RICE UNIVERSITY

Regression Survival Analysis With Dependent Censoring and a Change Point for the Hazard Rate: With Application to the Impact of the Gramm-Leach-Bliley Act to Insurance Companies' Survival

by

Nan (Jenny) Zhang

A Thesis Submitted
in Partial Fulfillment of the
Requirements for the Degree

Doctor of Philosophy

Approved, Thesis Committee:

Xuelin Huang, Co-Advisor
Associate Professor of Biostatistics at MD Anderson Cancer Center

Dennis D. Cox, Chairman and Co-Advisor
Professor of Statistics

James R. Thompson
Noah Harding Professor of Statistics

Wotao Yin
Assistant Professor of Computational and Applied Mathematics

Houston, Texas

December, 2009

Abstract

Regression Survival Analysis With Dependent Censoring and a Change Point for the Hazard Rate: With Application to the Impact of the Gramm-Leach-Bliley Act to Insurance Companies' Survival

by

Nan (Jenny) Zhang

This dissertation is aiming to find out the impact of the Gramm-Leach-Bliley Act on insurance companies' survival. The events of interest are bankruptcy and acquisition, which are correlated and censor each other. A statistical survival analysis method is developed first and then applied to the real insurance companies' survival data. In the methodology development, we first assess the effect of assuming independent censoring on the regression parameter estimates in Cox proportional hazard model. Then we apply the copula function to model the dependent censoring. Next, we propose an extended partial likelihood function maximized with an iteration algorithm to estimate the regression parameters and to derive the marginal survival functions

under a dependent censoring setting. Simulations are conducted to demonstrate the method's performance, and sensitivity analyses are performed to assess the impact of the dependent censoring on the regression parameter estimates. In the last part of methodology, we propose a method to test the existence and to identify the location of a change-point in a hazard function. The application of our methodology to real insurance companies' survival data discloses important influence of the GLB Act on insurance companies' survival.

Acknowledgements

Foremost, I would like to express my sincere gratitude to my advisor Dr. Xuelin Huang for his continuous support of my Ph.D study and research, for his patience, motivation, enthusiasm, and immense knowledge. His guidance helped me at all times in the research and writing of this thesis. I could not have imagined having a better advisor and mentor for my Ph.D study.

Besides my advisor, I would like to thank the rest of my thesis committee: Dr. Dennis D. Cox, Dr. James R. Thompson and Dr. Wotao Yin, for their encouragement, insightful comments, and questions.

My sincere thanks also goes to my husband, Shu Han, my parents, Guohua Zhang and Xinhua Wang, and my sister, Yi Zhang for their everlasting encouragement and support.

I thank my fellow classmates Hongxiao Zhu, Xiaowei Wu, and Bo Peng, and my colleague Erin English. They all offered sincere help on my thesis.

Contents

List of Figures

List of Tables

Chapter 1

Introduction

1.1 Overview

Insurance industry is one of the earliest industries that has been essentially relying on statistical mechanisms to harvest profit. The fundamental rules of statistics, such as the Law of Large Numbers, built the foundation for the industry's economics. In this research, we also utilize statistics to study the impact of insurance regulation, which has shaped the industry's competitive landscape and its practices since the inception of this industry.

Regulation has a long history in the insurance industry, and it reflects the everlasting government role in ensuring the integrity of the practitioners and protecting customers. Almost all the major changes in this industry were accompanied by the passage of major regulations. For instance, the Sherman Antitrust Act in 1890 put all

the insurance companies under the federal regulation, while the McCarran-Ferguson Act in 1945 demolished most federal regulations on insurance companies while re-affirming the governance of state regulation.

In 1999, the Gramm-Leach-Bliley (GLB) Act was enacted. It was expected to make a profound impact on the competitive landscape because it started a new era that allowed commercial banks and insurance companies to conduct each other's business. With the firewall between the two comprehensive business fields cracked down, it was expected that business integrations equipped with intensive capital flows would transform the two fields that would forever change the face of the insurance industry. Did it really happen? How significant is the effect? This curiosity motivated me to use comprehensive statistical techniques to pursue the answer. Statistics has been playing a key role in the insurance business. I hope this research can continue the legendary role of statistics in this industry.

Specifically, this research studies the changes in the survival of insurance companies before and after the passage of the GLB Act using a survival analysis framework. In order to effectively address the dependency issues between events and censoring, we utilize copula functions to model the dependent censoring. We also study whether there is a change-point in a hazard function and how to locate it.

1.2 Outline

Chapter 2 introduces the financial and statistical knowledge that underlines our research topic, and reviews recent research findings that may contribute to this research. Chapter 3 proposes our methodology to model the GLB Act's impact on insurance companies' survival, and conducts a simulation study to demonstrate the performance of the method. The analysis and results using real data are presented in Chapter 4. Built upon the findings in Chapters 3 and 4, Chapter 5 introduces the method to identify change-point(s) in hazard functions. Chapter 6 concludes this research and highlights possible future efforts.

Chapter 2

Background and Literature Review

In this chapter, we introduce the background and related history of regulations on United States' insurance industry that elicited our research topic. We also introduce concepts and methods in statistics that are used in our proposed method.

Section 2.1 presents the financial background and related literature review. Several statistical definitions and methods are given in Section 2.2. Sections 2.3 and 2.4 introduce survival analysis and copula function, respectively. The literature review of regression survival analysis is covered in Section 2.5.

2.1 Financial background and literature review

2.1.1 Financial background

Unlike most products, the value of an insurance product can not be determined until after claims are closed - when it is too late to decide whether a different insurer or a different product might be a better choice. For example, if an insurer goes bankrupt, the insured with unpaid claims will face a loss. The loss may even impact the insured's daily life if he or she does not have the financial capacity to pay the claim(s). In addition, insurers usually draft the insurance contract and decide the final premium. Insureds do not have much input about it. Such special aspects of insurance products add much uncertainty to the consumers.

In order to protect consumers, U.S. government started implementing regulation of insurance companies and agents in the States back to more than one hundred years ago. In 1869, there was an important U.S. Supreme Court case "Paul v. Virginia". The Supreme Court decided that the regulation of insurance should be through states. The use of bureau rates was encouraged, which led to rate-making in concert. As a result, local and regional bureaus were formed after 1877.

However, Congress passed the Sherman Antitrust Act in 1890 to eliminate rating bureaus. Similar Antitrust Acts were passed later on, such as Clayton Act, Federal Trade Commission and Robinson-Patman Acts. Meanwhile, there were voices from the opponents' side. For example, between 1910 and 1922, National Association of

Insurance Commissioners (NAIC) focused on the fire insurance rate reform. Merritt Committee was formed and provided persuasive rationale for an exception to the concept of anti-trust. The two opposite voices coexisted until the other important U.S. Supreme Court case "U.S. v. South-Eastern Underwriters Association (SEUA)" was closed in 1944. The Supreme Court decided that the Sherman Antitrust Act applied to insurance industry and bureau rate-making was not allowed.

However, State Governments did not like the idea since it impacted their revenue and taxation. They acted on it quickly and got the McCarran-Ferguson Act passed in 1945, which affirmed the state insurance regulation and taxation and exempted the business of insurance from federal antitrust laws with a few exceptions. The battle was continued for another fifteen years after McCarran-Ferguson, marked by judicial and regulatory activity to inhibit the use of bureau rates. For example, O'Mahoney Committee, formed in 1958, affirmed that competition should be the prime regulator of insurance.

The end of bureau rates happened in 1970's. Bureau rates became "advisory rates", served as a point of reference to insurers. The biggest bureau, Insurance Services Office (ISO) got challenged. In 1988, 20 states' attorneys general filed suit against ISO relating to policy language for Commercial General Liability. Finally in 1997, ISO became a "for profit" corporation. Insurance industry no longer controls ISO. To some extent, the new ISO was a symbol that the long transformation of insurance rating bureaus has reached a new level and probably will be stable for a

while. Figure 2.1 illustrates the evolution of insurance regulation.

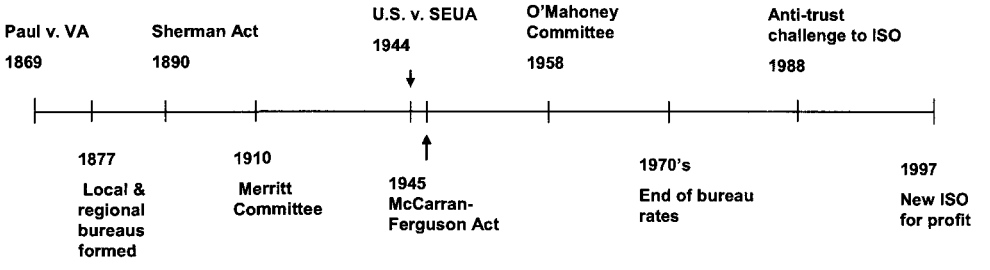

Figure 2.1: The Evolution of Insurance Regulation

Nowadays, the insurance industry is generally under state regulation with a few exceptions, although the supporters of federal insurance regulation have been continuing pushing for a change. Furthermore, competition starts playing a bigger role in the insurance regulation than it used to, although some states still have very restrictive insurance regulation.

Over time, the highly regulated environment became one of the main characteristics of the insurance industry. The function of the insurance regulation is two-fold. The government needs to ensure that insurers maintain a high level of solvency to meet the obligations of paying claims; and meanwhile stringently supervise the insurance sales, marketing, and policy terms to make sure that the consumers are treated fairly when they purchase insurance products and file claims.

Because of such strict regulatory environment of the insurance industry, government decisions can still cast a huge impact on insurance companies' practices in the twenty first century. For example, in 2006, in order to avoid consumer discrimina-

tion, California regulators started to not allow using zip codes as one of the factors to price car insurance any more. Given that territory is one of the most predictable independent variables in the premium prediction model, such regulation forced most insurers to reprice their car insurance products in California, and it added significant uncertainties for the car insurance companies' risk management.

The Gramm-Leach-Bliley Act in 1999

In 1999, U.S. government passed the GLB Act, and the purpose of such Act is to reduce the barriers among insurance companies, commercial banks and brokerage firms. Such barriers were set by several Acts throughout the history: In 1916, congressional passage of the National Banking Act prohibited national banks from selling insurance in any location where the population exceeded 5,000 inhabitants. In 1934, as a consequence of the Great Depression, the passage of the Glass-Steagall Act further separated the banking and insurance industry by stringently segmenting the whole financial industry into separate and unique insurance, banking, and investment sectors. In the following years, passages of additional legislation such as the Bank Holding Company Act and the Garn-St. Germain Act created a more awkward system of regulation among the sectors.

In the past two decades, these regulatory barriers restricting financial integration have been challenged in the courts and in the legislature. In 1998, the 105th Congress nearly succeeded in repealing Glass-Steagall Act when the House narrowly passed HR

10 while the Senate was unable to negotiate a compromise before the session ended. In the following year, 1999, both the House and Senate reached an agreement and passed the Gramm-Leach-Bliley Act. The passage of GLB allows the creation of financial holding companies to underwrite and sell both insurance and securities, to engage in commercial and merchant banking, and to develop real estate through subsidiaries. It also expedited the review of conflicts between state and federal regulators regarding insurance issues. State governments, however, remain as the functional regulators of insurance activities.

As expected, the passage of GLB led to a significant impact on many insurers' practices and thereby affected the market landscape of the insurance industry. There are many Sections within the GLB. This research focuses on the Sections 302 and 303 as shown in Table 2.1. With observations of U.S. insurance companies' major activities such as significant market expanding, merger, or bankruptcy, we are mainly interested in studying whether the GLB Act carries crucial effects to the survival of insurance firms.

Table 2.1: Gramm-Leach-Bliley Act (1999): TITLE III - INSURANCE

Subtitle A - State Regulation of Insurance
Sec. 302. Insurance underwriting in national banks.
Sec. 303. Title insurance activities of national banks and their affiliates.

2.1.2 Literature review

Previous research indicates that many regulations have made significant impacts on the value of financial services firms including insurers. Although not all the findings on such value changes are in the same direction, such changes surely affect the business activities of the firms. Cornett and Tehranian (1989) found that the passage of the Depository Institutions Deregulation and Monetary Control Act (DIDMCA) of 1980 had positive wealth effects for large commercial banks and a negative impact on savings and loans. Carow and Heron (1998) found that the passage of the Interstate Banking and Branch Efficiency Act (IBBEA) of 1994 resulted in positive wealth effects for large bank holding companies. Amoako-Adu and Smith (1995) concluded that financial services deregulation in Canada between 1984 and 1991 had a positive impact on insurance firms. Pacini and Marlett (2001) resulted that the legislative creation of the Florida Hurricane Catastrophe Fund had negative wealth effects for property-liability insurers.

Akhigbe and Whyte (2001) found that positive valuation effects for banks, brokerage firms, insurance companies, and all the financial institutions combined due to the passage of GLB; while Carow and Heron (2002) found negative returns for foreign banks, thrifts and finance companies, and found insignificant returns for banks and positive returns for investment banks and insurance companies.

Marlett et al. (2003)

Marlett et al. (2003) also studies the impact of GLB. However, it focuses on the different impact on life insurers and on property-liability insurers; and the market impact rather than the impact to insurers' survival rates.

The four hypotheses (H1 - H4) in this paper focus on different topics. H1 deals with returns; H2 focuses on trading volumes; H3 compares the abnormal returns of life insurers and property-liability insurers; and H4 examines any asymmetrical effects of the GLB. The specific hypotheses are:

H1: "The abnormal returns of insurers during the legislative enactment process of the GLB were not significantly different from zero."

H2: "The trading volume of insurers on legislative announcement days involving the GLB was not significantly different from the trading volume on non-announcement days."

H3: "The abnormal returns of life insurers were not significantly different from those of property-liability insurers on GLB announcement days."

H4: "The GLB legislative enactment process had no differential effect on the abnormal returns of insurers possessing different firm-specific characteristics."

For hypothesis one, a generalized least squares (GLS) portfolio approach and a nonparametric technique, Corrado's rank statistic (Corrado, 1989), are used. The dependent variable is the equally-weighted portfolio return for day t. Independent variables include the market return for day t, a dummy variable indicating a life insurer or not, and another indicator indicating the day t for the j^{th} event day, denoted

as D_{jt}. D_{jt} is the focus of hypothesis one. In the application of the Corrado's rank statistic, H1 is rejected. That is, the GLB legislative event is associated with various positive and negative share price reactions.

One-tail t-tests are used to assess hypothesis two. As a result, the H2 is rejected. That is, there is a raise in the trading volume due to the GLB announcement.

Based upon a t-test, H3 is rejected as well. In other words, GLB event disclosures caused significantly different share price reactions for both life and property-liability insurers.

For hypothesis four, a GLS cross-sectional rank regression model is used. The independent variables include the standardized rank of firms' size, liquidity, premiums-written-to-surplus ratio and variance of abnormal returns. The results show that the GLB has a greater impact on life insurers than on property-liability insurers. Therefore, H4 is rejected.

In summary, smaller life insurers with high liquidity and more leverage seemed to have the most positive share price reactions.

Mamun et al. (2004)

Mamun et al. (2004) examines the impact of GLB Act across three main sectors of the financial services industry: commercial banks, insurance companies and brokerage firms, taking into account the wealth effect associated with the announcement.

The four hypotheses tested in this paper are:

H1: "The GLB Act creates value for all sectors of the financial services industry."

H2: "The banking industry gains the most from the passage of the GLB Act."

H3: "The GLB Act reduces exposures to systematic risk across the industry."

H4: "The GLB Act is a de facto large-firm law."

Seemingly Unrelated Regressions methodology (Zellner, 1962) is used to assess the stock price reaction. To perform a cross-sectional analysis, models for banking, insurance and brokerage firms are established, respectively.

The results show that all three sectors of the financial services industry have gained from this law, and when normalized for the asset base it turns out that the banking industry benefits most among the three sectors, followed by the insurance industry (H1 & H2). Mamun et al. (2004) also finds out that the GLB Act creates diversification opportunities for the financial services industry and hence appears to reduce exposures to systematic risk (H3). Furthermore, larger firms benefit more in the banking and the insurance industries (H4).

The current literatures do not particularly study the GLB impact to insurance firms' survival, which is the focus of our research.

2.2 Basic statistical concepts

2.2.1 Indicator function

An indicator function is defined as in Casela and Berger (2002):

Definition 1. *The indicator function of a set A, most often denoted by $I_A(x)$, is the function*

$$I_A(x) = \begin{cases} 1 & , \quad x \in A, \\ 0 & , \quad x \notin A. \end{cases}$$

An alternative notation is $I(x \in A)$.

In Section 3.3.3 we explain how the proposed method has a different indicator function than the traditional definition.

2.2.2 Likelihood function

A likelihood function is defined as in Casela and Berger (2002):

Definition 2. *Let $f(x|\theta)$ denote the joint probability density function of the sample $X = (X_1, ..., X_n)$. Then, given that $X = x$ is observed, the function of θ defined by*

$$L(\theta|x) = f(x|\theta) \tag{2.1}$$

is called the likelihood function.

In this research, we maximize the extended partial likelihood function to get the optimal parameter estimators.

2.2.3 Bootstrapping method

Bootstrapping is a re-sampling technique. We retake samples from the original sample with replacement while following an approximating distribution. One standard

choice for an approximating distribution is the empirical distribution of the observed data. Bootstrapping can be used to calculate standard errors of parameter estimates. Furthermore, confidence intervals of parameter estimates can be obtained. When the original sample size is small, the bootstrapping method can be used to create more samples in order to infer to the population. The bootstrap was developed by Efron in the late 1970s.

In this research, the bootstrapping method is used to calculate the covariance matrices of parameter estimates. See details in Section 3.4.

2.2.4 Sensitivity analysis

Sensitivity analysis is used to determine how "sensitive" a model is to changes in the value of the parameters of the model or to changes in the structure of the model. In this research, we conducted a sensitivity analysis on parameter sensitivity.

Parameter sensitivity is usually performed as a series of tests in which the modeler sets different parameter values to see how a change in the parameter causes a change in the dynamic behavior of the models' outputs. By showing how the model behavior responds to changes in parameter values, sensitivity analysis is a useful tool in model building as well as in model evaluation. Sensitivity analysis helps build confidence in the model by studying the uncertainties that are often associated with parameters in models. See further details about how sensitivity analysis is used in this research in Section 3.1.2.

2.2.5 Newton's method

Newton's method (also known as the Newton-Raphson method) is one of the classic numeric analysis tools, named after Isaac Newton and Joseph Raphson. It is perhaps the best known method for finding successively better approximations to the roots of a real-valued function. Newton's method can often converge remarkably quickly, especially if the iteration begins "sufficiently near" the desired root.

Given a function f(x) and its derivative $f'(x)$, we start from a first guess x_0. A better approximation x_1 is

$$x_1 = x_0 - \frac{f(x_0)}{f'(x_0)} \tag{2.2}$$

The iteration process of the method is as follows: one starts with an initial guess which is reasonably close to the true root, then the function is approximated by its tangent line, and one computes the x-intercept of this tangent line. This x-intercept will typically be a better approximation to the function's root than the original guess. This process iterates until the new x-intercept converges to a value, which would be the best estimate of the root. Figure 2.2 is an illustration of one iteration of Newton's method.

By definition, the derivative at a given point is the slope of a tangent at that point. Furthermore, the Equation 2.2 can be extended to later iterations. That is, for n = 0, 1, 2, ..., we have

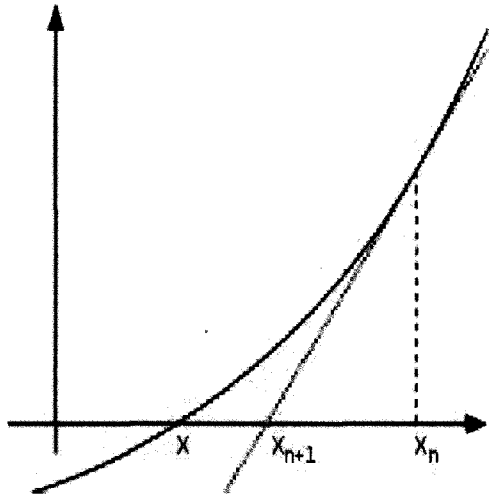

Figure 2.2: An Illustration of One Iteration of Newton's Method

The function f is shown in blue and the tangent line is in red. We see that x_{n+1} is a better approximation than x_n for the root x of the function f.

$$x_{n+1} = x_n - \frac{f(x_n)}{f'(x_n)}, \tag{2.3}$$

$$f'(x_n) = \frac{\Delta y}{\Delta x} = \frac{f(x_n) - 0}{x_n - x_{n+1}}. \tag{2.4}$$

The method will usually converge, provided the initial guess is close enough to the true root, and that $f'(x_0) \neq 0$.

Newton's method is used in this research to maximize the extended partial likelihood function in order to get optimal parameter estimates. See details in Section 3.3.4.

2.3 Introduction to survival analysis

According to the classical definition (Miller, 1981), survival analysis focuses on the time until the event of interest occurs. The event of interest should be defined to serve specific research topics. For example, when studying clinical trials, the event of interest can be defined as the death of a patient if the purpose is to analyze patients' survival; or it can be defined as a disease incidence if the purpose is to look into disease occurrences.

Depending on the research objective, the event of interest could be defined broadly. It could be the time to learn a skill, or it may not even be a time at all. For instance, it could be the number of dollars that an insurance company pays to a particular claim. This characteristic can be particularly useful. It allows us to apply the survival analysis techniques to other fields, with much broader applications.

2.3.1 Censoring

Censoring is a special feature under the survival analysis setting. When there is an incomplete observation of the failure time, we call it a **censoring**. For example, a patient lost to follow up or withdraw from a clinical trial. Censoring can be categorized in different ways.

Type I censoring occurs when a study ends when a certain time point is reached. **Type II censoring** occurs when a study ends when a certain number of failures occurs. In both cases, the event of interest is not observed.

Right censoring occurs when the exact survival time becomes incomplete at the right side of the follow-up period. For example, if a patient is still alive at the end of a study. **Left censoring** occurs when the survival time becomes incomplete at the left side of the follow-up period. For instance, when the event of interest occurs before this particular observation enters a study and the exact time is unknown.

Independent censoring, also called noninformative censoring, means the censoring is independent with the event of interest. For example, administrative censoring is one kind of independent censoring, e.g. a subject has not yet had an event due to the administrative termination of a study. By contrast, **dependent censoring**, also called informative censoring, means the censoring is dependent on the event.

If there is more than one risk that can trigger the event of interest, we call those risks **competing risks**. When one risk happens, it will prevent the other risk from happening.

2.3.2 Basic functions and curves

In order to model the time until the event occurs, several basic functions are introduced:

1. Probability density function (pdf),

$$f(t) = \lim_{\Delta t \to 0} \frac{1}{\Delta t} P\left[t \leq T < t + \Delta t\right]. \tag{2.5}$$

A probability density function indicates the instantaneous death rate at time t.

2. Cumulative distribution function (cdf),

$$F(t) = P[T \leq t] = \int_0^t f(s)ds. \tag{2.6}$$

A cumulative distribution function is a cumulative function of pdf.

3. Survival function,

$$S(t) = P[T > t] = 1 - F(t) = \int_t^\infty f(s)ds. \tag{2.7}$$

A survival function S(t) indicates the probability that a person survives longer than some specified time t. The range of a survival function is from 0 to 1. A survival function is a monotone function by time.

4. Hazard function,

$$\lambda(t) = \lim_{\Delta t \to 0} \frac{1}{\Delta t} P[t \leq T \leq t + \Delta t | T \geq t] = \frac{f(t)}{S(t)}. \tag{2.8}$$

A hazard function or hazard rate indicates instantaneous death rate at time t, given alive up to t.

5. Cumulative hazard function,

$$\Lambda(t) = \int_0^t \lambda(s)ds = -ln(S(t)). \tag{2.9}$$

The summary of the relationships among functions 1- 5 is shown below:

$$F(t) = \int_0^t f(s)ds = 1 - S(t); \tag{2.10}$$

$$S(t) = 1 - F(t) = \int_t^\infty f(s)ds = e^{-\int_0^t \lambda(s)ds} = e^{-\Lambda(t)}; \tag{2.11}$$

$$\lambda(t) = \frac{f(t)}{S(t)} = \frac{f(t)}{1 - F(t)} = \frac{\frac{dF(t)}{dt}}{1 - F(t)} = -\frac{\frac{dS(t)}{dt}}{S(t)}; \tag{2.12}$$

$$\Lambda(t) = \int_0^t \lambda(t)ds = -ln(S(t)) = -ln(1 - F(t)); \tag{2.13}$$

$$f(t) = \frac{dF(t)}{dt} = -\frac{dS(t)}{dt} = \lambda(t)S(t) = \lambda(t)(1 - F(t)); \tag{2.14}$$

In Equation 2.10, $F(t)$ is the probability that a person did not survive longer than time t. That is, it is the complement of $S(t)$ by definition.

Note that any one of $f(t)$, $F(t)$, $\lambda(t)$ or $\Lambda(t)$ is enough to specify the survival function, because knowing any one, you can calculate the other three functions. Figure 2.3 shows the pdf, hazard function, survival function and cdf for a Weibull distribution with the scale parameter $\rho = 1$ and the shape parameter $\gamma = 0.5$.

2.3.3 Kaplan-Meier estimator

The Kaplan-Meier (K-M) estimator is a typical way of estimating a survival function. It is defined below:

Let $t_0 = 0$,

$$\hat{S}(t) = \prod_{i:t_i \leq t} (1 - \frac{D_i}{N_i}) = \prod_{i:t_i \leq t} \frac{S_i}{N_i} \tag{2.15}$$

where $i = 0, 1, ..., n$; D_i indicates the number of death at time t_i; S_i indicates the number of survival at time t_i; N_i indicates the number of objects at risk at time t_i.

The survival function calculated by the K-M estimator is a step function. The K-M estimator is popular due to the simplicity of its calculation and its ease of interpretation.

2.3.4 Cox regression model

Cox regression model is also called the Relative Risk model. It is a hazard function including covariates, denoted as $Z_1, ..., Z_k$.

$$\begin{aligned} \lambda(t|Z_1, ..., Z_k) &= \lambda_0(t)e^{\beta_1 Z_1 + ... + \beta_k Z_k}; \\ &= \lambda_0(t)exp(Z'\beta); \end{aligned} \tag{2.16}$$

Partial likelihood function:

$$\begin{aligned} L(\beta) &= \prod_{i=1}^{n} \frac{\lambda_0(t)exp(Z'\beta)}{\sum_{j \in R(t_i)} \lambda_0(t)exp(Z_j'\beta)} \\ &= \prod_{i=1}^{n} \frac{exp(Z'\beta)}{\sum_{j \in R(t_i)} exp(Z_j'\beta)} \\ &= \prod_{i=1}^{n} \frac{exp(Z'\beta)}{[\sum_{k=1}^{n} exp(Z_k'\beta)]^{\delta_k}}; \end{aligned} \tag{2.17}$$

where δ_k is an indicator function. It indicates the number of failures at time k (with ties).

As shown above, $\lambda_0(t)$ cancels in the numerator and denominator. We extended the Cox partial likelihood function in Section 3.3.3.

2.3.5 Breslow's method

Breslow's method can be used to estimate the baseline cumulative hazard function. The Breslow's estimator is given by

$$\hat{\Lambda}_0(t) = \sum_{t_i<t} \hat{\lambda}_0(t_i) = \sum_{t_i<t} \frac{1}{\sum_{j \in R_i} e^{\beta x_j}} \tag{2.18}$$

This method is further utilized in Chapter 3 Equations 3.13 and 3.14.

2.4 Introduction to copulas

Copulas are "multivariate distribution functions whose one-dimensional margins are uniform on the interval $(0, 1)$" (Nelsen 2006). Copulas are frequently utilized to model the dependent structure among variables, and have broad applications in survival analysis and actuarial science. Especially since Li (2000) first introduced copulas into the modeling of default risks, researchers in the finance field have shown increasing interest in applying copulas to address extensive financial topics.

Basically, copulas provide a convenient way to model two or more dependent variables in terms that it does not require knowing explicitly either the function form of the variables' joint distribution or their marginal distributions. For any continuous variable, the cumulative distribution function follows a uniform distribution, thus

copulas are created based upon uniform distributed variables. Copulas naturally have wide applications in statistical modeling.

2.4.1 Different copula functions

In this research, we employ a commonly used two-dimensional copula, Frank (1979) copula. It is popular because unlike other copulas, Frank copula can model the full range of association, $\tau \in (-1, 1) \setminus \{0\}$. Frank copula is defined as below,

- Frank (1979) copula

$$H(u, v; \alpha) = \log_\alpha\{1 + \frac{(\alpha^u - 1)(\alpha^v - 1)}{\alpha - 1}\}, \alpha > 0, \alpha \neq 1, \tag{2.19}$$

where u and v represent known uniform variates within the range of 0 and 1. H represents a copula function. And α is the parameter of a copula.

Several other popular copulas are:

- Independent copula

$$H(u, v) = uv \tag{2.20}$$

- Clayton (1978) copula

$$H(u, v; \alpha) = (u^{-\alpha} + v^{-\alpha} - 1)^{-1/\alpha}, \alpha > 0. \tag{2.21}$$

- Gumbel-Hougaard copula (Gumbel, 1961; Hougaard, 1986)

$$H(u,v;\alpha) = exp[-\{(-\log u)^{\alpha} + (-\log v)^{\alpha}\}^{1/\alpha}], \alpha \geq 1. \tag{2.22}$$

2.4.2 Kendall's τ

In order to gauge the degree of association between X and Y, we use a measure known as Kendall's τ to measure the degree of association between variables X and Y. τ is defined as the probability of *concordance* minus the probability of *discordance*, as shown below:

$$\begin{aligned} \tau_{X,Y} &= Pr[(X_1 - X_2)(Y_1 - Y_2) > 0] - Pr[(X_1 - X_2)(Y_1 - Y_2) < 0] \\ &= Pr[(X_1 > X_2, Y_1 > Y_2) \text{ or } (X_1 < X_2, Y_1 < Y_2)] \\ &\quad -Pr[(X_1 > X_2, Y_1 < Y_2) \text{ or } (X_1 < X_2, Y_1 > Y_2)] \end{aligned} \tag{2.23}$$

Let X and Y be continuous random variables. Then we have,

$$\tau_{X,Y} = \tau_H = Q(H,H) = 4 \int \int_{I^2} H(u,v;\alpha) dH(u,v;\alpha) - 1 \tag{2.24}$$

The range of τ is $(-1,1) \setminus \{0\}$. -1 means a perfect negative correlation and 1 means a perfect positive correlation. Kendall's τ is invariant under strictly increasing transformations of the underlying random variables. Therefore, τ is independent from marginal distributions, as the Equation 2.24 shows.

2.4.3 Transformation between τ and α

The transformation relation between Kendall's τ and the parameter of a copula function α can be calculated. For the Frank copula, we have

$$\tau = 1 + 4(-\log\alpha)^{-1}\{D_1(-\log\alpha) - 1\}, \alpha > 0, \alpha \neq 1. \tag{2.25}$$

where D_k, the Debye function, is defined below:

$$D_k(x) = \frac{k}{x^k}\int_0^x \frac{t^k}{e^t - 1}dt. \tag{2.26}$$

Based upon Equations 2.25 and 2.26, we can transform the value of τ and the value of α, as Table 2.2 shows.

Table 2.2: Transformation Between τ and α.

	τ	α
1	-0.5	309.91
2	0.2	0.15554
3	0.5	0.003215
4	0.8	0.00000001258

2.5 Literature review of the dependent censoring and sensitivity analysis

The coexistence of dependent and independent censoring has been investigated by various researchers. Scharfstein et al (2001), Scharfstein and Robins (2002), and Tsiatis (1975) have shown that extra information regarding the censoring must be collected in order to precisely measure the impact due to dependent (informative) censoring.

There is rich literature on sensitivity analysis. In the setting of survival analysis, see, for example, Slud (1992) and references therein. Troxel, Ma, and Heitjan (2004) proposed an index of local sensitivity to non-ignorability. Zhang and Heitjan (2006), Siannis, Copas, and Lu (2005) and Siannis (2004) used parametric survival models to do sensitivity analysis.

However, none of these methods can be used to do a sensitivity analysis for the most widely used Cox (1972) proportional hazards models. Park, Tian and Wei (2006) proposed a sensitivity analysis method in the nonparametric setting without covariates. But the method can hardly be extended to general regression problems because it is incapable of identifying values of the regression parameters with dependent censoring under a nonparametric setting.

The latest work in this area was by Peng and Fine (2007). They did regression analysis for semi-competing risk data. It is not clear whether their estimation method

can be extended to the competing risk framework.

Zheng and Klein (1995)

Zheng and Klein (1995) show that when the copula function is known, the competing risks data is sufficient to identify the marginal survival functions, and construct a suitable and consistent estimator. When the event and censoring times are independent, the proposed estimator reduces to the Kaplan-Meier estimator.

Specifically, in a competing risks framework, let X be the time until an event of interest occurs, Y be the time until a competing risk occurs. That is, X can not be observed if a competing risk happens. Assume the copula of X and Y is known and a competing risk sample is denoted as (T, δ), where $T = min(X, Y)$ and $\delta = I(X \leq Y)$. I(A) is the indicator function of the set A.

Zheng and Klein (1995) prove several important theorems, which enable us to further utilize their proposed method.

Theorem 1. *"Suppose the marginal distribution functions of (X,Y) are continuous and strictly increasing in* $(0, \infty)$. *Suppose the copula, C, of (X,Y), is known, and* $\mu_c(E) > 0$ *for any open set E in [0,1]×[0,1]. Then F and G, the marginal distribution functions of X and Y, are uniquely determined by* $\{k(t), p_1(t), p_2(t), t > 0\}$."

Theorem 2. *"Suppose that two marginal distribution functions F, G, are continuous and strictly increasing on* $(0, \infty)$, *and the assumed copula has density function* $u(x, y) > 0$ *on [0,1]×[0,1]. Then* $\hat{F}_n$ *and* $\hat{G}_n$ *are strongly consistent for F and G. That*

is with probability 1 as $n \to \infty$, $\hat{F}_n(t) \to F(t)$ *and* $\hat{G}_n(t) \to G(t)$ *for all* $t \in [o, \infty)$."

Theorem 3. *"The copula-graphic estimator is a maximum likelihood estimator."*

Theorem 4. *"For the independence copula C(x,y)=xy, when* $t \leq t_n$, *the largest observed time, the copula-graphic estimates of marginal survival functions are exactly the Kaplan-Meier estimates."*

Note: Zheng and Klein (1995) was written before Zheng and Klein (1994), although it was published later than Zheng and Klein (1994).

Zheng and Klein (1994)

Zheng and Klein (1994) apply the copula method to construct an estimator of the marginal survival function based on dependent competing risk data.

In a competing risks framework, we define X, Y, T and δ same as in Zheng and Klein (1995). Zheng and Klein (1994) show that the marginal survival function can be estimated:

$$\hat{S}(t) = n^{-1}\left\{\sum_{i=1}^{n} I\left[t_i \geq t\right] + \sum_{t_i<t}(1-\delta_i)\hat{P}\left[X > t | X > t_i, Y = t_i\right]\right\} \quad (2.27)$$

$$\hat{R}(t) = n^{-1}\left\{\sum_{i=1}^{n} I\left[t_i \geq t\right] + \sum_{t_i<t}\delta_i\hat{P}\left[Y > t | Y > t_i, X = t_i\right]\right\}. \quad (2.28)$$

Here, both $\hat{S}(t)$ and $\hat{R}(t)$ are self-consistent estimators. When X and Y are dependent with a known copula $C(u, v)$, we have:

$$\hat{P}\left[X > t | X > t_i, Y = t_i\right] = \frac{1 - C_v\left[1 - \hat{S}(t), 1 - \hat{R}(t_i)\right]}{1 - C_v\left[1 - \hat{S}(t_i), 1 - \hat{R}(t_i)\right]}; \quad (2.29)$$

$$\hat{P}\left[Y > t | Y > t_i, X = t_i\right] = \frac{1 - C_u\left[1 - \hat{S}(t_i), 1 - \hat{R}(t)\right]}{1 - C_u\left[1 - \hat{S}(t_i), 1 - \hat{R}(t_i)\right]}; \quad (2.30)$$

where $C_u(a,b) = \frac{\partial C(u,v)}{\partial u}$ and $C_v(a,b) = \frac{\partial C(u,v)}{\partial v}$, evaluated at the point $(u,v) = (a,b)$. u and v represent uniform variates within the range of 0 and 1.

Zheng and Klein (1994) also conduct Monte Carlo simulation studies, which show that the self-consistent estimators $\hat{S}(t)$ and $\hat{R}(t)$ are reasonably robust to model misspecification. Zheng and Klein (1994)'s method is used in this research. See Section 3.3.2 for details.

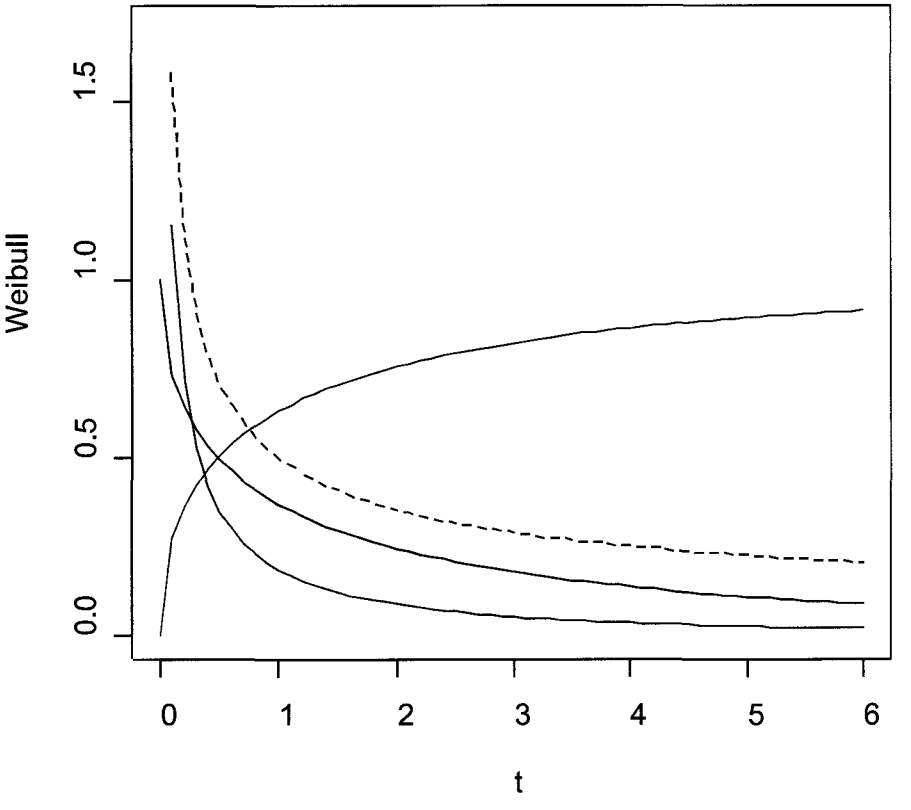

Figure 2.3: Functions for the Weibull Distribution with $\rho = 1$ and $\gamma = 0.5$

Blue line, PDF ; Red line, survival function; Green line, CDF; Dashed line, hazard function.

Chapter 3

Regression Survival Analysis with Dependent Censoring

In this chapter, we develop the methodology to study the GLB Act's impact both to the survival of insurance companies and to the acquisition transactions among those companies. Under a framework of regression survival analysis, the proposed method models both independent and dependent censoring, and takes covariates into consideration.

Section 3.1 introduces the definition of events and censoring with our research objective, and illustrates common concerns over dependent censoring and proposes a solution. Section 3.2 explains the advantages of copula approach and describes the selection process among candidate copula functions. Section 3.3 integrates the copula approach into survival analysis. Section 3.4 presents a simulation study demonstrating

the performance of the proposed method.

3.1 Dependent censoring and sensitivity analysis

3.1.1 Dependent censoring

Censoring is a particular issue requiring additional attention in survival analysis. A common assumption made in survival analysis is that the censoring is independent of the event of interest. This assumption may hold for the non-informative (independent) censoring, but it hardly makes sense for the informative (dependent) censoring.

The time horizon we considered for this research is from 1994 to 2005, which covers 6 years before the passage of the GLB Act and 6 years after. The data used in this research was originally collected in 2006 and was updated in 2007.

When analyzing the GLB Act's impact on the survival of insurance companies, we define an **event** as an insurance company filed bankruptcy, a **dependent censoring** as an insurance company got acquired (identified as being involved in a Mergers and Acquisitions (M&A) transaction and abandoning its original company name subsequently), and an **independent censoring** as otherwise. The reason for such definition of dependent censoring is, when the event is defined as filing bankruptcy, being acquired (giving up one's name in an M&A transaction) can be reasonably regarded as being dependent on the event of interest. In our real data set, the majority of the M&A cases is companies being acquired. Thus, in the application, we focus on

acquisitions rather than both mergers and acquisitions.

When studying acquisitions among insurance companies, we define an **event** as an insurance company got acquired, a **dependent censoring** as an insurance company filed bankruptcy, and an **independent censoring** as otherwise.

Due to the existence of both dependent and independent censoring, conventional survival analysis approaches, which only address the independent censoring, are no longer adequate for this research. The root reasons for such inadequacy are twofold: First, the degree and direction of the correlation between events and censoring (dependent censoring) can lead to biased estimations of survival rates. Specifically, the survival rates will be overestimated under the conventional approach when there is a positive correlation between events and censoring, and vice versa. Second, the percentage of observations being dependently censored will also affect the magnitude of the bias. Unless there is only a negligible amount of dependently censored data, the dependent censoring tends to have sizeable influence on the estimation bias.

3.1.2 Sensitivity analysis

The coexistence of dependent and independent censoring can be solved if the additional data regarding the dependent censoring can be obtained. However, in practice, more often than not, such data is unavailable. For example, as far as mergers and acquisitions are concerned, many of the genuine motivations behind the deals may only reside in companies' internal documents that are strictly confidential to outsiders.

Thus, without additional data, dependent censoring is not identifiable (Tsiatis, 1975). For instance, it is reasonable to believe that a company being acquired implies that it has limited ability to survive on its own. But the degree of such dependency is unclear.

One feasible way that researchers normally take under these circumstances is to conduct a sensitivity analysis, i.e., to evaluate model outputs under various scenarios formed by reasonable contemplations. In this research, we also take such approach to assess the impact of dependent censoring on the performance of the proposed model and to help answer questions such as,

1. Would some covariates become non-significant when dependent censoring is taken into account?

2. Would the parameter estimates be affected when the correlation between the dependent censoring and events changes?

Although there has been a wide range of sensitivity analysis methods being studied, we found many of them are difficult to apply to our research. It was Zheng and Klein's (1995) method that inspired us to utilize the copula approach to model the structure of dependent censoring. We extended their method to the sensitivity analysis under a general setting of the well-known Cox proportional hazards model.

3.2 Why copulas and which copulas?

3.2.1 Advantages of the copula approach

Among many approaches in the sensitivity analysis, the copula approach is particularly appealing to our research. One of the advantages is that by using copula functions, the joint distribution of two variables can be obtained without the requirement of knowledge of the marginal distributions. Such attribute is naturally suitable in estimating the joint distribution of events and dependent censoring, a critical step to derive the conditional probability of survival (See Section 3.3.2 for details).

Another advantage of copulas that benefits this research is their flexibilities. Extensive copula functions have been well studied (Nelson, 2006, and see Section 2.3 for details). The wide range of copula functions serves versatile needs on modeling joint distributions of events and dependent censoring. In addition, some fundamental research of applying copula functions to study financial data has been published recently (Li, 2000; Cherubini et al, 2004), which provides further insights into the ways of utilizing flexibilities of copulas.

3.2.2 Which copula function to use?

Previous studies demonstrate that, in contrast to choosing an incorrect copula function, falsely assuming the value of Kendall's τ leads to a much more severe bias in parameter estimates. For instance, setting τ's value as zero (assuming independent

censoring) under the circumstances of dependent censoring will lead to more bias than applying a different choice of copula does. Grethen's Ph.D. dissertation (2004) further concludes that, when evaluating the impact on the parameter estimation bias, the choices on the value of τ dominate the choices on copula functions. In other words, identifying the correct value of τ will essentially minimize the parameter estimation bias irrespective of the copula function used.

Based on such findings, we conveniently selected the Frank copula to model the data of events and dependent censoring. The main advantage of using Frank copula is that it enables Kendall's τ to take any value between -1 and 1 that covers the whole range of all possible associations.

3.3 Methodology - extended Cox proportional hazards models

3.3.1 Event time and dependent censoring time models

We denote the event time as T, the informative censoring time as C, and the non-informative censoring time as S. The observed survival time is denoted as $X = min(T, C, S)$. That is to say, only what happens first will be observed. We also define two indicator functions $\delta_1 = I(X = T)$ and $\delta_2 = I(X = C)$. The covariates Z and W are associated with events and informative censoring, with dimensions $p \times n$ and $q \times n$ respectively. They may be identical, overlapped, or completely distinct.

In our research, we simply assume that Z_i and W_i have the same value for the i^{th} subjects, $i = 1, ..., n$. The proposed method also applies when Z_i and W_i have different values.

Note that, the observation of T (or C) means that C (or T) will not be observed in the short term, and vice versa. Thus, mathematically, T and C can be treated as competing risks. Zheng and Klein (1995) shows that the competing risks are sufficient to identify the marginal survival functions and construct a suitable estimator. Here, we developed two sets of equations for T_i and C_i, which will be used to study insurance companies' "bankruptcy" and "acquisition", respectively, in both Chapters 4 and 6.

Since the proposed method is an extension of the classic Cox model, the basic marginal functions from the Cox model are adopted, as shown below:

- Hazard functions for the event time T_i and the dependent censoring time C_i are, respectively, assumed to be:

$$\lambda_i(t|Z_i, W_i) = \lambda_0(t)\exp(Z_i'\beta), \tag{3.1}$$

$$\psi_i(t|Z_i, W_i) = \psi_0(t)\exp(W_i'\beta_c), \tag{3.2}$$

where $\lambda_0(t)$ and $\psi_0(t)$ are unspecified baseline hazard functions; and β and β_c are unknown parameters with respective dimensions $p \times 1$ and $q \times 1$. Note that although only Z_i or W_i appears on the right side of the equation, both Z_i and W_i contribute to $\hat{\beta}$ and $\hat{\beta}_c$, estimates of β and β_c. Therefore, it is more accurate to write the hazard functions given both Z_i and W_i.

- Marginal cumulative distribution functions for T_i and C_i are:

$$\begin{aligned} F_i(t|Z_i, W_i) &= 1 - S_i(t|Z_i, W_i) \\ &= 1 - \exp\{-\Lambda_0(t)\exp(Z_i'\beta)\}, \\ G_i(t|Z_i, W_i) &= 1 - R_i(t|Z_i, W_i) \\ &= 1 - \exp\{-\Psi_0(t)\exp(W_i'\beta_c)\}, \end{aligned} \tag{3.3}$$

where $S_i(t)$ and $R_i(t)$ represent survival functions, and denote cumulative baseline hazard functions by $\Lambda_0(t)$ and $\Psi_0(t)$, respectively.

We will use $F_i(t)$ and $G_i(t)$ to denote above marginal cumulative distribution functions. For a given Copula H with parameter α, the joint cumulative distribution function of T_i and C_i can be modeled as follows:

$$\begin{aligned} J_i(t, c) &= Pr(T_i \leq t, C_i \leq c) \\ &= H\{F_i(t), G_i(c); \alpha\} \end{aligned} \tag{3.4}$$

The identifiability of the parameters in the above model has been shown by Heckman and Honore (1989) and several corresponding estimation methods have been proposed. However, those methods only work for some special cases. According to Peng and Fine (2007), a general estimation method is not yet seen in the literature.

3.3.2 Conditional survival probability functions

To fit the joint model, we extended the idea of "redistribution of mass" by Efron (1967) to deal with the dependent censoring problem. Before the occurrence of the

event, we assume each subject has unit mass, while after the occurrence, we assume each subject has zero mass. Next step before fitting the model is to sort all the observed time points in ascending order and let the smallest be on the far left. Assuming independent censoring, the mass of a censored subject is redistributed uniformly to all the event time points on its right. Applying this procedure to all (independently) censored subjects from left to right, the mass should be distributed on event time points only. Such procedure resulted in the Kaplan-Meier estimator.

Zheng and Klein's research (1994) applies this idea of "redistribution of mass" to obtain self-consistent estimators for the marginal distribution functions and modeled dependent competing risks under an assumption of a copula-based conditional survival probability function. Unfortunately, under the constraint of a copula function for the joint distribution, the redistribution of mass to the right is no longer uniform for dependent censoring. Here we further extend Zheng and Klein's method to the Cox proportional hazards model. For a dependently censored subject, we show below how its mass is redistributed to the right.

Assume that x_i, $i = 1, ..., n$, are sorted time points in ascending order without ties. If the subject i is censored (dependent censoring) at time x_i, then for each event time point $x_j > x_i$, we want to compute the probability that this subject i fails at time x_j. With some calculus (Zheng and Klein, 1994) under the joint distribution

assumption specified by the copula H, for $t > x_i$, we have

$$\begin{aligned} \Pr(T_i \geq t | T_i \geq x_i, C_i = x_i) &= \frac{\int_t^\infty f(x, x_i)dx}{\int_{x_i}^\infty f(x, x_i)dx} \\ &= \frac{\Pr(T_i \geq t, C_i = x_i)}{\Pr(T_i \geq x_i, C_i = x_i)} \\ &= \frac{1 - H_v\{F_i(t), G_i(x_i); \alpha\}}{1 - H_v\{F_i(x_i), G_i(x_i); \alpha\}} \end{aligned} \tag{3.5}$$

where $H_v(a, b; \alpha) = \frac{\partial H(u,v;\alpha)}{\partial v}\Big|_{(u,v)=(a,b)}$. Denote the above conditional survival probability by $P_i(t)$.

Then the piece of mass that the censored subject i loses at time x_j is denoted as $D_i(x_j)$, where x_j is the event time of subject j and $x_j > x_i$. Assuming that x_i, $i = 1, ..., n$ are sorted in ascending order without ties, we define $D_i(x_j)$ as follows:

$$D_i(x_j) = P_i(x_{j-1}) - P_i(x_j). \tag{3.6}$$

Similarly, all other subjects dependently censored before time x_j lose some mass at the event time point x_j, denoted as $D_i(x_j)$ as well.

As far as the dependent censoring is concerned, we also need estimates of β_c since β_c is included in the $P_i(t)$. As mentioned, we put events and dependent censoring in the competing-risks settings. That is to say, by treating dependent censoring as the event of interest, we can get the counterpart functions. At last, we get parameter estimates by maximizing likelihood functions for both events and dependent censoring, respectively.

Specifically, for a subject i fails at time x_i, we want to compute the probability

that this subject i is dependently censored through time c. For $c > x_i$, we have

$$\begin{aligned} Q_i(c) &= \frac{\int_c^\infty f(x, x_i)dx}{\int_{x_i}^\infty f(x, x_i)dx} \\ &= \Pr(C_i \geq c | C_i \geq x_i, T_i = x_i) \\ &= \frac{\Pr(C_i \geq c, T_i = x_i)}{\Pr(C_i \geq x_i, T_i = x_i)} \\ &= \frac{1 - H_u\{F_i(x_i), G_i(c); \alpha\}}{1 - H_u\{F_i(x_i), G_i(x_i); \alpha\}}. \end{aligned} \tag{3.7}$$

where $H_u(a, b; \alpha) = \frac{\partial H(u,v;\alpha)}{\partial u}\Big|_{(u,v)=(a,b)}$.

Let $E_i(x_j)$ represent the piece of mass that a failed subject i losses at dependent censoring time x_j. Again, for $x_j > x_i$, we define that

$$E_i(x_j) = Q_i(x_{j-1}) - Q_i(x_j). \tag{3.8}$$

Note that there are a few restrictions of above notations listed in Section 3.3.3.

3.3.3 Partial likelihood functions

In order to consider the event and dependent censoring simultaneously, we introduce copula-based indicator functions $D_i(x_j)$ and $E_i(x_j)$, which can take any value between 0 and 1; while traditional indicators can only take two values, either 0 or 1. For a subject with dependent censoring, we assume that its contribution to the likelihood function is decreasing gradually at each observed event time point, represented by $P_i(t)$ and $D_i(x_j)$. Intuitively, this subject is going to fail gradually rather than immediately, as the traditional survival analysis assumes.

We define an extended Cox partial likelihood function for an event as follows:

$$L_j^{(T)}(\beta) = \prod_{i=1}^{j} \left\{ \frac{P_i(x_j)\exp(Z_i'\beta)}{\sum_{k=1}^{n} P_k(x_j)\exp(Z_k'\beta)} \right\}^{D_i(x_j)} \tag{3.9}$$

and,

$$\begin{aligned} L^{(T)}(\beta) &= \prod_{j=1}^{n} L_j^{(T)}(\beta) \\ &= \prod_{j=1}^{n}\prod_{i=1}^{j} \left\{ \frac{P_i(x_j)\exp(Z_i'\beta)}{\sum_{k=1}^{n} P_k(x_j)\exp(Z_k'\beta)} \right\}^{D_i(x_j)}, \end{aligned} \tag{3.10}$$

where $L_j^{(T)}(\beta)$ is the likelihood function for the time point x_j.

To make the above equation well-defined, we need to make several adjustments. First, we set $P_k(x_j) = 1$ for $x_j \leq k$. Because by definition, $P_k(x_j)$ represents the probability of the subject k survived at time x_j for $x_j > k$, given that it survives until time k. That is for $x_j \leq k$, the subject k survives, i.e. $P_k(x_j) = 1$.

Second, please note that observations have been sorted in ascending order already. That is, for $i = 1, ..., n$, the subject i failed at time x_i. Therefore, for a failed subject i, we set $P_i(x_j) = 0$ for $j > i$.

Third, regarding $D_i(x_j)$ for a failed subject i, instead of using Equation 3.6, we set $D_i(x_i) = 1$ and $D_i(x_j) = 0$ for $j > i$. That is, the subject i losses all of its unit mass when it fails at time x_i. As a result, a failed subject contributes only one time in the extended Cox partial likelihood function.

The counterpart of the extended Cox partial likelihood function is shown below,

where we treat the previous dependent censoring as the event of interest:

$$\begin{aligned} L^{(C)}(\beta_c) &= \prod_{j=1}^{n} L_j^{(C)}(\beta_c) \\ &= \prod_{j=1}^{n}\prod_{i=1}^{j}\left\{\frac{Q_i(x_j)\exp(W_i'\beta_c)}{\sum_{k=1}^{n} Q_k(x_j)\exp(W_k'\beta_c)}\right\}^{E_i(x_j)} \end{aligned} \tag{3.11}$$

Similarly, we set $Q_k(x_j) = 1$ for $x_j \leq k$. For a dependently censored subject i, set $Q_i(x_j) = 0$, $E_i(x_i) = 1$, and $E_i(x_j) = 0$ for $j > i$.

We also need some additional setups for the independent censoring. For an independently censored subject i, due to the similar reasons mentioned above, we set $P_i(x_j) = Q_i(x_j) = 1$ for $j \leq i$, $P_i(x_j) = Q_i(x_j) = 0$ for $j > i$. For all j, we let $D_i(x_j) = E_i(x_j) = 0$. That is, for independent censoring, the subject i does not lose mass at the time point x_j.

In summary, the way we treat events and independent censoring by the proposed method is the same as the one by the traditional Cox method. Only dependent censoring is treated differently.

Next, the parameters β and β_c can be estimated by maximizing the following extended joint partial likelihood function,

$$L(\beta, \beta_c) = L^{(T)}(\beta)L^{(C)}(\beta_c). \tag{3.12}$$

Please note that the likelihood function $L^{(T)}(\beta)$ in Equation 3.10 depends not only on parameter β, but also implicitly on parameter β_c through the functions $P_i(x_j)$ and $D_i(x_j)$. So it is better written as $L^{(T)}(\beta, \beta_c)$. Similarly, the likelihood $L^{(C)}(\beta_c)$ in

Equation 3.11 is better written as $L^{(C)}(\beta, \beta_c)$. In this dissertation, we will simply denote $L^{(T)}(\beta, \beta_c)$ and $L^{(C)}(\beta, \beta_c)$ as $L^{(T)}(\beta)$ and $L^{(C)}(\beta_c)$.

3.3.4 Iteration steps

Because the extended joint partial likelihood function involves unknown quantities such as $P_i(x_j), Q_i(x_j)$, $D_i(x_j)$, $E_i(x_j)$ and etc., we use iterations to get final estimates for βs and β_cs. Note that we model events and dependent censoring as competing risks as explained in Section 3.3.1. Thus, we need to solve two sets of functions. We treat failures as the event of interest in one case and treat dependent censoring as the event of interest in the other case. The iteration flow is listed first, and then each step is explained in detail.

The iteration flow,

Step 1. Initialize $\hat{\beta}^{(0)}, \hat{\beta}_c^{(0)} \Rightarrow \hat{S}_0^{(0)}(t), \hat{R}_0^{(0)}(t)$

$\Rightarrow \hat{F}_i^{(0)}(\cdot), \hat{G}_i^{(0)}(\cdot) \Rightarrow \hat{P}_i^{(0)}(\cdot), \hat{Q}_i^{(0)}(\cdot) \Rightarrow \hat{D}_i^{(0)}(\cdot), \hat{E}_i^{(0)}(\cdot)$

$\Rightarrow$ Maximize $L(\beta, \beta_c)$ (see Equation 3.12) $\Rightarrow$ Let $m = 1$; $\hat{\beta}^{(1)}, \hat{\beta}_c^{(1)}$

Step 2, $\Rightarrow \hat{\beta}^{(m)}, \hat{\beta}_c^{(m)} \Rightarrow \hat{S}_0^{(m)}(t), \hat{R}_0^{(m)}(t)$

$\Rightarrow \hat{F}_i^{(m)}(\cdot), \hat{G}_i^{(m)}(\cdot) \Rightarrow \hat{P}_i^{(m)}(\cdot), \hat{Q}_i^{(m)}(\cdot) \Rightarrow \hat{D}_i^{(m)}(\cdot), \hat{E}_i^{(m)}(\cdot)$

Step 3, $\Rightarrow$ Again, maximize $L(\beta, \beta_c)$ (Equation 3.12) $\Rightarrow \hat{\beta}^{(m+1)}, \hat{\beta}_c^{(m+1)}$

Step 4, $\Rightarrow$ Keep updating $\hat{\beta}^{(m+1)}$ and $\hat{\beta}_c^{(m+1)}$ until they converge respectively

Step 5, $\Rightarrow$ Let $m = m+1$, repeat steps 2, 3 and 4 until $\hat{\beta}^{(m+1)}$ and $\hat{\beta}_C^{(m+1)}$ converge respectively.

Consequently, we could get estimated hazard functions $\hat{\Lambda}(\cdot)$ and $\hat{\Psi}(\cdot)$, and survival functions $\hat{S}(\cdot)$ and $\hat{R}(\cdot)$.

Specifically, at Step 1,

Assuming independent censoring, we fit two Cox proportional hazards models to get initial estimators $\hat{\beta}^{(0)}$ and $\hat{\beta}_c^{(0)}$ for β and β_c, respectively. Then we use the Breslow (1972) method to obtain estimators for baseline cumulative hazard functions, which give estimates of baseline survival functions.

- Baseline survival functions

For an event T and $i \leq k$, we have

$$\hat{S}_0^{(0)}(t) = exp\{-\hat{\Lambda}_0^{(0)}(t)\} = exp\{-\sum_{obs_i \leq t} \frac{\delta_{obs_i,1}}{\sum_{obs_k \geq obs_i} exp(Z_k' \hat{\beta}^{(0)})}\}; \tag{3.13}$$

For dependent censoring C and $i \leq k$, we have

$$\hat{R}_0^{(0)}(t) = exp\{-\hat{\Psi}_0^{(0)}(t)\} = exp\{-\sum_{obs_i \leq t} \frac{\delta_{obs_i,2}}{\sum_{obs_k \geq obs_i} exp(W_k' \hat{\beta}_c^{(0)})}\}; \tag{3.14}$$

- Marginal cumulative distribution functions

$$\begin{aligned}
\hat{F}_i^{(m)}(t) &= 1 - exp\{-\hat{\Lambda}_0^{(m)}(t)exp(Z_i'\hat{\beta}^{(m)})\} \\
&= 1 - \hat{S}_0^{(0)}(t)^{exp(Z_i'\hat{\beta}^{(m)})} \\
\hat{G}_i^{(m)}(t) &= 1 - exp\{-\hat{\Psi}_0^{(m)}(t)exp(W_i'\hat{\beta}_c^{(m)})\} \\
&= 1 - \hat{R}_0^{(0)}(t)^{exp(W_i'\hat{\beta}_c^{(m)})}
\end{aligned}$$

where $m = 0, 1, 2, ..., n$; (extended from Equation 3.3).

- $P_i(x_j)$ and $Q_i(x_j)$ represent the following:

$P_i(x_j)$ = Pr(subject i will survive through x_j| censored at x_i), (see Equation 3.5);

$Q_i(x_j)$ = Pr(subject i will not be dependently censored through x_j| failed at x_i), (see Equation 3.7).

For $x_j > x_i$, we have

$$\hat{P}_i^{(m)}(x_j) = \frac{1 - H_v\{\hat{F}_i^{(m)}(x_j), \hat{G}_i^{(m)}(x_i); \alpha\}}{1 - H_v\{\hat{F}_i^{(m)}(x_i), \hat{G}_i^{(m)}(x_i); \alpha\}}, \tag{3.15}$$

if subject i is censored; (extended from Equation 3.5).

$$\hat{Q}_i^{(m)}(x_j) = \frac{1 - H_u\{\hat{F}_i^{(m)}(x_i), \hat{G}_i^{(m)}(x_j); \alpha\}}{1 - H_u\{\hat{F}_i^{(m)}(x_i), \hat{G}_i^{(m)}(x_i); \alpha\}}, \tag{3.16}$$

if subject i is failed; (extended from Equation 3.7).

- Copula-based indicator functions

For $x_j > x_i$, we have

$$\hat{D}_i^{(m)}(x_j) = \hat{P}_i^{(m)}(x_{j-1}) - \hat{P}_i^{(m)}(x_j), \tag{3.17}$$

if subject i is censored; (extended from Equation 3.6).

$$\hat{E}_i^{(m)}(x_j) = \hat{Q}_i^{(m)}(x_{j-1}) - \hat{Q}_i^{(m)}(x_j), \tag{3.18}$$

if subject i is failed; (extended from Equation 3.8).

Using the above computation results and other specifications as described earlier, replace the unknown functions by their estimates at the initial step, and then maximize the likelihood functions in Equations 3.10 and 3.11 with respect to β and β_c, respectively. Denote the resulting estimators for β and β_c by $\hat{\beta}^{(1)}$ and $\hat{\beta}_c^{(1)}$.

Step 2,

Use $\hat{\beta}^{(m)}, \hat{\beta}_c^{(m)}, \hat{P}_i^{(m-1)}(\cdot), \hat{Q}_i^{(m-1)}(\cdot), \hat{D}_i^{(m-1)}(\cdot)$ and $\hat{E}_i^{(m-1)}(\cdot)$ to obtain $\hat{S}_0^{(m)}(\cdot)$ and $\hat{R}_0^{(m)}(\cdot)$.

$$\begin{aligned}
\hat{S}_0^{(m)}(t) &= exp\{-\hat{\Lambda}_0^{(m)}(t)\} \\
&= exp\{-\sum_{x_j \le t} \frac{\delta_{x_j,1}}{\sum_{k=1}^n \hat{P}_k^{(m-1)}(x_j) exp\left(Z_k' \hat{\beta}^{(m)}\right)} \\
&\quad -\sum_{x_j \le t} \frac{\sum_{x_i < x_j} \delta_{x_i,2} \hat{D}_{x_i}^{(m-1)}(x_j)}{\sum_{k=1}^n \hat{P}_k^{(m-1)}(x_j) exp\left(Z_k' \hat{\beta}^{(m)}\right)}\} \\
&= exp\{-\sum_{x_j \le t} \frac{\sum_{x_i \le x_j} \hat{D}_{x_i}^{(m-1)}(x_j)}{\sum_{k=1}^n \hat{P}_k^{(m)}(x_j) exp\left(Z_k' \hat{\beta}^{(m)}\right)}\}
\end{aligned} \tag{3.19}$$

$$\begin{aligned}
\hat{R}_0^{(m)}(t) &= exp\{-\hat{\Psi}_0^{(m)}(t)\} \\
&= exp\{-\sum_{x_j \le t} \frac{\delta_{x_j,2}}{\sum_{k=1}^n \hat{Q}_k^{(m-1)}(x_j) exp\left(W_k' \hat{\beta}_c^{(m)}\right)} \\
&\quad -\sum_{x_j \le t} \frac{\sum_{x_i < x_j} \delta_{x_i,1} \hat{E}_{x_i}^{(m-1)}(x_j)}{\sum_{k=1}^n \hat{Q}_k^{(m-1)}(x_j) exp\left(W_k' \hat{\beta}_c^{(m)}\right)}\} \\
&= exp\{-\sum_{x_j \le t} \frac{\sum_{x_i \le x_j} \hat{E}_{x_i}^{(m-1)}(x_j)}{\sum_{k=1}^n \hat{Q}_k^{(m)}(x_j) exp\left(W_k' \hat{\beta}_c^{(m)}\right)}\} \qquad (3.20)
\end{aligned}$$

Note that the proposed $\hat{S}_0^{(m)}(t)$ and $\hat{R}_0^{(m)}(t)$ are also copula-based, which are different from $\hat{S}_0^{(0)}(t)$ and $\hat{R}_0^{(0)}(t)$.

Similarly as Step 1, we then update corresponding $\hat{F}_i^{(m)}(\cdot)$, $\hat{G}_i^{(m)}(\cdot)$, $\hat{P}_i^{(m)}(\cdot)$, $\hat{Q}_i^{(m)}(\cdot)$, $\hat{D}_i^{(m)}(\cdot)$ and $\hat{E}_i^{(m)}(\cdot)$.

Step 3,

Again, maximize $L^{(T)}(\beta, \beta_c)$ and get the updated $\hat{\beta}^{(m+1)}$ and $\hat{\beta}_c^{(m+1)}$.

Step 4,

Iterates as below:

- Keep updating $\hat{\beta}^{(m+1)}$ and $\hat{\beta}_c^{(m+1)}$ using the Newton-Raphson Method until they converge.

- Following values are used to get new estimates of β and β_c: $\hat{\beta}^{(m+1)}, \hat{\beta}_c^{(m+1)}, \hat{P}_i^{(m)}(\cdot), \hat{Q}_i^{(m)}(\cdot), \hat{D}_i^{(m)}(\cdot)$ and $\hat{E}_i^{(m)}(\cdot)$. That is, after getting new

$\hat{\beta}^{(m+1)}$ and $\hat{\beta}_c^{(m+1)}$, following values are not updated within Step 4:

$\hat{P}_i^{(m)}(\cdot), \hat{Q}_i^{(m)}(\cdot), \hat{D}_i^{(m)}(\cdot)$ and $\hat{E}_i^{(m)}(\cdot)$.

Step 5,

Let $m = m + 1$, repeat Steps 2, 3 and 4 until $\hat{\beta}^{(m+1)}$ and $\hat{\beta}_c^{(m+1)}$ converge, respectively.

Furthermore, we could get estimated cumulative hazard functions $\hat{\Lambda}(\cdot)$ and $\hat{\Psi}(\cdot)$, survival functions $\hat{S}(\cdot)$ and $\hat{R}(\cdot)$ and their confidence intervals by combining results of simulations and bootstrapping. Details are listed in Section 3.4.

3.3.5 Remarks

Another approach of the iteration

There is another approach to complete the iteration. Instead of using a separate Step 4, we can combine Steps 2, 3 and 4 together to get the converged β and β_c. That is to say, always update $\hat{P}_i^{(m)}(\cdot), \hat{Q}_i^{(m)}(\cdot), \hat{D}_i^{(m)}(\cdot)$ and $\hat{E}_i^{(m)}(\cdot)$ after getting $\hat{\beta}^{(m+1)}$ and $\hat{\beta}_c^{(m+1)}$. We tested those two approaches of iterations. Both the results and the length of time used are quite similar. Thus, we randomly picked the first approach to conduct this research.

Assumption

An implicit assumption used in our method is that, the association between events and dependent censoring is the same among different subgroups defined by covariate values. It is plausible in most situations. This assumption would fail if, for instance,

those patients who withdraw are the sicker in Group A and the healthier in Group B. However, such phenomenon is unlikely, especially in a blinded randomized trial.

Consistency

Note that the estimators $\hat{\beta}$, $\hat{\beta}_c$, $\hat{\Lambda}(\cdot)$ and $\hat{\Psi}(\cdot)$ satisfy the definition of "self-consistent", which was first introduced by Efron (1967). Based upon such definition, many research projects were conducted, among which Tsai and Crowley (1985) discussed the theoretical properties of self-consistent estimators in general non-regression settings. They showed (1) the guaranteed convergence of the above iteration algorithm and its connection with the Expectation-Maximization (EM) algorithm (Dempster et al., 1977); (2) such a self-consistent estimator is actually a generalized maximum-likelihood estimator in the sense of Kiefer and Wolfowitz (1956); (3) the strong consistency of the self-consistent estimators; and (4) its weak convergence to a Gaussian process. Zheng and Klein (1994) also shows their $\hat{S}(t)$ and $\hat{R}(t)$ are self-consistent estimators (see Section 2.5 for details).

These results and the simulation studies in the next section indicate potentially good large sample properties of our estimators in the regression setting. However, further theoretical investigation will be helpful.

Covariance

The covariance matrices of the above estimators can be obtained by the bootstrap

method (Efron, 1979). By the algorithm in Section 3.3, it can be seen that the final survival estimator $\hat{S}_0(\cdot)$ has jumps at event time points only. And $\hat{R}(\cdot)$ has jumps at informative censored time points only.

Ties

When there are tied failure events, the proposed method (Equation of likelihood function 3.10 and 3.11) can handle them as in the Breslow's (1974) method. Actually, if we view the pieces of mass $D_i(x_j), i = 1, ..., j$, as ties at time x_j, then the formation of the $L^{(T)}(\beta)$ and $L^{(C)}(\beta_c)$ (in Equations 3.10 and 3.11) are obtained from Breslow's method. In the simulation process, ties can be avoided easily. In the application case, for simplicity, we trimmed the data before fitting the model to eliminate ties, which has no impact to the conclusion. See details in Chapter 4.

Values of the α

In this research, the parameter α was not estimated. Instead, it was assumed to be known. In reality, it is unknown. But when knowledge is available from experts or literature about the degree and direction of the association between events and informative censoring, our approach can be used to obtain less biased parameter estimates. If not much knowledge is available, the proposed method can be used to do a sensitivity analysis in a very conservative way, just as was done by Park, Tian and Wei (2006). That is, letting the Kendall's τ change from near -1 to near 1 to perform

the sensitivity analysis. In some situations, for instance, when censoring percentages are small and/or balanced across covariate groups, the estimated regression parameters and their confidence intervals may not change much. As a result, the typical assumption of independent censoring may still be valid and traditional approach in survival analysis can be used. In such situations, a sensitivity analysis can help us decide how much confidence can be put in the results of analysis.

3.4 Simulation study of dependent censoring

In order to evaluate the method proposed in Section 3.3, we conducted a simulation study, through which we compared the results from correctly assuming dependent censoring with the ones from the false assumptions (traditional Coxph method).

We run 300 simulations with sample size 200. And for 10 simulations, we run 30 times of bootstrapping each to get the covariance matrices. The statistical software package R is used to implement the simulation throughout this dissertation.

To generate data, we assume two covariates: B = 0 or 1 with equal probability and centered Age which follows Uniform (-5,5). Assuming Weibull distributions, which is a popular distribution to model the time to event data, we specified the marginal distributions for events and informative censoring times T and C, respectively, by the following survival functions.

For time t, we have

$$S(t) = \exp\{-.5t^2\exp(\beta_1 B + \beta_2 \text{Age})\}, \tag{3.21}$$

$$R(t) = \exp\{-.5t\exp(\beta_{1c} B + \beta_{2c} \text{Age})\}, \tag{3.22}$$

where $\beta_1 = .5, \beta_2 = .4, \beta_{1c} = .6$ and $\beta_{2c} = .3$.

The joint distribution of T and C is specified by the Frank copula as below:

$$J(t, c; \alpha) = \log_\alpha\{1 + \frac{(\alpha^{F(t)} - 1)(\alpha^{G(c)} - 1)}{\alpha - 1}\} \tag{3.23}$$

We tested Kendall's $\tau = 0.8$, corresponding to $\alpha = 0.00000001258$ with different percentages of events, dependent and independent censoring.

In order to generate the observations of event time and informative censoring time, we used the results by Nelsen (1986):

- Step 1, generate two independent Uniform(0,1) random variables u and v'.
- Step 2, let $v = \log_\alpha\{1 + \frac{v'(\alpha-1)}{v'+(1-v')\alpha^u}\}$.

Then we get the cumulative distribution functions u and v for two variables, which satisfy the relationship specified by the Frank copula. We get $S(t)$ and $R(t)$ through u and v. Plug them into Equations 3.21 and 3.22, and then get t_T and t_C as shown below:

$$t_T = \sqrt{\frac{-2\log(S(t))}{exp(Z_i'\beta)}}$$

$$t_C = \frac{-2\log(R(t))}{exp(W_i'\beta_c)}$$

Assuming α is known, we get the time of event T_i and dependent censoring C_i. In addition, we generated another random time variable S_i from a Uniform distribution, which is independent of everything else. This was the independent censoring time. Then, the obtained observations are $X_i = min(T_i, C_i, S_i)$. We also had indicators $\delta_1 = I(X = T)$ and $\delta_2 = I(X = C)$. After putting all the data in ascending order by the time of observations, we used the proposed method in Section 3.3 to get the estimators of β and β_c and their standard errors.

Test 1

In Test 1, a Uniform (2,12) is used to generate S_i. The percentage of events is 47%; the percentage of dependent censoring is 48.5%; and the percentage of independent censoring is 4.5%. The results by the proposed method are summarized in Table 3.1. In contrast, the results from falsely assumed independent censoring are shown in Table 3.2. Please note that in the proposed method, the parameter α was not estimated. Instead, it was assumed to be known as 0.00000001258, corresponding to $\tau = 0.8$.

From Tables 3.1 and 3.2, it can be seen that the proposed method outperforms the traditional Coxph method. The parameter estimates from the proposed method are closer to the true value than the ones from the Coxph method. The Standard Deviation (SD) and the Standard Error (SE) of estimators from the proposed method are close to the ones from the Coxph method. Similarly, the survival curves obtained

Table 3.1: Results Obtained Using the Proposed Method With τ=0.8; T, C and S are 47%, 48.5% and 4.5%.

	Parameters	*True Values*	*Estimates*	*SD of Est*	*SE of Est*
1	β_1	.5	0.543	0.169	0.204
2	β_2	.4	0.394	0.045	0.060
3	β_{1c}	.6	0.586	0.246	0.333
4	β_{2c}	.3	0.397	0.105	0.102

using the proposed method are closer to the true curves than the ones obtained using the Coxph method. These results are confirmed by Figure 3.1 and Figure 3.2.

Test 2

In Test 2, a Uniform (0,5) is used for S_i and $\tau = 0.8$. Percentages of T, C and S are 34.5%, 51.5%, and 14%, respectively, different from Test 1.

From Tables 3.3 and 3.4, it can be seen that the proposed method outperforms the traditional Coxph method. Similarly as in Test 1, this result is confirmed by Figure 3.3 and Figure 3.4.

In summary, when the dependent censoring exists with certain magnitude, the proposed method provides more accurate estimates of parameters and of survival curves than the traditional Coxph method. In fact, Coxph model assumes independent censoring.

Table 3.2: Results Obtained Using the Coxph Method With τ=0.8; T, C and S are 47%, 48.5% and 4.5%.

	Parameters	*True Values*	*Estimates*	*SD of Est*	*SE of Est*
1	β_1	.5	0.347	0.220	0.226
2	β_2	.4	0.489	0.065	0.060
3	β_{1c}	.6	0.703	0.205	0.207
4	β_{2c}	.3	0.283	0.042	0.041

Table 3.3: Results Obtained Using the Proposed Method With τ=0.8; T, C and S are 34.5%, 51.5%, and 14%.

	Parameters	*True Value*	*Estimates*	*SD of Est*	*SE of Est*
1	β_1	.5	0.515	0.167	0.195
2	β_2	.4	0.375	0.042	0.046
3	β_{1c}	.6	0.487	0.167	0.196
4	β_{2c}	.3	0.287	0.036	0.045

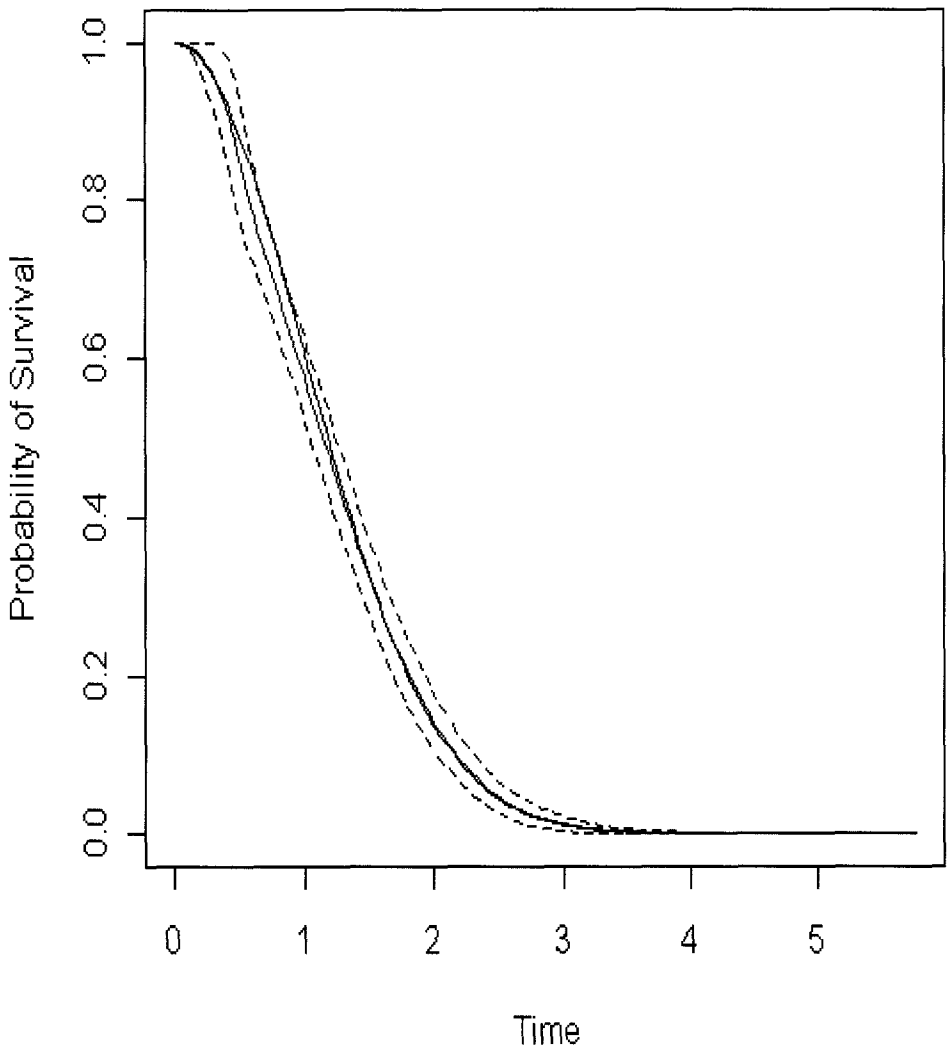

Figure 3.1: Survival Curves Obtained Using the Proposed Method From Test 1

Treat T: as the time of event; C: as the time of dependent censoring; S: as the time of independent censoring. The true curve (in Red), the estimated curve obtained using the proposed method (in Blue) and its Confidence Intervals (dashed lines).

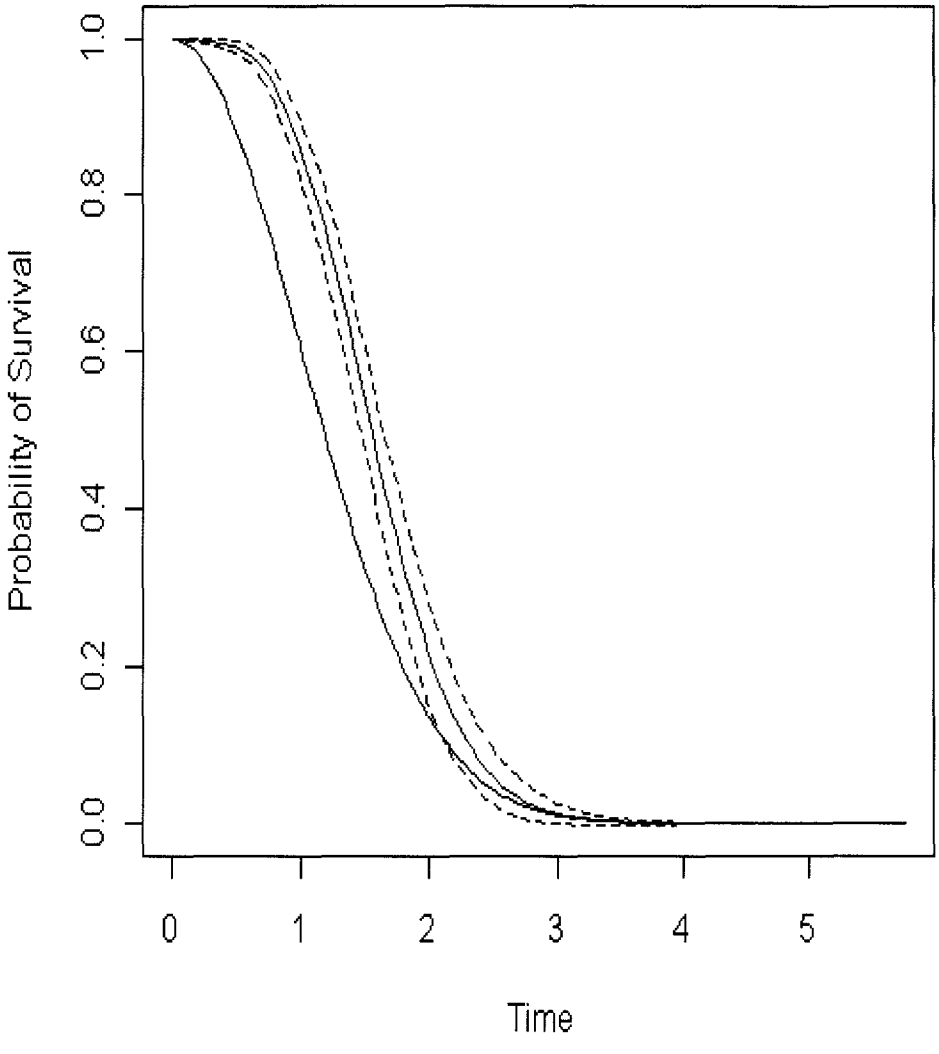

Figure 3.2: Survival Curves Obtained Using the Coxph Method From Test 1

Treat T: as the time of event; C: as the time of dependent censoring; S: as the time of independent censoring. The true curve (in Red), the estimated curve obtained using the Coxph method (in Blue) and its Confidence Intervals (dashed lines).

Table 3.4: Results Obtained Using the Coxph Method With τ=0.8; T, C and S are 34.5%, 51.5%, and 14%.

	Parameters	*True Value*	*Estimates*	*SD of Est*	*SE of Est*
1	β_1	.5	0.314	0.270	0.256
2	β_2	.4	0.507	0.069	0.064
3	β_{1c}	.6	0.707	0.213	0.207
4	β_{2c}	.3	0.280	0.042	0.041

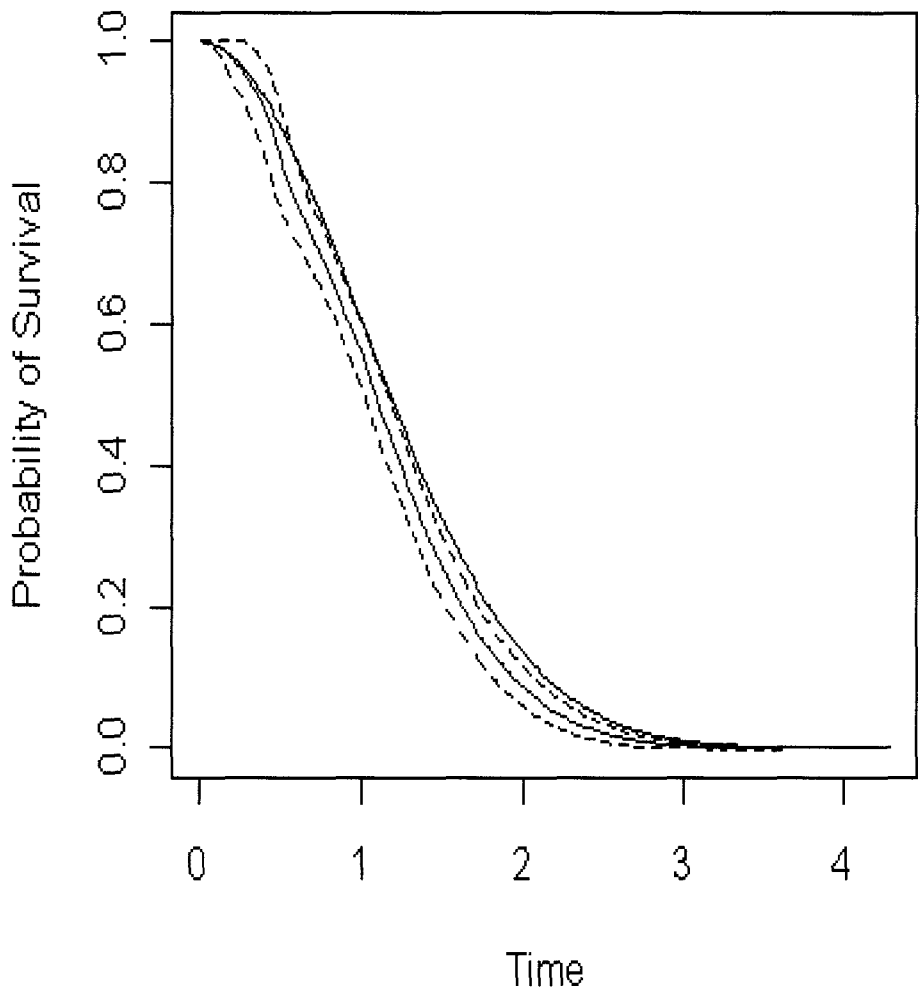

Figure 3.3: Survival Curves Obtained Using the Proposed Method From Test 2

Treat T: as the time of event; C: as the time of dependent censoring; S: as the time of independent censoring. The true curve (in Red), the estimated curve obtained using the proposed method (in Blue) and its Confidence Intervals (dashed lines).

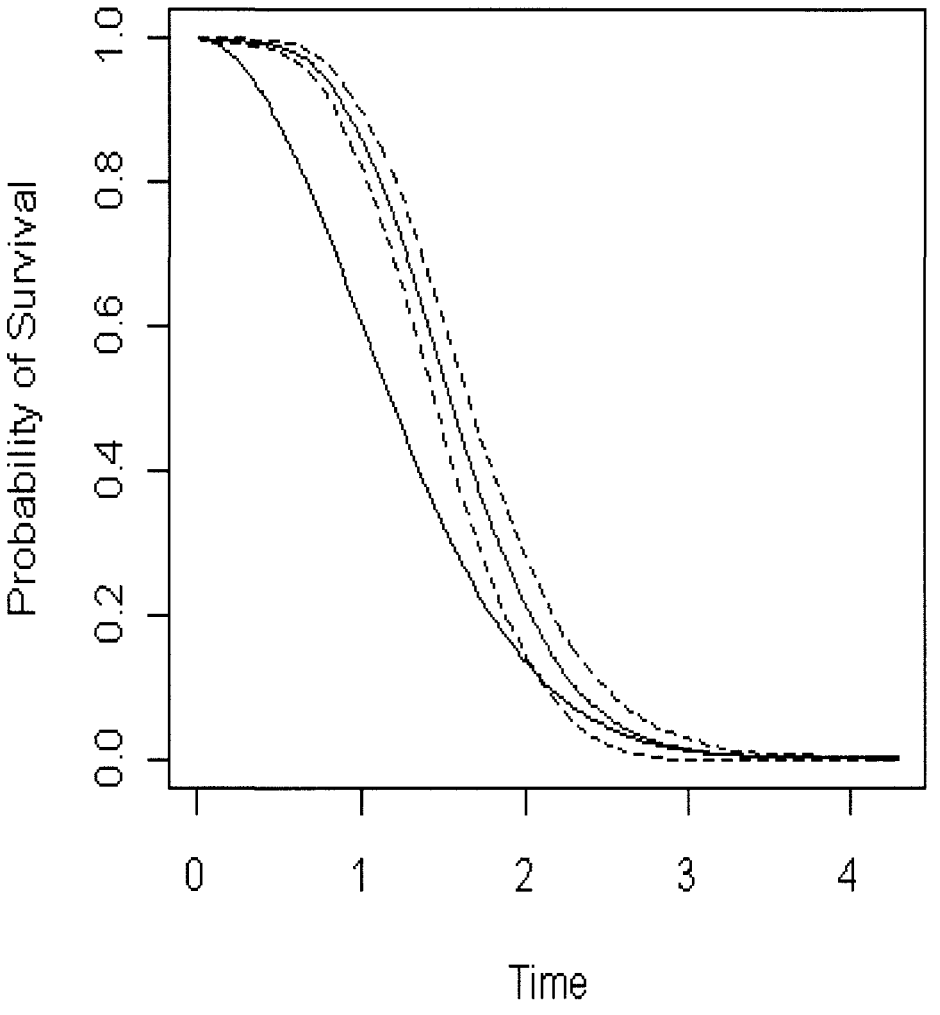

Figure 3.4: Survival Curves Obtained Using the Coxph Method From Test 2

Treat T: as the time of event; C: as the time of dependent censoring; S: as the time of independent censoring. The true curve (in Red), the estimated curve obtained using the Coxph method (in Blue) and its Confidence Intervals (dashed lines).

Chapter 4

Application: How Does the 1999 Gramm-Leach-Bliley Act Impact the Survival of Insurance Companies?

In this chapter, the methodology developed in Chapter 3 is applied to the real insurance company survival data. We also compare the performance of our method to the performance of the conventional Coxph method.

Section 4.1 lays out the framework for the application and explains the two scenarios that we are investigating. Section 4.2 introduces the data source and structure. Sections 4.3 and 4.4 show how covariates are determined and present results of the

sensitivity analysis.

The work described in this chapter built the foundation to test the hypothesis on whether a hazard function has a significant change-point, which is illustrated in Chapter 5.

4.1 Basic framework

We want to understand the impact of the GLB Act on insurance companies from two perspectives, i.e. to the companies' survival and to the Acquisition among insurance companies. When we looked at the data, the percentage of acquisitions is relatively high, which shows that being acquired is an important choice that insurers have when their survival ability is challenged. Hence, a study of acquisitions is essentially valuable for understanding the impact of the GLB Act. Given the respective definitions of an event and dependent censoring in Section 3.1, we applied the proposed method to two scenarios.

In the first scenario, the purpose is to identify what factors, if any, have significant impacts to insurance firms' survival, denoted as "for bankruptcies". Here, the event of interest is "going bankrupt".

In the second scenario, we focused on the insurance companies that are likely and willing to be acquired. The purpose is to identify what factors, if any, have significant impacts on insurance firms' acquisition, denoted as "for acquisitions". Here, the event of interest is "pursuing acquisitions". This scenario is not quite the same as

traditional definitions in survival analysis setting because the event is not necessarily bad events that people want to avoid completely. As mentioned in Section 3.1, we define an **event** as an insurance company being acquired in M&A deal(s), a **dependent censoring** as an insurance company filed bankruptcy, and an **independent censoring** as otherwise. Also, we are more interested in knowing whether there are any negative impacts of the GLB Act. Thus, in this research for M&A deal(s), we focus on "being acquired", which is dependent with going bankrupt.

Given different definitions of events and censoring, the methodology developed in Chapter 3 is applicable to both scenarios under the competing-risks setting.

4.2 Data source and data structure

Given our research interest, we used the combined database of the University of Chicago's Center for Research on Security Prices (CRSP) and Standard & Poor's COMPUSTAT, which is created by a division of Standard & Poor's called COMPUSTAT. Those data sets are available through Wharton Research Data Services (WRDS). The major insurance companies doing business in U.S., with Standard Industrial Classification (SIC) Codes 6311, 6321, 6331, and 6351, were extracted from the database. This choice of insurance companies is consistent with Marlett et al. (2003).

The industrial annual data with CRSP and COMPUSTAT combined is used to get the covariate information. Daily stocks data from CRSP is used to get the detailed

reasons of companies' elimination, which defines the bankruptcy and acquisition. "Permanent Numbers" (PERMNO) are used to merge different data sets. Since the missing values of covariates are relatively small, we simply deleted companies with missing covariates.

In the WRDS data set, companies are separated as "active companies" and "inactive companies". In this research, following the definitions in Chapter 3 on events and censoring, the active companies were cases of independent censoring, and inactive companies were further categorized into events or dependent censoring.

In total, 173 companies were included in the analysis, among which 21 companies went bankrupt, 68 are active companies, i.e. the independent censoring, and 84 were acquired by other companies and lost their original names. The percentages are 12.14%, 39.31% and 48.55%, respectively.

As mentioned, the time horizon we considered for this research is from 1994 to 2005, which covers 6 years before the passage of the GLB Act and 6 years after.

Under the proposed method, the observed time points are sorted in ascending order. Note that at the end of 2005, there are 68 active companies. Since their observed time points are tied, we added a random number of days from 1 to 30 days to each tied company to facilitate the analysis. This approach is reasonable since all independent censoring will contribute to the denominator of the partial likelihood function. The order of them will affect neither the parameter estimates, nor the results of this research. In addition, there are only a few companies that went bankrupt on

the same day. Their observed dates were randomly added or reduced by one day. As a result, there are no ties.

4.3 How covariates are determined

We selected covariates related to bankruptcy and acquisition using a stepwise selection algorithm assuming independent censoring. We set a p-value of 0.30 as the threshold for both variable entry and stay. That choice was based on the consideration of the potential change in p-values for covariates after accounting for dependent censoring. Outputs are shown in Table 4.1 for bankruptcy; and in Table 4.2 for acquisition (being acquired). A total of five covariates are selected for the model of bankruptcy and of acquisition. Two of them are significant for both bankruptcy and acquisition. They are size and liability. Age is a significant predictor in the model for bankruptcy; Growth and Profit are significant predictors for acquisition. Profit, growth and size were also used in Fama and French (2001) when studying disappearing dividends. Age was also used in Gretchen's Ph.D. dissertation (2004) regarding dividend initiation policy. Size was also used in Marlett et al. (2003) when studying the impact of the 1999 GLB Act. Intuitively, growth, profitability, size, age and liability are reasonable measures of a company's financial status. Therefore, they can be related to companies' survival.

Those five covariates are defined as follows,

$$Growth_i = \frac{Assets_i - Assets_{i-1}}{Assets_{i-1}}[\%] \quad (4.1)$$

$$Profitability_i = \frac{\#Shares_i \times \text{“}Earned \cdot Per \cdot Share_i\text{”}}{Assets_i}[\%] \quad (4.2)$$

$$Size = \text{“}Outstanding \cdot Shares\text{”} \times \text{“}Common \cdot Stock\text{”}[MM\$] \quad (4.3)$$

$$Age = 1994 - \text{“}Starting \cdot Year\text{”}[years] \quad (4.4)$$

$$Liability = Liability[MM\$], \quad (4.5)$$

where i indicates a particular calendar year.

Table 4.1: Results of P-values of Five Covariates for Bankruptcies

	Estimates	Exp of Est	SE of Est	z Statistic	p-value
Growth	0.116	1.123	0.440	0.263	0.790
Profit	0.218	1.243	0.456	0.477	0.630
Size	0.940	2.559	0.560	1.677	**0.093**
Age	-0.503	0.605	0.447	-1.125	**0.260**
Liability	1.545	4.686	0.695	2.223	**0.026**

Remarks

Note that the starting year represents the time when the company was first established. If the company was established before 1950, the starting year is counted as 1950, which is the earliest recorded year in the CRSP data set. In this analysis, all companies existed in 1994. Companies started after 1994 or companies that went

Table 4.2: Results of P-values of Five Covariates for Acquisitions

	Estimates	Exp of Est	SE of Est	z Statistic	p-value
Growth	0.335	1.398	0.228	1.468	**0.140**
Profit	0.255	1.290	0.236	1.079	**0.280**
Size	0.724	2.063	0.255	2.838	**0.005**
Age	0.100	1.105	0.232	0.429	0.670
Liability	-0.327	0.721	0.260	-1.257	**0.210**

bankrupt before 1994 are excluded from this research due to the potential bias. All covariates are redefined as binary variables categorized by the median before model fitting. Specifically, minimum to median is defined as 0 and median to maximum is defined as 1.

4.4 Sensitivity analysis and results

Note that the true value of Kendall's τ is unknown. In reality, we assume that the value of τ can be obtained by asking for experts' opinions. In this research, we applied different values of τ to conduct a sensitivity analysis. Specifically, we tested Kendall's $\tau = -0.5$, 0, 0.2, 0.5, 0.8 with corresponding values of α, as shown in Table 2.2.

To answer the two questions in Section 3.1.2, the results show that some covariates would become non-significant when dependent censoring is taken into account. For example, "Liability" in the first scenario. Some covariates would become non-

significant when the values of τ vary. For example, "Size" in both scenarios.

Furthermore, parameter estimates vary when the correlation between the dependent censoring and events changes. In some cases, the estimates vary slightly. For example, "Age" from the second scenario. The estimates may vary quite a bit. For instance, "Profit" in the first scenario. Detailed results are shown below.

- Tests 3, 4, 5 and 6

Tables 4.3 - 4.12 present parameter estimates and related standard errors and 95 percent confidence intervals for both bankruptcy and acquisition scenarios. Figures 4.1 - 4.10 illustrate such data graphically. In those figures, "Beta" represents the first scenario and "Betac" represents the second. P-values are calculated based on the fact that the following statistic $(\frac{\hat{\beta}}{SE(\hat{\beta})})^2$ follows the χ_1^2 distribution.

Figure 4.11 shows the cumulative hazard curve of bankruptcy obtained using the proposed method and its 95 percent confidence interval. It will be further discussed in Chapter 5.

Table 4.3: Results for Covariate 1 - Growth for Bankruptcies

	$\tau = -0.5$	$\tau = 0$	$\tau = 0.2$	$\tau = 0.5$	$\tau = 0.8$
Parameters	-0.870	0.116	0.204	0.169	0.166
SE of Est	1.127	0.440	0.232	0.247	0.208
Upper Bound	1.338	0.978	0.658	0.654	0.575
Lower Bound	-3.079	-0.746	-0.251	-0.315	-0.242

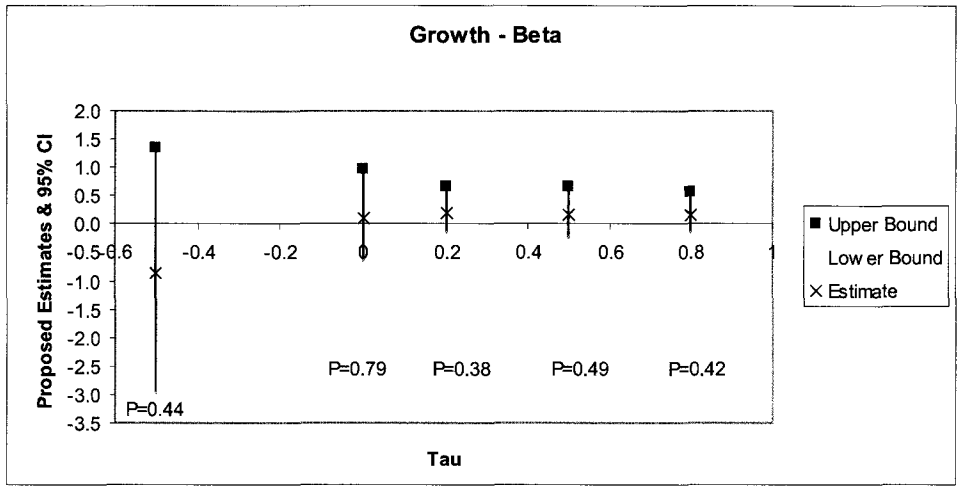

Figure 4.1: Parameter Estimates of Growth for Bankruptcies

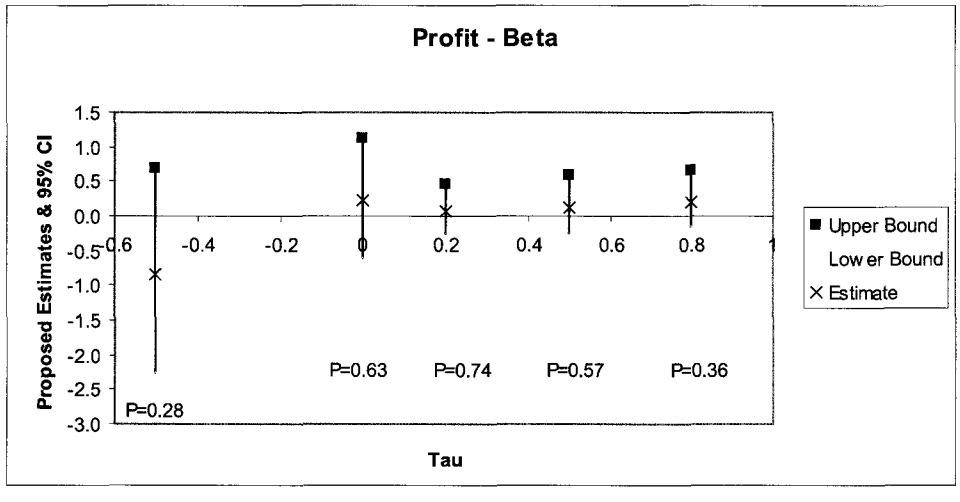

Figure 4.2: Parameter Estimates of Profitability for Bankruptcies

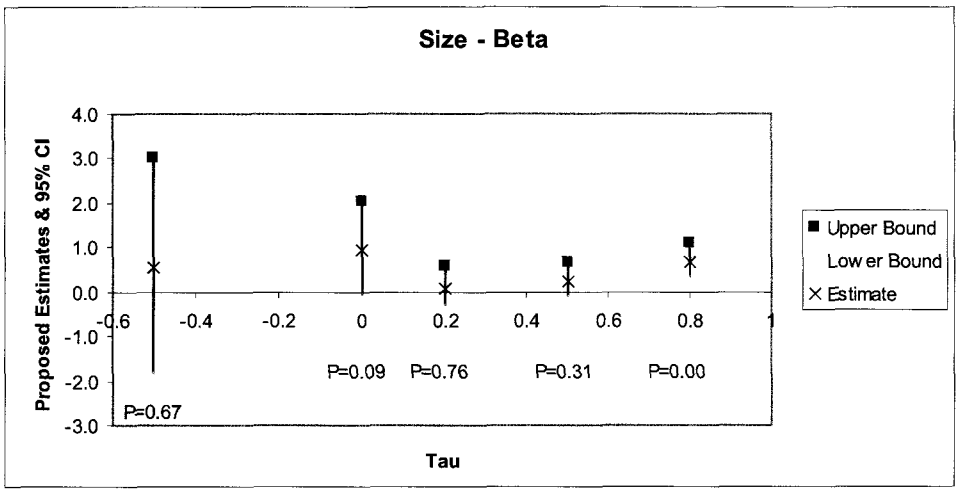

Figure 4.3: Parameter Estimates of Size for Bankruptcies

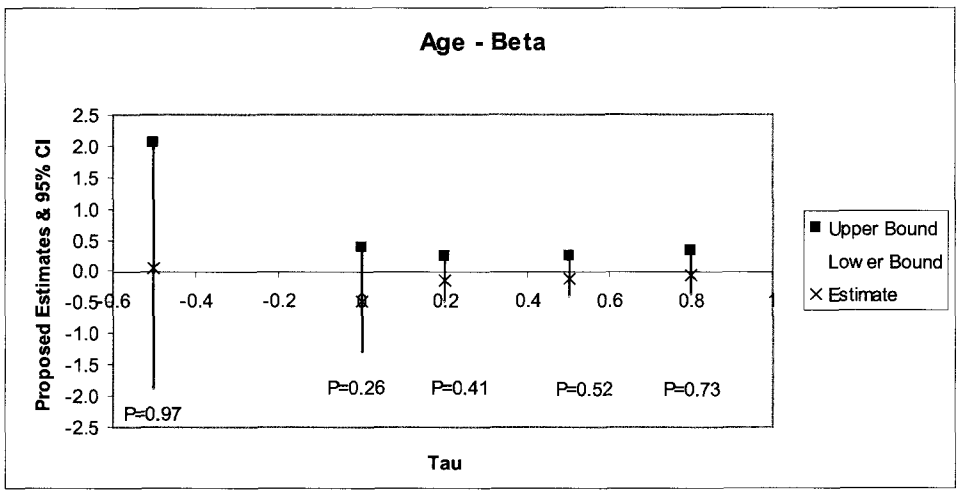

Figure 4.4: Parameter Estimates of Age for Bankruptcies

Table 4.4: Results for Covariate 2 - Profit for Bankruptcies

	$\tau = -0.5$	$\tau = 0$	$\tau = 0.2$	$\tau = 0.5$	$\tau = 0.8$
Parameters	-0.832	0.218	0.066	0.133	0.213
SE of Est	0.774	0.456	0.200	0.234	0.231
Upper Bound	0.685	1.112	0.458	0.591	0.666
Lower Bound	-2.349	-0.677	-0.327	-0.324	-0.241

Table 4.5: Results for Covariate 3 - Size for Bankruptcies

	$\tau = -0.5$	$\tau = 0$	$\tau = 0.2$	$\tau = 0.5$	$\tau = 0.8$
Parameters	0.546	0.940	0.075	0.223	0.653
SE of Est	1.262	0.560	0.249	0.220	0.225
Upper Bound	3.019	2.038	0.563	0.655	1.095
Lower Bound	-1.927	-0.158	-0.413	-0.209	0.212

Table 4.6: Results for Covariate 4 - Age for Bankruptcies

	$\tau = -0.5$	$\tau = 0$	$\tau = 0.2$	$\tau = 0.5$	$\tau = 0.8$
Parameters	0.042	-0.503	-0.166	-0.118	-0.070
SE of Est	1.020	0.447	0.203	0.186	0.206
Upper Bound	2.042	0.373	0.233	0.247	0.334
Lower Bound	-1.958	-1.379	-0.565	-0.483	-0.475

Table 4.7: Results for Covariate 5 - Liability for Bankruptcies

	$\tau = -0.5$	$\tau = 0$	$\tau = 0.2$	$\tau = 0.5$	$\tau = 0.8$
Parameters	0.438	1.545	0.039	0.249	-0.134
SE of Est	1.168	0.695	0.259	0.244	0.198
Upper Bound	2.727	2.906	0.547	0.726	0.254
Lower Bound	-1.851	0.183	-0.470	-0.229	-0.522

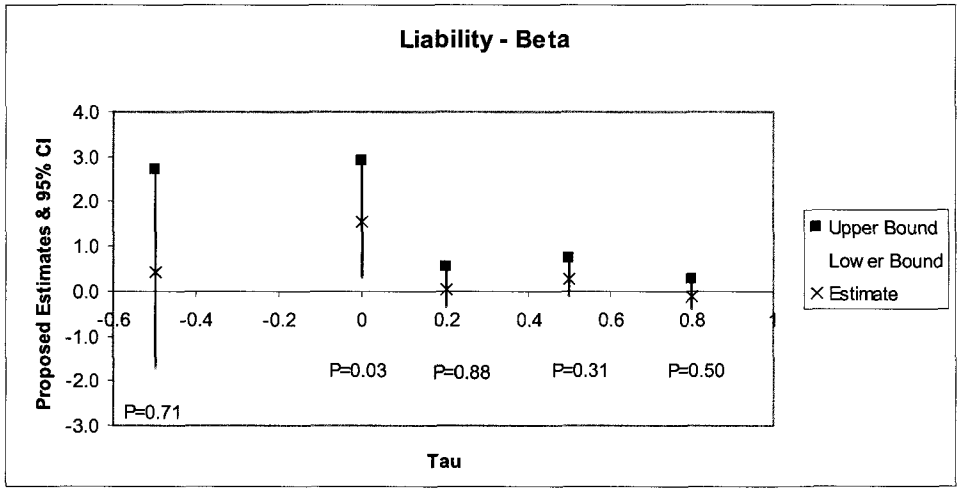

Figure 4.5: Parameter Estimates of Liability for Bankruptcies

Table 4.8: Results for Covariate 1 - Growth for Acquisitions

	$\tau = -0.5$	$\tau = 0$	$\tau = 0.2$	$\tau = 0.5$	$\tau = 0.8$
Parameters	0.061	0.335	0.224	0.172	0.173
SE of Est	0.207	0.228	0.190	0.216	0.198
Upper Bound	0.467	0.783	0.596	0.595	0.562
Lower Bound	-0.345	-0.112	-0.149	-0.251	-0.216

Table 4.9: Results for Covariate 2 - Profit for Acquisitions

	$\tau = -0.5$	$\tau = 0$	$\tau = 0.2$	$\tau = 0.5$	$\tau = 0.8$
Parameters	-0.154	0.255	0.157	0.148	0.215
SE of Est	0.186	0.236	0.171	0.200	0.224
Upper Bound	0.211	0.718	0.493	0.541	0.654
Lower Bound	-0.518	-0.208	-0.178	-0.244	-0.224

For the first scenario, the standard errors of estimates tend to be larger when $\tau = -0.5$ than the ones when τ takes other values. For the second scenario, the standard errors of estimates do not vary much as the values of τ change.

Note that the parameter estimates are not necessarily monotonically increasing or decreasing as the values of τ change.

Since it is the real data set, we loose the significant threshold of p-values to 0.15.

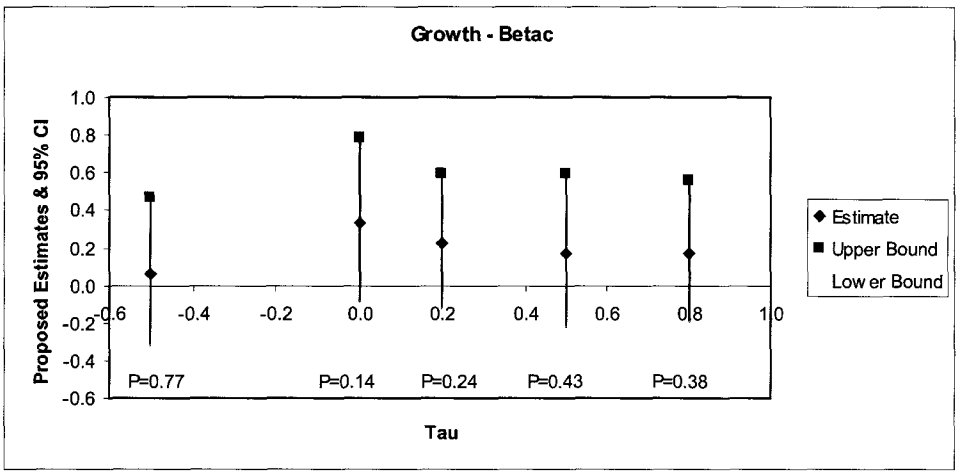

Figure 4.6: Parameter Estimates of Growth for Acquisitions

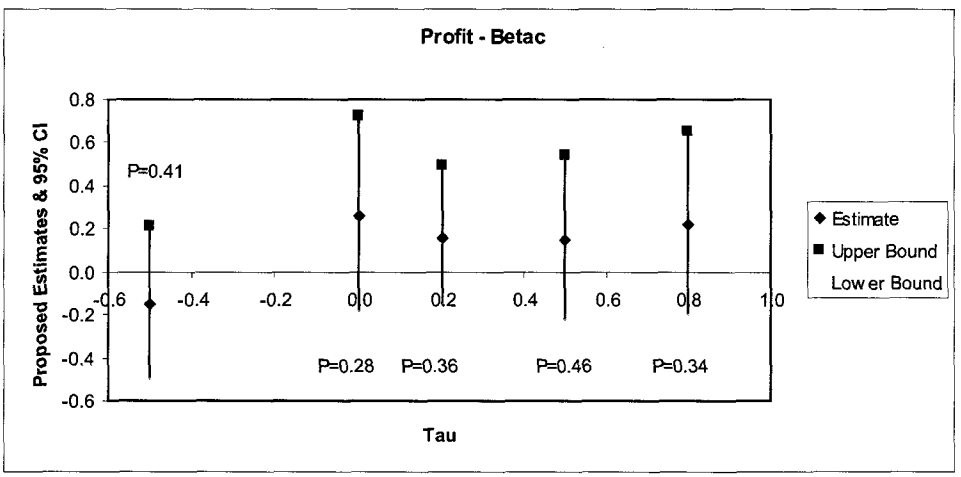

Figure 4.7: Parameter Estimates of Profitability for Acquisitions

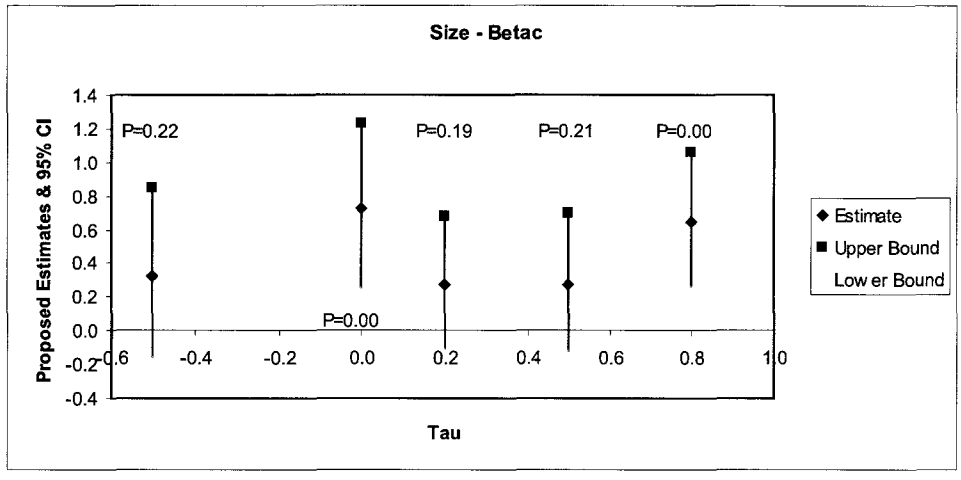

Figure 4.8: Parameter Estimates of Size for Acquisitions

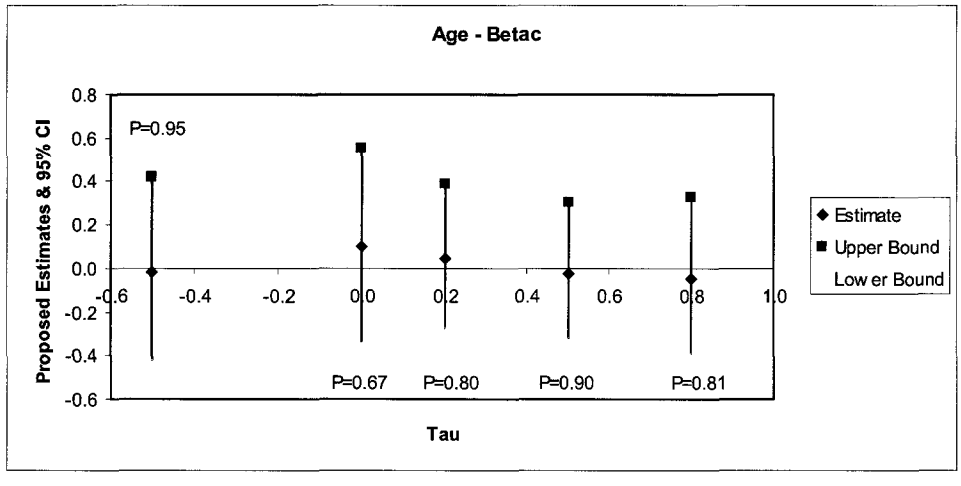

Figure 4.9: Parameter Estimates of Age for Acquisitions

Table 4.10: Results for Covariate 3 - Size for Acquisitions

	$\tau = -0.5$	$\tau = 0$	$\tau = 0.2$	$\tau = 0.5$	$\tau = 0.8$
Parameters	0.329	0.724	0.270	0.272	0.645
SE of Est	0.266	0.255	0.207	0.217	0.211
Upper Bound	0.850	1.225	0.677	0.697	1.057
Lower Bound	-0.193	0.224	-0.137	-0.152	0.232

Table 4.11: Results for Covariate 4 - Age for Acquisitions

	$\tau = -0.5$	$\tau = 0$	$\tau = 0.2$	$\tau = 0.5$	$\tau = 0.8$
Parameters	-0.013	0.100	0.043	-0.021	-0.045
SE of Est	0.220	0.232	0.174	0.163	0.189
Upper Bound	0.418	0.555	0.384	0.299	0.325
Lower Bound	-0.445	-0.356	-0.298	-0.342	-0.415

Conclusions

Our results show that both Size and Liability have significant impacts to both bankruptcy and acquisition. Besides those two factors, Growth has significant impacts to acquisition but not to bankruptcy. Specifically, large size and high liability might be associated with the high probability of bankruptcy. Large size, low liability, and high growth might be associated with the high probability of acquisition. Such association tends to vary with respect to the value of τ.

Table 4.12: Results for Covariate 5 - Liability for Acquisitions

	$\tau = -0.5$	$\tau = 0$	$\tau = 0.2$	$\tau = 0.5$	$\tau = 0.8$
Parameters	-0.361	-0.327	-0.268	0.056	-0.162
SE of Est	0.240	0.260	0.195	0.226	0.181
Upper Bound	0.110	0.183	0.114	0.498	0.192
Lower Bound	-0.832	-0.836	-0.650	-0.386	-0.516

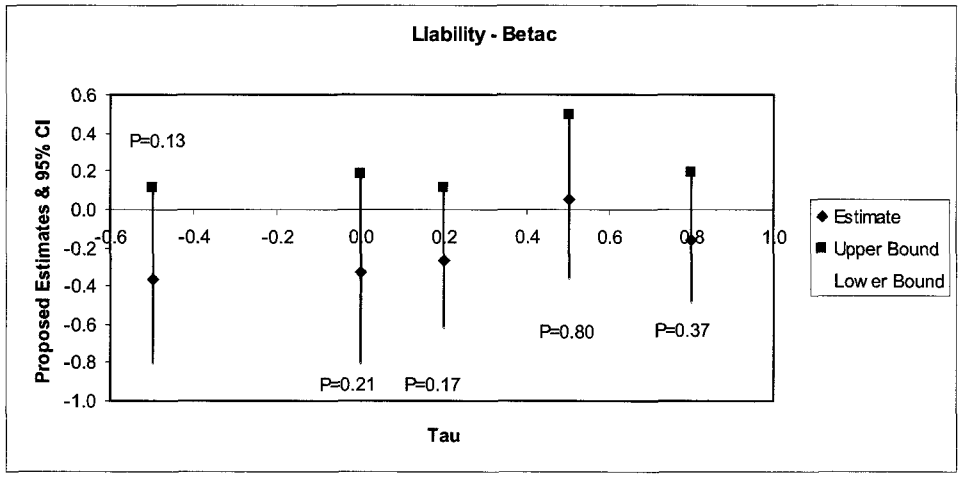

Figure 4.10: Parameter Estimates of Liability for Acquisitions

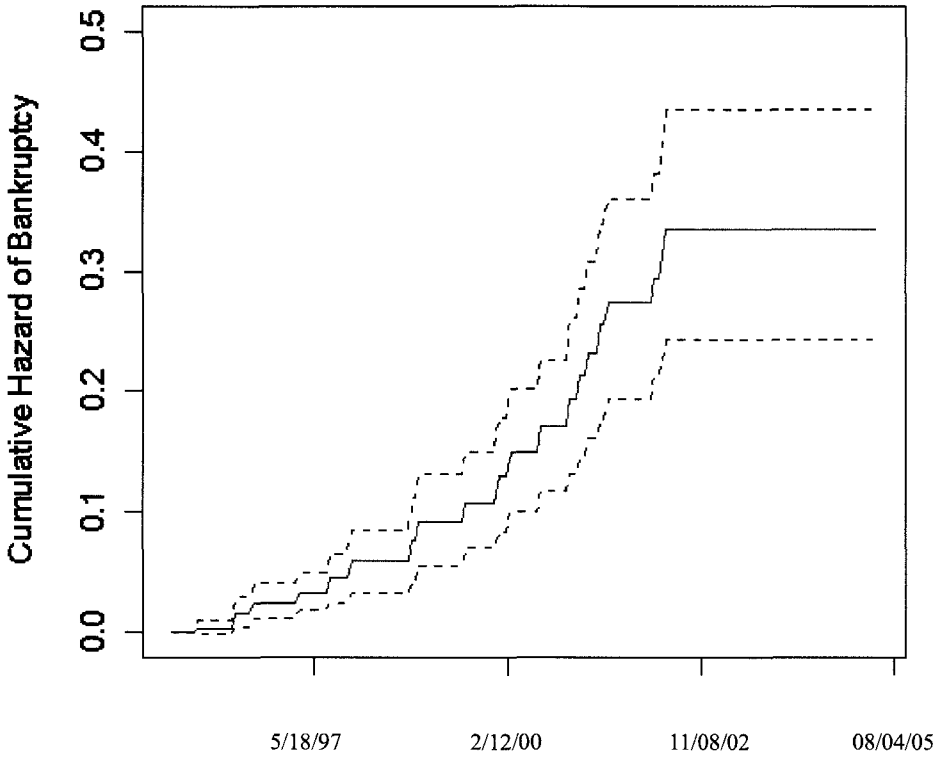

Figure 4.11: Cumulative Hazard Curve and CI Obtained Using the Proposed Method

Cumulative Hazard Curve (in Blue) and its CIs (dashed lines).

Chapter 5

A Change-Point Hazard Rate Model

When looking at Figure 4.11 of the derived cumulative hazard curve in Chapter 4, we can easily see a change-point on the curve. Two questions are naturally raised: is this change-point real (statistically significant)? If so, how to locate the change-point? The answers to these questions will be valuable to business decision making. Chapter 5 tries to provide such answers by testing the hypothesis on the existence of change-point(s) and offering effective ways to locate the change-point, when it exists.

We first review the related literature including both parametric and non-parametric approaches in Section 5.1; and then introduce our methodologies in Section 5.2. Also explained in this Section is how the iterations here differ from those in Chapter 3. Revisiting the application and results are shown in Section 5.3.

5.1 Literature review

There are both parametric and nonparametric change-point hazard rate models in literatures. In general, the parametric methods are likelihood based. Most existing methods consider neither covariates nor dependent censoring. Some relatively new methods, such as the one proposed by Wu (2003) and the one by Dupuy (2006), model covariates in the change-point hazard rate model. However, to our best knowledge, there is no published work that jointly considers independent and dependent censoring together with covariates in change-point hazard rate models.

5.1.1 The likelihood based approaches

Matthews and Farewell (1982)'s approach is popular in this field. The model they considered is:

$$\lambda(t) = \begin{cases} \lambda_1 & , \quad t \leq \theta, \\ \rho\lambda_1 & , \quad t > \theta. \end{cases}$$

This model has three parameters, λ_1, ρ and θ. Here θ represents a change-point; ρ represents the change in a hazard function before and after a change-point. Same as other early work, covariate was not included in the model. It is clear that the standard asymptotic likelihood inference on these parameters will not be applicable. Assume the null hypothesis is no change-point, i.e. $\theta = 0$. Let the log-likelihood statistic be denoted by $L(\lambda_1, \rho, \theta)$. Matthews and Farewell applied the likelihood ra-

tio test using the log-likelihood test statistic: $\Delta_0 = L(\hat{\lambda_1}, \rho\hat{\lambda}_1, \hat{\theta}) - L(\hat{\lambda}, \hat{\lambda}, 0)$, where $\hat{\lambda}$ is the maximum likelihood estimator of the failure rate in a simple exponential model. Let $\lambda_1 = 1.0$. The authors use simulations to get the asymptotic distribution of $2\Delta_0$. Both Weibull and Gamma distribution are tested using simulations. Matthews and Farewell suggest that this procedure should also be applicable to censored data. However, only independent censoring is considered.

Henderson (1990)

Henderson (1990) considers the same model as in Matthews and Farewell (1982). A test of $H_0 : \rho = 1$ against $H_1 : \rho \neq 1$ is considered with θ unknown. The Monte Carlo power and Mean Squared Error estimates are presented by simulations. Henderson shows that the adjusted log-likelihood method can be used when the likelihood ratio test is not sufficient. The adjusted method gives better power and smaller Mean Squared Error than unadjusted log-likelihood statistic.

Loader (1991)

Loader (1991) considers the following model:

$$\lambda(t) = \begin{cases} \lambda_0 & , \quad 0 \leq t < \theta, \\ \lambda_1 & , \quad t \geq \theta. \end{cases}$$

$\delta = log(\lambda_1/\lambda_0)$ is used for inference about θ and the size of the change. Loader uses the log-likelihood ratio process and the score process considered by Mathews et

al. (1985). He also extends the likelihood ratio method to find confidence regions for the change-point θ and joint confidence regions for (θ, δ). He applies the results to Stanford heart transplant data.

Dupuy (2006)

Dupuy (2006) extends the hazard function in Wu et al. (2003) by including the effect of covariates. He also allows for the time-dependent covariates. Thus, the hazard function he studies is:

$$\lambda(t|Z) = (\alpha + \theta I_{\{t>\tau\}} exp\left\{\left(\beta + \gamma I_{\{t>\tau\}}\right)^T Z(t)\right\}. \tag{5.1}$$

This paper deduces the format of the log-likelihood function for the above model. Furthermore, Newton-Raphson method can be used to get the parameter estimator for the change-point. Dupuy proved that the convergence of the estimators is warranted. And the estimators are shown to be consistent.

However, dependent censoring is not taken into consideration. Thus, Dupuy's method is not directly applicable to the question in this research.

5.1.2 Non-parametric approaches

Wu et al. (2003)

The change-point model considered in Wu et al. (2003) is:

$$\lambda(t) = (\alpha + \theta I_{\{t>\tau\}})\lambda_0(t;\gamma), \tag{5.2}$$

where $\lambda_0(\cdot;\gamma)$ is a baseline hazard function depending on an unknown parameter γ. This model is tested on many important distributions commonly used in survival analysis, for instance, exponential, Weibull, extreme and log-logistic models for which $\lambda_0(t;\gamma) = 1$, t^γ, $e^{\gamma t}$ and $t^{\gamma_1} / \left(1 + \gamma_2 t^{(\gamma_1+1)}\right)$, respectively.

Wu et al. (2003) provides a non-parametric estimator of the change-point in the context of counting process, based upon a function of Nelson-Aalen type estimator. The estimators for change-point and other parameters are shown to be consistent. Monte Carlo simulation tests are conducted, which show that the proposed procedure for estimating the change-point is effective.

Although independent censoring is considered in the proposed approach, the dependent censoring is not included. Thus, the method can not be applied to this dissertation directly.

5.2 Methodology

5.2.1 Models and parameter estimators

We are interested in testing the hypothesis on whether change-point(s) exist on the copula-based hazard curve; and if so, how to locate it.

We consider the change-point model as below:

$$\lambda(t) = \begin{cases} \lambda_0 \exp(Z_i'\beta) & , \quad 0 \le t < \theta, \\ \lambda_1 \exp(Z_i'\beta) & , \quad t \ge \theta. \end{cases}$$

It is a natural extension of Matthews and Farewell's (1982) model by including covariates. Instead of using one parameter ρ to represent the difference of hazard functions before and after the change-point, we use λ_0 and λ_1, which can be estimated in a relatively straightforward way.

To estimate the baseline hazard function, we propose:

$$\hat{\lambda}_0 = \frac{\hat{\Lambda}_0(\theta) - \hat{\Lambda}_0(\theta_{min})}{\theta - \theta_{min}} \tag{5.3}$$

$$\hat{\lambda}_1 = \frac{\hat{\Lambda}_0(\theta_{max}) - \hat{\Lambda}_0(\theta)}{\theta_{max} - \theta} \tag{5.4}$$

where θ is the change-point; and θ_{min} and θ_{max} are the starting and end time point of the study period.

$\hat{\lambda}_0$ and $\hat{\lambda}_1$ are used to represent the slope of the cumulative hazard curve. Because the hazard functions in the change-point model are special cases of the format of Equations 3.1 and 3.2, the estimates of cumulative hazard functions in Equations 3.19 and 3.20 from the regression survival analysis are used to estimate λs as a reasonable approximation. Therefore, both independent and dependent censoring can be included.

Similar to Chapter 4, when treating the dependent censoring as the event of interest, the counterparts of above equations are as below:

$$\lambda_c(t) = \begin{cases} \lambda_{0c} \exp(W_i' \beta_c) & , \quad 0 \leq t < \theta, \\ \lambda_{1c} \exp(W_i' \beta_c) & , \quad t \geq \theta. \end{cases}$$

$$\hat{\lambda}_{0c} = \frac{\hat{\Psi}_0(\theta) - \hat{\Psi}_0(\theta_{min})}{\theta - \theta_{min}} \tag{5.5}$$

$$\hat{\lambda}_{1c} = \frac{\hat{\Psi}_0(\theta_{max}) - \hat{\Psi}_0(\theta)}{\theta_{max} - \theta} \tag{5.6}$$

5.2.2 Iterations

During the iteration, in the first round, estimates of cumulative hazard functions can be obtained through Equations 3.13 and 3.14. Later on, they can be obtained through Equations 3.19 and 3.20.

The variance of $\hat{\lambda}_0$ and $\hat{\lambda}_1$ can be estimated by the bootstrap method. The corresponding 95 percent confidence interval will be $\hat{\lambda}_0 \pm 1.96 \times SE(\hat{\lambda}_0)$ and $\hat{\lambda}_1 \pm 1.96 \times SE(\hat{\lambda}_1)$, respectively. For $\hat{\lambda}_{0c}$ and $\hat{\lambda}_{1c}$, we have 95 percent confidence intervals $\hat{\lambda}_{0c} \pm 1.96 \times SE(\hat{\lambda}_{0c})$ and $\hat{\lambda}_{1c} \pm 1.96 \times SE(\hat{\lambda}_{1c})$.

In order to test the hypothesis $H_0 : \lambda_1 = \lambda_0; H_1 : \lambda_1 \neq \lambda_0$, we have the following "z_λ statistic" $= \frac{(\hat{\lambda}_1 - \hat{\lambda}_0)}{SE(\hat{\lambda}_1 - \hat{\lambda}_0)}$, which follows a normal (0,1) distribution. If the p-value is small enough (with certain significance level), we reject the null hypothesis. Otherwise, we fail to reject the null hypothesis $\lambda_1 = \lambda_0$.

5.3 Application: Revisiting the impact of the GLB Act

In this Section, we revisit the application in Chapter 4. We use the same data set but focus on identifying and locating a change-point. Section 5.3.1 explains the basic setups for this application. Results are shown in Section 5.3.2. Conclusions are included in Section 5.3.3.

5.3.1 Basic setups

Please note that the data structure and covariates used in Section 5.3 are the same as in Chapter 4.

In order to estimate λs, we let $\theta_{min} = Jan.\ 1994$ and $\theta_{max} = Dec.\ 2005$, which is the range of the time horizon considered. In order to locate the potential change-point θ, we tested December 1999. The reason is that from the cumulative hazard curve in Figure 4.11, it seems that there might be a change-point of the cumulative hazard function around December 1999. Since the true value of τ is unknown, we did a sensitivity analysis for $\tau = 0.5$ and $\tau = -0.5$, respectively.

5.3.2 Results of the tests

Test 7 with θ = December 99; τ =0.5

Table 5.1 shows the parameter estimates for both βs and λs. Note that the values

of λs are fairly small, because it represents the "daily" hazard rate of an insurance company.

The "z_λ statistic" is 0.984 for λs by coxph method; -8.603 for $\lambda_c s$ by coxph. We have $\hat{\lambda_0}$=0.00000807, $\hat{\lambda_1}$=0.0000123; $\hat{\lambda_{0c}}$=0.000140 and $\hat{\lambda_{1c}}$=0.0000328.

Similarly, the "z_λ statistic" is 1.149 for λs and -5.196 for $\lambda_c s$ by the proposed method. We have $\hat{\lambda_0}$=0.0000673, $\hat{\lambda_1}$=0.000103; $\hat{\lambda_{0c}}$=0.000197 and $\hat{\lambda_{1c}}$=0.0000489. Figure 5.1 shows the increase of the monthly hazard rate for bankruptcy based on the change-point model in Section 5.2.1, i.e. covariates are considered as well. And Figure 5.2 shows the decrease of the "monthly" hazard rate for acquisition based on the change-point model in Section 5.2.1.

Test 8 with θ = December 99; τ = -0.5

Table 5.2 shows the parameter estimates for both βs and λs.

The "z_λ statistic" is 0.919 for λs by coxph; -8.689 for $\lambda_c s$ by coxph. We have $\hat{\lambda_0}$=0.00000807, $\hat{\lambda_1}$=0.0000123; $\hat{\lambda_{0c}}$=0.000140 and $\hat{\lambda_{1c}}$=0.0000328. Note that in Table 5.2, the Standard Errors of Coxph estimates are slightly different from the results when τ=0.5 because the bootstrap method is used to calculate the SE of estimates.

Similarly, the "z_λ statistic" is 0.127 for λs and -4.955 for $\lambda_c s$ by the proposed method. We have $\hat{\lambda_0}$=0.0000313, $\hat{\lambda_1}$=0.0000342; $\hat{\lambda_{0c}}$=0.000250 and $\hat{\lambda_{1c}}$=0.0000523. Figure 5.3 shows the increase of the monthly hazard rate for bankruptcy based on the change-point model in Section 5.2.1, i.e. covariates are considered as well. And

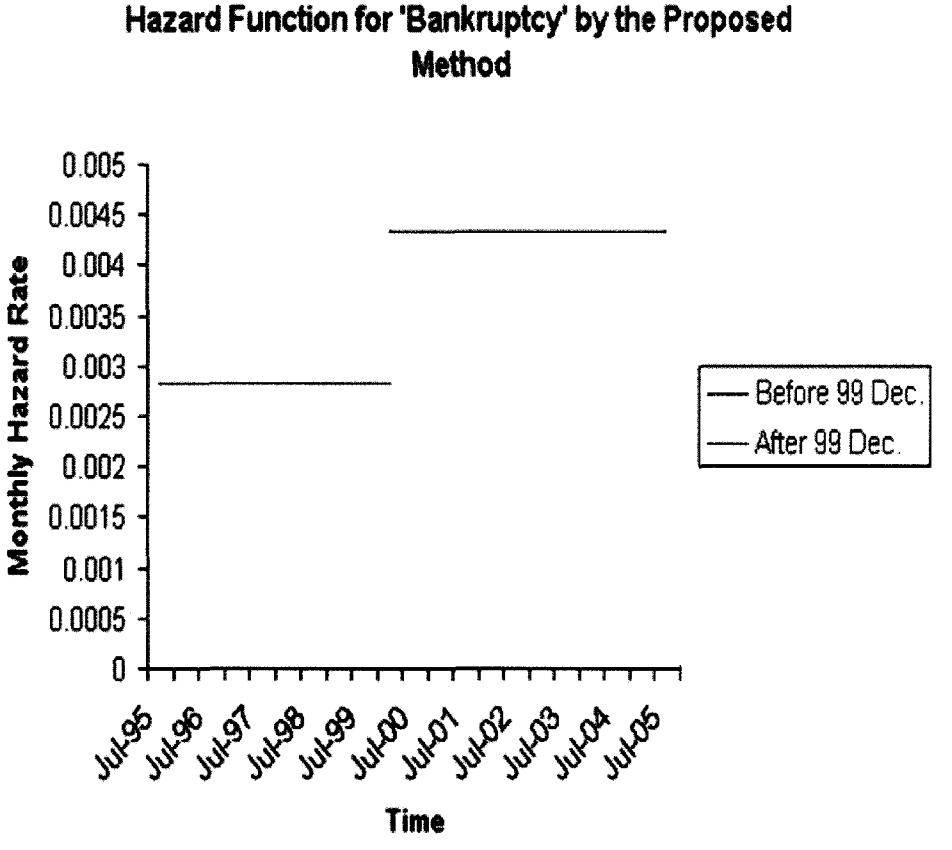

Figure 5.1: Hazard Function for Bankruptcy Obtained Using the Proposed Method with τ=0.5

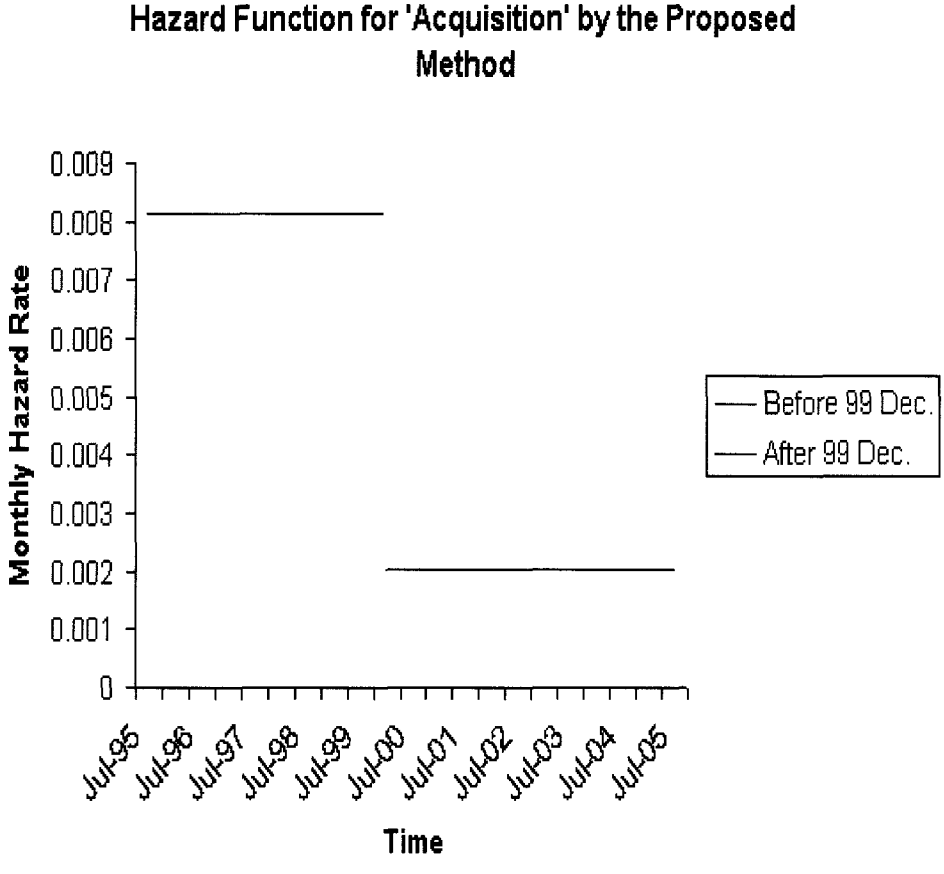

Figure 5.2: Hazard Function for Acquisition Obtained Using the Proposed Method with τ=0.5

Figure 5.4 shows the decrease of the "monthly" hazard rate for acquisition based on the change-point model in Section 5.2.1.

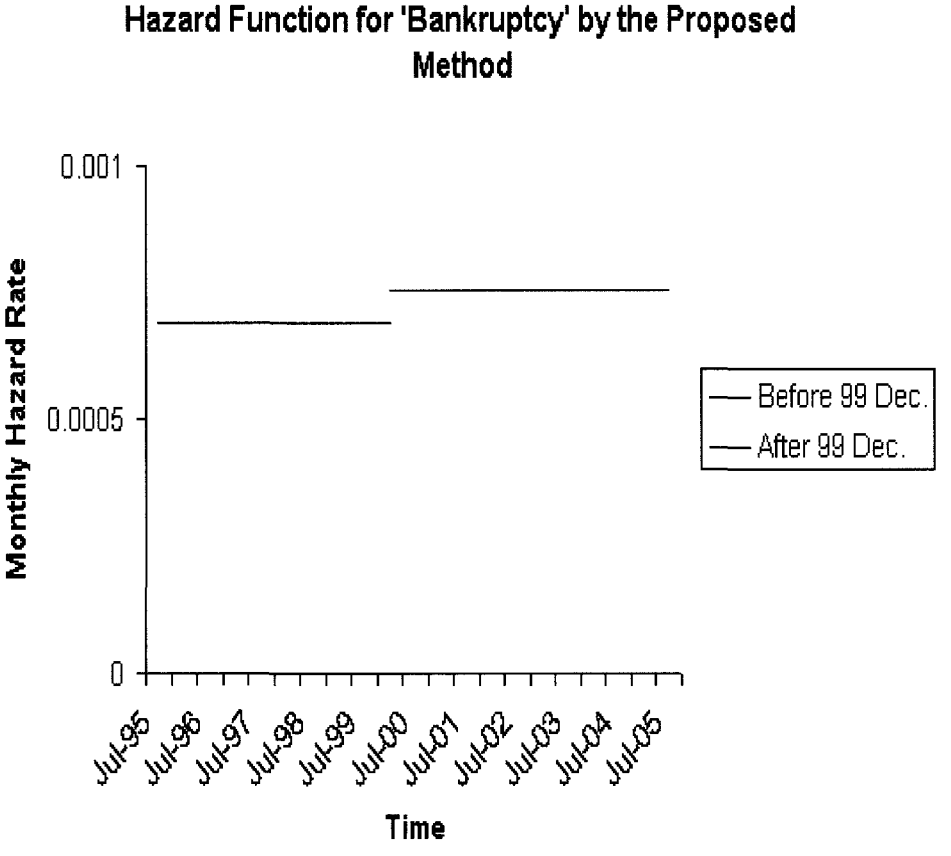

Figure 5.3: Hazard Function for Bankruptcy Obtained Using the Proposed Method with τ= - 0.5

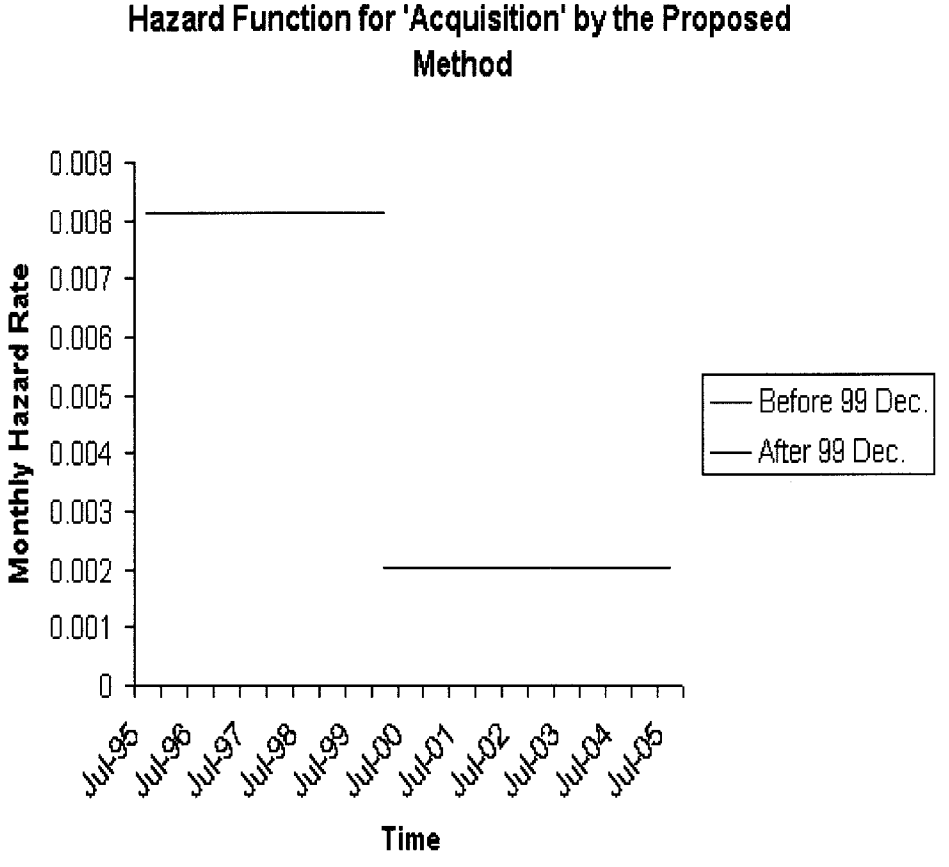

Figure 5.4: Hazard Function for Acquisition Obtained Using the Proposed Method with τ= - 0.5

5.3.3 Conclusions

For bankruptcies, the probability of going bankrupt slightly increased after the passage of the 1999 GLB Act. But the increase is not statistically significant.

For acquisitions, the probability of acquisition statistical significantly decreased after the GLB Act. Further discussions are included in Chapter 6.

Table 5.1: Results With $\tau = 0.5$ and $\theta = 99$ December

	Parameters	*Proposed Estimates*	*Proposed SE of Est*	*Coxph Estimates*	*Coxph SE of Est*
1	β_1	0.169	0.247	0.116	0.440
2	β_2	0.133	0.234	0.218	0.456
3	β_3	0.223	0.220	0.940	0.560
4	β_4	-0.118	0.186	-0.503	0.447
5	β_5	0.249	0.244	1.545	0.695
6	β_{1c}	0.172	0.216	0.335	0.228
7	β_{2c}	0.148	0.200	0.255	0.236
8	β_{3c}	0.272	0.217	0.724	0.255
9	β_{4c}	-0.021	0.163	0.100	0.232
10	β_{5c}	0.056	0.226	-0.327	0.260
11	λ_0	0.0000673	0.0000225	0.00000807	0.00000275
12	λ_1	0.000103	0.0000298	0.0000123	0.00000322
13	λ_{0c}	0.000197	0.0000346	0.000140	0.00000832
14	λ_{1c}	0.0000489	0.0000135	0.0000328	0.00000970

Table 5.2: Results with $\tau = -0.5$ and $\theta = 99$ December

	Parameters	*Proposed Estimates*	*Proposed SE of Est*	*Coxph Estimates*	*Coxph SE of Est*
1	β_1	-0.870	0.728	0.116	0.440
2	β_2	-0.832	0.564	0.218	0.456
3	β_3	0.546	0.798	0.940	0.560
4	β_4	0.042	0.691	-0.503	0.447
5	β_5	0.438	0.709	1.545	0.695
6	β_{1c}	0.061	0.218	0.335	0.228
7	β_{2c}	-0.154	0.182	0.255	0.236
8	β_{3c}	0.329	0.249	0.724	0.255
9	β_{4c}	-0.013	0.233	0.100	0.232
10	β_{5c}	-0.361	0.249	-0.327	0.260
11	λ_0	0.0000313	0.0000287	0.00000807	0.00000280
12	λ_1	0.0000342	0.0000355	0.0000123	0.00000322
13	λ_{0c}	0.000250	0.0000474	0.000140	0.00000834
14	λ_{1c}	0.0000523	0.0000187	0.0000328	0.00000928

Chapter 6

Conclusions and Future Work

6.1 Conclusions

As shown in Chapter 5, the probability of bankruptcy stays about the same after the passage of the 1999 GLB Act. On the other hand, the probability of acquisitions decreases.

There are a variety of possible reasons behind the conclusions we reached. For instance, as the business in either banking or insurance industry is complex in nature, not many bankers or insurers have the necessary knowledge to effectively operate the combined business. Thus, they are hesitant to initiate acquisitions across industries.

Another argument is that the major acquisitions had started prior to the passage of the 1999 GLB Act. For example, Citibank merged with Traveler's Group in 1998. It was speculated that the passage of the GLB Act was designed to justify those

acquisitions. If it is true, it might have diluted the number of acquisitions after the Act (Radcliffe, 2005).

Another viewpoint is raised by Yeager (2007). He argued that "It is probably too early to assess the long-term impacts of the financial modernization legislation ", even though "early indications suggest that the legislation will have only modest effects on the financial services industry".

6.2 Future work

Based on the findings of this dissertation, the future research can focus on studying more broadly about the change in the value of τ after the passage of the GLB Act. In addition, we can collect more data to assess the long term effects of GLB, as Yeager (2007) suggested.

Appendix A

Plots of Copula Functions

A.1 Frank Copula

A.2 Clayton Copula

A.3 Gumbel-Hougaard Copula

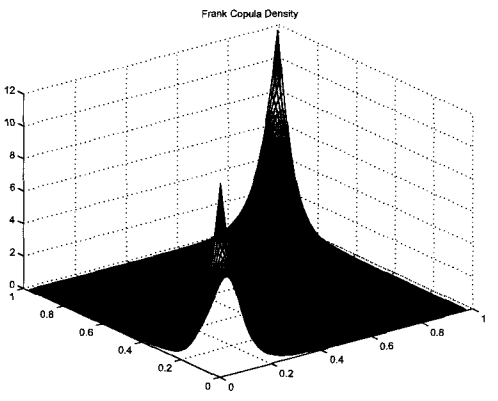

Figure A.1: Density Plot of a Frank Copula ($\tau = 0.8$)

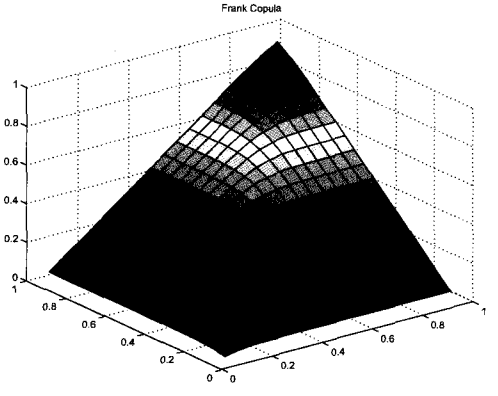

Figure A.2: Frank Copula CDF Plot ($\tau = 0.8$)

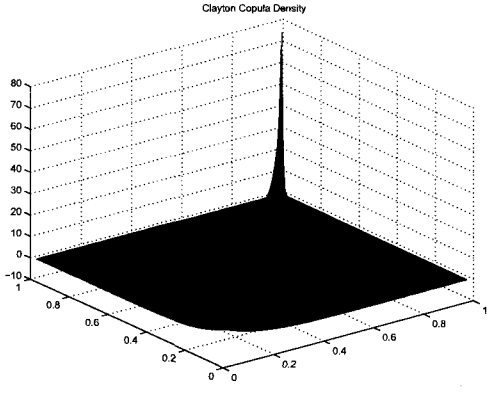

Figure A.3: Density Plot of a Clayton Copula ($\tau = 0.8$)

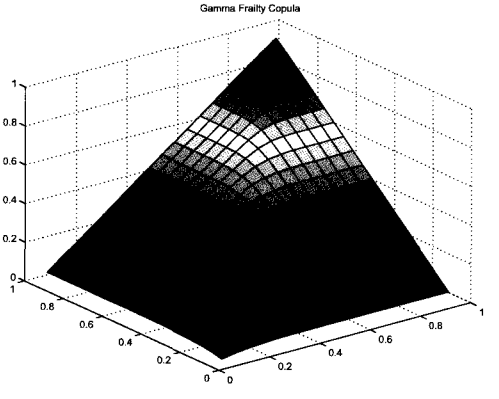

Figure A.4: Clayton Copula CDF Plot ($\tau = 0.8$)

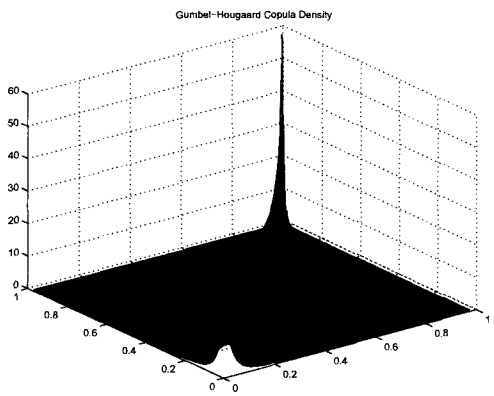

Figure A.5: Density Plot of a Gumbel-Hougaard Copula ($\tau = 0.8$)

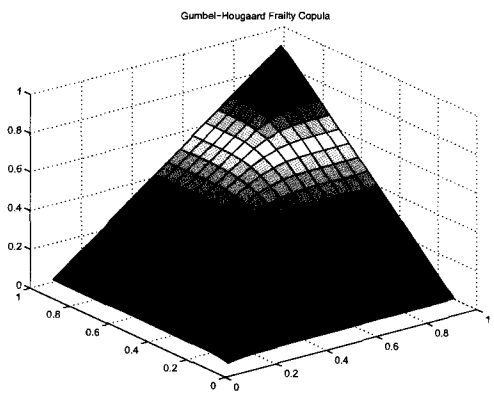

Figure A.6: Gumbel-Hougaard Copula CDF Plot ($\tau = 0.8$)

Appendix B

R Codes

B.1 Generate u and v' as shown in Section 3.4

```
## Copula simulation (From Nelsen (1986) p.3282)

copula.sim < − function(nn,copula.type,alphaa) {

if (copula.type == 1){ #Frank.copula.type=1

xx1 < − as.double(runif(nn,0,1))

vv < − as.double(runif (nn,0,1))

numm < − as.double(vv*(alphaa - 1))

denomm < − as.double((alphaa^xx1)*(1-vv)+ vv)

xx2 < − as.double(logb(1 + numm/denomm,alphaa)) }

par(mfrow=c(1,1))

plot(xx1,xx2,xlim=c(0,1),ylim=c(0,1),main"Simulated Copula Data")
```

```
x.mat < − cbind(xx1,xx2)

return(x.mat) }
```

B.2 Generate t_T and t_C as shown in Section 3.4

```
observ.sim < − function(CDFt,CDFr,beta.org,betac.org) { #1

denomt < − rep(NA,1,m)

denomtc < − rep(NA,1,m)

TT < − rep(NA,1,n)

TC < − rep(NA,1,n)

for (i in 1:n) { #2

denomt[i] < − exp((beta.org)% * %V[,i])

denomtc[i] < − exp((betac.org)% * %V[,i])

TT[i] < − sqrt((-2*(log(1-CDFt[i])))/(denomt[i]))

TC[i] < − (-2*(log(1-CDFr[i])))/(denomtc[i]) } #2

results < − rbind(TT,TC)

return(results) } #1
```

B.3 Calculate $\hat{S}_0^{(0)}(t)$ as shown in Equation 3.13

```
## base line survival function

surv.fnc < − function(obs.ordn,delta1n,delta2n,covv,betaa.f,betac.f){ #10
```

```
cum.hazd < − matrix(NA,1,n);

cum.hazdc < − matrix(NA,1,n);

item < − matrix(NA,1,n)

itemc < − matrix(NA,1,n)

item.den < − matrix (0,1,n)

surv.fnc.s < − matrix (NA,1,n)

surv.fnc.r < − matrix (NA,1,n)

for (i in 1:n) { #11 i represents t

if (delta1n[i]!=0){ #13

for (k in i:n){#14

item.den[i] < − item.den[i] + exp(betaa.fitem[i] < − delta1n[i]/(item.den[i]);

} else { #13 & 15

item[i] < − 0 } #15

if (delta2n[i]!=0){ #13

for (k in i:n){#14

item.denc[i] < − item.denc[i] + exp(betac.f% ∗ %covv[,k]) } #14

itemc[i] < − delta2n[i]/(item.denc[i])

} else { #13 & 15

itemc[i] < − 0} }#11

for (j in 1:n) {#16

cum.hazd[j] < − cumsum(item)[j] ;
```

```
cum.hazdc[j] < − exp(-(cum.hazd[j])) ;

surv.fnc.r[j] < − exp(-(cum.hazdc[j])) } #16

surv.fnc.f < − rbind(surv.fnc.s,surv.fnc.r)

return(surv.fnc.f) } #10
```

B.4 Calculate $\hat{P}_i^{(m)}(x_j)$ as shown in Equation 3.15

```
P.fnc < − function(obs.ordn,delta1n,delta2n,covv,rww,rwwc) { #50

# weight,weightc;P;Pc

PP.f < − matrix(NA,n,n)

PPc.f < − matrix(NA,n,n)

for ( i in 1:n){ #52

for ( j in 1:n) { #53 # time j

if (j >1 ) {#53.1

PP.f[i,j] < − rww[i,(j-1)] - rww[i,j];

PPc.f[i,j] < − rwwc[i,(j-1)] - rwwc[i,j];} #53.1

if (j ==1) {

PP.f[i,j] < − 1- rww[i,j];

PPc.f[i,j] < − 1- rwwc[i,j]; } } #53

if (delta1n[i]==1) {PP.f[i,i] < − 1 ; if (i ¡ n) PP.f[i,i+1] = 0}

if (delta2n[i]==1) {PPc.f[i,i] < − 1; if (i ¡ n) PPc.f[i,i+1] = 0} } #52

return(PP.f,PPc.f) } #50
```

where rww is calculated as below.

```
cy < − function(x,y,alphaa){
a < − alphaa^x
b < − alphaa^y
cyy < − b*(a-1)/(alphaa-1+(a-1)*(b-1))
return(cyy) }
cx < − function(x,y,alphaa){
a < − alphaa^x
b < − alphaa^y
cxx < − a*(b-1)/(alphaa-1+(a-1)*(b-1))
return(cxx) }
rw.fnc < − function(delta1n,delta2n,ff,gg,alphaa) {#30
rw.f < − matrix(NA,n,n)
rwc.f < − matrix(NA,n,n)
for(i in 1:n) {#33 # subject i
for (j in 1:n) {#34 # time j
if (i >= j) {#35
rw.f[i,j] < − 1;
rwc.f[i,j]< − 1}#35
if (i < j) { #36
if (delta1n[i]==1) {#37
```

```
rw.f[i,j] < − 0;
if (gg[i,j] <1){ #38
rwc.f[i,j] < − (1-cx(ff[i,i],gg[i,j],alphaa))/(1-cx(ff[i,i],gg[i,i],alphaa));
} else {rwc.f[i,j] < − 0} }#37
if (delta2n[i]==1){ #40
rwc.f[i,j] < − 0;
if (ff[i,j] < 1) { #41
rw.f[i,j] < − (1-cy(ff[i,j],gg[i,i],alphaa))/(1-cy(ff[i,i],gg[i,i],alphaa));
} else {rw.f[i,j] < − 0} }#40
if (delta1n[i]!=1 && delta2n[i]!=1){ #43
rw.f[i,j] < − 0;
rwc.f[i,j] < − 0; } } }}#33,34
return(rw.f,rwc.f) } #30
```

B.5 Calculate $\hat{S}_0^{(m)}(t)$ as shown in Equation 3.19

```
surv.fnc.n < − function(obs.ordn,delta1n,delta2n,covv,betaa.f,
betac.f,rww,rwwc,PP,PPc) {#10
cum.hazd < − matrix(NA,1,n);
cum.hazdc < − matrix(NA,1,n);
item < − matrix(NA,1,n)
itemc < − matrix(NA,1,n)
```

```
item.den < − matrix(0,1,n)

item.denc < − matrix(0,1,n)

item.num < − matrix(0,1,n)

item.numc < − matrix(0,1,n)

surv.fnc.s < − matrix(NA,1,n)

surv.fnc.r < − matrix(NA,1,n)

## using two terms for S(t) and R(t)

for (j in 1:n) { #11 j represents t

for (k in 1:n){#12

item.den[j] < − item.den[j] + (rww[k,j])*(exp(betaa.f% * %covv[,k]));

item.denc[j] < − item.denc[j] + (rwwc[k,j])*(exp(betac.f% * %covv[,k]))} #12

for (y in 1:j) { #13

item.num[j] < − item.num[j] + PP[y,j];

item.numc[j] < − item.numc[j] + PPc[y,j]; } #13

item[j] < − item.num[j]/item.den[j];

itemc[j] < − item.numc[j]/item.denc[j];

cum.hazd[j] < − cumsum(item)[j]

cum.hazdc[j] < − cumsum(itemc)[j]

surv.fnc.s[j] < − exp(-cum.hazd[j])

surv.fnc.r[j] < − exp(-cum.hazdc[j]) } #11

surv.fnc.f < − rbind(surv.fnc.s,surv.fnc.r)
```

```
return(surv.fnc.f) } #10
```

B.6 Calculate the likelihood function as shown in Equation 3.12

```
likel.h <- function(covv,delta1n,delta2n,betaa.f,betac.f,rww,rwwc,PP,PPc){#61
betaa.f <- matrix(betaa.f,1,m)
betac.f <- matrix(betac.f,1,m)
zbetaa <- rep(NA,n)
zbetacc <- rep(NA,n)
for (z in 1:n){
zbetaa[z] <- betaa.f% * %covv[,z]
zbetacc[z] <- betac.f% * %covv[,z] }
weight <- matrix(NA,n,n)
weightc <- matrix(NA,n,n)
for ( i in 1:n){ #62
for ( j in 1:n) { #63
weight[i,j] <- (exp(zbetaa[i]))*rww[i,j];
weightc[i,j] <- (exp(zbetacc[i]))*rwwc[i,j] } } #62
denom <- matrix(NA,n,n)
denom.term <- rep(0,n)
```

```
denomc < − matrix(NA,n,n)

denomc.term < − rep(0,n)

loglik < − 0

der.1 < − matrix(0,m,1)

der.2 < − matrix(0,m,m)

loglikc < − 0

der.1c < − matrix(0,m,1)

der.2c < − matrix(0,m,m)

for (j in 1:n){ #64 j-time

loglik.i < − 0

loglikc.i < − 0

der.1.i < − matrix(0,m,1)

der.2.i < − matrix(0,m,m)

der.1c.i < − matrix(0,m,1)

der.2c.i < − matrix(0,m,m)

zye < − matrix(0,m,1)

zzye < − matrix(0,m,m)

zyec < − matrix(0,m,1)

zzyec < − matrix(0,m,m)

for (k in 1:n){ #65.1

denom.term[j] < − denom.term[j] + weight[k,j]
```

```
denomc.term[j] <- denomc.term[j] + weightc[k,j]

zye <- zye+ covv[,k]*weight[k,j]

zzye <- zzye+ (covv[,k])%*%(t(covv[,k]))*weight[k,j]

zyec <- zyec + covv[,k]*weightc[k,j]

zzyec <- zzyec + (covv[,k])%*%(t(covv[,k]))*weightc[k,j] }#65.1

for (i in 1:j){ #65

denom[i,j] <- denom.term[j]

der.1.i <- der.1.i + PP[i,j]*(covv[,i]-(zye)/(denom[i,j]))

der.2.i <- (der.2.i + (PP[i,j]*((zye%*%t(zye))-(denom[i,j])*(zzye))

/((denom[i,j])^2)))

denomc[i,j] <- denomc.term[j]

der.1c.i <- der.1c.i + PPc[i,j]*(covv[,i]-(zyec)/(denomc[i,j]))

der.2c.i <- (der.2c.i + (PPc[i,j]*((zyec%*%t(zyec))-(denomc[i,j])*(zzyec))

/((denomc[i,j])^2))) } #65

loglik <- loglik + loglik.i

der.1 <- der.1 + der.1.i

der.2 <- der.2 + der.2.i

loglikc <- loglikc + loglikc.i

der.1c <- der.1c + der.1c.i

der.2c <- der.2c + der.2c.i }#64

return(der.1,der.2,der.1c,der.2c,loglik,loglikc) } #61
```

Bibliography

[1] A. Akhigbe and A. M. Whyte. The market's assessment of the financial services modernization act of 1999. *The Financial Review*, 36:119–138, 2001.

[2] B. Amoako-Adu and B. Smith. The wealth effects of deregulation of canadian financial institutions. *Journal of Banking and Finance*, 19:1211–1236, 1995.

[3] N. Breslow. Discussion of Dr. cox's paper "regression models and life tables". *Journal of the Royal Statistical Society, Series B*, 34:216–217, 1972.

[4] N. Breslow. Covariance analysis of censored survival data. *Biometrics*, 30:89–99, 1974.

[5] K. Carow and R. Heron. The interstate banking and branching efficiency act of 1994: a wealth event for acquisition targets. *Journal of Banking and Finance*, 22(2):175–196, 1998.

[6] K. Carow and R. Heron. Capital market reactions of the passage of the financial services modernization act of 1999. *Quarterly Review of Economics and Finance*, 42(summer):463–485, 2002.

[7] G. Casella and R.L. Berger. *Statistical inference, Second edition.* Duxbury, California, 2002.

[8] U. Cherubini, E. Luciano, and W. Vecchiato. *Copula methods in finance.* Wiley, New Jersey, 2004.

[9] M. M. Cornett and H. Tehranian. Stock market reactions to the depository institutions deregulation and monetary control act of 1980. *Journal of Banking and Finance*, 13:81–100, 1989.

[10] C. Corrado. A nonparametric test for abnormal security price performance in event studies. *Journal of Financial Economics*, 23(August):385–395, 1989.

[11] A. P. Dempster, N. M. Laird, and D. B. Rubin. Maximum likelihood estimation for incomplete data via the EM algorithms (with discussion). *Journal of the Royal Statistical Society, Series B*, 39:1–38, 1977.

[12] J. F. Dupuy. Estimation in a change-point hazard regression model. *Statistics & Probability Letters*, 76:182–190, 2006.

[13] B. Efron. *The two sample problems with censored data. In Proceedings of the 5th Berkeley Symposium (Volume 4).*

[14] B. Efron. Bootstrap methods: Another look at the jackknife. *The Annuals of Statistics*, 7:1–26, 1979.

[15] A.F. Gretchen. *Estimating marginal survival in the presence of dependent and independent censoring: with applications to dividend initiation policy.* Rice University, Texas, 2004.

[16] S. E. Harrington. Insurance rate regulation in the 20th century. *Journal of Insurance Regulation*, Winter:204–217, 2000.

[17] J. J. Heckman and B. E. Honore. The identifiability of the competing risks model. *Biometrika*, 76:325–330, 1989.

[18] X. L. Huang and N. Zhang. Regression survival analysis with an assumed copula for dependent censoring: a sensitivity analysis approach. *Biometrics*, 64:1090–1099, 2008.

[19] J.D. Kalbfleisch and Prentice R.L. *The statistical analysis of failure time data.* Wiley-Interscience, New Jersey, 2002.

[20] J. Kiefer and J. Wolfowitz. Consistency of the maximum likelihood estimator in the presence of infinitely many incidental parameters. *The Annals of Mathematical Statistics, Part A: Theory and Methods*, 15:3277–3285, 1956.

[21] J.P. Klein and M.L. Moeschberger. *Survival analysis.* Springer, New York, 2003.

[22] D. X. Li. On default correlation: a copula function approach. *Journal of Fixed Income*, 9, No. 4:43–54, March 2000.

[23] A. A. Mamun, M. K. Hassan, and V. S. Lai. The impact of the gramm-leach-bliley act on the financial services industry. *Journal of Economics and Finance*, 28 (Fall):333–347, 2004.

[24] D. Marlett, C. Pacini, and W. Hillison. The market impact of the financial services modernization act of 1999: differences between life insurers and property-liability insurers. *Journal of Insurance Issues*, 26:69–92, 2003.

[25] J. Mckinney. More bank for your buck? - how the gramm-leach-bliley act will affects personal finance management - industry overview. *http://www.highbeam.com/doc/1G1-61834972.html*, 2000.

[26] R.G. Miller Jr. *Survival analysis.* Wiley-Interscience, Massachusetts, 1998.

[27] R. B. Nelsen. Properties of a one-parameter family of bivariate distributions with specified marginals. *Communications in Statistics, Part A: Theory and Methods*, 15:3277–3285, 1986.

[28] R.B. Nelsen. *An introduction to copulas.* Springer, New York, 2006.

[29] C. Pacini and D. Marlett. The market valuation and trading volume effects of the creation of the Florida hurricane catastrophe fund on property-liability insurers. *Journal of Business Finance & Accounting*, 28 (3&4):407–445, 2001.

[30] Y. Park, L. Tian, and L. J. Wei. One- and two-sample nonparametric inference procedures in the presence of a mixture of independent and dependent censoring. *Biostatistics*, 7:252–267, 2006.

[31] L. Peng and J. P. Fine. Regression modeling of semi-competing risks data. *Biometrics*, 63:96–108, 2007.

[32] R. Radcliffe. Effect of the gramm-leach-bliley act on financial services. *http://www.helium.com/items/302259-effect-of-the-gramm-leach-bliley-act-on-financial-services*, 2005.

[33] D. O. Scharfstein and J. M. Robins. Estimation of the failure time distribution in the presence of informative censoring. *Biometrika*, 89:617–634, 2002.

[34] D. O. Scharfstein, J. M. Robins, W. Eddings, and A. Rotnitzky. Inference in randomized studies with informative censoring and discrete time-to-event endpoints. *Biometrics*, 57:404–413, 2001.

[35] F. Siannis. Applications of a parametric model for informative censoring. *Biometrics*, 60:704–714, 2004.

[36] F. Siannis, J. Copas, and G. Lu. Sensitivity analysis for informative censoring in parametric survival models. *Biostatistics*, 6:77–91, 2005.

[37] E. Slud. *Nonparametric identifiability of marginal survival distributions in the presence of dependent competing risks and a prognostic covariate, in Survival*

Analysis: State of the Art, J. P. Klein and P. K. Goel (Eds.), 355-368. Kluwer Academic, Dordrecht, 1992.

[38] A. B. Troxel, G. Ma, and D. F. Heitjan. An index of local sensitivity to nonignorability. *Statistica Sinica*, 14:1221–1237, 2004.

[39] W. Y. Tsai and J. Crowley. A large sample study of the generalized maximum likelihood estimators from incomplete data via self-consistency. *The Annals of Statistics*, 13:1317–1334, 1985.

[40] A. Tsiatis. A nonidentifiability aspect of the problem of competing risks. *Proceedings of the National Academy of Science, U.S.A.*, 72:20–22, 1975.

[41] T. Wagner. Insurance rating bureaus. *Journal of Insurance Regulation*, Winter:189–202, 2000.

[42] C. Q. Wu, L. C. Zhao, and Y. H. Wu. Estimation in change-point hazard function models. *Statistics & Probability Letters*, 63:41–48, 2003.

[43] T. Yeager. Impact of the gramm-leach-bliley act. *http://dailyheadlines.uark.edu/9957.htm*, 2007.

[44] A. Zellner. An efficient method of estimating seemingly unrelated regression and tests of aggregation bias. *Journal of American Statistical Association*, 57:348–368, 1962.

[45] J. Zhang and D. F. Heitjan. A simple local sensitivity analysis tool for nonignorable coarsening: Application to dependent censoring. *Biometrics*, 62:1260–1268, 2006.

[46] M. Zheng and J. P. Klein. A self-consistent estimator of marginal survival functions based on dependent competing risk data and an assumed copula. *Communications in Statistics, Part A - Theory and Methods*, 23:2299–2311, 1994.

[47] M. Zheng and J. P. Klein. Estimates of marginal survival for dependent competing risks based on an assumed copula. *Biometrika*, 82:127–138, 1995.