IMES DISCUSSION PAPER SERIES

Competing Risk Analysis of Japan's Small Financial Institutions

Xinghua Yu

Discussion Paper No. 2006-E-7

IMES

INSTITUTE FOR MONETARY AND ECONOMIC STUDIES

BANK OF JAPAN

C.P.O BOX 203 TOKYO

100-8630 JAPAN

You can download this and other papers at the IMES Web site:

http://www.imes.boj.or.jp

Do not reprint or reproduce without permission.

NOTE: IMES Discussion Paper Series is circulated in order to stimulate discussion and comments. Views expressed in Discussion Paper Series are those of authors and do not necessarily reflect those of the Bank of Japan or the Institute for Monetary and Economic Studies.

IMES Discussion Paper Series 2006-E-7
March 2006

Competing Risk Analysis of Japan's Small Financial Institutions

Xinghua Yu*

Abstract

This paper develops a dependent competing risks model to investigate the determinants of time to bankruptcy and time to merger jointly, and more importantly, to investigate their interdependence. This paper identifies strong interdependence between the bankruptcy and merger hazards, both through the correlation of the unobserved heterogeneities and through the preventive behavior of the individual firms. This paper shows that the common practice of assuming the independence of the competing risks would yield biased estimates and lower the predictive accuracy. In addition, this paper addresses important econometric-theoretic questions that arise with the empirical analysis, namely, identification and testing a hypothesis in a nonstandard situation.

Key words: Competing Risks Duration Models, Identification, Maximum Likelihood Estimator, Bankruptcy and Merger, Credit Cooperatives in Japan
JEL classification: C13, C24, C41, G21, G33, G34

* Stanford University, Department of Economics (E-mail: xhyu@stanford.edu)

This research was initiated when I visited the Institute of Monetary and Economic Studies (IMES) at the Bank of Japan in the summer of 2004. I have greatly benefited from deep discussions with my advisors Takeshi Amemiya, Michael Boskin and Edward Vytlacil. I am extremely grateful to Professor Takeshi Amemiya for his continuous guidance and encouragement. I am also grateful to Professor Peter Hansen, Professor Aprajit Mahajan, Professor Tetsuji Okazaki, Mr. Akira Ieda, Mr. Yutaka Kosai, Ms. Keiko Murata, Dr. Toshiaki Watanabe, Mr. Satoshi Yamaguchi, Mark Jacobson, Niny Khor, Ben Malin, Azeem Shaikh and Joanne Yoong for many valuable comments. I would like to thank IMES and the Economic and Social Research Institute (ESRI) at the Cabinet Office of Japanese government for providing the electronic version of the data for the empirical research in this paper. The views expressed in this paper are those of the author and do not necessarily reflect the official views of the Bank of Japan or IMES or ESRI.

Contents

Competing Risk Analysis of Japan's Small Financial Institutions

1 Introduction

Japan's banking system has experienced a dramatic restructuring and consolidation process since the burst of the asset bubble at the end of the 1980s. To the surprise of many, even some of the top financial institutions did not survive the financial turbulence. This ongoing restructuring process has had an especially profound impact on the small financial institutions. Among the approximately 400 credit cooperatives[1] existing at the beginning of Fiscal Year 1992, over 30% filed bankruptcy and another 25% merged with other financial institutions during the ten-year period from Fiscal Year 1992 to Fiscal Year 2001. Today, the total number of credit cooperatives has declined to less than 180. This consolidating process has both heightened the necessity of and provided the data for building an efficient early-warning system for the financial industry. To this end, this paper utilizes the panel data of credit cooperatives in Japan to investigate the bankruptcy and merger hazards with a dependent competing risks model. The method developed in this paper can be readily applied to data collected from other industries or countries.

Various models have been applied to study the failures and mergers of financial institutions. The most frequently used models include multivariate discriminant analysis, qualitative response models (such as probit and logit models) and duration models, in a roughly chronological order of their development. Examples of other statistical approaches are neural network analysis (see Atiya (2001) for a survey) and decision tree models (see Cheng et al. (2003)). Hooks (1995) points

[1] There are three major types of private deposit-taking financial institutions in Japan: commercial banks, credit associations (*shinyo kinko*, also translated as "shinkin banks") and credit cooperatives (*shinyo kumiai*, also translated as "community banks"). There is no significant difference among the three types in terms of areas of business, but their organizational structures and geographical restrictions are quite different. The latter two types are relatively small deposit-taking financial institutions based on membership systems. Both of these two types focus on local communities and are important financial sources for regional economy. One of the main differences between credit cooperatives and credit associations is that the regulatory authority of cooperatives belongs to local governments while associations are regulated by the Financial Supervisory Agency. As of April 1, 2005, credit cooperatives account for more than 27% of the total numbers of the insured deposit-taking financial institutions in Japan. In comparison, commercial banks and credit associations account for 24% and 46%, respectively. For a detailed comparison, please refer to Murata and Hori (2004).

out that the forecasting accuracy of the early-warning models with the indicator of bankruptcy as the dependent variable tends to fluctuate over time. Part of such instability lies in the fact that all of the models surveyed in the paper take static approaches. Duration models, however, have the capability of taking the entire history of a given firm into consideration dynamically, as well as being able to incorporate censored cases and time-varying covariates easily.

Most literature on the exit hazard of the banking industry focuses either on the failure hazard only (such as Lane et al. (1986)) or the merger and acquisition (M&A) hazard only (such as Hannan and Rhoades (1987)). In the few papers that jointly model the probability of being acquired and failing using the competing-risk framework, the independence of the two hazards is usually assumed (for example,Wheelock and Wilson (2000)); that is, when estimating the failure hazard, mergers are treated as censored at their dates of merger; similarly, when estimating the merger hazard, failed banks are treated as censored at their dates of bankruptcy. However, both intuitions and empirical observations suggest that the bankruptcy and merger processes are interrelated. For example, if a bank has poor performance and foresees itself running high risks of bankruptcy, it might increase its intensity to look for a counterpart to merge with in order to avoid closure. Such kind of firm behavior implies that higher risks of bankruptcy might lead to higher probability of merger, which induces a positive correlation between the bankruptcy and merger processes. Take Japan's credit cooperatives for example, during the past decade, we frequently observe the situation where a pair of cooperatives merged to avoid bankruptcy, sometimes under the suggestion of the regulatory authority. In addition, there might be some unobservable firm-specific characteristics that affect both the bankruptcy and merger processes. For example, it is hard to measure the effect of firm culture and management skills on a firm's life. But certain combination of such factors might make a firm more likely to expand and less likely to go bankrupt. Such kind of unobserved heterogeneities might induce either positive or negative correlations between the bankruptcy and merger processes. Despite these intuitions and observations, there have been no empirical studies, as far as we know, to explicitly account for the interdependence between the bankruptcy and merger processes. The above argument, however, implies that the independent competing risks model might not fully describe the failure and merger processes and may generate inconsistent estimates.

Besides the banking literature, many other empirical competing risks studies assume stochastic independence among risks (see, for example, Duffie et al. (2005)). In addition to computational convenience, one of the main reasons for researchers to restrict their attention to independent specification is the common misconception that dependent competing risks specifications are not identifiable. The frequently cited paper of Tsiatis (1975) proves that, for any joint survival function with arbitrary dependence between the competing risks, one can find a different survival function with independent survival times that has exactly the same cause-specific hazards. In other words, the data would not be able to distinguish a model with dependent competing risks from one with independent competing risks. If that is the case, then there is no point to complicate the model with dependence assumption because the data could not test for it anyway. However, Tsiatis (1975)'s argument is valid only if the sample is homogeneous, that is, if there is no heterogeneity which depends on either observable or unobservable factors. The problem of non-identifiability can be resolved by introducing heterogeneity through the variation of the covariates. In particular, Heckman and Honoré (1989) prove that the class of proportional and accelerated hazard competing risks models are non-parametrically identified under certain regularity conditions if there is sufficient variation of the latent failure times with the covariates. Abbring and van den Berg (2003) use a set of less restrictive assumptions to identify the smaller class of mixed proportional hazard (MPH) competing risks models non-parametrically. (See van den Berg (2001) for a survey of studies of multiple dependent durations.)

However, both Heckman and Honoré (1989) and Abbring and van den Berg (2003) assume that the covariates are constant over time, which is not likely to be true for many empirical studies. Their identification result critically relies on the ability to vary covariates X and time t separately. But if the covariates X also depend on time t, such argument would break down without necessary modification. To generalize the identification proof to the more realistic case of time-varying covariates, this paper considers the typical empirical setting where covariates are constant over discrete intervals. Identification is achieved by exploiting the variation of t in the neighborhood of zero and its separability from X in the first interval. Based on this identification result, we specify a dependent competing risks duration model to identify empirically the determinants of bankruptcy and merger hazards, and more importantly, the interdependence of these two processes, using the data of small financial institutions in Japan. In particular, our model allows the interdependence of these competing risks both through the correlation of the unobserved heterogeneity associated with each risk and through a "structural" dependence induced by the preventive behavior of individual institutions. As far as we know, this is the first paper to treat rigorously the interdependence among the hazards of the different exit modes for a given firm.

Although interdependence among competing risks is prevalent in other areas too, few empirical studies have addressed the issue explicitly. Two of such rare examples are Lillard (1993), who studies the simultaneity of the marriage duration and fertility timing; and Fallic and Ryu (2003), who investigate the interaction between the recall and new job hazards. Our model specification is similar to, but differs from theirs in three important aspects: first, Lillard (1993) assumes the heterogeneity components are jointly normal, while we estimate the distribution of heterogeneities non-parametrically. Second, unlike Lillard (1993), we allow the structural correlation to affect only a subgroup, more specifically, bankruptcy hazard only has a structural effect on passive mergers, but not on active mergers.[2] Third, Fallic and Ryu (2003) assume the structural effect of one competing risk on the other competing risk is only partial in the sense that only the observable part of one hazard enters the other hazard's reduced form equation. Given that we define unobserved heterogeneity to be only unobservable to the econometricians, but observable to the firms, such assumption is arbitrary and might lead to inconsistency. We, instead, allow the bankruptcy hazard rate, both the observable part and the unobservable part, to enter the passive merger equation, which keeps the consistency of the assumption on the unobservable heterogeneity.

In terms of covariate choice, this paper falls into the category of financial ratio analysis based primarily on financial statement information. Firm life studies incorporating market price information of the firm's debt and equity, rather than using only financial statement information, are considered superior in the sense that market information is forward-looking. To overcome the problem of lacking such information for non-publicly-traded firms, Moody's KMV developed a "Private Firm Model", using the market information of the competing publicly-traded firms (Nyberg et al. (2001)). However, market information is not available for our study. On one hand, credit cooperatives are of cooperative type and do not trade shares publicly; on the other hand, the competing publicly-traded counterparts – city banks, regional banks and trust banks combined –

[2]The distinction of active vs. passive mergers will be explained in Section 3.

are both much less in number and very different in organizational features. Our model, however, is not restricted merely to financial ratios; market information, when available, can be directly added in. In this paper, the selection of financial ratios is primarily based on the "PEARLS" evaluation system developed by the World Council of Credit Unions Inc. (Richardson (2002)), which covers a full spectrum of institution-specific characteristics. In addition, this paper studies the effect of several economy-wide and industry-wide factors on individual firms' exit hazards.

The remainder of the paper is organized as follows. Section 2 proves the semi-parametric identification of the dependent competing risks model with time-varying covariates and proposes a maximum likelihood estimator for the empirical analysis. Section 3 describes our data set and the selection of covariates, as well as presenting their summary statistics. Section 4 compares the estimating results from different model specifications and demonstrates the importance of allowing the interdependence of the competing risks. Section 5 concludes.

2 Econometric Model Specification

As mentioned in the previous section, duration models have the advantage of being able to incorporate time-varying covariates and censored observations easily. In particular, a dependent competing risks model allows us to estimate the bankruptcy and merger hazards jointly, accounting for the interdependence of these competing risks. In this section, we first prove the semi-parametric identification of the dependent competing risks model with the presence of time-varying covariates. Then we will specify the particular econometric model, as well as its different variations, which are applied in Section 4 to identify empirically the determinants of the exit hazards of Japan's small financial institutions.

2.1 Semi-parametric Identifiability of the Dependent Competing Risks Model with Time-varying Covariates

Heckman and Honoré (1989) and Abbring and van den Berg (2003) prove that by introducing enough covariate variation, the dependent competing risks duration model can be identified semi-parametrically. In both papers, one implicit assumption is that the marginal hazards are separable in terms of duration dependence and covariate variation. Without proper modification, their proof would fail if we introduce time-varying covariates. However, most economic studies, like ours, involve time-varying covariates. Therefore, it is worthwhile to generalize the identifiability argument to this more realistic situation.

First, consider the case with two competing risks. Let T_1 and T_2 be the two competing latent durations that we are interested in. Denote X as the related observable covariates and V_1 and V_2 as the unobserved non-negative random variables that are distributed independently of X. This paper considers a commonly used class of competing risks models defined by the following sub hazards:

$$h_1(t|x, v_1, v_2) = \lambda_1(t)\phi_1(x(t))v_1 \tag{1}$$

$$h_2(t|x, v_1, v_2) = \lambda_2(t)\phi_2(x(t))v_2, \tag{2}$$

where $\lambda_1 : \mathcal{R}_+ \to (0, \infty)$ and $\lambda_2 : \mathcal{R}_+ \to (0, \infty)$ are the baseline hazards; $\phi_1 : \mathcal{X} \to (0, \infty)$ and $\phi_2 : \mathcal{X} \to (0, \infty)$ are continuous regressor functions, with $\mathcal{X}$ as the support of $\{x(t); t \in \mathcal{R}_+\}$.

Strictly speaking, when the covariates X change with t, the model is no longer proportional in the original sense of the term. That is why the proof is not a trivial extension of Heckman and Honoré (1989) and Abbring and van den Berg (2003).

We assume that T_1 and T_2 are independent conditioning on the observed and unobserved heterogeneities (i.e. $T_1 \perp T_2 | (X, V_1, V_2)$), but notice that conditional on the observed heterogeneity X only, T_1 and T_2 can be dependent. Actually, in this model, interdependence between the two risks, conditioning on X, is generated solely through the correlation of the unobserved heterogeneities associated with the two risks. We will introduce "structural" dependence in the next section and the identification proof here will continue to be valid for the reduced form of that augmented model. Based on this assumption, we can derive the conditional survival function from Equations

(1) and (2) as

$$
\begin{aligned}
& Pr(T_1 > t_1, T_2 > t_2 | x, v_1, v_2) \\
= \ & \exp\left\{ -\int_0^{t_1} \lambda_1(s)\phi_1(x(s))ds \cdot v_1 - \int_0^{t_2} \lambda_2(s)\phi_2(x(s))ds \cdot v_2 \right\}.
\end{aligned} \tag{3}
$$

Following the notation in Abbring and van den Berg (2003), we denote the unknown distribution function of (V_1, V_2) by G and its Laplace transformation by $\mathcal{L}_G$ for convenience to simplify the notation:

$$
\mathcal{L}_G(s_1, s_2) = \int_0^\infty \int_0^\infty \exp\left(-s_1 v_1 - s_2 v_2\right) \, dG(v_1, v_2). \tag{4}
$$

The joint survival function of (T_1, T_2) conditioning on X is then defined as

$$
S(t_1, t_2|x) := Pr(T_1 > t_1, T_2 > t_2|x) = \mathcal{L}_G \left\{ \int_0^{t_1} \lambda_1(s)\phi_1(x(s))ds, \int_0^{t_2} \lambda_2(s)\phi_2(x(s))ds \right\}. \tag{5}
$$

The identification in Abbring and van den Berg (2003)'s proof is achieved by varying t and X separately. However, this argument breaks down when covariates X also depend on t. As argued in Sueyoshi (1992) and many other existing empirical and theoretical studies, most economic data are collected in discrete times and there usually is no prior knowledge about the path of the data between such discrete times, it is reasonable to assume that the time-varying covariates are constant over discrete intervals. In particular, we make the following assumption: assume covariates X are constant over discrete intervals with splitting times $t_0(=0), t_1, t_2, \cdots$. Define the mapping $k(t) : \mathcal{R}_+ \to 1, 2, 3, \cdots$, which maps t into an interval k. Let x_k be the constant value that X takes in the k^{th} interval. Then, we have $X(t) = x_{k(t)}$. Without loss of generality, we normalize the interval to be of length 1, that is, $X(t+\Delta) = X(t)$ for $t = 0, 1, 2, \cdots$ and $0 \le \Delta < 1$.

Based on this assumption, we can rewrite the joint survival function as

$$
S(t_1, t_2|x) = \mathcal{L}_G\{s_1(t_1, x), s_2(t_2, x)\}, \tag{6}
$$

where

$$
\begin{aligned}
s_1(t_1, x) &= \sum_{k=1}^{[t_1]} (\Lambda_1(k) - \Lambda_1(k-1)) \cdot \phi_1(x_k) + (\Lambda_1(t_1) - \Lambda_1([t_1])) \cdot \phi_1(x_{[t_1]+1}); \qquad (7) \\
s_2(t_2, x) &= \sum_{k=1}^{[t_2]} (\Lambda_2(k) - \Lambda_2(k-1)) \cdot \phi_2(x_k) + (\Lambda_2(t_2) - \Lambda_2([t_2])) \cdot \phi_2(x_{[t_2]+1}).^3 \qquad (8)
\end{aligned}
$$

Proposition 1. *Assume that* (T_1, T_2) *has the joint survival function as given in* (6)-(8)*. Then, the competing risks model (characterized by the functions* $\Lambda_1, \Lambda_2, \phi_1, \phi_2$ *and* $\mathcal{L}$*) is identified from the identified minimum* (T, I)[4] *under the following assumptions:*

1. *Variation of observed covariates: the value of* $\{\phi_1(x_1), \phi_2(x_1)\}$ *contains a non-empty open set* $\Phi \subset \Re_+^2$*.*

2. *Tail of the heterogeneity distribution:* $E(V_i) < \infty$*,* $i = 1, 2$*.*

3. *Normalization:* $\int_0^1 \lambda_i(s)ds = 1$ *and* $\phi_i(y^*) = 1$ *for some fixed point* $y^* \in \mathcal{X}$*,* $i = 1, 2$*.*

Please see Appendix A for proof. It is straightforward to apply the same argument in proving Proposition 1 to the identification of the dependent competing risks model with more than two risks.

2.2 Maximum Likelihood Estimator

This subsection proposes the maximum likelihood estimator for the dependent competing risks model used in the empirical analysis in Section 4. Before elaborating the model specification, we list the notation and assumptions to facilitate the measurement and analysis.

Let the non-negative random variable T be the duration until exit or the end of the sample period. In the current case, the exit event can be of two general types: bankruptcy (B) or Merger (M). Credit cooperatives involved in the merging activities are further classified as being bidder (active) or target (passive), denoted as MA and MP. The main reason to distinguish these two types of merger is that they have very different incentives to merge, which helps to disentangle the two different sources of interdependence between the bankruptcy and merger hazards. More specifically, passive merger usually occurs when a firm runs high risks of bankruptcy and intensively

[3] $[t]$ denotes the largest integer that is less than t. $\Lambda_i(t) := \int_0^t \lambda_i(s)\, ds$.

[4] $(T, I) := (\min\{T_1, T_2\}, \arg\min\{T_1, T_2\})$

looks for opportunity to become the target of a merger in order to avoid closure. Such dependence is labeled as "structural" dependence. In addition, interdependence of the bankruptcy and merger processes can be induced by the correlation of the associated unobserved heterogeneities. Such dependence exists for both passive merger and active merger. The distinction of active vs. passive merger will be explained in the data section. Let T_B, T_{MA}, T_{MP} and $\bar{T}$ be the duration until bankruptcy, active merger, passive merger and the end of the sample period, respectively. By definition, $T = \min(T_B, T_{MA}, T_{MP}, \bar{T})$. The indicator variables for exits by bankruptcy, active merger and passive merger are denoted as d_B, d_{MA} and d_{MP}, respectively. We study the exit event history by controlling for both observed heterogeneity, denoted as $X = \{(X_B(t), X_M(t))\}$ (exclusion restrictions are not necessary), and unobservable heterogeneity, denoted as (V_B, V_M), which has the joint density of $g(v_B, v_M)$. Assume that the time-varying covariate process X is exogenous[5] for duration T and independent from the unobserved heterogeneity. The cause-specific hazard function is defined by:

$$h^m(t|x, v_B, v_M) = \lim_{\Delta t \to 0^+} \frac{P(t \leq T_m < t + \Delta t | T \geq t, x, v_B, v_M)}{\Delta t} \tag{9}$$

for $m = B, MA, MP$.

In particular, we assume the following multiplicative form for the baseline hazards, observed and unobserved heterogeneity for our empirical analysis.

$$h^B(t|x, v_B, v_M) = \lambda_0^B(t) \cdot \exp\{x_B(t)\beta_B + v_B\} \tag{10}$$

$$h^{MA}(t|x, v_B, v_M) = \lambda_0^M(t) \cdot \exp\{x_M(t)\beta_{MA} + v_M\} \tag{11}$$

$$h^{MP}(t|x, v_B, v_M) = \lambda_0^M(t) \cdot \exp\{x_M(t)\beta_{MP} + v_M + \delta \cdot \log(h^B(t|x, v_B, v_M))\} \tag{12}$$

It is easy to show that the reduced form of this model is a competing risks model with three risks, whose identification is proved in Proposition 1. The identification of the structural parameters is achieved through the fact that baseline hazard function for the passive merger is the same as that for the active merger. This is true because in most of the merger events, both the bidding companies and the target companies are in our data set, and we fix the starting time to be the

[5]The covariate process $X(t)$ is defined to be exogenous for duration T if $P(X(t, t+dt)|T \geq t + dt, X(t)) = P(X(t, t+dt)|X(t))$, where $X(t, t+dt)$ is the covariate path over the interval of $(t, t+dt]$. In other words, we assume that the information that a firm survives up to time t does not help to predict the future path of the covariate process X given its history to t.

same (April 1 1992) for all observations, therefore we would observe high merging probabilities for both bidders and targets around similar time period. As a result, there is no need to assume different distributions of baseline hazards for acquiring and target credit cooperatives. We further assume that they also share the same distribution for unobserved heterogeneities, which is for purely practical reasons due to our relatively small sample size. In theory, however, we would be able to identify the three unobserved heterogeneities associated with the three competing risks separately. Since we have controlled for a rich set of covariates that have been used to distinguish bidders and targets in the existing literature, we think this restriction would have minor cost, if any.

Note that in the above model defined by Equations (10) to (12), the interdependence between the bankruptcy and merger hazards are accounted for in two ways, one is through the correlation of the unobserved heterogeneities, and the other is through the structural dependence. First, firms that file bankruptcy and firms that merge with others might share similar or distinctive unobserved features, which are identified by the positive or negative correlation of V_B and V_M in our model. More specifically, on one hand, "poorly" managed firms may be quicker than "well" managed firms to exit either through bankruptcy or through merger; therefore, the relationship between the bankruptcy and merger hazards induced directly by heterogeneity could be positive. On the other hand, firms that are likely to file bankruptcy are less attractive as targets for merging, and firms that are more likely to expand through merging might be less prone to bankruptcy risks; therefore, the relationships between the bankruptcy and merger hazards induced by the distinctive unobserved heterogeneities could be negative. Second, firms that are close to being insolvent might seek to merge with others to avoid failure. The Japanese regulatory authority also promotes merger as an alternative to closure if the credit cooperative is considered to have the potential to revive. Such preventive motivation induced correlation is captured by the structural dependence parameter δ in Equation (12). If the above hypothesis is true, we expect δ to be significant and positive.

The general model defined by Equations (10) to (12) nests three restricted models, of which the most restricted two are most frequently used in applied studies. To illustrate the advantage of the general model and the possible bias of the restricted models, we estimated all four models using the same data set. For convenience, we enumerate the models as model I to model IV in the decreasing order of restrictiveness, with model IV being the least restricted one defined by Equations (10) to (12). The first restriction we can apply to the general model is $\delta = 0$, which removes the "structural" dependence of the bankruptcy and merger hazards and only allows the interdependence between the exit risks through unobserved heterogeneities. This defines model III. Next, we can impose restriction on the unobserved heterogeneities by assuming $V_B \perp V_M$, which defines model II. Last, we can completely ignore unobserved heterogeneities by restricting that $g(v_B, v_M)$ degenerates at (0,0). This last model is denoted as model I. Models I and II are equivalent to estimating the bankruptcy hazard separately from the active and passive merging hazards, which greatly reduces the computational load and, therefore, is another reason that researchers prefer to use such model specification. As far as we know, only model I has ever been applied to firm exit studies (see, for example, Wheelock and Wilson (2000)). However, we are aware that there are applications using models II and III to study the prepayment and default of mortgages or job search behavior (see, for example, Deng et al. (2000)).

Based on the assumption that T_B, T_{MA} and T_{MP} are independent conditioning on $\{x, v_B, v_M\}$, we derive the conditional survival function as

$$\begin{aligned}
& S(t|x, v_B, v_M) \\
= \; & P(T \geq t|x, v_B, v_M) \\
= \; & P(T_B \geq t|x, v_B, v_M) \cdot P(T_{MA} \geq t|x, v_B, v_M) \cdot P(T_{MP} \geq t|x, v_B, v_M) \\
= \; & \exp\left\{-\int_0^t \left(h^B(s|x, v_B, v_M) + h^{MA}(s|x, v_B, v_M) + h^{MP}(s|x, v_B, v_M)\right) ds\right\}. \qquad (13)
\end{aligned}$$

Therefore, the likelihood function for a typical credit cooperative would be

$$L_i(t_i|x_i) = \int_{V_B}\int_{V_M} L(t_i|x_i, v_B, v_M) \cdot g(v_B, v_M)\, dv_B\, dv_M, \qquad (14)$$

where

$$\begin{aligned}
& L(t_i|x_i, v_B, v_M) \\
= \; & \left[h^B(t_i|x_i, v_B, v_M)\right]^{d_B} \cdot \left[h^{MA}(t_i|x_i, v_B, v_M)\right]^{d_{MA}} \cdot \left[h^{MP}(t_i|x_i, v_B, v_M)\right]^{d_{MP}} \cdot \\
& \exp\left\{-\int_0^{t_i} \left(h^B(s|x, v_B, v_M) + h^{MA}(s|x, v_B, v_M) + h^{MP}(s|x, v_B, v_M)\right) ds\right\}. \qquad (15)
\end{aligned}$$

To ensure the empirical identification of the model using the data set of Japanese credit cooperatives while allowing as much flexibility as possible, we further make the following assumptions when estimating:

Assumption 1: Baseline hazards are piecewise constant with splitting times $t_0(=0) < t_1 < t_2 < \cdots$ being the beginning of each fiscal year.

$$\lambda_0^m(t) = \sum_k \lambda_k^m 1_{t \in [t_k, t_{k+1})}, \qquad m = B,\ MA,\ MP. \tag{16}$$

Assumption 2: Unobserved heterogeneities are individual specific and follow a two-dimensional discrete distribution with $r \times r$ points of support.[6] One of the support points in each cause specific hazard is further normalized to be zero.[7]

$$\begin{array}{lll}
V_B & V_M & P = pr(V_B, V_M) \\
v_1^B(=0) & v_1^M(=0) & P_{11} \\
\vdots & \vdots & \vdots \\
v_1^B(=0) & v_r^M & P_{1r} \\
\vdots & \vdots & \vdots \\
v_r^B & v_1^M(=0) & P_{r1} \\
\vdots & \vdots & \vdots \\
v_r^B & v_r^M & P_{rr},
\end{array}$$

where $\sum_{i=1}^r \sum_{j=1}^r P_{ij} = 1$.

To ensure that the probabilities lie within $[0, 1]$ and sum up to 1, we make the following logistic transformation :

$$P_{ij} = \frac{\exp(q_{ij})}{\sum_{i=1}^r \sum_{j=1}^r \exp(q_{ij})} \qquad i, j = 1, 2, \cdots, r, \tag{17}$$

where $-\infty < q_{ij} < \infty$, for $i, j = 1, 2, \cdots, r$. Since this transformation is only identifiable up to a constant shift (that is, $\tilde{q}_{ij} = q_{ij} + k$ implies $\tilde{P}_{ij} = P_{ij}$), we further normalize $q_{11} = 0$. When reporting the result, we transform back to the P representation, with the standard errors estimated using the delta method.

[6] Heckman and Singer (1984) show that in the case of single risk, the non parametric maximum likelihood estimator of the distribution of the unobserved heterogeneity is a discrete distribution with a finite number of mass points.

[7] This normalization is necessary because there is already a constant in each of the cause specific hazard – the baseline hazard. An alternative way of normalization is to require $E(V^B) = E(V^M) = 0$, that is, $\sum_{i=1}^r V_i^B \left(\sum_{j=1}^r P_{ij}\right) = \sum_{j=1}^r V_j^M \left(\sum_{i=1}^r P_{ij}\right) = 0$ in the case of two-dimensional discrete distribution with $r \times r$ points of support. The estimation results do not depend on the normalization condition.

3 Data and Summary Statistics

Our data set combines the data of individual credit cooperatives with the macro level economic data. Annual balance sheet and income statement for each credit cooperative are gathered from various issues of “Financial Statements of Credit Cooperatives in Japan” (*Zenkoku Shinyo Kumiai Zaimu Shohyo*), edited and published annually by Financial Book Consultants, Ltd.[8] The history of individual credit cooperatives, including the date of establishment, the date of exit if not censored and whether the exit is the consequence of a bankruptcy, active merger or passive merger, was retrieved from various issues of “List of Credit Cooperatives in Japan” (*Zenkoku Shinyo Kumiai Meibo*) and data provided by IMES. The business conditions index time series are downloaded from the web site of ESRI.

3.1 Bankruptcy and Merger Data

In this study, exit is defined by the official announcement of bankruptcy or merger, instead of the actual reorganization.[9] When studying the bankruptcy and merger hazard, analysis based on the actual date of reorganization may draw a very misleading picture, especially in the current empirical study. In the case of bankruptcy, the cooperative is usually either placed in receivership or granted the right to reorganize under certain provisions to prevent closure.[10] On the actual date of resolution, the event might be recorded as an acquisition or a merger. Table 1 provides the information regarding the time lag in our sample between the announcement date and the actual execution date. The variation in such time lags depends on many regulation related issues and the decision of the board of the Deposit Insurance Corporation of Japan, which are beyond the scope of this paper. Clearly, using announcement data will result in a more timely and accurate analysis.

[8]We would like to thank the Institute of Monetary and Economic Studies (IMES) at the Bank of Japan and the Economic and Social Research Institute (ESRI) at the Cabinet Office of Japan for providing the electronic version of the data from fiscal years 1996 to 2001 and 1991 to 1995, respectively.

[9]In few cases when the announcement dates are not available, we replaced them with the actual reorganization dates.

[10]The Deposit Insurance Corporation of Japan (DICJ) offers financial assistance to facilitate the business transfer of the failed credit cooperatives to the assuming financial institutions. Please refer to http://www.dic.go.jp/english/e_katsudou/e_katsudou1.html for the details.

Exit Type	Obs	Max	Min	Mean	Std
Bankruptcy	118	929	101	363	223
Merge	93	562	42	223	109

Table 1: Time Lag (in days) between Announcement and Implementation

In this paper, we chose the sample period to be from April 1, 1992 to March 31, 2002, which covers the main banking crisis in Japan after the burst of the twin bubbles of the stock market and the real estate market, as well as a period of a gradual regulatory policy shift from blanket protection convoy system to market discipline (see Murata and Hori (2004)). Both the financial crisis and the implementation of the deregulation policy lead to an accelerating wave of restructuring and consolidation among the existing financial institutions at all levels in this ten-year sample period. Since credit cooperatives are smaller in size relative to other types of financial institutions, they could not handle negative economic shocks very well, which gives us a unique opportunity to collect many bankruptcy and merger events within a relatively short period of time. Such abundance of events greatly helps the identification of the interdependence between the bankruptcy and merger processes.

Most credit cooperatives were established after World War II and the total number of credit cooperatives were quite stable until 1990s. Exits due to either bankruptcy or merger were sporadic before 1990s. Ideally we would like to use the information of the entire history of a given credit cooperative since its establishment. Unfortunately, the financial statement information is not available until Fiscal Year 1991 and the earliest sample period we could choose is Fiscal Year 1992 due to the usage of lagged variables. In this paper, we assume that the information before Fiscal Year 1991 are relevant in predicting the probability and timing of exit during our sample period only in the sense that it affected the covariate path up to the beginning of our sample period. A theocratically rigorous treatment of left censoring problem (see Amemiya (1995)) would require prior knowledge of the distribution of the starting time and the value of the covariates since the starting time, which is not available in our data set. In addition, we are constrained from investigating the regime switch between the low exit period and the high exit period by the limited data set.

Our data set consists of all credit cooperatives that were operating in Japan at the beginning of the 1992 fiscal year[11] and each of them is followed up to the year when it exits due to either bankruptcy or merger, or until the end of our study period, whichever comes first. After deleting the cooperatives that did not provide complete financial statements, we have 378 cooperatives in total. Among them, 118 filed for bankruptcy during our sample period; 93 "disappeared" from the data set because they merged with other cooperatives or banks.[12] Based on their relative size and the headquarters of the post-merger cooperatives, as well as the announcement news of each merger, we further classify credit cooperatives that are involved in merger activities into active merging (bidding) cooperatives and passive merging (target) cooperatives. Usually the post-merger cooperative keeps the bidder's name and headquarter, but not always so. Such distinction is necessary because the two types have different characteristics and their incentives to merge also differ. Of the 93 cooperatives that merged with other financial institutions, about one-third are bidders and the remaining two-thirds are targets by our definition.

To obtain a general idea of the distribution of exits, we did a preliminary analysis without explanatory variables, using the standard nonparametric method of Kaplan-Meier estimators.[13] Figures 1 and 2 show the estimated Kaplan-Meier hazard function and survival function, respectively. In estimating the bankruptcy (merger) hazard, we treat merger (bankruptcy) cases as "censored" data, based on the implicit assumption that the bankruptcy and merger hazards are independent. Later, in our full-fledged analysis, we will relax this assumption. As can be seen in Figures 1 and 2, both bankruptcies and mergers occurred more often after fiscal year 1996, when the "big bang" policy took effect in Japan; however, the two types of exits have different time

[11] The fiscal year begins on April 1st in Japan.

[12] After a credit cooperative merges with other financial institution, we remove it from our data set with the concern that the post-merge cooperative may have systematically different characteristics from the other cooperatives in our data set. Such removal is also consistent with our definition of firm exit.

[13] Suppose that the events occur at the discrete moments of $t_1, t_2, \cdots, t_k$. At period t_i, n_i cooperatives are at risk (i.e., operating at the beginning of period t_i) and d_i cooperatives exit. The Kaplan-Meier estimator of the hazard function and survival function are

$$
\begin{aligned}
h_{KM}(t) &= \frac{d_i}{n_i}, \\
S_{KM}(t) &= \prod_{t_i \le t} \left[1 - \frac{d_i}{n_i}\right], \qquad for \quad t \ge t_1; \qquad = 1, \qquad for \quad t < t_1.
\end{aligned}
$$

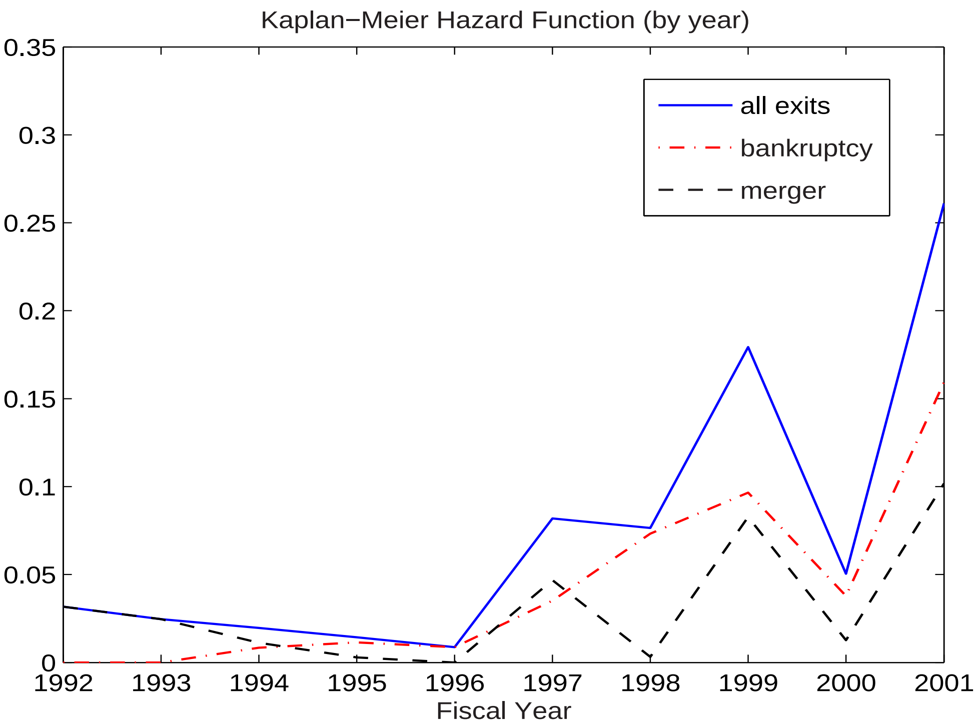

Figure 1: Kaplan-Meier Estimates of Hazard Functions

trends, partly due to the regulatory policy, and partly because of the difference in the determinating factors for the bankruptcy and merger hazards and their interdependence, which will be studied in this paper using the competing risks framework.

3.2 Selection of Covariates

In most financial ratio analysis of the banking industry, covariates are chosen according to the so-called "CAMEL" system, which is a uniform inter-agency banking rating system adopted by U.S. regulators in 1978. This rating system measures the relative soundness of a financial institution by assessing capital adequacy (C), asset quality (A), management efficiency (M), earning ability (E) and liquidity level (L), and is calculated on a 1-5 scale, with 1 being a strong performance. It has become a standard guide in the covariate selection of the bank failure literature.[14] However, this system has been criticized as failing to evaluate financial structure and growth, as

[14]See, for example, Wheelock and Wilson (2000).

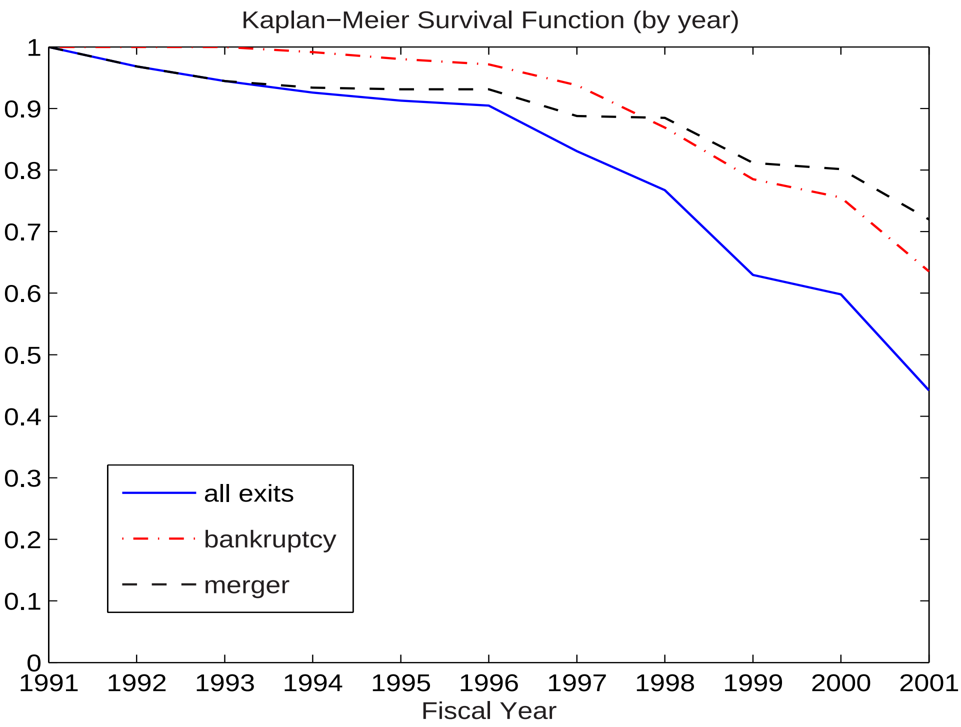

Figure 2: Kaplan-Meier Estimates of Survival Functions

well as lacking objectivity.[15] Since 1990, the World Council of Credit Unions Inc. has applied a new evaluation system called "PEARLS," which is the abbreviation of protection (P), effective financial structure (E), asset quality (A), rates of return and cost (R), liquidity (L) and signs of growth (S). Since credit cooperatives are closer to credit unions in terms of organization structure, we use "PEARLS" as the primary guide in choosing covariates.

In particular, we examine the determinants of bankruptcy and merger hazards as well as their interdependence using the following firm-specific financial ratios:

1. Protection: The variable P is calculated as the ratio of allowances for possible loan losses ($APLL$) to the total amount of outstanding loans. A more accurate measure would be using the total amount of non-performing loans (NPL) as the denominator. Unfortunately, we do not have annual data on NPL for each cooperative throughout our sample period. Notice that $P = \frac{APLL}{\text{Total Loans}} = \frac{APLL}{NPL} \cdot \frac{NPL}{\text{Total Loans}}$, that is, we can factorize P into two parts: the first factor represents how well non-performing loans are protected; the second represents the

[15] See Richardson (2002).

quality of the outstanding loans. While the first factor has a negative effect on bankruptcy hazard, the second factor increases bankruptcy hazard. Therefore, the direction of variable P's effect on bankruptcy hazards depends on which factor dominates. The effect on the merger hazards for both bidding and target firms are expected to be opposite to that on the bankruptcy hazard.

2. Financial structure: *CAPTA* denotes the ratio of institutional capital to total assets.[16] Since institutional capital has no explicit interest cost, it can be used to finance either non-earning assets or productive assets; it can also be used to absorb loan losses or operational deficits as a last resort. Therefore, everything else being equal, the higher the capital ratio is, the less likely the cooperative would go bankrupt. Adequate capital could also imply higher possibility to merge because it provides the financial strength for a credit cooperative to acquire others. In addition, capital adequacy might affect cooperatives' behavior through regulation. Since the introduction of 1988 Basel Accord, Japanese banking regulatory authority has gradually tightened the capital requirement rule. Although the requirement is only on banks at that time, small financial institutions, like credit cooperatives, might feel the pressure of meeting the minimum requirement. In fact, since fiscal year 1996, credit cooperatives have started to report capital adequacy ratio in their annual financial statements. The most preferable source of new capital in terms of raising cost is by accumulating retained earnings. However, this usually occurs over a long period of time. One way of quickly increasing institutional capital to meet the requirement is to merge with other financial institutions. This hypothesis of regulatory-induced merging implies that, everything else being equal, the lower the capital ratio, the more likely the cooperative would seek to merge with others. *TLTA* is the ratio of total liabilities to total assets and appears in most bank failure analysis. Since a higher leveraged cooperative is more prone to bankruptcy risk, we expect *TLTA* to affect the bankruptcy hazard positively. It has been proposed that merging sometimes serves as an instrument to change financial structures (see Gugler and Konrad (2002)). In particular, this hypothesis suggests that firms that have more extreme financial structures

[16] Notice that *CAPTA* is not adjusted by asset risks. We do not have risk-adjusted capital adequacy ratio until 97 fiscal year. Even then, this variable is not continuously reported. For consistency, we used plain capital asset ratio for the entire sample period. Moreover, Hooks (1995) shows that risk-based capital measure would not add substantially to the accuracy of bank failure prediction, which reassure us that the result would not change much if we used risk-adjusted capital measure instead.

are more likely to merge, everything else being equal. To test this hypothesis, we measure a firm's relative position in terms of financial structure by the absolute value of the deviation of each firm's *TLTA* relative to the industry average, denoted as "*TLTAabs*", and include it in the merging hazard equations.

3. Asset quality: *PREMTA* measures fixed assets, such as premises and equipment, as a percentage of total assets. Since these fixed assets do not generate income directly, a healthy financial institution should limit the size of its non-earning assets, relative to its total assets. We expect a positive correlation between *PREMTA* and the bankruptcy hazard, especially because the burst of real estate bubble greatly shrank the market value of the real estate owned by the institutions relative to their book value. Another important measurement of asset quality is the ratio of delinquent loan balances to the total outstanding loans. Unfortunately, we could not obtain information to calculate the delinquency ratio for each credit cooperative each year.

4. Rates of return and cost: *LNINTLN* is the ratio of net interest income on loans to total outstanding loans, which measures the earning generating capacity of the cooperative's core business. Higher earning power implies lower probability to fail and higher probability to expand by acquiring other credit cooperatives. However, the direction of the effect on the passive merger hazard is not determinate. On one hand, a financially healthier cooperative is more attractive to the potential bidders if immediate profitability is the incentive for the merger. On the other hand, poor profitability can be an indicator of poor managerial performance. The hypothesis that mergers serve to drive out bad management implies that cooperatives with lower earning power can be good candidates for merger if the bidder believes that its superior managerial skills can overturn the performance of the target cooperative. *GENEXPTA* measures the size of general and administrative expenses relative to the total assets. High administrative cost is one of the main reasons that many credit cooperatives are not profitable; hence, it is expected to be positively correlated with the bankruptcy hazard and negatively correlated with the active merger hazard. Similar to *LNINTLN*, depending on the motivation for a merger, *GENEXPTA* might have either positive or negative effect on the passive merger hazard.

5. Liquidity: Effective liquidity management is the core of risk management, especially when

a bank run might be triggered by the unsound performance of a major cooperative in the region. We measure liquidity by two variables: one is the ratio of liquid assets (including cash, dues from banks, short term loans to other financial institutions and government bonds) to total deposits, denoted as *LIQDPST*; the other is the interaction term of the percentage of short term deposits (including checking and ordinary deposits)and the number of failed cooperatives within the same prefecture, denoted as *SHORTFAIL*. The latter variable is intended to capture the effect of bank-run induced liquidity shortage. Since short term deposits do not subject to withdrawal penalty and can be withdrawn without notice in advance, they are always the first outflow from a financial institution if it is under the threat of bank run. Meanwhile, the more cooperatives fail in the local market, the more likely a bank run will occur. Hence, the interaction variable of the two forces is expected to be positively correlated with the bankruptcy hazard if the bank-run hypothesis is true.

6. Growth: Last, but not least, the continuous growth of total assets, savings deposits, membership and institutional capital reflects both sustained profitability and strong customer confidence. Since the four growth rates are highly correlated and share similar empirical distributions in our data set, we chose growth rate of savings deposits as a representative factor, denoted as "*DPSTGROW*".

Besides the above PEARLS variables, we also control for three other firm-specific characteristics (*SIZE, REGIONAL*, σ_{NITA}),[17] as well as market conditions (*BIZCONDI, NEARBYFAILSH*), defined as below.

In many empirical studies on firm exits, size is usually a significant covariate. However, this might not be true for Japan's credit cooperatives during this sample period. In particular, prior to 1996, the banking regulation policy in Japan was so called "too big to fail" (TBTF) policy. In other words, the government tended to protect financial institutions with relatively large size in order to maintain the stability of the financial system. However, after 1996, the regulation policy switched to the "big bang" deregulation policy, by which the size itself would not bring additional protection from the government besides the possible correlation of size and performance. Al-

[17] Many firm-life studies include firm age as a covariate. In our study, however, firm age is insignificant in all three hazards. For the sake of parsimony, we omitted it in the models reported in the paper.

though the TBTF policy does not directly apply to credit cooperatives, Murata and Hori (2004) show that cooperative members have gradually relinquished their special trust in larger size as the deregulation policy took place after 1996. Since most exits in our data set are observed after 1996, if the deregulation policy affects the bankruptcy and merger processes, we would expect that size might not be as influential as in other empirical studies on firm exits. To test the effect of such policy change, we include size as one of the covariates. As we know, credit cooperatives are restricted to operate within a given prefecture,[18] therefore, a meaningful measure of size would be the relative market share of a given credit cooperative in its prefecture, in terms of total assets, total number of employees, total number of members, or total amount of deposits. Since all of the above possible measures of relative size have similar empirical distribution in our data set, we define *SIZE* as the percentage of a given cooperative's total assets relative to all the credit cooperatives operating in the same prefecture.

As we mentioned earlier, credit cooperatives are membership based financial institutions. According to the common feature of their members, credit cooperatives are further classified into four types: regional (serving small and medium-sized enterprises (SMEs) and residents within certain region), industry-based (serving SMEs and individuals engaged in the same profession), occupation-based (serving individuals work at the same establishment), and foreign (organized by foreign residents and foreigner-owned SMEs). A cooperative with occupational or industry-wise concentrated customers tends to be affected by the same economic shocks from the local market. In contrast, a regional-based cooperative has better chance to diversify such risks. Therefore, we include a dummy variable of whether a credit cooperative is regional in modeling the bankruptcy and merger hazards, denoted as "*REGIONAL*". Everything else being equal, we expect regional credit cooperatives are less likely to exit through bankruptcy or being acquired, while more likely to expand through acquiring other institutions.

Another factor that we need to account for is the uncertainty or risk of a credit cooperative's profitability. Such uncertainty can be attributed to idiosyncratic risk associated with local market, or deliberate risk-taking behavior of a given credit cooperative. Higher uncertainty in profit increases the likelihood for a credit cooperative to go bankrupt, and may also decrease the

[18] Prefectures in Japan somewhat correspond to states in the U.S., although with much less political autonomy.

likelihood for it to merge with others either as a bidder or a target. We measure this risk factor by the volatility of the ratio of net income to total assets, denoted as σ_{NITA}.

In addition to the risks experienced by individual credit cooperatives, there is also a list of fundamental shocks that lead to massive failures or mergers. Macro level variables affect the performance of individual credit cooperatives both directly through changes in the values of their financial assets and interest margin, and indirectly through the performance of their clients. In this paper, we use the business conditions index (denoted as "*BIZCONDI*"),[19] to approximate business cycles and study its effect on the bankruptcy and merger risks. Individual credit cooperatives are more likely to experience financial stress during the downturn or at the trough of a business cycle. Hence, we expect the business condition indicator to be negatively correlated with bankruptcies and mergers that are intended to avoid such failures.

Another factor that is exogenous to an individual credit cooperative's choice set, but may force it to close, is self-fulfilling bank runs due to contagious loss of confidence following failures of other institutions. Since we do observe clusters of exit time in our data set, it is important to take the contagion effect into consideration. For this purpose, we construct a regional distress factor ("*NEARBYFAILSH*") by calculating, for each credit cooperative, the market share of the failed cooperatives within the same prefecture in each fiscal year. In theory, this factor might capture two opposite effects on the bankruptcy hazard of a given credit cooperative: on one hand, if many local credit cooperatives fail, the customers of the remaining credit cooperatives might lose confidence in their own cooperatives and withdraw their deposits, which might cause the remaining credit cooperatives more likely to go bankrupt. Such contagion effect implies a positive correlation of *NEARBYFAILSH* and the bankruptcy hazard. On the other hand, one can also argue that as more local credit cooperatives fail, the remaining cooperatives can gradually gain larger market share, therefore, become less likely to go bankrupt. The sign of the estimated parameter depends on which of these two effects dominates in the ten-year sample period.

[19] Also translated as "Composite Indices of Business Indicators". Business conditions index is compiled by aggregating the quantitative movement of indicators that are sensitive to business cycles and is used to measure the amplitude and rate (volume) of economic fluctuations. Generally speaking, when the business conditions index is rising, the economic cycle is considered to be in the expansion phase, and when declining, in the contraction phase. For more detailed description of how the index is calculated, please refer to http://www.esri.cao.go.jp/en/stat/di/di2e.html.

Following the common practice, both the micro and macro level variables are lagged by one period to ensure that the data is available to the market at the time of estimation (See, for example, Shumway (2001); Chava and Jarrow (2004)). This treatment is important to avoid the criticism of "back-casting" (See Ohlson (1980)). In addition, this also implies that all the explanatory variables are predictable processes in the sense of van den Berg (2001), which shows that with this property of explanatory variables, one can perform valid econometric inference using standard methods.

Summary statistics for these covariates are contained in Table 2. In calculating the statistics, each cooperative-year constitutes a separate observation, except for $REGIONAL$ and σ_{NITA}, for which each cooperative is counted as a separate observation. As shown in the table, there is significant variation in firm-specific characteristics, which enables us to identify the effects of different determinants.

4 Empirical Analysis

In this section, we investigate the determinants of the bankruptcy and merger hazards as well as their interdependence, using data of Japan's small financial institutions. Estimates without unobserved heterogeneities (model I), with independent unobserved heterogeneities (model II), with dependent unobserved heterogeneities (model III) and with structural dependence (model IV) are reported and compared.

The models are estimated using maximum likelihood method. We use the Kaplan-Meier estimates as the starting value for the baseline hazards of model I and set all the other coefficients at 0. Different starting values have been tried to ensure the global maximum. Various perturbation of the estimates from the simpler models is set as the starting value for the more complicated models. For models with unobserved heterogeneities, we have experimented with different numbers of mass points, but the improvement of the log likelihood by increasing the number of mass points is trivial after we introduce 2×2 points of support. More specifically, it is not worth the

Table 2: Summary Statistics of Credit Cooperatives' Characteristics

Variable	N	Mean	Std Error	Minimum	Median	Maximum
P	3206	0.0183	0.0318	0.0000	0.0070	0.4400
$CAPTA$	3206	0.0384	0.0228	0.0007	0.0318	0.1723
$TLTA$	3206	0.9501	0.0322	0.8101	0.9560	1.4515
$TLTAabs$	3206	0.0219	0.0260	0.0000	0.0147	0.4864
$PREMTA$	3206	0.0148	0.0109	0.0000	0.0135	0.0895
$LNINTLN$	3206	0.0523	0.0172	-0.0008	0.0510	0.1264
$GENEXPTA$	3206	0.0167	0.0051	0.0024	0.0168	0.0505
$LIQDPST$	3206	0.2732	0.1259	0.0038	0.2434	0.8679
$SHORTFAIL$	3206	0.1293	0.3252	0.0000	0.0000	4.9779
$DPSTGROW$	3206	1.0233	0.0927	0.4311	1.0130	2.8752
$SIZE$	3206	0.1204	0.1625	0.0003	0.0580	1.0000
$REGIONAL$	378	0.6323	0.4828	0.0000	1.0000	1.0000
σ_{NITA}	378	0.0192	0.0487	0.0000	0.0034	0.3397
$NEARBYFAIL$	3206	0.9339	2.0325	0.0000	0.0000	11.0000
$NEARBYFAILSH$	3206	0.0638	0.1452	0.0000	0.0000	1.0000

Note: Table 2 presents the summary statistics of Japan's credit cooperatives' characteristics during the period from fiscal years 1991 to 2000, which is lagged by one year from our sample period. In calculating the summary statistics for $REGIONAL$ and σ_{NITA}, each cooperative is counted as an observation; for all of the rest variables, each cooperative-year constitutes a separate observation.

loss of degrees of freedom by Akaike information criteria. Therefore, we only include 2×2 points of support in our empirical study. The results are reported in Tables 3 through 6.

Tables 3 and 5 present the annualized baseline hazards estimates from the four different models. As can be seen, they all share similar pattern as the Kaplan-Meier estimates, although smaller in magnitude because most of the variation in the exit hazards is explained by the variation of observed covariates and/or unobserved heterogeneities. The year to year variation does not follow any smoothed parametric distribution, which demonstrates the importance of estimating the baseline hazards semi-parametrically, since the estimates for coefficients β's and the distribution of the unobserved heterogeneity would be inconsistent if the parametric forms of the baseline hazards are incorrectly specified.

The estimated coefficients of the observed covariates, however, differ significantly across different model specifications, sometimes involving change of signs. We would first compare different models based on the modified and ordinary likelihood ratio tests (LRT), then analyze the coefficients from the most favored model specification.

A comparison of models I and II based on the modified likelihood ratio test strongly supports the presence of unobserved heterogeneity. The complication here is that the regularity conditions of the ordinary LRT are not satisfied in the current case: the null hypothesis of homogeneity, (i.e., $P_{11} = 1$), which is model I, lies on the boundary of the parameter space of model II. In addition, this hypothesis test runs into the Davies problem:[20] parameters v_2^B and v_2^M in model II are not identifiable under the null hypothesis because the two statements "$P_{11} = 1$" and "$v_2^B = v_2^M = 0$" are equivalent. Chen and Chen (2001) showed that because of the breakdown of the regularity condition, the asymptotic distribution of the ordinary LRT for homogeneity is very complicated and involves the supremum of a Gaussian process. Instead of using the ordinary LRT, we conducted a modified LRT proposed by Chen et al. (2001), which retains the power of the ordinary LRT while having a much simpler asymptotic distribution under the null hypothesis of homogeneity. More specifically, the asymptotic distribution of the modified LRT statistics is $\frac{1}{2}\chi_1^2 + \frac{1}{2}\chi_0^2$ under the null hypothesis, where χ_1^2 is a chi-square distribution with 1 degree of freedom and χ_0^2 is a degenerate distribution with all its mass at 0. Since models I and II assume the independence of the bankruptcy hazard and the merger hazards, we conducted two separate tests for heterogeneity. Please refer to Appendix B for the details of how to construct the modified LRT statistics. The values of the modified LRT statistics are 5.4544 (P-value = 0.0098) for the bankruptcy hazard and 6.1462 (P-value = 0.0066) for the active and passive merger hazards. Both tests strongly support the presence of heterogeneity.

A comparison of Models II and III shows a very surprising result. Almost no gain is recognized in terms of log likelihood by further relaxing the restriction of independence between the heterogeneities associated with the bankruptcy and merger hazards. In the case of bivariate discrete distribution with 2×2 mass points, as assumed in our model, independence assumption is equivalent to imposing $P_{11}P_{22} = P_{12}P_{21}$. It is easy to show that under this condition,

[20] See Davies (1977).

$Pr(v^m = v_i^m | v^n) = Pr(v^m = v_i^m), \quad m, n = B, M; \, i = 1, 2$. Under the null hypotheses of independent heterogeneity the ordinary LRT statistics $-2 \log \frac{\mathcal{L}_{\text{model II}}}{\mathcal{L}_{\text{model III}}}$ is distributed as χ^2 with 1 degree of freedom. In particular, the P-value of the calculated LRT statistics of model II against model III is as high as 0.3899. The estimated correlation of the unobserved heterogeneities in model III is not significantly different from 0. An immediate conclusion would be that the bankruptcy and merger hazards are independent. However, as we mentioned earlier, there are multiple sources of interdependence between the two processes. On one hand, credit cooperatives that are in danger of bankruptcy might seek to merge with others, which induces a positive correlation between the bankruptcy and passive merger hazards; on the other hand, credit cooperatives that are more likely to expand through merging might be less prone to bankruptcy risks, and credit cooperatives that have high bankruptcy risks are less attractive as merging targets, both of which imply a negative correlation between the unobserved heterogeneities associated with the bankruptcy and merger hazards. It is possible that model III fails to detect any correlation between the two heterogeneities simply because these two forces cancel each other.

To disentangle these two sources of interdependence, model IV introduces an additional term in the passive merger hazard function to capture the "structural" relationship induced by firms' preventive behavior. The calculated value of the test statistics $-2 \log \frac{\mathcal{L}_{\text{model III}}}{\mathcal{L}_{\text{model IV}}}$ is 58.34, with a P-value of 0.0000, which strongly supports the existence of interdependence between time to bankruptcy and time to merger. Unlike model III, model IV detects a statistically significant negative correlation between v^B and v^M at 1% level. In addition, the coefficient of the structural dependence, δ, is positive and statistically significant at 1% level. This result confirms that the earlier conclusion of independence is a consequence of neglecting the structural dependence. More specifically, the negative sign of the estimated correlation between the unobserved heterogeneities seems to suggest that there are some unmeasurable factors that make a firm more likely to merge and less likely to go bankrupt, and vice versa. The positive sign of δ seems to support the hypothesis of structural dependence induced by firms' preventive behavior.

Since all of the simpler specifications are rejected in favor of model IV, the following analysis on the effects of covariates is based on the result of model IV, as shown in Table 6.

First, we investigate the bankruptcy hazard. All of the coefficients conform to the hypothesized signs. In particular, the following performance-related factors significantly decrease the bankruptcy hazard of a given credit cooperative: capital adequacy (*CAPTA*), profitability of core business (*LNINTLN*), liquidity (*LIQDPST*), growth rate of deposits (*DPSTGROW*); in contrast, leverage ratio (*TLTA*), proportion of non-earning assets (*PREMTA*), ratio of general and administrative expense over total assets (*GENEXPTA*) and uncertainty of profitability (σ_{NITA}) significantly increases a credit cooperative's bankruptcy hazard. In addition, everything else being equal, a credit cooperative that is regional-based has lower bankruptcy hazard than the industrial, occupational or foreigner-based cooperatives, which confirms our hypothesis that diversification reduces bankruptcy hazards. Another interesting finding is that a credit cooperative's size, in terms of total assets relative to those of the local market, does not have a significant effect on the bankruptcy hazard. This seems to suggest that the deregulation policy introduced during our sample period has reduced, if not removed, the traditional trust in larger size, which is consistent with the finding of Murata and Hori (2004). We also detect the effects of Macro factors. In particular, *NEARBYFAILSH*, defined by the market share of the failed credit cooperatives within the same prefecture at each period, is statistically significant and carries a positive sign, which seems to suggest that during this period of time, contagion effect dominates market share effect to play an important role on bankruptcy hazard. In addition, business cycle affects bankruptcy hazard mainly through the changes of the credit cooperatives' asset value and the financial health of their clients. In view of correlation among bankruptcies, these two variables can be regarded as controls for contagion correlation and cyclical correlation. The result supports the well held view that bankruptcy risks are negatively correlated with the real economic activities over the business cycles.

Next, we turn to the merger hazards for bidders and targets. Some of the factors affects the two hazards in the same direction. More specifically, credit cooperatives that are well protected (P), adequately capitalized (*CAPTA*), with less extreme leverage ratio (*TLTAabs*), highly profitable (*LNINTLN*), and relatively stable in terms of profitability (σ_{NITA}) are more likely to merge with others, either as targets or bidders. Our empirical result does not support the hypothesis that cooperatives with more extreme financial structure tend to merge with others sooner. Instead, we find that cooperatives with more balanced financial structure are more likely to merge as targets or bidders. We did not find evidence that capital adequacy requirement might drive some credit cooperatives to merge. Like bankruptcy hazard, merger movement also appears to be countercyclical during our sample period. Other factors affect the two hazards in the opposite directions and deserve individual attention. First, cooperatives with more fixed assets (*PREMTA*) and less liquid assets (*LIQDPST*) tend to merge with cooperatives with less fixed assets and more liquid assets. This suggests that merger is used as an instrument to achieve efficiency through financial synergy. Second, cost inefficiency, measured by the ratio of general and administrative expenses to total assets (*GENEXPTA*), decreases active merger hazard, while increases the passive merger hazard. This confirms the hypothesis that mergers are partly motivated by the potential efficiency gain of driving out the bad management of the inefficient targets, as discussed in Hannan and Rhoades (1987). Third, cooperatives with slower rate of growth (*DPSTGROW*) tend to merge with cooperatives that are enjoying a fast growth, which implies that merger is also used as an instrument to boost growth when internal growth has slowed down.

The above analysis is based on the significance level and sign of each covariate, as shown in table 6. But to understand the relative importance of each variable to the determinants of the bankruptcy, active merger and passive merger hazards, respectively, we need calculate the changes of the three hazard functions due to one standard deviation change of each covariate, which are displayed in table 7. For the passive merger hazard, we calculate both the direct effect and the total effect after taking into consideration that some variables affect the bankruptcy hazard and therefore indirectly affect the passive merger hazard through the structural dependence parameter. We highlight the five most important factors for each hazard. As shown in the table, some factors, such as capital adequacy, are very important determinants of all three survival times, but in general, the variables that affect each hazard the most are quite different across the three competing risks.

Although we reject models I to III in favor of model IV, it is still worthwhile to compare estimates from those models to model IV in order to illustrate the difference. A quick comparison reveals that model IV identifies more factors that affect the bankruptcy and merger hazards both significantly and in the direction of what economic theories predict. For example, profitability (*LNINTLN*) is never a significant factor for the bankruptcy hazard in models I to III. Another example is cost efficiency (*GENEXPTA*), which carries an opposite sign for the passive merger hazard in models I to III comparing to model IV. Without listing all of the items, we end our comparison by noting that ignoring the interrelation of the bankruptcy and merger hazards would result in inconsistent estimates, sometimes involving change of signs.

Ideally, we would like to compare the forecast accuracy of the four models using out-of-sample data. Unfortunately, we only have data for the sample period covered in this paper. In addition, our sample size is not large enough for the experiment to estimate the models using part of the data and compare the out-of-sample predictive results using the rest of the data. Therefore, we can only compare the in-sample goodness of fit for different models using the available data. Figure 3 presents the cumulative accuracy rate of bankruptcy for all four models. For each firm each year from fiscal year 1992 to 2001, we apply the estimated parameters from each model to the covariates of the year before and calculate the fitted values of their bankruptcy hazard. We then sort all firm-years into 10 deciles based on their fitted values of the bankruptcy hazard, from the lowest to the highest, which are displayed along the horizontal axis. The vertical axis displays the cumulative percentage of bankrupt firms corresponding to each decile. To get a cumulative accuracy curve for each model, we accumulate the fraction of the bankrupt firms whose fitted values of bankruptcy hazard are equal to or lower than any given decile. The closer the curve to the bottom and right border of the box, the better the model fits the data. Such shaped curve implies that the model "anticipates" (in the sense that the hazard value is calculated using last year's covariates) most of the bankruptcies and assigns a much higher value of the bankruptcy hazard to those firms that actually went bankrupt the next year than the survived firms. The worst-case scenario is the dotted diagonal line, which implies that the model does nothing better than randomly assigning hazard values. Figures 4 and 5 use the same idea to compare the goodness of fit of different models in terms of active merger hazard and passive merger hazard.

As we can see from Figures 3 to 5, there is no significant difference among the four models in terms of fitting bankruptcy events. However, when fitting active merger events, model IV shows a slight advantage and model I performs the worst. Such contrast becomes the most significant in fitting the passive merger events, where model IV greatly outperforms the other three models. More specifically, at the highest decile, the improvement of model IV over the other three models is about 25%. These comparisons show that by fully accounting for the interdependence between the bankruptcy process and the merger process, one can significantly improve the accuracy of the model.

Table 3: Annualized Baseline Hazards: Independent Competing Risks

		Model I		Model II	
$\mathbf{\Lambda^B}$	Λ_1^B	0.0000	(0.0000)	0.0000	(0.0000)
	Λ_2^B	0.0000	(0.0000)	0.0000	(0.0000)
	Λ_3^B	0.0125	(0.0247)	0.0028	(0.0019) *
	Λ_4^B	0.0165	(0.0183)	0.0049	(0.0026) **
	Λ_5^B	0.0151	(0.0189)	0.0052	(0.0030) **
	Λ_6^B	0.0493	(0.0240) **	0.0297	(0.0080) ***
	Λ_7^B	0.0479	(0.0380)	0.0288	(0.0084) ***
	Λ_8^B	0.1066	(0.0441) ***	0.0268	(0.0067) ***
	Λ_9^B	0.0507	(0.0223) **	0.0193	(0.0073) ***
	Λ_{10}^B	0.1556	(0.0992) *	0.1188	(0.0268) ***
$\mathbf{\Lambda^M}$	Λ_1^M	0.0113	(0.0049) **	0.0020	(0.0007) ***
	Λ_2^M	0.0234	(0.0887)	0.0012	(0.0004) ***
	Λ_3^M	0.0176	(0.0424)	0.0005	(0.0003) **
	Λ_4^M	0.0050	(0.0054)	0.0002	(0.0002)
	Λ_5^M	0.0000	(0.0000)	0.0000	(0.0000)
	Λ_6^M	0.1184	(0.0425) ***	0.0102	(0.0025) ***
	Λ_7^M	0.0086	(0.0084)	0.0007	(0.0007)
	Λ_8^M	0.5164	(0.1145) ***	0.0201	(0.0039) ***
	Λ_9^M	0.0544	(0.1007)	0.0029	(0.0016) **
	Λ_{10}^M	0.4497	(0.2636) **	0.0447	(0.0117) ***

Note: This table reports the maximum likelihood estimators of annualized baseline hazards from model specifications I and II. $\Lambda_t^m = \int_{t-1}^t \lambda^m(s)\, ds, \quad m = B, M.$

Standard errors are in parentheses.

* Statistically significant at 10% level;

** Statistically significant at 5% level;

*** Statistically significant at 1% level;

Since $\{\Lambda_t^B\}_{t=1}^{10}$ and $\{\Lambda_t^M\}_{t=1}^{10}$ are non-negative by definition, the rejection rule is based on one-sided t tests.

Table 4: Bankruptcy and Merger Hazards: Independent Competing Risks

		Model I			Model II		
β_B	P	3.7087	(1.8015)	**	-0.9879	(1.3026)	
	$CAPTA$	-39.8492	(10.4165)	***	-33.9229	(2.5928)	***
	$TLTA$	5.2644	(0.2148)	***	15.0293	(0.2404)	***
	$PREMTA$	7.0037	(10.2404)		11.0143	(5.0387)	**
	$LNINTLN$	4.3576	(17.2945)		-10.9338	(7.9302)	
	$GENEXPTA$	-13.3868	(11.0176)		22.0397	(3.5400)	***
	$LIQDPST$	-2.5770	(0.6428)	***	-1.9586	(0.9021)	**
	$SHORTFAIL$	0.1491	(0.2851)		0.1653	(0.1538)	
	$DPSTGROW$	-2.2446	(0.6919)	***	-4.0848	(0.2709)	***
	$SIZE$	-0.7920	(1.8053)		-0.7246	(0.6435)	
	$REGIONAL$	-0.4463	(0.6329)		-0.6214	(0.1737)	***
	σ_{NITA}	2.8826	(1.8209)		3.0337	(1.1867)	**
	$NEARBYFAILSH$	2.0109	(1.5558)		1.9809	(0.3694)	***
	$BIZCONDI$	-1.3441	(0.4725)	***	-13.7785	(0.2630)	***
β_{MA}	P	11.2836	(56.6758)		25.7006	(4.9282)	***
	$CAPTA$	28.4064	(9.7663)	***	28.3457	(9.3710)	***
	$TLTAabs$	-38.1815	(19.1034)	**	-38.4447	(12.7003)	***
	$PREMTA$	7.4461	(86.6634)		-5.1517	(3.2380)	
	$LNINTLN$	43.0843	(10.9823)	***	49.3211	(0.8168)	***
	$GENEXPTA$	-115.6365	(37.9544)	***	-127.7903	(20.4752)	***
	$LIQDPST$	-5.2205	(1.7445)	***	-5.4661	(1.5273)	***
	$DPSTGROW$	-4.4786	(1.0551)	***	-5.0822	(0.4341)	***
	$SIZE$	0.5661	(4.8716)		0.4597	(0.8848)	
	$REGIONAL$	0.1594	(1.6183)		0.2890	(0.3877)	
	σ_{NITA}	-315.0718	(207.4510)		-338.3892	(34.1950)	***
	$BIZCONDI$	3.9992	(0.1713)		-7.4777	(0.2374)	
β_{MP}	P	4.5701	(10.6710)		22.0555	(3.7343)	***
	$CAPTA$	-17.9741	(7.0860)	**	-18.2962	(5.8046)	***
	$TLTAabs$	-2.8339	(11.1413)		-0.0465	(4.6954)	
	$PREMTA$	3.2983	(11.7290)		-7.4704	(6.2039)	
	$LNINTLN$	43.7340	(1.1202)	***	54.2050	(0.9755)	***
	$GENEXPTA$	-31.0498	(27.0699)		-26.4292	(11.6585)	**
	$LIQDPST$	-0.1310	(1.2405)		-0.4309	(0.2495)	*
	$DPSTGROW$	-3.8602	(1.5030)	**	-5.2054	(0.3625)	***
	$SIZE$	-1.6952	(1.1127)		-1.8676	(0.9439)	**
	$REGIONAL$	-0.5566	(1.0462)		-0.4779	(0.2110)	**

Continued on next page

Table 4 (continued)

		Model I			Model II		
	σ_{NITA}	-187.5041	(52.6955)	***	-230.2369	(8.7568)	***
	$BIZCONDI$	3.1153	(0.2278)		-8.1373	(0.3766)	
V	v_2^B	—			5.0969	(0.3250)	***
	v_2^M	—			14.7062	(0.2426)	***
	P_{11}	—			0.0122	(0.0075)	*
	P_{12}	—			0.0298	(0.0184)	*
	P_{21}	—			0.2782	(0.0180)	***
	P_{22}	—			0.6798	(0.0233)	***
$Log\mathcal{L}$		-646.75			-633.74		

Note: Standard errors are in parentheses. *, ** and *** represent that the estimate is statistically significant at 10%, 5% and 1% level, respectively; t tests for $P_{11}, \cdots, P_{22}$ are one-sided since the variables are non-negative by definition. t tests for all the other variables are two-sided.

Table 5: Annualized Baseline Hazards: Dependent Competing Risks

		Model III		Model IV	
$\mathbf{\Lambda^B}$	Λ_1^B	0.0000	(0.0000)	0.0017	(0.0011) *
	Λ_2^B	0.0000	(0.0000)	0.0006	(0.0004) *
	Λ_3^B	0.0027	(0.0028)	0.0010	(0.0007) *
	Λ_4^B	0.0047	(0.0031) *	0.0013	(0.0010) *
	Λ_5^B	0.0051	(0.0041)	0.0011	(0.0010)
	Λ_6^B	0.0288	(0.0188) *	0.0031	(0.0014) **
	Λ_7^B	0.0279	(0.0112) ***	0.0031	(0.0013) ***
	Λ_8^B	0.0259	(0.0120) **	0.0046	(0.0021) **
	Λ_9^B	0.0188	(0.0113) **	0.0033	(0.0014) ***
	Λ_{10}^B	0.1157	(0.0430) ***	0.0120	(0.0049) ***
$\mathbf{\Lambda^M}$	Λ_1^M	0.0020	(0.0009) **	0.0078	(0.0066)
	Λ_2^M	0.0012	(0.0006) **	0.0078	(0.0059) *
	Λ_3^M	0.0005	(0.0003) *	0.0022	(0.0019)
	Λ_4^M	0.0002	(0.0002)	0.0008	(0.0008)
	Λ_5^M	0.0000	(0.0000)	0.0000	(0.0000)
	Λ_6^M	0.0103	(0.0044) ***	0.0191	(0.0113) **
	Λ_7^M	0.0007	(0.0007)	0.0020	(0.0021)
	Λ_8^M	0.0203	(0.0075) ***	0.0265	(0.0139) **
	Λ_9^M	0.0029	(0.0019) *	0.0084	(0.0081)
	Λ_{10}^M	0.0452	(0.0179) ***	0.0435	(0.0262) **

Note: This table reports the maximum likelihood estimators of annualized baseline hazards from model specifications III and IV. $\Lambda_t^m = \int_{t-1}^t \lambda^m(s)\, ds, \quad m = B, M$. Standard errors are in parentheses. *, ** and *** represent that the estimate is statistically significant at 10%, 5% and 1% level, respectively; t tests are one-sided since these variables are non-negative by definition.

Table 6: Bankruptcy and Merger Hazards: Dependent Competing Risks

		Model III			Model IV		
β_B	P	-1.0814	(1.1371)		-1.2412	(1.0189)	
	$CAPTA$	-33.7024	(4.2944)	***	-39.9280	(1.0809)	***
	$TLTA$	15.1268	(0.3667)	***	14.0021	(0.8237)	***
	$PREMTA$	11.0194	(6.4281)	*	4.1963	(1.0115)	***
	$LNINTLN$	-10.7063	(9.1123)		-29.8984	(1.0220)	***
	$GENEXPTA$	21.7312	(10.2777)	**	16.9230	(3.1179)	***
	$LIQDPST$	-1.9253	(0.7548)	**	-2.4486	(0.8655)	***
	$SHORTFAIL$	0.1648	(0.1718)		-0.0740	(0.2004)	
	$DPSTGROW$	-4.0994	(0.3003)	***	-4.9620	(0.9282)	***
	$SIZE$	-0.7212	(0.6221)		-1.2679	(0.8304)	
	$REGIONAL$	-0.6156	(0.1830)	***	-0.7945	(0.3146)	**
	σ_{NITA}	3.0696	(1.2080)	**	1.7836	(1.0174)	*
	$NEARBYFAILSH$	1.9872	(0.4147)	***	3.0059	(0.3158)	***
	$BIZCONDI$	-13.8594	(0.0367)	***	-7.4426	(0.7038)	***
β_{MA}	P	25.7825	(6.1116)	***	26.6524	(1.0399)	***
	$CAPTA$	28.3952	(2.7436)	***	24.1645	(1.0175)	***
	$TLTAabs$	-38.4761	(3.0938)	***	-33.8954	(1.0803)	***
	$PREMTA$	-5.2020	(4.1129)		3.6478	(1.2020)	***
	$LNINTLN$	49.5637	(8.6750)	***	20.1224	(1.0901)	***
	$GENEXPTA$	-129.7245	(21.1270)	***	-104.7880	(0.9253)	***
	$LIQDPST$	-5.4991	(1.7540)	***	-4.8011	(0.9768)	***
	$DPSTGROW$	-5.0725	(0.6366)	***	-2.7815	(1.0883)	**
	$SIZE$	0.4524	(1.0975)		0.4957	(0.9922)	
	$REGIONAL$	0.3049	(0.3594)		0.1227	(0.5956)	
	σ_{NITA}	-339.2083	(9.1556)	***	-334.9142	(1.0179)	***
	$BIZCONDI$	-7.5247	(0.6712)	***	-1.1818	(0.4449)	***
β_{MP}	P	22.1073	(4.2599)	***	19.7416	(1.1447)	***
	$CAPTA$	-18.3130	(3.7805)	***	47.8471	(1.5035)	***
	$TLTAabs$	0.0995	(5.8541)		-7.2013	(1.0154)	***
	$PREMTA$	-7.5855	(5.1966)		-46.2139	(1.0309)	***
	$LNINTLN$	54.3805	(4.2959)	***	133.5790	(1.0877)	***
	$GENEXPTA$	-27.4656	(12.7880)	**	8.6248	(1.1667)	***
	$LIQDPST$	-0.4393	(1.6443)		2.2205	(0.9957)	**
	$DPSTGROW$	-5.2195	(0.7354)	***	1.1128	(0.3298)	***
	$SIZE$	-1.8811	(1.0995)	*	-0.9477	(1.0058)	
	$REGIONAL$	-0.4647	(0.3400)		0.0952	(0.6271)	

Continued on next page

Table 6 (continued)

		Model III			Model IV		
	σ_{NITA}	-230.8059	(16.9547)	***	-281.8690	(1.0008)	***
	$BIZCONDI$	-8.1736	(0.4493)	***	-8.5881	(0.7395)	***
	v_2^B	5.1040	(0.3050)	***	4.2405	(0.4174)	***
	v_2^M	14.7572	(0.3796)	***	6.3012	(0.6336)	***
	P_{11}	0.0000	(0.0000)		0.0000	(0.0000)	**
V	P_{12}	0.0451	(0.0319)	*	0.0825	(0.0368)	**
	P_{21}	0.2972	(0.1805)	*	0.2581	(0.0886)	***
	P_{22}	0.6577	(0.1985)	***	0.6593	(0.0950)	***
	$corr(v^B, v^M)$	-0.1413	(0.0985)		-0.1769	(0.0585)	***
δ	h_B	—			1.4114	(0.1241)	***
$Log\mathcal{L}$		-633.61			-604.44		

Note: Standard errors are in parentheses. *, ** and *** represent that the estimate is statistically significant at 10%, 5% and 1% level, respectively; t tests for $P_{11}, \cdots, P_{22}$ are one-sided since the variables are non-negative by definition. t tests for all the other variables are two-sided.

Table 7: Change of Hazards due to 1σ Change of Covariates: Model IV

Variables	Bankruptcy		Active Merger		Passive Merger		
					Direct		Total
P	-0.04		1.33	***	0.87	***	0.77
$CAPTA$	-0.60	***	0.73	***	1.98	***	-0.18
$TLTA$	0.57	***	–		–		0.89
$TLTAabs$	–		-0.59	***	-0.17	***	-0.17
$PREMTA$	0.05	***	0.04	***	-0.40	***	-0.36
$LNINTLN$	-0.40	***	0.41	***	8.95	***	3.82
$GENEXPTA$	0.09	***	-0.41	***	0.04	***	0.18
$LIQDPST$	-0.27	***	-0.45	***	0.32	**	-0.14
$SHORTFAIL$	-0.02		–		–		-0.03
$DPSTGROW$	-0.37	***	-0.23	**	0.11	***	-0.42
$SIZE$	-0.19		0.08		-0.14		-0.36
$REGIONAL$	-0.32	**	0.06		0.05		-0.39
σ_{NITA}	0.09	*	-1.00	***	-1.00	***	-1.00
$NEARBYFAILSH$	0.55	***	–		–		0.85
$BIZCONDI$	-0.31	***	-0.06	***	-0.35	***	-0.62

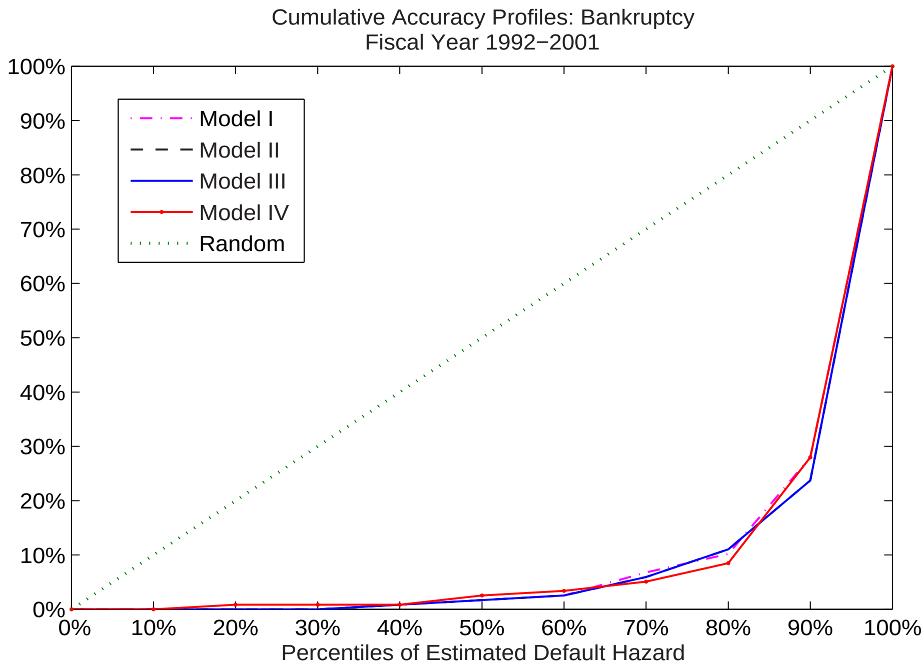

Figure 3: Cumulative Accuracy Curves: Bankruptcy

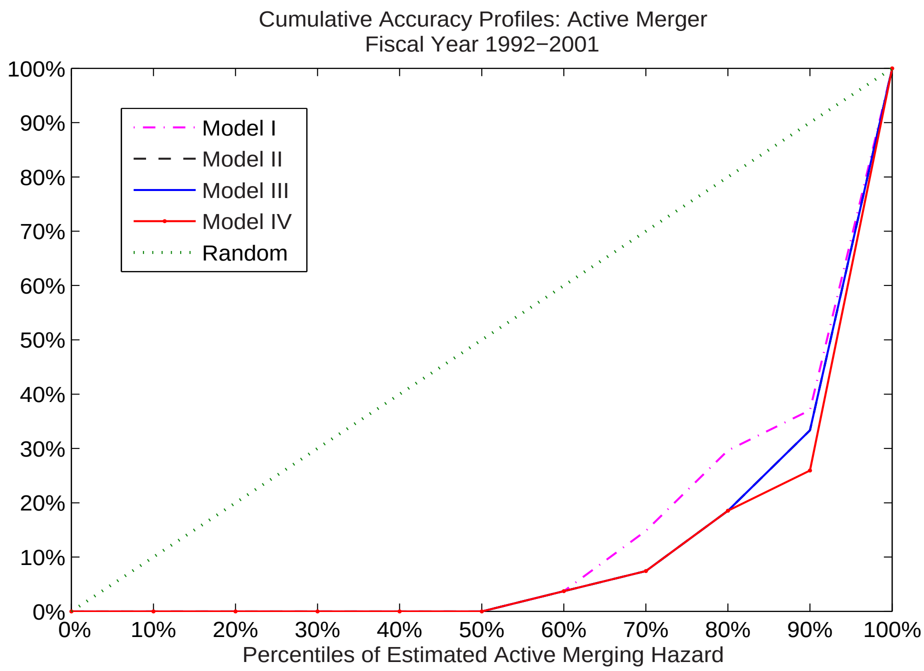

Figure 4: Cumulative Accuracy Curves: Active Merger

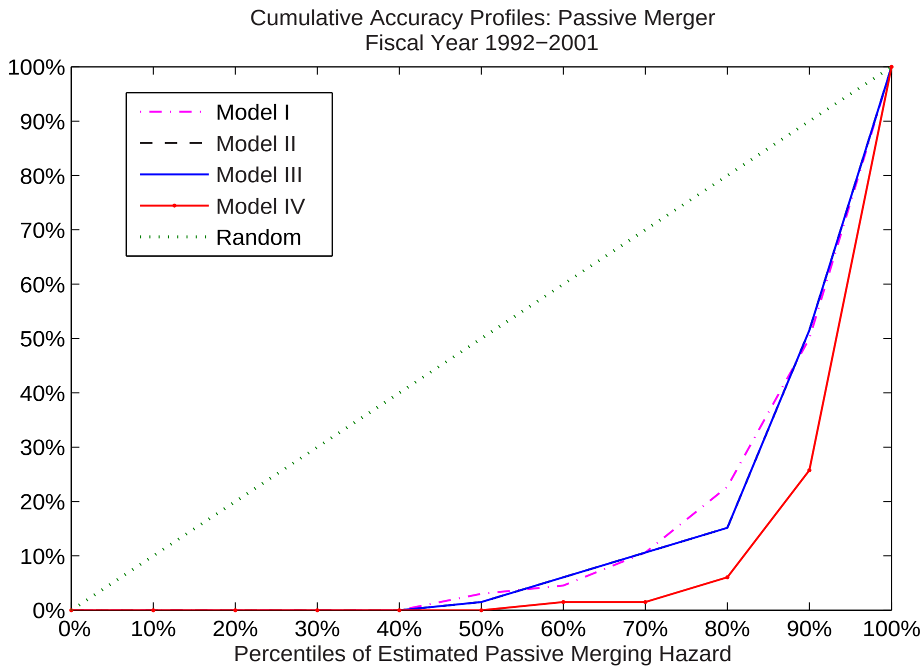

Figure 5: Cumulative Accuracy Curves: Passive Merger

5 Conclusion

This paper specifies a dependent competing risks model to investigate the determinants of the bankruptcy and merger (both active and passive) hazards as well as their interdependence using a panel data of Japan's small financial institutions. The model developed in this paper can be applied to any application that involves dynamically dependent competing risks with time-varying covariates. Our model has three important features:

First and most importantly, we incorporates both dependence through unobserved heterogeneities and dependence due to firms' preventive behavior. This general model is consistent with economic intuition and nests three other more restricted models, including independent risks without unobserved heterogeneities, independent risks with unobserved heterogeneities and dependent risks with unobserved heterogeneities. To illustrate the importance of fully capturing the interdependence between the bankruptcy and merger risks, we estimated and compared all four models using the data of the Japan's small financial institutions. Both the likelihood ratio tests and the comparison of the cumulative accuracy curves imply that more sophisticated models are superior to the simpler models, with the most general model outperforms all the restricted ones. Our estimation results demonstrate that ignoring the interdependence of the competing risks will significantly bias the estimates. Meanwhile, it is important to take structural dependence into consideration explicitly. Using only the unobserved heterogeneities to capture the interdependence might fail to detect any correlation because the different driving forces of the interdependence might cancel each other if structural dependence is not accounted for in the model.

Second, we allow much flexibility in estimating the baseline hazards and the distribution of the unobserved heterogeneities. Our result shows that the baseline hazard does not follow any smoothed parametric distribution.

Third, we include both the financial ratios of individual credit cooperatives, and some macro variables to control for external environment. In particular, we test the existence of contagion effect by constructing a distress factor based on the relative market share of the failed credit cooperatives within the same prefecture as the credit cooperative under consideration. We also study the effect of business cycles on the firm exit hazards.

The unique data set of Japan's small financial institutions, which has abundant cases of bankruptcies and mergers, provides us with the rare opportunity to investigate the influence of financial, regulatory and macro economic factors on time to bankruptcy and time to merger (both for bidding and target firms) in financial industry, as well as the interdependency of the latent survival times. In particular, we find that cooperatives with the following characteristics are more prone to bankruptcy risk: inadequate institutional capital, high leverage ratio, large proportion of non-earning and/or illiquid assets, low profitability, less cost efficiency, slow and unstable growth, and less diversification. In addition, we were able to identify correlation of bankruptcies among credit cooperatives through both cyclical and contagion factors. Credit cooperatives that are well protected, adequately capitalized with a relatively balanced financial structure, highly profitable and relatively stable in terms of profitability are more likely to merge with others. Our empirical findings also support the views that merging serves as an instrument to achieve efficiency through operating and financial synergy. Merging is also found as a way to stimulate growth.

The most important empirical finding of this paper is that time to bankruptcy and time to merger are interdependent: on one hand, firms that are likely to expand through merging are less prone to bankruptcy risk due to unobserved heterogeneity; on the other hand, firms with high bankruptcy hazard tend to increase their intensity to seek for merger in order to avoid failure. In Japan, the latter effect is further enhanced through the regulatory authority who promotes merger as an alternative to closure. Although both intuition and empirical observation suggest that such kind of interdependence among different modes of firm exit exists, it has not been explicitly accounted for in the empirical literature. Our findings demonstrate the importance and feasibility of taking explicit account of the interdependence among the competing risks. It is hoped that this paper will raise attention to the possible bias due to the independence assumption made in many empirical studies, as well as providing a method that relaxes such assumption.

This paper raises at least two immediate research questions, one empirical and the other theoretical. First, it would be interesting to investigate whether such interdependence between the bankruptcy and merger processes also exists for other industries and other time periods. The data set used in this paper is unique in that there are many events of bankruptcies and mergers within a relatively short period of time. This leads to the second question: whether the dependent competing risks model developed in this paper can cope with the more common situation where the events of interest are less frequent. In a separate paper coauthored with professor Takeshi Amemiya (Amemiya and Yu (2006)), we investigate the endogenous sampling method when the events of interest are rare. However, that paper only considers single risks duration models. In future research, it would be both theoretically challenging and empirically important to study the effects of introducing endogenous sampling to the dependent competing risks model in order to deal with the situation where the events of interest are rare.

Appendices

A Proof of Proposition 1

Proof. As pointed out in Tsiatis (1975) and noted by Abbring and van den Berg (2003), the joint distribution of the identified minimum (T, I) is fully characterized by the subsurvival functions

$$\mathcal{Q}_1(t|x) := Pr(T_1 > t, T_2 > T_1|x), \quad \mathcal{Q}_2(t|x) := Pr(T_2 > t, T_1 > T_2|x).$$

Since our goal is to identify the joint distribution of (T_1, T_2) out of that of (T, I), it is appropriate to assume $\mathcal{Q}_1(t|x)$ and $\mathcal{Q}_2(t|x)$ to be known for all t and $x_k \in \mathcal{X}$, $k = 1, 2, \cdots$ in the following proof.

According to Theorem 1 of Tsiatis (1975), we have for almost all t

$$\begin{aligned} \mathcal{Q}_i'(t|x) &= \left[\frac{\partial S(t_1, t_2|x)}{\partial t_i}\right]_{t_1=t_2=t} \\ &= \lambda_i(t)\phi_i(x_{[t]+1})D_i\mathcal{L}_G\{s_1(t, x), s_2(t, x)\}, \qquad i = 1, 2, \end{aligned} \tag{A-1}$$

where $\mathcal{Q}_i'(t|x) := \partial\mathcal{Q}_i(t|x)/\partial t$ and $D_i\mathcal{L}_G(s_1, s_2) := \partial\mathcal{L}_G(s_1, s_2)/\partial s_i$.

1. Identification of ϕ_1 and ϕ_2:

   $\forall\, t$, pick an arbitrary $x \in \{x : x_k \in \mathcal{X}, \quad k = 1, 2, \cdots, [t]+1\}$ and consider $x^* \in \{x : x_{[t]+1} = y^*\}$.

   $$\frac{\mathcal{Q}_1'(t|x)}{\mathcal{Q}_1'(t|x^*)} = \phi_1(x_{[t]+1})\frac{D_1\mathcal{L}_G\{s_1(t, x), s_2(t, x)\}}{D_1\mathcal{L}_G\{s_1(t, x^*), s_2(t, x^*)\}} \tag{A-2}$$

   As $t \downarrow 0$, Equation (A-2) $\to \phi_1(x_1)$ because $D_1\mathcal{L}_G\{s_1(0, x), s_2(0, x)\} = -E(V_1) > -\infty$ by Assumption 2. Since x is arbitrary, so is x_1. Therefore, ϕ_1 can be identified through the variation of x_1. ϕ_2 can be identified in the same way.

2. Identification of $\mathcal{L}_G$:

   Notice that $S(t, t|x) = \mathcal{Q}_1(t|x) + \mathcal{Q}_2(t|x)$ is known through the knowledge of $\mathcal{Q}_i(t|x)$. According to the normalization assumption 3, we have $S(1, 1|x) = \mathcal{L}_G\{\phi_1(x_1), \phi_2(x_1)\}$. Identification of $\mathcal{L}_G(\cdot,\cdot)$ follows from Proposition 1 of Abbring and van den Berg (2003).

3. Identification of $\{\Lambda_i(t),\, i=1,2\}$:

   Following Abbring and van den Berg (2003), we can rewrite Equation (A-1) as a system of differential equations $f=(f_1,f_2)$ of $\Lambda=(\Lambda_1,\Lambda_2)$ with initial conditions $\Lambda_1(0)=\Lambda_2(0)=0$:

$$\Lambda'(t) = f\{t,\Lambda(t)\}, \tag{A-3}$$

$$\text{where } f_i\{t,\Lambda(t)\} = \mathcal{Q}_i'(t|x)\cdot\left[\phi_i(x_{[t]+1})D_i\mathcal{L}_G\{s_1(t,x),s_2(t,x)\}\right]^{-1}, \quad i=1,2.$$

   In particular, $\forall t\in(0,1]$, we have

$$f_i\{t,\Lambda(t)\} = \mathcal{Q}_i'(t|x)\cdot\left[\phi_i(x_1)D_i\mathcal{L}_G\{\Lambda_1(t)\phi_1(x_1),\Lambda_2(t)\phi_2(x_1)\}\right]^{-1}, \quad i=1,2.$$

   Since f is known, same arguments as in Abbring and van den Berg (2003) would imply that Equation (A-3) has a unique solution Λ on (0,1].

   $\forall t\in(1,2]$, we have

$$f_i\{t,\Lambda(t)\} = \mathcal{Q}_i'(t|x)\cdot\left[\phi_i(x_1)D_i\mathcal{L}_G\left\{\begin{array}{c}\Lambda_1(1)(\phi_1(x_1)-\phi_1(x_2))+\Lambda_1(t)\phi_1(x_2),\\ \Lambda_2(1)(\phi_2(x_1)-\phi_2(x_2))+\Lambda_2(t)\phi_2(x_2)\end{array}\right\}\right]^{-1}, \quad i=1,2.$$

   Since $(\Lambda_1(1),\Lambda_2(1))$ is identified, f becomes a known function, and Equation (A-3) has a unique solution Λ on (1,2].

   Carrying out the same procedure , we can identify $\{\Lambda_i(t),\, i=1,2\}$ for all t iteratively.

The above three steps identify all the elements of the joint survival function defined by Equations (6)-(8). $\square$

B Modified Likelihood Ratio Test for Homogeneity

This section provides details about how to construct the modified likelihood ratio test statistics, which is proposed by Chen et al. (2001).

Since both model I and model II assume independency between bankruptcy hazard and merger hazards, as defined in Section 2.2, we can estimate and test the heterogeneity of bankruptcy hazards separately from that of active and passive merger hazards.

Suppose the heterogeneity V has the following discrete distribution:

$$V = \begin{cases} v^1 & \text{with probability } p \\ v^2 & \text{with probability } (1-p). \end{cases} \tag{B-1}$$

Let β be the parameter vector excluding (p, v^1, v^2). For $0 < p < 1$, the modified (penalized) log likelihood function is defined as:

$$l_n(p, v^1, v^2, \beta) = \sum_{i=1}^{N} \log\left\{(1-p)L(t_i|x_i, V = v^1) + pL(t_i|x_i, V = v^2)\right\} + C\log\left\{4p(1-p)\right\}, \tag{B-2}$$

where the kernel function $L(t_i|x_i, v)$ is the conditional likelihood function for the duration variable of our interest;[21] $C > 0$ is a constant that is used to control the level of modification.[22]

Under the null hypothesis of homogeneity, i.e. $H_0 : v^1 = v^2 = 0$,[23] the maximum likelihood estimator $\hat{\beta}^{H_0}$ can be estimated from the null likelihood function $l_n(\frac{1}{2}, 0, 0, \beta)$, in which the penalty term is equal to 0.

[21] In particular,

$$\begin{aligned} L(t_i|x_i, v_B) &= \left(\lambda_0^B(t_i) \cdot \exp\{x_B(t_i)\beta_B + v_B\}\right)^{d_B} \exp\left\{-\int_0^{t_i} \lambda_0^B(s) \cdot \exp\{x_B(s)\beta_B + v_B\}\, ds\right\} &\text{(B-3)} \\ L(t_i|x_i, v_M) &= \left(\lambda_0^M(t) \cdot \exp\{x_M(t)\beta_{MA} + v_M\}\right)^{d_{MA}} \left(\lambda_0^M(t) \cdot \exp\{x_M(t)\beta_{MP} + v_M\}\right)^{d_{MP}} \\ &\quad \exp\left\{-\int_0^{t_i} \left(\lambda_0^M(s) \cdot \exp\{x_M(s)\beta_{MA} + v_M\} + \lambda_0^M(s) \cdot \exp\{x_M(s)\beta_{MP} + v_M\}\right) ds\right\} &\text{(B-4)} \end{aligned}$$

[22] Chen et al. (2001) suggest to use $C = \log(M)$ for $V \in [-M, M]$. Their simulation result shows that the test is not sensitive to the values of C, which is confirmed by our application. The values are negligibly smaller (larger) with a larger (smaller) C. In particular, We report the values of the test statistics when $C = 4$.

[23] v^1 is normalized at 0 for reasons explained in footnote 7.

The modified likelihood ratio test statistics is then defined as

$$M_n = 2\left\{ l_n(\hat{p}, 0, \hat{v^2}, \hat{\beta}^{H_1}) - l_n(\frac{1}{2}, 0, 0, \hat{\beta}^{H_0}) \right\} \tag{B-5}$$

Theorem 1 in Chen et al. (2001) shows that the asymptotic null distribution of the above statistics M_n is $\frac{1}{2}\chi_0^2 + \frac{1}{2}\chi_1^2$, where χ_0^2 is a degenerate distribution with all its mass at 0 and χ_1^2 is a chi-square distribution with 1 degree of freedom.

References

Abbring, J. H. and van den Berg, G. J. (2003). The identifiability of the mixed proportional hazards competing risks model. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, 65(3):701–710.

Altman, E. (1968). Financial ratios, discriminant analysis and the prediction of corporate bankruptcy. *Journal of Finance*, 23(4):589–609.

Amemiya, T. (1985). *Advanced Econometrics*. Harvard University Press, Cambridge, Mass.

Amemiya, T. (1995). A note on left censoring. In Hsiao, C., Lahiri, K., Lee, L.-F., and Pesaran, M., editors, *Analysis of Panels and Dependent Variable Models*, pages 7–22. Cambridge University Press, Cambridge, England.

Amemiya, T. and Yu, X. (2006). Endogenous sampling and matching method in duration models. IMES Discussion Paper Series, No. 2006-E-6, Institute for Monetary and Economic Studies, Bank of Japan.

Atiya, A. F. (2001). Bankruptcy prediction for credit risk using neural networks: A survey and new results. *IEEE Transactions on Neural Networks*, 12(4):929–935.

Carling, K. and Jacobson, T. (1995). Modeling unemployment duration in a dependent competing risks framework: Identification and estimation. *Lifetime Data Analysis*, 1:111–122.

Chava, S. and Jarrow, R. A. (2004). Bankruptcy prediction with industry effects. *Review of Finance*, 8(4):537–569.

Chen, H. and Chen, J. (2001). The likelihood ratio test for homogeneity in the finite mixture models. *Canadian Journal of Statistics*, 29:201–215.

Chen, H., Chen, J., and Kalbfleisch, J. D. (2001). A modified likelihood ratio test for homogeneity in finite mixture models. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, 63:19–29.

Cheng, P. Y. K., Koh, H.-C., and Low, C.-K. (2003). Understanding U.S. bank failures and developing an early warning system with decision tree models. FMA International Annual Meeting.

Davies, R. B. (1977). Hypothesis testing when a nuisance parameter is present only under the alternative. *Biometrika*, 64:247–254.

Demirguc-Kunt, A. (1989). Deposit-institution failures: a review of empirical literature. *Economic Review - Federal Reserve Bank of Cleveland*, 25(4):2–18.

Deng, Y., Quigley, J., and van Order, R. (2000). Mortgage terminations, heterogeneity and the exercise of mortgage options. *Econometrica*, 68(2):275–307.

Duffie, D. J., Saita, L., and Wang, K. (2005). Multi-period corporate failure prediction with stochastic covariates. Working paper, Graduate School of Business, Stanford University, available at http://www.stanford.edu/ duffie/ddkw.pdf.

Fallic, B. and Ryu, K. (2003). The recall and new job search of laid-off workers: a bivariate proportional hazard model with unobserved heterogeneity. *Federal Reserve Board Finance and Economics Discussion Series*, 2003-22.

Focarelli, D., Panetta, F., and Salleo, C. (2002). Why do banks merge? *Journal of Money, Credit and Banking*, 34(4):1047–1066.

Gugler, K. and Konrad, K. A. (2002). Merger target selection and financial structure. Working paper, University of Vienna and Wissenschaftszentrum Berlin, available at http://homepage.univie.ac.at/Klaus.Gugler/public/rio.pdf.

Han, A. and Hausman, J. (1990). Flexible parametric estimation of duration and competing risk models. *Journal of Applied Econometrics*, 5:1–28.

Hannan, T. and Rhoades, S. (1987). Acquisition targets and motives: the case of the banking industry. *Review of Economics and Statistics*, 69(1):67–74.

Heckman, J. and Singer, B. (1984). A method for minimizing the impact of distributional assumption in econometric models for duration data. *Econometrica*, 52(2):271–320.

Heckman, J. J. and Honoré, B. E. (1989). The identifiability of the competing risks model. *Biometrika*, 76(2):325–330.

Hooks, L. M. (1995). Bank asset risk: Evidence from early warning models. *Contemporary Economic Policy*, 13(4):36–50.

Kalbfleisch, J. and Prentice, R. (2002). *The Statistical Analysis of Failure Time Data*. John Wiley & Sons, New York, second edition.

Lancaster, T. (1990). *The Econometric Analysis of Transition Data*. Cambridge University Press, Cambridge, UK.

Lane, W. R., Looney, S. W., and Wansley, J. W. (1986). An application of the Cox proportional hazards model to bank failure. *Journal of Banking and Finance*, 10:511–531.

Lillard, L. A. (1993). Simultaneous equations for hazard: marriage duration and fertility timing. *Journal of Econometrics*, 56:189–217.

McCall, B. P. (1996). Unemployment insurance rules, joblessness, and part-time work. *Econometrica*, 64(3):647–682.

Murata, K. and Hori, M. (2004). End of the convoy system and the surge of market discipline: Evidence from Japanese small financial institutions. Economic and Social Research Institute discussion papers No. 105.

Nyberg, M., Sellers, M., and Zhang, J. (2001). Private firm model: introduction to the modeling methodology. Technical Report, Moody's KMV.

Ohlson, J. (1980). Financial ratios and the probabilistic prediction of bankruptcy. *Journal of Accounting Research*, 18(1):109–131.

Richardson, D. (2002). Pearls monitoring system. World Council of Credit Unions Toolkit Series No. 4.

Shumway, T. (2001). Forecasting bankruptcy more accurately: a simple hazard model. *Journal of Business*, 74(1):101–124.

Sorensen, D. (2000). Characteristics of merging firms. *Journal of Economics and Business*, 52(5):423–433.

Sueyoshi, G. T. (1992). Semiparametric proportional hazards estimation of competing risks models with time-varying covariates. *Journal of Econometrics*, 51:25–58.

Tsiatis, A. (1975). A nonidentifiability aspect of the problem of competing risks. *Proceedings of the National Academy of Sciences*, 72(1):20–22.

van den Berg, G. J. (2001). Duration models: specification, identification and multiple durations. In Heckman, J. J. and Leamer, E., editors, *Handbook of Econometrics, Volume V*, chapter 55. North-Holland, Amsterdam.

van den Berg, G. J., van Lomwel, A. G. C., and van Ours, J. C. (2003). Nonparametric estimation of a dependent competing risks model for unemployment durations. IZA Discussion Paper Series No. 898.

Wheelock, D. C. and Wilson, P. W. (2000). Why do banks disappear? the determinants of U.S. bank failures and acquisitions. *Review of Economics and Statistics*, 82:127–138.

Worthington, A. C. (2004). Determinants of merger and acquisition activity in australian cooperative deposit-taking institutions. *Journal of Business Research*, 57:47–57.