Distribution Agreement

In presenting this thesis or dissertation as a partial fulfillment of the requirements for an advanced degree from Emory University, I hereby grant to Emory University and its agents the non-exclusive license to archive, make accessible, and display my thesis or dissertation in whole or in part in all forms of media, now or hereafter known, including display on the world wide web. I understand that I may select some access restrictions as part of the online submission of this thesis or dissertation. I retain all ownership rights to the copyright of the thesis or dissertation. I also retain the right to use in future works (such as articles or books) all or part of this thesis or dissertation.

Signature:

__________________________  ______________
Ming Zhu                    Date
Survival Analysis with Covariate Measurement Error

By
Ming Zhu
Doctor of Philosophy
Biostatistics

Yijian Huang, Ph.D.
Advisor

Nancy G. Kutner, Ph.D.
Committee Member

Brent A. Johnson, Ph.D.
Committee Member

Robert H. Lyles, Ph.D.
Committee Member

Accepted:

Lisa A. Tedesco, Ph.D.
Dean of the James T. Laney School of Graduate Studies

Date
Survival Analysis with Covariate Measurement Error

By

Ming Zhu
B.S., University of Science and Technology of China, 2002
M.S., National University of Singapore, 2004

Advisor: Yijian Huang, Ph.D.

An abstract of
A dissertation submitted to the Faculty of the
James T. Laney School of Graduate Studies of Emory University
in partial fulfillment of the requirements for the degree of
Doctor of Philosophy
in Biostatistics
2011
Abstract

In many medical research studies, survival time is typically the primary outcome of interest. The Cox proportional hazards model and the accelerated failure time model are two popular methods to investigate the relationship between covariates and possibly right-censored survival time. However, in many clinical trials, the true covariates may not always be accurately measured due to natural biological fluctuation or instrument error. For regression analysis in general, naively using mismeasured covariates in conventional inference procedures may incur substantial estimation bias. In this dissertation research, we aim to resolve several issues in survival data analysis with covariate measurement error, with particular emphasis on the aforementioned two models.

The first topic focuses on the analysis of recurrent events data, in which the event of interest may occur more than once for each subject during the follow up period. We propose an estimation procedure for recurrent events data under the accelerated failure time model in which some covariates are not accurately measured. With replicated mismeasured covariates available, the proposed estimation procedure requires no distributional assumptions on either the true covariates or the error except for the boundedness of the latter. The resulting regression coefficient estimators are shown to be consistent and asymptotically normal. The performance of the proposed procedure is investigated by extensive numerical studies with practical sample size. In addition, an application to data from a clinical trial is provided to illustrate the proposed method.

The second topic considers the Cox proportional hazards model for univariate survival data. In the presence of covariate measurement error, several functional modeling methods have been proposed under the situation where the distribution of the measurement error is known. Among them are parametric corrected score (Nakamura, 1992) and conditional score (Tsiatis & Davidian, 2001). Although both methods are consistent, each suffers from severe problem of multiple roots or absence of appropriate root when the measurement error is substantial. The problem persists even when the sample size is practically large. We conduct a detailed investigation on the pathological behaviors of parametric corrected score and propose an approach of incorporating additional estimating functions to remedy these pathological behaviors. The estimation and inference are then accomplished by means of quadratic inference function. Extensive simulation studies are conducted to evaluate the performance of proposed method.

In the third topic, we consider the Cox proportional hazards model with covariate measurement error where the error distribution is completely unspecified, but replicated mismeasured covariates are available instead. A consistent nonparametric corrected score (Huang & Wang, 2000) has been proposed for Cox proportional hazards model with replicated mismeasured covariates. But it also suffers from finite-
sample pathological behaviors similar to that of the parametric corrected score when the measurement error is substantial. To address this issue, we develop a similar technique as in the second topic for the nonparametric corrected score and evaluate its performance by extensive simulation study.
Survival Analysis with Covariate Measurement Error

By

Ming Zhu
B.S., University of Science and Technology of China, 2002
M.S., National University of Singapore, 2004

Advisor: Yijian Huang, Ph.D.

A dissertation submitted to the Faculty of the
James T. Laney School of Graduate Studies of Emory University
in partial fulfillment of the requirements for the degree of
Doctor of Philosophy
in Biostatistics
2011
Acknowledgement

I would like to express my deepest gratitude to my advisor, Dr. Yijian Huang, for his guidance, encouragement and great patience over the past three years. I always feel fortunate to have him as my mentor and advisor who showed me the ways of conducting scientific research, and who guided me through many difficult times. Without his unconditional support and enormous help, I would never be able to reach where I am.

I would like to thank Dr. Nancy Kutner, Dr. Brent Johnson, and Dr. Robert Lyles for their time in serving on my dissertation committee. Their constructive comments and suggestions led to substantial improvement in this work.

The faculty and staff in the Department of Biostatistics and Bioinformatics provided great support in my graduate study. I am indebted to Dr. Limin Peng, who provided me with my first research experience at Emory, to Dr. Qi Long and Dr. Lawrence Phillips, who provided financial support and exciting collaboration experience in diabetic research, and to Kirk Easley, who guided me on statistical consulting. I also want to thank all the administrative staff in the department, who have helped take care of many things for me.

Finally, my most heartfelt gratitude goes to my parents and my wife Li Sun for their faith in me and their support for the past five years, without which this work would not have been possible.
Contents

1 Introduction ........................................ 1
   1.1 Background ...................................... 1
   1.2 Motivating Examples ............................... 3
      1.2.1 Nutritional Prevention of Cancer (NPC) Trial ........ 3
      1.2.2 AIDS Clinical Trial Group (ACTG) 175 Study .......... 5
   1.3 Covariate Measurement Error .................... 7
   1.4 Functional Modeling and Structural Modeling .......... 9
   1.5 Measurement Error Techniques in Survival Analysis ...... 12
      1.5.1 Recurrent Events Data ......................... 12
      1.5.2 Univariate Survival Data ....................... 13
   1.6 Outline ........................................... 14

2 Accelerated Failure Time Model for Recurrent Events With Errors in Covariates 16
   2.1 Introduction ....................................... 16
   2.2 Inference Procedure ............................... 18
      2.2.1 The Model and Data ............................. 18
      2.2.2 Proposed Estimation Procedure ................. 19
      2.2.3 Asymptotic Properties ......................... 21
4.2 Nonparametric Corrected Score and Pathological Behaviors 76
   4.2.1 Nonparametric Corrected Score 76
   4.2.2 Pathological Behaviors 77
4.3 Improving Nonparametric Corrected Score 80
4.4 Numerical Studies 83
   4.4.1 Simulations 83
   4.4.2 Application to ACTG 175 data 88
4.5 Discussion 92

5 Summary and Future Work 93
   5.1 Summary 93
   5.2 Future Work 95
List of Figures

1.1 Replicated baseline CD4 count measurements from 1,036 antiretroviral-naive patients in the ACTG 175 study. .......................... 6

3.1 Observed root patterns of the parametric corrected score $\eta(b)$. The true $\beta$ is -1 and these two corrected curves correspond to the same profile score (with true covariates). Portion of a corrected curve is thickened to indicate negative derivative. .......................... 47

3.2 Observed root patterns of the conditional score estimating function. The true $\beta$ is -1. .......................... 49

3.3 Parametric corrected score and additional estimation function based on equation (3.7) for a single-covariate model with $\beta = -1$. .......................... 54

3.4 Quadratic inference functions for the two datasets in Figure 3.1. .......................... 57

3.5 Quantile-quantile plots for $\beta$ in the single-covariate models with 20% censoring rate, where $\beta = -1$. Red, yellow, green, blue, and black correspond to sample sizes 100, 200, 400, 800, and 1,600. .......................... 63

3.6 Quantile-quantile plots for $\beta$ in the single-covariate models with 60% censoring rate, where $\beta = -1$. Red, yellow, green, blue, and black correspond to sample sizes 100, 200, 400, 800, and 1,600. .......................... 64

4.1 Observed root patterns of the nonparametric corrected score estimating function. The true $\beta$ is -1. .......................... 79

4.2 Quadratic inference functions for the data sets in Figure 4.1. .......................... 82

4.3 Quantile-quantile plots for $\beta$ in the single-covariate models with 20% censoring rate, where $\beta = -1$. Red, yellow, green, blue, and black correspond to sample sizes 100, 200, 400, 800, and 1,600. .......................... 86
4.4 Quantile-quantile plots for $\beta$ in the single-covariate models with 60% censoring rate, where $\beta = -1$. Red, yellow, green, blue, and black correspond to sample sizes 100, 200, 400, 800, and 1,600.
List of Tables

2.1 Simulation summary statistics of the estimators in the single-covariate model ........................................... 25
2.2 Simulation summary statistics of the estimators in the multi-covariate model with normal error ........................................... 27
2.3 Simulation summary statistics of the estimators in the multi-covariate model with $\chi^2$ error ........................................... 28
2.4 NPC Trial SCC Data ........................................... 30
3.1 Prevalence (%) of single-root pattern for the parametric corrected score with $E(X) = 0$, $\text{Var}(X) = 1$, $\beta = -1$, and $\varepsilon \sim \text{Normal}(0,1).$ .......................... 47
3.2 Prevalence (%) of root-finding failure for the conditional score with $E(X) = 0$, $\text{Var}(X) = 1$, $\beta = -1$, and $\varepsilon \sim \text{Normal}(0,1).$ .......................... 50
3.3 Simulation summary statistics for the single-covariate models with 20% censoring rate: Ideal, naive(NV), regression calibration (RC), re-defined conditional score (ConS), re-defined parametric corrected score (CS), and first k augmented parametric corrected score (ACS:k), k = 2, 3, 4. ........................................... 61
3.4 Simulation summary statistics for the single-covariate models with 60% censoring rate: Ideal, naive(NV), regression calibration (RC), re-defined conditional score (ConS), re-defined parametric corrected score (CS), and first k augmented parametric corrected score (ACS:k), k = 2, 3, 4. ........................................... 62
3.5 Simulation summary statistics for the double-covariate models with 20% censoring rate: Ideal, naive(NV), regression calibration (RC), re-defined conditional score (ConS), re-defined parametric corrected score (CS), and augmented parametric corrected score (ACS:1/3, ACS:2/3, ACS:2) .......................................................... 66

3.6 Simulation summary statistics for the double-covariate models with 60% censoring rate: Ideal, naive(NV), regression calibration (RC), re-defined conditional score (ConS), re-defined parametric corrected score (CS), and augmented parametric corrected score (ACS:1/3, ACS:2/3, ACS:2) .......................................................... 67

3.7 Coverage of 95% confidence interval for the second order augmented parametric corrected score with 20% censoring rate. C (chi-square distribution), BC (bootstrap calibration) and W (Wald-type) indicate the type of confidence interval. .......................................................... 68

3.8 Coverage of 95% confidence interval for the second order augmented parametric corrected score with 60% censoring rate. C (chi-square distribution), BC (bootstrap calibration) and W (Wald-type) indicate the type of confidence interval. .......................................................... 69

3.9 Comparison of regression coefficient estimators in the ACTG 175 data .......................................................... 70

4.1 Prevalence (%) of root-finding failure for the nonparametric corrected score with $E(X) = 0$, $\text{Var}(X) = 1$, $\beta = -1$, and $\varepsilon \sim \text{Normal}(0,1)$ .......................................................... 80

4.2 Simulation summary statistics for the single-covariate models with 20% censoring rate: Ideal, naive(NV), regression calibration (RC), re-defined nonparametric corrected score (NPC), and second order augmented corrected score (ACS:2). .......................................................... 84

4.3 Simulation summary statistics for the single-covariate models with 60% censoring rate: Ideal, naive(NV), regression calibration (RC), re-defined nonparametric corrected score (NPC), and second order augmented corrected score (ACS:2). .......................................................... 85

4.4 Simulation summary statistics for the double-covariate models with 20% censoring rate: Ideal, naive(NV), regression calibration (RC), re-defined nonparametric corrected score (NPC), and second order augmented corrected score (ACS:2). .......................................................... 89
4.5 Simulation summary statistics for the double-covariate models with 60% censoring rate: Ideal, naive(NV), regression calibration (RC), re-defined nonparametric corrected score (NPC), and second order augmented corrected score (ACS:2).

4.6 Coverage of 95% confidence interval for the second order augmented nonparametric corrected score with 20% censoring rate. C (chi-square distribution), BC (bootstrap calibration) and W (Wald-type) indicate the type of confidence interval.

4.7 Coverage of 95% confidence interval for the second order augmented nonparametric corrected score with 60% censoring rate. C (chi-square distribution), BC (bootstrap calibration) and W (Wald-type) indicate the type of confidence interval.

4.8 Comparison of regression coefficient estimators in the ACTG 175 data
Chapter 1

Introduction

1.1 Background

In many clinical research studies, survival time is typically the primary outcome of interest. The proportional hazards model (Cox, 1972) is one of the most popular methods to investigate the relationship between covariates and possibly right-censored survival times. An important alternative to Cox model is the accelerated failure time (AFT) model. In conventional statistical inference procedures, it is common to assume that all covariates are observed and measured accurately without any error. However, in medical research fields, the concerns of measurement errors often arise because of natural biological fluctuation or instrument error. For regression analysis in general, it is well known that naively using mismeasured covariates in conventional inference procedures may incur substantial estimation bias and several statistical methods have been suggested to address covariate measurement error; see Section 1.3
for a brief review and the monograph of Carroll et al. (2006) for a more comprehensive summary.

The goal of this dissertation research is to develop statistical methods to address several issues on covariate measurement error in survival data analysis. We focus on two types of data: recurrent events data and univariate survival data.

The first topic studies the recurrent events data under the accelerated failure time model. The accelerated failure time model is an important alternative to the Cox model and is appealing in that the model has a direct physical interpretation. But to our knowledge, there is little development on the accelerated failure time model for recurrent events data with covariate measurement error. In this topic, we develop an estimation procedure for recurrent events data under the accelerated failure time model, with the availability of replicated mismeasured covariates.

The second and third topics concern the Cox proportional hazards model for univariate survival data. In the presence of covariate measurement error, several consistent approaches have been proposed for Cox model, namely conditional score (Tsiatis & Davidian, 2001), parametric corrected score (Nakamura, 1990), and non-parametric corrected score (Huang & Wang, 2000). But each of these methods suffers from finite-sample pathological behaviors when the measurement error is substantial. Nonetheless, neither the nature nor the severity of these pathological behaviors are well understood in the literature. In these two topics, we investigate the pathological behaviors of these methods and propose new estimation procedure to tackle the pathological behaviors. In the second topic, we consider the situation that the distribution of measurement error is known. In the third topic, the error distribution is
no longer known but replicated mismeasured covariates are available.

In this chapter, we first present two real data sets that motivate our dissertation research, the Nutritional Prevention of Cancer (NPC) trial and the AIDS Clinical Trial Group (ACTG) 175 study. After that, we will give a brief introduction of general measurement error problems and existing methods in dealing with them. Reviews of relevant existing literature specific to survival data will also be provided. Finally, we will present an outline of this dissertation.

1.2 Motivating Examples

The methods proposed in this dissertation are motivated by two problems. The first one, which motivates our research on the accelerated failure time model for recurrent events data, is the Nutritional Prevention of Cancer (NPC) trial (Clark et al., 1996). Our research on the Cox proportional hazards model is motivated by the AIDS Clinical Trial Group (ACTG) 175 study.

1.2.1 Nutritional Prevention of Cancer (NPC) Trial

The Nutritional Prevention of Cancer (NPC) trial is a randomized double-blind, placebo-controlled clinical trial to evaluate the long term safety and efficacy of daily 200-µg supplement of selenium (Se) in preventing two types of skin cancers: basal cell carcinoma (BCC) and squamous cell carcinoma (SCC). From 1983 through 1991, a total of 1,312 patients with previous histories of BCC or SCC were recruited and
randomized into the treatment group in which patients were supposed to take daily 200-µg supplement of selenium (Se), or the control group in which placebo pills were given. Patients were then followed for a period of over twelve years. Follow up visits to clinics were scheduled for every 6 months or more frequent if necessary to exam for new dermatological problems or selenium toxicity. The primary end points of this trial were new occurrences of BCC or SCC. Possible censoring events include death and end of routine dermatological exams.

Clark et al. (1996) reported an adverse but statistically non-significant effect of selenium supplement in prevention of new occurrence of BCC or SCC. Their findings are contrary to prior expectation since highly significant positive benefits were found on selenium supplement in preventing a number of other types of cancers. The findings of Clark et al. (1996) might be contributed to two reasons. First, they did not consider recurrent BCC or SCC. During the study, each patient might experience multiple occurrences of BCC or SCC. But in the analysis, they considered only time to the first occurrence of BCC or SCC. Second, the baseline plasma Se level as an important prognostic risk factor for BCC and SCC is subject to substantial measurement error. But the authors did not consider the possible effect of measurement error on statistical analysis.

One natural question arises is how to incorporate these additional information of time to multiple occurrences of BCC or SCC and consistently estimate the covariate effect for recurrent events data under the circumstance of covariate measurement error. We will develop methods in Chapter 2 to address this question.
1.2.2 AIDS Clinical Trial Group (ACTG) 175 Study

AIDS Clinical Trial Group (ACTG) 175 study is a randomized clinical trial to evaluate four treatments, zidovudine alone (ZDV), zidovudine plus didanosine (ZDV + ddI), zidovudine plus zalcitabine (ZDV + ddC), and zalcitabine alone (ddC), in HIV-infected subjects with an initial screening CD4 count between 200 and 500. As a note, a healthy individual without HIV infection usually has a CD4 count of between 800 and 1200. A total of 2,467 HIV-infected volunteers participated in the ACTG 175 study and, among them, 1,067 of whom had never taken any antiretroviral therapy at study entry; see Hammer et al. (1996) for a more detailed description of this study.

We are interested in assessing the effect of baseline CD4 count on time to AIDS or death in patient. The Cox proportional hazards model is the most popular regression model for right-censored survival data and it has been widely applied in a large number of HIV/AIDS studies. However, the issue of covariate measurement error arises often in practice. For example, the CD4 count has no gold standard measurement and is subject to substantial measurement error. In ACTG 175 study, 1,036 antiretroviral naive patients had two duplicated baseline CD4 count measurements prior to the start of treatment and within 3 weeks of randomization. The duplicated baseline CD4 count measurements were taken from different blood samples. Figure 1.1 illustrates the issue of measurement error on ACTG 175 study.

Although there exists a rich literature on Cox regression with covariate measurement error, all available consistent methods suffer from severe finite-sample pathological behaviors. For example, the estimating functions may have multiple zero-crossings, no zero-crossing, or a single zero-crossing that is far away from the true
Figure 1.1: Replicated baseline CD4 count measurements from 1,036 antiretroviral-naive patients in the ACTG 175 study.
parameter. Though the pathological behaviors have been noticed in the literature, they were never well understood or systematically studies.

The above observation prompts the second research question of how to improve the finite-sample behaviors of an estimating function for Cox regression with substantial covariate measurement error.

## 1.3 Covariate Measurement Error

In literature, there are two broad classes of measurement error models, namely, *classical measurement error model* and *Berkson measurement error model* (Berkson, 1950).

Suppose covariates $Z = (Z_a^T, Z_e^T)^T$ where $Z_a$ are those covariates that can be accurately measured and $Z_e$ are covariates prone to measurement error and cannot be accurately measured. Though $Z_e$ cannot be measured directly, we can observe them through their surrogates, $W_e$. Under the classical additive measurement error model,

$$W_e = Z_e + \epsilon_e$$

where $\epsilon_e$ is a mean-zero random noise. In most applications, it is assumed that $\epsilon_e$ is independent of both $Z_a$ and $Z_e$, and $\epsilon_e \sim N(0, \sigma_e^2)$.

On the other hand, Berkson measurement error model is defined as

$$Z_e = W_e + \epsilon_e$$
The true covariate equals mismeasured covariates plus measurement error. Same as in classical error model, $\epsilon$ are usually assumed to be independent of $Z_a$ and $Z_e$.

In this dissertation research, we will consider the classical measurement error model. Let $W = (Z_a, W_e)$, it is well known that naively replacing $Z$ by $W$ in regression analysis could lead to substantial bias in estimation of some or all regression coefficients (Carroll et al., 2006; Fuller, 1987).

Here is a simple example to illustrate the effect of measurement error in naive analysis. Consider a simple linear regression with classical additive error, $Y = \beta_0 + \beta_z Z + \epsilon$, where $\text{Var}(Z) = \sigma^2_z$ and $\text{Var}(\epsilon) = \sigma^2_\epsilon$. Covariate $Z$ is unobservable; instead, a surrogate $W = Z + U$ is observed. $E(U|Z) = 0$, and $\text{Var}(U|Z) = \sigma^2_u$. If we fit a naive regression model of $Y$ on $W$, then the estimator is a consistent estimator of $\beta_{z*} = \lambda \beta_z$, where

$$\lambda = \frac{\sigma^2_z}{\sigma^2_z + \sigma^2_u} < 1$$

and

$$\text{Var}(Y|W) = \sigma^2_\epsilon + \frac{\beta^2_z \sigma^2_u \sigma^2_z}{\sigma^2_z + \sigma^2_u} > \sigma^2_\epsilon$$

The naive estimation procedure gives an estimator that is attenuated to zero. As the example shows, measurement error not only produces a biased parameter estimate but also increase the residual variance. In a more complex model, the effect of measurement error in statistical inference will be much more complicated.

The simple measurement error models introduced above can be extended to accommodate more complex applications. For example, Kipnis et al. (1999) introduced a complex measurement error model for some nutritional studies which allows for
both bias and variance components.

To perform measurement error analysis, some additional information is needed. This additional information could be known error distribution, or in the case when error distribution is unknown, some forms of extra data. Depending on the source, these extra data can be categorized as internal data which are obtained for a subset of primary study and external data which are obtained from independent studies. In each category, there are three types of data: validation data, in which we can observe \((Z, W)\) directly for a subset of data; replication data, in which replicates of \(W\) are available for each subject; instrumental data, in which an additional instrumental variable \(T\) is observed in addition to \(W\) (correlated with \(Z\), but may or may not be unbiased for \(Z\)). In medical research fields, replication data are more commonly available and there exists some literature on measurement error problems with replication data.

1.4 Functional Modeling and Structural Modeling

A large number of statistical models have been proposed to tackle various measurement error problems. Traditionally, based on the property of the unobserved true covariates \(Z_e\), a distinction was made between classical functional models, in which the \(Z_e\) are regarded as unknown fixed constants or parameters, and classical structural models, in which the \(Z_e\) are regarded as random variables. Carroll et al. (2006) instead suggested to make a distinction between functional modeling, in which \(Z_e\) may be either fixed or random, but in the latter case, no or only minimal assumptions
are made about the distribution of \( Z_e \), and *structural modeling* in which parametric
distributional models are imposed on \( Z_e \).

The structural modeling approach plays an important role in application (Carroll et al., 2006, chapter 8), but the robustness of inference to parametric model assumptions is of concern. The functional modeling approach is more appealing in the sense
that the estimation procedures does not need specify a true covariate distribution and
thus not subject to misspecification.

Available functional modeling methods might be categorized into three classes. The first includes the conditional score method (Stefanski & Carroll, 1987) and locally
efficient score method (Tsiatis & Ma, 2004). Conditional score requires the
measurement error to be normally distributed and the idea is to condition away the
obtained an unbiased score function for generalized linear measurement error models.
Later, Tsiatis & Ma (2004) proposed a class of semiparametric estimators that require
a pilot distribution of unobserved error-prone covariates to be specified. It is shown
that such estimators are consistent no matter what the pilot distribution is and are
efficient if computed under the truth. It was shown later by Ma & Tsiatis (2006) that
when implemented in generalized liner model and with normal measurement error,
the locally efficient score method (Tsiatis & Ma, 2004) is equivalent to conditional
score. Therefore, their method could be viewed as an extension of conditional score.

The second class is the corrected score method (Nakamura, 1990; Stefanski, 1989;
Huang & Wang, 2001); it is also referred to as parametric correction in the literature.
The corrected score method requires the existence of an estimating function that
produces consistent estimators in the absence of measurement error. It is called the original or reference estimating function. A corrected score is an estimating function based on observed error-contaminated data that has the same limit as the reference estimating function. If the reference estimating function admits consistent estimates only, the corrected estimating function inherits this property in a compact parameter space containing the true value.

The third class is nonparametric correction (Huang & Wang, 1999, 2000, 2001, 2006). Similar to the parametric correction approach, one starts with a reference estimating function based on the underlying true covariates and then construct an estimating function with error-contaminated covariates which share the same limit as reference. The most important difference, however, is that the parametric correction method requires parametric distributional assumptions on the measurement error, whereas the nonparametric correction spares them with the availability of replicated mismeasured covariates.

To avoid any confusion, we will refer to the second class as parametric corrected score and the third class as nonparametric corrected score later in this dissertation.
1.5 Measurement Error Techniques in Survival Analysis

1.5.1 Recurrent Events Data

For recurrent events data without covariate measurement error, a number of regression models have been proposed in the literature as an extension of the Cox proportional hazards model. Andersen & Gill (1982) introduced the counting process model with a Cox-type intensity function for recurrent events and established an elegant large sample asymptotic theory through martingale theory. Wei et al. (1989) proposed a marginal approach for the analysis of recurrent failures. Later, Pepe & Cai (1993) and Lawless et al. (1997) proposed multiplicative models for the rate and mean functions of arbitrary counting processes. More recently, Lin et al. (1998) proposed to extend the univariate accelerated failure time model for the analysis of recurrent events. They developed a class of consistent and asymptotic normal rank estimators. Comparing to the Cox model, the accelerated failure time model has an appealing feature of a direct physical interpretation.

In the presence of covariate measurement error, Jiang et al. (1999) developed a method using replicated measurements of error-prone covariates and applied their method to the NPC trial data. Their method is based on a discrete-time proportional means model (see Lawless & Nadeau, 1995). Parametric assumptions are imposed on both the true covariates and the errors. Later, Hu & Lin (2004) considered an extended proportional hazards model. Though no parametric assumption is imposed on the covariates, their approach requires the errors to be symmetrically distributed.
To our knowledge, there is little development under the accelerated failure time model framework for recurrent events data with covariate measurement error.

### 1.5.2 Univariate Survival Data

There exist a large collection of literature dealing with the problem of measurement error in the analysis of Cox-type models for univariate survival data. Prentice (1982) introduced the induced hazard function and proposed an induced partial likelihood function for the Cox regression model with rare events and utilized the regression calibration approach to estimate the parameters. Clayton (1991) proposed a modification of regression calibration that does not require events to be rare. Another regression calibration method was later developed by Wang et al. (1997).

Though regression calibration is used frequently to yield a reasonable approximation, it is well known that the regression calibration method is inconsistent and may have large bias under certain circumstances. Zhou & Pepe (1995) developed a consistent estimator for regression coefficients. But the requirement of discrete covariates and the availability of a validation set greatly limit the applicability of their method in practice. Another approach is to correct the partial score or related estimating functions. Nakamura (1992) developed a parametric corrected score approach for normal measurement errors and it was shown by Kong & Gu (1999) that, the basic estimator of Nakamura (1992) is consistent and asymptotically normal. Tsiatis & Davidian (2001) proposed a conditional score for normal measurement error, which is asymptotically equivalent to Nakamura’s parametric corrected score with normal error. For data with at least two replicates for mismeasured covariates, Huang &
Wang (2000) proposed a consistent nonparametric estimator based on a correction of the partial score function.

More recently, Song & Huang (2005) investigated the finite-sample performance of parametric corrected score and conditional score and found that they all might suffer from pathological behaviors as the error magnitude increases. They suggested a number of refinements, but the improvement is fairly modest.

1.6 Outline

Chapter 2 deals with parameter estimation in the accelerated failure time model for recurrent events when covariates are measured with error. We consider the classical additive measurement error model and propose a consistent estimation procedure for parameter of interest. With replicated mismeasured covariates available, the proposed estimation procedure requires no distributional assumptions on either the true covariates or the error except for the boundedness of the latter. The resulting estimators are proven to be consistent and asymptotically normal. Simulation studies indicate that the proposed model works well for practical sample sizes and moderate measurement error. An illustration with application to the NPC trial is also provided.

Chapter 3 considers the Cox proportional hazards model. In the presence of covariate measurement error, several functional modeling methods have been proposed for the proportional hazards model under the situation where the distribution of the measurement error is known. In this chapter, we study the finite-sample pathological behaviors of parametric corrected score and conditional score. To address the
issue of pathological behaviors, we propose an approach to incorporate additional estimating functions. Extensive simulation studies are conducted to investigate the performance of the proposed method. Furthermore, we illustrate the practical utility of the proposed methods via an application to the ACTG 175 study.

In Chapter 4, we propose an extension of the method in Chapter 3 by allowing the error distribution to be completely unspecified. Huang & Wang (2000) developed a nonparametric corrected score for proportional hazards model when replicated measurements for mismeasured covariates are available. But the nonparametric corrected score suffers from similar finite-sample pathological behaviors as the parametric corrected score when the measurement error is substantial. We extend the method described in Chapter 3 to nonparametric corrected score and the simulation study result shows that the proposed estimation procedure is promising in resolving those finite-sample pathological behaviors.

In Chapter 5, we provide a summary and discuss some future work for the dissertation.
Chapter 2

Accelerated Failure Time Model for Recurrent Events With Errors in Covariates

2.1 Introduction

In clinical research studies, the event of interest may occur more than once for each subject during the follow up period. One example is the Nutritional Prevention of Cancer (NPC) trial (Clark et al., 1996; Duffield-Lillico et al., 2003), which involved multiple recurrences of skin cancers, basal cell carcinoma (BCC) and squamous cell carcinoma (SCC). For such recurrent events, a number of statistical methods have been developed to extend the univariate Cox proportional hazards model, including Andersen & Gill (1982), Wei et al. (1989), Pepe & Cai (1993) and Lawless et al.
More recently, Lin et al. (1998) proposed to extend the univariate accelerated failure time model for the analysis of recurrent events. As an appealing and well recognized feature, the accelerated failure time model has a direct physical interpretation. All these inference procedures presume the conventional setting that covariates are observed and measured accurately. However, this assumption may not be appropriate in some studies, including the aforementioned NPC trial. The baseline plasma selenium as an important prognostic risk factor for BCC and SCC is subject to substantial measurement error, due to natural biological fluctuation and instrument error.

For recurrent events data with covariate measurement error, available approaches include Jiang et al. (1999)’s method which is based on a discrete-time proportional means model utilizing replicated measurements of error-prone covariates, and an extended proportional hazards model developed by Hu & Lin (2004). But for the former approach, parametric assumptions are imposed on both the true covariates and the errors, whereas for the latter, though no parametric assumption is imposed on the covariates, it requires the errors to be symmetrically distributed. To our knowledge, there is little development on the accelerated failure time model for recurrent events data with covariate measurement error.

In this chapter, we propose a consistent estimation procedure for the accelerated failure time model based on replicated mismeasured covariates. The proposed approach bears similarity, in minimal assumption requirement, to the nonparametric corrected score method used in the other regression models (Huang & Wang, 2000, 2001, 2006). But this model of interest is not amendable to the correction strategy
and the proposed estimating function is developed on the basis of a novel identity. In Section 2.2, we present the proposed estimation procedure. Consistency and asymptotic normality of the proposed estimator will also be established. Simulation studies with practical sample size are reported in Section 2.3, along with an illustration with the NPC trial data. Further discussion is given in Section 2.4. Technical details and proofs for the large-sample study on the proposed inference procedure are collected in Section 2.5.

2.2 Inference Procedure

2.2.1 The Model and Data

Let $\tilde{N}^*(t)$ denotes the number of recurrent events by time $t$ in the absence of censoring and $Z$ the p-vector covariates of interest. The accelerated failure time model (Lin et al., 1998) postulates their relationship as

$$E\{\tilde{N}^*(t)|Z\} = \mu_0(e^{-\beta_0'Zt}),$$

(2.1)

where $\beta_0$ is an unknown p-vector of parameter of interest, and $\mu_0(\cdot)$ is an unspecified baseline rate function. Let $C$ be the censoring time, and the conditionally independent censoring mechanism is adopted, i.e., $C \perp \tilde{N}^*(\cdot)|Z$, where $\perp$ denotes statistical independence. In the presence of censoring, $\tilde{N}^*(\cdot)$ is not fully observed but only through $\tilde{N}(\cdot) = \tilde{N}^*(\cdot \land C)$.

In the case of covariate measurement error, some elements of $Z$ cannot be accu-
rately measured. Split covariates $Z = (Z_a^T, Z_e^T)^T$, where $Z_a$ are accurately measured and $Z_e$ are error prone. Though $Z_e$ cannot be measured directly, we can observe it through their surrogates, $W_e$. Under the additive measurement error model, $R \geq 2$ surrogates $\{W_e^{(m)} : m = 1, \ldots, R\}$ are observed, where

$$W_e^{(m)} = Z_e + \epsilon_e^{(m)}, m = 1, \ldots, R. \quad (2.2)$$

These errors $\epsilon_e^{(m)}$ are iid replicates of $\epsilon_e$ and are independent of all other random variables. No distributional assumption is imposed on covariates $Z$, nor on $\epsilon_e$ except for boundedness. The observed data, $\{\tilde{N}_t(\cdot); C_i; Z_{ai}; R_i; W_e^{(m)} : m = 1, \ldots, R_i\}, i = 1, \ldots, n$, consist of $n$ iid replicates of $\{\tilde{N}(\cdot); C; Z_a; R; W_e^{(m)} : m = 1, \ldots, R\}$.

### 2.2.2 Proposed Estimation Procedure

Pick arbitrarily two replicates $W_e^{(1)}$ and $W_e^{(2)}$ from $\{W_e^{(m)} : m = 1, \ldots, R\}$ and denote

$W^{(1)} = (Z_a^T, W_e^{(1)})^T$ and $W^{(2)} = (Z_a^T, W_e^{(2)})^T$; $R(R-1)$ different permutations can be formed. We assume that $\epsilon_e$ is bounded. Therefore, there exists a function of $\beta$, say $\xi(\beta)$, such that

$$\xi(\beta) \geq |\beta'(W^{(1)} - W^{(2)})|, a.s. \quad (2.3)$$

Define counting process $N(t; Z, \beta) = \tilde{N}\{\exp(\beta'Z)t\}$, and at-risk process $Y(t; Z, \beta) = I\{C \geq \exp(\beta'Z)t\}$. To motivate our estimation procedure, we obtain an important
Identity:

\[
E\{N(t; W^{(1)}, \beta_0) | Z, Y(e^{\xi(t)}; W^{(2)}, \beta_0) = 1, W^{(2)}\} \\
= E\{\tilde{N}^*(e^{\beta_0 W^{(1)} t} \land C) | Z, C \geq e^{\beta_0 W^{(2)} + \xi(t)} \}
= E\{\tilde{N}^*(e^{\beta_0 W^{(1)}}) | Z\}
= E\{\mu_0(e^{\beta_0 t})\} \equiv \Lambda_0(t), \tag{2.4}
\]

where we have used the fact \( \beta_0' W^{(2)} + \xi(\beta_0) \geq \beta_0' W^{(1)} \) and \( \Lambda_0(t) \) is a quantity independent of \( Z \).

Identity (2.4) naturally leads to the following estimating function of both \( \beta_0 \) and \( \Lambda_0(t) \):

\[
\frac{1}{n} \sum_{i=1}^{n} \int_0^\tau A \left( Y_i(e^{\xi(\beta)} t; W_i^{(2)}, \beta) \right) \left( \frac{Y_i(e^{\xi(\beta)} t; W_i^{(2)}, \beta)}{A(W_i^{(2)}, \beta)} \right) d\{N_i(t; W_i^{(1)}, \beta) - \Lambda(t)\} = 0,
\]

where \( \tau \) is a prespecified constant, and \( A \) denotes the operator averaging over all the different permutations of \( W^{(1)} \) and \( W^{(2)} \). Further profiling out \( \Lambda(t) \), we have the following estimating function for \( \beta_0 \):

\[
\Psi(\beta) = \frac{1}{n} \sum_{i=1}^{n} \int_0^\tau A \left[ W_i^{(2)} - \frac{\sum_{j=1}^{n} A(W_j^{(2)}, \beta) Y_j(e^{\xi(\beta)} t; W_j^{(2)}, \beta)}{\sum_{j=1}^{n} A(Y_j(e^{\xi(\beta)} t; W_j^{(2)}, \beta))} \right] Y_i(e^{\xi(\beta)} t; W_i^{(2)}, \beta) dN_i(t; W_i^{(1)}, \beta) \tag{2.5}
\]

In the absence of covariate measurement error, \( W^{(1)} = W^{(2)} = Z \) and we may set \( \xi(\beta) = 0 \). Then, the above estimating function reduces to Lin et al. (1998)'s
estimating function.

It is clear that $\xi(\beta)$ results in additional censoring in the proposed estimation procedure. Therefore, this function should be kept as small as possible. Suppose that $Z_e$ is $q$-dimensional. Let $\epsilon^{(1)}_{[k]}, \epsilon^{(2)}_{[k]}, \beta_{[k]}, k = 1, \ldots, q,$ be the $k$th element of corresponding vectors. Let $M_{[k]} = \max|\epsilon^{(1)}_{[k]} - \epsilon^{(2)}_{[k]}|$, the function $\xi(\beta)$ can be set as $\xi(\beta) = \sum_{k=1}^{q} |\beta_{[k]}| M_{[k]}$. Of course, $M_{[k]}$ is typically unknown. For practical purpose, it may be empirically determined; see Section 2.3.

2.2.3 Asymptotic Properties

To allow for a rigorous large-sample study, we consider the following mild regularity conditions:

C1. Counting process $\tilde{N}(\cdot)$ and covariates $Z$ are bounded;

C2. $C$ has a bounded density function and $\mu_0$ has a bounded second derivative;

C3. Support of $\epsilon_e$ is bounded;

C4. Time limit $\tau$ satisfies $\lim_{n \to \infty} n^{-1} \sum_{i=1}^{n} Y_i(\tau; Z_i, \beta_0) > 0$.

**Theorem 2.1.** Under conditions C1-C4, almost surely, a zero-crossing of $\Psi(\beta)$, say $\hat{\beta}$, exists and converges to $\beta_0$.

The estimating function $\Psi(\beta)$ is a piecewise constant function of $\beta$. Furthermore, the estimating function is not generally monotone. Therefore, the computation is not simple. When the dimension of covariates is small, $\hat{\beta}$ can be obtained by using an
iterative bisection method (Huang, 2002) or direct grid search. When the dimension of covariates is high, specialized numerical methods such as simulated annealing (Lin & Geyer, 1992) might be more efficient.

**Theorem 2.2.** Under conditions $C1$-$C4$, $n^{1/2} (\hat{\beta} - \beta_0)$ is asymptotically normal with mean 0.

In terms of the interval estimation for $\beta_0$, there are two general approaches available. One is to use bootstrap resampling which is computational intensive. An alternative method was proposed by Huang (2002) who developed a sample-based variance estimation procedure for the estimators based on nonsmooth estimation functions in general. Compared to the bootstrap resampling, Huang (2002)'s approach takes much less computing time. To adopt Huang (2002)'s approach, a consistent estimate of $\Psi(\beta_0)$ is given in Section 2.5.

### 2.3 Numerical Studies

We investigated the performance of the proposed estimation procedure under practical sample size via extensive simulation studies. Also the procedure will be illustrated through an application to the NPC study.

For reference and comparison, the ideal, naive, and regression calibration estimators were also studied. The ideal estimator used the procedure of Lin et al. (1998) with the true covariates and, of course, it is not a realistic estimator. The naive approach uses the average of replicated surrogates in replace of true covariates in the procedure of Lin et al. (1998), whereas the regression calibration approach uses the
best linear approximation by replication data given in Carroll et al. (2006, chap. 4). For all these estimators, we used the approach of Huang (2002) for interval estimation.

For all these estimators, one needs to specify the time limit $\tau$ for their estimating function. We chose $\tau$ large enough to include all the follow-up time, as is standard practice in survival analysis.

Typically, the bound of $\epsilon$ is unknown. For practical purpose, we suggest to use an empirical version of $M_{[k]}$ given in Section 2.2.2 in order to specify $\xi(\beta)$ function. Let $W_{[k]}^{(1)}, W_{[k]}^{(2)}, k = 1, \ldots, q$, be the $k$th element of corresponding vectors. An empirical version of $M_{[k]}$ is $\max_{i=1,\ldots,n} |W_{i[k]}^{(1)} - W_{i[k]}^{(2)}|$. The operator max here also applies to all possible pairs of $W^{(1)}$ and $W^{(2)}$.

### 2.3.1 Simulations

Systematic numerical experiments were conducted to investigate the finite-sample performance of the proposed estimation procedure.

We first considered the accelerated failure time model with a single and error-prone covariate and set $\beta_0=0.5$. Recurrent events were generated from Poisson processes with unit baseline rate. To evaluate the performance of the proposed approach in general applications, we considered both symmetric and asymmetric distributions of the covariate and measurement error. For the true covariate, we considered i) a standard normal distribution and ii) a chi-square distribution location-shifted to mean 0 and scale-changed to variance 1. The surrogate $W$ of $Z$ was generated from the additive error model $W = Z + \epsilon$, where $\epsilon$ was i) normal with mean zero and variance
\( \sigma_i^2 = 0.25, \text{ or } 0.5, \text{ or ii) modified chi-square distribution with mean 0 and variance 0.25 or 0.5. For ii), the modification was made on the chi-square distribution with 1 degree of freedom by truncation at 5 and then by location shift and scale change. Two replicates of } W \text{ were generated for each } Z. \text{ Censoring times } C \text{ were generated from the uniform distribution on } [0, \lambda], \text{ where } \lambda \text{ was chosen to yield an average of 10 events per subject. With a sample size of 400, 1,000 data sets were generated.}

Table 2.1 summarizes the simulation results. For each scenario, the mean bias, standard deviation, averaged standard error, and coverage probability of 95% Wald-type confidence interval were calculated. For the proposed estimator, median bias, a robustified standard deviation defined as inter quartile range (IQR) divided by 1.349, and median of estimated standard error were also calculated. In all different settings, naive estimator was biased toward 0 with the bias increasing as the magnitude of measurement error increasing. Naive estimator also had poor coverage probability under all scenarios. Regression calibration estimator performed well in both bias and coverage probability when the underlying true covariate was normal. However, when the true covariate was modified chi-square distribution, regression calibration estimator was biased toward 0 and the coverage probabilities were lower than nominal level. The proposed estimator performed well with small bias when the true covariate was normally distributed. However, when true covariate followed the modified chi-square distribution, the mean bias could be substantial especially with increasing error variance although median bias was still small in these circumstances. Nevertheless, the distribution of the proposed estimator was quite skewed, which is similar to other
Table 2.1: Simulation summary statistics of the estimators in the single-covariate model

<table>
<thead>
<tr>
<th>$\sigma^2_\varepsilon$</th>
<th>I</th>
<th>NV</th>
<th>RC</th>
<th>Prop</th>
<th>I</th>
<th>NV</th>
<th>RC</th>
<th>Prop</th>
</tr>
</thead>
<tbody>
<tr>
<td>Z, $\varepsilon$ ∼ Normal</td>
<td>0.25</td>
<td>B</td>
<td>5</td>
<td>-561</td>
<td>0</td>
<td>5</td>
<td>-5</td>
<td>-11</td>
</tr>
<tr>
<td>SD</td>
<td>179</td>
<td>187</td>
<td>218</td>
<td>337</td>
<td>311</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SE</td>
<td>181</td>
<td>197</td>
<td>222</td>
<td>385</td>
<td>361</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CP</td>
<td>95.4</td>
<td>19.3</td>
<td>95.3</td>
<td>94.8</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Z, $\varepsilon$ ∼ Modified $\chi^2$</td>
<td>0.25</td>
<td>B</td>
<td>5</td>
<td>-561</td>
<td>0</td>
<td>5</td>
<td>-5</td>
<td>-11</td>
</tr>
<tr>
<td>SD</td>
<td>277</td>
<td>272</td>
<td>297</td>
<td>2044</td>
<td>814</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SE</td>
<td>286</td>
<td>273</td>
<td>308</td>
<td>1755</td>
<td>1060</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CP</td>
<td>94.5</td>
<td>1.7</td>
<td>46.1</td>
<td>93.6</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Z ∼ Normal, $\varepsilon$ ∼ Modified $\chi^2$</td>
<td>0.5</td>
<td>B</td>
<td>5</td>
<td>-561</td>
<td>0</td>
<td>5</td>
<td>-5</td>
<td>-11</td>
</tr>
<tr>
<td>SD</td>
<td>277</td>
<td>278</td>
<td>341</td>
<td>4623</td>
<td>1546</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SE</td>
<td>286</td>
<td>279</td>
<td>351</td>
<td>3086</td>
<td>2103</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CP</td>
<td>94.5</td>
<td>0</td>
<td>28.3</td>
<td>90.4</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Z ∼ Modified $\chi^2$, $\varepsilon$ ∼ Normal</td>
<td>0.25</td>
<td>B</td>
<td>-8</td>
<td>-304</td>
<td>11</td>
<td>-2</td>
<td>-29</td>
<td>11</td>
</tr>
<tr>
<td>SD</td>
<td>182</td>
<td>185</td>
<td>198</td>
<td>305</td>
<td>308</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SE</td>
<td>287</td>
<td>250</td>
<td>283</td>
<td>864</td>
<td>703</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CP</td>
<td>95.7</td>
<td>0.1</td>
<td>30.1</td>
<td>96.2</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Z ∼ Modified $\chi^2$</td>
<td>0.5</td>
<td>B</td>
<td>-8</td>
<td>-304</td>
<td>11</td>
<td>-2</td>
<td>-29</td>
<td>11</td>
</tr>
<tr>
<td>SD</td>
<td>282</td>
<td>236</td>
<td>314</td>
<td>2404</td>
<td>898</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SE</td>
<td>287</td>
<td>242</td>
<td>303</td>
<td>1944</td>
<td>1270</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CP</td>
<td>95.7</td>
<td>0</td>
<td>8.9</td>
<td>94.3</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Note: I, ideal method; NV, naive estimator; RC, regression calibration; Prop, the proposed estimator. Except for the second column under Prop, B: mean bias ($\times 10^4$); SD: standard deviation($\times 10^4$); SE: averaged standard error ($\times 10^4$); CP, coverage probability(%) of the 95% Wald confidence interval. For the second column under Prop, B: median bias ($\times 10^4$); SD: robustified standard deviation($\times 10^4$); SE: median standard error ($\times 10^4$).
correction-type estimators (e.g. Nakamura, 1990). This skewness is also reflected in standard deviation versus robustified standard deviation and averaged standard error versus median standard error. The coverage probabilities of the proposed estimator were close to nominal level under all settings.

We also considered an accelerated failure time model with two covariates. The first covariate is subject to measurement error, with two surrogate replicates available, and the second covariate is measured precisely. The true coefficients $\beta = (\beta_1, \beta_2)^T = (0.5, 0.5)^T$. We present the results from two scenarios. In the first scenario, the true covariates followed the standard bivariate normal distribution with correlation coefficient $\rho$. The measurement error follows a normal distribution with mean 0 and variance 0.25 or 0.5. In the second scenario, the true covariates were generated from the standard bivariate normal distribution upon a marginal increasing transformation and that the first covariate has the $\chi^2(1)$ distribution location-shifted to 0 and scale-changed to 1. The measurement errors were generated from $\chi^2(1)$ distribution truncated at 5 and were transformed to have mean 0 and variance $= 0.25$ or 0.5 by scale change and location shift. For both scenarios, we considered different correlation parameter $\rho$ between two true covariates. Recurrent events were generated from Poisson processes. Censoring times were generated from the uniform distribution on $[0, \lambda]$, where $\lambda$ was chosen to yield an average of 5 events per subject. With a sample size of 400, 1,000 data sets were generated.

Table 2.2 and 2.3 summarize the simulation results for multivariate cases. As
Table 2.2: Simulation summary statistics of the estimators in the multi-covariate model with normal error

<table>
<thead>
<tr>
<th>$\rho$</th>
<th>$\sigma^2_\epsilon$</th>
<th>B $\beta_1$</th>
<th>$\rho$</th>
<th>$\sigma^2_\epsilon$</th>
<th>B $\beta_2$</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>I</td>
<td>NV</td>
<td>RC</td>
<td>Prop</td>
</tr>
<tr>
<td>0.5</td>
<td>0.25</td>
<td>-23</td>
<td>-741</td>
<td>-24</td>
<td>36</td>
</tr>
<tr>
<td></td>
<td></td>
<td>SD</td>
<td>285</td>
<td>304</td>
<td>365</td>
</tr>
<tr>
<td></td>
<td></td>
<td>SE</td>
<td>302</td>
<td>303</td>
<td>361</td>
</tr>
<tr>
<td></td>
<td></td>
<td>CP</td>
<td>95.7</td>
<td>31.5</td>
<td>94.2</td>
</tr>
<tr>
<td>0.5</td>
<td>B</td>
<td>-23</td>
<td>-1277</td>
<td>-19</td>
<td>44</td>
</tr>
<tr>
<td></td>
<td></td>
<td>SD</td>
<td>285</td>
<td>304</td>
<td>443</td>
</tr>
<tr>
<td></td>
<td></td>
<td>SE</td>
<td>302</td>
<td>299</td>
<td>414</td>
</tr>
<tr>
<td></td>
<td></td>
<td>CP</td>
<td>95.7</td>
<td>1.9</td>
<td>92.7</td>
</tr>
<tr>
<td>0</td>
<td>0.25</td>
<td>4</td>
<td>-556</td>
<td>4</td>
<td>46</td>
</tr>
<tr>
<td></td>
<td></td>
<td>SD</td>
<td>260</td>
<td>265</td>
<td>307</td>
</tr>
<tr>
<td></td>
<td></td>
<td>SE</td>
<td>270</td>
<td>279</td>
<td>315</td>
</tr>
<tr>
<td></td>
<td></td>
<td>CP</td>
<td>95.7</td>
<td>48.9</td>
<td>94.6</td>
</tr>
<tr>
<td>0.5</td>
<td>B</td>
<td>4</td>
<td>-1001</td>
<td>10</td>
<td>45</td>
</tr>
<tr>
<td></td>
<td></td>
<td>SD</td>
<td>260</td>
<td>266</td>
<td>358</td>
</tr>
<tr>
<td></td>
<td></td>
<td>SE</td>
<td>270</td>
<td>281</td>
<td>354</td>
</tr>
<tr>
<td></td>
<td></td>
<td>CP</td>
<td>95.7</td>
<td>5.4</td>
<td>94.6</td>
</tr>
<tr>
<td>-0.5</td>
<td>0.25</td>
<td>-3</td>
<td>-718</td>
<td>6</td>
<td>56</td>
</tr>
<tr>
<td></td>
<td></td>
<td>SD</td>
<td>307</td>
<td>301</td>
<td>365</td>
</tr>
<tr>
<td></td>
<td></td>
<td>SE</td>
<td>378</td>
<td>366</td>
<td>456</td>
</tr>
<tr>
<td></td>
<td></td>
<td>CP</td>
<td>98.1</td>
<td>49.3</td>
<td>98.6</td>
</tr>
<tr>
<td>0.5</td>
<td>B</td>
<td>-3</td>
<td>-1253</td>
<td>14</td>
<td>261</td>
</tr>
<tr>
<td></td>
<td></td>
<td>SD</td>
<td>307</td>
<td>291</td>
<td>432</td>
</tr>
<tr>
<td></td>
<td></td>
<td>SE</td>
<td>378</td>
<td>351</td>
<td>529</td>
</tr>
<tr>
<td></td>
<td></td>
<td>CP</td>
<td>98.1</td>
<td>4.6</td>
<td>98.1</td>
</tr>
</tbody>
</table>

Note: Same as that of Table 2.1.
Table 2.3: Simulation summary statistics of the estimators in the multi-covariate model with $\chi^2$ error

<table>
<thead>
<tr>
<th>$\rho$</th>
<th>$\sigma^2_c$</th>
<th>B</th>
<th></th>
<th></th>
<th>I</th>
<th>NV</th>
<th>RC</th>
<th>Prop</th>
<th></th>
<th>I</th>
<th>NV</th>
<th>RC</th>
<th>Prop</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.5</td>
<td>0.25</td>
<td>12</td>
<td>-1</td>
<td>-1212</td>
<td>-862</td>
<td>144</td>
<td>10</td>
<td>11</td>
<td>322</td>
<td>151</td>
<td>-10</td>
<td>-29</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>529</td>
<td>450</td>
<td>491</td>
<td>1073</td>
<td>956</td>
<td>301</td>
<td>314</td>
<td>323</td>
<td>452</td>
<td>449</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>613</td>
<td>471</td>
<td>532</td>
<td>1544</td>
<td>1191</td>
<td>255</td>
<td>287</td>
<td>279</td>
<td>509</td>
<td>470</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>97.4</td>
<td>26.5</td>
<td>63.4</td>
<td>96.0</td>
<td>90.5</td>
<td>78.7</td>
<td>57.6</td>
<td>95.7</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>0.5</td>
<td>-12</td>
<td>-1872</td>
<td>-1294</td>
<td>328</td>
<td>0</td>
<td>11</td>
<td>501</td>
<td>217</td>
<td>-54</td>
<td>5</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>529</td>
<td>418</td>
<td>497</td>
<td>2017</td>
<td>1238</td>
<td>301</td>
<td>318</td>
<td>337</td>
<td>644</td>
<td>548</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>613</td>
<td>427</td>
<td>529</td>
<td>2187</td>
<td>1794</td>
<td>255</td>
<td>299</td>
<td>293</td>
<td>714</td>
<td>636</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>97.4</td>
<td>0.9</td>
<td>30.2</td>
<td>95.3</td>
<td>90.5</td>
<td>62.1</td>
<td>85.3</td>
<td>95.6</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>0</td>
<td>0.25</td>
<td>9</td>
<td>-720</td>
<td>-419</td>
<td>68</td>
<td>10</td>
<td>6</td>
<td>17</td>
<td>17</td>
<td>19</td>
<td>29</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>400</td>
<td>354</td>
<td>377</td>
<td>633</td>
<td>652</td>
<td>264</td>
<td>274</td>
<td>274</td>
<td>388</td>
<td>362</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>412</td>
<td>368</td>
<td>395</td>
<td>826</td>
<td>724</td>
<td>257</td>
<td>266</td>
<td>265</td>
<td>401</td>
<td>394</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>95.2</td>
<td>48.5</td>
<td>80.2</td>
<td>96.3</td>
<td>93.8</td>
<td>94.3</td>
<td>94.5</td>
<td>94.9</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>0.5</td>
<td>-1207</td>
<td>-675</td>
<td>140</td>
<td>29</td>
<td>6</td>
<td>20</td>
<td>20</td>
<td>18</td>
<td>10</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>400</td>
<td>339</td>
<td>384</td>
<td>866</td>
<td>825</td>
<td>264</td>
<td>283</td>
<td>284</td>
<td>468</td>
<td>463</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>412</td>
<td>353</td>
<td>404</td>
<td>1707</td>
<td>1074</td>
<td>257</td>
<td>272</td>
<td>271</td>
<td>571</td>
<td>479</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>95.2</td>
<td>8.0</td>
<td>58.8</td>
<td>96.4</td>
<td>93.8</td>
<td>94.4</td>
<td>93.8</td>
<td>95.4</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>-0.5</td>
<td>0.25</td>
<td>29</td>
<td>-646</td>
<td>-244</td>
<td>58</td>
<td>10</td>
<td>9</td>
<td>-220</td>
<td>-22</td>
<td>10</td>
<td>-10</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>388</td>
<td>341</td>
<td>376</td>
<td>606</td>
<td>583</td>
<td>296</td>
<td>298</td>
<td>307</td>
<td>442</td>
<td>449</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>475</td>
<td>422</td>
<td>491</td>
<td>975</td>
<td>858</td>
<td>264</td>
<td>273</td>
<td>276</td>
<td>484</td>
<td>430</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>97.8</td>
<td>66.2</td>
<td>96.5</td>
<td>98.1</td>
<td>91.7</td>
<td>84.0</td>
<td>91.4</td>
<td>94.0</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>0.5</td>
<td>-1115</td>
<td>-396</td>
<td>140</td>
<td>-39</td>
<td>9</td>
<td>-343</td>
<td>-39</td>
<td>38</td>
<td>-10</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>388</td>
<td>389</td>
<td>461</td>
<td>1054</td>
<td>820</td>
<td>296</td>
<td>461</td>
<td>496</td>
<td>613</td>
<td>557</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>475</td>
<td>401</td>
<td>512</td>
<td>1605</td>
<td>1232</td>
<td>264</td>
<td>283</td>
<td>292</td>
<td>770</td>
<td>620</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>97.8</td>
<td>17.6</td>
<td>91.8</td>
<td>98.1</td>
<td>91.7</td>
<td>69.5</td>
<td>91.6</td>
<td>92.9</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Note: Same as that of Table 2.1.
shown, when two covariates are correlated, the measurement error generally has impact not only on the error-prone covariate but also on that of the accurately measured one. The relative performance is similar to what was observed in the single covariate case.

2.3.2 Illustration with the NPC Trial Data

The NPC trial is a randomized double-blind, placebo-controlled clinical trial to evaluate the long term safety and efficacy of a daily 200-µg supplement of selenium (Se) in preventing two types of skin cancers, BCC and SCC. Patients with previous histories of BCC or SCC were recruited and randomized into the treatment group in which patients were supposed to take daily 200-µg supplement of selenium (Se), or the control group in which placebo pills were given. One of the primary end points of this trial was newly diagnosed SCC lesions.

In this analysis, we are interested in the effect of baseline plasma Se on new SCC lesions. In the NPC trial, measurements of plasma Se were taken from routine clinic visits at approximately six months intervals. For those in the control group, repeated measurements for each patient might represent replicated measurements of baseline plasma Se under the assumption of stationarity (Jiang et al., 1999; Hu & Lin, 2004). This, however, would not be true for patients in the Se group since the plasma Se level was likely to change due to the treatment. As an illustration of the proposed estimation procedure, we will restrict our attention to the patients in the control
We considered selenium measurements within 24 months of randomization and included 589 patients in the control group with at least 2 such measurements in the analysis. The average number of replicates was 4.4 with a maximum of 12 replicates. One hundred sixty seven (28.4%) patients had at least one occurrence of SCC after randomization. More specifically, 83 (14.1%), 32 (5.4%), 19 (3.2%) and 33 (5.6%) patients had 1, 2, 3 and 4 or more occurrences. The average duration of follow-up was 67.7 months and the average number of occurrences was 0.67 per patient.

Table 2.4: NPC Trial SCC Data

<table>
<thead>
<tr>
<th>Method</th>
<th>log(Se) Estimate</th>
<th>SE</th>
</tr>
</thead>
<tbody>
<tr>
<td>Naive I</td>
<td>1.156</td>
<td>0.535</td>
</tr>
<tr>
<td>Naive II</td>
<td>1.280</td>
<td>0.634</td>
</tr>
<tr>
<td>RC</td>
<td>1.693</td>
<td>0.976</td>
</tr>
<tr>
<td>Proposed</td>
<td>1.743</td>
<td>1.292</td>
</tr>
</tbody>
</table>

Note: Baseline Se measurement (Naive I), Average of all Se measurements (Naive II), Regression Calibration (RC), and Proposed Estimator (Proposed)

The covariate of interest is the logarithm of the true baseline selenium. Table 2.4 shows the analysis results based on the proposed method, regression calibration method and two naive methods, one using only the baseline selenium measurement and the other using the average of all selenium measurements within 24 months of randomization. In comparison, the naive approaches yield coefficient estimates
of log(Se) with smaller magnitude. Both the regression calibration and proposed estimates are larger in magnitude. Given the fact that the distribution of error is symmetric and close to normal (Hu & Lin, 2004), this result was expected, and is consistent with our simulations.

2.4 Discussion

In this chapter, we have proposed a consistent estimation procedure for the accelerated failure time model with recurrent events data in the presence of covariate measurement error. Under the accelerated failure time model and additive measurement error model, this method yields consistent regression coefficient estimation with only mild boundedness assumption on the measurement error. Simulation studies showed satisfactory performance of the proposed estimation procedure in samples of moderate size and measurement error, even when the measurement error was not bounded.

Lin et al. (1998) considered a class of weighted estimating functions, including the log-rank estimating function and Gehan estimating function as special cases. Our proposed estimation procedure should be able to be extended along the same line. With a proper chosen weight function, the efficiency of estimation procedure might be improved.

The proposed estimation procedure falls into the category of functional modeling
since it does not impose any parametric distribution on covariates. In particular, the proposed estimation procedure resembles the nonparametric correction method in that both approaches are free of distributional assumptions on underlying covariates and measurement error. However, the proposed method is distinct in its motivation, as being based on a novel identity (2.4). In contrast, the nonparametric correction method relies on an original estimating function (in the absence of measurement error) to construct a corrected estimating function with the same limit.

This method requires replicated measurements on error-prone covariates, which might not always be available. For instance, patients in Se group of the NPC trial did not have replicated baseline plasma Se measurements. This limitation warrants future research.

2.5 Appendix: Technical Details

2.5.1 Proof of the Asymptotic Theory

Proof of Theorem 2.1

We first prove the uniform consistency of the estimating function \( \Psi(\beta) \) given by
(2.5). By algebra, we can write $\Psi(\beta)$ as a function of four empirical processes

$$
\Psi(\beta) = \tilde{E}A \int_{0}^{\tau} W^{(2)}Y(e^{\xi(\beta)}s; W^{(2)}, \beta)dN(s; W^{(1)}, \beta)
$$

$$
- \int_{0}^{\tau} \tilde{E}A\{W^{(2)}Y(e^{\xi(\beta)}t; W^{(2)}, \beta)\}
$$

$$
\times d\tilde{E}A \int_{0}^{t} Y(e^{\xi(\beta)}s; W^{(2)}, \beta)dN(s; W^{(1)}, \beta)
$$

$$
\equiv A_{1}^{(n)}(\beta) - \int_{0}^{\tau} \frac{A_{2}^{(n)}(t, \beta)}{A_{3}^{(n)}(t, \beta)}dA_{4}^{(n)}(t, \beta)
$$

where $\tilde{E}$ represents sample empirical mean, e.g., $\tilde{E}A\{W^{(2)}Y(e^{\xi(\beta)}t; W^{(2)}, \beta)\} = n^{-1} \sum_{i=1}^{n} A\{W^{(2)}_{i}Y_{i}(e^{\xi(\beta)}t; W^{(2)}_{i}, \beta)\}.$

Using the technique of Kosorok (2008), we will show in the following that all four empirical processes $A_{1}^{(n)}$ to $A_{4}^{(n)}$ are P-Donsker classes and thus each of them will converge to their limits uniformly in $t$ and $\beta$.

We first show $A_{3}^{(n)}$ is P-Donsker class. First consider the case when $\beta$ is a scalar, we have $\xi(\beta) = |\beta|M_{[1]}$ where $M_{[1]}$ is defined in Section 2.2.2.

Empirical process

$$
Y(e^{\xi(\beta)}t; W^{(2)}, \beta)
$$

$$
= I(C \geq e^{\beta W^{(2)} + \xi(\beta)t})
$$

$$
= I\{\log C \geq \beta W^{(2)} + \xi(\beta) + \log t\}
$$

$$
= I\{\log C \geq \beta(W^{(2)} + M_{[1]}) + \log t\} \times I\{\log C \geq \beta(W^{(2)} - M_{[1]}) + \log t\}$$
\[ I(\log C \geq \beta(W^{(2)} + M_{[1]}) + \log t) \] and \[ I(\log C \geq \beta(W^{(2)} - M_{[1]}) + \log t) \] are indicator functions of half-space in \( \mathbb{R}^2 \) and thus belongs to Vapnik-Červonenkis-class with index 4 (Kosorok, 2008, Lemma 9.12). Following the proof of Kosorok (2008, Lemma 8.12), we can show that these functions are pointwise measurable classes. Applying Theorem 9.3, Theorem 8.19 of Kosorok (2008) and the fact that indicator functions are bounded, we can see that \( I(\log C \geq \beta(W^{(2)} + M_{[1]}) + \log t) \) and \( I(\log C \geq \beta(W^{(2)} - M_{[1]}) + \log t) \) are P-Donsker classes indexed by \( \beta \) and \( t \). We now have that \( Y(e^{\xi(\beta)}t; W^{(2)}, \beta) \) is Donsker since products of bounded Donsker classes are also Donsker. When \( \beta \) is a \( p \)-element vector, \( Y(e^{\xi(\beta)}t; W^{(2)}, \beta) \) can be written as the product of \( 2^p \) indicator functions. Each of these indicator functions can be shown as Donsker class. Therefore we can show \( Y(e^{\xi(\beta)}t; W^{(2)}, \beta) \) is Donsker with similar argument. Thus \( A_{3}^{(n)}(t, \beta) = \hat{\mathcal{E}} \mathcal{A}(Y(e^{\xi(\beta)}t; W^{(2)}, \beta)) \) is Donsker since sums of bounded Donsker classes are Donsker.

From Condition C3, \( W^{(2)} \) is P-Donsker. \( W^{(2)}Y(e^{\xi(\beta)}t; W^{(2)}, \beta) \) is the product of two bounded P-Donsker classes and therefore it is also P-Donsker. Donsker property of \( A_{2}^{(n)}(t, \beta) \) is shown.

Now turn to \( A_{4}^{(n)}(t, \beta) = \hat{\mathcal{E}} \mathcal{A} \int_{0}^{t} Y(e^{\xi(\beta)}s; W^{(2)}, \beta) dN(s; W^{(1)}, \beta) \). Let \( T_j, j = 1, 2, \ldots \), be the \( j \)th event time for the subject, and \( B \) be the upper bound of counting
process $\tilde{N} (\cdot)$ (Condition C1),

$$\int_0^t Y(e^{\xi(s)}; W^{(2)}, \beta) dN(s; W^{(1)}, \beta) = N^*(t \wedge Ce^{-\beta^\prime W^{(2)} - \xi(\beta)}; W^{(1)}, \beta)$$

$$= \tilde{N}^*(e^{\beta^\prime W^{(1)} t \wedge Ce^{-\beta^\prime W^{(2)} - \xi(\beta)} + \beta^\prime W^{(1)})}$$

$$= \sum_{k=1}^{B} I(T_k \leq e^{\beta^\prime W^{(1)} t \wedge Ce^{-\beta^\prime W^{(2)} - \xi(\beta)} + \beta^\prime W^{(1)})}$$

is the summation of a finite number of P-Donsker classes. Therefore

$$\int_0^t Y(e^{\xi(s)}; W^{(2)}, \beta) dN(s; W^{(1)}, \beta)$$

is P-Donsker. $A_4^{(n)}(t, \beta)$ and $A_1^{(n)}(\beta)$ are also P-Donsker.

With the uniform convergence of $A_1^{(n)}$, $A_2^{(n)}$, $A_3^{(n)}$ and $A_4^{(n)}$ and the fact that the limit of $A_3^{(n)}$ is bounded away from 0, the uniform convergence of estimating function $\Psi(\beta)$ can be established. $\Psi(\beta)$ will converge to its limit $\psi(\beta)$ uniformly in $\beta$. Note that the limit function $\psi(\beta)$ has a zero-crossing at $\beta_0$. Thus the uniform convergence of the estimating function $\Psi(\beta)$ implies that there exists one zero-crossing of $\Psi(\beta)$ converging to $\beta_0$.

Proof of Theorem 2.2

First, we show the asymptotic normality of the estimating function $\Psi(\beta_0)$.

Estimating function $\Psi(\beta)$ is a mapping from four empirical processes which have been shown to be P-Donsker. It can be shown by Lemma 12.2, Lemma 12.3 and Lemma 6.19 (chain rule) of Kosorok (2008) that the mapping is Hadamard differ-

entiable. Applying functional delta method, we obtain that $n^{1/2}(\Psi(\beta_0) - \psi(\beta_0))$
converges weakly to a zero mean Gaussian process.

We then establish the asymptotic linearity of $\Psi(\beta)$ in a neighborhood of $\beta_0$. Let $A_1$ to $A_4$ denote the limit of $A_1^{(n)}$ to $A_4^{(n)}$ respectively. We will first show the following property

$$\lim_{n \to 0} \limsup_{n \to \infty} \Pr \left\{ \sup_{\|\beta - \beta_0\| \leq \eta} n^{1/2} \| A_k^{(n)}(\beta, t) - A_k(\beta, t) - A_k^{(n)}(\beta_0, t) + A_k(\beta_0, t) \| > \epsilon \right\} = 0$$

(2.6)

for $k = 1, \cdots, 4$.

First consider $k = 3$. Let $M(\beta) = (\text{sgn}(\beta[1])M_{[1]}, \ldots, \text{sgn}(\beta[\eta])M_{[\eta]})$. $\xi(\beta)$ can be re-written as $\beta'M(\beta)$. Therefore $A\{Y(e^{\xi(\beta)}t; W^{(2)}, \beta)\} = A\{Y(e^{\beta'M(\beta)}t; W^{(2)}, \beta)\}$.

We have

$$\text{Var} \left[ A\{Y(e^{\beta'M(\beta)}t; W^{(2)}, \beta) - Y(e^{\beta_0'M(\beta_0)}t; W^{(2)}, \beta_0)\} \right]$$

$$= A(\text{Var} \left[ I\{\log C \geq \beta'U(\beta) + \log t \} - I\{\log C \geq \beta_0'U(\beta_0) + \log t \} \right]) \text{ (let } U(\beta) = W^{(2)} + M(\beta))$$

$$= A(\text{Var} \left[ I\{\beta'U(\beta) + \log t \leq \log C \leq \beta_0'U(\beta_0) + \log t \} \right]$$

$$+ I\{\beta_0'U(\beta_0) + \log t \leq \log C \leq \beta'U(\beta) + \log t \} \right\}$$

$$= A(\text{Var} \left[ I\{\beta'U(\beta) + \log t \leq \log C \leq \beta_0'U(\beta_0) + \log t \} \right]$$

$$+ \text{Var} \left[ I\{\beta_0'U(\beta_0) + \log t \leq \log C \leq \beta'U(\beta) + \log t \} \right]$$

$$\leq A(E \left[ I\{\beta'U(\beta) + \log t \leq \log C \leq \beta_0'U(\beta_0) + \log t \} \right]$$

$$+ \text{E} \left[ I\{\beta_0'U(\beta_0) + \log t \leq \log C \leq \beta'U(\beta) + \log t \} \right]$$

$$= A(P_1 + P_2)$$
$\beta'U(\beta)$ is continuous in $\beta$ at $\beta_0$. When $||\beta - \beta_0|| \leq \delta$ and $\delta \to 0$, we have $P_1$ and $P_2$ $\to 0$. By Condition C2, then

$$\text{Var} \left[ A\{Y(e^{\xi(\beta)}t; W^{(2)}, \beta) - Y(e^{\xi(\beta)}t; W^{(2)}, \beta_0)\} \right] \to 0$$

The Donsker property of $A\{Y(e^{\xi(\beta)}t; W^{(2)}, \beta)\}$ implies the tightness of the process.

Thus $\text{Var} \left[ A\{Y(e^{\xi(\beta)}t; W^{(2)}, \beta) - Y(e^{\xi(\beta)}t; W^{(2)}, \beta_0)\} \right] \to 0$ implies that equation (2.6) holds (Kosorok, 2008, p.15).

Now we have shown equation (2.6) holds for $k = 3$. Similarly, we can also show that equation (2.6) holds for $k = 1, 2$ and $4$.

Then with some simple algebra, one can show that

$$\lim_{\xi \to 0} \limsup_{n \to \infty} \Pr \left\{ \sup_{||\beta - \beta_0|| \leq \xi} n^{1/2} ||\Psi(\beta, t) - \psi(\beta, t) - \Psi(\beta_0, t) + \psi(\beta_0, t)|| > \epsilon \right\} = 0$$

(2.7)

After proving (2.7), together with the asymptotic normality of $\Psi(\beta)$, the asymptotic normality of $\hat{\beta}$ follows via fairly standard arguments.

### 2.5.2 Asymptotic Variance of Estimating Function

For abbreviation of notation, we will use $Y^{(1)}$ to denote $Y(e^{\xi(\beta)}s; W^{(1)}, \beta)$ and $N^{(1)}$ to denote $N(t; W^{(1)}, \beta)$. Similar abbreviations will also be used for other notations.
With functional delta method, straightforward algebra gives

\[ n^{1/2} \Psi(\beta) = n^{-1/2} \sum_{i=1}^{n} [B_{i1} - B_{i2}] + o_p(1) \]  \hspace{1cm} (2.8) \]

where

\[
B_{i1} = A \int_0^\tau \left\{ \frac{W_i^{(2)} - \mathcal{E}W^{(2)}Y^{(2)}}{\mathcal{E}Y^{(2)}} \right\} Y_i^{(2)} dN_i^{(1)} \\
B_{i2} = A \int_0^\tau \left\{ \frac{W_i^{(2)}Y_i^{(2)}}{\mathcal{E}Y^{(2)}} - \frac{\mathcal{E}W^{(2)}Y^{(2)}Y_i^{(2)}}{(\mathcal{E}Y^{(2)})^2} \right\} d\mathcal{E} \int Y^{(2)} dN^{(1)}
\]

Thus, \( n^{1/2} \Psi(\beta) \) is asymptotically a sum of iid random variables.

For fixed \( \beta \), \( n^{1/2}(\Psi(\beta) - \psi(\beta)) \) is asymptotically normal with mean 0 and a covariance matrix \( \Omega(\beta) \) that can be consistently estimated by

\[
\hat{\Omega}(\beta) = n \sum_{i=1}^{n} \{ \omega_i(\beta) - \bar{\omega}(\beta) \} \{ \omega_i(\beta) - \bar{\omega}(\beta) \}'
\]

where \( \omega_i(\beta) = n^{-1} (B_{i3} - B_{i4}) \), \( \bar{\omega}(\beta) = n^{-1} \sum_{i=1}^{n} \omega_i(\beta) \), and \( B_{i3} \) and \( B_{i4} \) are defined as

\[
B_{i3} = A \int_0^\tau \left\{ \frac{W_i^{(2)} - \tilde{\mathcal{E}}AW^{(2)}Y^{(2)}}{\tilde{\mathcal{E}}AY^{(2)}} \right\} Y_i^{(2)} dN_i^{(1)} \text{, and} \\
B_{i4} = A \int_0^\tau \left\{ \frac{W_i^{(2)}Y_i^{(2)}}{\tilde{\mathcal{E}}AY^{(2)}} - \frac{(\tilde{\mathcal{E}}AW^{(2)}Y^{(2)})Y_i^{(2)}}{(\tilde{\mathcal{E}}AY^{(2)})^2} \right\} \times d\tilde{\mathcal{E}}A \int Y^{(2)} dN^{(1)}. \]
Chapter 3

Augmented Parametric Corrected Score for Proportional Hazards Model with Covariate Measurement Error

3.1 Introduction

The proportional hazards model is one of the most popular models to investigate the relationship between time to failure and covariates. However, in many clinical trials, the true covariates may not always be accurately measured. In some studies,
the magnitude of measurement error could be substantial to the extent that it is comparable to or even larger than that of the true underlying covariate. An example of substantial measurement error is the HIV viral load in HIV/AIDS studies.

As introduced in Chapter 1, available functional modeling methods for Cox proportional hazards model include the conditional score (Tsiatis & Davidian, 2001), the parametric corrected score (Nakamura, 1992; Kong & Gu, 1999), and the non-parametric corrected score (Huang & Wang, 2000, 2006; Hu & Lin, 2004). The idea of the conditional score is to condition away the nuisance parameters based on certain sufficient statistics whereas the last two classes adopt a correction strategy by constructing a corrected estimating function with error-contaminated covariates that shares the same limit as a reference estimating function with true underlying covariates. If the reference estimating function admits consistent estimates only, the corrected estimating function shall inherit this property in a compact parameter space containing the true value. Conditional score and parametric corrected score are generally different. But in the case of the Cox proportional hazards model and normal measurement error, the conditional score and parametric corrected score estimators are asymptotically equivalent.

Although all three aforementioned methods produce consistent estimators, they all suffer from finite-sample pathological behaviors especially when the measurement error is substantial. Though noticed in some literature, the pathological behaviors of these estimation procedures were never well understood. Recently, Huang (2011)
proposed an approach to incorporate additional estimating functions which constrain the derivatives of the parametric corrected score for loglinear model. This approach effectively remedies those pathological behaviors and also considerably improves the estimation efficiency. Huang’s approach provides a promising general strategy to handle similarly ill-behaved estimating functions. Motivated by this approach, we will develop an estimation procedure for the Cox proportional hazards model with covariate measurement error.

In this chapter, we first conduct a detailed investigation on pathological behaviors of parametric corrected score and conditional score. After that, we propose an augmented estimation procedure in which additional estimating functions are added to the parametric corrected score for the proportional hazards model. In Section 3.2, we briefly describe the parametric corrected score and conditional score for the proportional hazards model and present the investigation results on the pathological behaviors when covariate measurement error is substantial. The proposed approach of incorporating additional estimating functions for the parametric corrected score is presented in Section 3.3. Simulation studies with practical sample size are reported in Section 3.4 together with an application to the ACTG 175 clinical trial data. Further discussion is given in Section 3.5. Technical details is collected in Section 3.6.
3.2 Parametric Corrected Score and Conditional Score and Their Pathological Behaviors

The proportional hazards model postulates that the cumulative hazard function \( \Lambda(\cdot) \) of survival time \( T \) of an individual with a \( p \)-vector of covariate \( Z \) has the form

\[
\Lambda(dt|Z) = \exp(\beta'Z)\Lambda_0(dt)
\]

where \( \beta \) is a \( p \)-vector parameter of interest and \( \Lambda_0(\cdot) \) is an unspecified baseline cumulative hazard function. Let \( C \) denotes the censoring time and adopt the usual independent censoring mechanism: given \( Z, C \) is independent of \( T \).

The observed data, \( (X_i, \Delta_i, Z_i, i = 1, \ldots, n) \), consist of \( n \) iid replicates of \( \{X \equiv T \wedge C, \Delta \equiv I(T \leq C), Z\} \). The standard inference procedure for Cox proportional hazards model is then to maximize the partial likelihood or, equivalently, to solve estimating function

\[
\xi^*(b, \tilde{\Lambda}_0(\cdot)) = n^{-1} \sum_{i=1}^{n} \int_0^\tau \left( \frac{1}{Z_i} \right) \left\{ dN_i(t) - Y_i(t) \exp(b'Z_i)d\tilde{\Lambda}_0(t) \right\}, \quad (3.1)
\]

where \( N_i(t) = I(X_i \leq t, \Delta_i = 1) \) is the counting process, \( Y_i(t) = I(X_i \geq t) \) is the at-risk process and \( \tau \) is a positive constant such that \( \Pr(T \geq \tau) > 0 \). Profiling out
\( \tilde{\Lambda}_0(\cdot) \), the estimating function for \( \beta \) alone is

\[
\xi(b) = n^{-1} \sum_{i=1}^{n} \int_{0}^{\tau} \left\{ Z_i - \frac{\sum_{j=1}^{n} Y_j(t)Z_j \exp(b'Z_j)}{\sum_{j=1}^{n} Y_j(t) \exp(b'Z_j)} \right\} dN_i(t).
\]

(3.2)

Estimating function (3.2) is actually the usual partial score function.

### 3.2.1 Parametric corrected score and conditional score

Split covariates \( Z = (Z_a^T, Z_e^T)^T \) where \( Z_a \) are those covariates that can be accurately measured and \( Z_e \) are covariates prone to measurement error and cannot be accurately measured. Though \( Z_e \) cannot be measured directly, we can observe them through their surrogates \( W_e \). Under the classical additive measurement error model, \( W_e = Z_e + \varepsilon_e \), where \( \varepsilon_e \) is the error vector and \( \varepsilon_e \) is assumed to be independent of \((T, C, Z)\).

In this chapter, we consider the situation that the distribution of \( \varepsilon_e \) is known; The situation where distribution of \( \varepsilon_e \) is unknown is studied in Chapter 4.

Let \( W = (Z_a, W_e) \) and \( \varepsilon = (0, \varepsilon_e) \). The observed data now consist of \((X_i, \Delta_i, W_i, i = 1, \ldots, n)\) in the presence of covariate measurement error. It is well known that naively replacing \( Z \) by \( W \) in estimating functions (3.1) or (3.2) could incur substantial estimation bias. Denote the cumulant-generating function of \( \varepsilon \) as \( \Omega(b) \equiv \log \mathbb{E} \{ \exp(b'\varepsilon) \} \) and its derivative \( \dot{\Omega}(b) \equiv \partial \Omega(b)/\partial b \). The parametric corrected score estimating
function is given by

$$\eta^*(b, \tilde{\Lambda}_0(\cdot)) = \frac{1}{n-1} \sum_{i=1}^{n} \int_0^\tau \left( W_i - \hat{\Omega}(0) \right) dN_i(t)$$

$$\left. - \int_0^\tau Y_i(t) \exp\{b'W_i - \Omega(b)\} \left( W_i - \hat{\Omega}(b) \right) \right| d\tilde{\Lambda}_0(t).$$

(3.3)

Further profiling out \( \tilde{\Lambda}_0(\cdot) \), we obtain the corrected score for \( \beta \) (Nakamura, 1992):

$$\eta(b) = \frac{1}{n-1} \sum_{i=1}^{n} \int_0^\tau \left\{ W_i + \hat{\Omega}(b) - \hat{\Omega}(0) - \frac{\sum_{j=1}^{n} Y_j(t) W_j \exp(b'W_j)}{\sum_{j=1}^{n} Y_j(t) \exp(b'W_j)} \right\} dN_i(t)$$

(3.4)

which has the same limit as reference (3.2) asymptotically for each and every finite \( b \). The estimation is then to find the zero crossing of the above estimating function. The consistency and asymptotic normality of corrected score estimator are later established by Kong & Gu (1999).

A root-consistent estimating function is an estimating function such that every zero-crossing is consistent and a normalized estimating function is root-consistent if its limit has a unique root at the estimand (Huang and Wang, 1999). By definition, reference (3.2) is a root-consistent estimating function and the new estimating function (3.4) shall inherit the root-consistency from (3.2). The root-consistency of (3.4) assures that in a compact parameter space containing the true parameter \( \beta \), the parametric corrected score will admit a unique root asymptotically and the root is consistent and asymptotically normal.

As for the conditional score estimating function for the Cox proportional haz-
ards model (Tsiatis & Davidian, 2001), when the measurement error is normally distributed and the variance matrix is $\Sigma$, it can be written as:

$$\eta_{\text{con}}(b) = n^{-1} \sum_{i=1}^{n} \int_{0}^{\tau} \left\{ W_i + \Sigma b \right\}$$

$$\sum_{j=1}^{n} \left[ W_j(t) \left( W_j + \Sigma b dN_j(t) \right) \exp(b'\left[ W_j + \Sigma b dN_j(t) \right]) \right] dN_i(t)$$

In fact, it can be shown that estimators from parametric corrected score and conditional score are asymptotically equivalent in the case of normal measurement error (Tsiatis & Davidian, 2001).

### 3.2.2 Pathological behaviors

When the magnitude of measurement error is small, the asymptotic results of parametric corrected score and conditional score provide a good approximation for practical purposes. However, when the measurement error increases, the pathological behaviors may start to arise (Song & Huang, 2005). These pathological behaviors include multiple zero-crossings or a single wrong zero-crossing that is inappropriate. These pathological behaviors may cause serious concerns when the measurement error is substantial and limit the applicability of parametric corrected score and conditional score in practice. But these pathological behaviors were never well understood or thoroughly investigated in the literature. In this section, we will conduct a detailed investigation of pathological behaviors for these two methods.
We first consider the parametric corrected score. Consider a single-covariate model with normal measurement error. In the absence of measurement error, the partial likelihood score function $\xi(b)$ is monotonically decreasing and has a unique root. The asymptotic result suggests that the parametric corrected score $\eta(b)$ should be monotonically decreasing as well in a compact parameter space containing the true parameter when the sample size is large. One may speculate the parametric corrected score to have an overall decreasing trend over the entire parameter space. But surprisingly, the overall trend of parametric corrected score in an unbounded parameter space is increasing. Function $\eta(b)$ takes a value of $-\infty$ when $b = -\infty$ and a value of $+\infty$ when $b = +\infty$. This observation suggests that the parametric corrected score $\eta(b)$ has an odd number of roots. In our numerical studies, only single- and triple-root patterns have been observed and two typical plots of parametric corrected score are illustrated in Figure 3.1. The same root patterns were observed in Huang’s (2011) investigation of loglinear model.

An alternative way of constructing corrected score is to first construct corrected partial likelihood and then take derivative. So if we characterize a root by increase or decrease of $\eta(\cdot)$ around it, the increasing and decreasing roots correspond to local minimizers and maximizers of corresponding corrected partial likelihood function, respectively. Therefore, an increasing root is considered as an inappropriate one. In the case of single-root pattern, the only root is increasing and thus inappropriate. For the triple-root pattern, an appropriate root exists since there is only one decreasing root. The single-root pattern is considered as root-finding failure.
Figure 3.1: Observed root patterns of the parametric corrected score $\eta(b)$. The true $\beta$ is -1 and these two corrected curves correspond to the same profile score (with true covariates). Portion of a corrected curve is thickened to indicate negative derivative.

Table 3.1: Prevalence (%) of single-root pattern for the parametric corrected score with $E(X) = 0$, $\text{Var}(X) = 1$, $\beta = -1$, and $\varepsilon \sim \text{Normal}(0,1)$.

<table>
<thead>
<tr>
<th>Censoring Rate</th>
<th>Distribution of X</th>
<th>Size</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>100</td>
</tr>
<tr>
<td>20%</td>
<td>Normal</td>
<td>68.0</td>
</tr>
<tr>
<td></td>
<td>Modified Chi-square</td>
<td>65.8</td>
</tr>
<tr>
<td></td>
<td>Uniform</td>
<td>64.5</td>
</tr>
<tr>
<td>40%</td>
<td>Normal</td>
<td>62.4</td>
</tr>
<tr>
<td></td>
<td>Modified Chi-square</td>
<td>64.4</td>
</tr>
<tr>
<td></td>
<td>Uniform</td>
<td>60.4</td>
</tr>
<tr>
<td>60%</td>
<td>Normal</td>
<td>61.0</td>
</tr>
<tr>
<td></td>
<td>Modified Chi-square</td>
<td>62.7</td>
</tr>
<tr>
<td></td>
<td>Uniform</td>
<td>58.5</td>
</tr>
</tbody>
</table>
We conduct a simulation study to exam the prevalence of single-root pattern. We consider a single covariate Z and generate it from various distributions: A) standard normal distribution, B) modified chi-square distribution, and C) uniform distribution with mean 0 and variance 1. To generate the modified chi-square distribution, the chi-square distribution with 1 degree of freedom was first truncated at 5 and then location-shifted to mean 0 and rescaled to variance 1. The measurement error follows standard normal distribution. The true coefficient was taken to be $\beta = -1$ and the baseline hazard is constant 1. Censoring was generated from a uniform distribution on $[0, \mu]$, where $\mu$ is chosen so that the censoring rate is ranged from 20% to 60%. These set-ups represent a practical scenario with substantial error contamination on the covariate. The results based on 1,000 iterations are reported in Table 3.1. The prevalence of single-root patten is similar across different scenarios with various censoring rates. When the sample size is 100, the percentage of single-root pattern under all three distributions are close to or over 60%. Even when the sample size increases to 800, the prevalence of single-root pattern is still quite high.

For the conditional score, the patterns are more complicated. When the absolute value of $\beta$ gets large, the estimating function fluctuates around zero and finally approaches zero as $\beta$ goes to infinity. Therefore the conditional score may have many zero-crossings. When the $\beta$ is not so extreme, two general patterns for the conditional scores are observed and plots from two simulation data sets are shown in Figure 3.2. As explained in previous section, conditional score and corrected score are asymptotically equivalent. Therefore an appropriate zero-crossing for conditional
score should be a decreasing one as well. In the first pattern, the conditional score has one zero-crossing close to the true parameter. This single zero-crossing is decreasing and thus an appropriate one. In the second pattern, the conditional score appears to have no proper zero-crossing near the truth though it may have multiple zero-crossings at extreme values of $\beta$. For conditional score, we define the estimator as the local point having the smallest $\ell^2$-norm in the neighborhood of the naive estimation. If the estimator is not a zero-crossing, we consider it as root-finding failure. Table 3.2 summarizes the prevalence of root-finding failure for the conditional score. The same simulation set up as in the corrected score is used. With a sample size of 100, the root-finding failure rate for conditional score varies from 3% to 5% for normal covariate and from 14% to 25% for modified chi-square covariate. As the sample size increases to 800, the failure rate drops to 0.2% for normal covariate but remains at

![Figure 3.2: Observed root patterns of the conditional score estimating function. The true $\beta$ is -1.](image-url)
least 14% for modified chi-square covariate.

Table 3.2: Prevalence (%) of root-finding failure for the conditional score with $E(X) = 0$, $\text{Var}(X) = 1$, $\beta = -1$, and $\varepsilon \sim \text{Normal}(0,1)$.

<table>
<thead>
<tr>
<th>Censoring Rate</th>
<th>Distribution of X</th>
<th>Size</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>100</td>
</tr>
<tr>
<td>20%</td>
<td>Normal</td>
<td>2.7</td>
</tr>
<tr>
<td></td>
<td>Modified Chi-square</td>
<td>14.4</td>
</tr>
<tr>
<td></td>
<td>Uniform</td>
<td>3.1</td>
</tr>
<tr>
<td>40%</td>
<td>Normal</td>
<td>4.6</td>
</tr>
<tr>
<td></td>
<td>Modified Chi-square</td>
<td>20.0</td>
</tr>
<tr>
<td></td>
<td>Uniform</td>
<td>3.1</td>
</tr>
<tr>
<td>60%</td>
<td>Normal</td>
<td>4.7</td>
</tr>
<tr>
<td></td>
<td>Modified Chi-square</td>
<td>24.6</td>
</tr>
<tr>
<td></td>
<td>Uniform</td>
<td>5.7</td>
</tr>
</tbody>
</table>

The above investigation results show that, in the presence of substantial measurement error, both parametric corrected score and conditional score suffer from severe finite-sample pathological behaviors. Therefore improvements are required for these methods to have practical applicability. We observe that an appropriate zero-crossing of an estimating function should be a decreasing one. This observation suggests that the trend of estimating function is also informative and could be taken into account in the estimate determination. By Taylor expansion, the trend of estimating function may be quantified by its derivative. Recognizing this feature, Huang (2011) proposed an approach to incorporate additional estimating functions which constrain the derivatives of the corrected score for the loglinear model. The estimation and inference are then accomplished by means of empirical likelihood. This approach effectively remedies the pathological behaviors of corrected score for loglinear model.
and also considerably improves the estimation efficiency. However, in the case of the Cox proportional hazards model, we are unable to construct additional estimating functions that effectively constrain the derivative of the parametric corrected score or conditional score because of the very nature of these two estimating functions. Nevertheless, Huang’s approach provides an insight into a new approach to address pathological behaviors of estimating functions. If we could identify additional estimating functions that do not share the same wrong roots as the original estimating function, then by combining the original and additional estimating functions, pathological behaviors could be reduced or eliminated. But if either the original estimating function or the additional estimating functions vanish to zero, then wrong root sharing would easily arise. We have shown in the simulation that the conditional score would vanish to zero when the absolute value of parameter becomes large. Thus, the trend pattern of the parametric corrected score is more desirable than that of the conditional score and our method will be developed for the parametric corrected score.

3.3 Improving Corrected Score

3.3.1 Augmented Estimation Method

Motivated by Huang’s (2011) trend-constrained corrected score, we first establish the following result:
Theorem 3.1. Under the proportional hazards model and the classical additive measurement error model,

\[
\mathcal{E} \left[ \int_0^\tau \frac{\partial^{k_1+\cdots+k_p} \exp \{ b'W - \Omega(b) \}}{\partial b_1^{k_1} \cdots b_p^{k_p}} dN_i(t) \right]_{b=0} = 0,
\]

\[
\int_0^\tau \frac{\partial^{k_1+\cdots+k_p} Y(t) \exp \{ b'W - \Omega(b) \}}{\partial b_1^{k_1} \cdots b_p^{k_p}} d\Lambda_0(t) \bigg|_{b=0} = 0,
\]

for \( k_l \geq 0, l = 1, \ldots, p \), where \( b_l, l = 1, \ldots, p \), is the \( l \)th element of \( b \).

Proof. Given

\[
\mathcal{E}[\exp \{ b'W - \Omega(b) \} | Z] = \exp(b'Z)
\]

under additive measurement error model, one may obtain

\[
\mathcal{E} \left[ \int_0^\tau Y \frac{\partial^{k_1+\cdots+k_p} \exp \{ b'W - \Omega(b) \}}{\partial b_1^{k_1} \cdots b_p^{k_p}} d\Lambda_0(t) \bigg| Z \right] = \prod_{l=1}^p Z_i^{k_l} \int_0^\tau Y(t) \exp(b'Z)d\Lambda_0(t)
\]

and

\[
\mathcal{E} \left[ \int_0^\tau \frac{\partial^{k_1+\cdots+k_p} \exp \{ b'W - \Omega(b) \}}{\partial b_1^{k_1} \cdots b_p^{k_p}} dN(t) \bigg| Z \right] = \prod_{l=1}^p Z_i^{k_l} N(\tau).
\]

Then given the fact that \( M(t) = N(t) - \int_0^t Y(u) \exp(\beta'Z)d\Lambda_0(u) \) is a mean zero martingale, equation (3.6) is implied by the above two equations.

Equation (3.6) is useful in constructing additional estimating equations for \( \beta \).

When \( \sum_{l=1}^p k_l = 0 \) and 1, one may obtain the usual parametric corrected score. When \( \sum_{l=1}^p k_l = 2 \), the additional estimating functions are the upper triangular elements of
the following symmetric matrix:

$$n^{-1} \sum_{i=1}^{n} \int_{0}^{T} \left( \{W_i - \hat{\Omega}(0)\} \otimes 2 - \ddot{\Omega}(0) \right) dN_i(t)$$

$$- \int_{0}^{T} Y_i(t) \exp\{b'W_i - \Omega(b)\} \left[ \{W_i - \hat{\Omega}(b)\} \otimes 2 - \ddot{\Omega}(b) \right] d\tilde{\Lambda}_0(t).$$

By profiling out $\tilde{\Lambda}_0(\cdot)$, we obtain

$$n^{-1} \sum_{i=1}^{n} \int_{0}^{T} \left[ \{W_i - \hat{\Omega}(0)\} \otimes 2 - \ddot{\Omega}(0) \right]$$

$$- \frac{\sum_{j=1}^{n} Y_j(t) \exp\{b'W_j - \Omega(b)\} \left[ \{W_j - \hat{\Omega}(b)\} \otimes 2 - \ddot{\Omega}(b) \right]}{\sum_{i=j}^{n} Y_j(t) \exp\{b'W_j - \Omega(b)\}} dN_i(t)$$

(3.7)

The additional estimating functions would be helpful if both parametric corrected score and additional estimating function are close to 0 around the truth (not necessarily having roots) and the additional estimating function is not close to 0 when the parametric corrected score is close to 0 at any point far away from the truth. Figure 3.3 shows four typical patterns of parametric corrected score and corresponding additional estimating function based on equation (3.7) for a single-covariate model with true parameter $\beta = -1$. Plot (a) is the ideal scenario. The parametric corrected score have three zero-crossings. The additional estimating function shares the same decreasing zero-crossing as the parametric corrected score. Moreover, they do not share any wrong zero-crossings. In plot (b), the parametric corrected score have three zero-crossings and two of them are close to each other. In this case, discrimi-
Figure 3.3: Parametric corrected score and additional estimation function based on equation (3.7) for a single-covariate model with $\beta = -1$. 
nating the two roots around the truth is not very important since they are close to each other anyway. In plots (c) and (d), the parametric corrected score has a single wrong zero-crossing but the additional estimating function is not close to 0 near this wrong zero-crossing. Both estimating functions are close to 0 around the truth.

### 3.3.2 Estimation and inference

With the additional estimating functions, we have more estimating functions than the number of parameters. Available methods to synthesize estimating functions that exceed the number of parameters include empirical likelihood (Qin & Lawless, 1994) and quadratic inference function (QIF) method (Lindsay & Qu, 2003). In this research, we shall use the quadratic inference function method to determine the estimate since the estimating functions are not sums of iid terms, thus the empirical likelihood would be computational difficult.

Let $\varphi(b)$ denotes the estimating functions. $\varphi(b)$ is comprised of the original parametric corrected score and additional estimating functions. The quadratic inference function takes the form

$$Q(b; \hat{C}) = \varphi'(b)\hat{C}^{-1}(b)\varphi(b),$$

(3.8)

where $\hat{C}(b)$ is any consistent estimator for the asymptotic variance of $\varphi(b)$. Then the estimator is defined as the minimizer of (3.8) and is consistent for the true value
of \( \mathbf{b} \). Furthermore, the estimator is asymptotically efficient in the class of consistent estimators based on linear combination of parametric corrected score and additional estimating functions. The construction of a quadratic inference function helps to solve both aspects of pathological behaviors of the parametric corrected score. Firstly, in the case of multiple zero-crossings, the introduction of additional estimating functions helps to pick up the right zero-crossing out from multiple ones if the additional estimating functions do not share the same wrong roots as the original parametric corrected score. Secondly, the problem of no appropriate zero-crossing could be solved by minimizing the quadratic inference function.

We name the method to incorporate additional estimating functions based on (3.6) as the \textit{augmented parametric corrected score}. One important special case is the method incorporating the upper triangular elements of matrix (3.7) and was termed the \textit{second order augmented parametric corrected score}. Augmented parametric corrected scores with higher order are also available, with additional estimating functions corresponding to \( k_l \) such that \( \sum_{l=1}^{p} k_l > 2 \). In Section 3.4, we will conduct extensive simulation studies to evaluate the performance of the augmented parametric corrected score and its applicability in practice.

Figure 3.4 plots quadratic inference functions corresponding to the two datasets in Figure 3.1. The second order augmented parametric corrected score is adopted. The formula for a consistent estimator of the asymptotic variance of \( \varphi(\mathbf{b}) \) is given in Section 3.6.
Figure 3.4: Quadratic inference functions for the two datasets in Figure 3.1.

The interval estimation can be easily achieved by inverting the hypothesis testing statistics. As an inference function, $Q(b; \hat{C})$ has similar properties as the log-likelihood function (Lindsay & Qu, 2003):

(a) $Q(\beta_0) - Q(\hat{\beta})$ is asymptotically chi-squared with degree of freedom $p$;

(b) the profile test statistics $Q(\psi_0, \hat{\lambda}_0) - Q(\hat{\psi}, \hat{\lambda})$, where $(\psi, \lambda)$ is a partition of the parameter of $\beta$ and $\hat{\lambda}_0$ is the profiled estimate of $\lambda$ given $\psi = \psi_0$, is asymptotically chi-squared as a test of $H : \psi = \psi_0$, with degree of freedom equal to the dimension of $\psi$. 
3.4 Numerical Studies

3.4.1 Simulations

Extensive simulations studies were conducted to evaluate the performance of augmented parametric corrected score. For reference and comparison, the ideal, naive, regression calibration, conditional score, and parametric corrected score were also studied. The ideal estimator used the ordinary partial score function with true covariates and, of course, it is not a realistic estimator. The naive approach uses the mismeasured surrogates in place of true covariates in the partial score function. For the regression calibration method, $Z_e$ is replaced by $E(Z_e | Z_a, W_e)$ in the partial score function. For the proposed approach, the optimization algorithm of Nelder & Mead (1965) will be used.

As shown in previous section, both conditional score and parametric corrected score have high prevalence of root-finding failure. To permit a fair comparison, we propose the following re-defined conditional score and re-defined parametric corrected score. If an appropriate zero-crossing could be found, the estimators will take the value of zero-crossing. But if root-finding failure occurs, the estimators will be defined as the local minimizer in the $\ell^2$-norm of the estimating functions closest to the naive estimator. Operationally, we will use the following modified Newton-Raphson algorithm. We start with the naive estimator and calculate the Newton-Raphson step size. Since the goal is to find a root or local minimizer, we need prevent overshooting.
In our simulation, we cap the step size at 0.2. During each iteration, we compare the $\ell^2$-norm evaluated at new estimator to that evaluated at current estimator. If the $\ell^2$-norm evaluated at new estimator is smaller, then the new estimator will be accepted and the algorithm continues to the next iteration. Otherwise, we will halve the step size and calculate the new estimator again. Iterations will be repeated until that i) the absolute value of estimating function is less than $10^{-6}$ or ii) the step size has been halved for more than 10 time during any single iteration. If criteria i) is satisfied, the algorithm converges to a zero-crossing and a root is identified in this case. If criteria ii) is satisfied, the algorithm converges to a local minimizer and root-finding failure occurs. Comparing to the original definition that estimators take the values of zero-crossings of estimating function, this new definition actually benefits the conditional score and corrected score. As shown in Figure 3.1 and 3.2, in the case of root-finding failure, zero-crossings of conditional score and corrected score are at extreme and far away from the true value.

In the simulation, we consider both single- and double-covariate models. In the single-covariate models, the true covariate $X$ is of mean 0 and variance 1. The true regression coefficient was set to be -1 and the baseline hazard is constant 1. The measurement error follows the standard normal distribution. Two different distributions of $X$ were studied: A) standard normal distribution and B) modified chi-square distribution. To generate the modified chi-square distribution, the chi-square distribution with 1 degree of freedom was first truncated at 5 and then location-shifted to mean 0 and rescaled to variance 1. Censoring time was generated from a uniform
distribution on $[0, \mu]$ and we will consider two different settings of censoring rate at 20% and 60%.

In the double-covariate models, true covariate $X$ follows bivariate normal distribution with mean $(-1, 1)$ and correlation coefficient of 0.5. The first covariate was subject to a standard normal measurement error, whereas the second covariate was accurately measured. The regression coefficients were set to $(-1, 1)$ and the baseline hazard is constant 1. Censoring time was also generated from a uniform distribution on $[0, \mu]$, where $\mu$ is chosen so that the censoring rate is 20% or 60%.

Sample sizes 100, 200, 400, 800, and 1,600 were investigated. For each scenario, 1,000 samples were simulated. We report the results on point and interval estimation separately.

Table 3.3 and 3.4 summarize the simulation results on the estimators in the single-covariate models with censoring rates of 20% and 60% respectively. The quantile-quantile plots are shown in Figure 3.4 and 3.5. For each scenario, the mean bias, and standard deviation were calculated. For augmented parametric corrected score, three sets of additional estimating functions were considered where $k = 2, 3, 4$. As expected, the naive estimator has substantial bias under both scenarios. The regression calibration estimator shows moderate bias, with larger bias in modified chi-square covariate case than normal covariate case. The re-defined conditional score shows slight bias under both scenarios, probably due to its left skewness. The quantile-quantile plots show that the re-defined conditional score deviates from normality considerably even
Table 3.3: Simulation summary statistics for the single-covariate models with 20% censoring rate: Ideal, naive (NV), regression calibration (RC), re-defined conditional score (ConS), re-defined parametric corrected score (CS), and first k augmented parametric corrected score (ACS:k), k = 2, 3, 4.

<table>
<thead>
<tr>
<th>size</th>
<th>Ideal B SD</th>
<th>NV B SD</th>
<th>RC B SD</th>
<th>ConS F B SD</th>
<th>CS F B SD</th>
<th>ACS: 2 B SD</th>
<th>ACS: 3 B SD</th>
<th>ACS: 4 B SD</th>
</tr>
</thead>
<tbody>
<tr>
<td>100</td>
<td>-15 154</td>
<td>593 99</td>
<td>148 254</td>
<td>2.7 -376</td>
<td>920 68.0</td>
<td>-11 287</td>
<td>46 431</td>
<td>-43 414</td>
</tr>
<tr>
<td>200</td>
<td>-8 104</td>
<td>602 65</td>
<td>189 163</td>
<td>1.5 -251</td>
<td>719 52.1</td>
<td>-89 273</td>
<td>-60 330</td>
<td>-130 348</td>
</tr>
<tr>
<td>400</td>
<td>-5 73</td>
<td>601 48</td>
<td>195 113</td>
<td>0.6 -202</td>
<td>575 38.0</td>
<td>-146 282</td>
<td>-99 287</td>
<td>-153 313</td>
</tr>
<tr>
<td>800</td>
<td>1 49</td>
<td>604 32</td>
<td>206 76</td>
<td>0.3 -99</td>
<td>351 20.0</td>
<td>-129 267</td>
<td>-87 246</td>
<td>-83 226</td>
</tr>
<tr>
<td>1600</td>
<td>1 36</td>
<td>605 23</td>
<td>0</td>
<td>-43 200</td>
<td>5.6 -75</td>
<td>206 -45 163</td>
<td>-60 178</td>
<td>-62 166</td>
</tr>
<tr>
<td>100</td>
<td>-25 204</td>
<td>639 91</td>
<td>242 247</td>
<td>14.4 -485</td>
<td>1221 65.8</td>
<td>145 241</td>
<td>110 387</td>
<td>41 409</td>
</tr>
<tr>
<td>200</td>
<td>-5 136</td>
<td>640 63</td>
<td>267 155</td>
<td>12.2 -396</td>
<td>1022 57.9</td>
<td>47 241</td>
<td>-10 351</td>
<td>-81 381</td>
</tr>
<tr>
<td>400</td>
<td>-6 98</td>
<td>642 43</td>
<td>280 101</td>
<td>15.1 -329</td>
<td>895 52.3</td>
<td>-21 237</td>
<td>-10 285</td>
<td>-76 306</td>
</tr>
<tr>
<td>800</td>
<td>-3 70</td>
<td>641 32</td>
<td>281 72</td>
<td>13.9 -249</td>
<td>740 40.1</td>
<td>-70 245</td>
<td>-34 241</td>
<td>-81 237</td>
</tr>
<tr>
<td>1600</td>
<td>0 47</td>
<td>644 21</td>
<td>12.5 -139</td>
<td>486 30.3</td>
<td>-82 242</td>
<td>-9 152</td>
<td>-29 133</td>
<td>-42 139</td>
</tr>
</tbody>
</table>

Note: F: root-finding failure (%); B: mean bias (×10^3); SD: standard deviation(×10^3).
Table 3.4: Simulation summary statistics for the single-covariate models with 60% censoring rate: Ideal, naive (NV), regression calibration (RC), re-defined conditional score (ConS), re-defined parametric corrected score (CS), and first k augmented parametric corrected score (ACS:k), k = 2, 3, 4.

<table>
<thead>
<tr>
<th>size</th>
<th>Ideal</th>
<th>NV</th>
<th>RC</th>
<th>ConS</th>
<th>CS</th>
<th>ACS: 2</th>
<th>ACS: 3</th>
<th>ACS: 4</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>B   SD</td>
<td>B   SD</td>
<td>B   SD</td>
<td>F   B   SD</td>
<td>F   B   SD</td>
<td>F   B   SD</td>
<td>B   SD</td>
<td>B   SD</td>
</tr>
<tr>
<td>100</td>
<td>-24 206</td>
<td>550 134</td>
<td>57 335</td>
<td>4.7 -429 1014</td>
<td>61.0 -27 301</td>
<td>94 478</td>
<td>-24 502</td>
<td>-90 551</td>
</tr>
<tr>
<td>200</td>
<td>-15 142</td>
<td>562  88</td>
<td>109 206</td>
<td>2.5 -272 738</td>
<td>46.5 -101 303</td>
<td>10 298</td>
<td>-99 347</td>
<td>-188 395</td>
</tr>
<tr>
<td>400</td>
<td>-8   97</td>
<td>563  63</td>
<td>119 142</td>
<td>0.7 -197 578</td>
<td>32.1 -139 305</td>
<td>-27 235</td>
<td>-96 257</td>
<td>-152 294</td>
</tr>
<tr>
<td>800</td>
<td>-2   65</td>
<td>567  42</td>
<td>132  96</td>
<td>0.3 -100 379</td>
<td>16.0 -113 271</td>
<td>-33 189</td>
<td>-72 199</td>
<td>-103 223</td>
</tr>
<tr>
<td>1600</td>
<td>1    46</td>
<td>567  30</td>
<td>0    -51  238</td>
<td>4.9 -68 205</td>
<td>-17 128</td>
<td>-38 125</td>
<td>-57 148</td>
<td></td>
</tr>
</tbody>
</table>

Case B

<table>
<thead>
<tr>
<th>size</th>
<th>Ideal</th>
<th>NV</th>
<th>RC</th>
<th>ConS</th>
<th>CS</th>
<th>ACS: 2</th>
<th>ACS: 3</th>
<th>ACS: 4</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>B   SD</td>
<td>B   SD</td>
<td>B   SD</td>
<td>F   B   SD</td>
<td>F   B   SD</td>
<td>B   SD</td>
<td>B   SD</td>
<td>B   SD</td>
</tr>
<tr>
<td>100</td>
<td>-66 351</td>
<td>687 124</td>
<td>351 288</td>
<td>24.6 -259 1154</td>
<td>62.7 247 297</td>
<td>196 459</td>
<td>219 531</td>
<td>260 568</td>
</tr>
<tr>
<td>200</td>
<td>-24 239</td>
<td>691   84</td>
<td>370 190</td>
<td>21.7 -291 1042</td>
<td>57.3 133 287</td>
<td>81 359</td>
<td>44 429</td>
<td>54 496</td>
</tr>
<tr>
<td>400</td>
<td>-16 159</td>
<td>693   57</td>
<td>382 124</td>
<td>25.7 -202 841</td>
<td>50.9 57 268</td>
<td>37 322</td>
<td>-22 368</td>
<td>-52 423</td>
</tr>
<tr>
<td>800</td>
<td>-9 112</td>
<td>693   40</td>
<td>385  86</td>
<td>25.0 -172 684</td>
<td>42.6 -8 263</td>
<td>1 280</td>
<td>-55 276</td>
<td>-94 306</td>
</tr>
<tr>
<td>1600</td>
<td>0    78</td>
<td>697  28</td>
<td>21.5 -77 412</td>
<td>34.0 -29 258</td>
<td>-3 232</td>
<td>-29 194</td>
<td>-57 207</td>
<td></td>
</tr>
</tbody>
</table>

Note: Same as in that of Table 3.3
Figure 3.5: Quantile-quantile plots for $\beta$ in the single-covariate models with 20% censoring rate, where $\beta = -1$. Red, yellow, green, blue, and black correspond to sample sizes 100, 200, 400, 800, and 1,600.
Figure 3.6: Quantile-quantile plots for $\beta$ in the single-covariate models with 60% censoring rate, where $\beta = -1$. Red, yellow, green, blue, and black correspond to sample sizes 100, 200, 400, 800, and 1,600.
when the sample size is 1,600. It also has a much larger standard deviation comparing to the re-defined parametric corrected score and the augmented parametric corrected score. The re-defined parametric corrected score is unbiased for both scenarios and the standard deviation is small. All three augmented corrected scores are consistent under both scenarios, and they become less biased as sample size increases. In comparison with higher-order augmented corrected scores, the second order augmented corrected score seems more favorable overall in terms of bias and standard error. Compared to the re-defined parametric corrected score, the second order augmented corrected score has a larger standard deviation when the sample size is small. But the standard deviation of augmented corrected score decreases rapidly as the sample size increases and is smaller than that of the re-defined corrected score when then sample size is 1,600.

Table 3.5 and 3.6 show the simulation results for the double-covariate models. As expected, when two covariates are correlated, the measurement error generally has impact not only on the mismeasured covariate but also on that of the accurately measured one. The relative performance of all estimators is similar to what was observed in the single-covariate models. For multiple-covariate models, each set of additional estimating functions contains more than one element. For example, when $k = 2$, the additional estimating functions includes three elements in the upper triangle of matrix (3.7). Various augmented corrected score could be constructed depending on the additional estimating functions chosen. Since the second order augmented corrected score shows a more favorable overall performance in the single-covariates
models, we consider only the second order augmented corrected score and augmented corrected scores using a subset of the three elements in this simulation. Simulation results show that the second order augmented corrected score performs better than two other augmented corrected scores. The bias of the second order augmented corrected score reduced quickly as sample size increased. Meanwhile, it also has the smallest standard deviation.

Table 3.5: Simulation summary statistics for the double-covariate models with 20% censoring rate: Ideal, naive(NV), regression calibration (RC), re-defined conditional score (ConS), re-defined parametric corrected score (CS), and augmented parametric corrected score (ACS:1/3, ACS:2/3, ACS:2)

<table>
<thead>
<tr>
<th>size</th>
<th>Ideal B SD</th>
<th>NV B SD</th>
<th>RC B SD</th>
<th>ConS B SD</th>
<th>CS B SD</th>
<th>ACS: 1/3 B SD</th>
<th>ACS: 2/3 B SD</th>
<th>ACS: 2 B SD</th>
</tr>
</thead>
<tbody>
<tr>
<td>100</td>
<td>β₁ -23 175</td>
<td>637 101</td>
<td>66 381</td>
<td>57.5 215 348</td>
<td>72.0 267 322</td>
<td>170 532 245</td>
<td>514 308 538</td>
<td></td>
</tr>
<tr>
<td></td>
<td>β₂  23 171</td>
<td>-400 142</td>
<td>-113 246</td>
<td>-67 336</td>
<td>-100 327 60</td>
<td>-456 97 404</td>
<td>-142 399</td>
<td></td>
</tr>
<tr>
<td>200</td>
<td>β₁ -13 114</td>
<td>648 66</td>
<td>136 224</td>
<td>46.9 155 315</td>
<td>62.8 186 294</td>
<td>-18 382 42</td>
<td>375 92 339</td>
<td></td>
</tr>
<tr>
<td></td>
<td>β₂  16 117</td>
<td>-412 101</td>
<td>-154 161</td>
<td>-35 310</td>
<td>-47 286 54</td>
<td>349 18 327</td>
<td>-20 276</td>
<td></td>
</tr>
<tr>
<td>400</td>
<td>β₁ -5 82</td>
<td>649 48</td>
<td>168 144</td>
<td>33.4 65 297</td>
<td>46.2 71 280</td>
<td>-89 301 56</td>
<td>277 -28 255</td>
<td></td>
</tr>
<tr>
<td></td>
<td>β₂  4 80</td>
<td>-416 68</td>
<td>-177 103</td>
<td>5 257</td>
<td>15 243 85</td>
<td>269 60 233</td>
<td>40 210</td>
<td></td>
</tr>
<tr>
<td>800</td>
<td>β₁ -2 55</td>
<td>651 33</td>
<td>181 93</td>
<td>20.0 22 224</td>
<td>30.5 -3 252</td>
<td>-99 282 80</td>
<td>262 -56 220</td>
<td></td>
</tr>
<tr>
<td></td>
<td>β₂  1 56</td>
<td>-420 52</td>
<td>-185 74</td>
<td>16 186</td>
<td>44 203 88</td>
<td>255 72 230</td>
<td>52 186</td>
<td></td>
</tr>
<tr>
<td>1600</td>
<td>β₁ -2 40</td>
<td>652 23</td>
<td>8.0 13 178</td>
<td>10.5 -43 207</td>
<td>-71 219 67</td>
<td>204 -61 190</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>β₂  1 39</td>
<td>-419 34</td>
<td>21 134</td>
<td>45 155 59</td>
<td>180 57 163</td>
<td>51 145</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

F: root-finding failure (%); B: mean bias (×10³); SD: standard deviation(×10³).
ACS: 1/3 and ACS: 2/3 correspond to the additional estimating functions containing the (1,1) element and the first row elements of matrix (3.7), respectively. ACS: 2 is the second order augmented parametric corrected score.

Table 3.8 and 3.8 report the coverage of three types of 95% confidence intervals
Table 3.6: Simulation summary statistics for the double-covariate models with 60% censoring rate: Ideal, naive(NV), regression calibration (RC), re-defined conditional score (ConS), re-defined parametric corrected score (CS), and augmented parametric corrected score (ACS:1/3, ACS:2/3, ACS:2)

<table>
<thead>
<tr>
<th>size</th>
<th>Ideal</th>
<th>NV</th>
<th>RC</th>
<th>ConS</th>
<th>CS</th>
<th>ACS: 1/3</th>
<th>ACS: 2/3</th>
<th>ACS: 2</th>
</tr>
</thead>
<tbody>
<tr>
<td>100</td>
<td>β₁</td>
<td>-32</td>
<td>242</td>
<td>-139</td>
<td>613</td>
<td>43.3</td>
<td>96.4</td>
<td>64.8</td>
</tr>
<tr>
<td></td>
<td>β₂</td>
<td>33</td>
<td>245</td>
<td>-403</td>
<td>361</td>
<td>-27</td>
<td>72.362</td>
<td>-50</td>
</tr>
<tr>
<td>200</td>
<td>β₁</td>
<td>-20</td>
<td>160</td>
<td>5926</td>
<td>36.1</td>
<td>29.393</td>
<td>57.8</td>
<td>145.3</td>
</tr>
<tr>
<td></td>
<td>β₂</td>
<td>21</td>
<td>161</td>
<td>-812</td>
<td>24</td>
<td>-8301</td>
<td>-7310</td>
<td>-4126</td>
</tr>
<tr>
<td>400</td>
<td>β₁</td>
<td>-6</td>
<td>109</td>
<td>97182</td>
<td>21.5</td>
<td>-2293</td>
<td>40.4</td>
<td>45.279</td>
</tr>
<tr>
<td></td>
<td>β₂</td>
<td>6</td>
<td>111</td>
<td>106137</td>
<td>31225</td>
<td>36254</td>
<td>79301</td>
<td>35246</td>
</tr>
<tr>
<td>800</td>
<td>β₁</td>
<td>-1</td>
<td>72</td>
<td>62145</td>
<td>112120</td>
<td>10.82349</td>
<td>25.716257</td>
<td>-73267</td>
</tr>
<tr>
<td></td>
<td>β₂</td>
<td>2</td>
<td>76</td>
<td>36969</td>
<td>11594</td>
<td>34198</td>
<td>55214</td>
<td>72236</td>
</tr>
<tr>
<td>1600</td>
<td>β₁</td>
<td>-2</td>
<td>53</td>
<td>62233</td>
<td>7.624186</td>
<td>11.346212</td>
<td>55193</td>
<td>40171</td>
</tr>
<tr>
<td></td>
<td>β₂</td>
<td>0</td>
<td>53</td>
<td>-37146</td>
<td>28142</td>
<td>47161</td>
<td>46162</td>
<td>36140</td>
</tr>
</tbody>
</table>

Note: Same as in that of Table 3.5
in the single- and double-covariate models. All three scenarios use the second order augmented corrected score. In constructing the confidence interval, we use two different approaches: inverting the hypothesis testing statistics as introduced in Section 3.2.2, and the Wald-type confidence interval. For the former, we use two critical values based on the asymptotic chi-square distribution and the bootstrap calibration (Efron & Tibshirani, 1993, chap. 12). Bootstrap size of 500 is used for the bootstrap calibration. The Wald-type confidence interval and test based one with chi-square distribution critical values have poor coverage when the sample size is small, but improve with larger sample size. The coverage probability of test based confidence interval using bootstrap-calibrated critical value is close to the nominal level of 95% for all sample sizes.

Table 3.7: Coverage of 95% confidence interval for the second order augmented parametric corrected score with 20% censoring rate. C (chi-square distribution), BC (bootstrap calibration) and W (Wald-type) indicate the type of confidence interval.

<table>
<thead>
<tr>
<th></th>
<th>size</th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>100</td>
<td>200</td>
<td>400</td>
<td>800</td>
<td>1600</td>
<td></td>
</tr>
<tr>
<td></td>
<td>C</td>
<td>BC</td>
<td>W</td>
<td>C</td>
<td>BC</td>
<td>W</td>
</tr>
<tr>
<td></td>
<td>90.2</td>
<td>96.8</td>
<td>87.0</td>
<td>87.1</td>
<td>96.6</td>
<td>91.1</td>
</tr>
<tr>
<td></td>
<td>92.9</td>
<td>88.8</td>
<td>92.2</td>
<td>87.6</td>
<td>94.3</td>
<td>85.6</td>
</tr>
<tr>
<td></td>
<td>82.0</td>
<td>94.5</td>
<td>78.4</td>
<td>86.3</td>
<td>95.4</td>
<td>87.3</td>
</tr>
<tr>
<td></td>
<td>75.8</td>
<td>94.0</td>
<td>83.3</td>
<td>86.8</td>
<td>93.8</td>
<td>92.3</td>
</tr>
<tr>
<td></td>
<td>79.2</td>
<td>93.0</td>
<td>83.9</td>
<td>86.8</td>
<td>93.8</td>
<td>92.3</td>
</tr>
</tbody>
</table>

Single-covariate: Case A
Single-covariate: Case B
Double-covariate
Table 3.8: Coverage of 95% confidence interval for the second order augmented parametric corrected score with 60% censoring rate. C (chi-square distribution), BC (bootstrap calibration) and W (Wald-type) indicate the type of confidence interval.

<table>
<thead>
<tr>
<th>size</th>
<th>100</th>
<th>200</th>
<th>400</th>
<th>800</th>
<th>1600</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>C</td>
<td>BC</td>
<td>W</td>
<td>C</td>
<td>BC</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Single-covariate: Case A</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>β</td>
<td>89.5</td>
<td>97.6</td>
<td>86.8</td>
<td>94.2</td>
<td>96.7</td>
</tr>
<tr>
<td>Single-covariate: Case B</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>β</td>
<td>87.5</td>
<td>93.7</td>
<td>73.9</td>
<td>89.7</td>
<td>92.4</td>
</tr>
<tr>
<td>Double-covariate</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>β₁</td>
<td>92.0</td>
<td>92.6</td>
<td>80.7</td>
<td>93.8</td>
<td>96.6</td>
</tr>
<tr>
<td>β₂</td>
<td>92.0</td>
<td>93.9</td>
<td>87.3</td>
<td>93.5</td>
<td>94.5</td>
</tr>
</tbody>
</table>

3.4.2 Application to ACTG 175 data

We apply the proposed approach to the AIDS Clinical Trial Group (ACTG) 175 study, a randomized clinical trial to evaluate four treatments in HIV-infected patients with an initial screening CD4 counts of between 200 and 500 per cubic millimeter. A total of 2,467 patients were enrolled and an almost equal number of patients were randomized into each of the four treatment groups: zidovudine alone (ZDV), zidovudine plus didanosine (ZDV + ddI), zidovudine plus zalcitabine (ZDV + ddC), and zalcitabine alone (ddC). We are interested in assessing the effect of baseline CD4 count on time to AIDS or death in antiretroviral-naive patients. Among all study patients, 1,067 had no prior antiretroviral therapy at enrollment, among which 1,036 patients had two CD4
measurements prior to the start of treatment and within 3 weeks of randomization.

For this analysis, we will consider the subset of 1,036 patients. The median length of follow-up was 32 months, and 85 events were observed.

Table 3.9: Comparison of regression coefficient estimators in the ACTG 175 data

<table>
<thead>
<tr>
<th></th>
<th>log(CD4)</th>
<th>ZDV + ddl</th>
<th>ZDV + ddC</th>
<th>ddl</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Est</td>
<td>Var</td>
<td>Est</td>
<td>Var</td>
</tr>
<tr>
<td>NV</td>
<td>-1.838</td>
<td>.1183</td>
<td>-.652</td>
<td>.0881</td>
</tr>
<tr>
<td>ConS</td>
<td>-2.172</td>
<td>.1698</td>
<td>-.659</td>
<td>.0916</td>
</tr>
<tr>
<td>CS</td>
<td>-2.177</td>
<td>.1678</td>
<td>-.659</td>
<td>.0900</td>
</tr>
<tr>
<td>ACS</td>
<td>-2.177</td>
<td>.1422</td>
<td>-.657</td>
<td>.0991</td>
</tr>
<tr>
<td></td>
<td>.1678</td>
<td>.0905</td>
<td>.0964</td>
<td>.0811</td>
</tr>
</tbody>
</table>

Note: For proposed estimator, the first row of variance estimator is obtained by inverting hypothesise testing statistics with bootstrap critical value; the second row of values is from sandwich variance estimator. Est: Estimated coefficient; Var: Variance.

We consider a Cox regression mode with 4 covariates: the true baseline log(CD4) and three indicators for the four treatments with ZDV group as the reference. We define the baseline log(CD4) as the average of the two log(CD4) measurements. From the duplicated measurements, we estimated the variance for error and true underlying log(CD4) to be 0.033 and 0.076 respectively. Note that the variance of measurement error is estimated using two replicated measurements of baseline log(CD). Therefore there is an additional estimating function for the variance of measurement error. Table 3.9 shows the estimators based on the naive, conditional score, parametric corrected score, and the proposed augmented corrected score. In comparison, the naive approach gives an coefficient estimator of log(CD4) with substantially smaller
magnitude. All the other approaches have similar estimates for all coefficients.

3.5 Discussion

For proportional hazards model with covariate measurement error, several consistent methods have been proposed under the functional modeling framework, including the conditional score and the parametric corrected score. However, when the measurement error is substantial to the extent that the errors are comparable to the true covariates in variance, both methods might experience pathological behaviors and root finding failure. Recently, Huang (2011) developed a novel approach to incorporate additional estimating functions which constrain the derivatives of the parametric corrected score. That approach proves effective and eliminates finite sample pathological behaviors of parametric corrected score for the loglinear model. Motivated by Huang’s (2011) approach, we conduct an investigation on the pathological behaviors of parametric corrected score and conditional score and propose an augmented parametric corrected score for the proportional hazards model by incorporating additional estimating functions to the original parametric corrected score. Results of simulation studies show the proposed approach is effective in eliminating pathological behaviors even with small sample size and substantial measurement error. The variance of proposed estimator appears to be larger than the parametric corrected score when sample size is smaller than 400, but it decreases rapidly and become smaller than the parametric corrected score as the sample size increases to 400.
In this chapter, we have only considered the situation where the distribution of the measurement error is known. With additional data available on the measurement error, the parametric distribution imposed on the measurement error may be spared. With the availability of replicated mismeasured covariates, Huang & Wang (2000) developed a nonparametric corrected score method for the proportional hazards model. In Chapter 4, we will extend the approach of incorporating additional estimating functions to the nonparametric corrected score.

3.6 Appendix: Asymptotic Variance of Estimating Function

For simplicity, we consider the single-covariate model given in Section 3.2.2. The second order augmented corrected score has the form

\[
\varphi(b) = n^{-1} \sum_{i=1}^{n} \int_0^{\tau} \left\{ \frac{(W_i + b\sigma^2)}{W_i^2 - \sigma^2} - \frac{\sum_{j=1}^{n} Y_j(t) \exp(bW_j)(W_j - (\sum_{j=1}^{n} W_j) - \sigma^2)}{\sum_{j=1}^{n} Y_j(t) \exp(bW_j)} \right\} dN_i(t).
\]

With functional delta method (Huang & Wang, 2000), straightforward algebra gives

\[
n^{1/2} \varphi(b) = n^{-1/2} \sum_{i=1}^{n} (B_{i1} - B_{i2}) + o_p(1)
\]
where

\[
B_{i1} = \int_0^\tau \left\{ \frac{W_i + b\sigma^2}{W_i^2 - \sigma^2} - \frac{E \{ Y(t) \exp(bW) \} \{ W_i \}}{E \{ Y(t) \exp(bW) \}} \right\} dN_i(t), \quad \text{and}
\]

\[
B_{i2} = \int_0^\tau \left\{ \frac{\exp(bW_i) \{ W_i \}}{E \{ Y(t) \exp(bW) \}} - \frac{E \{ Y(t) \exp(bW) \} \{ W_i \}}{[E \{ Y(t) \exp(bW) \}]^2} \right\} Y_i(t) d\hat{E}N(t).
\]

Thus, \( n^{1/2} \varphi(b) \) is asymptotically a sum of iid random variables.

For fixed \( b \), \( n^{1/2} \varphi(b) \) is asymptotically normal with a covariance matrix \( \Sigma(b) \) that can be consistently estimated by

\[
\hat{\Sigma}(b) = n \sum_{i=1}^n \{ \omega_i(b) - \bar{\omega}(b) \} \{ \omega_i(b) - \bar{\omega}(b) \}^t
\]

where \( \omega_i(b) = n^{-1} (B_{i3} - B_{i4}) \), \( \bar{\omega}(b) = n^{-1} \sum_{i=1}^n \omega_i(b) \), and \( B_{i3} \) and \( B_{i4} \) are defined as

\[
B_{i3} = \int_0^\tau \left\{ \frac{W_i + b\sigma^2}{W_i^2 - \sigma^2} - \frac{\hat{E} \{ Y(t) \exp(bW) \} \{ W_i \}}{\hat{E} \{ Y(t) \exp(bW) \}} \right\} dN_i(t), \quad \text{and}
\]

\[
B_{i4} = \int_0^\tau \left\{ \frac{\exp(bW_i) \{ W_i \}}{\hat{E} \{ Y(t) \exp(bW) \}} - \frac{\hat{E} \{ Y(t) \exp(bW) \} \{ W_i \}}{[\hat{E} \{ Y(t) \exp(bW) \}]^2} \right\} Y_i(t) d\hat{E}N(t).
\]

Asymptotic variance of other augmented corrected scores could be derived similarly.
Chapter 4

Augmented Nonparametric Corrected Score for Proportional Hazards Model with Covariate Measurement Error

4.1 Introduction

In Chapter 3, to address the issue of finite-sample pathological behaviors of the parametric corrected score, we propose an approach to incorporate additional estimating functions. In that topic, we consider the situation where the distribution of mea-
Sure measurement error is known. But in practice, it is common that the measurement error distribution is unknown but rather replication data are available for error-prone covariates. Based on replicated mismeasured covariates, Huang & Wang (2000) proposed a nonparametric corrected score for the Cox proportional hazards model under additive measurement error model and the resulting regression coefficient estimators are shown to be consistent and are asymptotically normal.

In this chapter, we conduct an investigation on the pathological behaviors of nonparametric corrected score and extend the approach developed in Chapter 3 to the nonparametric corrected score for the proportional hazards model. In Section 4.2, we briefly describe the nonparametric corrected score method and investigate its finite sample behaviors. The proposed approach of incorporating additional estimating functions is presented in Section 4.3. Simulation study results and an application to the ACTG 175 clinical trial data is summarized in Section 4.4. Further discussion is given in Section 4.5.
4.2 Nonparametric Corrected Score and Pathological Behaviors

4.2.1 Nonparametric Corrected Score

In this section, we will briefly describe the nonparametric corrected score of Huang & Wang (2000) and study its pathological behaviors when the measurement error is substantial.

First note that the partial score function (3.2) can be rewritten as a function of four empirical processes

\[ \xi(b) = \int_0^T \left[ d\hat{E}\{ZN(t)\} - \frac{\hat{E}\{Y(t)Z\exp(b'Z)\}}{\hat{E}\{Y(t)\exp(b'Z)\}} d\hat{E}\{N(t)\} \right] \]  

where \( \hat{E} \) represents sample empirical mean as defined in Chapter 2. With the functional representation of \( \xi(b) \), we can see clearly its limit:

\[ \tilde{\xi}(b) = \int_0^T \left[ d\mathcal{E}\{ZN(t)\} - \frac{\mathcal{E}\{Y(t)Z\exp(b'Z)\}}{\mathcal{E}\{Y(t)\exp(b'Z)\}} d\mathcal{E}\{N(t)\} \right] \]  

Split covariates \( Z = (Z_e^T, Z_a^T)^T \), where \( Z_a \) are accurately measured and \( Z_e \equiv (Z_1, \ldots, Z_L)^T \) are subject to an additive measurement error \( \epsilon_e = (\epsilon_1, \ldots, \epsilon_L)^T \) are mutually independent and independent of all other variables. Under the additive measurement error model, for each \( Z_l, l = 1, \ldots, L \), a finite number of \( R_l \geq 2 \)
surrogates $W_i^{[R_l]} \equiv \{W_{lm} : m = 1, \ldots, R_l\}$ are observed, where $W_{lm} = Z_l + \epsilon_{lm}$.

These error $\epsilon_{lm}$ are iid replicate of $\epsilon_l$. The observed data,

$$\{X_i; Y_i; W_i^{[R_l]} : l = 1 \ldots, L; Z_{ai}\}, i = 1, \ldots, n,$$

consist of $n$ iid replicates of $\{X; Y; W_i^{[R_l]} : l = 1 \ldots, L; Z_a\}$.

Pick arbitrarily two replicates from each $W_i^{[R_l]}$ where $l = 1, \ldots, L$ to form two vectors $W^{(r)} \equiv (W_1^{(r)T}, \ldots, W_L^{(r)T}, Z_a^T)^T$, $r = 1, 2$. Since the selection of $W^{(1)}$ and $W^{(2)}$ is arbitrary, $\prod_{l=1}^L R_l (R_l - 1)$ different permutations can be formed. The following nonparametric corrected estimating function is proposed by Huang & Wang (2000):

$$\hat{\xi}(b) = \int_0^T \left[ d\hat{\mathcal{E}}\{ AW^{(1)}N(t)\} - \frac{\hat{\mathcal{E}}\{ AY(t)W^{(1)} \exp(b'W^{(2)})\}}{\hat{\mathcal{E}}\{ AY(t) \exp(b'W^{(1)})\}} d\hat{\mathcal{E}}\{ N(t)\} \right] \quad (4.3)$$

where $\mathcal{A}$ denotes the operator averaging over all the different permutations of $W^{(1)}$ and $W^{(2)}$. Estimating function (4.3) converges to the same limit as reference (4.1) and the resulting regression coefficient estimators are shown to be consistent and are asymptotic normal.

### 4.2.2 Pathological Behaviors

Consider a single covariate model. In the absence of measurement error, the partial score function is monotonically decreasing and has a unique zero-crossing. But with
measurement error, we observe five different patterns for the nonparametric corrected score estimating functions and typical plots from simulated data sets are shown in Figure 4.1.

Pattern (a) in Figure 4.1 has no zero-crossing and represents a case of root-finding failure. Pattern (b) is monotonically decreasing with one zero-crossing and is considered as the ideal case. A correct root can be easily identified for this pattern. Though pattern (c) also has only one zero-crossing, it is an increasing root and is also considered as root finding failure. Pattern (d) has two zero-crossings that are far away from each other. With a properly chosen numerical algorithm and initial value, the correct root could be identified. Pattern (e) has two zeros-crossings that are both close to the true parameter.

We conducted a simulation study to investigate the prevalence of root-finding failure in the presence of substantial measurement error. We use the same simulation set up, definition of root-finding failure and modified Newton-Raphson algorithm as described in Chapter 3. For each covariates subject to measurement error, two replicated surrogates were generated. As shown in Table 4.1, in the single-covariate models, the failure rate is about 11-14% for normal covariate model and 27-29% for modified chi-square covariate model when the sample size is 100. The root-finding failure rate decreases as the sample size increases. But even when the sample size increases to 800, the failure rate is still 11-16% for modified chi-square covariate model.
Figure 4.1: Observed root patterns of the nonparametric corrected score estimating function. The true $\beta$ is -1.
Table 4.1: Prevalence (%) of root-finding failure for the nonparametric corrected score with $E(X) = 0$, $\text{Var}(X) = 1$, $\beta = -1$, and $\varepsilon \sim \text{Normal}(0,1)$.

<table>
<thead>
<tr>
<th>Censoring Rate</th>
<th>Distribution of $X$</th>
<th>Size 100</th>
<th>200</th>
<th>400</th>
<th>800</th>
</tr>
</thead>
<tbody>
<tr>
<td>20%</td>
<td>Normal</td>
<td>10.7</td>
<td>7.4</td>
<td>2.6</td>
<td>1.1</td>
</tr>
<tr>
<td></td>
<td>Modified Chi-square</td>
<td>27.0</td>
<td>20.9</td>
<td>14.0</td>
<td>11.2</td>
</tr>
<tr>
<td></td>
<td>Uniform</td>
<td>12.5</td>
<td>7.9</td>
<td>3.8</td>
<td>1.3</td>
</tr>
<tr>
<td>40%</td>
<td>Normal</td>
<td>10.9</td>
<td>6.8</td>
<td>2.7</td>
<td>1.1</td>
</tr>
<tr>
<td></td>
<td>Modified Chi-square</td>
<td>29.2</td>
<td>23.5</td>
<td>17.3</td>
<td>14.5</td>
</tr>
<tr>
<td></td>
<td>Uniform</td>
<td>12.3</td>
<td>8.4</td>
<td>3.5</td>
<td>1.3</td>
</tr>
<tr>
<td>60%</td>
<td>Normal</td>
<td>13.8</td>
<td>7.5</td>
<td>2.6</td>
<td>1.0</td>
</tr>
<tr>
<td></td>
<td>Modified Chi-square</td>
<td>28.4</td>
<td>23.9</td>
<td>22.0</td>
<td>15.7</td>
</tr>
<tr>
<td></td>
<td>Uniform</td>
<td>13.9</td>
<td>10.0</td>
<td>4.9</td>
<td>2.2</td>
</tr>
</tbody>
</table>

4.3 Improving Nonparametric Corrected Score

Let $Z = (Z_1, Z_2, \ldots, Z_p)^T$, where $Z_l$ is the $l$th element of vector $Z$. Consider the following class of estimating equations:

$$\int_0^T \left[ d\hat{E}\{Z_1^{k_1} \cdots Z_p^{k_p} N(t)\} - \frac{\hat{E}\{Y(t)Z_1^{k_1} \cdots Z_p^{k_p} \exp(b'Z)\}}{\hat{E}\{Y(t) \exp(b'Z)\}} \right] d\hat{E}\{N(t)\}, \text{where } \sum_{l=1}^p k_l = k. \tag{4.4}$$

Equations (4.4) present a class of estimating functions for parameter $b$ in the absence of measurement error. When $k = 1$, (4.4) reduces to the usual partial score function. Using (4.4) as a class of reference functions, their corresponding corrected version could serve the purpose of providing additional estimating functions to the nonparametric corrected score function.
Let $W^{(r)} = (W_1^{(r)}, \ldots, W_p^{(r)})^T, r = 1, 2$. Mimicking the argument in Huang & Wang (2000), we obtain the following additional estimating functions with mismeasured covariates sharing the same limit as reference (4.4):

$$
\int_0^\tau \left[ d\hat{E}\{A(W_1^{(1)})^{k_1} \cdots (W_p^{(1)})^{k_p} N(t) \}
- \frac{\hat{E}\{AY(t)(W_1^{(1)})^{k_1} \cdots (W_p^{(1)})^{k_p} \exp(b'W^{(2)})\}}{\hat{E}\{AY(t) \exp(b'W^{(1)})\}} d\hat{E}\{N(t)\} \right], \text{where } \sum_{l=1}^p k_l = k.
$$

Quadratic inference function method will then be used to combine nonparametric corrected score and these additional estimating functions. We will name this method the \textit{augmented nonparametric corrected score}. Specifically, when $\sum_{l=1}^p k_l = k$, this method will be called the $k$-th order augmented nonparametric corrected score. Figure 4.2 illustrates the quadratic inference functions for the five typical plots in Figure 4.1.
Figure 4.2: Quadratic inference functions for the data sets in Figure 4.1.
4.4 Numerical Studies

4.4.1 Simulations

The simulation set up is same as in that of Chapter 3. For each covariate subject to measurement error, two replicated surrogates were generated. For reference and comparison, the ideal, naive, regression calibration, and nonparametric corrected score estimators are also presented. The nonparametric corrected score is similarly re-defined and solved using a modified Newton-Raphson algorithm as in Chapter 3. For naive estimator, the average of two replicated measurements will be used in place of the true covariate in the usual partial score function. For regression calibration, the best linear approximation given in Carroll et al. (2006, chap. 4) will be used.

Table 4.2 and 4.3 summarize the simulation results on the estimators in the single-covariate models. Quantile-quantile plots are shown in Figure 4.3 and 4.4. Under each scenario, the mean bias, and standard deviation were calculated. In addition, median bias and a robustified standard deviation (IQR divided by 1.349) are also reported for nonparametric corrected score and the proposed approach. As expected, naive estimator incurs substantial bias under all scenarios. The regression calibration estimator shows small bias for normal covariate. But it shows substantial bias when the covariate is modified chi-square. The nonparametric corrected score estimator is consistent and shows little bias. For the single-covariate models with 60% censoring rate, the finding failure rate is about 11% for normal covariate when the sample size
Table 4.2: Simulation summary statistics for the single-covariate models with 20% censoring rate: Ideal, naive (NV), regression calibration (RC), re-defined nonparametric corrected score (NPC), and second order augmented corrected score (ACS:2).

<table>
<thead>
<tr>
<th>size</th>
<th>Ideal B SD</th>
<th>NV B SD</th>
<th>RC B SD</th>
<th>NPC F B SD</th>
<th>ACS: 2 B SD</th>
</tr>
</thead>
<tbody>
<tr>
<td>Case A</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>100</td>
<td>-11 152</td>
<td>421 119</td>
<td>111 212</td>
<td>10.7 -349 -97 842 466</td>
<td>-633 -105 2121 442</td>
</tr>
<tr>
<td>200</td>
<td>-5 106</td>
<td>429 79</td>
<td>134 139</td>
<td>7.4 -214 -64 565 330</td>
<td>-385 -67 1510 322</td>
</tr>
<tr>
<td>400</td>
<td>-4 72</td>
<td>435 55</td>
<td>148 95</td>
<td>2.6 -92 -28 332 191</td>
<td>-167 -20 1025 199</td>
</tr>
<tr>
<td>800</td>
<td>-2 51</td>
<td>437 37</td>
<td>155 61</td>
<td>1.1 -33 -9 151 128</td>
<td>-35 -7 171 126</td>
</tr>
<tr>
<td>1600</td>
<td>-2 37</td>
<td>436 27</td>
<td>152 44</td>
<td>0.1 -21 -9 101 96</td>
<td>-22 -14 103 97</td>
</tr>
<tr>
<td>Case B</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>100</td>
<td>-28 202</td>
<td>492 108</td>
<td>212 198</td>
<td>27.0 -299 -41 882 484</td>
<td>-804 -90 2373 543</td>
</tr>
<tr>
<td>200</td>
<td>-10 136</td>
<td>496 75</td>
<td>236 127</td>
<td>20.9 -194 -46 626 368</td>
<td>-333 -51 1009 350</td>
</tr>
<tr>
<td>400</td>
<td>-9 92</td>
<td>500 52</td>
<td>244 88</td>
<td>14.0 -175 -24 583 304</td>
<td>-222 -17 1102 233</td>
</tr>
<tr>
<td>800</td>
<td>-6 67</td>
<td>498 37</td>
<td>245 62</td>
<td>11.2 -115 -35 346 244</td>
<td>-126 -24 646 161</td>
</tr>
<tr>
<td>1600</td>
<td>-2 46</td>
<td>500 26</td>
<td>250 43</td>
<td>5.2 -68 -14 241 161</td>
<td>-53 -12 241 115</td>
</tr>
</tbody>
</table>

Note: F: root-finding failure (%); B: bias (×10³); SD: standard deviation(×10³). For NPC and ACS: 2, The first column under B is mean bias and the second column is median bias; The first column under SD is the usual standard deviation and the second column is robustified standard deviation defined as IQR/1.349.
Table 4.3: Simulation summary statistics for the single-covariate models with 60% censoring rate: Ideal, naive (NV), regression calibration (RC), re-defined nonparametric corrected score (NPC), and second order augmented corrected score (ACS:2).

<table>
<thead>
<tr>
<th>size</th>
<th>Ideal B SD</th>
<th>NV B SD</th>
<th>RC B SD</th>
<th>NPC B SD</th>
<th>F B SD</th>
<th>ACS: 2 B SD</th>
</tr>
</thead>
<tbody>
<tr>
<td>Case A</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>100</td>
<td>-24 206 382 160</td>
<td>52 273</td>
<td>13.8 -338 -98 888 494</td>
<td>-463 -112 1799 422</td>
<td></td>
<td></td>
</tr>
<tr>
<td>200</td>
<td>-15 142 384 105</td>
<td>66 180</td>
<td>7.5 -205 -81 534 347</td>
<td>-271 -78 1025 304</td>
<td></td>
<td></td>
</tr>
<tr>
<td>400</td>
<td>-8 97 395 72</td>
<td>89 120</td>
<td>2.6 -79 -23 303 202</td>
<td>-85 -27 413 186</td>
<td></td>
<td></td>
</tr>
<tr>
<td>800</td>
<td>-2 65 400 51</td>
<td>99 83</td>
<td>1.0 -30 -5 173 139</td>
<td>-24 -2 173 128</td>
<td></td>
<td></td>
</tr>
<tr>
<td>1600</td>
<td>1 46 398 36</td>
<td>95 57</td>
<td>0 -21 -13 109 94</td>
<td>-14 -9 106 88</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Case B</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>100</td>
<td>-66 351 545 149</td>
<td>294 255</td>
<td>28.4 -215 55 922 523</td>
<td>-410 -22 1488 598</td>
<td></td>
<td></td>
</tr>
<tr>
<td>200</td>
<td>-24 239 557 102</td>
<td>329 161</td>
<td>23.9 -134 34 673 401</td>
<td>-298 -39 1106 474</td>
<td></td>
<td></td>
</tr>
<tr>
<td>400</td>
<td>-16 159 562 70</td>
<td>338 111</td>
<td>22.0 -155 -11 606 377</td>
<td>-355 -38 1169 405</td>
<td></td>
<td></td>
</tr>
<tr>
<td>800</td>
<td>-9 112 562 52</td>
<td>342 81</td>
<td>15.7 -108 9 426 309</td>
<td>-181 -16 720 254</td>
<td></td>
<td></td>
</tr>
<tr>
<td>1600</td>
<td>0 78 564 37</td>
<td>345 57</td>
<td>12.2 -81 -2 305 236</td>
<td>-107 -10 500 179</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Note: Same as in that of Table 4.2.
Figure 4.3: Quantile-quantile plots for $\beta$ in the single-covariate models with 20% censoring rate, where $\beta = -1$. Red, yellow, green, blue, and black correspond to sample sizes 100, 200, 400, 800, and 1,600.
Figure 4.4: Quantile-quantile plots for $\beta$ in the single-covariate models with 60% censoring rate, where $\beta = -1$. Red, yellow, green, blue, and black correspond to sample sizes 100, 200, 400, 800, and 1,600.
is 100. The failure rate deceases to 1% as the sample size increases to 800. The failure
rate is much higher for modified chi-square covariates. Even with a sample size of
800, the failure rate is around 11%. Similar root-finding failure rates are observed
for models with 20% censoring rate. The proposed augmented corrected score is
left skewed with substantial mean bias and large standard deviation. But the median
bias is close to zero and the robustified standard deviation is comparable to or smaller
than that of nonparametric corrected score. The quantile-quantile plots show that
the proposed estimator suffers from substantial skewness.

Table 4.4 and 4.5 report the simulation results with the double-covariate models.
The relative performance of most estimators are as expected and follow a similar
pattern to that in the single-covariate models.

The coverage of the 95% confidence intervals are summarized in Table 4.6 and
4.7. The test based confidence interval using chi-square distribution critical value
and Wald-type one appear to have poor coverage probability, whereas the one using
bootstrap-calibrated critical value has coverage probability close to nominal level of
95%.

4.4.2 Application to ACTG 175 data

We apply the proposed approach to the AIDS Clinical Trial Group (ACTG) 175 study.
Among 1,067 antiretroviral-naive patients at enrollment, 1,036 patients had two CD4
Table 4.4: Simulation summary statistics for the double-covariate models with 20% censoring rate: Ideal, naive (NV), regression calibration (RC), re-defined nonparametric corrected score (NPC), and second order augmented corrected score (ACS:2).

<table>
<thead>
<tr>
<th>size</th>
<th>Ideal</th>
<th>NV</th>
<th>RC</th>
<th>NPC</th>
<th>ACS: 2</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>B SD</td>
<td>B SD</td>
<td>B SD</td>
<td>F B SD</td>
<td>B SD</td>
</tr>
<tr>
<td>100</td>
<td>β₁ -23 175</td>
<td>463 122</td>
<td>61 271</td>
<td>30.5 -47</td>
<td>69 518 383</td>
</tr>
<tr>
<td></td>
<td>β₂ 23 171</td>
<td>-289 154</td>
<td>-88 211</td>
<td>100 -41</td>
<td>512 335</td>
</tr>
<tr>
<td>200</td>
<td>β₁ -13 114</td>
<td>477 85</td>
<td>100 174</td>
<td>21.0 -24</td>
<td>38 386 278</td>
</tr>
<tr>
<td></td>
<td>β₂ 16 117</td>
<td>-307 105</td>
<td>-122 140</td>
<td>57 -16</td>
<td>373 235</td>
</tr>
<tr>
<td>400</td>
<td>β₁ -5 82</td>
<td>485 58</td>
<td>137 115</td>
<td>9.4 -9 10</td>
<td>217 196</td>
</tr>
<tr>
<td></td>
<td>β₂ 4 80</td>
<td>-313 74</td>
<td>-137 94</td>
<td>23 -8</td>
<td>188 167</td>
</tr>
<tr>
<td>800</td>
<td>β₁ -2 55</td>
<td>482 39</td>
<td>135 78</td>
<td>4.5 -28</td>
<td>-14 166 150</td>
</tr>
<tr>
<td></td>
<td>β₂ 1 56</td>
<td>-313 50</td>
<td>-138 64</td>
<td>28 15</td>
<td>134 119</td>
</tr>
<tr>
<td>1600</td>
<td>β₁ -2 40</td>
<td>485 28</td>
<td>140 54</td>
<td>0.6 -25</td>
<td>-11 119 106</td>
</tr>
<tr>
<td></td>
<td>β₂ 1 39</td>
<td>-313 37</td>
<td>-141 46</td>
<td>20 13</td>
<td>93 92</td>
</tr>
</tbody>
</table>

Note: Same as in that of Table 4.2.

Table 4.5: Simulation summary statistics for the double-covariate models with 60% censoring rate: Ideal, naive (NV), regression calibration (RC), re-defined nonparametric corrected score (NPC), and second order augmented corrected score (ACS:2).

<table>
<thead>
<tr>
<th>size</th>
<th>Ideal</th>
<th>NV</th>
<th>RC</th>
<th>NPC</th>
<th>ACS: 2</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>B SD</td>
<td>B SD</td>
<td>B SD</td>
<td>F B SD</td>
<td>B SD</td>
</tr>
<tr>
<td>100</td>
<td>β₁ -32 242</td>
<td>429 172</td>
<td>2 347</td>
<td>27.8 -78</td>
<td>28 556 373</td>
</tr>
<tr>
<td></td>
<td>β₂ 33 245</td>
<td>-252 204</td>
<td>-38 270</td>
<td>113 -9</td>
<td>551 343</td>
</tr>
<tr>
<td>200</td>
<td>β₁ -20 160</td>
<td>437 119</td>
<td>40 230</td>
<td>17.3 -73</td>
<td>13 452 291</td>
</tr>
<tr>
<td></td>
<td>β₂ 21 161</td>
<td>-260 144</td>
<td>-62 186</td>
<td>91 9</td>
<td>432 265</td>
</tr>
<tr>
<td>400</td>
<td>β₁ -6 109</td>
<td>451 81</td>
<td>80 152</td>
<td>7.7 -35</td>
<td>9 304 230</td>
</tr>
<tr>
<td></td>
<td>β₂ 6 111</td>
<td>-268 103</td>
<td>-80 127</td>
<td>43 1</td>
<td>279 188</td>
</tr>
<tr>
<td>800</td>
<td>β₁ -1 72</td>
<td>452 54</td>
<td>81 101</td>
<td>2.8 -37</td>
<td>-16 183 167</td>
</tr>
<tr>
<td></td>
<td>β₂ 2 76</td>
<td>-273 68</td>
<td>-88 82</td>
<td>27 13</td>
<td>143 131</td>
</tr>
<tr>
<td>1600</td>
<td>β₁ -2 53</td>
<td>449 38</td>
<td>83 68</td>
<td>0.6 -32</td>
<td>-20 128 116</td>
</tr>
<tr>
<td></td>
<td>β₂ 0 53</td>
<td>-268 49</td>
<td>-84 57</td>
<td>24 15</td>
<td>99 89</td>
</tr>
</tbody>
</table>

Note: Same as in that of Table 4.2.
Table 4.6: Coverage of 95% confidence interval for the second order augmented non-parametric corrected score with 20% censoring rate. C (chi-square distribution), BC (bootstrap calibration) and W (Wald-type) indicate the type of confidence interval.

<table>
<thead>
<tr>
<th>size</th>
<th>100</th>
<th>200</th>
<th>400</th>
<th>800</th>
<th>1600</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>C</td>
<td>BC</td>
<td>W</td>
<td>C</td>
<td>BC</td>
</tr>
</tbody>
</table>
| Single-covariate: Case A  
\( \beta \)    | 79.6 93.9 86.7 | 81.5 96.2 90.5 | 85.3 96.2 90.6 | 86.7 95.5 91.3 | 88.4 95.2 90.6 |
| Single-covariate: Case B  
\( \beta \)    | 82.7 96.4 84.6 | 83.3 96.9 87.9 | 85.3 96.8 90.7 | 85.1 96.1 91.0 | 86.2 95.8 91.6 |
| Double-covariate  
\( \beta_1 \)  | 72.1 95.1 75.2 | 78.6 96.8 88.3 | 82.7 95.4 88.4 | 85.3 96.2 90.3 | 84.8 96.1 88.5 |
| \( \beta_2 \)  | 73.7 93.0 77.9 | 79.4 95.2 87.5 | 84.0 96.9 88.5 | 85.3 95.7 90.5 | 86.0 95.4 89.6 |

---

Table 4.7: Coverage of 95% confidence interval for the second order augmented non-parametric corrected score with 60% censoring rate. C (chi-square distribution), BC (bootstrap calibration) and W (Wald-type) indicate the type of confidence interval.

<table>
<thead>
<tr>
<th>size</th>
<th>100</th>
<th>200</th>
<th>400</th>
<th>800</th>
<th>1600</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>C</td>
<td>BC</td>
<td>W</td>
<td>C</td>
<td>BC</td>
</tr>
</tbody>
</table>
| Single-covariate: Case A  
\( \beta \)    | 84.0 94.0 89.7 | 84.6 95.8 93.5 | 89.6 95.4 94.1 | 88.9 95.0 93.9 | 89.3 95.0 92.0 |
| Single-covariate: Case B  
\( \beta \)    | 88.2 96.2 84.6 | 90.9 98.3 90.5 | 89.2 96.6 90.2 | 88.8 94.7 91.0 | 89.4 94.7 91.0 |
| Double-covariate  
\( \beta_1 \)  | 81.2 95.6 84.6 | 81.0 97.2 89.5 | 86.8 93.6 90.3 | 88.5 96.3 91.8 | 89.0 96.1 91.7 |
| \( \beta_2 \)  | 83.0 94.7 86.6 | 84.8 95.8 89.2 | 86.4 94.5 91.8 | 89.1 95.7 92.0 | 87.6 95.4 89.8 |
measurements prior to the start of treatment and within 3 weeks of randomization. For this analysis, we will consider the subset of 1,036 patients and use the two CD4 counts as replicated measurements of true baseline CD4 count.

We consider a Cox regression mode with 4 covariates: the true baseline log(CD4) and three indicators for the four treatments with ZDV group as the reference. Table 4.8 shows the estimators based on the naive, regression calibration, nonparametric-correction and the proposed approaches. In comparison, the naive approach gives an coefficient estimator of log(CD4) with substantially smaller magnitude. All the other approaches have similar estimates for all coefficients.

Table 4.8: Comparison of regression coefficient estimators in the ACTG 175 data

<table>
<thead>
<tr>
<th></th>
<th>log(CD4)</th>
<th>ZDV + ddl</th>
<th>ZDV + ddC</th>
<th>ddl</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Est</td>
<td>Var</td>
<td>Est</td>
<td>Var</td>
</tr>
<tr>
<td>NV</td>
<td>-1.838</td>
<td>.1183</td>
<td>-.652</td>
<td>.0881</td>
</tr>
<tr>
<td>RC</td>
<td>-2.235</td>
<td>.1756</td>
<td>-.649</td>
<td>.0882</td>
</tr>
<tr>
<td>NPC</td>
<td>-2.200</td>
<td>.1795</td>
<td>-.653</td>
<td>.0906</td>
</tr>
<tr>
<td>Proposed</td>
<td>-2.186</td>
<td>.2135</td>
<td>-.652</td>
<td>.0970</td>
</tr>
<tr>
<td></td>
<td>.1820</td>
<td>.0901</td>
<td>.0999</td>
<td>.0826</td>
</tr>
</tbody>
</table>

Note: For proposed estimator, the first row of variance estimator is obtained by inverting hypotheese testing statistics with bootstrap critical value; the second row of values is from sandwich variance estimator. Est: Estimated coefficient; Var: Variance.
4.5 Discussion

In this chapter, we have extended the approach of incorporating additional estimating functions developed in Chapter 3 into the nonparametric corrected score for Cox proportional hazards model. With replicated measurements of error-prone covariates available, this approach does not require any assumptions in addition to the Cox proportional hazards model and the additive measurement error model. This is an appealing feature because many assumptions can be difficult to verify in practice. Simulation studies show that the proposed approach produces small bias and is promising in eliminating root finding failure. The proposed estimator suffers from left skewness, especially in the case of modified chi-square covariate. In general, the variance of proposed estimator appears to be larger than the nonparametric corrected score when sample size is smaller than 400, but it becomes smaller than the nonparametric corrected score for sample size greater than 400.
Chapter 5

Summary and Future Work

5.1 Summary

In this dissertation research, we study three measurement error problems in the analysis of survival data.

We first study the analysis of recurrent events data under the accelerated failure time model in the presence of covariate measurement error. With replicated mis-measured covariates available, we propose an estimation procedure based on a novel identity. The proposed estimation procedure requires no distributional assumptions on either the true covariates or the error except for the boundedness of the latter. The resulting regression coefficient estimators are shown to be consistent and asymptotically normal. Simulation studies show the proposed procedure performance well
with practical sample size and moderate measurement error. We apply the proposed method to NPC clinical trial to illustrate its practical utility.

We then study the Cox proportional hazards model assuming the distribution of the measurement error is known. Both existing consistent methods, parametric corrected score (Nakamura, 1992) and conditional score (Tsiatis & Davidian, 2001), suffer from severe finite-sample pathological behaviors when the measurement error is substantial. We study the finite sample pathological behaviors of these two estimators and propose an augmented parametric corrected score by incorporating additional estimating functions to the original parametric corrected score. The estimation and inference are then accomplished by means of quadratic inference function. Simulation studies show the proposed approach is effective in eliminating finite-sample pathological behaviors even with small sample size and substantial measurement error. An application to ACTG 175 study is presented.

Furthermore, we consider the Cox proportional hazards model when the error distribution is completely unspecified but replicated mismeasured covariates are available. We applied the technique in the second topic to the nonparametric corrected score and the simulation study result shows that the proposed estimation procedure is promising in resolving those pathological behaviors.

For all three measurement error problems, the naive estimators are biased, which is obvious and well expected. Prediction and hypothesis testing are, however, different stories. In general, the presence of measurement error has no effect on prediction
problem. The surrogate $W$ is error-free as a measurement of itself. So if a model is built on the mismeasured covariate, predicting the response from the mismeasured covariate does not involve any bias. The effect of measurement error on hypothesis testing is much more complicated. See the monograph of Carroll et al. (2006) for a detailed discussion on the hypothesis testing problem. In general, the naive test of no effects due to $Z_e$ is still valid. But the naive test of no effects due to $Z_a$ is not valid except for under some restrictive assumptions.

### 5.2 Future Work

In this subsection we discuss some future research topics and possible extensions of this dissertation work.

In the first topic, the proposed method requires replicated measurements on error-prone covariates, which might not always be available in practice. For example, patients in Se group of the NPC trial did not have replicated baseline plasma Se measurements. In some other applications, due to financial constrains or various other reasons, the replication data may only be available for a subset of all subjects. This limits the applicability of the proposed method. One direction of future research is to relax this requirement and to develop an estimating function that can utilize the whole data set.

The method of incorporating additional estimating functions in the last two topics
are motivated by Huang’s (2011) investigation of loglinear model. But it turns out that we have a weaker claim than in Huang’s (2011) paper. In the case of loglinear model, the additional estimating functions effectively impose constrains on the derivative of corrected score. But for Cox regression model, we do not have such nice property because of the form of partial score function. Another issue is the efficiency. The proposed estimator should be asymptotically more efficient than the corrected score estimator since it involves more estimating functions. However, this research shows that having additional estimating function might not always provide better efficiency when sample size is small or moderate. How to improve the estimation efficiency is a tough problem to solve but an interesting topic for future research.

In this dissertation research, we adopt the classical measurement error model. In many applications, the measurement error models may be much more complex. For example, in nutritional studies, the observed surrogate may be a linear function of the true underlying covariate and several other covariates (Kipnis et al., 1999, 2001, 2003). One potential future research topic is to extend the proposed approaches to accommodate different measurement error models.
Bibliography


