

ESTIMATION OF SEMIVARYING COEFFICIENT MODELS FOR COUNTING
PROCESS WITH APPLICATIONS

by

Liqui Deng

A dissertation submitted to the faculty of
The University of North Carolina at Charlotte
in partial fulfillment of the requirements
for the degree of Doctor of Philosophy in
Applied Mathematics

Charlotte

2019

Approved by:

Dr. Yanqing Sun

Dr. Yang Li

Dr. Weihua Zhou

Dr. Weidong Tian

ABSTRACT

LIQIU DENG. Estimation of Semivarying Coefficient Models for Counting Process with Applications. (Under the direction of DR. YANQING SUN)

Recurrent events are very common in many different fields, including biological, medical, engineering and finance. Existing research have developed methodologies to model constant covariate effects and time-dependent covariate effects. However, in reality, for instance medical cases, covariate effects can be depending on other covariates as well. Therefore, in this dissertation, we investigate a semiparametric model for recurrent events, which incorporates both time-varying covariate effects and covariate-varying effect.

In our model, we use fixed parameters to model constant covariate effects, while we assume both time-dependent effects and covariate-varying effects to be unknown functions. An estimation procedure is proposed to estimate the unknow parameters and functions. Local linear smoothing method is adopted in our estimation procedure. Detailed computation is carried out by using Newton-Raphson iterative method. The asymptotic properties including asymptotic normality and consistency are established for the proposed estimators.

In order to assess the finite-sample performance of the proposed estimators and estimation procedure, simulation studies are conducted for different cases. The simulation results show that the proposed estimators perform very well with small bias and an empirical coverage probability close to its nominal level 95%.

In addition, the proposed model and methodologies are applied on the data from

the African American Study of Kidney Disease and Hypertension (AASK) trial. The data application is aiming to examining the treatment effects of three different design in the study.

ACKNOWLEDGMENTS

It has been a long journey before I get to this point. I would like to acknowledge the roles of all individuals that have helped me complete this dissertation and get through this long journey.

I would like to express my most sincere gratitude to my advisor Dr. Yanqing Sun, who encouraged me to pursue this study and provided me insightful guidance from the beginning to the finishing line. I am deeply grateful for her understanding, patience, enthusiasm to research, encouragement and inspiration, without any of which this dissertation would not have been completed. The weekly meetings we had in the past five years helped me go further in my research, while Friday lunches with Dr. Sun helped me grow to be a better woman. What I have learned from her will benefit me for the rest of my life.

To my committee members, Dr. Weihua Zhou, Dr. Yang Li, and Dr. Weidong Tian, I am deeply grateful for their valuable comments, advise, time and encouragement throughout my dissertation.

I would like to thank Dr. Yuanan Diao, Dr. Shaozhong Deng and Dr. Mohammad Kazemi for being extremely supportive and understanding during my Ph.D study. As a Ph.D student with two young kids, my life would have been much harder without your support.

I would like to thank student fellows at the department, Fei, Peilin, Yuehan, Li, Dan, Yetong and Sha, for your encouragement. Thank you for babysitting my kids from time to time when I needed help.

I would like to thank my parents and my brothers for their unconditional support and love. Special thanks to my children Nelson and Olivia. The love, smiles, hugs and laughters from them helped me get through the darkness during this journey.

Last of all, I owe my thanks to my husband Siyuan Huang for encouraging me in all of my pursuits and supporting me emotionally and financially. Thank you for your patience, company, inspiration, and endless love throughout this journey and my life.

TABLE OF CONTENTS

LIST OF FIGURES	ix
LIST OF TABLES	x
CHAPTER 1: INTRODUCTION	1
1.1. A Motivating Example	2
1.1.1. AASK Trial	2
1.1.2. Existing Research on AASK Trial	9
1.2. Literature Review	10
CHAPTER 2: SEMIPARAMETRIC MODEL WITH NONPARAMETRIC COVARIATE-VARYING EFFECTS	13
2.1. Model Description	13
2.2. Estimation Procedure	16
2.3. Computational Algorithm	19
2.4. Bandwidth Selection	20
CHAPTER 3: ASYMPTOTIC PROPERTIES	23
3.1. Notations	23
3.2. Asymptotic Properties	24
CHAPTER 4: SIMULATION STUDIES	27
4.1. Simulation on Single Event Data - Survival Analysis	27
4.1.1. Generating Single Event Data	27
4.1.2. Simulation Example	29
4.2. Simulation on Recurrent Events Data	38
4.2.1. Generating Recurrent Events Data	38

	viii
4.2.2. Simulation Example 1	39
4.2.3. Simulation Example 2	47
REFERENCES	55
APPENDIX A: PROOFS OF THE THEOREMS	59

LIST OF FIGURES

- FIGURE 1: Plots for Bias, CP, SSE and ESE for $n=400, 600, 800,$ and 1000 with $h_t = 0.4, h_u = 0.4$ for $\alpha_0(t) = -1.5 + 0.8t$ and $\alpha_1(t) = t$ for $0 \leq t \leq 2$. 36
- FIGURE 2: Plots for bias, CP, SSE and ESE for $n=400, 600, 800,$ and 1000 with $h_t = 0.4, h_u = 0.4$ for $\gamma(u) = -0.5u$. 37
- FIGURE 3: Plots for bias, CP, SSE and ESE for $n=400, 600, 800$ with $h_t=0.3, h_u=0.3$ for $\alpha_0(t) = 1.5 - \log(1 + t)$ for $0 \leq t \leq 5$. 45
- FIGURE 4: Plots for bias, CP, SSE and ESE for $n=400, 600, 800$ with $h_t=0.3, h_u=0.3$ for $\gamma(u) = \sqrt{u} - 2$. 46
- FIGURE 5: Plots for bias and CP for $n=400, 600, 800$ with $h_t = 0.3, h_u = 0.3$. Left panel is for $\alpha_1(t) = 2 - \log(1 + t)$. Right panel is for $\alpha_2(t) = \sin(0.2t)$. 53
- FIGURE 6: Plots for bias and CP for $n=400, 600, 800$ with $h_t = 0.3, h_u = 0.3$. Left panel is for $\gamma_1(u) = u - 1$. Right panel is for $\gamma_2(u) = \sqrt{u} - 2$. 54

LIST OF TABLES

TABLE 1: Blood pressure goals in AASK study and corresponding MAP range.	5
TABLE 2: Anti-hypertensive regimens and different dose levels in AASK study.	6
TABLE 3: Allocation of participants in different treatment groups in AASK study.	7
TABLE 4: Follow up visit schedule and measurement schedule for patients in AASK study.	8
TABLE 5: Summary of Bias, SSE, ESE and CP for $\hat{\beta}$ under model (4.6).	34
TABLE 6: Summary of RMSEs for $\hat{\alpha}_0(t)$, $\hat{\alpha}_1(t)$ and $\hat{\gamma}(u)$ under model (4.6).	35
TABLE 7: Summary of Bias, SSE, ESE and CP for $\hat{\beta}$ under model (4.7).	43
TABLE 8: Summary of RMSEs for $\hat{\alpha}_0(t)$ and $\hat{\gamma}(u)$ under model (4.7).	44
TABLE 9: Summary of Bias, SSE, ESE and CP for $\hat{\beta}$ under model (4.8).	51
TABLE 10: Summary of RMSEs for $\hat{\alpha}_1(t)$, $\hat{\alpha}_2(t)$, $\hat{\gamma}_1(u)$ and $\hat{\gamma}_2(u)$ under model (4.8).	52

CHAPTER 1: INTRODUCTION

Survival analysis aims at modeling time to certain event. It is of interest in many biological, medical, engineering and financial applications. For example, animals life in control and experimental groups, time to relapse of a certain disease after different treatments, first time failure of a mechanical part, or time to bankruptcy of a financial institution. By assuming survival time follows a probability distribution (e.g. Exponential, Gamma, Weibull, or Log-Normal etc.) that depends on covariates, a variety of parametric regression models are developed in order to investigate how survival time distributes given the covariates, as well as exploring how strongly survival time relates to the covariates. Another approach developed by Cox (1972) is a semi-parametric, which assumes a parametric and proportional structure that relates covariates to hazard rate while leaves the non-parametric part completely unspecified. Since Cox model does not require distribution assumption, it soon becomes one of the most popular survival regression methods.

The ordinary Cox regression model suffices in survival analysis in which events are assumed conditionally independent given covariates. However, it is common to observe that the event of interest occurs multiple times on one single subject, which is referred as recurrent events. Typical examples include recurrent episodes of a disease in patients, multiple admissions to hospitals, multiple ear infections on young age children, multiple delinquencies of credit card customers, earthquakes in one city,

repairs in machines and automobiles, etc. It is of interest to identify risk factors that are associated with the frequencies of recurrent events. In this dissertation, we aim at developing new methodologies for analyzing recurrent events data.

1.1 A Motivating Example

This dissertation is motivated by the African American Study of Kidney Disease and Hypertension (AASK) trial. In this section, we start with a description for the background of this clinical trial, experimental design and primary/secondary outcomes of this trial, followed by a short summary of existing research on this clinical trial study and our research question with this dataset.

1.1.1 AASK Trial

Studies have shown that hypertensive nephrosclerosis is the second leading cause of renal failure in African Americans and the leading cause of renal failure or end-stage renal disease (ESRD) (Institute of Medicine (US) Committee for the Study of the Medicare End-Stage Renal Disease Program, 1991). The AASK trial is a randomized clinical trial study, which aims to investigate the potential risk factors and the use of antihypertensive treatments on the progression of chronic kidney disease (CKD) on African Americans with hypertension.

A total of 2802 African American participants were recruited for the AASK trial. Physical measurements were taken for a screening process to determine patients' eligibility for this trial. Based on the results from a pre-trial screening process, 1708 participants were excluded from this study due to a variety of reasons, for instance, Patient refusal, medical exclusion, study team preference, etc. A total of 1094 par-

ticipants were enrolled in this clinical trial and randomized into different treatment groups.

A 3*2 experimental design was applied to randomize participants into different groups. The treatments considered in this trial include two levels of blood pressure (BP) control and three different anti-hypertensive regimens. The BP controls were defined by mean arterial pressure (MAP). Details for the two levels of BP controls are given by Table 1. Participants were randomized to one of the two BP control groups. The BP controls allocation was not blinded to neither investigators nor participants. The anti-hypertensive regimens considered in this trial include Angiotensin Converting Enzyme Inhibitors (ACEI), Calcium Channel Blocker (DHPCCB) and β -Blocker (BB). Different dose levels for each of those three regimens were used to maintain the patient's blood pressure within the assigned goal range. Table 2 below shows the different dose levels used in this trial. The allocation process of regimens was double blinded to both investigators and patients. A so-called "Double Dummy" system was used to accomplish medication masking in this trial. Each patient takes one tablet from either placebo or BB plus one capsule from either ACEI, DHPCCB or placebo. Table 3 summarizes the allocation of participants into different treatment groups.

The enrollment dates of patients range from March 1995 to September 1998. Patients were followed up until the end of this clinical trial September 2001. During follow up visits, a variety of physical measurements were recorded by AASK-certified personnel, including glomerular filtration rate (GFR), central serum measurements (for instance, sodium, urea nitrogen, phosphorus, etc.), central 24-hour urine measurements, fasting lipid profiles and quality of life measurements. Table 4 shows the

schedule for taking different measurements during follow up visits.

The primary outcome of AASK trial is GFR measure, which can measure how well a kidney is functioning. Specifically, GFR estimates the amount of blood that passes through the glomeruli each minute. Waste from the blood can be filtered by glomeruli in the kidneys. Besides the primary outcome, AASK trial defined several items as secondary outcomes, including the reduction in GFR measure by 50% or by 25 ml/min/1.73m², ESRD, and death. In addition, the AASK Clinical Outcome Committee also reviewed hospitalizations and deaths and then determined those events as cardiovascular (CV) events.

The objectives of AASK trial included the following.

1. To identify risk factors that are associated with progression of CKD on African Americans with hypertension.
2. To determine the effects of the anti-hypertensive drugs and the effects of lower than usual blood pressure control on African Americans with hypertension.

In addition, it is of interest to model the secondary recurrent CV events and explore the risk factors that are associated with such CV events. The following subsection summarizes some of the existing studies based on the data from AASK clinical trial.

Table 1: Blood pressure goals in AASK study and corresponding MAP range.

BP goal	MAP range
lower goal	$MAP 102 - 107 mmHg$ ($BP 135/85$ to $140/90 mmHg$)
usual goal	$MAP < 92 mmHg$ ($BP 115/80 mmHg$)

Table 2: Anti-hypertensive regimens and different dose levels in AASK study.

Regimen	Dose Level (mg)		
ACEI	2.5	5	10
DHPCCB	5	5	10
BB	50	100	200

Table 3: Allocation of participants in different treatment groups in AASK study.

BP Goal	ACEI	DHPCCB	BB	Total
low goal	215	215	110	540
usual goal	221	226	107	554
Total	436	441	227	1094

Table 4: Follow up visit schedule and measurement schedule for patients in AASK study.

Measurements	Schedule
GFR	3, 6, 12 months, then every 6 months thereafter
serum measurements	every 6 months
urine measurements; fasting lipid profiles; quality of life measurement	annually

1.1.2 Existing Research on AASK Trial

Many studies have been conducted based on this clinical trial (Norris et al., 2006; Beddhu et al., 2009; Ku et al., 2016; Wright et al., 2002; Lea et al., 2008; Brooks et al., 2008; Lea et al., 2003; Sika et al., 2007; Contreras et al., 2005; Raphael et al., 2010; Hong et al., 2015; Astor et al., 2008; Wang et al., 2006; Lea et al., 2005; Lewis et al., 2004; Gassman et al., 2003; Chen et al., 2015; Ku et al., 2018). In particular, Norris et al. (2006) examined 38 baseline risk factors by applying Cox regression and found a subset of factors associated with increased risk for progression to ESRD, including baseline proteinuria, serum creatinine, urea nitrogen, and phosphorus. By using Cox model, the research conducted by Beddhu et al. (2009) investigated the associations of serum alkaline phosphatase with cardiovascular (CV) events and all-cause death. Positive association between them were found to be significant. Ku et al. (2016) evaluated the association between APOL1 genotype and mortality risk. Their results show that Strict blood pressure control is significantly associated with decreased risk for death by APOL1 genotype.

The aforementioned studies focused on identifying risk factors for progression of CKD and CV events or the effects of treatments (drugs and blood pressure level) on progression of CKD or death. However, it is natural for one to wonder if the treatment effects would change after a specific event happens, for example patients enter ESRD, GFR drops by 20%, etc. If yes, one would further wonder how the treatment effects change after that specific time point? This dissertation aims to answer these questions. Therefore, a model that can incorporate not only covariate

effects but also covariate-varying effects is desired.

1.2 Literature Review

In the recent decades, it has become a great interest to model the occurrence of recurrent events. Intensive studies have been conducted to develop methodologies to analyze recurrent event data. For example, the hazard functions of gap times between recurrent events were considered by Prentice and others (1981) and marginal hazard function of each recurrent event by Wei and others (1989). Andersen and Gill (1982) proposed a multiplicative intensity model, in which recurrent events were considered as non-homogeneous Poisson process. Later on, Pepe and Cai (1993), Lawless and Nadeau (1995), Lin et al. (2000, 2001) studied a mean/rate model with the Poisson assumption removed.

$$E\{dN_i^*(t) \mid X_i(t)\} = \exp\{\beta^T X_i(t)\} \lambda_{0i}(t) dt, \quad (1.1)$$

The aforementioned models assume that the covariate coefficients are constant over time. For many applications, however, the covariate effects change over time instead of staying constant. For example, in clinical study, the treatment effects may change over time, and the temporal effects of treatment may be of interest. Chiang and Wang (2007) modified the above the model to allow the coefficients to be time-dependent.

$$E\{dN_i^*(t) \mid X_i(t)\} = \exp\{\beta(t)^T X_i(t)\} \lambda_{0i}(t) dt, \quad (1.2)$$

Amorim et al. (2008) developed a semiparametric rates model allowing covariate effects to be time-dependent. Regression splines techniques are incorporated in its estimation procedure for estimating the time varying covariate coefficients. More

recently, Sun et al. (2011) proposed a marginal rates model, which allowed some covariate coefficients to be time-varying while others constant.

$$E\{dN_i^*(t) \mid X_i(t), Z_i(t)\} = \exp\{\beta_0(t)^T X_i(t) + \gamma_0^T Z_i(t)\}d\mu_0(t), \quad (1.3)$$

where $\beta_0(t)$ is an unknown p -vector of time-varying regression coefficients, γ_0 is a q -vector of unknown constant regression coefficients, and $\mu_0(t)$ is an unspecified baseline mean function.

In many applications, for instance the motivating example mentioned above, the covariate effects can not only be time dependent but also vary with other covariates. Thus, a model that can consider such covariate-varying coefficients is desired. However, to our best knowledge, existing models and methodologists have not yet considered such covariate-varying effects on recurrent event data. In this dissertation, we propose a semiparametric model which incorporates both time-varying covariate effects and covariate-varying effects motivated by our motivating example.

The remainder of the dissertation is organized as follows. In Chapter 2, we introduce the proposed semiparametric model, estimation procedure for estimating the unknown parameters and functions in the model followed by a detailed computational algorithm, and bandwidth selection for smoothing. Chapter 3 defines some notations and presents the asymptotic properties of the proposed parameter estimates, the proofs for which are given in Appendix. In Chapter 4, the finite-sample performance of the proposed method is evaluated by two types of simulation studies, one of which is survival analysis with single-event data and the other is for recurrent event data. Chapter 5 is a real data application of the proposed method, followed by a discussion

of our findings on the application.

CHAPTER 2: SEMIPARAMETRIC MODEL WITH NONPARAMETRIC COVARIATE-VARYING EFFECTS

In Chapter 2, the contents are organized as follows. We introduce the proposed model in Section 2.1, along with related notations. In Section 2.2, we propose a procedure to estimate the nonparametric parameters and parametric parameters contained in our model. The detailed algorithm procedure for estimation is given in Section 2.3. Since bandwidth selection is involved in our estimation procedure, we use cross validation method to choose the optimal bandwidth, details of which are presented in Section 2.4.

2.1 Model Description

In this section, we start with introducing the notations that are used in our proposed model and methodology.

Let n be the total number of subjects in a random sample, t be the follow up time, and $t \in [0, \tau]$. Suppose the event of interest can be repeated for the same subject i and the recurrent event times can be recorded by T_{ij} , which means the j^{th} event time for the i^{th} subject. Let K_i be the total number of recurrent events for subject i , then we have the event times $0 \leq T_{i1} < T_{i2} < \dots < T_{iK_i} \leq \tau$ for subject i .

During the follow up period, it is common that subjects could drop out at somepoint due to a variety of reasons. For instance, in clinical trial, patients could die due to unrelated causes or lose interest and thus stop following up with the study. Let C_i be

the minimum of end of the study τ and censoring time for subject i . For subject i , only observations at time points before C_i are possible, which means that the observation for subject i is terminated at C_i . Thus, if we define $N_i^*(t) = \sum_{j=1}^{n_i} I(T_{ij} \leq t)$ to be the number of events recorded for subject i by time t , then the observed event process can be written as $N_i(t) = N_i^*(t \wedge C_i)$. Note that $I(\cdot)$ is the indicator function, and $a \wedge b$ means the minimum of a and b .

Suppose $Q_i(t)$ and $U_i(t)$ are covariates that are associated with subject i . $U_i(t)$ is a scalar process and has support \mathcal{U} . $Q_i(t)$ consists three parts $X_i(t)$, $Z_i(t)$ and $W_i(t)$, the dimensions of which are p_1 , p_2 and p_3 , respectively. Each part of $Q_i(t) = (X_i^T(t), Z_i^T(t), W_i^T(t))^T$ is allowed to be either time-dependent or time-independent over the time interval $[0, \tau]$. Note that A^T means the transpose of matrix/vector A .

We assume that the random processes $\{N_i(\cdot), Q_i(\cdot), U_i(\cdot), i = 1, \dots, n\}$ are independent identically distributed (iid). In this dissertation, we only consider the case that the censoring is caused by the termination of the study or random loss of follow up, therefore it is reasonable to assume that the censoring time C_i is noninformative and independent of event time T_{ij} . In other words, we have $E\{dN_i^*(t) | Q_i(t), U_i(t), C_i \geq t\} = E\{dN_i^*(t) | Q_i(t), U_i(t)\}$. However, the censoring time C_i is allowed to depend on covariates $Q_i(t)$ and $U_i(t)$. Some researchers, for instance Ghosh and Lin (2003), have developed methodologies to deal with recurrent events data in the presence of dependent censoring.

Let \mathcal{F}_{t-} denote the filtration, which is the σ -algebra generated by the observed information. Let $\lambda_i(t) = E\{dN_i^*(t) | \mathcal{F}_{t-}\}$ be the intensity or the mean rate $\lambda_i(t) = E\{dN_i^*(t) | X_i(t), U_i(t)\}$. We propose the following semiparametric varying-coefficients

intensity model:

$$\lambda_i(t) = \exp\{\alpha^T(t)X_i(t) + \beta^T Z_i(t) + \gamma^T(U_i(t))W_i(t)\}, \quad (2.1)$$

for $0 \leq t \leq \tau$, where

1. $\alpha(t) = (\alpha_1(t), \alpha_2(t), \dots, \alpha_{p_1}(t))^T$ is a vector with the same dimension p_1 as $X_i(t)$, each element of which is an unspecified function over the time period $0 \leq t \leq \tau$. By setting the first element of $X_i(t)$ to be identity vector I , the corresponding $\alpha_1(t)$ is the nonparametric baseline function.
2. $\beta = (\beta_1, \beta_2, \dots, \beta_{p_2})^T$ is a vector of unknown time-independent parameters with the same dimension p_2 as $Z_i(t)$;
3. $\gamma(U_i(t)) = (\gamma_1(U_i(t)), \gamma_2(U_i(t)), \dots, \gamma_{p_3}(U_i(t)))^T$ is a p_3 -dimensional vector of functions of covariate $U_i(t)$.

In the proposed model, we model the time-varying effects by the non-parametric part $\alpha(t)$, constant effects by the parametric part β , and covariate-varying effects by the non-parametric part $\gamma(u)$. $\gamma(u)$ models the effects of covariate $W_i(t)$ at the level u of another covariate $U_i(t)$. Compared with model (1.3) proposed by Sun et al. (2011), our model adds the ability to model the covariate-varying effects, which is common in reality. For example, in a clinical trial, it is common that patients switch to another treatment due to a variety of reasons. In this case, it is of interest to model the effects of certain covariates after the patient switches treatments. Let $S_i(t)$ denote the time point that patient i switches treatment, then we can define $U_i(t) = t - S_i(t)$ to model the treatment switching effects.

2.2 Estimation Procedure

In this section, we describe the estimating procedure for the parameters in our model. Since the proposed model contains both non-parametric functions $\alpha(t)$ and $\gamma(u)$ and parametric parameters β , we divide our estimating procedure into two parts. First, we estimate the non-parametric part by using Taylor expansion and local linear approximation. Secondly, we estimate the parametric part by using profile likelihood method. Details are as follows.

In order to use Taylor expansion, we first assume that $\alpha(\cdot)$ and $\gamma(\cdot)$ are smooth functions such that their first and second derivatives $\dot{\alpha}(\cdot)$, $\dot{\gamma}(\cdot)$, $\ddot{\alpha}(\cdot)$ and $\ddot{\gamma}(\cdot)$ exist. With this assumption, $\alpha(\cdot)$ and $\gamma(\cdot)$ can be approximated by the first order Taylor expansion as follows.

For any t that belongs to a neighborhood of t_0 , $t \in \mathcal{N}_{t_0}$, we have

$$\alpha(t) = \alpha(t_0) + \dot{\alpha}(t_0)(t - t_0) + O((t - t_0)^2),$$

and similarly for any u that belongs to a neighborhood of u_0 $u \in \mathcal{N}_{u_0}$, we have

$$\gamma(u) = \gamma(u_0) + \dot{\gamma}(u_0)(u - u_0) + O((u - u_0)^2).$$

Therefore, for $t \in \mathcal{N}_{t_0}$ and $U_i(t) \in \mathcal{N}_{u_0}$, the proposed model (2.1) can be approximated by the following

$$\lambda_i^*(t, \vartheta^*, \beta; t_0, u_0) = \exp\{\vartheta^{*T}(t_0, u_0)Q_i^*(t, t_0, u_0) + \beta Z_i(t)\}, \quad (2.2)$$

where $\vartheta^*(t_0, u_0)$ consists a total of four parts, not only the functions themselves but also their first derivatives, as follows.

$$\vartheta^*(t_0, u_0) = (\alpha^T(t_0), \gamma^T(u_0), \dot{\alpha}^T(t_0), \dot{\gamma}^T(u_0))^T$$

And similarly for $Q_i^*(t, t_0, u_0)$, we have

$$Q_i^*(t, t_0, u_0) = (X_i^T(t), W_i^T(t), X_i^T(t) \times (t - t_0), W_i^T(t) \times (U_i(t) - u_0))^T$$

In our estimation procedure, we adopt kernel smoothing method. Suppose

1. $K_1(\cdot)$ and $K_2(\cdot)$ are kernel functions,
2. h_t and h_u are bandwidth parameters,
3. $K_{h_t}(\cdot) = K_1(\cdot/h_t)/h_t$ and $K_{h_u}(\cdot) = K_2(\cdot/h_u)/h_u$.

Then, at each t_0 and u_0 , we define $K_{h_t, h_u}(t, U_i(t); t_0, u_0) = K_{h_t}(t - t_0)K_{h_u}(U_i(t) - u_0)$ as a two dimensional product kernel function.

For fixed β , at each t_0 and u_0 , by Cook and Lawless (2007), the local log-likelihood function for $\alpha(\cdot)$ and $\gamma(\cdot)$ can be written as follows.

$$\ell_{\vartheta}(\vartheta^*; \beta, t_0, u_0) = \sum_{i=1}^n \int_0^{\tau} [\log(\lambda_i^*(t, \vartheta^*, \beta; t_0, u_0)) dN_i(t) - \lambda_i^*(t, \vartheta^*, \beta; t_0, u_0) dt] K_{h_t, h_u}(t, U_i(t); t_0, u_0). \quad (2.3)$$

By taking the first derivative for (2.3) with respect to ϑ^* , at each t_0 and u_0 , the local score function for $\alpha(\cdot)$ and $\gamma(\cdot)$ for fixed β is

$$U_{\vartheta}(\vartheta^*; \beta, t_0, u_0) = \sum_{i=1}^n \int_0^{\tau} [dN_i(t) - \lambda_i^*(t, \vartheta^*, \beta; t_0, u_0) dt] Q_i^*(t, t_0, u_0) K_{h_t, h_u}(t, U_i(t); t_0, u_0) \quad (2.4)$$

By setting $U_{\vartheta}(\vartheta^*; \beta, t_0, u_0) = 0$, we can solve it and denote the solution by $\tilde{\vartheta}^*(\beta, t_0, u_0)$.

We define the following notations:

1. $\tilde{\vartheta}(\beta, t, u)$ be the first $p_1 + p_3$ components of $\tilde{\vartheta}^*(\beta, t_0, u_0)$,
2. $\tilde{Q}_i(t)$ be the first $p_1 + p_3$ components of $Q_i^*(t)$, i.e., $\tilde{Q}_i(t) = ((X_i(t))^T, (W_i(t))^T)^T$,
3. $\tilde{\lambda}_i(t, \beta) = \exp\{(\tilde{\vartheta}(\beta, t, U_i(t)))^T \tilde{Q}_i(t) + \beta^T Z_i(t)\}$.

Then the profile likelihood function can be written as follow:

$$\ell_\beta(\beta) = \sum_{i=1}^n \int_{t_1}^{t_2} \left[\log(\tilde{\lambda}_i(t, \beta)) dN_i(t) - \tilde{\lambda}_i(t, \beta) dt \right]. \quad (2.5)$$

By maximizing the profile likelihood function (2.5), we can obtain the profile maximum likelihood estimator $\hat{\beta}$ of β .

By taking derivative of (2.5) with respect to β , the profile estimating equation for β can be obtained by $U_\beta(\beta) = \frac{\partial \ell_\beta(\beta)}{\partial \beta}$ as follows:

$$U_\beta(\beta) = \sum_{i=1}^n \int_{t_1}^{t_2} \left[dN_i(t) - \tilde{\lambda}_i(t, \beta) dt \right] \left\{ \frac{\partial \tilde{\vartheta}(\beta, t, U_i(t))}{\partial \beta} \tilde{Q}_i(t) + Z_i(t) \right\}, \quad (2.6)$$

where $\frac{\partial \tilde{\vartheta}(\beta, t, U_i(t))}{\partial \beta}$ is the first $p_1 + p_3$ columns of

$$\frac{\partial \tilde{\vartheta}^*(\beta, t, U_i(t))}{\partial \beta} = - \left\{ \frac{\partial U_\vartheta(\vartheta^*; \beta, t, U_i(t))}{\partial \vartheta^*} \right\}^{-1} \frac{\partial U_\vartheta(\vartheta^*; \beta, t, U_i(t))}{\partial \beta} \Bigg|_{\vartheta^* = \tilde{\vartheta}^*(\beta, t, U_i(t))}.$$

By setting $U_\beta(\beta) = 0$, we can solve it and denote the solution by $\hat{\beta}$.

In this dissertation, we adopt the Newton-Raphson iterative method to find the estimators of the nonparametric components $\hat{\vartheta}(t_0, u_0)$ and the parametric components $\hat{\beta}$.

We define the following notations.

1. $\hat{\vartheta}(t_0, u_0) = (\hat{\alpha}(t_0, u_0), \hat{\gamma}(t_0, u_0))^T$

2. $\hat{\alpha}(t_0, u_0)$ as the first p_1 elements of $\hat{\vartheta}^*(t_0, u_0)$ corresponding to the position of $\alpha(t_0)$ in $\vartheta^*(t_0, u_0)$.
3. $\hat{\gamma}(t_0, u_0)$ as the p_3 elements of $\hat{\vartheta}^*(t_0, u_0)$ corresponding to the position of $\gamma(u_0)$ in $\vartheta^*(t_0, u_0)$.

When we estimate $\alpha(t_0)$ by using $\hat{\alpha}^*(t_0, u_0)$, only those local observations with $U_i(t)$ within the neighborhood of u for t within the neighborhood of t_0 are utilized during estimation process. Thus, the estimator $\hat{\alpha}^*(t_0, u_0)$ is not efficient. Similar issue applies to the estimator of $\gamma(u_0)$. In order to improve the efficiency, an aggregation step can be adopted to have more efficient estimators. The following are the proposed estimators $\alpha(t_0)$ and $\gamma(u_0)$, respectively:

$$\hat{\alpha}(t_0) = \frac{1}{n} \sum_{j=1}^n \hat{\alpha}(t_0, U_j(t_0)), \quad \hat{\gamma}(u_0) = \frac{1}{n_{u_0}} \sum_{j=1}^{n_{u_0}} \hat{\gamma}(t_{u_0,j}, u_0), \quad (2.7)$$

where $t_{u_0,j} \in U_j^{-1}(u_0) = \{t : U_j(t) = u_0\}$, and n_{u_0} is the number of points in the union $\cup_{j=1}^n \{U_j^{-1}(u_0)\}$.

2.3 Computational Algorithm

The previous section derives the estimators $\hat{\alpha}(t_0)$ and $\hat{\gamma}(u_0)$ for the non-parametric functions and $\hat{\beta}$ for the parametric parameters in our model. In this section, we sketch the detailed computational algorithm to accomplish those estimators by using Newton-Raphson iterative method.

First, we define some notations that will be used in the algorithm as follows.

- Let $\hat{\vartheta}(t, u)^{\{0\}}$ be the initial values of $\hat{\vartheta}(t, u)$ and $\hat{\beta}^{\{0\}}$ be the initial values for $\hat{\beta}$.

- Let $\hat{\vartheta}^{\{k\}}(t, u)$ be the k^{th} step estimator of $\vartheta(t, u)$ and $\hat{\beta}^{\{k\}}$ be the k^{th} step estimator of β

The steps of computational algorithm are given as follows.

1. Generate equally spaced grid points (t_0, u_0) over t and u .
2. Initialize $\hat{\vartheta}(t, u)^{\{0\}}$ and $\hat{\beta}^{\{0\}}$ by using arbitrary values;
3. For each grid point (t, u) , plug $\hat{\beta}^{\{k-1\}}$, the $(k-1)^{\text{th}}$ step estimator of β , into the local score function (2.4). Find the root and denote it as the k^{th} step estimator $\hat{\vartheta}^{*\{k\}}(t, u) = \hat{\vartheta}^*(t, u, \hat{\beta}^{\{k-1\}})$. The k^{th} step estimator satisfies that $U_{\vartheta}(\hat{\vartheta}^{*\{k\}}(t, u); \hat{\beta}^{\{k-1\}}, t, u) = 0$
4. Obtain the estimates $\hat{\alpha}^{\{k\}}(t_0)$ and $\hat{\gamma}^{\{k\}}(u_0)$ by doing aggregation through (2.7) such that the estimated curves are smooth enough.
5. Plug in $\hat{\alpha}^{\{k\}}(t_0)$ and $\hat{\gamma}^{\{k\}}(u_0)$ to (2.5). The k^{th} step estimator $\hat{\beta}^{\{k\}}$ can be obtained by maximizing the profile likelihood (2.5).
6. Repeat step 3, 4, and 5 and update the estimators $\hat{\vartheta}^{*\{k\}}(t, u)$ and $\hat{\beta}^{\{k\}}$ at each iteration until the convergence criteria is met. The estimator $\hat{\beta}$ is $\hat{\beta}^{\{k\}}$ at the convergence.

2.4 Bandwidth Selection

Kernel smoothing method is adopted in our estimation procedure, which involves selecting the optimal bandwidths. Cross-validation method has been commonly used for bandwidth selection. In this dissertation, we choose the optimal bandwidths for

estimating the non-parametric functions $\alpha(t)$ and $\gamma(u)$ by using the K-fold cross-validation method, where K represents the number of groups that a given sample is to be split into.

In our estimation procedure, we use a two dimensional product kernel function, which involves two bandwidths h_t and h_u parameters. We use K-fold cross-validation method to choose the optimal bandwidths $h_{t,K}^*$ and $h_{u,K}^*$ parameters for h_t and h_u , respectively. Briefly, we go through each combination of (h_t, h_u) to see which one results in the least negative log-likelihood define as follows. The optimal bandwidths combination $(h_{t,K}^*, h_{u,K}^*)$ is the one corresponding to the least negative log-likelihood.

Suppose we have a given sample. The detailed computational procedure for carrying out the K-fold cross-validation are given as follows.

1. Create combinations for h_t and h_u .
2. Shuffle the sample randomly, split it into K groups and denote them as (G_1, G_2, \dots, G_K) .
3. For each combination of $(h_{t,K}, h_{u,K})$, do the following:
 - 3.1. Hold out group G_k ($k = 1, 2, \dots, K$) as a test data set and take the remaining groups as a training data set.
 - 3.2. Fit the proposed model on the training set and evaluate it on the test data set by calculating the negative log-likelihood.
 - 3.3. Retain the evaluation score and discard the model.
 - 3.4. repeat Steps 3.1 - 3.3 until each group gets a turn to be the test data set.

- 3.5. Take the simple average of all K evaluation scores from the loop, and denote it as the Score for current testing bandwidth combination.
4. Repeat Step 3 until each bandwidth combination is used to fit the model.
5. Compare all scores corresponding to different bandwidth combinations and denote the one corresponding to the lowest score as the optimal bandwidth combination $(h_{t,K}^*, h_{u,K}^*)$

CHAPTER 3: ASYMPTOTIC PROPERTIES

In Chapter 2, we propose estimators $\hat{\alpha}$, $\hat{\gamma}$ and β for the non-parametric functions $\alpha(t)$ and $\gamma(u)$ and the parameter β , respectively. It is naturally of interest to explore the asymptotic properties for those estimators. Therefore, in Chapter 3, we will establish the asymptotic properties for those proposed estimators, including asymptotic normality, consistency, etc. Chapter 3 are organized as follows. In Section 3.1, we define all related notations. Section 2 presents all theorems that we establish regarding the asymptotic properties of our estimators.

3.1 Notations

We define the notations as follows.

- Let $\mathcal{I}_1 = \{\mathcal{I}_{ij}\}_{p_1 \times (p_1+p_3)}$ be a matrix with elements like the following.

$$\mathcal{I}_{ij} = \begin{cases} 1 & \text{for } i = 1, \dots, p_1, i = j \\ 0 & \text{otherwise} \end{cases}$$

- Let $\mathcal{I}_3 = \{\mathcal{I}_{ij}\}_{p_3 \times (p_1+p_3)}$ be a matrix with elements like the following.

$$\mathcal{I}_{ij} = \begin{cases} 1 & \text{for } i = 1, \dots, p_3, j = i + p_1 \\ 0 & \text{otherwise} \end{cases}$$

- Let $\alpha_0(t)$, β_0 and $\gamma_0(u)$ be the true values of $\alpha(t)$, β and $\gamma(u)$ under model (2.1),

respectively.

- Let $\lambda_{i0}(t) = \exp\{\alpha_0(t)X_i(t) + \beta_0 Z_i(t) + \gamma_0(U_i(t))W_i(t)\}$
- Let $\hat{\lambda}_i(t) = \exp\{\hat{\vartheta}^T(t, U_i(t))\tilde{Q}_i(t) + \beta^T Z_i(t)\}$
- Let $f_U(t, u)$ be the density function of $U(t)$ evaluated at u

- Define

$$e_{11}(t, u) = E \left[(-\lambda_{i0}(t)dt) \{\tilde{Q}_i(t)\}^{\otimes 2} \mid U_i(t) = u \right] f_U(t, u)$$

and

$$e_{12}(t, u) = E \left[(-\lambda_{i0}(t)dt) Z_i(t) \{\tilde{Q}_i(t)\}^{\otimes 2} \mid U_i(t) = u \right] f_U(t, u)$$

- Define

$$\hat{E}_{11}(t_0, u_0) = \frac{1}{n} \sum_{i=1}^n \int_0^\tau K_h(t - t_0) K_b(U_i(t) - u_0) \left[-\hat{\lambda}_i(t)dt \right] \{\tilde{Q}_i(t)\}^{\otimes 2},$$

and

$$\hat{E}_{12}(t_0, u_0) = \frac{1}{n} \sum_{i=1}^n \int_0^\tau K_h(t - t_0) K_b(U_i(t) - u_0) \left[-\hat{\lambda}_i(t)dt \right] \tilde{Q}_i(t) (Z_i(t))^T.$$

3.2 Asymptotic Properties

In this section, we establish the asymptotic properties for the proposed estimators $\hat{\alpha}(t)$, $\hat{\gamma}(t)$ and $\hat{\beta}$ in this section. Three theorems are given as follows.

Theorem 3.1. *Assuming the conditions given in Appendix are satisfied, then*

$$\sqrt{n}(\hat{\beta} - \beta_0) \longrightarrow N(0, A_\beta^{-1} \Sigma_\beta A_\beta^{-1})$$

in distribution, where

$$A_\beta = E \left[\int_0^\tau \{Z_i(t) - (e_{12}(t, U_i(t)))^T (e_{11}(t, U_i(t)))^{-1} \tilde{Q}_i(t)\}^{\otimes 2} dt \right]$$

and

$$\Sigma_\beta = E \left[\int_0^\tau \{Z_i(t) - (e_{12}(t, U_i(t)))^T (e_{11}(t, U_i(t)))^{-1} \tilde{Q}_i(t)\} dM_i(t) \right]^{\otimes 2}.$$

The matrices A_β and Σ_β can be consistently estimated by the following two formulas, respectively.

$$\hat{A}_\beta = \frac{1}{n} \sum_{i=1}^n \int_{t_1}^{t_2} \{Z_i(t) - (\hat{E}_{12}(t, U_i(t)))^T (\hat{E}_{11}(t, U_i(t)))^{-1} \tilde{Q}_i(t)\}^{\otimes 2} dt$$

and

$$\hat{\Sigma}_\beta = \frac{1}{n} \sum_{i=1}^n \left(\int_{t_1}^{t_2} \{dN_i(t) - \hat{\lambda}_i(t) dt\} \{Z_i(t) - (\hat{E}_{12}(t, U_i(t)))^T (\hat{E}_{11}(t, U_i(t)))^{-1} \tilde{Q}_i(t)\} \right)^{\otimes 2},$$

where $0 < t_1 < t_2 < \tau$.

Theorem 3.2. *Assuming the conditions given in Appendix are satisfied, then*

- (1) $\sup_{t \in [0, \tau]} |\hat{\alpha}(t) - \alpha_0(t)| = o_p(1)$;
- (2) $\sqrt{nh_t}(\hat{\alpha}(t) - \alpha_0(t) - \frac{1}{2}h_t^2\nu_2\ddot{\alpha}(t)) \xrightarrow{\mathcal{D}} N(0, \Sigma_\alpha(t))$,

where

$$\Sigma_\alpha(t) = \lim_{n \rightarrow \infty} h_t E \left[\int_0^\tau \{dN_i(s) - \lambda_i(s) ds\} \mathcal{A}_1 e_{11}(t, U_i(s))^{-1} \tilde{Q}_i(s) K_{h_t}(s - t) \right]^{\otimes 2}.$$

The limiting variance-covariance matrix $\Sigma_\alpha(t)$ can be consistently estimated by the

following formula.

$$\hat{\Sigma}_\alpha(t) = \frac{h_t}{n} \sum_{i=1}^n \left[\int_0^\tau \left\{ dN_i(s) - \hat{\lambda}_i(s) ds \right\} \mathcal{I}_1 \hat{E}_{11}(t, U_i(s))^{-1} \tilde{Q}_i(s) K_{h_t}(s-t) \right]^{\otimes 2}.$$

Theorem 3.3. *Assuming the conditions given in Appendix are satisfied, then*

- (1) $\sup_{u \in [u_1, u_2]} |\hat{\gamma}(u) - \gamma_0(u)| = o_p(1);$
- (2) $\sqrt{nh_u}(\hat{\gamma}(u) - \gamma_0(u) - \frac{1}{2}h_u^2 \nu_2 \ddot{\gamma}(u)) \xrightarrow{\mathcal{D}} N(0, \Sigma_\gamma(u)).$

where

$$\Sigma_\gamma(u) = \lim_{n \rightarrow \infty} h_u E \left[\int_0^\tau \left\{ dN_i(s) - \lambda_i(s) ds \right\} \mathcal{I}_3 e_{11}(u, U_i(s))^{-1} \tilde{Q}_i(s) K_{h_u}(s-u) \right]^{\otimes 2}.$$

The limiting variance-covariance matrix $\Sigma_\gamma(u)$ can be consistently estimated by

$$\hat{\Sigma}_\gamma(u) = \frac{h_u}{n} \sum_{i=1}^n \left[\int_0^\tau \left\{ dN_i(s) - \hat{\lambda}_i(s) ds \right\} \mathcal{I}_3 \hat{E}_{11}(u, U_i(s))^{-1} \tilde{Q}_i(s) K_{h_u}(s-u) \right]^{\otimes 2}.$$

CHAPTER 4: SIMULATION STUDIES

In order to assess the finite sample performance of the proposed model and estimation procedure, we conduct a few simulation studies in this chapter. First, we apply the methodologies on a survival analysis, which is a single-event case. Second, the methodologies are applied to recurrent event case. Details for simulations for both cases will be presented in the following subsections.

4.1 Simulation on Single Event Data - Survival Analysis

In this section, we illustrate our method on single event data and check the finite sample performance of the proposed method. This section first introduces the method that is used to generate single event counting process data, followed by a simulation example with a specific model by using the proposed method.

4.1.1 Generating Single Event Data

Some researchers had studied on how to generate single event data for survival analysis. For instance, both Bender et al. (2005) and Qian et al. (2010) proposed a method to simulate survival data with Cox model proposed by Cox (1972).

In the Cox model, the intensity function is written as follows.

$$\lambda(t|X) = \lambda_0(t) \exp\{\beta^T X\}, \quad (4.1)$$

where $\lambda_0(t)$ is the baseline intensity function, β is the parameter vector and X is

a time independent covariate vector with the same dimension as β . The cumulative intensity function $\Lambda(t|X)$ can be written as the integral of $\lambda(t|X)$ from 0 to t as follows.

$$\Lambda(t|X) = \int_0^t \lambda(s|X) ds, \quad (4.2)$$

The survival function has the following relationship with cumulative intensity function.

$$S(t|X) = \exp\{-\Lambda(t|X)\}, \quad (4.3)$$

Let $F(t|X) = 1 - S(t|X)$, then we have

$$F(t|X) = 1 - \exp\{-\Lambda(t|X)\}, \quad (4.4)$$

Supposed Y is a random variable, which follows uniform distribution on $[0, 1]$. By setting $F(t|X) = Y$, we can derive the corresponding T by solving the inverse function as below. T is the desired time to event.

$$T = F^{-1}(Y) = \Lambda^{-1}(-\ln(1 - Y)), \quad (4.5)$$

where $F^{-1}()$ and $\Lambda^{-1}()$ are the inverse functions of $F()$ and cumulative intensity function $\Lambda()$. In the real world, observations can be censored. For example, patients do not experience death by the end of the study period and thus are censored. In order to take into account censoring, we set the study period to be τ . If the T generated by the equation above is greater than τ , the corresponding subject is marked as censored.

With a given intensity model, a survival dataset with single event can be generated by following the steps.

1. Set the true value for parameter vector β , true function for $\lambda_0(t)$, study period

- τ , and the total number of observations n (aka. sample size).
2. Simulate the time independent covariates vector X .
 3. Plug in all information from step 1 and 2 to (4.1) and calculate (4.1) through (4.4).
 4. Simulate a number y from the distribution $Uniform[0, 1]$.
 5. Plug y into (4.5), find T .
 6. Compare T with τ , if T is less than τ , set the censoring indicator to be 0, otherwise set it to be 1.
 7. Repeat step 2 through step 6 for n times to simulate data for n subjects.

4.1.2 Simulation Example

By following the procedure presented in the previous subsection, a single-event sample is generated to be used for evaluating the finite sample performance of our model. In this example, we use the following model to illustrate our method.

$$\lambda(t) = \exp\{\alpha_0(t) + \alpha_1(t)X + \beta Z + \gamma(t - S)W\}, \quad (4.6)$$

for $0 \leq t \leq \tau$ with $\tau = 2$, with the following settings.

- $\alpha_0(t) = -1.5 + 0.8t$, $\alpha_1(t) = t$, $\beta = -0.5$, $\gamma(U_i(t)) = -0.5u$;
- X_i is generated from truncated normal distribution $N(-0.5, 0.5, 0, 1)$, Z_i is an uniform random variable on $[-0.5, 0.5]$ and W_i is generated from the distribution $Binary(0.5)$;

- $U_i(t) = t - S_i$, where S_i is generated from the uniform distribution $U[0, 0.5]$;
- Censoring time C_i is generated from the uniform distribution $U[1, 3]$.

In the generated sample, about 67 % of subjects are censored. Approximately 33 % of subjects experience an event during the study period $[0, 2]$.

During estimation, cross-validation method described in Chapter 2 is applied for preliminary bandwidth selection. In this example, three sets of bandwidth combination for h_t and h_u are selected to reflect different levels of smoothness, including

- $h_t = h_u = 0.30$
- $h_t = h_u = 0.40$
- $h_t = h_u = 0.50$

Boundaries effect is taking into consideration in all simulation examples in this study, thus we set $t_1 = h_t$ and $t_2 = \tau - h_t$ in the estimating functions in chapter 2. For all simulations examples in this study, we use the Epanechnikov kernel for smoothing, which is given as follows.

$$K(x) = .75(1 - x^2)I(|x| \leq 1)$$

For all simulation examples in this dissertation, we consider some criterias to evaluate the performance of the proposed estimators, which are presented as follows.

In order to assess the performance of the estimator $\hat{\beta}$, we measure the following items in all simulation examples.

- Bias = estimate - true value

- the sample standard error of the estimates (SSE)
- the sample mean of the estimated standard errors (ESE)
- the 95% empirical coverage probability (CP)

In order to assess the performance of the estimators $\hat{\alpha}_0(t)$ for baseline, $\hat{\alpha}_1(t)$ for time varying covariate effects and $\hat{\gamma}(u)$ for covariate-varying covariate effect, we measure their pointwise Bias, SSE, ESE, and CP at different fixed time points t and u , respectively.

Besides Bias, SSE, ESE and CP, to better assess the overall performance for those estimators for unknown functions, we calculate the square root of integrated mean square error (RMSE) for each of them respectively.

Suppose

- N is the total number of repetitions;
- $\alpha_{00}(t)$ and $\alpha_{10}(t)$ are the true function values of $\alpha_0(t)$ and $\alpha_1(t)$ at each time point $t \in [0, \tau]$, respectively;
- $\gamma_{10}(u)$ and $\gamma_{20}(u)$ are the true function values of $\gamma_1(u)$ and $\gamma_2(u)$ at each point $u \in [0, \tau]$, respectively;

then, the RMSEs for each of estimators in this chapter are defined as follows.

$$RMSE_{\alpha_0} = \left\{ \frac{1}{N(\tau - 2h_t)} \sum_{j=1}^N \int_{h_t}^{\tau-h_t} (\hat{\alpha}_{0j}(t) - \alpha_{00}(t))^2 dt \right\}^{1/2},$$

$$RMSE_{\alpha_1} = \left\{ \frac{1}{N(\tau - 2h_t)} \sum_{j=1}^N \int_h^{\tau-h} (\hat{\alpha}_{1j}(t) - \alpha_{10}(t))^2 dt \right\}^{1/2},$$

$$RMSE_{\gamma_1} = \left\{ \frac{1}{N(\tau - 2h_u)} \sum_{j=1}^N \int_{h_u}^{\tau-h_u} (\hat{\gamma}_{1j}(u) - \gamma_{10}(u))^2 du \right\}^{1/2},$$

$$RMSE_{\gamma_2} = \left\{ \frac{1}{N(\tau - 2h_u)} \sum_{j=1}^N \int_{h_u}^{\tau-h_u} (\hat{\gamma}_{2j}(u) - \gamma_{20}(u))^2 du \right\}^{1/2},$$

where $\hat{\alpha}_{0j}(t)$, $\hat{\alpha}_{1j}(t)$, $\hat{\gamma}_{1j}(u)$ and $\hat{\gamma}_{2j}(u)$ are the j^{th} estimate of $\alpha_0(t)$, $\alpha_1(t)$, $\gamma_1(u)$ and $\gamma_2(u)$, respectively, for $j = 1, \dots, N$.

Sample sizes $n = 400, 600, 800, 1000$ are considered in this example. All results presented below are calculated based on 500 simulation repetitions. Table 5 summarizes the Bias, SSE, ESE and CP for the fixed covariate coefficient estimator $\hat{\beta}$ under model (4.6). The results show the following.

1. The bias for the estimator $\hat{\beta}$ are small among different sample sizes and bandwidth combinations, which indicates that the estimates are unbiased.
2. Both empirical and estimated standard errors presented on Table 5 are reasonably close to each other, and thus the coverage probabilities are close to the nominal level 95%.

Note that for a particular sample size, when the bandwidth gets larger, Bias decreases while both empirical and estimated standard errors increases. Fan and Gijbels (1996) find that bigger bandwidth results in larger variance but smaller bias. Our results are consistent with their finding. However, the results show that coverage probabilities are not sensitive to bandwidth selection. To conclude, Table 5 indicates that the proposed estimator $\hat{\beta}$ for fixed covariate effect performs well under model (4.6).

Figure 1 and Figure 2 summarize the results produced with the bandwidth $h_t = 0.40$ and $h_u = 0.40$. The results are averaged on 500 simulation repetitions. The red curve corresponds to the results for sample size $n = 400$, while blue curve for $n = 600$, green curve for $n = 800$ and black curve for $n = 1000$, respectively.

The left panel of Figure 1 presents the Bias, SSEs, ESEs and CPs at different fixed time points for $\hat{\alpha}_0(t)$, while the right panel for $\hat{\alpha}_1(t)$. The results show that the bias for both estimators are small. The coverage probability fluctuates around the nominal level 95%. In addition, the plots show that when sample size is 400 or 600, the estimated standard errors are not stable for the latter half of study period. However, when sample size increases ($n = 800, 1000$), SSE and ESE agree to each other very well. This could be due to data sparsity when sample size is not considerably large. The estimators perform better when sample size increases.

Figure 1 show that the pointwise bias for $\hat{\gamma}(u)$ are very small, which indicates that the pointwise estimates are unbiased. An agreement is observed between pointwise SSE and ESE, thus the coverage probability curves slightly fluctuate around the line of 97%, which shows reasonable performance of the estimator $\hat{\gamma}(u)$.

Table 6 summarizes the RMSEs based on 500 simulation repetitions for $\hat{\alpha}_0(t)$, $\hat{\alpha}_1(t)$ and $\hat{\gamma}(u)$ under model 4.6. The results show that the RMSEs for all those four estimators decrease when sample size increases. The same trend is observed for all three selected bandwidth combinations.

Table 5: Summary of Bias, SSE, ESE and CP for $\hat{\beta}$ under model (4.6).

n	h_t	h_u	Bias	SSE	ESE	CP
400	0.30	0.30	-0.0333	0.3481	0.3511	0.966
	0.40	0.40	-0.0241	0.3630	0.3855	0.966
	0.50	0.50	-0.0097	0.4109	0.4385	0.972
600	0.30	0.30	0.0055	0.4871	0.2822	0.956
	0.40	0.40	-0.0047	0.2939	0.3066	0.958
	0.50	0.50	-0.0014	0.3315	0.3517	0.968
800	0.30	0.30	-0.0159	0.2283	0.2347	0.964
	0.40	0.40	-0.0128	0.2488	0.2622	0.964
	0.50	0.50	-0.0103	0.2902	0.3020	0.964
1000	0.30	0.30	-0.0168	0.1989	0.2079	0.968
	0.40	0.40	-0.0136	0.2180	0.2330	0.966
	0.50	0.50	-0.0146	0.2493	0.2687	0.970

Table 6: Summary of RMSEs for $\hat{\alpha}_0(t)$, $\hat{\alpha}_1(t)$ and $\hat{\gamma}(u)$ under model (4.6).

n	h_t	h_u	RMSE $_{\alpha_0}$	RMSE $_{\alpha_1}$	RMSE $_{\gamma}$
400	0.30	0.30	0.2701	0.6550	1.0922
	0.40	0.40	0.2134	0.5475	0.3863
	0.50	0.50	0.1849	0.4778	0.3384
600	0.30	0.30	0.2244	0.5154	0.6548
	0.40	0.40	0.1648	0.4418	0.2806
	0.50	0.50	0.1457	0.3874	0.2384
800	0.30	0.30	0.1659	0.4422	1.1642
	0.40	0.40	0.1404	0.3819	0.2240
	0.50	0.50	0.1239	0.3374	0.1934
1000	0.30	0.30	0.1451	0.3996	0.5083
	0.40	0.40	0.1243	0.3488	0.2057
	0.50	0.50	0.1111	0.3114	0.1800

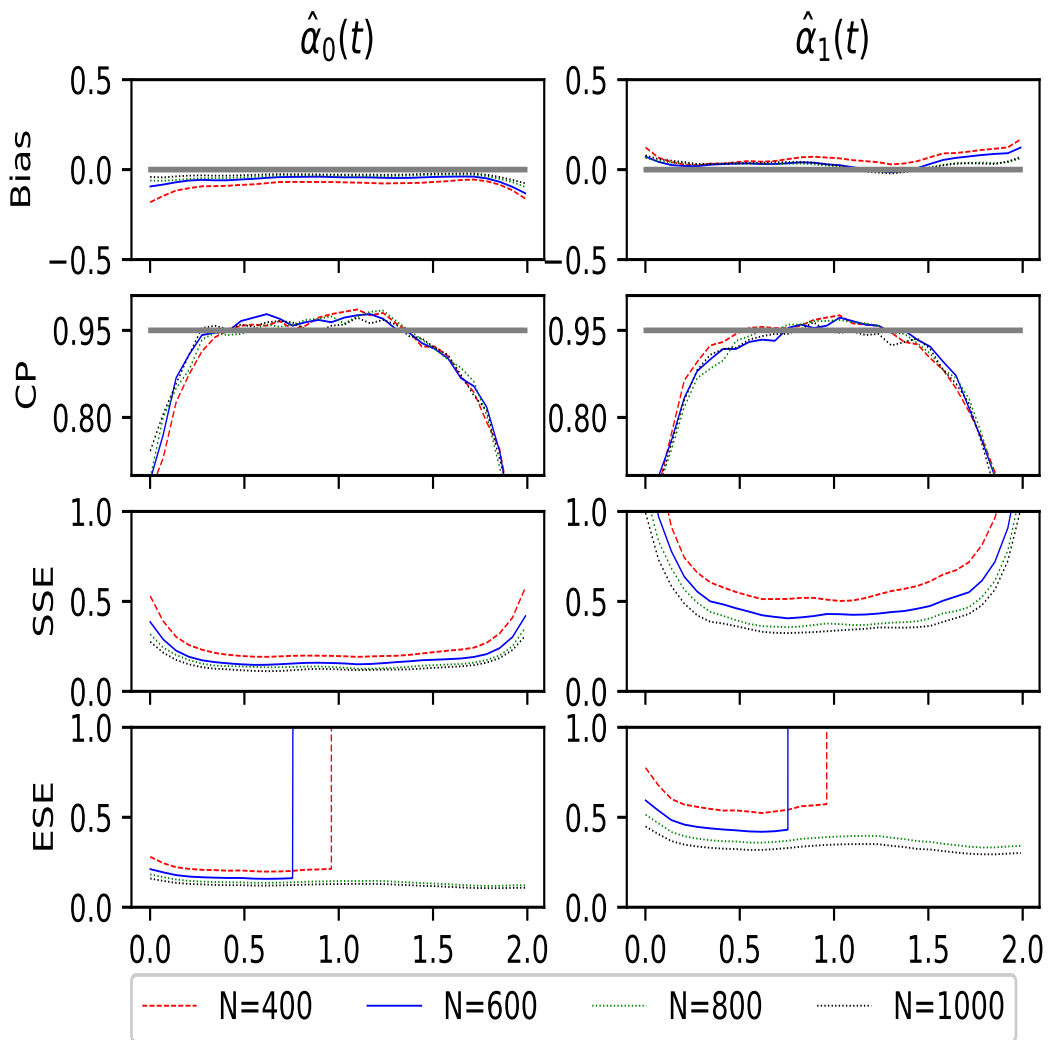


Figure 1: Plots for Bias, CP, SSE and ESE for $n=400, 600, 800,$ and 1000 with $h_t = 0.4, h_u = 0.4$ for $\alpha_0(t) = -1.5 + 0.8t$ and $\alpha_1(t) = t$ for $0 \leq t \leq 2$.

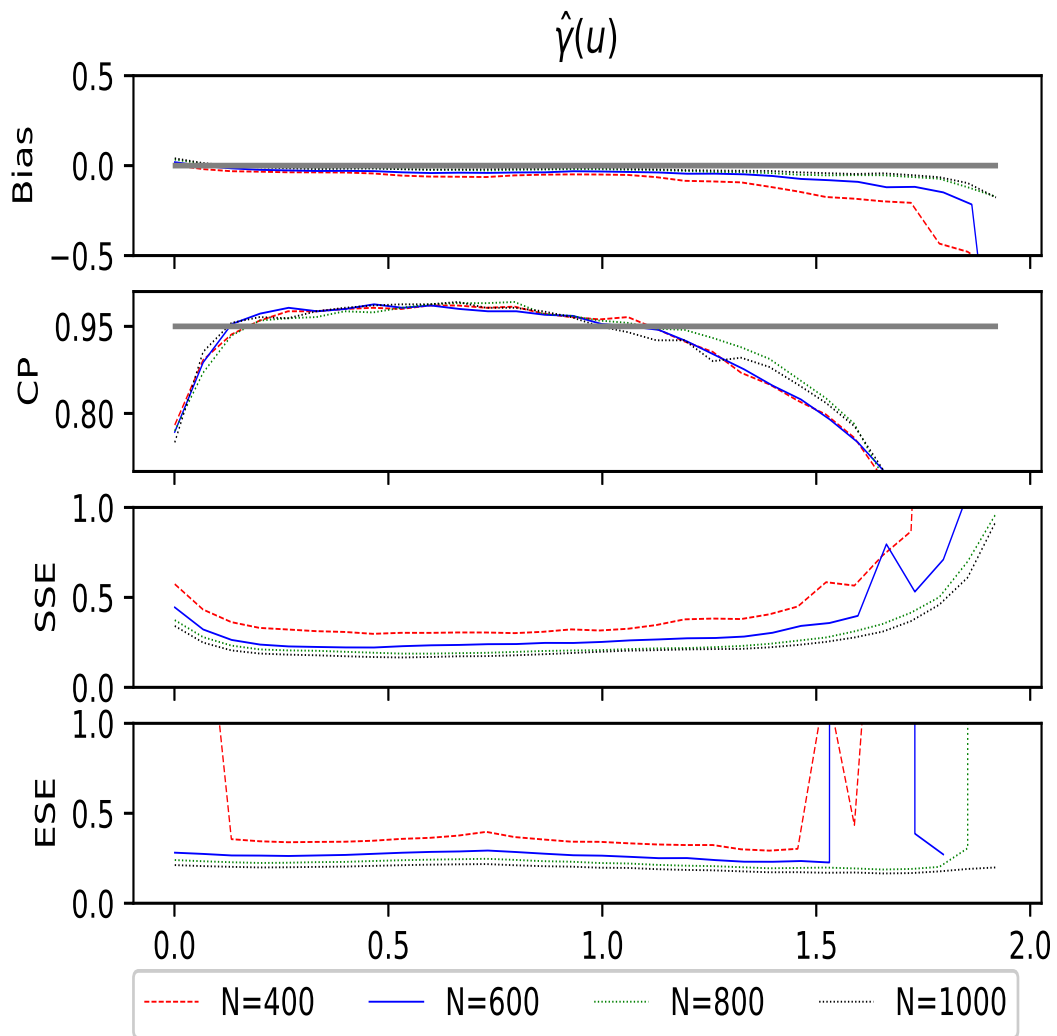


Figure 2: Plots for bias, CP, SSE and ESE for $n=400, 600, 800,$ and 1000 with $h_t = 0.4, h_u = 0.4$ for $\gamma(u) = -0.5u$.

4.2 Simulation on Recurrent Events Data

The previous section show that the proposed model and estimation procedure work well under the survival analysis framework. In this section, the goal is to check the finite sample performance of the proposed methodology when the event of interest can be repeated for each subject. This section starts with a description on how to generate recurrent events data by thinning method, followed by two simulation examples under different models.

4.2.1 Generating Recurrent Events Data

There are many different existing methods for generating a non-Homogeneous Poisson Process (NHPP). For instance, a NHPP can be generated through time-scale transformation of a Homogeneous Poisson Process (HPP) or it can be generated by using order statistics. Lewis and Shedler (1976) generated a NHPP with log linear rate function. Another approach for generating a NHPP is the thinning method proposed by Lewis and Shedler (1979). Compared with other methods, the thinning method has many advantages. For instance, the thinning method does not require numerical integration of the rate function, ordering of points, or generating Poisson variates. Thus, in this dissertation, we adopt the thinning method for simulating recurrent events data.

Given the intensity function $\lambda(t)$ for $0 \leq t \leq \tau$, choose a constant $\bar{\lambda}$ s.t. $\lambda(t) \leq \bar{\lambda}$ for all t . The detailed procedure for generating recurrent events by using thinning method are given as follows. By repeating the procedure below, we can generate the recurrent event times for each subject i in a sample. The recurrent event times are

recorded by T_{ij} ($i = 1, 2, \dots, n$ and $j = 1, 2, \dots, K_i$)

1. Set $T_0 = 0$, $T^* = 0$, and $j = 1$.
2. Generate an random variable V from exponential distribution $\exp(1/\bar{\lambda})$
3. Update $T^* = T^* + V$.
4. If $T^* > \tau$, stop; otherwise generate a random variable R from uniform distribution $U(0, 1)$.
5. Compare R with $\lambda(T^*)/\bar{\lambda}$. If $R \leq \lambda(T^*)/\bar{\lambda}$, then accept the arrival time, store it by $T_{ij} = T^*$, and $j = j + 1$; otherwise reject the arrival time and return to Step 2.

In the following subsections, this procedure is used to generate recurrent event samples to conduct simulation studies.

4.2.2 Simulation Example 1

We start with a simple model to illustrate the proposed method and evaluate the finite sample performance of the proposed estimators. The model given as follows considers an intercept term as the baseline, time-independent effect and covariate-varying effect.

$$\lambda_i(t) = \exp\{\alpha_0(t) + \beta Z_i + \gamma(U_i(t))W_i\}, \quad (4.7)$$

for $0 \leq t \leq \tau$ with $\tau = 5$, with the following settings.

- $\alpha_0(t) = 1.5 - \log(1 + t)$, $\beta = 1.5$, $\gamma(U_i(t)) = \sqrt{U_i(t)} - 2$;

- Z_i is an uniform random variable on $[-0.5, 0.5]$ and W_i is generated from truncated normal distribution $N(-0.5, 0.5, 0, 1)$.
- $U_i(t) = t - S_i$, where S_i is generated from the uniform distribution $U[0, 0.5]$.
- Censoring time C_i for the i^{th} subject is generated from the uniform distribution $U[4, 9]$.

In the generated sample, about 20% of subjects are censored. Approximately a total of 12 recurrent events are observed per subject during the study period $[0, 5]$.

Three sample sizes ($n = 400, 600, 800$) are considered in this example. All results presented below are calculated based on 500 simulation repetitions. Cross-validation method is applied for bandwidth selection. The following three sets of bandwidth combinations for h_t and h_u are selected to reflect different levels of smoothness.

- $h_t = h_u = 0.25$
- $h_t = h_u = 0.30$
- $h_t = h_u = 0.35$

Table 7 summarizes the Bias, SSE, ESE, and CP for the fixed covariate coefficient estimator $\hat{\beta}$ under model (4.7) based on 500 simulation repetitions. The results show the following.

1. The bias for the estimator $\hat{\beta}$ are small among different sample sizes and bandwidth combinations, which indicates that the estimates are unbiased.

2. Both empirical and estimated standard errors presented on Table 7 agree to each other, which results coverage probabilities that are close to the nominal level 95%. In addition, coverage probabilities are found to be not sensitive to bandwidth selection.
3. Bias, empirical standard errors, and estimated standard errors all decrease when sample size increases.
4. For a particular sample size, when the bandwidth gets larger, Bias decreases while both empirical and estimated standard errors increases, which is again consistent with the finding by Fan and Gijbels (1996).

To conclude, Table 7 indicates that the proposed estimator $\hat{\beta}$ for fixed covariate effect performs well under model (4.7).

Figure 3 plots the pointwise bias, empirical standard errors, estimated standard errors and coverage probabilities for the estimate of the baseline function $\alpha_0(t)$, while Figure 4 shows the plots for $\hat{\gamma}(u)$. All plots are generated based on the results from 500 simulation repetitions. The bandwidth combination used to generate these plots is $h_t=0.3$, $h_u=0.3$. The red curve represents result for sample size $n = 400$, while blue for $n = 600$ and green for $n = 800$. The plots reveal the following findings.

1. The pointwise bias for $\hat{\alpha}_0(t)$ is reasonably small. When sample size increases, bias decreases. Bias tends to become relatively larger at the end of study period, which could be the boundary effects.
2. The pointwise bias for $\hat{\gamma}(u)$ is reasonably small and the curve fluctuates around

the zero line.

3. For both $\hat{\alpha}_0(t)$ and $\hat{\gamma}(u)$, the pointwise empirical standard error and estimated standard errors are very close to each other, and consequently the curves for pointwise coverage probabilities fluctuate around the nominal level 95%.

Besides pointwise bias, SSE, ESE and CP, RMSEs are also calculated to assess the overall performance of both $\hat{\alpha}_0(t)$ and $\hat{\gamma}(u)$. Table 8 shows the results under model (4.7) with different sample sizes and bandwidth combinations. It appears that RMSEs for both $\hat{\alpha}_0(t)$ and $\hat{\gamma}(u)$ drops when sample size increases, regardless of bandwidth combination.

To conclude, based on the results given by Figure 3, 4 and Table 8, the proposed estimator $\hat{\alpha}_0(t)$ and $\hat{\gamma}(u)$ show satisfied performance under model (4.7).

Table 7: Summary of Bias, SSE, ESE and CP for $\hat{\beta}$ under model (4.7).

n	h_t	h_u	Bias	SSE	ESE	CP
400	0.25	0.25	0.0173	0.0724	0.0697	0.934
	0.30	0.30	-0.0016	0.0736	0.0702	0.938
	0.35	0.35	0.0088	0.0754	0.0722	0.938
600	0.25	0.25	0.0112	0.0580	0.0567	0.944
	0.30	0.30	-0.0051	0.0579	0.0572	0.950
	0.35	0.35	0.0063	0.0596	0.0589	0.954
800	0.25	0.25	0.0070	0.0498	0.0490	0.936
	0.30	0.30	-0.0089	0.0509	0.0495	0.938
	0.35	0.35	0.0031	0.0529	0.0510	0.940

Table 8: Summary of RMSEs for $\hat{\alpha}_0(t)$ and $\hat{\gamma}(u)$ under model (4.7).

n	h_t	h_u	RMSE $_{\alpha_0}$	RMSE $_{\gamma}$
400	0.25	0.25	0.0862	0.2981
	0.30	0.30	0.0781	0.2770
	0.35	0.35	0.0728	0.2643
600	0.25	0.25	0.0683	0.2395
	0.30	0.30	0.0625	0.2230
	0.35	0.35	0.0584	0.2126
800	0.25	0.25	0.0584	0.2089
	0.30	0.30	0.0536	0.1934
	0.35	0.35	0.0502	0.1846

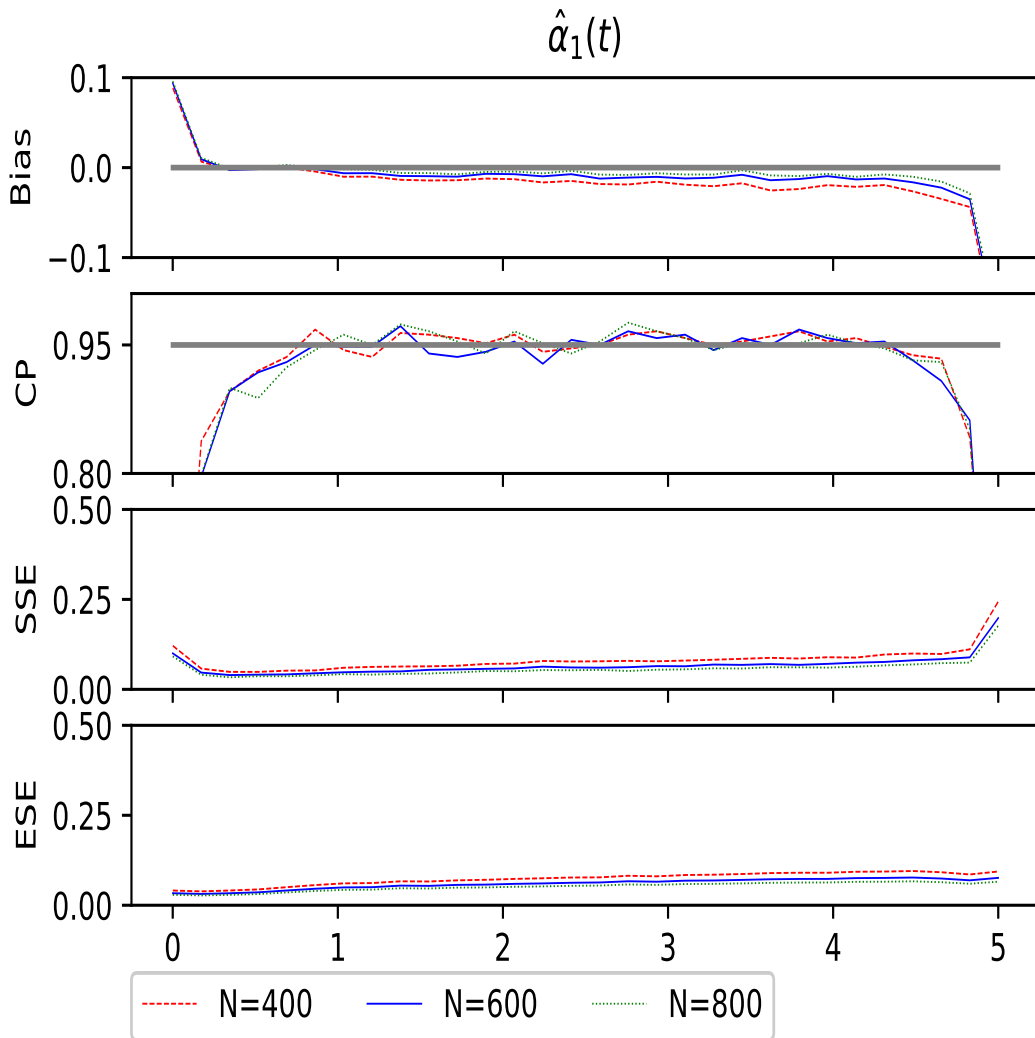


Figure 3: Plots for bias, CP, SSE and ESE for $n=400, 600, 800$ with $h_t=0.3, h_u=0.3$ for $\alpha_0(t) = 1.5 - \log(1 + t)$ for $0 \leq t \leq 5$.

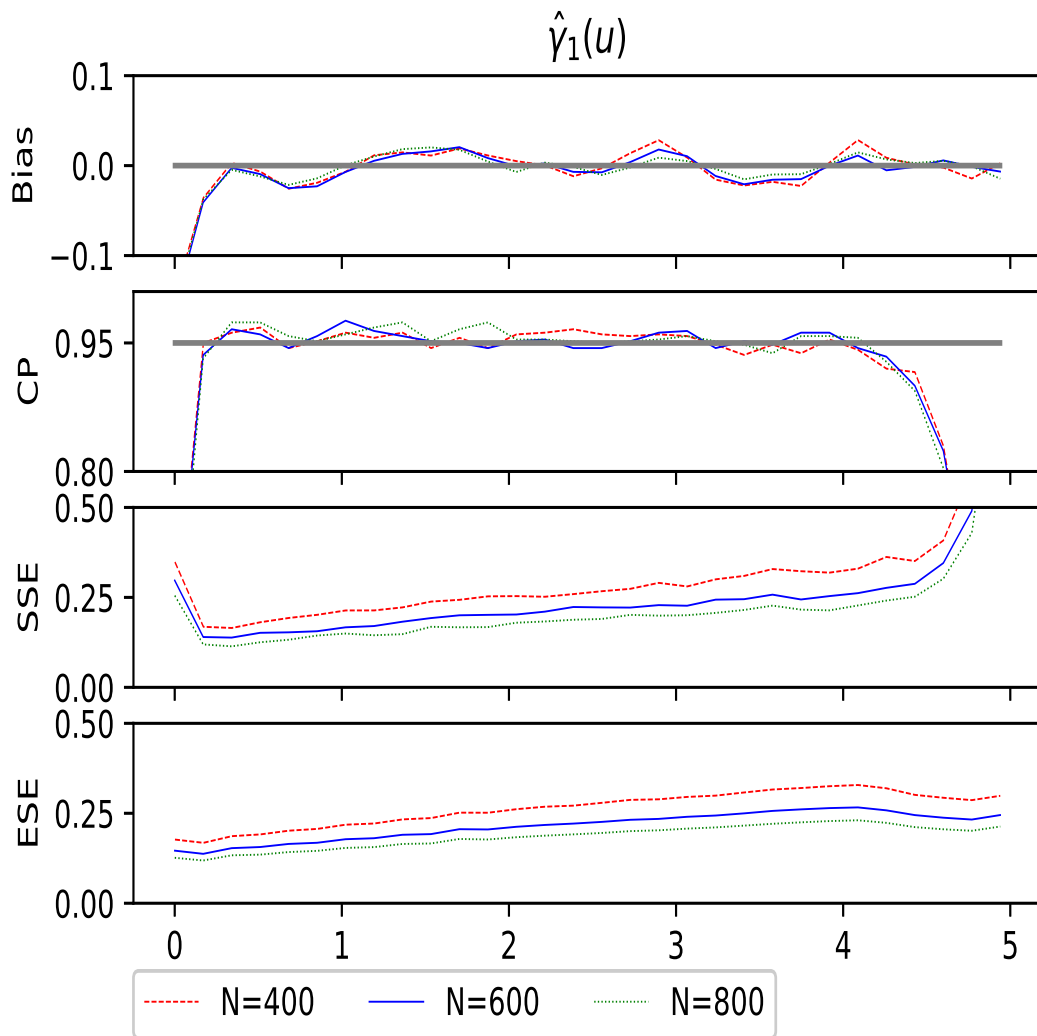


Figure 4: Plots for bias, CP, SSE and ESE for $n=400, 600, 800$ with $h_t=0.3, h_u=0.3$ for $\gamma(u) = \sqrt{u} - 2$.

4.2.3 Simulation Example 2

The previous simulation example considers fixed covariate effects and covariate-varying effects. The results show that the proposed method performs well. In example 2, a more complicated model is considered to illustrate the proposed method, which incorporates fixed covariate effects, time-varying effects and covariate-varying effects. The model is given as follows.

$$\lambda_i(t) = \exp\{\alpha_1(t) + \alpha_2(t)X_i(t) + \beta Z_i + \gamma_1(t)W_{1i} + \gamma_2(U_i(t))W_{2i}\}, \quad (4.8)$$

for $0 \leq t \leq \tau$ and $\tau = 5$, with the following settings.

- $\alpha_1(t) = 2 - \log(1 + t)$, $\alpha_2(t) = \sin(0.2t)$;
- $\beta = 1.5$;
- $\gamma_1(U_i(t)) = U_i(t) - 1$, $\gamma_2(U_i(t)) = \sqrt{U_i(t)} - 2$;
- Covariate $X_i(t)$ is time-dependent; it is generated from truncated normal distribution $N(-0.5, 0.5, 0, 1)$;
- Z_i is generated from uniform distribution $U[-0.5, 0.5]$;
- W_{1i} and W_{2i} are generated from truncated bivariate normal with marginal $N(-0.5, 0.5, 0, 1)$ and correlation $\rho = 0.2$;
- Censoring time C_i is generated from an uniform distribution $U[4, 9]$;
- $U_i(t) = t - S_i$, where S_i is generated from an uniform distribution $U[0, 0.5]$.

With the above settings, about 20% of subjects are censored. Approximately a total of 12 recurrent events are observed per subject during the study period $[0, 5]$.

Similar to previous examples, we set $t_1 = h_t$ and $t_2 = \tau - h_t$ in the estimating functions to deal with boundary effects. Cross-validation method is applied for bandwidth selection, and the following three sets of bandwidth combination for h_t and h_u are selected to reflect different levels of smoothness.

- $h_t = h_u = 0.25$
- $h_t = h_u = 0.30$
- $h_t = h_u = 0.35$

Three sample sizes ($n = 400, 600, 800$) are considered in this study. All results presented below are calculated based on 500 simulation repetitions.

Table 9 summarizes the Bias, SSE, ESE, and CP for the fixed covariate coefficient estimator $\hat{\beta}$ under model (4.8), averaging over 500 simulation repetitions. The results show that the bias for the estimator $\hat{\beta}$ are small among different sample sizes and bandwidth combinations, which indicates that the estimates are unbiased. Both empirical and estimated standard errors presented on Table 9 agree to each other, and thus the coverage probabilities are close to 95%. In addition, Bias, empirical standard errors, and estimated standard errors all decrease when sample size increases. However, the results show that coverage probabilities are not sensitive to bandwidth selection. To conclude, Table 9 indicates that the proposed estimator $\hat{\beta}$ for fixed covariate effect performs well under model (4.8).

To assess the performance of the estimators for time varying covariate effects $\hat{\alpha}_1(t)$ and $\hat{\alpha}_2(t)$ and those for covariate-varying covariate effects $\hat{\gamma}_1(u)$ and $\hat{\gamma}_2(u)$, we calculate their Bias, SSE, ESE, and CP at different fixed time points t and u , respectively. Figure 5 and Figure 6 summarize the results produced with the bandwidth $h_t = 0.30$ and $h_u = 0.30$. The results are averaged on 500 simulation repetitions. The left panel of Figure 5 presents the Bias, SSEs, ESEs and CPs at different fixed time points for $\hat{\alpha}_1(t)$, while the right panel for $\hat{\alpha}_2(t)$. The left panel of Figure 6 shows the Bias, SSEs, ESEs and CPs at different fixed time points for $\hat{\gamma}_1(t)$, while the right panel for $\hat{\gamma}_2(t)$.

The plots from Figure 5 and Figure 6 show the following findings.

1. The pointwise bias for all four sets of estimates $\hat{\alpha}_1(t)$, $\hat{\alpha}_2(t)$, $\hat{\gamma}_1(u)$ and $\hat{\gamma}_2(u)$ are very small, thus the pointwise estimates are unbiased. In addition, bias decreases along with increasing sample size.
2. For each of all four estimators, an agreement is observed between pointwise empirical standard error and estimated standard error. The coverage probability curves slightly fluctuate around the line of nominal level 95%.

To assess the overall performance of $\hat{\alpha}_1(t)$, $\hat{\alpha}_2(t)$, $\hat{\gamma}_1(u)$ and $\hat{\gamma}_2(u)$, the RMSE is calculated for each of them. Table 10 summarizes the RMSEs based on 500 simulation repetitions for $\hat{\alpha}_1(t)$, $\hat{\alpha}_2(t)$, $\hat{\gamma}_1(u)$ and $\hat{\gamma}_2(u)$ under model 4.8. The results show that the RMSEs for all those four estimators decrease when sample size increases. The same trend is observed for all three selected bandwidths.

To conclude, the proposed estimators $\hat{\alpha}_1(t)$, $\hat{\alpha}_2(t)$, $\hat{\gamma}_1(u)$ and $\hat{\gamma}_2(u)$ for unknown non-parametric functions shows strong performance under model 4.8.

In this chapter, in order to illustrate our proposed method and assess the finite sample performance, two simulation examples are conducted on survival analysis framework and another two simulation examples on recurrent events data framework. All examples show reasonably small bias for our estimates and an agreement between empirical standard errors and estimated standard errors, which result in coverage probabilities that are close to the nominal level 95%. Therefore, the proposed method performs very well with finite samples.

Table 9: Summary of Bias, SSE, ESE and CP for $\hat{\beta}$ under model (4.8).

n	h_t	h_u	Bias	SSE	ESE	CP
400	0.25	0.25	0.0164	0.0532	0.0521	0.930
	0.30	0.30	-0.0018	0.0539	0.0525	0.936
	0.35	0.35	0.0082	0.0559	0.0540	0.936
600	0.25	0.25	0.0099	0.0436	0.0422	0.934
	0.30	0.30	-0.0067	0.0444	0.0426	0.932
	0.35	0.35	0.0046	0.0459	0.0439	0.940
800	0.25	0.25	0.0072	0.0371	0.0364	0.942
	0.30	0.30	-0.0082	0.0380	0.0368	0.938
	0.35	0.35	0.0035	0.0396	0.0379	0.938

Table 10: Summary of RMSEs for $\hat{\alpha}_1(t)$, $\hat{\alpha}_2(t)$, $\hat{\gamma}_1(u)$ and $\hat{\gamma}_2(u)$ under model (4.8).

n	h_t	h_u	RMSE $_{\alpha_1}$	RMSE $_{\alpha_2}$	RMSE $_{\gamma_1}$	RMSE $_{\gamma_2}$
400	0.25	0.25	0.0722	0.2051	0.2190	0.2091
	0.30	0.30	0.0643	0.1891	0.2016	0.1948
	0.35	0.35	0.0597	0.1786	0.1917	0.1855
600	0.25	0.25	0.0567	0.1672	0.1787	0.1685
	0.30	0.30	0.0511	0.1549	0.1654	0.1566
	0.35	0.35	0.0476	0.1463	0.1574	0.1491
800	0.25	0.25	0.0482	0.1451	0.1528	0.1451
	0.30	0.30	0.0437	0.1349	0.1417	0.1353
	0.35	0.35	0.0407	0.1278	0.1349	0.1288

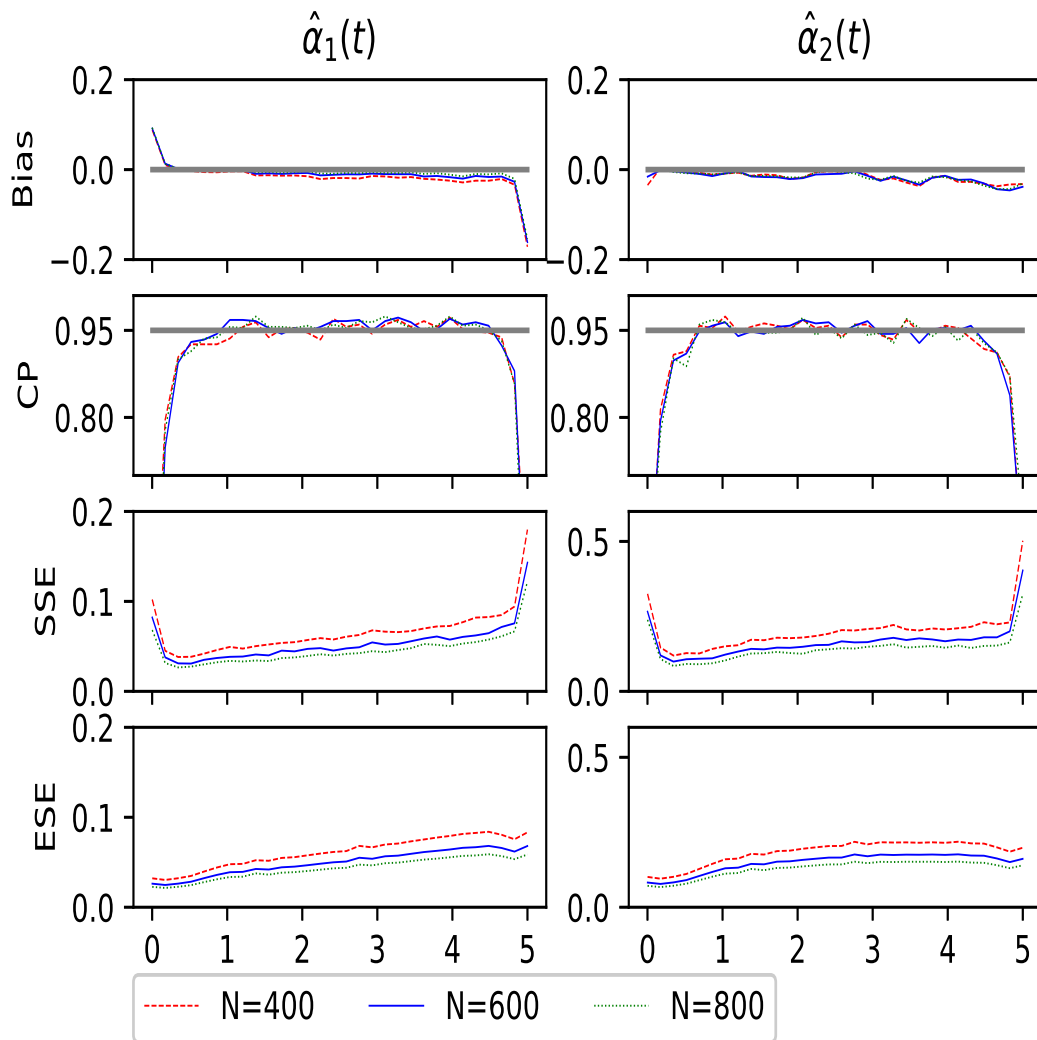


Figure 5: Plots for bias and CP for $n=400, 600, 800$ with $h_t = 0.3, h_u = 0.3$. Left panel is for $\alpha_1(t) = 2 - \log(1 + t)$. Right panel is for $\alpha_2(t) = \sin(0.2t)$.

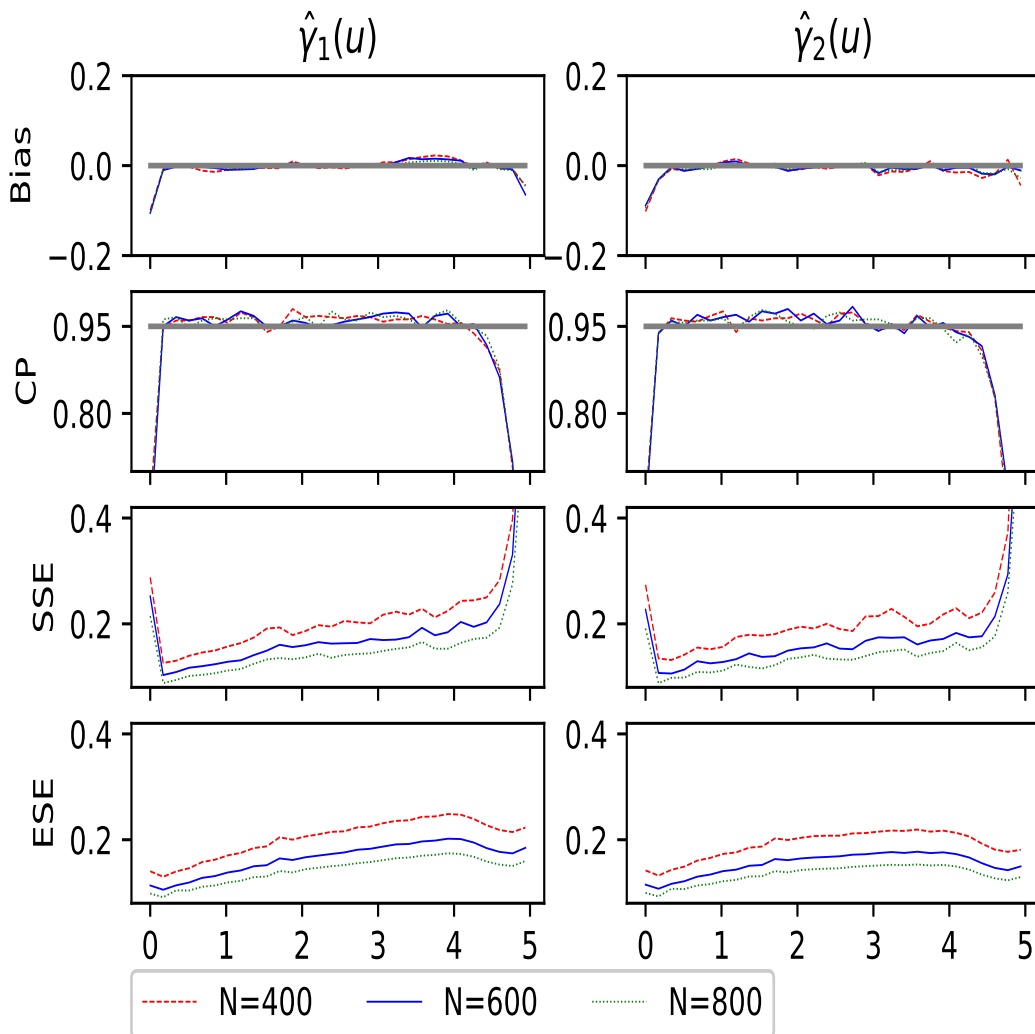


Figure 6: Plots for bias and CP for $n=400, 600, 800$ with $h_t = 0.3, h_u = 0.3$. Left panel is for $\gamma_1(u) = u - 1$. Right panel is for $\gamma_2(u) = \sqrt{u} - 2$.

REFERENCES

- L. D. Amorim, J. Cai, D. Zeng, and M. L. Barreto. Regression splines in the time-dependent coefficient rates model for recurrent event data. *Statistics in Medicine*, 27:5890–5906, 2008.
- P. K. Andersen and R. D. Gill. Coxs regression model counting process: A large sample study. *Annals of Statistics*, 10:1100–1120, 1982.
- B.C. Astor, S. Yi, L. Hiremath, T. Corbin, V. Pogue, B. Wilkening, G. Peterson, J. Lewis, J.P. Lash, F. Van Lente, J. Gassman, X. Wang, G. Bakris, L.J. Appel, and G. Contreras. N-terminal prohormone brain natriuretic peptide as a predictor of cardiovascular disease and mortality in blacks with hypertensive kidney disease: the african american study of kidney disease and hypertension (aask). *Circulation*, 117(13):1685–1692, 2008.
- S. Beddhu, X. Ma, B. Baird, A.K. Cheung, and T. Greene. Serum alkaline phosphatase and mortality in african americans with chronic kidney disease. *Clin J Am Soc Nephrol*, 4(11):1805–1810, 2009.
- R. Bender, T. Augustin, and M. Blettner. Generating survival times to simulate cox proportional hazards models. *Statistics in Medicine*, 24(11):1713–1723, 2005.
- D. Brooks, J. Charleston, D. Dowie, A. Gabriel, Y.B. Hall, L. Hiremath, T. Lightfoot, M. Sika, W.C. Smith, and X. Wang. Predictors of participant adherence and retention in the african american study of kidney disease and hypertension. *Nephrol Nurs Journal*, 35(2):133–142, 2008.
- T.K. Chen, M.J. Choi, W.H. Kao, B.C. Astor, J.J. Scialla, L.J. Appel, L. Li, M.S. Lipkowitz, M. Wolf, R.S. Parekh, C.A. Winkler, M.M. Estrella, and D.C. Crews. Examination of potential modifiers of the association of apoll alleles with ckd progression. *Clin J Am Soc Nephrol*, 10(12):2128–2135, 2015.
- G. Contreras, T. Greene, L.Y. Agodoa, D. Cheek, G. Junco, D. Dowie, J. Lash, M. Lipkowitz, E.R.3rd Miller, A. Ojo, M. Sika, B. Wilkening, and R.D. Toto. Blood pressure control, drug therapy, and kidney disease. *Hypertension*, 46(1):44–50, 2005.
- R.J. Cook and J. Lawless. *The Statistical Analysis of Recurrent Events*. Springer-Verlag New York, 2007.
- D. R. Cox. Regression models and life tables. *Journal of the Royal Statistical Society, Series B*, 20:187–220, 1972.
- J. Fan and I. Gijbels. *Local polynomial modelling and its applications: Monographs on statistics and applied probability*. CRC Press, 1996.

- J.J. Gassman, T. Greene, J.T.Jr Wright, L. Agodoa, G.J. Bakris, G. Beck, J. Douglas, K. Jamerson, J. Lewis, M. Kutner, O.S. Randall, and S.R. Wang. Design and statistical aspects of the african american study of kidney disease and hypertension (aask). *J Am Soc Nephrol*, 14(7):154–165, 2003.
- D. Ghosh and D.Y. Lin. Semiparametric analysis of recurrent events data in the presence of dependent censoring. *Biometrics*, 59:877–885, 2003.
- K. Hong, P. Muntner, I. Kronish, D. Shilane, and T.I. Chang. Medication adherence and visit-to-visit variability of systolic blood pressure in african americans with chronic kidney disease in the aask trial. *J Hum Hypertens*, 30(1):73–78, 2015.
- E. Ku, M.S. Lipkowitz, L.J. Appel, A. Parsa, J. Gassman, D.V. Glidden, M. Smogorzewski, and C.Y. Hsu. Strict blood pressure control associates with decreased mortality risk by apoll genotype. *Kidney Int*, 91(2):443–450, 2016.
- E. Ku, J.D. Kopple, K.L. Johansen, C.E. McCulloch, A.S. Go, D. Xie, F. Lin, L.L. Hamm, J. He, J.W. Kusek, S.D. Navaneethan, A.C. Ricardo, H. Rincon-Choles, M. Smogorzewski, C.Y. Hsu, and CRIC Study Investigators. Longitudinal weight change during ckd progression and its association with subsequent mortality. *Am J Kidney Dis*, 71(5):657–665, 2018.
- J. F. Lawless and C. Nadeau. Some simple robust methods for the analysis of recurrent events. *Technometrics*, 37:158–168, 1995.
- J. Lea, T. Greene, L. Hebert, M. Lipkowitz, S. Massry, J. Middleton, S.G. Rostand, E. Miller, W. Smith, and G.L. Bakris. The relationship between magnitude of proteinuria reduction and risk of end-stage renal disease: results of the african american study of kidney disease and hypertension. *Arch Intern Med*, 165(8):947–953, 2005.
- J. Lea, D. Cheek, D. Thornley-Brown, L. Appel, L. Agodoa, G. Contreras, J. Gassman, J. Lash, Randall O. Miller, E. R. 3rd, X. Wang, and W. McClellan. Metabolic syndrome, proteinuria, and the risk of progressive ckd in hypertensive african americans. *Am J Kidney Dis*, 51(5):732–740, 2008.
- J.P. Lea, D.T. Brown, M. Lipkowitz, J. Middleton, K. Norris, and AASK Study Group. Preventing renal dysfunction in patients with hypertension: clinical implications from the early aask trial results. *Am J Cardiovasc Drugs*, 3(3):193–200, 2003.
- J. Lewis, T. Greene, L. Appel, G. Contreras, J. Douglas, J. Lash, R. Toto, F. Van Lente, X. Wang, J.T.Jr Wright, and AASK Study Group. A comparison of iothalamate-gfr and serum creatinine-based outcomes: acceleration in the rate of gfr decline in the african american study of kidney disease and hypertension. *J Am Soc Nephrol*, 15(12):3175–3183, 2004.

- P.A.W Lewis and G.S. Shedler. Simulation of nonhomogeneous poisson processes with log linear rate function. *Biometrika*, 63(3):501–505, 1976.
- P.A.W Lewis and G.S. Shedler. Simulation of nonhomogeneous poisson processes by thinning. *Naval Research Logistics Quarterly*, 26(3):403–413, 1979.
- D. Y. Lin, L. J. Wei, I. Yang, and Z. Ying. Semiparametric regression for the mean and rate function of recurrent events. *Journal of the Royal Statistical Society: Series B*, 62:711–730, 2000.
- D. Y. Lin, L. J. Wei, and Z. Ying. Semiparametric transformation models for point processes. *Journal of the American Statistical Association*, 96:620–628, 2001.
- K.C. Norris, T. Greene, J. Kopple, J. Lea, J. Lewis, M. Lipkowitz, A. Miller, P. Richardson, S. Rostand, X. Wang, and L.J. Appel. Baseline predictors of renal disease progression in the african american study of hypertension and kidney disease. *J Am Soc Nephrol*, 17(10):2928–2936, 2006.
- M. S. Pepe and J. Cai. Some graphical displays and marginal regression analyses for recurrent failure times and time dependent covariates. *Journal of the American Statistical Association*, 88:881–820, 1993.
- J. Qian, B. Li, and P. Chen. Generating survival data in the simulation studies of cox model. *2010 Third International Conference on Information and Computing*, 2010.
- K.L. Raphael, G. Wei, B.C. Baird, T. Greene, and S. Beddhu. Higher serum bicarbonate levels within the normal range are associated with better survival and renal outcomes in african americans. *Kidney Int*, 79(3):356–362, 2010.
- M. Sika, J. Lewis, J. Douglas, T. Erlinger, D. Dowie, M. Lipkowitz, J. Lash, D. Cornish-Zirker, G. Peterson, R. Toto, J. Kusek, L. Appel, C. Kendrick, J. Gassman, and AASK Study Group. Baseline characteristics of participants in the african american study of kidney disease and hypertension (aask) clinical trial and cohort study. *Am J Kidney Dis*, 50(1):78–89, 2007.
- L. Sun, X. Zhou, and S. Guo. Marginal regression models with time-varying coefficients for recurrent event data. *Statistics in Medicine*, 30:2265–2277, 2011.
- A.W. Van Der Vaart. *Asymptotic Statistics*. Cambridge University Press, 1998.
- X. Wang, J. Lewis, L. Appel, D. Cheek, G. Contreras, M. Faulkner, H. Feldman, J. Gassman, J. Lea, J. Kopple, M. Sika, R. Toto, and T. Greene. Validation of creatinine-based estimates of gfr when evaluating risk factors in longitudinal studies of kidney disease. *J Am Soc Nephrol*, 17(10):2900–2909, 2006.
- J.T.Jr Wright, G. Bakris, T. Greene, L.Y. Agodoa, L.J. Appel, J. Charleston, D. Cheek, J.G. Douglas-Baltimore, J. Gassman, R. Glassock, L. Hebert, K. Jamerson, J. Lewis, R.A. Phillips, R.D. Toto, J.P. Middleton, and S.G. Rostand. Effect

of blood pressure lowering and antihypertensive drug class on progression of hypertensive kidney disease: results from the aask trial. *JAMA*, 288(19):2421–2431, 2002.

G. Yin, H. Li, and D. Zeng. Partially linear additive hazards regression with varying coefficients. *Journal of the American Statistical Association*, 103(483), 2008.

APPENDIX A: PROOFS OF THE THEOREMS

Conditions

The following conditions are needed for our derivation of asymptotic properties in this study.

- The censoring time C_i is noninformative in the sense that $E\{dN_i^*(t)|Q_i(t), U_i(t), C_i \geq t\} = E\{dN_i^*(t)|Q_i(t), U_i(t)\}$, while the censoring time C_i is allowed to depend on the left continuous covariate process $Q_i(\cdot)$;
- The processes $Q_i(t)$ and $\lambda_i(t)$, $0 \leq t \leq \tau$, are bounded and their total variations are bounded by a constant; $E|N_i(t_2) - N_i(t_1)|^2 \leq L(t_2 - t_1)$ for $0 \leq t_1 \leq t_2 \leq \tau$, where $L > 0$ is a constant; $E|N_i(t+h) - N_i(t-h)|^{2+v} = O(h)$, for some $v > 0$;
- The kernel function $K(\cdot)$ is symmetric with compact support on $[-1, 1]$ and Lipschitz continuous; Bandwidths $h_t \asymp h_u$; $h_t \rightarrow 0$; $nh_t^2 \rightarrow \infty$ and nh_t^5 is bounded;
- $\alpha_0(t)$, $\gamma_0(u)$, $e_{11}(t)$ and $e_{12}(t)$ are twice differentiable; $(e_{11}(t))^{-1}$ is bounded over $0 \leq t \leq \tau$; the matrices A_β and Σ_β are positive definite;
- The following two limits exist and are finite.

$$\lim_{n \rightarrow \infty} h_t E \left[\int_0^\tau \{dN_i(s) - \lambda_i(s)ds\} \mathcal{I}_1 e_{11}(t, U_i(s))^{-1} \tilde{Q}_i(s) K_{h_t}(s-t) \right]^{\otimes 2},$$

and

$$\lim_{n \rightarrow \infty} h_u E \left[\int_0^\tau \frac{\dot{\lambda}_i(s)}{\lambda_i(s)} \{dN_i(s) - \lambda_i(s)ds\} \mathcal{I}_3 e_{11}(u, U_i(s))^{-1} \tilde{X}_i(s) K_{h_u}(s-u) \right]^{\otimes 2}.$$

Proof of Theorems

Proof of Theorem 3.1

By Lemma 1 and Lemma 3 in (add citation) and application of the Glivenko-Cantelli theorem to the estimating function defined in (2.6), we have

$$\begin{aligned}
& \frac{1}{n} U_\beta(\beta) \\
&= \frac{1}{n} \sum_{i=1}^n \int_0^\tau [dN_i(t) - \tilde{\lambda}_i(t, \beta) dt] \left\{ \frac{\partial \tilde{\vartheta}(\beta, t, U_i(t))}{\partial \beta} \tilde{Q}_i(t) + Z_i(t) \right\} \\
&\xrightarrow{\mathcal{P}} E \int_0^\tau [dN_i(t) - \lambda \{ \vartheta_\beta^T(t, U_i(t)) \tilde{Q}_i(t) + \beta^T Z_i(t) \} dt] \\
&\quad \times \{ Z_i(t) - (e_{\beta,12}(t, U_i(t)))^T (e_{\beta,11}(t, U_i(t)))^{-1} \tilde{Q}_i(t) \} \\
&= E \int_0^\tau [\varphi \{ \vartheta_0^T(t, U_i(t)) \tilde{Q}_i(t) + \beta_0^T Z_i(t) \} dt - \varphi \{ \vartheta_\beta^T(t, U_i(t)) \tilde{Q}_i(t) + \beta^T Z_i(t) \} dt] \\
&\quad \times \{ Z_i(t) - (e_{\beta,12}(t, U_i(t)))^T (e_{\beta,11}(t, U_i(t)))^{-1} \tilde{Q}_i(t) \} \\
&= u(\beta), \tag{A.1}
\end{aligned}$$

where β_0 is the unique root of $u(\beta)$. Then by Theorem 5.9 of Van Der Vaart (1998),

$$\hat{\beta} \xrightarrow{\mathcal{P}} \beta_0.$$

By Glivenko-Cantelli theorem and Lemma 3 of (add citation),

$$\begin{aligned}
& -\frac{1}{n} \frac{\partial U_\beta(\beta)}{\partial \beta} \Big|_{\beta=\beta_0} \\
&= \frac{1}{n} \sum_{i=1}^n \int_0^\tau [dN_i(t) - \lambda \{\tilde{\vartheta}^T(t, U_i, \beta_0) \tilde{Q}_i(t) + \beta_0^T Z_i(t)\} dt] \\
&\quad \times \left\{ \frac{\partial \tilde{\vartheta}(t, U_i(t), \beta_0)}{\partial \beta} \tilde{Q}_i(t) + Z_i(t) \right\} \\
&+ \frac{1}{n} \sum_{i=1}^n \int_0^\tau [-\dot{\lambda} \{\tilde{\vartheta}^T(t, U_i, \beta_0) \tilde{Q}_i(t) + \beta_0^T Z_i(t)\}] \left\{ \frac{\partial \tilde{\vartheta}(t, U_i(t), \beta_0)}{\partial \beta} \tilde{Q}_i(t) + Z_i(t) \right\}^{\otimes 2} dt \\
&+ \frac{1}{n} \sum_{i=1}^n \int_0^\tau [dN_i(t) - \lambda \{\tilde{\vartheta}^T(t, U_i(t), \beta_0) \tilde{Q}_i(t) + \beta_0^T Z_i(t)\} dt] \left\{ \frac{\partial^2 \tilde{\vartheta}(t, U_i(t), \beta_0)}{\partial^2 \beta} \tilde{Q}_i(t) \right\}
\end{aligned} \tag{A.2}$$

The first and third terms go to zero as $n \rightarrow \infty$ by Lemma (add citation) and the Glivenko-Cantelli theorem. Thus, it follows that

$$\begin{aligned}
& -\frac{1}{n} \frac{\partial U_\beta(\beta)}{\partial \beta} \Big|_{\beta=\beta_0} \xrightarrow{\mathcal{P}} E \int_0^\tau \dot{\lambda} \{\tilde{\vartheta}^T(t, U_i, \beta_0) \tilde{Q}_i(t) + \beta_0^T Z_i(t)\} \\
&\quad \times \{Z_i(t) - (e_{12}(t, U_i(t)))^T (e_{11}(t, U_i(t)))^{-1} \tilde{Q}_i(t)\}^{\otimes 2} dt \equiv A_\beta
\end{aligned} \tag{A.3}$$

Now we show that $n^{-1/2} U_\beta(\beta_0)$ converges in distribution to a normal distribution. By

Taylor expansion,

$$\begin{aligned}
& \exp\{\tilde{\vartheta}^T(t, U_i(t), \beta_0) \tilde{Q}_i(t) + \beta_0^T Z_i(t)\} - \exp\{\vartheta_0^T(t, U_i(t)) \tilde{Q}_i(t) + \beta_0^T Z_i(t)\} \\
&= \exp\{\vartheta_0^T(t, U_i(t)) \tilde{Q}_i(t) + \beta_0^T Z_i(t)\} [\tilde{\vartheta}^T(t, U_i(t), \beta_0) - \vartheta_0^T(t, U_i(t))] \tilde{Q}_i(t) \\
&+ O_p(\|\tilde{\vartheta}(t, U_i(t), \beta_0) - \vartheta_0(t, U_i(t))\|^2)
\end{aligned} \tag{A.4}$$

We define $dM_i(t) = dN_i(t) - \exp\{\vartheta_0^T(t, U_i(t))\tilde{Q}_i(t) + \beta_0^T Z_i(t)\}dt$, then we have

$$\begin{aligned}
\frac{1}{\sqrt{n}}U_\beta(\beta_0) &= \frac{1}{\sqrt{n}}\sum_{i=1}^n\int_0^\tau[dN_i(t) - \exp\{\tilde{\vartheta}^T(t, U_i(t), \beta_0)\tilde{Q}_i(t) + \beta_0^T Z_i(t)\}] \\
&\quad \times \left\{ \frac{\partial\tilde{\vartheta}(t, U_i(t), \beta_0)}{\partial\beta}\tilde{Q}_i(t) + Z_i(t) \right\} \\
&= \frac{1}{\sqrt{n}}\sum_{i=1}^n\int_0^\tau \left\{ \frac{\partial\tilde{\vartheta}(t, U_i(t), \beta_0)}{\partial\beta}\tilde{Q}_i(t) + Z_i(t) \right\} dM_i(t) \\
&\quad - \frac{1}{\sqrt{n}}\sum_{i=1}^n\int_0^\tau [\exp\{\tilde{\vartheta}^T(t, U_i(t), \beta_0)\tilde{Q}_i(t) + \beta_0^T Z_i(t)\} - \varphi\{\vartheta_0^T(t, U_i(t))\tilde{Q}_i(t) + \beta_0^T Z_i(t)\}] \\
&\quad \times \left\{ \frac{\partial\tilde{\vartheta}(t, U_i(t), \beta_0)}{\partial\beta}\tilde{Q}_i(t) + Z_i(t) \right\} dt
\end{aligned} \tag{A.5}$$

By Lemma 1 in Lin et al. (2001), the second term equals to the following

$$\begin{aligned}
&\frac{1}{\sqrt{n}}\sum_{i=1}^n\int_0^\tau \exp\{\tilde{\vartheta}^T(t, U_i, \beta_0)\tilde{Q}_i(t) + \beta_0^T Z_i(t)\}[\tilde{\vartheta}^T(t, U_i(t), \beta_0) - \vartheta_0^T(t, U_i(t))]\tilde{Q}_i(t)dt \\
&\quad \times \left\{ \frac{\partial\tilde{\vartheta}(t, U_i(t), \beta_0)}{\partial\beta}\tilde{Q}_i(t) + Z_i(t) \right\} = o_p(1)
\end{aligned} \tag{A.6}$$

Thus, we have

$$\begin{aligned}
&\frac{1}{\sqrt{n}}U_\beta(\beta_0) \\
&= \frac{1}{\sqrt{n}}\sum_{i=1}^n\int_0^\tau \left\{ \frac{\partial\tilde{\vartheta}(t, U_i(t), \beta_0)}{\partial\beta}\tilde{Q}_i(t) + Z_i(t) \right\} dM_i(t) + o_p(1) \\
&= \frac{1}{\sqrt{n}}\sum_{i=1}^n\int_0^\tau \{Z_i(t) - (e_{12}(t, U_i(t)))^T(e_{11}(t, U_i(t)))^{-1}\tilde{Q}_i(t)\}dM_i(t)
\end{aligned} \tag{A.7}$$

which converges in distribution to $N(0, \Sigma_\beta)$ by Central Limit Theorem, where

$$\Sigma_\beta = E \left(\int_0^\tau \{Z_i(t) - (e_{12}(t, U_i(t)))^T (e_{11}(t, U_i(t)))^{-1} \tilde{Q}_i(t)\} dM_i(t) \right)^{\otimes 2}.$$

By Taylor expansion, we have

$$U_\beta(\hat{\beta}) = U_\beta(\beta_0) + \frac{\partial U_\beta(\beta)}{\partial \beta} \Big|_{\beta=\beta_0} (\hat{\beta} - \beta_0) + O_p(\|\hat{\beta} - \beta_0\|^2).$$

Thus,

$$\sqrt{n}(\hat{\beta} - \beta_0) = \left(-\frac{1}{n} \frac{\partial U_\beta(\beta)}{\partial \beta} \Big|_{\beta=\beta_0} \right)^{-1} \times \frac{1}{\sqrt{n}} U_\beta(\beta_0)$$

Hence, by Slutsky Theorem, we have $\sqrt{n}(\hat{\beta} - \beta_0) \xrightarrow{\mathcal{D}} N(0, A_\beta^{-1} \Sigma_\beta A_\beta^{-1})$. \square

Proof of Theorem 3.2

(a) Since $\hat{\vartheta}(t_0, u_0) = \tilde{\vartheta}(t_0, u_0, \hat{\beta})$, we have $\hat{\vartheta}(t_0, u_0) \xrightarrow{\mathcal{P}} \vartheta_0(t_0, u_0)$ uniform in $t \in [0, \tau]$

and $u \in [u_1, u_2]$ by Lemma 1 in (add citation) and Theorem 1. Then

$$\begin{aligned} \sup_{t_0 \in [t_1, t_2]} |\hat{\vartheta}(t_0) - \vartheta_0(t_0)| &= \sup_{t_0 \in [t_1, t_2]} |n^{-1} \sum_{j=1}^n \{\hat{\vartheta}(t_0, U_j(t_0)) - \vartheta_0(t_0, U_j(t_0))\}| \\ &\leq \sup_{t_0 \in [t_1, t_2], u_0 \in [u_1, u_2]} |\hat{\vartheta}(t_0, u_0) - \vartheta_0(t_0, u_0)| = o_p(1). \end{aligned}$$

(b) Following the proof of Lemma 4 in (add citation), we have

$$\begin{aligned} &\sqrt{nh_t h_u} \{\tilde{\alpha}(t_0, u_0, \beta_0) - \alpha_0(t_0, u_0)\} \\ &= -\mathcal{J}_1 e_{11}^{-1}(t_0, u_0) \sqrt{\frac{h_t h_u}{n}} \sum_{i=1}^n \int_0^\tau \{dN_i(t) - \lambda_i(t) dt\} \tilde{X}_i(t) K_{h_t}(t - t_0) K_{h_u}(U_i(t) - u_0) \\ &+ \frac{1}{2} \sqrt{nh_t h_u} \nu_2 e_{11}^{-1}(t_0, u_0) b_\alpha(t_0, u_0) + \frac{1}{2} \sqrt{nh_t h_u} \nu_2 e_{11}^{-1}(t_0, u_0) b_\gamma(t_0, u_0) + o_p(\sqrt{nh_t h_u} (h^2 + b^2)) \end{aligned}$$

Note that $e_{11}^{-1}(t_0, u_0)b_\gamma(t_0, u_0)$ is zero for the first p_1 components and $e_{11}^{-1}(t_0, u_0)b_\alpha(t_0, u_0)$ is $\ddot{\alpha}(t_0)$ for the first p_1 components. Then

$$\begin{aligned}
& \sqrt{nh_t}\{\hat{\alpha}(t_0) - \alpha_0(t_0)\} \\
&= -\sqrt{\frac{h_t}{n}} \sum_{i=1}^n \int_0^\tau \{dN_i(t) - \lambda_i(t)dt\} \{n^{-1} \sum_{j=1}^n \mathcal{I}_1 e_{11}^{-1}(t_0, U_j(t_0)) Q_i(t) K_{h_u}(U_i(t) - U_j(t_0))\} \\
&\quad \times K_{h_t}(t - t_0) \\
&+ \sqrt{\frac{h_t}{n}} n^{-1} \sum_{j=1}^n \{e_{11}(t_0, U_j(t_0))^{-1} e_{12}(t_0, U_j(t_0))\} (\hat{\beta} - \beta_0) \\
&+ \frac{1}{2} \sqrt{nh_t} h_t^2 \nu_2 \ddot{\alpha}(t_0).
\end{aligned}$$

By Lemma A.1 in Yin et al. (2008),

$$\frac{1}{n} \sum_{j=1}^n e_{11}^{-1}(t_0, U_j(t_0)) K_{h_u}(u - U_j(t_0)) = e_{11}^{-1}(t_0, u) + O_p\left(\frac{\log h_u}{\sqrt{nh_u}}\right) + O(h_u^2)$$

uniformly in $t \in [t_1, t_2]$ and $u \in [u_1, u_2]$. It follows that

$$\begin{aligned}
& \sqrt{nh_t}\{\hat{\alpha}(t_0) - \alpha_0(t_0) - \frac{1}{2}h_t^2\nu_2\ddot{\alpha}(t_0)\} \\
&= \sqrt{\frac{h_t}{n}} \sum_{i=1}^n \int_0^\tau \{dN_i(t) - \lambda_i(t)dt\} \mathcal{I}_1 e_{11}^{-1}(t_0, U_i(t)) Q_i(t) K_{h_t}(t - t_0) + o_p(1) \\
&= n^{-1/2} \sum_{i=1}^n g_i(t_0) + o_p(1),
\end{aligned}$$

where $g_i(t_0) = h_t^{1/2} \int_0^\tau \{dN_i(t) - \lambda_i(t)dt\} \mathcal{I}_1 e_{11}^{-1}(t_0, U_i(t)) Q_i(t) K_{h_t}(t - t_0)$. Following the arguments of Lemma 2 of Sun (2010),

$$\sqrt{nh_t}(\hat{\alpha}(t) - \alpha_0(t) - \frac{1}{2}h_t^2\nu_2\ddot{\alpha}(t_0)) \xrightarrow{\mathcal{D}} N(0, \Sigma_\alpha(t)) \tag{A.8}$$

where

$$\Sigma_\alpha(t_0) = \lim_{n \rightarrow \infty} h_t E \left[\int_0^\tau \{dN_i(t) - \lambda_i(t)dt\} \mathcal{S}_1 e_{11}^{-1}(t_0, U_i(t)) Q_i(t) K_{h_t}(t - t_0) \right]^{\otimes 2}.$$

□

Proof of Theorem 3.3

Following the same argument as the proof of Theorem 3.2, we have $\hat{\gamma}(u) \xrightarrow{\mathcal{P}} \gamma_0(u)$ uniformly in $u \in [u_1, u_2]$, and $\sqrt{nh_u}(\hat{\gamma}(u) - \gamma_0(u) - \frac{1}{2}h_u^2 \nu_2 \ddot{\gamma}(u)) \xrightarrow{\mathcal{D}} N(0, \Sigma_\gamma(u))$.

□