that of $\hat{\vartheta}_2{}^{(m)}$.

We see from Table7 and Figure12 that there is a close relation between the distribution of bootstrap estimates and the number of data. The number of Toyota's bonds being issued on April 15, 2004 was only 3. That explains why the estimated bootstrap distribution of Toyota is quite different from normal distribution. It seems the larger amount of data(bonds) being used, the more like the bootstrap distribution is normal distribution. The bootstrap distribution of Mitsui & Co. looks like to have two peaks. The number of data and the existence of outlier data may explain this.

Company	Data		Estimating results		
	Rating	Observations	$\hat{\theta}_1$	$\hat{\theta}_2$	mean(|residuals |)
Toyota	AAA	3	0.0011	1.0749	0.000013
NTT Data	AA	8	0.0165	1.9923	0.001664
Mitsui & Co.	AA	7	0.0202	1.8881	0.000490
Asahi Breweries	A	10	0.0207	1.8545	0.000744
Kajima	A	11	0.0197	1.5656	0.000812
Kintetsu	BBB	15	0.0218	1.5364	0.001514
JAL	BBB	9	0.0297	1.6017	0.001521
Cosmo Oil	BB	9	0.0242	1.3502	0.000099

Table 7: Information about the data and the parameter estimates of 8 individual companies. The items "Rating" and "Observations" are the issuer's rating on April 15, 2004 and the number of a company's bonds being issued on April 15, 2004, respectively. "$\hat{\theta}_1$" and "$\hat{\theta}_2$" are parameter estimates and the rightmost item is the mean of the absolute values of residuals.

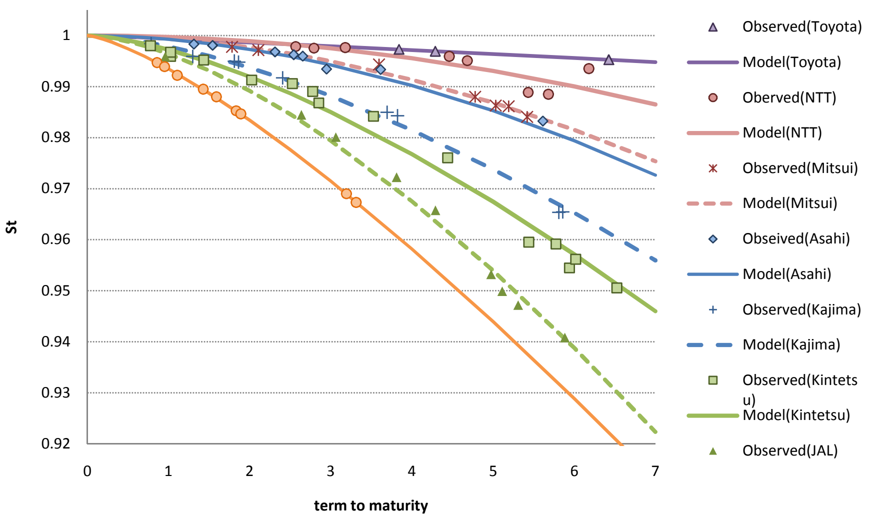

Figure 11: Estimated survival function and 'Observed' survival probabilities of 8 individual companies. "Observed" plots are survival probabilities calculated from interest yield and "Model" line is estimated survival function of our estimating model.

	Toyota	NTT Data	Mitsui	Asahi
$\tilde{\theta}_1^{(Boots)}$	0.00108	0.01824	0.02040	0.02078
$\hat{\theta}_1$	0.00108	0.01653	0.02023	0.02067
$\widetilde{Bias}(\hat{\theta}_1)^{(Boots)}$	0.00000	0.00171	0.00017	0.00011
$\widetilde{Var}(\hat{\theta}_1)^{(Boots)}$	0.00000	0.00012	0.00001	0.00001
$\widetilde{SD}_{\hat{\theta}_1}(t)^{(Boots)}$	0.00008	0.01113	0.00319	0.00353
$\tilde{\theta}_2^{(Boots)}$	1.07462	2.04300	1.89376	1.85723
$\hat{\theta}_2$	1.07492	1.99226	1.88807	1.85452
$\widetilde{Bias}(\hat{\theta}_2)^{(Boots)}$	-0.00030	0.05074	0.00569	0.00271
$\widetilde{Var}(\hat{\theta}_2)^{(Boots)}$	0.00021	0.30568	0.01594	0.01484
$\widetilde{SD}_{\hat{\theta}_2}(t)^{(Boots)}$	0.01435	0.55291	0.12627	0.12182

	Kajima	Kintetsu	JAL	Cosmo Oil
$\tilde{\theta}_1^{(Boots)}$	0.01968	0.02181	0.02976	0.03553
$\hat{\theta}_1$	0.01965	0.02176	0.02971	0.02418
$\widetilde{Bias}(\hat{\theta}_1)^{(Boots)}$	0.00003	0.00005	0.00005	0.01135
$\widetilde{Var}(\hat{\theta}_1)^{(Boots)}$	0.00000	0.00000	0.00001	0.00000
$\widetilde{SD}_{\hat{\theta}_1}(t)^{(Boots)}$	0.00182	0.00195	0.00248	0.00159
$\tilde{\theta}_2^{(Boots)}$	1.56630	1.53760	1.60337	1.59484
$\hat{\theta}_2$	1.56563	1.53641	1.60171	1.35018
$\widetilde{Bias}(\hat{\theta}_2)^{(Boots)}$	0.00067	0.00119	0.00166	0.24466
$\widetilde{Var}(\hat{\theta}_2)^{(Boots)}$	0.00309	0.00324	0.00430	0.00099
$\widetilde{SD}_{\hat{\theta}_2}(t)^{(Boots)}$	0.05564	0.05697	0.06560	0.03145

Table 8: The parameter estimate of the bootstrap estimation for 8 individual companies. The notations are given in Appendix.

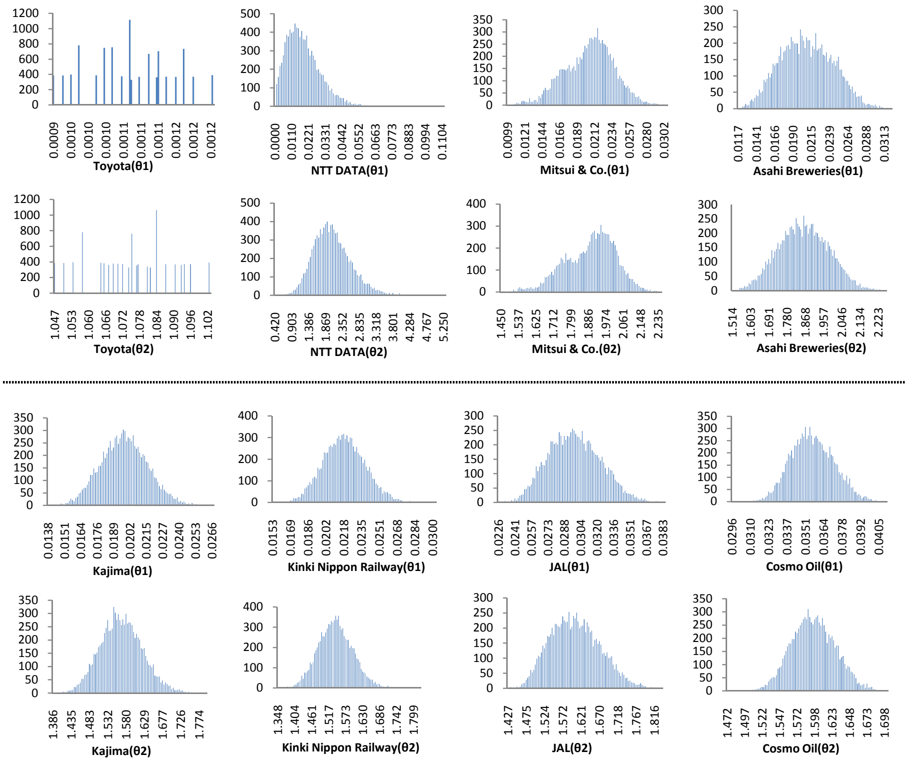

Figure 12: Distributions of bootstrap parameter estimates for 8 individual companies. The upper graph shows the histogram of the bootstrap estimate $\hat{\vartheta}_1^{(m)}$ and the lower graph shows that of $\hat{\vartheta}_2^{(m)}$.

	Data	Estimating results		
Class	Observations	$\hat{\theta}_1$	$\hat{\theta}_2$	mean(|residuals |)
AAA	8	0.0018	1.1348	0.001123
AA	523	0.0010	0.9297	0.002211
A	563	0.0140	1.4778	0.003447
BBB	381	0.0121	1.2030	0.449742
BB	45	0.0263	1.2351	0.936139

Table 9: Information about the data and the representative parameter estimates of 5 rating classes. The items "Class" and "Observations" are the rating class and the number of corporate bonds belonging to the class on April 15, 2004, respectively. "$\hat{\theta}_1$" and "$\hat{\theta}_2$" are parameter estimates and the rightmost item is the mean of the absolute values of residuals.

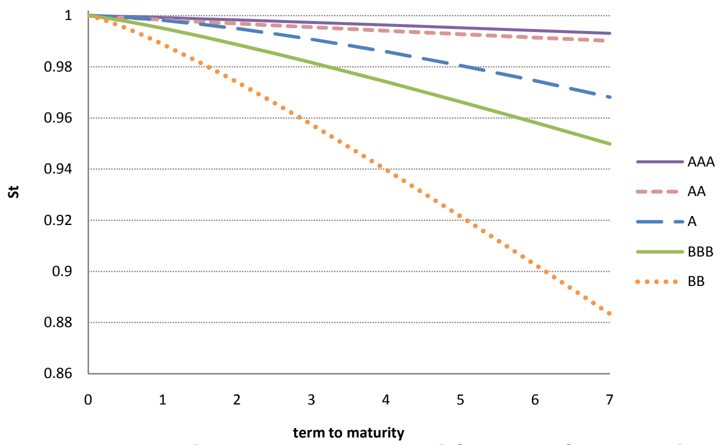

Figure 13: Estimated representative survival function of 5 rating classes.

4.7 Estimating results of 5 rating classes

In this subsection, we present estimating results for the data of 5 rating classes. As in the previous subsection, we set α to be $\frac{1}{3}$ and the recovery rate δ to be 0.3, and use statistical model of Weibull survival function under RM type recovery. We analyze 5 rating classes; AAA, AA, A, BBB, BB. Table 9 indicates the information about the data and the representative parameter estimates of 5 rating classes. Figure 13 shows the estimated representative survival function of 5 rating classes. Figure 14 contains 5 separate graphs of the estimated representative survival functions of rating classes. "Observed" plots indicate the survival probabilities calculated from the interest yields and the "Model" line is the estimated representative survival function of rating classes.

Table 10 indicates the bootstrap parameter estimates of 5 rating classes. Figure 15 shows the distribution of bootstrap parameter estimates of 5 rating classes. The upper graph shows the histogram of the bootstrap estimate $\hat{\vartheta}_1{}^{(m)}$ and the lower graph shows that of $\hat{\vartheta}_2{}^{(m)}$.

As seen in the individual companies' cases, there seems to be a close relation between the distribution of bootstrap estimates and the number of data. It seems the larger amount of data(bonds) being used, the more like the bootstrap distribution is normal distribution. Though many bonds (523 bonds) were issued by the "AA" class's companies, the bootstrap distribution of "AA" class looks like to have some warp. (See Figure 15.) This may be explained by the existence of outlier data. When we look Figure 14 carefully, we

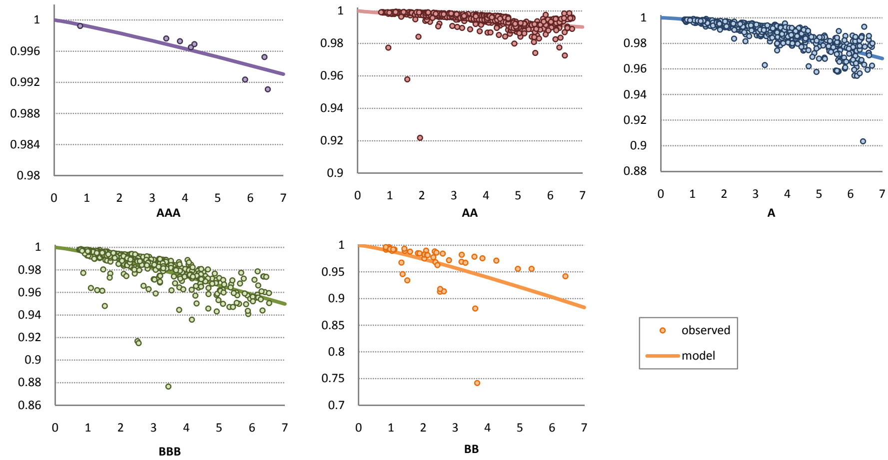

Figure 14: 5 separate graphs of the estimated representative survival functions of rating classes. ”Observed” plots indicate the survival probabilities calculated from the interest yields and the ”Model” line is the estimated representative survival function of rating classes.

	AAA	AA	A	BBB	BB
$\tilde{\theta}_1^{(Boots)}$	0.00238	0.00120	0.01406	0.01222	0.02775
$\hat{\theta}_1$	0.00180	0.00100	0.01402	0.01211	0.02634
$\widetilde{Bias}(\hat{\theta}_1)^{(Boots)}$	0.00059	0.00021	0.00003	0.00011	0.00141
$\widetilde{Var}(\hat{\theta}_1)^{(Boots)}$	0.00000	0.00000	0.00000	0.00001	0.00022
$\widetilde{SD}_{\hat{\theta}_1}(t)^{(Boots)}$	0.00223	0.00091	0.00102	0.00235	0.01480
$\tilde{\theta}_2^{(Boots)}$	1.15810	0.93420	1.47868	1.20364	1.23848
$\hat{\theta}_2$	1.13481	0.92966	1.47782	1.20300	1.23508
$\widetilde{Bias}(\hat{\theta}_2)^{(Boots)}$	0.02329	0.00454	0.00086	0.00064	0.00340
$\widetilde{Var}(\hat{\theta}_2)^{(Boots)}$	0.04835	0.01479	0.00133	0.00482	0.08222
$\widetilde{SD}_{\hat{\theta}_2}(t)^{(Boots)}$	0.21991	0.12164	0.03650	0.06941	0.28676

Table 10: The parameter estimate of the bootstrap estimation for 5 rating classes. The notations are given in Appendix.

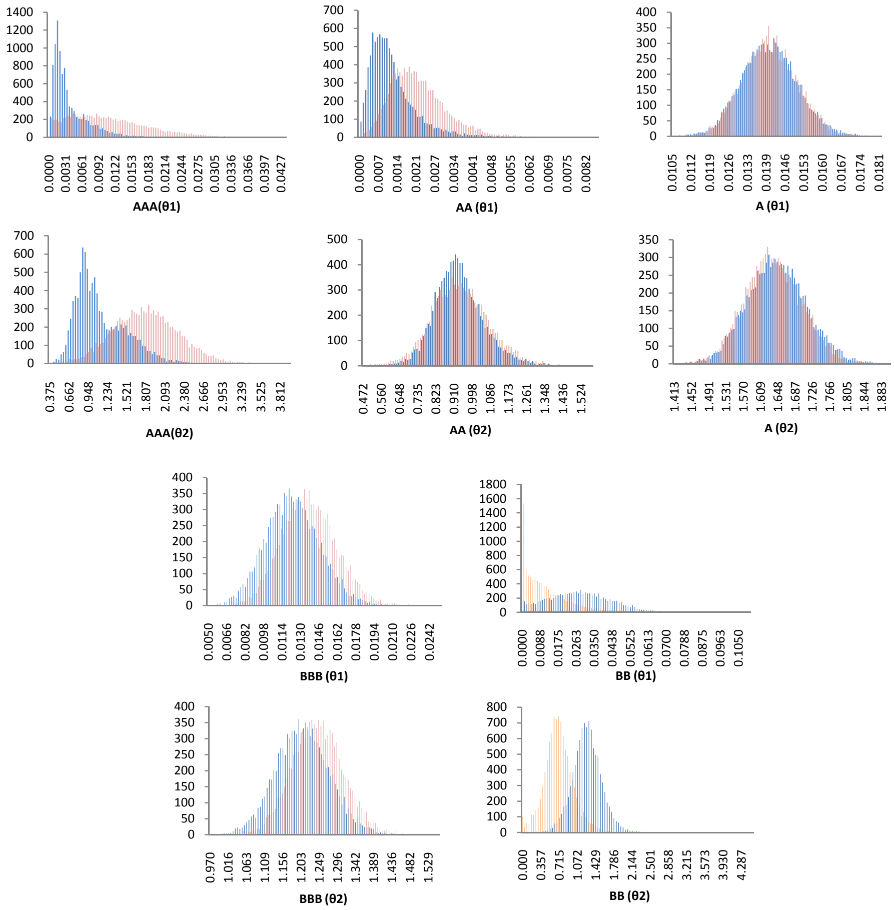

Figure 15: Distributions of bootstrap parameter estimates for 5 rating classes. Blue(deep) distribution is for statistical model and pink(pale) is for calibration

Data		Estimating results			
Data	Observations	Method	$\hat{\theta}_1$	$\hat{\theta}_2$	mean(|residuals |)
Original	523	Statistical model	0.00100	0.930	0.002210
		Calibration	0.00241	1.106	0.002076
No Outlier	520	Statistical model	0.00455	1.293	0.001750
		Calibration	0.00431	1.275	0.001749

Table 11: The representative parameter estimates for "AA" original data and for "AA" no outlier data. Original data are the same as "AA" in Table9. Outlier data are excluded 3 bonds' price data from the original data.

	Original data		No outlier data	
	Statistical model	Calibration	Statistical model	Calibration
$\tilde{\theta}_1^{(Boots)}$	0.00120	0.00257	0.00458	0.00439
$\hat{\theta}_1$	0.00100	0.00241	0.00455	0.00431
$\widetilde{Bias}(\hat{\theta}_1)^{(Boots)}$	0.00021	0.00016	0.00003	0.00008
$\widetilde{Var}(\hat{\theta}_1)^{(Boots)}$	0.00000	0.00000	0.00000	0.00000
$\widetilde{SD}_{\hat{\theta}_1}(t)^{(Boots)}$	0.00091	0.00115	0.00073	0.00100
K-S statistic	0.12476	0.06334	0.02554	0.02927
K-S test	reject	reject	reject	reject
$\tilde{\theta}_2^{(Boots)}$	0.93420	1.10797	1.29289	1.27692
$\hat{\theta}_2$	0.92966	1.10592	1.29266	1.27502
$\widetilde{Bias}(\hat{\theta}_2)^{(Boots)}$	0.00454	0.00205	0.00023	0.00190
$\widetilde{Var}(\hat{\theta}_2)^{(Boots)}$	0.01479	0.01287	0.00253	0.00568
$\widetilde{SD}_{\hat{\theta}_2}(t)^{(Boots)}$	0.12164	0.11345	0.05034	0.07534
K-S statistic	0.04810	0.01928	0.01272	0.00910
K-S test	reject	reject	not reject	not reject

Table 12: The parameter estimate of the bootstrap estimation for "AA" original data and for "AA" no outlier data. The results of K-S test at 0.05 level are given in the bottom 2 lines.

notice that "Observed" plots of "AA" class contains outlier.

We next show the influence of outlier data on our estimating method using data of "AA" class. In the "AA" graph of Figure 14, we may see 3 outlier plots among "Observed" plots. We excluded these 3 bonds' price data from the original data. Table 11 indicates the representative parameter estimates for "AA" original data and for "AA" no outlier data. Original data are the same as "AA" in Table9 and outlier data are excluded 3 bonds' price data from the original data.

Table 12 indicates the bootstrap parameter estimates for "AA" original data and for "AA" no outlier data. Figure 16 shows the distribution of bootstrap parameter estimates for "AA" original data and for "AA" no outlier data. The upper graph shows the histogram of the bootstrap estimate $\hat{\vartheta}_1{}^{(m)}$ and the lower graph shows that of $\hat{\vartheta}_2{}^{(m)}$.

We see from Figure 16 that the warps are reduced by the exclusion of the outlier data(Compare the left graphs and the middle graphs). We also see that the distributions of bootstrap parameter estimates using no outlier data look like normal distribution(middle graphs). We tested these results for normality of the distribution. One of the most useful tests for normality of the distribution is Kolmogorov-Smirnov test. Given M estimates $\hat{\vartheta}^{(m)}\quad i = 1,2, m = 1,\ldots,M$, the Kolmogorov-Smirnov statistic for cumulative normal

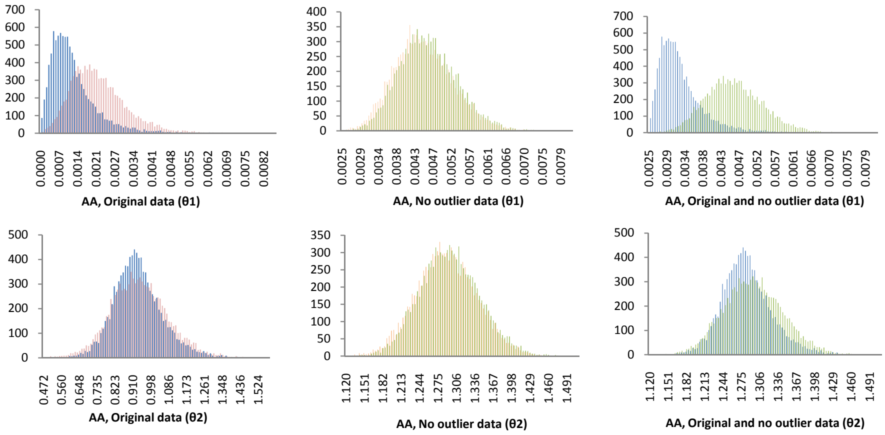

Figure 16: Distributions of bootstrap parameter estimates for "AA" original data and no outlier data. The left graphs are for original data. Blue(deep) distribution is for statistical model and pink(pale) is for calibration. The middle graphs are for no outlier data. 3 bonds' data are excluded. Green(deep) distribution is for statistical model and orange(pale) is for calibration. The right graphs are the distributions using statistical model for original data and no outlier data. Blue(deep) distribution is for original data and green(pale) is for no outlier data.

distribution function $F(x)$ is

$$D_M = sup_x |F_m(x) - F(x)|$$

where sup_x is the supremum of the set of distances and

$$F_M(x) = \frac{1}{M}\sum_{i=1}^{M} I_{\hat{\vartheta}^{(m)}<x},$$

where $I_{\hat{\vartheta}^{(m)}<x}$ is the indicator function, equal to 1 if $\hat{\vartheta}^{(m)} < x$ and equal to 0 otherwise. The null hypothesis that $\hat{\vartheta}^{(m)} \quad i = 1,2, m = 1,\ldots,M$ is normal distributed is rejected at level 0.05 if

$$\sqrt{M}D_M > 1.36.$$

M is 10000 in our case.

The Kolmogorov-Smirnov statistic D_{10000} and the results of K-S test are in Table12. We got better results by excluding only 3 bonds' price data from the original 523 bonds' data,

4.8 Estimation for Time Series Data

So far, we have estimated the survival probabilities of one day. In this subsection, we estimate the time series survival probabilities. The data we used is daily price data of the corporate bond and the government bond.

Since we are interested in the impact of the bankruptcy of Lehman Brothers(September 15, 2008), we use the data from January 4, 2008 to March 31, 2009. U.S. house sales prices peaked in mid-2006 and began their steep decline forthwith. The subprime mortgage crisis began to affect the financial sector in 2007. Although Japanese financial firms held a few securities backed with subprime mortgages, Japanese economy was affected by subprime mortgage crisis. We select five companies, which might be affected by the bankruptcy. Toyota Motor Corporation is a multinational automaker and one of the world's largest automobile manufacturers. Toyota has a rating of AAA throughout our analysis. Mitsui & Co., LTD. also known as Mitsui Bussan, is one of the largest sogo shosha in Japan. Mizuho Corporate Bank, Ltd. is the corporate and investment banking subsidiary of Mizuho Financial Group. Tokyu Land Corporation is the 4th biggest real estate agency in Japan. Japan Airlines Co., Ltd. is an airline in Japan. The airline filed for bankruptcy protection on January 19, 2010, after losses of nearly 100 billion yen in a single quarter.

Figure 17 shows the historical survival probabilities of the 5 companies. The survival probabilities of Tokyu Land Corporation began to decline slowly before the bankruptcy of Lehman Brothers. Although the risk of subprime mortgages was pointed before the bankruptcy, the decline is considered by the cause of its own. In 2008, housing values in Japan decreased by 10 percent with the previous year and the stock price of Tokyu Land declined all through 2008. The survival probability curves in the graph begin to drop sharply at November 2009. The bond buyers were likely to avoid the risk in the real estate market.

We next contrast the graphs of Toyota, Mitsui and Mizuho. We see that the all 3 companies' survival probabilities were declining after the bankruptcy, but Mizuho's graph began to drop first. Since Mizuho Corporate Bank is the corporate and investment banking subsidiary and was at the core of the financial crisis, the bankruptcy was seem to make an immediate effect. One month later of the Mizuho's survival probability curve drop, the survival probabilities of Toyota and Mitsui began to decline more slowly. It seems the financial crisis spreaded over gradually.

JAL's steep curve stands out in the graph. The 5 year survival probability varied 80% to 64% for only 3 months. It is because JAL was confronted not only with financial crisis, but also with the rise of oil price and slow cost cut. JAL's prospects seem to have been considered gloomy by the bond buyers.

5 Conclusion

In this paper, we study estimating method of implied survival probability. We focus on the statistical estimating model introduced by Takahashi(2011). The advantage of the statistical estimating model is its ability to discuss asymptotic properties. Our contributions are the following three points. First, we provide a complete proof of the consistency of the estimator in the statistical estimating model. Second, we verify regularity conditions for asymptotic properties of the estimator in the Weibull survival function case. Third, we provide an empirical analysis of implied survival probability estimating. When we estimate implied survival probability, instead of using Duffie & Singleton(1999) model directly, as in previous studies, we apply statistical estimating model which contains some error term. In the empirical analysis, we discuss asymptotic properties using bootstrap method.

Several interesting observations can be made. When we apply Weibull distribution to parametric survival function, the estimating result is much better than that of exponential distribution. The number of the data and the existence of the outlier data have a major effect on the bootstrap estimating results. When we analyze the impact of the bankruptcy of Lehman Brothers, we observe that the influences differ with the 5 individual companies.

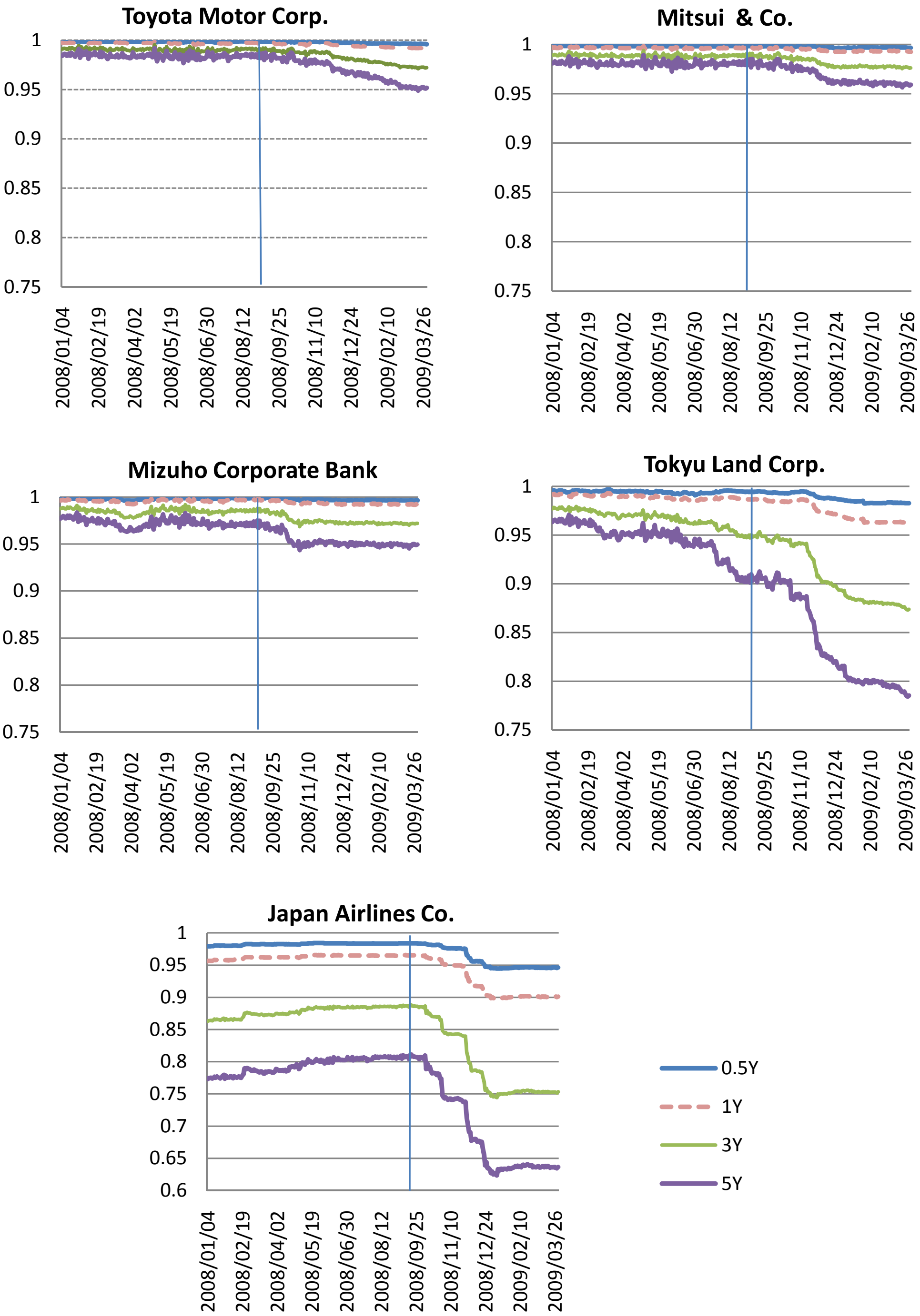

Figure 17: Time series default probabilities of Toyota, Mitsui & Co, Mizuho, Tokyu Land, JAL

Acknowledgements The author is grateful to Hajime Takahashi for his much valuable advice. The author is also grateful to Center for Financial Engineering Education (CFEE) of Hitotsubashi University for their support. Data for this study were provided through CFEE.

AppendixA Bootstrap Method

When analyzing the real data, Takahashi(2011) recommends to use the bootstrap method rather than asymptotic theory. We introduce the bootstrap method of Takahashi(2011), which is used on the empirical research in section 4.

[Bootstrap Method of Takahashi(2011)] We consider bootstrap estimate of $\hat{\vartheta}$. We first consider the case of statistical model under RM type recovery. We let $\{x_{t_i} = x(t_i) \;\; i = 1, \ldots, n\}$ be a sequence of the pseudo data, where $x_t = x(t) = \frac{{s_t^1}^\delta}{\rho_t}$. We estimate the base line estimate $\hat{\vartheta}$ by

$$\hat{\vartheta} = \arg_{\theta\in\Theta}\left[\min \sum_{i=1}^{n} \{\Xi(t_i,\theta) - x_{t_i}\}^2\right]$$

where

$$\Xi(t,\theta) = \frac{G(t,\theta)^{1-\delta}}{\rho_t}.$$

We estimate $\{\sigma\varepsilon_{t_i}, i = 1, \ldots, n\}$ by

$$e^*_{t_i} = \Xi(t,\hat{\vartheta}) - x_{t_i} \qquad i = 1, \ldots, n.$$

We define the set E of standerdized residuals, from which the bootstrap sampling will be made (cf.Wu(1986)):

$$E = \{e_{t_1}, \ldots, e_{t_n}\}$$

where

$$e_{t_i} = \sqrt{\frac{n}{n-1}}\left(\overline{e^*_{t_i}} - e^*\right)$$

and

$$\overline{e^*} = \frac{1}{n}\sum_{i=1}^{n} e^*_{t_i}.$$

$0^{th} step$ Set

$$m = 1$$

$1^{st} step$ We choose n elements at random from E with replacement, denote it $E^{(m)}$

$$E^{(m)} = \{e^{(m)}_{t_1}, \ldots, e^{(m)}_{t_n}\}$$

$2^{nd} step$ Constract a bootstrap pseudo sample by

$$x^{(m)}_{t_i} = \Xi(t_i,\hat{\vartheta}) - e^{(m)}_{t_i} \qquad i = 1, \ldots, n.$$

$3^{rd}step$ Estimate θ from the pseudo data $\{x_{t_i}^{(m)}, i = 1, \ldots, n\}$ and denote it as $\hat{\vartheta}^{(M)}$

$$\hat{\vartheta}^{(m)} = \arg_{\theta\in\Theta}\left[\min\sum_{i=1}^{n}\left\{\Xi(t_i,\theta) - x_{t_i}^{(m)}\right\}^2\right].$$

$4^{th}step$ Repeat $1^{st}step$ to $3^{rd}step$ M times, and denote the estimate by

$$\hat{\vartheta}^{(1)}, \ldots, \hat{\vartheta}^{(M)}$$

We define bootstrap estimate by

$$\tilde{\vartheta}^{(Boots)} = \frac{1}{M}\sum_{m=1}^{M}\hat{\vartheta}^{(m)}.$$

We also define the estimate of bias and variance of $\tilde{\vartheta}$ by

$$Bias^{(Boots)}(\hat{\vartheta}) = \tilde{\vartheta}^{(Boots)} - \hat{\vartheta}$$

and

$$Var^{(Boots)}(\hat{\vartheta}) = \frac{1}{M-1}\sum_{m=1}^{M}\left(\hat{\vartheta}^{(m)} - \tilde{\vartheta}^{(Boot)}\right)\left(\hat{\vartheta}^{(m)} - \tilde{\vartheta}^{(Boot)}\right)'$$

respectively.

We next consider bootstrap estimate when using statistical model under RT type recovery. The baseline estimate $\hat{\vartheta}$ is given by

$$\hat{\vartheta}_{RT} = \arg_{\theta\in\Theta}\left[\min\sum_{i=1}^{n}\left\{\Xi^{\sharp}(t_i,\theta) - x_{t_i}^{\sharp}\right\}^2\right]$$

where

$$x^{\sharp}(t) = \frac{(1-\delta)s_t}{\rho_t}.$$

and

$$\Xi^{\sharp}(t,\theta) = \frac{(1-\delta)G(t,\hat{\vartheta}_{RT})}{\rho_t}.$$

We estimate $\{\sigma\varepsilon_{t_i}, i = 1, \ldots, n\}$ by

$$e_{t_i}^{\sharp *} = \Xi^{\sharp}(t,\hat{\vartheta}_{RT}) - x_{t_i}^{\sharp} \qquad i = 1, \ldots, n.$$

Then we constract standerdized residuals $E^{\sharp} = \{e_{t_1}^{\sharp}, \ldots, e_{t_n}^{\sharp}\}$ and bootstrap estimate $\tilde{\vartheta}^{(Boots)}$ in the same way as that of RM type recovery.

References

[1] Duffie, D. and K. Singleton (1999) ”Modeling Term Structures of Defaultable Bonds”, Review of Financial Studies, 12, 678-720

[2] Ito, S. (1963) ”Introduction to the Lebesgue Integral”, Shokabo Publishing (in Japanese)

[3] Jarrow, R. and S. Turnbull (1995) ”Pricing Options on Financial Securities Subject to Default Risk”, Journal of Finance, 50, 53-86

[4] Newey, W. K. and D. McFadden (1994) ”Large Sample Estimation and Hypothesis Testing”, Handbook of Econometrics, 4, 2111- 2245, Elsevier Science B.V.

[5] Takahashi, H. (2011) ”On a Statistical Analysis of Implied Data ”, Asia Pacific Financial Markets, 18, 231-266

[6] Wald, A. (1949) ”Note on the Consistency of the Maximum Likelihood Estimate”, Annals of Mathematical Statistics, 20, 595-601

[7] Wu, C. F. J. (1986) ”Jackknife, Bootstrap and Other Resampling Methods in Regression Analysis”, Annals of Statistics 14, 1261-1295