SEMIPARAMETRIC TIME-VARYING COEFFICIENT REGRESSION MODEL
FOR LONGITUDINAL DATA WITH CENSORED TIME ORIGIN

by

Qiong Shou

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

2012

Approved by:

Dr. Yanqing Sun

Dr. Zongwu Cai

Dr. Jiancheng Jiang

Dr. Ron Sass
QIONG SHOU. Semiparametric time-varying coefficient regression model for longitudinal data with censored time origin. (Under the direction of DR. YANQING SUN)

In preventive HIV vaccine efficacy trials, thousands of HIV-negative volunteers are randomized to receive vaccine or placebo, and are monitored for HIV infection. The primary objective is to assess vaccine efficacy to prevent HIV infection. An important aspect of vaccine efficacy trials is to assess whether vaccine decreases secondary transmission of HIV and ameliorates HIV disease progression in vaccine recipients who become infected.

This thesis investigates the vaccine effect on the post HIV longitudinal biomarkers (e.g., viral loads and CD4 counts) over time since the actual HIV acquisition. The method applies to the situation when the time of the actual HIV acquisition may be missing or censored.

The problem is investigated under the semiparametric additive time-varying coefficient model where the influences of some covariates vary nonparametrically with time while the effects of the other covariates remain constant. The weighted profile least squares estimators are developed for the unknown parameters as well as for the nonparametric coefficient functions. The method uses the expectation maximization approach to deal with the censored time origin. The asymptotic properties of both the parametric and nonparametric estimators are derived and the consistent estimates of the asymptotic variances are given. The numerical simulations are conducted to examine finite sample properties of the proposed estimators. The method is also applied to a real data from the STEP study with MITT cases.
ACKNOWLEDGMENTS

Upon the completion of this thesis I would sincerely gratefully express my thanks to many people. First of all I would like to show my respect and gratitude to my supervisor, Dr. Yanqing Sun who was greatly helpful and offered invaluable guidance to me on both academic performance and personal life during my study at University of North Carolina at Charlotte. Her attitude to research work and her attitude to life deeply engraved in my heart and memory. Special thanks also go to the members of the supervisory committee, Dr. Zongwu Cai, Dr. Jiancheng Jiang and Dr. Ron Sass without whose solid wisdom and abundant assistance I would not have accomplished my doctoral study. Also I would never forget the generous support from Dr. Xiyuan Qian at East China University of Science and Technology and Dr. Peter B. Gilbert at University of Washington and Fred Hutchinson Cancer Research Center. Without their professional knowledge and endless patience the application would not have been realized so successfully. Not forgetting to my honorable professors in the Department of Mathematics and Statistics department who supported me on such an unforgettable and unique study experience for five years.

Here allow me to express my full love and gratitude to my beloved families. Speechless thanks for their understanding, their support and their love during so many years from my birth. Finally, deep gratitude also due to my graduate friends through the duration of my study in this university. Thank all of them for sharing their time with me and priceless assistance.
TABLE OF CONTENTS

LIST OF TABLES vi

LIST OF FIGURES vii

CHAPTER 1: INTRODUCTION 1
  1.1 A motivating example 1
  1.2 Existing works 3

CHAPTER 2: ESTIMATION APPROACH THROUGH EM ALGORITHM 4
  2.1 Preliminaries 4
  2.2 Estimation Procedures 6
  2.3 Computational algorithm 9
  2.4 Cross-validation bandwidth selection 12

CHAPTER 3: ASYMPTOTIC PROPERTIES 13

CHAPTER 4: A SIMULATION STUDY 18

CHAPTER 5: REAL DATA APPLICATION 29

REFERENCES 39

APPENDIX A: PROOFS OF LEMMA AND THEOREM 45
  A.1 Preliminaries 45
  A.2 Some Lemmas 47
  A.3 Proof of Theorems 62
LIST OF TABLES

TABLE 4.1: Summary statistics for the estimators $\hat{\gamma}$ and $\hat{\beta}(t)$ with no right censoring 22

TABLE 4.2: Summary statistics for the estimators $\hat{\gamma}$ and $\hat{\beta}(t)$ with 30% right censoring rate 23

TABLE 4.3: Summary statistics for the estimator $\hat{\gamma}$ with misplaced time origin and 50% left censoring in the presence ($c_R = 30\%$) and absence ($c_R = 0\%$) of right censoring 24

TABLE 4.4: Summary statistics for the estimator $\hat{\beta}(t)$ for misplaced time origin and 50% left censoring in the presence ($c_R = 30\%$) and absence ($c_R = 0\%$) of right censoring 25
LIST OF FIGURES

FIGURE 1.1: Time since actual HIV acquisition in case of Ab+ and PCR+. 2

FIGURE 4.1: The averages of the estimator $\beta(t) = (\beta_0(t), \beta_1(t))^T$ for $n = 300$, $h = 0.4$ and 30% right censoring rate. The solid lines are for $\beta_1(t)$ and the dashed lines are for $\beta_0(t)$. The grey lines are the true curves. Figure (a), (b) and (c) shows the averages in the case of 0%, 20% and 50% left censoring rate of $S_i$ respectively. Figure (d) shows the results in the case of misplaced time origin by ignoring $S_i$. 26

FIGURE 4.2: The sample and estimated standard errors of the estimator $\beta(t) = (\beta_0(t), \beta_1(t))^T$ for $n = 300$, $h = 0.4$ and 30% right censoring rate. The solid lines are for $\beta_1(t)$ and the dashed lines are for $\beta_0(t)$. The grey lines are the estimated standard error and the black ones are the sample standard error. Figure (a), (b) and (c) shows the results in the case of 0%, 20% and 50% left censoring rate of $S_i$ respectively. Figure (d) shows the results in the case of misplaced time origin by ignoring $S_i$. 27

FIGURE 4.3: The coverage probabilities of 95% pointwise confidence intervals of $\beta(t) = (\beta_0(t), \beta_1(t))^T$ for $n = 300$, $h = 0.4$ and 30% right censoring rate. The solid lines are for $\beta_1(t)$ and the dashed lines are for $\beta_0(t)$. Figure (a), (b) and (c) shows the averages in the case of 0%, 20% and 50% left censoring rate of $S_i$ respectively. Figure (d) shows the results in the case of misplaced time origin by ignoring $S_i$. 28

FIGURE 5.1: Histogram of times $T_{ij}$ from the first positive Elisa confirmed by Western Blot or RNA to subsequent visits. 33

FIGURE 5.2: Histograms of times $S_i$ from actual HIV acquisition to the first positive Elisa confirmed by Western Blot or RNA. Figure (a) shows the histogram of observed $S_i$'s ($R_i = 1$) while figure (b) shows the counts of censored $S_i$'s ($R_i = 0$). 34

FIGURE 5.3: Histograms of times $C_i$ from the first positive Elisa confirmed by Western Blot or RNA to ART initiation or censoring. 35

FIGURE 5.4: The Kaplan Meier estimator of the distribution function of the time from actual HIV acquisition to the first positive Elisa confirmed by Western Blot or RNA. 36
FIGURE 5.5: Estimation of $\beta(t) = (\beta_0(t), \beta_1(t), \beta_2(t))^T$ based on the data from STEP study with MITT cases. Figure (a) shows the estimated intercept, $\beta_0(t)$ and its 95% pointwise confidence interval. Figure (b) shows the estimated effect of the square root of CD4 effect, $\beta_1(t)$ and its 95% pointwise confidence interval. Figure (c) shows the estimated treatment effect, $\beta_2(t)$ and its 95% pointwise confidence interval. The solid curves are the estimated curves and the dashed curves are the confidence intervals.

FIGURE 5.6: The scatter plot of residuals of the subjects with $R_t = 1$. 

37
CHAPTER 1: INTRODUCTION

1.1 A motivating example

In preventive HIV vaccine efficacy trials, thousands of HIV-negative volunteers are randomized to receive vaccine or placebo, and are monitored for HIV infection. The primary objective is to assess vaccine efficacy to prevent HIV infection. An important aspect of vaccine efficacy trials is to assess whether vaccine decreases secondary transmission of HIV and ameliorates HIV disease progression in vaccine recipients who become infected (cf., Clemens et al., 1997; Halloran et al., 1997; Clements-Mann, 1998; Nabel, 2001; Shiver et al., 2002; HVTN, 2004; IAVI, 2004).

We propose to investigate the vaccine effect on the post HIV longitudinal biomarkers (e.g., viral loads and CD4 counts). Viral load and CD4 counts have been found to be highly prognostic for both secondary transmission and progression to clinical disease in observational studies (cf., Mellors et al., 1997; HIV Surrogate Marker Collaborative Group, 2000; Quinn et al., 2000; Gray et al., 2001). All previous analyses of HIV vaccine efficacy trials assessed these biomarkers based on the time from HIV positive diagnosis. However, it is biologically meaningful to assess whether vaccination modifies or accelerates the development of these biomarkers over time since the actual HIV acquisition. This assessment can be challenging since exact times of actual HIV acquisition are often unobtainable for trial participants. A brief description of HIV vaccine efficacy trial’s diagnosis algorithm is given in the following.

HIV vaccine trials test volunteers for anti-HIV antibodies at periodic intervals (e.g., every 3 or 6 months); these antibody-based tests have near-perfect sensitivity to detect infections that occurred at least four weeks ago but otherwise may miss
the infection. For all subjects with an HIV antibody positive (Ab+) test, a “look-
back” procedure is applied wherein earlier available blood samples are tested for HIV
infection using a more sensitive antigen-based HIV-specific PCR assay, which has
near-perfect sensitivity if the infection occurred at least one week ago. Therefore,
each infected subject is classified into one of two groups, defined by whether the
earliest HIV positive sample is Ab- and PCR+ or is Ab+ and PCR+. The actual
HIV acquisition time is approximated well by the time at Ab- and PCR+, while
actual infection time occur approximately between the first Ab+ and earlier Ab- test
times in the case of Ab+ and PCR+. The Ab+ and PCR+ cases occur in between
20% and 70% of infected subjects, with the rate depending on the frequency of HIV
testing.

![Diagram](image)

Figure 1.1: Time since actual HIV acquisition in case of Ab+ and PCR+.

Consider the \( i = 1, \ldots, n \) subjects who become HIV infected during the HIV
vaccine efficacy trial. Let \( O_i \) be the time of actual HIV acquisition, \( D_i \) the HIV
positive diagnosis time based on the trial’s diagnosis algorithm (first Ab+ test time)
and \( L_i \) the last Ab- test time. Post-infection biomarkers are measured at times
\( T_{i1}, \ldots, T_{in_j} \), where \( T_{ij} \) is the time between the first Ab+ and the time at which the \( j \)th
measurement is taken. Let \( S_i \) be the gap between HIV acquisition and the diagnosis,
\( S_i = D_i - O_i \). If subject \( i \) has an acute sample (Ab- and PCR+), the actual infection
time can be well approximated by \( L_i \) and in this case let \( S_i = D_i - L_i \). Otherwise, \( S_i \)
is less than \( D_i - L_i \). The \( S_i \) (time origin) is left censored by \( D_i - L_i \) with censoring
indicator \( R_i: R_i = 1 \) if \( S_i \) is observed and \( R_i = 0 \) if \( S_i \) is less than \( D_i - L_i \). The time
from actual HIV acquisition to the $j$th sampling time is then $T_{ij}^o = S_i + T_{ij}$. Figure 1.1 illustrates the set-up.

1.2 Existing works

The sampling times $T_{ij}^o = S_i + T_{ij}$ from the actual HIV acquisition are known when $S_i$ is completely observed. In this case many existing statistical methods can be used to analyze model (2.1). Among others, recent works in this area include semiparametric methods by Moyeed & Diggle (1994), Zeger & Diggle (1994), and Liang, Wu & Carroll (2003), nonparametric methods by Hoover, Rice, Wu & Yang (1998), Wu, Chiang & Hoover (1998), Scheike & Zhang (1998), Wu & Zhang (2002), Wu & Liang (2004) and Sun & Wu (2003). Martinussen & Scheike (1999, 2000, 2001) and Lin & Ying (2001) considered time-varying coefficients regression models for longitudinal data and successfully integrated counting process techniques into the analysis of longitudinal data, providing further bridging between survival analysis, recurrent events, and time-dependent observations. Sun and Wu (2005) developed weighted least squares estimation procedure which avoids modeling of the sampling times is asymptotically more efficient than a single nearest neighbor smoothing which depends on estimation of the sampling model.
CHAPTER 2: ESTIMATION APPROACH THROUGH EM ALGORITHM

2.1 Preliminaries

Suppose that there is a random sample of \( n \) subjects. For subject \( i \), let \( Y_i(t) \) be the response process and let \( X_i(t) \) and \( Z_i(t) \) be the possibly time-dependent covariates of dimensions \((p + 1) \times 1\) and \( q \times 1\), respectively, where \( t \) is the time since actual HIV acquisition. The proposed general semiparametric time-varying coefficients regression model assumes that

\[
Y_i(t) = \beta^T(t)X_i(t) + \gamma^TZ_i(t) + \epsilon_i(t), \quad i = 1, \ldots, n
\]

(2.1)

where \( \beta(t) \) is an unspecified \((p + 1) \times 1\) vector of smooth regression functions, \( \gamma \) is a \( q \times 1 \) dimensional vector of parameters, and \( \epsilon_i(t) \) is a mean-zero process. The notation \( x^T \) represents transpose of a vector or matrix \( x \). The first component of \( X(t) \) is specified as 1 in general, giving to a model with a nonparametric baseline. The effect of \( X(t) \) is modeled nonparametrically while the effect of \( Z(t) \) follows a given parameter.

The observations of \( Y_i(t) \) are taken at time points \( T_{i1}^0 < T_{i2}^0 < \cdots < T_{in_i}^0 \), where \( n_i \) is the total number of observations on the \( i \)th subject. The observation times \( T_{ij}^0 \) can be decomposed in two parts \( T_{ij}^0 = S_i + T_{ij} \) as shown in Figure 1.1, where \( S_i \) is the time from actual HIV acquisition to the first positive diagnosis test and \( T_{ij} \) is the time from the first positive diagnosis test to the \( j \)th visit for the \( i \)th subject. The number of observations taken on the \( i \)th subject by time \( t \) is \( N_i^o(t) = \sum_{j=1}^{n_i} I(T_{ij}^0 \leq t) \), where \( I(\cdot) \) is the indicator function. Let \( C_i \) be the end of follow-up time or censoring time for the \( i \)th subject starting at HIV positive diagnosis (Ab+ test time). The censoring
time $C_i$ will be allowed to depend on the covariates $X_i(\cdot)$ and $Z_i(\cdot)$. The responses for the $i$th subject can only be observed at the time points before $C_i$. The censoring time since the actual time origin (HIV acquisition) is $S_i + C_i$.

Let the conditional mean rate of the observation times $\alpha_i(t)$ for subject $i$ be defined as

$$E\{dN_i(t) \mid X_i(t), Z_i(t)\} = \alpha(t, U_i(t)) \, dt \equiv \alpha_i(t) \, dt, \quad i = 1, \ldots, n, \quad (2.2)$$

where $U_i(t)$, a $m \times 1$ vector, is the part of the covariates $(X_i(t), Z_i(t))$ that affects the potential sampling times. The function $\alpha(t, u)$ is an unspecified nonnegative smooth function.

The time $S_i$ from actual HIV acquisition to HIV positive diagnosis may be left censored by the censoring time $V_i$. Let $R_i = I(S_i \geq V_i)$ be the censoring indicator. For the application concerned in this paper, the censoring time $V_i$ (e.g. $D_i - L_i$) is assumed to be observed for all subjects. Let $\mathcal{D}_i = \{V_i, C_i, A_i, T_{ij}, X_i(T_{ij}^0), Z_i(T_{ij}^0), Y_i(T_{ij}^0), j = 1, \ldots, n_i\}$, where $A_i$ is a collection of possible auxiliary variables that are not of interest in the modelling of $Y_i(t)$ but may be useful in predicting the distribution of $S_i$. The observed data for subject $i$ can be expressed as $\mathcal{X}_i = \{R_i S_i, R_i, \mathcal{D}_i\}$. The observation is $\{S_i, \mathcal{D}_i\}$ if $R_i = 1$ and $\mathcal{D}_i$ if $R_i = 0$. Although exact times $T_{ij}^0$ may be unobtainable, the values $X_{ij} = X_i(T_{ij}^0)$, $Z_{ij} = Z_i(T_{ij}^0)$ and $Y_{ij} = Y_i(T_{ij}^0)$ at $T_{ij}^0$ are known. Denote the observed data by $\mathcal{X} = \{\mathcal{X}_i, i = 1, 2, \ldots, n\}$.

Assume that $S_i$ and $V_i$ are independent, and that the censoring time $C_i$ is non-informative in the sense that $E\{dN_i^0(t) \mid X_i(t), Z_i(t), S_i + C_i \geq t\} = E\{dN_i(t) \mid X_i(t), Z_i(t)\}$ and $E\{Y_i(t) \mid X_i(t), Z_i(t), S_i + C_i \geq t\} = E\{Y_i(t) \mid X_i(t), Z_i(t)\}$. Assume also that $Y_i(t)$ and $N_i^0(t)$ are independent conditional on $X_i(t)$, $Z_i(t)$ and $S_i + C_i \geq t$. This assumption implies that, conditional on covariate processes, sampling times are non-informative for the response. Note that dependence between response and sampling times as well dependence between sampling times and the censoring time $C_i$ is often
induced by ignoring certain covariates (cf., Miloslavsky et al., 2004 and Zeng, 2005). The stated conditional independence assumptions make the proposed methods applicable to situations where dependence may exist among response process, sampling times and censoring time $C_i$ but becoming independent by including appropriate additional covariates. A recent work by Sun and Lee (2011) on testing independent censoring for longitudinal data provides needed procedures for checking such assumptions. Let $N_i(t) = \sum_{j=1}^{n_i} I(T_{ij} \leq t)$. Assume $E\{Y_i(t)|X_i(\cdot), Z_i(\cdot), S_i, V_i, C_i\} = E\{Y_i(t)|X_i(\cdot), Z_i(\cdot)\}$.

When all $S_i$’s are observed, the existing statistical methods cited in Section 1.2 can be used to analyze model (2.1). However, none of these methods address the problem in which the time origin may be censored. We propose to extend the investigation of model (2.1) to accommodate this situation.

### 2.2 Estimation Procedures

It is important to note that if the unobserved or censored $S_i$ is treated as missing, then $S_i$ is not missing at random in the sense of Robin (1976). The inverse probability weighting of complete-cases method of Horvitz and Thompson (1952) and the augmented inverse probability weighted complete-case method of Robins, Rotnitzky and Zhao (1994), which have been successfully adapted in Sun and Gilbert (2011), Sun, Wang and Gilbert (2011) and by many other authors, will not work in this situation. We propose an estimation procedure based on the missing-data principle using the EM-algorithm. The EM-algorithm has been applied by Scheike and Sun (2007) to develop maximum likelihood estimation for tied survival data under Cox regression model.

Let $F_S(s|D_i)$ be the conditional distribution of $S_i$ given $D_i$. The conditional distribution of $S_i$ given $D_i$ and $R_i = 0$, $F_S(s|D_i, R_i = 0)$, equals $F_S(s|D_i)/F_S(V_i|D_i)$ for $s \leq V_i$ and 1 for $s > V_i$. Assume that $\max\{S_i, V_i\}$ is bounded by a predetermined constant $c$. This is reasonable since for the application concerned here $\max\{S_i, V_i\}$
is less than the time interval between two consecutive testing times which is usually between 3 and 6 months. The distribution of $S_i$ based on the left censored data can be estimated by using the right censored data through the transformation $\{\min\{c - S_i, c - V_i\}, R_i = I(c - S_i \leq c - V_i)\}$. Thus, the Kaplan-Meier estimator can be used to estimate the distribution of $S_i$ when $S_i$ is independent of $D_i$. Otherwise, a failure time regression model such as the Cox model (Cox, 1972) can be used to estimate the conditional distribution $F_S(s|D_i)$. Observing the censoring time $V_i$ for all subjects is a key factor in the estimation of $F_S(s|D_i, R_i = 0)$. Otherwise $F_S(s|D_i, R_i = 0)$ is not identifiable.

Let $\hat{F}_S(s|D_i)$ be the estimated conditional distribution of $F_S(s|D_i)$. The probability $\pi_i = P(R_i = 1|D_i) = P(S_i \geq V_i|D_i)$ can be estimated by $\hat{\pi}_i = 1 - \hat{F}_S(V_i|D_i)$. Let $dN^c_i(t) = I(S_i + C_i \geq t)dN^o_i(t)$. The estimation of model (2.1) will be based on targeting to minimize the following objective function:

$$l(\beta, \gamma) = \sum_{i=1}^{n} R_i \int_0^\tau W_i(u)\{Y_i(u) - \beta^T(u)X_i(u) - \gamma^TZ_i(u)\}^2 dN^c_i(u)$$

$$+ \sum_{i=1}^{n} (1 - R_i) \hat{E}_S\left\{ \int_0^\tau W_i(u)\{Y_i(u) - \beta^T(u)X_i(u) - \gamma^TZ_i(u)\}^2 dN^c_i(u)|\mathcal{X} \right\}, \quad (2.3)$$

where $W_i(\cdot)$ is a nonnegative weight function, and $\hat{E}_S\{\cdot|\mathcal{X}\}$ is the estimate of the conditional expectation, $E_S\{\cdot|\mathcal{X}\}$, of a function of $S_i$ given $\mathcal{X}$. For $R_i = 0$ and a smooth random function $G_n(t, X_i(t), Z_i(t), Y_i(t))$, $\hat{E}_S\{\int_0^\tau G_n(u, X_i(u), Z_i(u), Y_i(u))dN^c_i(u)|\mathcal{X}\}$ equals

$$\sum_{j=1}^{n_i} \hat{E}_S\{G_n(S_i + T_{ij}, X_i(T_{ij}^o), Z_i(T_{ij}^o), Y_i(T_{ij}^o))I(C_i \geq T_{ij})I(S_i + T_{ij} \leq \tau)|\mathcal{X}\}$$

$$= \sum_{j=1}^{n_i} \hat{E}_S\{G_n(S_i + T_{ij}, X_{ij}, Z_{ij}, Y_{ij})I(C_i \geq T_{ij})I(S_i + T_{ij} \leq \tau)|\mathcal{X}\}$$
where \( \int U \) The root of the equation

This leads to the following estimating function

where and hereafter, the notation \( \ll H_i(t) \gg_R = R_i H_i(t) + (1 - R_i) \hat{E}S \{ H_i(t) | X \} \)

is used for a random function \( H_i(t) \), and \( K_h(t) = h^{-1} K(t/h) \), \( K(t) \) is a symmetric
kernel function with a compact support and \( h \) is the bandwidth depending on \( n \).

Taking derivative of \( \tilde{I}_t(\beta, \gamma) \) with respect to \( \beta \) for a fixed \( \gamma \) yields

This leads to the following estimating function

The root of the equation \( U_i(\beta, \gamma) = 0 \) is denoted by \( \tilde{\beta}(t, \gamma) \). Let \( \tilde{E}_{yx}(t) = n^{-1} \sum_{i=1}^n \ll \int_0^T K_h(u-t)Z_i(u)X_i^T(u) dN^c_i(u) \gg_R \).

The \( \tilde{E}_{yx}(t) \) and \( \tilde{E}_{xx}(t) \) are similarly defined by replacing \( Z_i \) with \( Y_i \) and \( X_i \) respectively. The local least squares estimator for \( \beta(t) \)
for fixed \( \gamma \) is then given by

where \( \hat{Y}_x(t) = \tilde{E}_{yx}(t)(\tilde{E}_{xx}(t))^{-1} \) and \( \hat{Z}_x(t) = \tilde{E}_{xx}(t)(\tilde{E}_{xx}(t))^{-1} \). Replacing \( \tilde{\beta}(t; \gamma) \) for
\(\beta(t)\) in (2.3) and taking derivative with respect to \(\gamma\), we obtain the profile estimating

equation for \(\gamma\):

\[
\sum_{i=1}^{n} \int_{t_1}^{t_2} \ll W_i(t)\{Z_i(t) - \tilde{Z}_x(t)X_i(t)\}\{Y_i(t) - X_i^T(t)\tilde{\beta}(t; \gamma)
- Z_i^T(t)\gamma\} \, dN^c_i(t) > R = 0.
\]

(2.8)

Here \([t_1, t_2]\) is taken as a subinterval of \([0, \tau]\) to avoid boundary problems in the theoretical justifications. In practice, \([t_1, t_2]\) can be taken as \([0, \tau]\). From (2.8), we solve for \(\gamma\) to get \(\hat{\gamma}\) which equals \(\hat{\beta}(t; \hat{\gamma})\).

When \(S_i\) is observed for all subjects, \(R_i = 1\). The estimators for \(\beta(t)\) and \(\gamma\) are reduced to those under Sun and Wu (2005). However, when \(S_i\) is censored, the estimating equations (2.5) and (2.8) are weighted according to the conditional distribution of \(S_i\) so that the estimated covariate effects correspond to those at the time since the actual time origin. A key factor for this procedure to work is that the censoring time \(V_i\) is observed for all subjects so that the estimation of \(F_S(s \mid D_i, R_i = 0)\) is possible.

2.3 Computational algorithm

The boundary effects on the estimation of \(\beta(t)\) and the covariance matrix of its estimator can be reduced by applying the equivalent kernel for the local linear approach; see Fan & Gijbels (1996).

Suppose the binary data \((T_1, B_1), (T_2, B_2), \ldots, (T_n, B_n)\) which are \(n\) independent and identically distributed copies from \((T, B)\). To estimate \(m(t_0) = E(B \mid T = t_0)\) is of interest. Suppose we use symmetric kernel \(K(x)\) in local constant method. Then
the local constant estimator of \( m(t) \) at point \( t_0 \) will be

\[
\hat{m}_C = \frac{n^{-1} \sum_{i=1}^{n} K_h(T_i - t_0)B_i}{n^{-1} \sum_{i=1}^{n} K_h(T_i - t_0)}.
\]

To get the equivalent kernel, we will mimic some notations in Fan & Gijbels (1996).

\[
S_{n,j}(t_0) = \sum_{i=1}^{n} K_h(T_i - t_0)(T_i - t_0)^j, \quad j = 0, 1, 2.
\]

Then

\[
S_n(t_0) = \begin{pmatrix}
S_{n,0}(t_0) & S_{n,1}(t_0) \\
S_{n,1}(t_0) & S_{n,2}(t_0)
\end{pmatrix}.
\]

Meanwhile the inverse can be written as

\[
S_n^{-1}(t_0) = \frac{1}{S_{n,0}(t_0)S_{n,2}(t_0) - S_{n,1}^2(t_0)}
\begin{pmatrix}
S_{n,2}(t_0) & -S_{n,1}(t_0) \\
-S_{n,1}(t_0) & S_{n,0}(t_0)
\end{pmatrix}.
\]

As stated in the Section 3.2.2 of Fan & Gijbels (1996), the equivalent kernel for local linear approach is

\[
K_h^*(t - t_0) = e_1^T S_n^{-1}(t_0)(1\ t - t_0)^T K_h(t - t_0),
\]

where \( e_1 = (1\ 0)^T \). Thus we can simplify the equivalent kernel as follows.

\[
K_h^*(t - t_0) = e_1^T S_n^{-1}(t_0)(1\ t - t_0)^T K_h(t - t_0) \\
= \frac{K_h(t - t_0)(1\ 0)}{S_{n,0}(t_0)S_{n,2}(t_0) - S_{n,1}^2(t_0)}
\begin{pmatrix}
S_{n,2}(t_0) & -S_{n,1}(t_0) \\
-S_{n,1}(t_0) & S_{n,0}(t_0)
\end{pmatrix}
\begin{pmatrix}
1 \\
t - t_0
\end{pmatrix} \\
= \frac{\{S_{n,2}(t_0) - S_{n,1}(t_0)(t - t_0)\} K_h(t - t_0)}{S_{n,0}(t_0)S_{n,2}(t_0) - S_{n,1}^2(t_0)}.
\]

Therefore, the local linear estimator \( \hat{m}_L \) at point \( t_0 \) under the model \( B = m(T) + \epsilon \) is

\[
\frac{n^{-1} \sum_{i=1}^{n} K_h^*(T_i - t_0)B_i}{n^{-1} \sum_{i=1}^{n} K_h^*(T_i - t_0)} = \frac{\sum_{i=1}^{n} \{S_{n,2}(t_0) - S_{n,1}(t_0)(T_i - t_0)\} K_h(T_i - t_0)B_i}{\sum_{i=1}^{n} \{S_{n,2}(t_0) - S_{n,1}(t_0)(T_i - t_0)\} K_h(T_i - t_0)}.
\]
Compared to the local constant estimator above, if we use the following kernel

\[ W_h (T_i - t_0) = \{ S_{n,2}(t_0) - S_{n,1}(t_0)(T_i - t_0)\} K_h (T_i - t_0) \]  \hspace{1cm} (2.9)

instead of \( K_h (T_i - t_0) \), we simply obtain the local linear estimator.

Let \( f(t) \) be the density function of \( T \). For a interior point \( t_0 \), the local linear estimator is asymptotically equivalent to the local constant estimator as \( h \to 0 \) and \( nh^5 = O(1) \) since for a symmetric kernel, \( \int K(x)x \, dx = 0 \). Then

\[ n^{-1} ES_{n,j}(t_0) = EK_h(T_i - t_0)(T_i - t_0)^j = \int K_h(t-t_0)(t-t_0)^j f(t) \, dt \]

\[ = \int K(x)h^j x^j f(t_0 + hx) \, dx = h^j (f(t_0) + o(h)) \int K(x)x^j \, dx = o(h). \]

Especially note that \( n^{-1} ES_{n,1}(t_0) = 0 \). Hence

\[ \hat{m}_L = \frac{\sum_{i=1}^n \{ S_{n,2}(t_0) - S_{n,1}(t_0)(T_i - t_0)\} K_h (T_i - t_0) B_i}{\sum_{i=1}^n \{ S_{n,2}(t_0) - S_{n,1}(t_0)(T_i - t_0)\} K_h (T_i - t_0)} \]

\[ \approx \frac{n^{-1} \sum_{i=1}^n \{ n^{-1} S_{n,2}(t_0) - n^{-1} S_{n,1}(t_0) h \frac{T_i - t_0}{h} \} K_h (T_i - t_0) B_i}{n^{-1} \sum_{i=1}^n n^{-1} S_{n,2}(t_0) - n^{-1} S_{n,1}(t_0) h \frac{T_i - t_0}{h} K_h (T_i - t_0)} \]

\[ = \hat{m}_C + o_p(h^2). \]

Thus \((nh)^{1/2}(\hat{m}_L - \hat{m}_C) = o_p((nh^5)^{1/2})\), which means the asymptotic distributions for the local linear estimator and the local constant estimator are the same for an interior point \( t_0 \) as \( h \to 0 \) and \( nh^5 = O(1) \). This enables using the equivalent kernel for the boundary time points while using the kernel in local constant approach for the interior time points.

In estimating \( \beta(t) \), time points \( T \) may be unknown since \( S_i \) is left censored by \( V_i \). Then we cannot simply use \( S_{n,j}(t_0) \) defined above. Let

\[ S_{n,j}(t) = \sum_{i=1}^n \ll \int_0^\tau K_h(u-t)(u-t)^j dN_i^e(u) \gg, \quad j = 0, 1, 2. \]

Now under the new definition of \( S_{n,j}(t_0) \), we still have the form of equivalent kernel
in (2.9) for local linear approach of estimating $\beta(t)$.

2.4 Cross-validation bandwidth selection

The optimal theoretical bandwidth is difficult to achieve since it would involve estimating the second derivative of $\beta(t)$ with respect to $t$, $\beta''(t)$; see Fan and Gijbels (1996) and Cai and Sun (2002). In practice, the appropriate bandwidth selection can be based on a cross-validation method. This approach is widely used in nonparametric function estimation literature; see Rice and Silverman (1991) for leave-one-subject-out cross-validation approach and Tian, Zucker and Wei (2005) for $K$-fold cross-validation approach.

An analog of the $K$-fold cross-validation approach in the current setting is to divide the data into $K$ equal-sized groups. Let $D_k$ denote the $k$th subgroup of data, then the $k$th prediction error is given by

$$PE_k(h) = \sum_{i \in D_k} \ll \int_{t_1}^{t_2} \left[ Y_i(t) - \left( \hat{\beta}_{(-k)}(t) \right)^T X_i(t) - \hat{\gamma}_{(-k)}^T Z_i(t) \right]^2 dN_i^c(t) \gg_R, \quad (2.10)$$

for $k = 1, \ldots, K$, where $\hat{\beta}_{(-k)}(t)$ and $\hat{\gamma}_{(-k)}$ are the estimators of $\beta(t)$ and $\gamma$ based on the data without the subgroup $D_k$. The data-driven bandwidth selection based on the $K$-fold cross-validation is to choose the bandwidth $h$ that minimizes the total prediction error $PE(h) = \sum_{k=1}^{K} PE_k(h)$. This bandwidth selection procedure will be further studied and tested empirically through simulations.
CHAPTER 3: ASYMPTOTIC PROPERTIES

In this section we will explore the asymptotic properties of the proposed estimators. First we will introduce some notations. Let

\[ e_{zx}(t) = E(\xi_i(t)\alpha_i(t)Z_i(t)X_i^T(t)), \]

where \( \xi_i(t) = I(S_i + C_i + c_1 \geq t) \) and \( \alpha_i(t) \) is the conditional mean rate of \( N_i^o(t) \) defined in (2.2). \( e_{xx}(t) \) and \( e_{yx}(t) \) are similarly defined. Let \( y_x(t) = e_{yx}(t)(e_{xx}(t))^{-1} \) and \( z_x(t) = e_{xx}(t)(e_{xx}(t))^{-1} \). Let \( \gamma_0 \) and \( \beta_0(t) \) be the true values of \( \gamma \) and \( \beta(t) \) respectively. In additional to the conditional independence assumptions and noninformative censoring assumptions stated in Section 2.1 we assume the following conditions for the asymptotic results to hold.

**Conditions (I).** Assume that the \( \{n_i\} \) are bounded; the \( \{S_i\} \) are bounded by a large enough value \( L \) and independent of \( D_i \); the kernel function \( K(\cdot) \) is symmetric with compact support on \([-1, 1]\); the processes \( X_i(t), Z_i(t) \) and \( \alpha_i(t), 0 \leq t \leq \tau \), are bounded by a constant, continuous and their total variations are bounded by a constant; the values of the \( j \)th measurement \( X_{ij} \) and \( Z_{ij} \) are also bounded; \( (e_{xx}(t))^{-1} \) for \( 0 \leq t \leq \tau \) are bounded; the weight function \( W_i(t) \) can be written as a difference of two monotone functions and each converges to a deterministic function so that \( W_i(t) \) converges to \( w(t) \) for all \( i \).

Under Conditions (I), it follows from Lemma A.2.3 that \( \tilde{E}_{zx}(t) \xrightarrow{P} e_{zx}(t) \) uniformly in \( t \in [t_1, t_2] \subset [0, \tau] \). Similar asymptotic results hold for \( \tilde{E}_{yx}(t) \) and \( \tilde{E}_{xx}(t) \). By continuous mapping theorem, the above results lead to the conclusion that \( \tilde{Y}_x(t) \) and \( \tilde{Z}_x(t) \) converge to \( y_x(t) \) and \( z_x(t) \) uniformly in \( t \in [t_1, t_2] \) respectively.
Theorem 3.1 and Theorem 3.2 state that both the estimators \( \hat{\gamma} \) and \( \tilde{\beta}(t) \) are consistent. Note that \( \hat{\gamma} \) is the minimizer of \( n^{-1}\tilde{I}(\gamma) = n^{-1}I(\tilde{\beta}(\cdot; \gamma), \gamma) \) which equals

\[
n^{-1} \sum_{i=1}^{n} \ll \int_{0}^{\tau} W_i(s)\{Y_i(s) - \hat{Y}_x(s)X_i(s) + \gamma T(\hat{Z}_x(s)X_i(s) - Z_i(s))\}^2 dN_i(s) \gg_R.
\]

In the proof of Theorem 3.1, we show that \( n^{-1}\tilde{I}(\gamma) \) converges uniformly to a deterministic function of \( \gamma \) that minimizes at \( \gamma = \gamma_0 \). Then the consistency of \( \hat{\gamma} \) follows by Theorem 5.7 of van der Vaart (1998).

**Theorem 3.1:** (Consistency of \( \hat{\gamma} \)) Under Condition (I), \( \hat{\gamma} = D^{-1}\tilde{W} \) converges to its true value \( \gamma_0 \) in probability as \( n \to \infty \).

The consistency of \( \tilde{\beta}(t) \) follows from Lemma A.2.3 and Theorem 3.1,

\[
\tilde{\beta}(t) = \hat{Y}_x^T(t) - \hat{Z}_x^T(t)\hat{\gamma} \xrightarrow{p} y_x^T(t) - \varepsilon_x^T(t)\gamma_0
\]

\[
= (e_{xx}(t))^{-1}e_{yx}(t) - (e_{xx}(t))^{-1}e_{xx}(t)\gamma_0
\]

\[
= (e_{xx}(t))^{-1}[E(\xi(t)\alpha_i(t)X_i(t)Y_i^T(t)) - E(\xi(t)\alpha_i(t)X_i(t)Z_i^T(t))]\gamma_0
\]

\[
= (e_{xx}(t))^{-1}E(\xi(t)\alpha_i(t)X_i(t)[Y_i^T(t) - Z_i^T(t)\gamma_0])
\]

\[
= (e_{xx}(t))^{-1}E(\xi(t)\alpha_i(t)X_i(t)[X_i^T(t)\beta_0(t) + \varepsilon^T(t)]
\]

\[
= (e_{xx}(t))^{-1}e_{xx}(t)\beta_0(t) + E(E[\xi(t)\alpha_i(t)X_i(t)\varepsilon^T(t) | X_i(t), Z_i(t), S_i + C_i \geq t])
\]

\[
= \beta_0(t) + E(\xi(t)\alpha_i(t)X_i(t)E[\varepsilon^T(t) | X_i(t), Z_i(t), S_i + C_i \geq t])
\]

\[
= \beta_0(t) + E(\xi(t)\alpha_i(t)X_i(t)E[\varepsilon^T(t) | X_i(t), Z_i(t)]) = \beta_0(t).
\]

**Theorem 3.2:** (Consistency of \( \tilde{\beta}(t) \)) Under Condition (I), \( \tilde{\beta}(t) = \tilde{\beta}(t; \hat{\gamma}) \) converges to \( \beta_0(t) \) in probability uniformly on \( [t_1, t_2] \) as \( n \to \infty \), where \( 0 \leq t_1 \leq t_2 \leq \tau \).

The details of the proofs of Theorem 3.1 and 3.2 are given in the Appendix A.3.

In Section 2.2, \( \hat{\gamma} \) is the solution of (2.8). So denote \( U(\gamma) \) as

\[
\sum_{i=1}^{n} \int_{t_1}^{t_2} \ll W_i(t)\{Z_i(t) - \tilde{Z}_x(t)X_i(t)\}\{Y_i(t) - X_i^T(t)\tilde{\beta}(t; \gamma) - Z_i^T(t)\gamma\} dN_i^c(t) \gg_R
\]
which is usually called the score function. Then the Taylor expansion of $U(\hat{\gamma})$ at $\gamma_0$ is

$$n^{1/2}(\hat{\gamma} - \gamma_0) = -\left( n^{-1} \frac{\partial U(\gamma^*)}{\partial \gamma^T} \right)^{-1} \left[ n^{-1/2} U(\gamma_0) \right],$$

where $\gamma^*$ is on the line segment between $\hat{\gamma}$ and $\gamma_0$. To prove the asymptotic normality of $n^{1/2}(\hat{\gamma} - \gamma_0)$, it is sufficient to prove the convergence in probability to a non-singular matrix of $n^{-1} \frac{\partial U(\gamma^*)}{\partial \gamma^T}$, and the weak convergence of $n^{-1/2} U(\gamma_0)$. The convergence in probability can be easily obtained by applying Lemma A.2.2. And $n^{-1/2} U(\gamma_0)$ can be derived to equal to

$$n^{-1/2} \sum_{i=1}^{n} \int_{t_1}^{t_2} w(t) \{ Z_i(t) - z_x(t)X_i(t) \} \epsilon_i(t) [ R_i dN^c_i(t) \]
\[+ E_s \{ (1 - R_i) dN^c_i(t) \mid \mathcal{D}_i, R_i = 0 \} \] + o_p(1).$$

Then applying theorem 5.21 (van der Vaart, 1998) to the sore function, the asymptotic normality of $\hat{\gamma}$ is presented in the following theorem.

**Theorem 3.3:** (Asymptotic Normality of $\hat{\gamma}$) Under Condition (I), $n^{1/2}(\hat{\gamma} - \gamma_0) \xrightarrow{D} \mathcal{N}(0, D^{-1} V D^{-1})$ as $n \to \infty$ where

$$D = E \left( \int_{t_1}^{t_2} w(t) \{ Z_i(t) - z_x(t)X_i(t) \} \right)^{\otimes 2} dN^c_i(t),$$

$$V = E \left\{ \int_{t_1}^{t_2} [ R_i w(t) (Z_i(t) - z_x(t)X_i(t)) \epsilon_i(t) dN^c_i(t) \]
\[+ (1 - R_i) E_s \{ w(t)(Z_i(t) - z_x(t)X_i(t)) \epsilon_i(t) dN^c_i(t) \mid \mathcal{D}_i, R_i = 0 \} \}^{\otimes 2} \right\}. $$

Based on the equations (A.9) and (A.11), the asymptotic variance above can be estimated by $\hat{D}^{-1} \hat{V} \hat{D}^{-1}$ where

$$\hat{D} = n^{-1} \sum_{i=1}^{n} \int_{t_1}^{t_2} W_i(t) \{ Z_i(t) - \tilde{Z}_x(t)X_i(t) \} \right)^{\otimes 2} dN^c_i(t) \gg R,$$

$$\hat{V} = n^{-1} \sum_{i=1}^{n} \left\{ \int_{t_1}^{t_2} W_i(t)(Z_i(t) - \tilde{Z}_x(t)X_i(t)) \epsilon_i(t) dN^c_i(t) \gg R \right\}^{\otimes 2}$$

and $\hat{\epsilon}_i(t) = Y_i(t) - \hat{\beta}(t)^T X_i(t) - \hat{\gamma}^T Z_i(t)$. This estimator is consistent estimator of
the asymptotic variance by the consistency of \( \hat{D} \) and \( \hat{V} \). The proof of Theorem 3.3 is given in the Appendix A.3.

Before demonstrating the asymptotic normality of \( \hat{\beta}(t) \) at each fixed time point \( t \), we first introduce the following notations. Note that \( N_i^c(t) = \int_0^t I(S_i + C_i \geq u) \, dN_i^u(u) \) is a counting process. Let the filtration \( \mathcal{F}_t^c = \sigma \{ N_i^c(s), R_i, X_i(\cdot), Z_i(\cdot), Y_i(\cdot), 0 \leq s \leq t \} \). By the Doob-Meyer decomposition theorem, there is a unique pair of martingale \( M_i^c(t) \) and compensator \( A_i^c(t) \) with respect to the filtration \( \mathcal{F}_t^c \) such that \( N_i^c(t) = A_i^c(t) + M_i^c(t) \). Let \( Y_i^c(t) = \sum_{j=1}^{n_i} I(T_{ij}^0 \geq t) \). The intensity of \( N_i^c(t) \) is \( Y_i^c(t) \alpha_i^c(t) \). It follows that \( A_i^c(t) = \int_0^t Y_i^c(s) \alpha_i^c(s) \, ds \).

Let \( \tilde{e}_{yx}(t) = \int_0^T K_h(u-t) e_{yx}(u) \, du \). Similar definitions can be defined for \( \tilde{e}_{xx}(t) \) and \( \tilde{e}_{xx}(t) \). Let \( \beta^*(t) = \gamma_0 \). Similar facts hold for \( \tilde{e}_{xx}(t) \) and \( \tilde{e}_{xx}(t) \) too. The transpose of the matrix is denoted by changing the order of the subscripts.

Applying Taylor expansion for \( \hat{\beta}(t) = \tilde{\beta}(t, \hat{\gamma}) \) at \( \gamma_0 \), we have

\[
(nh)^{1/2}(\tilde{\beta}(t) - \beta^*(t)) = (nh)^{1/2}(\tilde{\beta}(t; \gamma_0) - \beta^*(t)) + (nh)^{1/2}(\hat{\gamma} - \gamma_0) \frac{\partial \tilde{\beta}(t; \gamma_0)}{\partial \gamma} + O_p(n^{-1/2}h^{1/2})
\]

From (A.15) in the proof of Theorem 3.4,

\[
\beta^*(t) = \beta_0(t) + (1/2) \mu_2 h^2 (e_{xx}(t))^{-1} (e_{xy}(t) - e_{xx}(t) \gamma_0 - e_{xx}(t) \beta_0(t)) + o(h^2),
\]

where \( e_{xy}(t) \) is the second derivative of \( e_{xy}(t) \) with respect to \( t \). The asymptotic normality of \( \hat{\beta}(t) \) is given in the following theorem and its proof is given in the Appendix A.3.

**Theorem 3.4:** (Asymptotic Normality of \( \hat{\beta}(t) \)) Under Condition (I), \( ((nh)^{1/2}(\tilde{\beta}(t) - \beta_0(t) - \beta_{Bias}(t))) \xrightarrow{D} \mathcal{N}(0, \mu_0 \Sigma(t)) \) for each fixed time point \( t \) as \( n \to \infty, h \to 0, \)
\( nh \rightarrow \infty \) and \( nh^5 = O(1) \). Here \( \mu_0 = \int_{-1}^{1} K^2(u)du, \mu_2 = \int_{-1}^{1} u^2 K^2(u)du, \)

\[
\beta_{\text{Bias}}(t) = (1/2)\mu_2 t^2 (e_{xx}(t))^{-1}[e_{xy}(t) - e_{xx}'(t) \gamma_0 - e_{xx}''(t) \beta_0(t) + 2e_{xx}'(t) \beta_0'(t) + e_{xx}(t) \beta_0''(t)],
\]

and \( \Sigma(t) \) is a positive semidefinite matrix.

Based on the equation (A.14), the covariance matrix of \( \hat{\beta}(t) \) can be estimated by

\[
n^{-2}(\hat{E}_{xx}(t))^{-1} \left[ \sum_{i=1}^{n} \left( \ll \int_{0}^{\tau} K_h(u - t) X_i(u) t_i(u) dN_i^c(u) \gg R \right) \right]^2 (\hat{E}_{xx}(t))^{-1},
\]

which is a consistent estimator based on the derivation in the Appendix A.3.

Note that

\[
(nh)^{1/2}(\hat{\beta}(t) - \beta_0(t) - \beta_{\text{Bias}}(t)) = (nh)^{1/2}(\hat{\beta}(t; \gamma_0) - \beta_0(t) - \beta_{\text{Bias}}(t)) + (nh)^{1/2}(\hat{\gamma} - \gamma_0) \frac{\partial \hat{\beta}(t; \gamma_0)}{\partial \gamma} + O_p(n^{-1/2}h^{1/2})
\]

\[
= n^{-1/2} \sum_{i=1}^{n} h^{1/2} \left[ (e_{xx}(t))^{-1} \left( R_i \int_{0}^{\tau} K_h(u - t) X_i(u) \epsilon_i(u) dN_i^c(u) \\
+ (1 - R_i) E_s \left( \int_{0}^{\tau} K_h(u - t) X_i(u) \epsilon_i(u) dN_i^c(u) \mid D_i, R_i = 0 \right) \right) \\
- D^{-1} \left( \int_{t_1}^{t_2} W(t) \{Z_i(t) - \tilde{z}_x(t) X_i(t)\} \epsilon_i(t) \mid D_i, R_i = 0 \right) \tilde{z}_x(t) \\
+ O(h^{1/2}) + O_p(n^{-1/2}h^{5/2}) + O_p(n^{-1/2}h^{1/2}).
\]

An adjusted estimation of the covariance matrix of \( \hat{\beta}(t) \) is given as

\[
n^{-2} \sum_{i=1}^{n} \left( (\hat{E}_{xx}(t))^{-1} \ll \int_{0}^{\tau} K_h(u - t) X_i(u) \epsilon_i(u) dN_i^c(u) \gg R \right) \\
- \hat{D}^{-1} \ll \int_{t_1}^{t_2} W(t) \{Z_i(t) - \tilde{Z}_x(t) X_i(t)\} \epsilon_i(t) dN_i^c(t) \gg R \tilde{Z}_x(t) \right]^2.
\]

(3.1)
A numerical study is conducted to illustrate the feasibility and validity of the proposed methods. The performances of the estimator for $\gamma$ are measured through the bias (Bias), the sample standard error of the estimates (SSE), the estimated standard error of $\hat{\gamma}$ (ESE) and the coverage probability of a 95% confidence interval for $\gamma$. The overall performance of the estimator for the $j$th component $\beta_j(\cdot)$ on the interval $[0, \tau]$ is evaluated through the square root of integrated average square error

$$ RASE(\hat{\beta}_j(\cdot)) = \left\{ \frac{1}{\tau} \int_0^\tau (\hat{\beta}_j(t) - \beta_j(t))^2 \, dt \right\}^{1/2}, $$

where $\hat{\beta}_j(t)$ is the estimate of $\beta_j(t)$. The simulation uses the unit weight function. The interval $[t_1, t_2] = [0, 15, \pi]$ is taken to be $[0, \tau]$ in the estimating functions (2.8).

The performance of the proposed estimators are examined under the following selected setting of model (2.1). Let $Y_i(t)$ follow the semiparametric additive model:

$$ Y_i(t) = \beta_0(t) + \beta_1(t)X_i + \gamma Z_i + \epsilon_i(t), \quad i = 1, \ldots, n, $$

where $\beta_0(t) = 1 - t$, $\beta_1(t) = 5 \sin(t)$, $\gamma = 8$, $X_i$ is uniformly distributed on $[0, 1]$, and $Z_i$ is a Bernoulli random variable with $P(Z_i = 1) = 0.5$. The error process $\epsilon_i(t)$ has a normal distribution with mean $\phi_i$ and variance 1 for subject $i$ where $\phi_i$ follows a standard normal distribution.

For subject $i$, $S_i$ is generated from the uniform distribution on $[0, 0.8]$. The first sampling point is set as $T_{i1} = 0$, and the rest $T_{ij}$'s are generated from a Poisson process $N_i(t)$ with the intensity rate of $\lambda_0 \exp(\eta_1 X_i + \eta_2 Z_i)$ where $\lambda_0 = 0.4$, $\eta_1 = 1$ and $\eta_2 = 0.3$. Let $Y_{ij}$ be the responses $Y_i(t)$ at time points $T_{ij}^o = T_{ij} + S_i$ following model
(4.1). The censoring time $C_i$ is exponentially distributed with the parameter adjusted to give an approximately 0% or 30% censoring in the time interval $[0, \tau] = [0, 4]$, which is the probability of $\max_{1 \leq j \leq n_i} \{T_{ij}^o \wedge \tau\} > S_i + C_i$, denoted as $c_R$. The average number of observations in the interval $[0, \tau] = [0, 4]$ per subject is about 3.48.

The following four cases, including three different left censoring percentages for $S_i$, denoted as $c_L$, and the one that ignores $S_i$ by mistreating $T_{ij}$ as the measurement times since the actual time origin, are conducted to examine the behavior of both estimators: (1) $c_L = 0\%$ which means $\{S_i\}$ are observed for all the subjects; (2) $c_L = 20\%$; (3) $c_L = 50\%$; and (4) the last case treats $T_{ij}$ as the time since the actual time origin and $Y_{ij} = Y_i(T_{ij}^o)$ as the response at $t = T_{ij}$. The censoring time $V_i$ is generated from an uniform distribution $[a, b]$ with the parameters $a$ and $b$ adjusted to yield desired percentages of left censoring for $S_i$.

The simulation presented in the following is carried out using local linear approach. As discussed in Section 2.3, to reduce the time consumption of simulations, the Epanechnikov kernel $K(u) = 0.75(1 - u^2)I(|u| \leq 1)$ is used for the inner points of time interval, i.e. $(3h, \tau - 3h)$ while the equivalent kernel in (2.9) is applied for the boundary points in $[0, 3h] \cup [\tau - 3h, \tau]$.

For sample sizes $n = 200, 300$ and $500$, and bandwidths $h = 0.3, 0.4$ and $0.5$, Table 4.1 shows the biases (Bias), the sample standard errors (SSE), the estimated standard errors (ESE) of $\gamma$, the coverage probabilities of a 95% confidence interval for $\gamma$ and also the square root of integrated average square error (RASE) of both components of $\hat{\beta}(t)$ for the first three cases based on 500 simulations when there is no right censoring. While Table 4.2 shows the same criterions for the first three cases based on 500 simulations when there is 30% of subjects right-censored during the time scale. The biases of $\hat{\gamma}$ for the first three cases using the proposed method are small. The sample standard errors of $\hat{\gamma}$ are close to its estimated standard errors. Both standard errors reduce as the sample size increases. When the left censoring
percentage of $S_i$ goes up, the standard errors rise a tiny bit since the increase of percentage means more unknown information of $S_i$. The coverage probabilities of $\hat{\gamma}$ are slightly around 0.95 as expected. The square root of integrated average square error of $\hat{\beta}_0(t)$ is smaller than that of $\hat{\beta}_1(t)$ because $\beta_0(t)$ is a straight line while $\beta_1(t)$ is more curvy. Both RASE’s increase together with the left censoring percentage of $S_i$. Furthermore, as the bandwidth $h$ changes, the $RASE(\hat{\beta}_0(\cdot))$ and $RASE(\hat{\beta}_1(\cdot))$ varies a little, which indicates that the choice of three bandwidths will not quite affect the estimates of $\beta_0(t)$ and $\beta_1(t)$.

Table 4.3 present the biases, sample standard errors, estimated standard errors and the coverage probabilities related to $\gamma$ in the case of mistreating $T_{ij}$ as the measurement times since the actual time origin. Although both the standard errors of $\hat{\gamma}$ increase compared to the third case with the same left censoring percentage, the biases are also small, the coverage probabilities are close to 0.95 and two types of standard errors are also close. This means even the time origin is mistreated, we can still get an unbiased estimator of $\gamma$ since $\gamma$ is time-independent.

Table 4.4 compare the RASE’s in the two cases when the left censoring percentage of $S_i$ is 50%. An obvious reduction of both RASE’s is shown in the table.

Figure 4.1 shows the average estimates of $\beta(t) = (\beta_0(t), \beta_1(t))^T$ based on 500 simulations under four cases proposed above. Figure 4.1 (a), (b) and (c) are the plots of the average of the estimates based on the proposed method corresponding to 0%, 20% and 50% left censoring for $S_i$, and Figure 4.1 (d) corresponds to the fourth case. Figure 4.1 (a), (b) and (c) show that the estimated curves fit the true curve quite well. There is an obvious time shift for the covariate effect of $X_i$ in Figure 4.1 (d).

Figure 4.2 shows both the standard errors of $\beta(t) = (\beta_0(t), \beta_1(t))^T$ based on 500 simulations under four cases proposed above. Figure 4.2 (a), (b) and (c) are the plots based on the proposed method corresponding to 0%, 20% and 50% left censoring for $S_i$, and Figure 4.2 (d) corresponds to the fourth case. In all four plots, the sample
standard error curves are quite close to the estimated standard error curve. In the first three cases there are big variation at the beginning time while in the fourth case there are large variation at the end of the time scale. It is related to the amount of data. According to the generation of data, for each subject the first measure is taken at $T_{ij} = 0$. Then in fourth case there are most data at the beginning while least data at the end. On the other hand, in the first three cases the time point is $T_{ij}^0 = S_i + T_{ij}$ which results in a time shift of length $S_i$. Then there are less data near the beginning and more data near $t = 4$ than in the fourth case.

Figure 4.3 shows the coverage probability of a pointwise 95% confidence interval for each component of $\beta(t) = (\beta_0(t), \beta_1(t))^T$ at each time point $t$ based on 500 simulations under four cases proposed above. Figure 4.3 (a), (b) and (c) are the plots based on the proposed method corresponding to 0%, 20% and 50% left censoring for $S_i$, and Figure 4.3 (d) corresponds to the fourth case. The dotted line in all four plots are the line when coverage probability is 95%. It is quite clear that all the coverage probabilities are close to 0.95.
Table 4.1: Summary statistics for the estimators $\hat{\gamma}$ and $\hat{\beta}(t)$ with no right censoring

<table>
<thead>
<tr>
<th>$c_L$</th>
<th>$n$</th>
<th>$h$</th>
<th>Bias</th>
<th>SSE</th>
<th>ESE</th>
<th>CP</th>
<th>RASE($\hat{\beta}_0(t)$)</th>
<th>RASE($\hat{\beta}_1(t)$)</th>
</tr>
</thead>
<tbody>
<tr>
<td>0%</td>
<td>200</td>
<td>0.3</td>
<td>-0.0090</td>
<td>0.1794</td>
<td>0.1780</td>
<td>0.958</td>
<td>0.0205</td>
<td>0.0479</td>
</tr>
<tr>
<td></td>
<td>0.4</td>
<td>-0.0082</td>
<td>0.1794</td>
<td>0.1786</td>
<td>0.948</td>
<td>0.0172</td>
<td>0.0596</td>
<td></td>
</tr>
<tr>
<td></td>
<td>0.5</td>
<td>-0.0078</td>
<td>0.1794</td>
<td>0.1790</td>
<td>0.954</td>
<td>0.0161</td>
<td>0.0858</td>
<td></td>
</tr>
<tr>
<td>300</td>
<td>0.3</td>
<td>-0.0009</td>
<td>0.1386</td>
<td>0.1450</td>
<td>0.966</td>
<td>0.0182</td>
<td>0.0500</td>
<td></td>
</tr>
<tr>
<td></td>
<td>0.4</td>
<td>0.0011</td>
<td>0.1385</td>
<td>0.1454</td>
<td>0.966</td>
<td>0.0163</td>
<td>0.0639</td>
<td></td>
</tr>
<tr>
<td></td>
<td>0.5</td>
<td>0.0013</td>
<td>0.1384</td>
<td>0.1457</td>
<td>0.968</td>
<td>0.0160</td>
<td>0.0907</td>
<td></td>
</tr>
<tr>
<td>500</td>
<td>0.3</td>
<td>-0.0083</td>
<td>0.1117</td>
<td>0.1134</td>
<td>0.950</td>
<td>0.0104</td>
<td>0.0323</td>
<td></td>
</tr>
<tr>
<td></td>
<td>0.4</td>
<td>-0.0083</td>
<td>0.1116</td>
<td>0.1136</td>
<td>0.952</td>
<td>0.0064</td>
<td>0.0445</td>
<td></td>
</tr>
<tr>
<td></td>
<td>0.5</td>
<td>-0.0081</td>
<td>0.1116</td>
<td>0.1137</td>
<td>0.950</td>
<td>0.0056</td>
<td>0.0724</td>
<td></td>
</tr>
<tr>
<td>20%</td>
<td>200</td>
<td>0.3</td>
<td>-0.0064</td>
<td>0.1809</td>
<td>0.1781</td>
<td>0.948</td>
<td>0.0279</td>
<td>0.0686</td>
</tr>
<tr>
<td></td>
<td>0.4</td>
<td>-0.0064</td>
<td>0.1808</td>
<td>0.1788</td>
<td>0.946</td>
<td>0.0256</td>
<td>0.0758</td>
<td></td>
</tr>
<tr>
<td></td>
<td>0.5</td>
<td>-0.0062</td>
<td>0.1810</td>
<td>0.1793</td>
<td>0.944</td>
<td>0.0241</td>
<td>0.0959</td>
<td></td>
</tr>
<tr>
<td>300</td>
<td>0.3</td>
<td>0.0022</td>
<td>0.1426</td>
<td>0.1450</td>
<td>0.960</td>
<td>0.0314</td>
<td>0.0772</td>
<td></td>
</tr>
<tr>
<td></td>
<td>0.4</td>
<td>0.0027</td>
<td>0.1427</td>
<td>0.1454</td>
<td>0.960</td>
<td>0.0310</td>
<td>0.0864</td>
<td></td>
</tr>
<tr>
<td></td>
<td>0.5</td>
<td>0.0033</td>
<td>0.1426</td>
<td>0.1457</td>
<td>0.960</td>
<td>0.0289</td>
<td>0.1061</td>
<td></td>
</tr>
<tr>
<td>500</td>
<td>0.3</td>
<td>-0.0059</td>
<td>0.1127</td>
<td>0.1135</td>
<td>0.942</td>
<td>0.0182</td>
<td>0.0704</td>
<td></td>
</tr>
<tr>
<td></td>
<td>0.4</td>
<td>-0.0058</td>
<td>0.1127</td>
<td>0.1137</td>
<td>0.944</td>
<td>0.0154</td>
<td>0.0759</td>
<td></td>
</tr>
<tr>
<td></td>
<td>0.5</td>
<td>-0.0057</td>
<td>0.1127</td>
<td>0.1139</td>
<td>0.944</td>
<td>0.0147</td>
<td>0.0914</td>
<td></td>
</tr>
<tr>
<td>50%</td>
<td>200</td>
<td>0.3</td>
<td>-0.0061</td>
<td>0.1821</td>
<td>0.1784</td>
<td>0.952</td>
<td>0.0905</td>
<td>0.2187</td>
</tr>
<tr>
<td></td>
<td>0.4</td>
<td>-0.0055</td>
<td>0.1822</td>
<td>0.1795</td>
<td>0.952</td>
<td>0.0897</td>
<td>0.1960</td>
<td></td>
</tr>
<tr>
<td></td>
<td>0.5</td>
<td>-0.0051</td>
<td>0.1822</td>
<td>0.1800</td>
<td>0.952</td>
<td>0.0547</td>
<td>0.1608</td>
<td></td>
</tr>
<tr>
<td>300</td>
<td>0.3</td>
<td>0.0051</td>
<td>0.1418</td>
<td>0.1451</td>
<td>0.962</td>
<td>0.0725</td>
<td>0.1798</td>
<td></td>
</tr>
<tr>
<td></td>
<td>0.4</td>
<td>0.0058</td>
<td>0.1417</td>
<td>0.1458</td>
<td>0.964</td>
<td>0.0672</td>
<td>0.1743</td>
<td></td>
</tr>
<tr>
<td></td>
<td>0.5</td>
<td>0.0060</td>
<td>0.1417</td>
<td>0.1461</td>
<td>0.962</td>
<td>0.0585</td>
<td>0.1626</td>
<td></td>
</tr>
<tr>
<td>500</td>
<td>0.3</td>
<td>-0.0050</td>
<td>0.1132</td>
<td>0.1138</td>
<td>0.942</td>
<td>0.0557</td>
<td>0.1824</td>
<td></td>
</tr>
<tr>
<td></td>
<td>0.4</td>
<td>-0.0039</td>
<td>0.1147</td>
<td>0.1145</td>
<td>0.948</td>
<td>0.0544</td>
<td>0.1711</td>
<td></td>
</tr>
<tr>
<td></td>
<td>0.5</td>
<td>-0.0041</td>
<td>0.1135</td>
<td>0.1143</td>
<td>0.942</td>
<td>0.0431</td>
<td>0.1615</td>
<td></td>
</tr>
</tbody>
</table>
Table 4.2: Summary statistics for the estimators $\hat{\gamma}$ and $\hat{\beta}(t)$ with 30% right censoring rate

<table>
<thead>
<tr>
<th>$c_L$</th>
<th>$n$</th>
<th>$h$</th>
<th>Bias</th>
<th>SSE</th>
<th>ESE</th>
<th>CP</th>
<th>RASE($\beta_0(t)$)</th>
<th>RASE($\beta_1(t)$)</th>
</tr>
</thead>
<tbody>
<tr>
<td>0%</td>
<td>200</td>
<td>0.3</td>
<td>-0.0131</td>
<td>0.1871</td>
<td>0.1836</td>
<td>0.946</td>
<td>0.0213</td>
<td>0.0479</td>
</tr>
<tr>
<td></td>
<td></td>
<td>0.4</td>
<td>-0.0121</td>
<td>0.1873</td>
<td>0.1843</td>
<td>0.946</td>
<td>0.0179</td>
<td>0.0569</td>
</tr>
<tr>
<td></td>
<td></td>
<td>0.5</td>
<td>-0.0113</td>
<td>0.1872</td>
<td>0.1848</td>
<td>0.950</td>
<td>0.0172</td>
<td>0.0823</td>
</tr>
<tr>
<td>300</td>
<td>0.3</td>
<td>-0.0011</td>
<td>0.1436</td>
<td>0.1500</td>
<td>0.962</td>
<td>0.0243</td>
<td>0.0548</td>
<td></td>
</tr>
<tr>
<td></td>
<td>0.4</td>
<td>-0.0009</td>
<td>0.1434</td>
<td>0.1504</td>
<td>0.968</td>
<td>0.0226</td>
<td>0.0667</td>
<td></td>
</tr>
<tr>
<td></td>
<td>0.5</td>
<td>-0.0006</td>
<td>0.1432</td>
<td>0.1507</td>
<td>0.968</td>
<td>0.0223</td>
<td>0.0921</td>
<td></td>
</tr>
<tr>
<td>500</td>
<td>0.3</td>
<td>-0.0092</td>
<td>0.1154</td>
<td>0.1173</td>
<td>0.948</td>
<td>0.0123</td>
<td>0.0334</td>
<td></td>
</tr>
<tr>
<td></td>
<td>0.4</td>
<td>-0.0092</td>
<td>0.1152</td>
<td>0.1175</td>
<td>0.946</td>
<td>0.0076</td>
<td>0.0415</td>
<td></td>
</tr>
<tr>
<td></td>
<td>0.5</td>
<td>-0.0089</td>
<td>0.1152</td>
<td>0.1177</td>
<td>0.944</td>
<td>0.0066</td>
<td>0.0677</td>
<td></td>
</tr>
<tr>
<td>20%</td>
<td>200</td>
<td>0.3</td>
<td>-0.0084</td>
<td>0.1874</td>
<td>0.1835</td>
<td>0.944</td>
<td>0.0330</td>
<td>0.0745</td>
</tr>
<tr>
<td></td>
<td>0.4</td>
<td>-0.0085</td>
<td>0.1875</td>
<td>0.1844</td>
<td>0.950</td>
<td>0.0306</td>
<td>0.0784</td>
<td></td>
</tr>
<tr>
<td></td>
<td>0.5</td>
<td>-0.0083</td>
<td>0.1879</td>
<td>0.1850</td>
<td>0.952</td>
<td>0.0290</td>
<td>0.0962</td>
<td></td>
</tr>
<tr>
<td>300</td>
<td>0.3</td>
<td>0.0015</td>
<td>0.1468</td>
<td>0.1500</td>
<td>0.960</td>
<td>0.0376</td>
<td>0.0796</td>
<td></td>
</tr>
<tr>
<td></td>
<td>0.4</td>
<td>0.0019</td>
<td>0.1469</td>
<td>0.1504</td>
<td>0.962</td>
<td>0.0381</td>
<td>0.0874</td>
<td></td>
</tr>
<tr>
<td></td>
<td>0.5</td>
<td>0.0024</td>
<td>0.1470</td>
<td>0.1507</td>
<td>0.962</td>
<td>0.0362</td>
<td>0.1065</td>
<td></td>
</tr>
<tr>
<td>500</td>
<td>0.3</td>
<td>-0.0066</td>
<td>0.1160</td>
<td>0.1174</td>
<td>0.942</td>
<td>0.0181</td>
<td>0.0773</td>
<td></td>
</tr>
<tr>
<td></td>
<td>0.4</td>
<td>-0.0064</td>
<td>0.1158</td>
<td>0.1176</td>
<td>0.942</td>
<td>0.0148</td>
<td>0.0806</td>
<td></td>
</tr>
<tr>
<td></td>
<td>0.5</td>
<td>-0.0063</td>
<td>0.1157</td>
<td>0.1178</td>
<td>0.942</td>
<td>0.0152</td>
<td>0.0924</td>
<td></td>
</tr>
<tr>
<td>50%</td>
<td>200</td>
<td>0.3</td>
<td>-0.0081</td>
<td>0.1897</td>
<td>0.1835</td>
<td>0.950</td>
<td>0.0921</td>
<td>0.2330</td>
</tr>
<tr>
<td></td>
<td>0.4</td>
<td>-0.0077</td>
<td>0.1897</td>
<td>0.1847</td>
<td>0.952</td>
<td>0.0873</td>
<td>0.2065</td>
<td></td>
</tr>
<tr>
<td></td>
<td>0.5</td>
<td>-0.0072</td>
<td>0.1898</td>
<td>0.1854</td>
<td>0.952</td>
<td>0.0565</td>
<td>0.1688</td>
<td></td>
</tr>
<tr>
<td>300</td>
<td>0.3</td>
<td>0.0042</td>
<td>0.1467</td>
<td>0.1500</td>
<td>0.962</td>
<td>0.0773</td>
<td>0.1844</td>
<td></td>
</tr>
<tr>
<td></td>
<td>0.4</td>
<td>0.0047</td>
<td>0.1468</td>
<td>0.1507</td>
<td>0.960</td>
<td>0.0736</td>
<td>0.1789</td>
<td></td>
</tr>
<tr>
<td></td>
<td>0.5</td>
<td>0.0052</td>
<td>0.1467</td>
<td>0.1512</td>
<td>0.962</td>
<td>0.0653</td>
<td>0.1672</td>
<td></td>
</tr>
<tr>
<td>500</td>
<td>0.3</td>
<td>-0.0057</td>
<td>0.1164</td>
<td>0.1175</td>
<td>0.942</td>
<td>0.0557</td>
<td>0.1935</td>
<td></td>
</tr>
<tr>
<td></td>
<td>0.4</td>
<td>-0.0051</td>
<td>0.1163</td>
<td>0.1179</td>
<td>0.944</td>
<td>0.0491</td>
<td>0.1852</td>
<td></td>
</tr>
<tr>
<td></td>
<td>0.5</td>
<td>-0.0049</td>
<td>0.1163</td>
<td>0.1182</td>
<td>0.942</td>
<td>0.0442</td>
<td>0.1683</td>
<td></td>
</tr>
</tbody>
</table>
Table 4.3: Summary statistics for the estimator $\hat{\gamma}$ with misplaced time origin and 50% left censoring in the presence ($c_R = 30\%$) and absence ($c_R = 0\%$) of right censoring

<table>
<thead>
<tr>
<th>$c_R$</th>
<th>$n$</th>
<th>$h$</th>
<th>Bias</th>
<th>SSE</th>
<th>ESE</th>
<th>CP</th>
</tr>
</thead>
<tbody>
<tr>
<td>0%</td>
<td>200</td>
<td>0.3</td>
<td>-0.0016</td>
<td>0.2126</td>
<td>0.2119</td>
<td>0.946</td>
</tr>
<tr>
<td></td>
<td>0.4</td>
<td>-0.0005</td>
<td>0.2122</td>
<td>0.2127</td>
<td>0.950</td>
<td></td>
</tr>
<tr>
<td></td>
<td>0.5</td>
<td>0.0006</td>
<td>0.2121</td>
<td>0.2134</td>
<td>0.946</td>
<td></td>
</tr>
<tr>
<td>300</td>
<td></td>
<td></td>
<td>0.0019</td>
<td>0.1746</td>
<td>0.1733</td>
<td>0.944</td>
</tr>
<tr>
<td></td>
<td>0.4</td>
<td>0.0026</td>
<td>0.1746</td>
<td>0.1738</td>
<td>0.944</td>
<td></td>
</tr>
<tr>
<td></td>
<td>0.5</td>
<td>0.0033</td>
<td>0.1745</td>
<td>0.1742</td>
<td>0.948</td>
<td></td>
</tr>
<tr>
<td>500</td>
<td></td>
<td>0.3</td>
<td>-0.0066</td>
<td>0.1410</td>
<td>0.1349</td>
<td>0.932</td>
</tr>
<tr>
<td></td>
<td>0.4</td>
<td>-0.0058</td>
<td>0.1407</td>
<td>0.1352</td>
<td>0.932</td>
<td></td>
</tr>
<tr>
<td></td>
<td>0.5</td>
<td>-0.0052</td>
<td>0.1404</td>
<td>0.1354</td>
<td>0.932</td>
<td></td>
</tr>
<tr>
<td>30%</td>
<td>200</td>
<td>0.3</td>
<td>0.0003</td>
<td>0.2251</td>
<td>0.2262</td>
<td>0.946</td>
</tr>
<tr>
<td></td>
<td>0.4</td>
<td>0.0014</td>
<td>0.2251</td>
<td>0.2272</td>
<td>0.946</td>
<td></td>
</tr>
<tr>
<td></td>
<td>0.5</td>
<td>0.0027</td>
<td>0.2250</td>
<td>0.2281</td>
<td>0.946</td>
<td></td>
</tr>
<tr>
<td>300</td>
<td></td>
<td>0.3</td>
<td>0.0019</td>
<td>0.1853</td>
<td>0.1865</td>
<td>0.938</td>
</tr>
<tr>
<td></td>
<td>0.4</td>
<td>0.0029</td>
<td>0.1848</td>
<td>0.1871</td>
<td>0.940</td>
<td></td>
</tr>
<tr>
<td></td>
<td>0.5</td>
<td>0.0040</td>
<td>0.1845</td>
<td>0.1876</td>
<td>0.946</td>
<td></td>
</tr>
<tr>
<td>500</td>
<td></td>
<td>0.3</td>
<td>-0.0106</td>
<td>0.1487</td>
<td>0.1449</td>
<td>0.936</td>
</tr>
<tr>
<td></td>
<td>0.4</td>
<td>-0.0096</td>
<td>0.1486</td>
<td>0.1452</td>
<td>0.936</td>
<td></td>
</tr>
<tr>
<td></td>
<td>0.5</td>
<td>-0.0086</td>
<td>0.1480</td>
<td>0.1455</td>
<td>0.942</td>
<td></td>
</tr>
</tbody>
</table>
Table 4.4: Summary statistics for the estimator $\hat{\beta}(t)$ for misplaced time origin and 50% left censoring in the presence ($c_R = 30\%$) and absence ($c_R = 0\%$) of right censoring

<table>
<thead>
<tr>
<th>$c_R$</th>
<th>$n$</th>
<th>$h$</th>
<th>Our method</th>
<th>Misplaced time origin</th>
<th>Our method</th>
<th>Misplaced time origin</th>
</tr>
</thead>
<tbody>
<tr>
<td>0%</td>
<td>200</td>
<td>0.3</td>
<td>0.0905</td>
<td>0.3959</td>
<td>0.2187</td>
<td>1.4308</td>
</tr>
<tr>
<td></td>
<td></td>
<td>0.4</td>
<td>0.0897</td>
<td>0.3979</td>
<td>0.1960</td>
<td>1.4260</td>
</tr>
<tr>
<td></td>
<td></td>
<td>0.5</td>
<td>0.0547</td>
<td>0.3991</td>
<td>0.1608</td>
<td>1.4229</td>
</tr>
<tr>
<td>300</td>
<td>0.3</td>
<td>0.0725</td>
<td>0.3860</td>
<td>0.1798</td>
<td>1.4345</td>
<td></td>
</tr>
<tr>
<td></td>
<td>0.4</td>
<td>0.0672</td>
<td>0.3870</td>
<td>0.1743</td>
<td>1.4309</td>
<td></td>
</tr>
<tr>
<td></td>
<td>0.5</td>
<td>0.0585</td>
<td>0.3871</td>
<td>0.1626</td>
<td>1.4284</td>
<td></td>
</tr>
<tr>
<td>500</td>
<td>0.3</td>
<td>0.0557</td>
<td>0.4050</td>
<td>0.1824</td>
<td>1.4057</td>
<td></td>
</tr>
<tr>
<td></td>
<td>0.4</td>
<td>0.0544</td>
<td>0.4063</td>
<td>0.1711</td>
<td>1.4024</td>
<td></td>
</tr>
<tr>
<td></td>
<td>0.5</td>
<td>0.0431</td>
<td>0.4070</td>
<td>0.1615</td>
<td>1.3995</td>
<td></td>
</tr>
<tr>
<td>30%</td>
<td>200</td>
<td>0.3</td>
<td>0.0921</td>
<td>0.4009</td>
<td>0.2330</td>
<td>1.4222</td>
</tr>
<tr>
<td></td>
<td></td>
<td>0.4</td>
<td>0.0873</td>
<td>0.4007</td>
<td>0.2065</td>
<td>1.4099</td>
</tr>
<tr>
<td></td>
<td></td>
<td>0.5</td>
<td>0.0565</td>
<td>0.4044</td>
<td>0.1688</td>
<td>1.4035</td>
</tr>
<tr>
<td>300</td>
<td>0.3</td>
<td>0.0773</td>
<td>0.3799</td>
<td>0.1844</td>
<td>1.4373</td>
<td></td>
</tr>
<tr>
<td></td>
<td>0.4</td>
<td>0.0736</td>
<td>0.3805</td>
<td>0.1789</td>
<td>1.4329</td>
<td></td>
</tr>
<tr>
<td></td>
<td>0.5</td>
<td>0.0653</td>
<td>0.3812</td>
<td>0.1672</td>
<td>1.4305</td>
<td></td>
</tr>
<tr>
<td>500</td>
<td>0.3</td>
<td>0.0557</td>
<td>0.4061</td>
<td>0.1935</td>
<td>1.3985</td>
<td></td>
</tr>
<tr>
<td></td>
<td>0.4</td>
<td>0.0491</td>
<td>0.4071</td>
<td>0.1852</td>
<td>1.3940</td>
<td></td>
</tr>
<tr>
<td></td>
<td>0.5</td>
<td>0.0442</td>
<td>0.4077</td>
<td>0.1683</td>
<td>1.3917</td>
<td></td>
</tr>
</tbody>
</table>
Figure 4.1: The averages of the estimator $\beta(t) = (\beta_0(t), \beta_1(t))^T$ for $n = 300$, $h = 0.4$ and 30% right censoring rate. The solid lines are for $\beta_1(t)$ and the dashed lines are for $\beta_0(t)$. The grey lines are the true curves. Figure (a), (b) and (c) shows the averages in the case of 0%, 20% and 50% left censoring rate of $S_i$ respectively. Figure (d) shows the results in the case of misplaced time origin by ignoring $S_i$. 
Figure 4.2: The sample and estimated standard errors of the estimator $\beta(t) = (\beta_0(t), \beta_1(t))^T$ for $n = 300$, $h = 0.4$ and 30% right censoring rate. The solid lines are for $\beta_1(t)$ and the dashed lines are for $\beta_0(t)$. The grey lines are the estimated standard error and the black ones are the sample standard error. Figure (a), (b) and (c) shows the results in the case of 0%, 20% and 50% left censoring rate of $S_i$ respectively. Figure (d) shows the results in the case of misplaced time origin by ignoring $S_i$. 

Figure (a)

Figure (b)

Figure (c)

Figure (d)
Figure 4.3: The coverage probabilities of 95% pointwise confidence intervals of $\beta(t) = (\beta_0(t), \beta_1(t))^T$ for $n = 300$, $h = 0.4$ and 30% right censoring rate. The solid lines are for $\beta_1(t)$ and the dashed lines are for $\beta_0(t)$. Figure (a), (b) and (c) shows the averages in the case of 0%, 20% and 50% left censoring rate of $S_i$ respectively. Figure (d) shows the results in the case of misplaced time origin by ignoring $S_i$. 
In this chapter a real data from the STEP study (cf., Buchbinder et al., 2008; Fitzgerald et al., 2011) is analyzed by applying the methods discussed in previous chapters. The step study was a multicenter, double-blind, randomized, placebo-controlled, phase II test-of-concept study to determine whether the MRKAd5 HIV-1 gag/pol/nef vaccine, which elicits T cell immunity, is capable to result in controlling the replication of the Human immunodeficiency virus among the participants who got HIV-infected after vaccination. This study opened in December 2004 and was conducted at 34 sites in North America, the Caribbean, South America, and Australia. Three thousand HIV-1 negative participants aged from 18 to 45 who were at high risk of HIV-infection were enrolled and randomly assigned to receive vaccine or placebo in ratio 1:1, stratified by sex, study site if North America or Australia and 0 otherwise) and adenovirus type 5 (Ad5) antibody titer at baseline. Some of the participant were fully adherent to vaccinations while others not.

The analysis in this chapter includes a subset of the 3000 participants which involves all 174 MITT cases as of September 22, 2009. MITT cases stand for modified intention-to-treat subjects who became HIV infected during the trial. The modified intention to treat refers to all randomized subjects, excluding the few that were found to be HIV infected at entry. It is recommended to study males only, for the entire analysis to avoid the effect of sex since there are only 15 females that are < 10% of the sample. There were 159 HIV-infected males. Each participant had the records of the first positive diagnosis (the dates of their first positive Elisa confirmed by Western Blot or RNA, illustrated as $D_i$'s in Figure 1.1), the dates of their first evidence of 
infection (determined by the dates of the first positive RNA (PCR) test), and the estimated dates of infection. The estimated dates of infection is considered as the midpoint between last RNA negative visit date ($L_i$’s in Figure 1.1) and the date of first evidence of infection. The last RNA negative visit date can be computed by the estimated date of infection and the dates of their first evidence of infection. As such, we calculate $V_i$ by $V_i = D_i - L_i = D_i - 2 \times \text{estimated infection dates} + \text{the date of first evidence of infection}$. The indicator of whether the actual acquisition of $i$th subject is observed or not is denoted by $R_i$; $R_i = 1$ if the actual HIV acquisition date can be determined by using the RNA test, and in this case the duration between actual HIV acquisition and the first positive diagnosis date $S_i = V_i$; otherwise $R_i = 0$ and $S_i < V_i$.

After the participant was infected, there were 18 scheduled post-infection visit per subject at weeks 0, 1, 2, 8, 12, 26, and every 26 weeks thereafter through week 338. However, the actual times and dates of visits may vary due to each individual. During $j$th visit, the $i$th subject received tests to have the measurements of HIV virus load and CD4 cell counts before the subject started the antiretroviral therapy (ART) or was censored. And the time from the first positive Elisa to the $j$th visit for $i$th subject is $T_{ij}$ in the above chapters. The time between the first positive Elisa and ART initiation or censoring is the right censoring time. In the analysis time is measured in years. Let $Y$ be the common logarithm of HIV virus load, $X_1$ be the square root of CD4 counts, $X_2$ be the treatment indicator ($X_2 = 1$ if the subject received vaccine and 0 if receiving placebo), $Z_1$ be the site indicator ($Z_1 = 1$ if North America or Australia and 0 otherwise), $Z_2$ be the natural logarithm of Ad5 and $Z_3$ be the pre-protocol indicator ($Z_3 = 1$ if the subject was fully adherent to vaccinations and 0 otherwise). Our main interest is to see the effect of vaccine on the HIV virus load response.

In the data 159 males made a total of 791 pre-ART visits. Among them there
are 156 missing in CD4 cell counts and 5 missing in HIV virus load. Since there are no missing in CD4 and virus load at the same time, we could use a simple imputation model to create a complete data set. At each time point separately, we use a linear regression model linking log_{10}(viral load) to square root of CD4 count (for those with data on both), and use the viral load value for those with missing data to fill in the missing CD4 cell count or predict missing virus load data by CD4 values. However, at three time points there are no complete data for conducting the linear regression model fitting; at two other points there are only one complete data which is unable to complete the linear model fitting; at another time point one predicted value of virus load is relatively far beyond the range of other values of virus load and may affect the analysis results. Therefore, we delete these six visits to get the complete data for the entire analysis.

Now in this complete data set there are 159 subjects with 785 visits. 97 Of all the participants were in the vaccine group while 62 received the placebo. 122 subjects participate in the study in North America or Australia and the rest are residents in the other sites mentioned at the beginning of this chapter. The left censoring rate of \( S_i \) is 70.44\% and the right censoring rate of \( T_{ij} \) is 69.81\%. Figure 5.1 to Figure 5.3 are further exploration of the data. It is easy to figure out that there are few data after time point 2.5. Therefore, we will choose \( t_1 = 0 \) and \( t_2 = 2.5 \) to estimate \( \gamma \), and also plot the estimators of \( \beta(t) \)'s for the time points in the interval \([0, 2.5]\). Finally, Figure 5.4 shows the Kaplan Meier estimator of the distribution of \( S_i \). Note that the smallest observed \( S_i \) is 0.14. Before that time we do not have enough information to get the estimator of the distribution. However, since time is always nonnegative, the probability of \( S_i \) reduce to 0 at \( S_i = 0 \).

After preliminary exploration of the data, we propose the following model for virus load response of the \( i \)th subject in this study:

\[
Y_i(t) = \beta_0(t) + \beta_1(t)X_{1i}(t) + \beta_2(t)X_{2i} + \gamma_1Z_{1i} + \gamma_2Z_{2i} + \gamma_3Z_{3i} + \epsilon_i(t). \quad (5.1)
\]
By the study of simulation and several tries of different bandwidths, a possible reasonable choice of the bandwidth for this data set is 0.5. And we still consider the unit weight for the analysis. The estimates of $\gamma_1$, $\gamma_2$ and $\gamma_3$ are 0.0302, $-0.1467$ and 0.1956, with the standard deviations 0.0389, 0.1492 and 0.1540, respectively. The $p$-values for testing $H_0 : \gamma_1 = 0$, $H_0 : \gamma_2 = 0$ and $H_0 : \gamma_3 = 0$ are equal to 0.4375, 0.3255 and 0.2042, respectively, which indicates that there are no significant effects of baseline Ad5 titer, study sites or the pre-protocol on the HIV viral load level. The estimates of time-dependent effects and their 95% pointwise confidence intervals are shown in Figure 5.5. From the graph the effects of vaccine or CD4 cell count on the HIV viral load level are not significant, either. Further hypothesis test study will be done in the future. Finally Figure 5.6 shows the scatter plot of the residuals from fitting the model (5.1).
Figure 5.1: Histogram of times ($T_{ij}$) from the first positive Elisa confirmed by Western Blot or RNA to subsequent visits.
Figure 5.2: Histograms of times \( (S_i) \) from actual HIV acquisition to the first positive Elisa confirmed by Western Blot or RNA. Figure (a) shows the histogram of observed \( S_i \)'s \( (R_i = 1) \) while figure (b) shows the counts of censored \( S_i \)'s \( (R_i = 0) \).
Figure 5.3: Histograms of times ($C_i$) from the first positive Elisa confirmed by Western Blot or RNA to ART initiation or censoring.
Figure 5.4: The Kaplan Meier estimator of the distribution function of the time from actual HIV acquisition to the first positive Elisa confirmed by Western Blot or RNA.
Figure 5.5: Estimation of $\beta(t) = (\beta_0(t), \beta_1(t), \beta_2(t))^T$ based on the data from STEP study with MITT cases. Figure (a) shows the estimated intercept, $\beta_0(t)$ and its 95% pointwise confidence interval. Figure (b) shows the estimated effect of the square root of CD4 effect, $\beta_1(t)$ and its 95% pointwise confidence interval. Figure (c) shows the estimated treatment effect, $\beta_2(t)$ and its 95% pointwise confidence interval. The solid curves are the estimated curves and the dashed curves are the confidence intervals.
Figure 5.6: The scatter plot of residuals of the subjects with $R_i = 1$. 
REFERENCES


HVTN. 2004. The pipeline project. (HVTN stands for HIV Vaccine Trials Network.) Available at: http://www.hvtn.org/.


Now we will show the detailed proofs of five lemmas and four theorems we present in Chapter 3. In Section A.2, Lemma A.2.1 is used to prove Lemma A.2.2. The results of Lemma A.2.2 and Lemma A.2.3 states the consistent properties of our proposed notation $\ll R$. Lemma A.2.4 is the basis of getting Lemma A.2.5. We will repeatedly apply Lemmas A.2.2, A.2.3 and A.2.5 in proofs of theorems in Section A.3.

A.1 Preliminaries

Preparing for future application in this section, we first derive the martingale decomposition of the Kaplan-Meier estimator of the survival function for the left censored data.

In general, we have the i.i.d. data structure of the left censored data as follows,

$$\{T_i = \max(S_i, C_i), \delta_i = I(S_i \geq C_i)\},$$

where $S_i$ is the failure time censored by $C_i$, $T_i$ is observed time and $\delta_i$ is the indicator of non-censorship for $i$th subject. Suppose $L$ be a large enough number so that all $S_i < L$. Then

$$\{L - T_i = \min(L - S_i, L - C_i), \delta_i = I(L - S_i \leq L - C_i)\}$$

is the corresponding right censored data structure. Let $S(t) = P(S_i > t)$ and $S^R(t) = P(L - S_i > t)$ be the survival functions of the failure time for the left and right censored data respectively. And $\hat{S}(t)$, $\hat{S}^R(t)$ are the Kaplan-Meyer estimators of the survival functions respectively. Now define the counting process $N^R_i(t) = I(L - T_i \leq t, \delta_i = 1)$. 

Martingale

For the left censored data, by the Doob-Meyer decomposition, there is a compensator \( \int_0^t Y_i^R(s) d\Lambda^R(s) \) and a martingale \( M_i^R(t) \) so that \( N_i^R(t) = \int_0^t Y_i^R(s) d\Lambda^R(s) + M_i^R(t) \). Here \( Y_i^R(t) = I(L-T_i \geq t) \) is the at risk indicator and \( \Lambda^R(t) \) is the cumulative hazard function. Let \( N^R(t) = \sum_{i=1}^n N_i^R(t) \), \( M^R(t) = \sum_{i=1}^n M_i^R(t) \) and \( Y^R(t) = \sum_{i=1}^n Y_i^R(t) = \sum_{i=1}^n I(T_i \leq L - t) \). Assume that \( Y^R(t)/n \xrightarrow{P} Y(t) \). Hence according to Equation (2.11) in Chapter 3 on Page 98 of Fleming & Harrington (1991), we have the decomposition

\[
n^{1/2}(\hat{S}^R(t) - S^R(t)) = -n^{1/2}S^R(t) \int_0^t \frac{\hat{S}^R(s-)}{S^R(s)} \frac{I(Y^R(s) > 0)}{Y^R(s)} dM^R(s) + o_p(1).
\]

Since

\[
S(t) = P(S_i > t) = P(L - S_i < L - t) = 1 - P(L - S_i \geq L - t) = 1 - S^R((L - t)^-),
\]

then for the left censored data

\[
n^{1/2}(\hat{S}(t) - S(t))
= -n^{1/2}[\hat{S}^R((L - t)^-) - S^R((L - t)^-)]
= n^{1/2}S^R((L - t)^-) \int_0^{(L-t)^-} \frac{\hat{S}^R(s-)}{S^R(s)} \frac{I(Y^R(s) > 0)}{Y^R(s)} dM^R(s) + o_p(1)
= n^{-1/2}(1 - S(t)) \int_0^{(L-t)^-} \frac{1 - \hat{S}(L - s)}{1 - S((L - s)^-)} \frac{I(Y^R(s) > 0)}{Y^R(s)} dM^R(s) + o_p(1)
= n^{-1/2}(1 - S(t)) \int_0^{(L-t)^-} \frac{1}{y^R(s)} dM^R(s) + o_p(1).
\]

(A.1)

Now let us define the following notations for the future use.

\[
X_{x_1}^I(u) = \int_0^u [R_i Z_i(w) X_i^T(w) dN_i^c(w) - E(R_i E_i(w) \alpha_i(w) Z_i(w) X_i^T(w)) dw],
\]

\[
X_{x_1}^H(t) = \int_0^\infty \int_0^L \int_{t_1}^t E \left\{ (1 - R_i) Z_i(u) X_i^T(u) \alpha_i^*(u - s) \frac{I(x \leq (L - V_i)^-)}{F_s(V_i)} \right\}
\cdot (e_{xx}(u))^{-1} du dF_s(s) \frac{dM^R(x)}{y^R(x)} - \int_0^L \int_0^{(L - x)^-} \int_{t}^t E \left\{ (1 - R_i) Z_i(u) X_i^T(u) \alpha_i^*(u - s) \right\}
\cdot (e_{xx}(u))^{-1} du dF_s(s) \frac{dM^R(x)}{y^R(x)}.
\]
Let a random function $X$ and $t$

However

$$X^{II}(u) = \int_0^u (E_s\{(1 - R_i)Z_i(w)X_i^T(w)dN_i^c(w) | D_i, R_i = 0\} - E\{(1 - R_i)\xi_i(w)\alpha_i(w)Z_i(w)X_i^T(w)\}dw),$$

and

$$X^{II}(u) = n^{-1/2} \sum_{i=1}^n X^{II}_z(u), X^{II}(t) = n^{-1/2} \sum_{i=1}^n X^{II}_z(t), X^{II}_z(u) = n^{-1/2} \sum_{i=1}^n X^{II}_z(u).$$

Similarly, we can define $X^{I}_y(u), X^{II}_y(t), X^{II}_y(u), X^{I}_y(u), X^{II}_y(u), X^{I}_y(u), X^{II}_z(u), X^{I}_z(u), X^{II}_z(u)$ by replacing $Z_i(\cdot)$ above with $Y_i(\cdot)$ and $X_i(\cdot)$ respectively.

However

$$X^{II}_z(t) = \int_0^L \int_{t_1}^t \beta^T(u)E\{(1 - R_i)X_i(u)X_i^T(u)\alpha_i^{*}(u - s)\frac{I(x \leq (L - V_i) -)}{F_s(V_i)}\}$$

$$\cdot (e_{xx}(u))^{-1}du dF_s(s) \frac{dM_i^R(x)}{y^R(x)}$$

$$- \int_0^L \int_{t_1}^{(L-x)-} \int_{t_1}^t \beta^T(u)E\{(1 - R_i)X_i(u)X_i^T(u)\alpha_i^{*}(u - s)\}$$

$$\cdot (e_{xx}(u))^{-1}du dF_s(s) \frac{dM_i^R(x)}{y^R(x)}$$

$$+ \int_{t_1}^L \int_{t_1}^t \beta^T(u)E\{(1 - R_i)X_i(u)X_i^T(u)\alpha_i^{*}(u - s)\}$$

$$\cdot (e_{xx}(u))^{-1}du dF_s(s) \frac{dM_i^R((L - s) -)}{y^R((L - s) -)}.$$ 

Then $X^{II}_z(t) = n^{-1/2} \sum_{i=1}^n X^{II}_z(t).$

A.2 Some Lemmas

**Lemma A.2.1:** Let a random function $g_i(t) = g(t, X_i(t), Z_i(t), Y_i(t))$. Then under Conditions (I), for $t \in [t_1, t_2] \subset [0, \tau]$,

$$n^{-1} \sum_{i=1}^n (1 - R_i) \hat{E}_s\{\int_{t_1}^{t_2} g_i(u)dN_i^c(u) | D_i, R_i = 0\} \xrightarrow{P} E\{(1 - R_i) \int_{t_1}^{t_2} g_i(u)dN_i^c(u)\}$$
as \( n \to \infty \).

**Proof.** As mentioned in Section 2.2,

\[
n^{-1} \sum_{i=1}^{n} (1 - R_i) \bar{E}_s \left\{ \int_{t_1}^{t_2} g_i(u) dN_i^c(u) \mid D_i, R_i = 0 \right\}
\]

\[= n^{-1} \sum_{i=1}^{n} (1 - R_i) \int_0^L \sum_{j=1}^{n_i} g_i(s + T_{ij}) I(C_i \geq T_{ij}) \frac{d\hat{F}_s(s \mid D_i)}{\hat{F}_s(V_i \mid D_i)}
\]

\[= n^{-1} \sum_{i=1}^{n} (1 - R_i) \int_0^L \sum_{j=1}^{n_i} g_i(s + T_{ij}) I(C_i \geq T_{ij}) \frac{d\hat{F}_s(s \mid D_i)}{\hat{F}_s(V_i \mid D_i)}
\]

\[+ n^{-1} \sum_{i=1}^{n} (1 - R_i) \int_0^L \sum_{j=1}^{n_i} g_i(s + T_{ij}) I(C_i \geq T_{ij}) \left( \frac{d\hat{F}_s(s \mid D_i)}{\hat{F}_s(V_i \mid D_i)} - \frac{dF_s(s \mid D_i)}{F_s(V_i \mid D_i)} \right)
\]

\[+ n^{-1} \sum_{i=1}^{n} (1 - R_i) \int_0^L \sum_{j=1}^{n_i} g_i(s + T_{ij}) I(C_i \geq T_{ij}) \frac{d\hat{F}_s(s \mid D_i) - dF_s(s \mid D_i)}{\hat{F}_s(V_i \mid D_i)}
\]

(A.2)

If \( \hat{F}_s(s \mid D_i) \) is the Kaplan-Meier estimator of conditional survival function, we still have \( \hat{F}_s(s \mid D_i) \xrightarrow{P} F_s(s \mid D_i) \), \( \hat{F}_s(V_i \mid D_i) \xrightarrow{P} F_s(V_i \mid D_i) \). Then by continuous theorem, \( 1/\hat{F}_s(V_i \mid D_i) \xrightarrow{P} 1/F_s(V_i \mid D_i) \). So under the Conditions (I) the second term in (A.2) which is equal to

\[n^{-1} \sum_{i=1}^{n} (1 - R_i) \int_0^L \sum_{j=1}^{n_i} g_i(s + T_{ij}) I(C_i \geq T_{ij}) \left( \frac{1}{\hat{F}_s(V_i \mid D_i)} - \frac{1}{F_s(V_i \mid D_i)} \right) dF_s(s \mid D_i)
\]

converges to zero in probability. Since \( S_i \) is independent of \( D_i \) and remind that \( N_i(t) = \sum_{j=1}^{n_i} I(T_{ij} \leq t) \), the third term in (A.2) is equal to

\[n^{-1} \sum_{i=1}^{n} (1 - R_i) \int_0^L \sum_{j=1}^{n_i} g_i(s + T_{ij}) I(C_i \geq T_{ij}) \frac{d\hat{F}_s(s) - dF_s(s)}{F_s(V_i \mid D_i)} + o_p(1)
\]

\[= \int_0^L \left[ n^{-1} \sum_{i=1}^{n} (1 - R_i) \int_{t_1 - s}^{t_2 - s} g_i(s + v) I(C_i \geq v) dN_i(v) \right] \frac{d(\hat{F}_s(s) - F_s(s))}{F_s(V_i \mid D_i)}
\]

\[+ o_p(1)
\]

\[= \int_0^L \left[ n^{-1} \sum_{i=1}^{n} (1 - R_i) \int_{t_1 - s}^{t_2 - s} g_i(s + v) I(C_i \geq v) dN_i(v) \right] \frac{1}{F_s(V_i \mid D_i)}
\]

\[d(\hat{F}_s(s) - F_s(s)) + o_p(1)
\]
Let
\[ H_n(s) = n^{-1} \sum_{i=1}^{n} (1 - R_i) \int_{t_{1-s}}^{t_2-s} g_i(s + v) I(C_i \geq v) dN_i(v) \frac{1}{F_s(V_i \mid D_i)}. \]

So the absolute value of the third term in (A.2) equals
\[
\left| \int_0^L H_n(s) d(\hat{F}_s(s) - F_s(s)) \right|
= \left| H_n(L)(\hat{F}_s(L) - F_s(L)) - H_n(0)(\hat{F}_s(0) - F_s(0)) - \int_0^L (\hat{F}_s(s) - F_s(s)) dH_n(s) \right|
\leq |H_n(L)(\hat{F}_s(L) - F_s(L))| + |H_n(0)(\hat{F}_s(0) - F_s(0))| + \left| \int_0^L (\hat{F}_s(s) - F_s(s)) dH_n(s) \right|
\leq |H_n(L)(\hat{F}_s(L) - F_s(L))| + |H_n(0)(\hat{F}_s(0) - F_s(0))|
+ \sup_{s \in [0,L]} |\hat{F}_s(s) - F_s(s)| \int_0^L |dH_n(s)|
\]

Under Conditions (I), by the uniform consistency of \(\hat{F}_s(s)\) and the convergence of \(\hat{F}_s(s)\) at point \(s = 0\), or \(s = L\), the third term converges to zero in probability uniformly in \(s\) as \(n \to \infty\). Therefore,
\[
(A.2) \quad \xrightarrow{P} \quad E\left\{ (1 - R_i) \int_0^L \sum_{j=1}^{n_i} g_i(s + T_{ij}) I(C_i \geq T_{ij}) \frac{dF_s(s \mid D_i)}{F_s(V_i \mid D_i)} \right\}
= E\left\{ (1 - R_i) E_s \left( \int_{t_1}^{t_2} g_i(u) dN_i^c(u) \mid D_i, R_i = 0 \right) \right\}
= E\left\{ I(R_i = 0) E_s \left( \int_{t_1}^{t_2} g_i(u) dN_i^c(u) \mid D_i, R_i = 0 \right) \right\}
= E\left\{ E_s \left( I(R_i = 0) \int_{t_1}^{t_2} g_i(u) dN_i^c(u) \mid D_i, R_i = 0 \right) \right\}
= E\left\{ (1 - R_i) \int_{t_1}^{t_2} g_i(u) dN_i^c(u) \right\}
\]

The proof of Lemma A.2.1 is completed. \(\square\)

Based on the above lemma, we can easily prove the following lemma.

**Lemma A.2.2:** Let a random function \(g_i(t) = g(t, X_i(t), Z_i(t), Y_i(t))\). Then under
Conditions (I), for \( t \in [t_1, t_2] \subset [0, \tau] \),
\[
n^{-1} \sum_{i=1}^{n} \ll \int_{t_1}^{t_2} g_i(u) dN_i^c(u) \gg R \quad \xrightarrow{P} \quad E \left\{ \int_{t_1}^{t_2} g_i(u) dN_i^c(u) \right\}
\]
as \( n \to \infty \).

Proof. Applying Lemma A.2.1,
\[
n^{-1} \sum_{i=1}^{n} \ll \int_{t_1}^{t_2} g_i(u) dN_i^c(u) \gg R
\]
\[
= n^{-1} \sum_{i=1}^{n} R_i \int_{t_1}^{t_2} g_i(u) dN_i^c(u) + n^{-1} \sum_{i=1}^{n} (1 - R_i) \hat{E}_2 \left\{ \int_{t_1}^{t_2} g_i(u) dN_i^c(u) \mid \mathcal{X} \right\}
\]
\[
= n^{-1} \sum_{i=1}^{n} R_i \int_{t_1}^{t_2} g_i(u) dN_i^c(u)
\]
\[
+ n^{-1} \sum_{i=1}^{n} (1 - R_i) \hat{E}_2 \left\{ \int_{t_1}^{t_2} g_i(u) dN_i^c(u) \mid \mathcal{D}_i, R_i = 0 \right\}
\]
\[
\xrightarrow{P} \quad E \left\{ R_i \int_{t_1}^{t_2} g_i(u) dN_i^c(u) \right\} + E \left\{ (1 - R_i) \int_{t_1}^{t_2} g_i(u) dN_i^c(u) \right\}
\]
\[
= E \left\{ R_i \int_{t_1}^{t_2} g_i(u) dN_i^c(u) \right\} + E \left\{ (1 - R_i) \int_{t_1}^{t_2} g_i(u) dN_i^c(u) \right\}
\]
\[
= E \left\{ \int_{t_1}^{t_2} g_i(u) dN_i^c(u) \right\}
\]
Lemma A.2.2 is proved. \( \square \)

Lemma A.2.3: Let a random function \( g_i(t) = g(t, X_i(t), Z_i(t), Y_i(t)) \). Then under

Conditions (I), for \( t \in [t_1, t_2] \subset [0, \tau] \), \( \xi_i(t) = I(S_i^* + C_i \geq t) \),
\[
n^{-1} \sum_{i=1}^{n} \ll \int_{t_1}^{t_2} K_h(u - t) g_i(u) dN_i^c(u) \gg R \quad \xrightarrow{P} \quad E(\xi_i(t) \alpha_i(t) g_i(t))
\]
as \( n \to \infty \), \( h \to 0 \) and \( nh^2 \to \infty \).

Proof. By the definition,
\[
n^{-1} \sum_{i=1}^{n} \ll \int_{t_1}^{t_2} K_h(u - t) g_i(u) dN_i^c(u) \gg R
\]
\[
= n^{-1} \sum_{i=1}^{n} R_i \int_{t_1}^{t_2} K_h(u - t) g_i(u) dN_i^c(u)
\]
+n^{-1} \sum_{i=1}^{n} (1 - R_i) \mathcal{E}_s \left\{ \int_{t_1}^{t_2} K_h(u - t)g_i(u) dN^c_i(u) \mid \mathcal{X} \right\}

By the independence of subjects, the second term can be written as

\begin{align*}
n^{-1} \sum_{i=1}^{n} (1 - R_i) \mathcal{E}_s \left\{ \int_{t_1}^{t_2} K_h(u - t)g_i(u) dN^c_i(u) \mid \mathcal{D}_i, R_i = 0 \right\} \\
= n^{-1} \sum_{i=1}^{n} (1 - R_i) \int_{t_1}^{t_2} K_h(u - t) d \left( \int_{0}^{u} \mathcal{E}_s\{g_i(v) dN^c_i(v) \mid \mathcal{D}_i, R_i = 0 \} \right). \quad (A.3)
\end{align*}

Note that the limits of integration in Lemma A.2.1 can be replaced by 0 and u, and the convergence is uniform in u. We have

\begin{align*}
n^{-1} \sum_{i=1}^{n} (1 - R_i) \left[ \int_{0}^{u} \mathcal{E}_s\{g_i(v) dN^c_i(v) \mid \mathcal{D}_i, R_i = 0 \} - \int_{0}^{u} E_s\{g_i(v) dN^c_i(v) \mid \mathcal{D}_i, R_i = 0 \} \right]
\end{align*}

converges to zero in probability uniformly in \( u \in [t_1, t_2] \). So

\begin{align*}
(A.3) &= \int_{t_1}^{t_2} K_h(u - t) d \left( n^{-1} \sum_{i=1}^{n} (1 - R_i) \int_{0}^{u} E_s\{g_i(v) dN^c_i(v) \mid \mathcal{D}_i, R_i = 0 \} \right) \\
&\quad + o_p(1) \\
&= \int_{t_1}^{t_2} K_h(u - t) d \left( E \left[ (1 - R_i) \int_{0}^{u} E_s\{g_i(v) dN^c_i(v) \mid \mathcal{D}_i, R_i = 0 \} \right] \right) + o_p(1) \\
&= \int_{t_1}^{t_2} K_h(u - t) d \left( \int_{0}^{u} E \{ (1 - R_i)g_i(v) dN^c_i(v) \} \mid \mathcal{D}_i, R_i = 0 \} \right) + o_p(1) \\
&= \int_{t_1}^{t_2} K_h(u - t) E \{ (1 - R_i)g_i(u) dN^c_i(u) \} + o_p(1).
\end{align*}

According to the argument on Page 37 of Sun & Wu (2005), the first term at the beginning of this proof is equal to

\[ \int_{t_1}^{t_2} K_h(u - t) E \{ R_i g_i(u) dN^c_i(u) \} + O_p(n^{-1/2}h^{-1}). \]

So the whole expression equals

\[ \int_{t_1}^{t_2} K_h(u - t) E \{ R_i g_i(u) dN^c_i(u) \} + \int_{t_1}^{t_2} K_h(u - t) E \{ (1 - R_i)g_i(u) dN^c_i(u) \} \]
\[ + O_p(n^{-1/2}h^{-1}) + o_p(1) \]

\[ = \int_{t_1}^{t_2} K_h(u-t)E\{g_i(u) dN^c_1(u)\} + O_p(n^{-1/2}h^{-1}) + o_p(1) \]

\[ = \int_{t_1}^{t_2} K_h(u-t)E\{g_i(u) \xi(t) dN^0_1(u)\} + O_p(n^{-1/2}h^{-1}) + o_p(1) \]

\[ = \int_{t_1}^{t_2} K_h(u-t)E\{E[\xi_i(u)g_i(u) dN^0_1(u) \mid X_i(u), Z_i(u), S_i^* + C_i \geq t]\} + O_p(n^{-1/2}h^{-1}) + o_p(1) \]

\[ = \int_{t_1}^{t_2} K_h(u-t)E[\xi_i(u)E[g_i(u) \mid X_i(u), Z_i(u), S_i^* + C_i \geq t]] + O_p(n^{-1/2}h^{-1}) + o_p(1) \]

\[ = \int_{t_1}^{t_2} K_h(u-t)E[\xi_i(u)E[g_i(u) \mid X_i(u), Z_i(u), E[dN^0_1(u) \mid X_i(u), Z_i(u), S_i^* + C_i \geq t]] + O_p(n^{-1/2}h^{-1}) + o_p(1) \]

\[ = \int_{t_1}^{t_2} K_h(u-t)E[\xi_i(u)E[g_i(u) \mid X_i(u), Z_i(u), E[dN^0_1(u) \mid X_i(u), Z_i(u)]] + O_p(n^{-1/2}h^{-1}) + o_p(1) \]

\[ = \int_{t_1}^{t_2} K_h(u-t)E[\xi_i(u)E[g_i(u) \mid X_i(u), Z_i(u), E[dN^0_1(u) \mid X_i(u), Z_i(u)]] + O_p(n^{-1/2}h^{-1}) + o_p(1) \]

\[ = \int_{t_1}^{t_2} K_h(u-t)E[\xi_i(u)E[g_i(u) \mid X_i(u), Z_i(u), E[dN^0_1(u) \mid X_i(u), Z_i(u)]] + O_p(n^{-1/2}h^{-1}) + o_p(1) \]

\[ = \int_{t_1}^{t_2} K_h(u-t)E[\xi_i(u)E[g_i(u) \mid X_i(u), Z_i(u), E[dN^0_1(u) \mid X_i(u), Z_i(u)]] + O_p(n^{-1/2}h^{-1}) + o_p(1) \]

\[ = E(\xi_i(t)\alpha_i(t)g_i(t)) + O(h^2) + O_p(n^{-1/2}h^{-1}) + o_p(1) \xrightarrow{P} E(\xi_i(t)\alpha_i(t)g_i(t)) \]

as \( n \to \infty, h \to 0 \) and \( nh^2 \to \infty \). Lemma A.2.3 is proved. \( \square \)

**Lemma A.2.4:**

\[ n^{1/2} \int_{t_1}^{t} (\tilde{E}_{xx}(u) - e_{xx}(u))(e_{xx}(u))^{-1} du \]

\[ = n^{-1/2} \sum_{i=1}^{n} \left\{ \int_{t_{i-1}}^{t_i} d(X_{zi}^I(v) + X_{zi}^{III}(v))((e_{xx}(v))^{-1} + O(h^2)) \right\} + X_{zi}^{III}(t) \]

\[ + O_p(n^{-1/2}h^2 + n^{1/2}h^2) + o_p(1) \]

converges weakly to a vector of mean-zero Gaussian processes with continuous paths

as \( n \to \infty, h \to 0 \) and \( nh^4 \to 0 \). Similar results hold for

\[ n^{1/2} \int_{t_1}^{t} (\tilde{E}_{yx}(u) - e_{yx}(u))(e_{xx}(u))^{-1} du \]
\begin{align*}
&= n^{-1/2} \sum_{i=1}^{n} \left\{ \int_{t_1-h}^{t+h} \left[ d(X_{yi}^T(v) + X_{yi}^{III}(v))((e_{xx}(v))^{-1} + O(h^2)) \right] + X_{yi}^{II}(t) \right\} \\
&\quad + O_p(n^{-1/2}h^2 + n^{1/2}h^2) + o_p(1), \\
&\quad
\end{align*}

\begin{align*}
&= n^{1/2} \int_{t_1}^{t} \beta^T(u)(\tilde{E}_{xx}(u) - e_{xx}(u))(e_{xx}(u))^{-1} du \\
&\quad - n^{-1/2} \sum_{i=1}^{n} \left\{ \int_{t_1-h}^{t+h} \left[ (\beta^T(u) + O(h^2))d(X_{zi}^T(v) + X_{zi}^{III}(v))((e_{xx}(v))^{-1} + O(h^2)) \right] \\
&\quad + X_{zi}^{II}(t) \right\} + O_p(n^{-1/2}h^2 + n^{1/2}h^2) + o_p(1).
\end{align*}

\textit{Proof.} By the definitions,

\begin{align*}
&= n^{1/2} \int_{t_1}^{t} \left( n^{-1} \sum_{i=1}^{n} \ll \int_{0}^{T} K_h(v-u)Z_i(v)X_i^T(u)dN_i^c(v) \gg_R \\
&\quad - E\{\xi_i(u)\alpha_i(u)Z_i(u)X_i^T(u)\} \right)(e_{xx}(u))^{-1} du \\
&\quad - E\{\xi_i(u)\alpha_i(u)Z_i(u)X_i^T(u)\} \right)(e_{xx}(u))^{-1} du \\
&\quad -(1-R_i)\hat{\mathcal{E}}_s\left\{ \int_{0}^{T} K_h(v-u)Z_i(v)X_i^T(u)dN_i^c(v) \mid \mathcal{X} \right\} \\
&\quad - E\{(1-R_i)\xi_i(u)\alpha_i(u)Z_i(u)X_i^T(u)\} \right)(e_{xx}(u))^{-1} du \\
&\quad -(1-R_i)\hat{\mathcal{E}}_s\left\{ \int_{0}^{T} K_h(v-u)Z_i(v)X_i^T(u)dN_i^c(v) \mid \mathcal{D}_i, R_i = 0 \right\} \\
&\quad - E\{(1-R_i)\xi_i(u)\alpha_i(u)Z_i(u)X_i^T(u)\} \right)(e_{xx}(u))^{-1} du \\
&\quad -(1-R_i)\hat{\mathcal{E}}_s\left\{ \int_{0}^{T} K_h(v-u)Z_i(v)X_i^T(u)dN_i^c(v) \mid \mathcal{D}_i, R_i = 0 \right\} \\
&\quad - E\{(1-R_i)\xi_i(u)\alpha_i(u)Z_i(u)X_i^T(u)\} \right)(e_{xx}(u))^{-1} du \\
&\quad -(1-R_i)\hat{\mathcal{E}}_s\left\{ \int_{0}^{T} K_h(v-u)Z_i(v)X_i^T(u)dN_i^c(v) \mid \mathcal{D}_i, R_i = 0 \right\} \\
&\quad - E\{(1-R_i)\xi_i(u)\alpha_i(u)Z_i(u)X_i^T(u)\} \right)(e_{xx}(u))^{-1} du \\
&\quad -(1-R_i)\hat{\mathcal{E}}_s\left\{ \int_{0}^{T} K_h(v-u)Z_i(v)X_i^T(u)dN_i^c(v) \mid \mathcal{D}_i, R_i = 0 \right\} \\
&\quad - E\{(1-R_i)\xi_i(u)\alpha_i(u)Z_i(u)X_i^T(u)\} \right)(e_{xx}(u))^{-1} du \\
&\quad -(1-R_i)\hat{\mathcal{E}}_s\left\{ \int_{0}^{T} K_h(v-u)Z_i(v)X_i^T(u)dN_i^c(v) \mid \mathcal{D}_i, R_i = 0 \right\} \\
&\quad - E\{(1-R_i)\xi_i(u)\alpha_i(u)Z_i(u)X_i^T(u)\} \right)(e_{xx}(u))^{-1} du
\end{align*}
\[ + n^{-1/2} \sum_{i=1}^{n} \int_{t_1}^{t} (1 - R_i) \left[ \hat{E}_s \left\{ \int_0^r K_h(v-u)Z_i(v)X_i^T(v) \, dN_i^c(v) \mid \mathcal{D}_i, R_i = 0 \right\} \right] (e_{xx}(u))^{-1} \, du \\
- n^{-1/2} \sum_{i=1}^{n} \int_{t_1}^{t} \left[ (1 - R_i) E_s \left\{ \int_0^r K_h(v-u)Z_i(v)X_i^T(v) \, dN_i^c(v) \mid \mathcal{D}_i, R_i = 0 \right\} \right] (e_{xx}(u))^{-1} \, du \\
+ n^{-1/2} \sum_{i=1}^{n} \int_{t_1}^{t} (1 - R_i) \int_0^r \sum_{j=1}^{n} K_h(s + T_{ij} - u)Z_{ij}X_{ij}^T \, dN_i^c(v) \\
- E \left\{ R_i \xi_i(u) \alpha_i(u)Z_i(u)X_i^T(u) \right\} (e_{xx}(u))^{-1} \, du \\
+ n^{-1/2} \sum_{i=1}^{n} \int_{t_1}^{t} \left[ (1 - R_i) E_s \left\{ \int_0^r K_h(v-u)Z_i(v)X_i^T(v) \, dN_i^c(v) \mid \mathcal{D}_i, R_i = 0 \right\} \right] (e_{xx}(u))^{-1} \, du \tag{A.4} \]

Now let us look at them summation by summation. The first summation of (A.4) equals

\[ n^{-1/2} \sum_{i=1}^{n} \int_{t_1}^{t} \left[ \int_0^r K_h(v-u)R_iZ_i(v)X_i^T(v) \, dN_i^c(v) \right] (e_{xx}(u))^{-1} \, du \\
- \int_0^r K_h(v-u)E \left\{ R_i \xi_i(v) \alpha_i(v)Z_i(v)X_i^T(v) \right\} dv + O(h^2) \]
\[-E\{R_i\xi_i(w)\alpha_i(w)Z_i(w)X_i^T(w)\}dw]\right) (e_{xx}(u))^{-1}du + O_p(n^{1/2}h^2).

Let

\[X_{zn}^I(v) = n^{-1/2} \sum_{i=1}^n \int_0^v \left[ R_iZ_i(w)X_i^T(w) - E\{R_i\xi_i(w)\alpha_i(w)Z_i(w)X_i^T(w)\} \right] dw.\]

Under Condition (I) \(X_{zn}^I(v)\) converges to a vector of mean zero Gaussian processes, saying \(X_I^I(v)\) uniformly in \(v\). Then also by the compactness of \(K(\cdot)\) and the application of the continuous mapping theorem the first summation above equals

\[
\int_{t_1}^t \int_{t_1-h}^{t+h} \frac{K_h(v-u)dX_{zn}^I(v)(e_{xx}(u))^{-1}du}{O_p(n^{1/2}h^2)} = \int_{t_1}^t \int_{t_1-h}^{t+h} dX_{zn}^I(v)h^{-1}K\left(\frac{v-u}{h}\right)(e_{xx}(u))^{-1}du + O_p(n^{1/2}h^2)
\]

\[
= \int_{t_1}^t \int_{t_1-h}^{t+h} dX_{zn}^I(v)((e_{xx}(v))^{-1} + O(h^2)) + O_p(n^{1/2}h^2)
\]

\[
\overset{D}{\to} \int_{t_1}^t dX_I^I(v)((e_{xx}(v))^{-1})
\]

as \(n \to \infty, h \to 0\) and \(nh^4 \to 0\).

Then the third summation in (A.4) is equal to

\[
n^{-1/2} \sum_{i=1}^n \int_{t_1}^t \left[ \int_0^\tau K_h(v-u)E_s\{(1 - R_i)Z_i(v)X_i^T(v)\}dN_i^c(v) | D_i, R_i = 0 \right]
\]

\[
- E\{(1 - R_i)\xi_i(u)\alpha_i(u)Z_i(u)X_i^T(u)\} (e_{xx}(u))^{-1}du
\]

\[
= n^{-1/2} \sum_{i=1}^n \int_{t_1}^t \left[ \int_0^\tau K_h(v-u)E_s\{(1 - R_i)Z_i(v)X_i^T(v)\}dN_i^c(v) | D_i, R_i = 0 \right]
\]

\[
- \int_0^\tau K_h(v-u)E\{(1 - R_i)\xi_i(v)\alpha_i(v)Z_i(v)X_i^T(v)\}dv + O(h^2)] (e_{xx}(u))^{-1}du
\]

\[
= \int_{t_1}^t \left[ \int_0^\tau K_h(v-u)\left\{ n^{-1/2} \sum_{i=1}^n (E_s\{(1 - R_i)Z_i(v)X_i^T(v)\}dN_i^c(v) | D_i, R_i = 0 \right\}
\]

\[
- E\{(1 - R_i)\xi_i(v)\alpha_i(v)Z_i(v)X_i^T(v)\}dv \right\} (e_{xx}(u))^{-1}du + O_p(n^{1/2}h^2)
\]
\[
= \int_{t_1}^{t} \left[ \int_{0}^{\tau} K_h(v-u) d\left\{ \sum_{i=1}^{n} \left( 1 - R_i \right) Z_i(w) X_i^T(w) dN_i^c(w) \mid \mathcal{D}_i, R_i = 0 \right\} \right] (e_{xx}(u))^{-1} du + O_p(n^{1/2}h^2).
\]

Let
\[
X_{zn}^{III}(v) = n^{-1/2} \sum_{i=1}^{n} \int_{0}^{v} \left( 1 - R_i \right) Z_i(w) X_i^T(w) dN_i^c(w) \mid \mathcal{D}_i, R_i = 0 \right\} - E\{ (1 - R_i) \xi_i(w) \alpha_i(w) Z_i(w) X_i^T(w) \} dw).
\]

Under Condition (I) \( X_{zn}^{III}(v) \) converges to a vector of mean zero Gaussian processes, saying \( X_{zn}^{III}(v) \) uniformly in \( v \). Now follow the argument in discussing the first summation, we know
\[
\int_{t_1}^{t} \int_{0}^{\tau} K_h(v-u) dX_{zn}^{III}(v)(e_{xx}(u))^{-1} du + O_p(n^{1/2}h^2)
\]
\[
= \int_{t_1-h}^{t+h} \left[ dX_{zn}^{III}(v)((e_{xx}(v))^{-1} + O(h^2)) \right] + O_p(n^{1/2}h^2)
\]
\[
\rightarrow D \int_{t_1}^{t} \left[ dX_{zn}^{III}(v)((e_{xx}(v))^{-1}) \right]
\]
as \( n \to \infty, h \to 0 \) and \( nh^4 \to 0 \).

Under the assumption that \( \{S_i\} \) are independent of \( \mathcal{D}_i \) and defining the counting process
\[
N_i^*(t) = \sum_{j=1}^{n_i} I(T_{ij} \leq t) I(C_i \geq t)
\]
with the mean rate
\[
E\{dN_i^*(t) \mid R_i, X_i(t), Y_i(t), Z_i(t), V_i \} = \alpha_i^*(t) dt,
\]
the second summation of (A.4) equals
\[
n^{-1/2} \sum_{i=1}^{n} \int_{t_1}^{t} (1 - R_i) \int_{0}^{L} \sum_{j=1}^{n_i} K_h(s + T_{ij} - u) Z_{ij} X_{ij}^T I(C_i \geq T_{ij}) \left[ \frac{d\hat{F}_s(s)}{\hat{F}_s(V_i)} \right]
\]
\[- \frac{dF_s(s)}{F_s(V_i)} \left( e_{xx}(u) \right)^{-1} du \]

\[= n^{-1/2} \sum_{i=1}^{n} \int_{t_{1}}^{t} (1 - R_i) \int_{0}^{L} \int_{0}^{\tau} K_h(v - u)Z_i(v)X_i^T(v)dN_s^*(v - s) \left[ \left( \frac{1}{\hat{F}_s(V_i)} \right. \right. \]

\[= n^{-1/2} \sum_{i=1}^{n} \int_{t_{1}}^{t} (1 - R_i) \int_{0}^{L} \int_{0}^{\tau} K_h(v - u)Z_i(v)X_i^T(v)dN_s^*(v - s) \frac{F_s(V_i) - \hat{F}_s(V_i)}{F_s^2(V_i)} \]

\[+ o_p(1) \]

\[= n^{-1} \sum_{i=1}^{n} \int_{t_{1}}^{t} \int_{0}^{L} \int_{0}^{\tau} (1 - R_i)K_h(v - u)Z_i(v)X_i^T(v)dN_s^*(v - s) \]

\[\frac{n^{1/2}(\hat{S}_s(V_i) - S_s(V_i))}{F_s^2(V_i)} dF_s(s)(e_{xx}(u))^{-1} du \] (A.5)

\[- n^{-1} \sum_{i=1}^{n} \int_{t_{1}}^{t} \int_{0}^{L} \int_{0}^{\tau} K_h(v - u)Z_i(v)X_i^T(v)dN_s^*(v - s) \]

\[\frac{d[n^{1/2}(\hat{S}_s(s) - S_s(s))]}{F_s(V_i)} (e_{xx}(u))^{-1} du \] (A.6)

\[+ o_p(1) \]

Plugging (A.1) into both (A.5) and (A.6), we have

\[(A.5) = n^{-1} \sum_{i=1}^{n} \int_{t_{1}}^{t} \int_{0}^{L} \int_{0}^{\tau} (1 - R_i)K_h(v - u)Z_i(v)X_i^T(v)dN_s^*(v - s) \frac{n^{-1/2}F_s(V_i)}{F_s^2(V_i)} \]

\[\int_{0}^{L-(V_i))} \frac{dM_R(x)}{y_R(x)} dF_s(s)(e_{xx}(u))^{-1} du + o_p(1) \]

\[= \int_{t_{1}}^{t} \int_{0}^{L} n^{-1} \sum_{i=1}^{n} \int_{0}^{\tau} (1 - R_i)K_h(v - u)Z_i(v)X_i^T(v)dN_s^*(v - s) \frac{n^{-1/2}}{F_s(V_i)} \]

\[\int_{0}^{\infty} \frac{I(x \leq (L - (V_i))] - \frac{dM_R(x)}{y_R(x)} dF_s(s)(e_{xx}(u))^{-1} du + o_p(1) \]

\[= n^{-1/2} \sum_{i=1}^{n} \int_{t_{1}}^{t} \int_{0}^{\tau} K_h(v - u)n^{-1} \sum_{i=1}^{n} (1 - R_i)Z_i(v)X_i^T(v)dN_s^*(v - s) \]

\[+ o_p(1) \]
\[
I(x \leq (L - (V_i))- \frac{1}{F_s(V_i)} (e_{xx}(u))^{-1} du F_s(s) \frac{dM^R(x)}{y^R(x)} + o_p(1),
\]
and
\[
(A.6) = -n^{-1} \sum_{i=1}^{n} \int_{t_1}^{t} (1 - R_i) \int_{0}^{L} \int_{0}^{\tau} K_h(v - u) Z_i(v) X^T_i(v) dN^*_i(v - s) F_s(V_i) d[1/2(S_s(s) - S_s(s))][e_{xx}(u)]^{-1} du \\
- n^{-1} \sum_{i=1}^{n} \int_{t_1}^{t} (1 - R_i) \int_{0}^{L} \int_{0}^{\tau} K_h(v - u) Z_i(v) X^T_i(v) dN^*_i(v - s) F_s(V_i) \\
\int_{0}^{(L-s)} \frac{dM^R(x)}{y^R(x)} (e_{xx}(u))^{-1} du \\
+ n^{-1} \sum_{i=1}^{n} \int_{t_1}^{t} (1 - R_i) \int_{0}^{L} \int_{0}^{\tau} K_h(v - u) Z_i(v) X^T_i(v) dN^*_i(v - s) n^{-1/2} F_s(s) \\
\int_{0}^{(L-s)} \frac{dM^R((L-s)-)}{y^R((L-s)-)} (e_{xx}(u))^{-1} du \\
- n^{-1/2} \int_{0}^{L} \int_{0}^{(L-s)} \int_{t_1}^{t} \int_{0}^{\tau} K_h(v - u) n^{-1} \sum_{i=1}^{n} (1 - R_i) \frac{Z_i(v) X^T_i(v)}{F_s(V_i)} \\
dN^*_i(v - s) (e_{xx}(u))^{-1} du dF_s(s) \frac{dM^R(x)}{y^R(x)} \\
+ n^{-1/2} \int_{0}^{L} \int_{t_1}^{t} \int_{t_1}^{\tau} K_h(v - u) n^{-1} \sum_{i=1}^{n} (1 - R_i) \frac{Z_i(v) X^T_i(v) dN^*_i(v - s)}{F_s(V_i)} F_s(s) \\
\int_{0}^{(L-s)-} \frac{dM^R((L-s)-)}{y^R((L-s)-)}.\]

Since
\[
\int_{0}^{\tau} K_h(v - u) n^{-1} \sum_{i=1}^{n} (1 - R_i) Z_i(v) X^T_i(v) dN^*_i(v - s) I(x \leq (L - (V_i))- \frac{1}{F_s(V_i)} \\
\int_{0}^{t} K_h(v - u)d \left( n^{-1} \sum_{i=1}^{n} \int_{0}^{L} (1 - R_i) Z_i(w) X^T_i(w) dN^*_i(w - s) \frac{I(x \leq (L - (V_i))-}{F_s(V_i)} \right)
\]
\begin{align*}
&= \int_0^\tau K_h(v-u)dE\left\{ \int_0^v (1-R_i)Z_i(w)X_i^T(w)\frac{I(x \leq (L-(V_i))^-)}{F_s(V_i)} \right\} + o_p(1) \\
&= \int_0^\tau K_h(v-u)dE\left\{ E\left[ \int_0^v (1-R_i)Z_i(w)X_i^T(w)\frac{I(x \leq (L-(V_i))^-)}{F_s(V_i)} \right| R_i, X_i(t), Y_i(t), Z_i(t), V_i \right\} + o_p(1) \\
&= \int_0^\tau K_h(v-u)dE\left\{ \int_0^v (1-R_i)Z_i(w)X_i^T(w)E[\frac{I(x \leq (L-(V_i))^-)}{F_s(V_i)}] \right\} + o_p(1) \\
&= \int_0^\tau K_h(v-u)dE\left\{ \int_0^v (1-R_i)Z_i(w)X_i^T(w)\alpha_i^*(w-s)dw \frac{I(x \leq (L-(V_i))^-)}{F_s(V_i)} \right\} + o_p(1) \\
&= \int_0^\tau K_h(v-u)dE\left\{ (1-R_i)Z_i(v)X_i^T(v)\alpha_i^*(v-s) \frac{I(x \leq (L-(V_i))^-)}{F_s(V_i)} \right\} dv + o_p(1) \\
&= E\left\{ (1-R_i)Z_i(u)X_i^T(u)\alpha_i^*(u-s) \frac{I(x \leq (L-(V_i))^-)}{F_s(V_i)} \right\} + O(h^2) + o_p(1)
\end{align*}

and similarly

\begin{align*}
&= \int_0^\tau K_h(v-u)n^{-1}\sum_{i=1}^n (1-R_i)Z_i(v)X_i^T(v)\frac{dN_i^*(v-s)}{F_s(V_i)} \\
&= E\left\{ (1-R_i)Z_i(u)X_i^T(u)\alpha_i^*(u-s) \frac{I(x \leq (L-(V_i))^-)}{F_s(V_i)} \right\} + O(h^2) + o_p(1),
\end{align*}

then

\begin{align*}
(A.5) &= n^{-1/2} \int_0^\infty \int_0^L \int_{t_1}^t E\left\{ (1-R_i)Z_i(u)X_i^T(u)\alpha_i^*(u-s) \frac{I(x \leq (L-(V_i))^-)}{F_s(V_i)} \right\} \\
&\quad (e_{xx}(u))^{-1}du\cdot df_s(s)\frac{dM^R(x)}{y^R(x)} + o_p(n^{-1/2}h^2) + o_p(1), \\
(A.6) &= -n^{-1/2} \int_0^L \int_{(L-x)^-}^{(L-x)^-} \int_{t_1}^t E\left\{ (1-R_i)Z_i(u)X_i^T(u)\alpha_i^*(u-s) \frac{I(x \leq (L-(V_i))^-)}{F_s(V_i)} \right\} (e_{xx}(u))^{-1} \\
&\quad du\cdot df_s(s)\frac{dM^R(x)}{y^R(x)} \\
&\quad + n^{-1/2} \int_0^L \int_{t_1}^t E\left\{ (1-R_i)Z_i(u)X_i^T(u)\alpha_i^*(u-s) \frac{I(x \leq (L-(V_i))^-)}{F_s(V_i)} \right\} F_s(s)(e_{xx}(u))^{-1} \\
&\quad du\cdot df_s(s)\frac{dM^R((L-s)^-)}{y^R((L-s)^-)} + o_p(n^{-1/2}h^2) + o_p(1)
\end{align*}
Thus the second summation of (A.4) equals

\[ n^{-1/2} \left[ \int_{0}^{\infty} \int_{0}^{L} \int_{t_1}^{t} E \left\{ (1 - R_i)Z_i(u)X_i^{T}(u)\alpha_i^*(u - s) \frac{I(x \leq (L - V_i) -)}{F_s(V_i)} \right\} \right. \]

\[ \left. (e_{xx}(u))^{-1} du dF_s(s) \frac{dM^R(x)}{y^R(x)} \right. \]

\[ - \int_{0}^{L} \int_{0}^{(L-x)-} \int_{t_1}^{t} E \left\{ (1 - R_i) \frac{Z_i(u)X_i^{T}(u)\alpha_i^*(u - s)}{F_s(V_i)} \right\} (e_{xx}(u))^{-1} du \]

\[ dF_s(s) \frac{dM^R(x)}{y^R(x)} \]

\[ + \int_{0}^{L} \int_{t_1}^{t} E \left\{ (1 - R_i) \frac{Z_i(u)X_i^{T}(u)\alpha_i^*(u - s)}{F_s(V_i)} \right\} F_s(s)(e_{xx}(u))^{-1} du \]

\[ \frac{dM^R((L-s)-)}{y^R((L-s)-)} \right] + O_p(n^{-1/2}h^2) + o_p(1). \]

By the multivariate martingale central limit theorem, we know that the above three terms converge weakly to Wiener processes since the integrants of the martingale integral are deterministic functions.

Above all, Equation (A.4) weakly converges to a vector of mean zero Gaussian processes with continuous paths as \( n \to \infty, h \to 0 \) and \( nh^4 \to 0 \).

Recall the definitions in Section A.1. We can have the following lemma.

**Lemma A.2.5:**

\[ n^{1/2} \int_{t_1}^{t} \left\{ \beta^T(u, \gamma_0) - \beta^T_0(u) \right\} du \]

\[ = n^{-1/2} \sum_{i=1}^{n} \left\{ X_{yi}(t) - X_{yi}^{II}(t) - X_{zi}^{II}(t) \right\} \]

\[ + \int_{t_1-h}^{t+h} d(X_{yi}(v) + X_{yi}^{III}(v) - X_{zi}^{I}(v) - X_{zi}^{III}(v))((e_{xx}(v))^{-1} + O(h^2)) \]

\[ - \int_{t_1-h}^{t+h} (\beta^T(v) + O(h^2)) d(X_{zi}(v) + X_{zi}^{III}(v))((e_{xx}(v))^{-1} + O(h^2)) \}

\[ + O_p(n^{-1/2}h^2 + n^{1/2}h^2) + o_p(1) \]

converges weakly to a vector of mean zero Gaussian processes with continuous paths as \( n \to \infty, h \to 0 \) and \( nh^4 \to 0 \).
Proof. By the definitions,

\[
\begin{align*}
&n^{1/2} \int_{t_1}^{t} \{ \tilde{\beta}^T(u, \gamma_0) - \beta_0^T(u) \} du \\
&= \int_{t_1}^{t} n^{1/2} \{ \tilde{Y}_x(u) - \gamma_0^T \tilde{Z}_x(u) - (y_x(u) - \gamma_0^T z_x(u)) \} du \\
&= n^{1/2} \int_{t_1}^{t} \{ \tilde{Y}_x(u) - y_x(u) \} du - \gamma_0^T n^{1/2} \int_{t_1}^{t} \{ \tilde{Z}_x(u) - z_x(u) \} du
\end{align*}
\]

By the continuous mapping theorem, it is sufficient to prove that

\[
\left( n^{1/2} \int_{t_1}^{t} \{ \tilde{Y}_x(u) - y_x(u) \} du, n^{1/2} \int_{t_1}^{t} \{ \tilde{Z}_x(u) - z_x(u) \} du \right) \tag{A.7}
\]

converges weakly to a vector of mean zero Gaussian processes with continuous sample paths. And

\[
\begin{align*}
n^{1/2} \int_{t_1}^{t} \{ \tilde{Y}_x(u) - y_x(u) \} du \\
&= n^{1/2} \int_{t_1}^{t} \{ \tilde{E}_{yx}(u)(\tilde{E}_{xx}(u))^{-1} - e_{yx}(u)(e_{xx}(u))^{-1} \} du \\
&= n^{1/2} \int_{t_1}^{t} \{ [\tilde{E}_{yx}(u) - e_{yx}(u)](\tilde{E}_{xx}(u))^{-1} - e_{yx}(u)(\tilde{E}_{xx}(u))^{-1}[\tilde{E}_{xx}(u) \\
&\quad - e_{xx}(u)](e_{xx}(u))^{-1} \} du \\
&= n^{1/2} \int_{t_1}^{t} \{ [\tilde{E}_{yx}(u) - e_{yx}(u)](e_{xx}(u))^{-1} - e_{yx}(u)(e_{xx}(u))^{-1}[\tilde{E}_{xx}(u) \\
&\quad - e_{xx}(u)](e_{xx}(u))^{-1} \} du + o_p(1)
\end{align*}
\]

\[
n^{1/2} \int_{t_1}^{t} \{ \tilde{Y}_x(u) - y_x(u) \} du \text{ has a similar decomposition. Under Condition (I), applying Lemma A.1 of Lin & Ying (2001) and Lemma A.2.4 above,}
\]

\[
\begin{align*}
n^{1/2} \int_{t_1}^{t} \{ \tilde{Y}_x(u) - y_x(u) \} du \text{ and } n^{1/2} \int_{t_1}^{t} \{ \tilde{Z}_x(u) - z_x(u) \} du
\end{align*}
\]

converges weakly to a mean zero Gaussian process respectively. So using the Wald device, we could have the joint weak convergence of (A.7) which leads to the weak convergence of \( n^{1/2} \int_{t_1}^{t} \{ \tilde{\beta}^T(u, \gamma_0) - \beta_0^T(u) \} du \) with zero mean. This completes the proof. □
Lemma A.2.3, we have

\[ \tilde{\beta}(t; \gamma) = Y_x^T(t) - \tilde{Z}_x^T(t) \gamma \rightarrow_P y_x^T(t) - z_x^T(t) \gamma \]

uniformly in \( t \in [t_1, t_2] \) as \( n \rightarrow \infty, h \rightarrow 0 \). Since \( \beta_0(t) = y_x^T(t) - z_x^T(t) \gamma_0 \), by using (2.7), replace \( \beta(s) \) in (2.3) and Applying Lemma A.2.2 We have \( n^{-1} \tilde{l}(\gamma) \) equals

\[
\begin{align*}
&\sum_{i=1}^{n} n^{-1} \int_{0}^{\tau} W_i(s) \{ Y_i(s) - (\tilde{Y}_x(s) - \gamma^T \tilde{Z}_x(s))X_i(s) - \gamma^T Z_i(s) \}^2 dN_i^c(s) \\
&+ n^{-1} \sum_{i=1}^{n} (1 - R_i) \mathbb{E}_\mathcal{S} \left[ \int_{0}^{\tau} W_i(s) \{ Y_i(s) - (\tilde{Y}_x(s) - \gamma^T \tilde{Z}_x(s))X_i(s) - \gamma^T Z_i(s) \}^2 dN_i^c(s) \big| \mathcal{X} \right] \\
&= \sum_{i=1}^{n} n^{-1} \int_{0}^{\tau} W_i(s) \{ Y_i(s) - (\tilde{Y}_x(s) - \gamma^T \tilde{Z}_x(s))X_i(s) - \gamma^T Z_i(s) \}^2 dN_i^c(s) \gg_R \\
&= \sum_{i=1}^{n} n^{-1} \int_{0}^{\tau} W_i(s) \{ Y_i(s) - \tilde{Y}_x(s)X_i(s) + \gamma^T (\tilde{Z}_x(s))X_i(s) \\
&\quad - Z_i(s)) \}^2 dN_i^c(s) \gg_R
\end{align*}
\]

where

\[
\begin{align*}
&\int_{0}^{\tau} W_i(s) \{ Y_i(s) - \tilde{Y}_x(s)X_i(s) + \gamma^T (\tilde{Z}_x(s))X_i(s) - Z_i(s)) \}^2 dN_i^c(s) \\
&= \int_{0}^{\tau} W_i(s) \{ Y_i(s) - \tilde{Y}_x(s)X_i(s) + \gamma^T (\tilde{Z}_x(s))X_i(s) - Z_i(s)) \}^2 - \{ Y_i(s) \\
&\quad - y_x(s)X_i(s) + \gamma^T (z_x(s))X_i(s) - Z_i(s)) \}^2 dN_i^c(s) \\
&\quad + \int_{0}^{\tau} W_i(s) \{ Y_i(s) - y_x(s)X_i(s) + \gamma^T (z_x(s))X_i(s) - Z_i(s)) \}^2 dN_i^c(s) \\
&= \int_{0}^{\tau} \{ -(\tilde{Y}_x(s) - y_x(s)) + \gamma^T (\tilde{Z}_x(s) - z_x(s)) \} X_i(s)W_i(s) [2Y_i(s) \\
&\quad - (\tilde{Y}_x(s) + y_x(s))X_i(s) + \gamma^T ((\tilde{Z}_x(s) + z_x(s))X_i(s) - 2Z_i(s))]dN_i^c(s) \\
&\quad + \int_{0}^{\tau} W_i(s) \{ Y_i(s) - y_x(s)X_i(s) + \gamma^T (z_x(s))X_i(s) - Z_i(s)) \}^2 dN_i^c(s)
\end{align*}
\]
The first term equals 

\[
\begin{align*}
\sum_{i=1}^{n} & \left\{ -(\bar{Y}_x(s) - y_x(s)) + \gamma^T(\bar{Z}_x(s) - z_x(s)) \right\} X_i(s) W_i(s) \nonumber \\
& + \gamma^T(\bar{Z}_x(s) - z_x(s)) X_i(s) + 2y_x(s)X_i(s) + 2Y_i(s) \\
& + \gamma^T(2z_x(s)X_i(s) - 2Z_i(s)) \right\} dN^c_i(s) \\
& + \int_0^\tau W_i(s) \{ Y_i(s) - y_x(s) \} X_i(s) + \gamma^T(z_x(s)X_i(s) - Z_i(s)) \right\}^2 dN^c_i(s) \\
& = \int_0^\tau \left\{ -(\bar{Y}_x(s) - y_x(s)) + \gamma^T(\bar{Z}_x(s) - z_x(s)) \right\} X_i(s) W_i(s) \\
& + \gamma^T(\bar{Z}_x(s) - z_x(s)) X_i(s) + 2y_x(s)X_i(s) + 2Y_i(s) \\
& + \gamma^T(2z_x(s)X_i(s) - 2Z_i(s)) \right\} dN^c_i(s) \\
& + \int_0^\tau W_i(s) \{ Y_i(s) - y_x(s) \} X_i(s) + \gamma^T(z_x(s)X_i(s) - Z_i(s)) \right\}^2 dN^c_i(s)
\end{align*}
\]

So by the linearity of the operation \( \ll \gg \),

\[
n^{-1} \tilde{I}(\gamma) = n^{-1} \sum_{i=1}^{n} \ll \int_0^\tau \left\{ -(\bar{Y}_x(s) - y_x(s)) + \gamma^T(\bar{Z}_x(s) - z_x(s)) \right\} X_i(s) W_i(s) \\
& + \gamma^T(\bar{Z}_x(s) - z_x(s)) X_i(s) + 2y_x(s)X_i(s) + 2Y_i(s) \\
& + \gamma^T(2z_x(s)X_i(s) - 2Z_i(s)) \right\} dN^c_i(s) \gg_R \\
+ n^{-1} \sum_{i=1}^{n} \ll \int_0^\tau 2\left\{ -(\bar{Y}_x(s) - y_x(s)) + \gamma^T(\bar{Z}_x(s) - z_x(s)) \right\} X_i(s) W_i(s) \\
& \{ y_x(s)X_i(s) + Y_i(s) + \gamma^T(z_x(s)X_i(s) - Z_i(s)) \} dN^c_i(s) \gg_R \\
+ n^{-1} \sum_{i=1}^{n} \ll \int_0^\tau W_i(s) \{ Y_i(s) - y_x(s) \} X_i(s) + \gamma^T(z_x(s)X_i(s) - Z_i(s)) \right\}^2 dN^c_i(s) \gg_R + o_p(1)
\]

The first term equals

\[
\begin{align*}
n^{-1} \sum_{i=1}^{n} & R_i \int_0^\tau \left\{ -(\bar{Y}_x(s) - y_x(s)) + \gamma^T(\bar{Z}_x(s) - z_x(s)) \right\} X_i(s) W_i(s) dN^c_i(s) \\
& + n^{-1} \sum_{i=1}^{n} (1 - R_i) \tilde{E}_i \left\{ \int_0^\tau \left\{ -(\bar{Y}_x(s) - y_x(s)) + \gamma^T(\bar{Z}_x(s) - z_x(s)) \right\} X_i(s) \right\}^2 W_i(s) dN^c_i(s) \\
& dN^c_i(s) \mid X
\end{align*}
\]

\[
= n^{-1} \sum_{i=1}^{n} R_i \int_0^\tau \left\{ -(\bar{Y}_x(s) - y_x(s)) + \gamma^T(\bar{Z}_x(s) - z_x(s)) \right\} X_i(s) W_i(s) dN^c_i(s)
\]
\[+ n^{-1} \sum_{i=1}^{n} (1 - R_i) E_s \left\{ \int_{0}^{\tau} \left\{ \left[ - (\tilde{Y}_x(s) - y_x(s)) + \gamma^T (\tilde{Z}_x(s) - z_x(s)) \right] X_i(s) \right\} ^2 W_i(s) \right. \\
\left. \quad dN_i^c(s) \mid \mathcal{X} \right\} + o_p(1)\]

\[= \int_{0}^{\tau} \left\{ - (\tilde{Y}_x(s) - y_x(s)) + \gamma^T (\tilde{Z}_x(s) - z_x(s)) \right\} \left( n^{-1} \sum_{i=1}^{n} R_i X_i(s) X_i(s)^T W(s) \right. \\
\left. \quad dN_i^c(s) \right\} \left[ - (\tilde{Y}_x(s) - y_x(s)) + \gamma^T (\tilde{Z}_x(s) - z_x(s)) \right] ^T + o_p(1)\]

\[+ E_s \left\{ \int_{0}^{\tau} \left\{ - (\tilde{Y}_x(s) - y_x(s)) + \gamma^T (\tilde{Z}_x(s) - z_x(s)) \right\} \left( n^{-1} \sum_{i=1}^{n} (1 - R_i) X_i(s) X_i(s)^T W_i(s) \right. \\
\left. \quad dN_i^c(s) \right\} \left[ - (\tilde{Y}_x(s) - y_x(s)) + \gamma^T (\tilde{Z}_x(s) - z_x(s)) \right] ^T \mid \mathcal{X} \right\} + o_p(1)\]

Since

\[n^{-1} \sum_{i=1}^{n} \int_{0}^{\tau} R_i X_i(u) X_i(u)^T W_i(u) dN_i^c(u) \quad \xrightarrow{P} \quad E \left\{ \int_{0}^{\tau} R_i X_i(u) X_i(u)^T W(u) dN_i^c(u) \right\},\]

\[n^{-1} \sum_{i=1}^{n} \int_{0}^{\tau} (1 - R_i) X_i(u) X_i(u)^T W_i(u) dN_i^c(u) \quad \xrightarrow{P} \quad E \left\{ \int_{0}^{\tau} (1 - R_i) X_i(u) X_i(u)^T W_i(u) dN_i^c(u) \right\}\]

and by the uniform convergence of \(\tilde{Y}_x(s)\) and \(\tilde{Z}_x(s)\) which lead to \(- (\tilde{Y}_x(s) - y_x(s)) + \gamma^T (\tilde{Z}_x(s) - z_x(s)) \xrightarrow{P} 0\), the first term converges to zero in probability.
Similarly to the first term, the second term converges to zero in probability.

\[
\begin{align*}
 n^{-1} \sum_{i=1}^{n} R_i \int_{0}^{r} 2\{- (\bar{Y}_x(s) - y_x(s)) + \gamma^T(\bar{Z}_x(s) - z_x(s))\} X_i(s) W_i(s) \{y_x(s) X_i(s) + Y_i(s) + \gamma^T[z_x(s) X_i(s) - Z_i(s)]\} dN^c_i(s)
 &+ n^{-1} \sum_{i=1}^{n} (1 - R_i) \bar{E}_s \left\{ \int_{0}^{r} 2\{- (\bar{Y}_x(s) - y_x(s)) + \gamma^T(\bar{Z}_x(s) - z_x(s))\} X_i(s) W_i(s) \{y_x(s) X_i(s) + Y_i(s) + \gamma^T[z_x(s) X_i(s) - Z_i(s)]\} dN^c_i(s) \right\} X
 &+ o_p(1).
\end{align*}
\]

Also

\[
\begin{align*}
 n^{-1} \sum_{i=1}^{n} \int_{0}^{s} R_i X_i(u) W_i(u) \{y_x(u) X_i(u) + Y_i(u)
 &+ \gamma^T[z_x(u) X_i(u) - Z_i(u)]\} dN^c_i(u)
 \xrightarrow{P} E \left\{ \int_{0}^{s} R_i X_i(u) W_i(u) \{y_x(u) X_i(u) + Y_i(u)
 &+ \gamma^T[z_x(u) X_i(u) - Z_i(u)]\} dN^c_i(u) \right\},
\end{align*}
\]

\[
\begin{align*}
 n^{-1} \sum_{i=1}^{n} \int_{0}^{s} (1 - R_i) X_i(u) W(u) \{y_x(u) X_i(u) + Y_i(u)
 &+ \gamma^T[z_x(u) X_i(u) - Z_i(u)]\} dN^c_i(u)
 \xrightarrow{P} E \left\{ \int_{0}^{s} (1 - R_i) X_i(u) W(u) \{y_x(u) X_i(u) + Y_i(u)
 &+ \gamma^T[z_x(u) X_i(u) - Z_i(u)]\} dN^c_i(u) \right\}.
\end{align*}
\]

Similarly to the first term, the second term converges to zero in probability.
Therefore according to our lemma A.2.2,

\[
n^{-1} \tilde{I}(\gamma) = n^{-1} \sum_{i=1}^{n} \ll \int_{0}^{\tau} W_i(s) \{Y_i(s) - y_z(s)X_i(s) + \gamma^T(z_x(s)X_i(s) - Z_i(s))\}^2 dN_i^c(s) \gg_R + o_p(1)
\]

\[
\xrightarrow{P} E\left\{ \int_{0}^{\tau} w(s) \{Y_i(s) - y_z(s)X_i(s) + \gamma^T(z_x(s)X_i(s) - Z_i(s))\}^2 dN_i^c(s) \right\}
\]

\[
= E\left\{ \int_{0}^{\tau} w(s) \{Y_i(s) - (y_z(s) - \gamma_0 z_x(s))X_i(s) - \gamma_0 Z_i(s)\}^2 dN_i^c(s) \right\}
\]

\[
= E\left\{ \int_{0}^{\tau} w(s) \{\epsilon_i(s) + (\gamma - \gamma_0)^T(z_x(s)X_i(s) - Z_i(s))\}^2 dN_i^c(s) \right\}
\]

\[
= E\left\{ \int_{0}^{\tau} w(s) \{\epsilon_i^2(s) + 2\epsilon_i(s)[(\gamma - \gamma_0)^T(z_x(s)X_i(s) - Z_i(s))]\} + [((\gamma - \gamma_0)^T(z_x(s)X_i(s) - Z_i(s))]^2 dN_i^c(s) \right\}
\]

\[
= E\left\{ \int_{0}^{\tau} w(s) \{\epsilon_i^2(s) + [(\gamma - \gamma_0)^T(z_x(s)X_i(s) - Z_i(s))]^2\} dN_i^c(s) \right\}
\]

\[
+ \int_{0}^{\tau} E\{E[2w(s)\epsilon_i(s)(\gamma - \gamma_0)^T(z_x(s)X_i(s) - Z_i(s))dN_i^c(s) | X_i(s), Z_i(s)]\}
\]

\[
= E\left\{ \int_{0}^{\tau} w(s) \{\epsilon_i^2(s) + [(\gamma - \gamma_0)^T(z_x(s)X_i(s) - Z_i(s))]^2\} dN_i^c(s) \right\}
\]

\[
+ \int_{0}^{\tau} E\{2w(s)(\gamma - \gamma_0)^T(z_x(s)X_i(s) - Z_i(s))E[\epsilon_i(s)]dN_i^c(s) | X_i(s), Z_i(s)]\}
\]

\[
E[dN_i^c(s) | X_i(s), Z_i(s)]
\]

\[
= E\left\{ \int_{0}^{\tau} w(s) \{\epsilon_i^2(s) + [(\gamma - \gamma_0)^T(z_x(s)X_i(s) - Z_i(s))]^2\} dN_i^c(s) \right\}
\]
uniformly in \( \gamma \) in \( \Gamma \). Let \( d(\gamma, \gamma_0) \) be the Euclidean distance between \( \gamma \) and \( \gamma_0 \). Therefore, for every \( \epsilon > 0 \),

\[
\sup_{\gamma : d(\gamma, \gamma_0) \geq \epsilon} (-l_0(\gamma)) = -\inf_{\gamma : d(\gamma, \gamma_0) \geq \epsilon} l_0(\gamma)
\]

\[
= -\inf_{\gamma : d(\gamma, \gamma_0) \geq \epsilon} E\left\{ \int_0^\tau w(s)\{\epsilon_t^2(s) + [(\gamma - \gamma_0)^T(z_x(s)X_i(s) - Z_i(s))]^2\} \, dN_i^c(s) \right\}
\]

\[
< -\inf_{\gamma : d(\gamma, \gamma_0) \geq \epsilon} E\left\{ \int_0^\tau w(s)\{\epsilon_t^2(s)\} \, dN_i^c(s) \right\} = -\inf_{\gamma : d(\gamma, \gamma_0) \geq \epsilon} l_0(\gamma)
\]

\[
= \sup_{\gamma : d(\gamma, \gamma_0) \geq \epsilon} (-l_0(\gamma)).
\]

Then according to Theorem 5.7 of van der Vaart (1998), we have \( \hat{\gamma} \xrightarrow{P} \gamma_0 \). \( \square \)

**Proof of Theorem 3.2**

By continuous mapping theorem, the asymptotic uniform consistency of \( \hat{\beta}(t) \) on \([t_1, t_2]\) can be easily obtained by the consistency of \( \hat{\gamma} \), the uniform consistency of \( \hat{Y}_x(t) \) and \( \hat{Z}_x(t) \) since \( \hat{\beta}(t) = \hat{Y}_x^T(t) - \hat{Z}_x^T(t)\hat{\gamma} \). \( \square \)

**Proof of Theorem 3.3**

Recall the score function \( U(\gamma) \) and the Taylor expansion of \( U(\hat{\gamma}) \) at \( \gamma_0 \)

\[
n^{1/2}(\hat{\gamma} - \gamma_0) = -\left( n^{-1}\frac{\partial U(\gamma^*)}{\partial \gamma^T} \right)^{-1} [n^{-1/2}U(\gamma_0)], \tag{A.8}
\]

where \( \gamma^* \) is on the line segment between \( \hat{\gamma} \) and \( \gamma_0 \).

By plugging (2.7) into the score function (2.8) we will have

\[
U(\gamma) = \sum_{i=1}^n \int_{t_1}^{t_2} W_i(t) \{ Z_i(t) - \tilde{Z}_x(t)X_i(t) \} \{ Y_i(t) - X_i^T(t)\hat{Y}_x(t)
\]

\[
-\tilde{Z}_x^T(t)\gamma - Z_i^T(t)\gamma \} \, dN_i^c(t) \gg_R
\]

\[
\]

\[
= \sum_{i=1}^n \int_{t_1}^{t_2} W_i(t) \{ Z_i(t) - \tilde{Z}_x(t)X_i(t) \} \{ Y_i(t) - X_i^T(t)\hat{Y}_x(t) + (X_i^T(t)\tilde{Z}_x(t)
\]

\[
-\tilde{Z}_x^T(t)\gamma \} \, dN_i^c(t) \gg_R .
\]
Then take the partial derivative with respect to $\gamma$, we get

$$n^{-1} \frac{\partial U(\gamma^*)}{\partial \gamma^T} = -n^{-1} \sum_{i=1}^{n} \int_{t_1}^{t_2} \ll W_i(t) \{Z_i(t) - \tilde{Z}_x(t)X_i(t)\} \otimes_2 dN_i^c(t) \gg_R. \quad (A.9)$$

According to the similar argument we discussed in the proof of consistency of $\tilde{\gamma}_i$, $\tilde{Z}_x(t)$ and $W_i(t)$ can be replaced by their limits $z_x(t)$ and $w(t)$ respectively, and this change only contributes a $o_p(1)$ difference to the above equation. Thus by Lemma A.2.2

$$n^{-1} \frac{\partial U(\gamma^*)}{\partial \gamma^T} = -n^{-1} \sum_{i=1}^{n} \ll \int_{t_1}^{t_2} w(t) \{Z_i(t) - z_x(t)X_i(t)\} \otimes_2 dN_i^c(t) \gg_R + o_p(1)$$

$$\quad \xrightarrow{p} -E \left( \int_{t_1}^{t_2} w(t) \{Z_i(t) - z_x(t)X_i(t)\} \otimes_2 dN_i^c(t) \right) = -D.$$

Now we define $B(t) = \int_{t_1}^{t} \beta_0(s)ds$ and a mean zero process

$$M_i(t; B, \gamma, \alpha) = \int_{t_1}^{t} \{Y_i(s) - \gamma^T Z_i(s)\} dN_i^c(s) - \xi_i(s)\alpha_i(s)X_i^T(s)dB(s). \quad (A.10)$$

For simplicity, we use $M_i(t) = M_i(t; B, \gamma_0, \alpha)$. Also let $O_i(t) = N_i^c(t) - \int_{0}^{t} \xi_i(s)\alpha_i(s)ds$.

Hence

$$n^{-1/2} U(\gamma_0) = n^{-1/2} \sum_{i=1}^{n} \int_{t_1}^{t_2} W_i(t) \{Z_i(t) - \tilde{Z}_x(t)X_i(t)\} \{Y_i(t) - X_i^T(t)\tilde{\beta}(t; \gamma_0)$$

$$- Z_i^T(t)\gamma_0\} dN_i^c(t) \gg_R$$

$$= n^{-1/2} \sum_{i=1}^{n} \int_{t_1}^{t_2} W_i(t) \{Z_i(t) - \tilde{Z}_x(t)X_i(t)\} \{dM_i(t)$$

$$+ \xi_i(t)\alpha_i(t)X_i^T(t)dB(t)\} \gg_R$$

$$- n^{-1/2} \sum_{i=1}^{n} \int_{t_1}^{t_2} W_i(t) \{Z_i(t) - \tilde{Z}_x(t)X_i(t)\}X_i^T(t)\tilde{\beta}(t; \gamma_0) dN_i^c(t) \gg_R$$

$$= n^{-1/2} \sum_{i=1}^{n} \int_{t_1}^{t_2} W_i(t) \{Z_i(t) - \tilde{Z}_x(t)X_i(t)\} \{dM_i(t)$$

$$- \beta_0^T(t)X_i(t)dO_i(t)\} \gg_R$$

$$- n^{-1/2} \sum_{i=1}^{n} \int_{t_1}^{t_2} W_i(t) \{Z_i(t) - \tilde{Z}_x(t)X_i(t)\} \{\tilde{\beta}^T(t; \gamma_0)$$

$$- \beta_0^T(t)\}X_i(t) dN_i^c(t) \gg_R$$

$$+ n^{-1/2} \sum_{i=1}^{n} \int_{t_1}^{t_2} W_i(t) \{Z_i(t) - \tilde{Z}_x(t)X_i(t)\} \{\beta_0^T(t)X_i^T(t)dO_i(t)$$

$$- \beta_0^T(t)\}X_i(t) dN_i^c(t) \gg_R.$$
\[ + \xi(t) \alpha_i(t) X_i^T(t) \beta(t) dt - \beta_0^T(t) X_i^T(t) dN_i^c(t) \] \[ \gg_R \]

By the definition of \( O_i(t) \), the third term above is equal to zero. Let \( \eta \) be the second term. Hence

\[ \eta = n^{-1/2} \sum_{i=1}^{n} \int_{t_1}^{t_2} \ll W_i(t) \{ Z_i(t) - \tilde{Z}_x(t) X_i(t) \} \{ \tilde{\beta}^T(t; \gamma_0) - \beta_0^T(t) \} X_i(t) dN_i^c(t) \gg_R \]

\[ = n^{-1/2} \sum_{i=1}^{n} \int_{t_1}^{t_2} \ll W_i(t) \{ Z_i(t) - \tilde{Z}_x(t) X_i(t) \} \{ \tilde{\beta}(t; \gamma_0) - \beta_0(t) \} \}

\[ + \xi(t) \alpha_i(t) dt \gg_R \]

\[ = n^{-1/2} \sum_{i=1}^{n} \int_{t_1}^{t_2} \ll W_i(t) \{ Z_i(t) - \tilde{Z}_x(t) X_i(t) \} X_i^T(t) \{ \tilde{\beta}(t; \gamma_0) - \beta_0(t) \} dO_i(t) \gg_R . \]

Denote

\[ \eta_1 = n^{-1/2} \sum_{i=1}^{n} \int_{t_1}^{t_2} \ll W_i(t) \{ Z_i(t) - \tilde{Z}_x(t) X_i(t) \} X_i^T(t) \}

\[ \{ \tilde{\beta}(t; \gamma_0) - \beta_0(t) \} dt \gg_R , \]

\[ \eta_2 = n^{-1/2} \sum_{i=1}^{n} \int_{t_1}^{t_2} \ll W_i(t) \{ \xi(t) \alpha_i(t) Z_i(t) - \tilde{Z}_x(t) \xi(t) \alpha_i(t) X_i(t) X_i^T(t) \}

\[ \{ \tilde{\beta}(t; \gamma_0) - \beta_0(t) \} dO_i(t) \gg_R . \]

In the following statement we will prove that both terms converge to zero in probability.

\[ \eta_1 = n^{-1/2} \sum_{i=1}^{n} \int_{t_1}^{t_2} \ll W_i(t) \{ \xi(t) \alpha_i(t) Z_i(t) - \tilde{Z}_x(t) \xi(t) \alpha_i(t) X_i(t) X_i^T(t) \}

\[ \{ \tilde{\beta}(t; \gamma_0) - \beta_0(t) \} dt \gg_R \]

\[ - n^{-1/2} \sum_{i=1}^{n} \int_{t_1}^{t_2} \ll W_i(t) ( \tilde{Z}_x(t) - \xi(t) \alpha_i(t) X_i(t) X_i^T(t) \{ \tilde{\beta}(t; \gamma_0) - \beta_0(t) \} dt \gg_R \]

\[ - \beta_0(t) dt \gg_R . \]
By the $\mathcal{X}$-measurability of the random functions $\tilde{\beta}(\cdot; \gamma_0)$, $\tilde{Z}_x(\cdot)$, $X_i(\cdot)$, $Z_i(\cdot)$ $R_i$ and $\xi_i(\cdot)$, then

$$\eta = n^{-1/2} \sum_{i=1}^{n} \int_{t_1}^{t_2} R_i W(t) \{\xi_i(t)\alpha_i(t)Z_i(t)X_i^T(t) - z_x(t)\xi_i(t)\alpha_i(t)X_i(t)X_i^T(t)\} \{\tilde{\beta}(t; \gamma_0) - \beta_0(t)\} dt$$

$$+ n^{-1/2} \sum_{i=1}^{n} \int_{t_1}^{t_2} (1 - R_i) \tilde{E}_s W_i(t) \{\xi_i(t)\alpha_i(t)Z_i(t)X_i^T(t) - z_x(t)\xi_i(t)\alpha_i(t)X_i(t)X_i^T(t)\} \{\tilde{\beta}(t; \gamma_0) - \beta_0(t)\} dt$$

$$- n^{-1/2} \sum_{i=1}^{n} \int_{t_1}^{t_2} R_i W(t)(\tilde{Z}_x(t) - z_x(t))\xi_i(t)\alpha_i(t)X_i(t)X_i^T(t)(\tilde{\beta}(t; \gamma_0) - \beta_0(t)) dt$$

$$- n^{-1/2} \sum_{i=1}^{n} \int_{t_1}^{t_2} (1 - R_i) \tilde{E}_s W_i(t)(\tilde{Z}_x(t) - z_x(t))\xi_i(t)\alpha_i(t)X_i(t)X_i^T(t)(\tilde{\beta}(t; \gamma_0) - \beta_0(t)) dt$$

$$= n^{-1/2} \sum_{i=1}^{n} \int_{t_1}^{t_2} W_i(t) \{\xi_i(t)\alpha_i(t)Z_i(t)X_i^T(t) - z_x(t)\xi_i(t)\alpha_i(t)X_i(t)X_i^T(t)\} \{\tilde{\beta}(t; \gamma_0) - \beta_0(t)\} dt$$

$$- n^{-1/2} \sum_{i=1}^{n} \int_{t_1}^{t_2} W_i(t)(\tilde{Z}_x(t) - z_x(t))\xi_i(t)\alpha_i(t)X_i(t)X_i^T(t)(\tilde{\beta}(t; \gamma_0) - \beta_0(t)) dt$$

$$= \int_{t_1}^{t_2} W_i(t) n^{-1} \sum_{i=1}^{n} \{\xi_i(t)\alpha_i(t)Z_i(t)X_i^T(t) - z_x(t)\xi_i(t)\alpha_i(t)X_i(t)X_i^T(t)\} d\left(n^{1/2} \int_{t_1}^{t} (\tilde{\beta}(s; \gamma_0) - \beta_0(s)) ds\right)$$
\[-\int_{t_1}^{t_2} W_i(t)(\tilde{Z}_x(t) - z_x(t))n^{-1} \sum_{i=1}^{n} \xi_i(t)\alpha_i(t)X_i(t)X_i^T(t) \left(n^{1/2} \int_{t_1}^{t} (\tilde{\beta}(s; \gamma_0) - \beta_0(s))ds \right).\]

By the consistency of the \(\tilde{Z}_x(t)\), the convergence of \(W_i(t)\), the application of Lemma A.2.5 and Lemma A.1 of Lin & Ying (2001), and the facts that

\[n^{-1} \sum_{i=1}^{n} \{\xi_i(t)\alpha_i(t)Z_i(t)X_i^T(t) - z_x(t)\xi_i(t)\alpha_i(t)X_i(t)X_i^T(t)\} \rightarrow P E\{\xi_i(t)\alpha_i(t)Z_i(t)X_i^T(t) - z_x(t)\xi_i(t)\alpha_i(t)X_i(t)X_i^T(t)\} = \]

\[E\{\xi_i(t)\alpha_i(t)Z_i(t)X_i^T(t)\} - z_x(t)E\{\xi_i(t)\alpha_i(t)X_i(t)X_i^T(t)\} =
\]

\[e_{xx}(t) - z_x(t)e_{xx} = e_{xx}(t) - e_{xx}(t)(e_{xx}(t))^{-1} = 0\]

and

\[n^{-1} \sum_{i=1}^{n} \xi_i(t)\alpha_i(t)X_i(t)X_i^T(t) \rightarrow P E\{\xi_i(t)\alpha_i(t)X_i(t)X_i^T(t)\} = e_{xx}(t),\]

we have \(\eta_1 \rightarrow P 0\).

\[\eta_2 = n^{-1/2} \sum_{i=1}^{n} \int_{t_1}^{t_2} R_i W_i(t)\{Z_i(t) - \tilde{Z}_x(t)X_i(t)\}X_i^T(t)\{\tilde{\beta}(t; \gamma_0) - \beta_0(t)\}dO_i(t)\]

\[+ n^{-1/2} \sum_{i=1}^{n} \int_{t_1}^{t_2} (1 - R_i) \tilde{E}_s\{W_i(t)\{Z_i(t) - \tilde{Z}_x(t)X_i(t)\}X_i^T(t)\{\tilde{\beta}(t; \gamma_0) - \beta_0(t)\}\}
\]

\[+ \tilde{\eta}_2 = \sum_{i=1}^{n} \int_{t_1}^{t_2} [R_i W_i(t)\{Z_i(t) - \tilde{Z}_x(t)X_i(t)\}X_i^T(t)dO_i(t)\{\tilde{\beta}(t; \gamma_0) - \beta_0(t)\}]
\]

\[+ n^{-1/2} \sum_{i=1}^{n} \int_{t_1}^{t_2} [(1 - R_i) \tilde{E}_s\{W_i(t)\{Z_i(t) - \tilde{Z}_x(t)X_i(t)\}X_i^T(t)\}dO_i(t) | \mathcal{X}]
\]

\[\{\tilde{\beta}(t; \gamma_0) - \beta_0(t)\}].\]

The first term of \(\eta_2\)

\[n^{-1/2} \sum_{i=1}^{n} \int_{t_1}^{t_2} [R_i W_i(t)\{Z_i(t) - \tilde{Z}_x(t)X_i(t)\}X_i^T(t)dO_i(t)\{\tilde{\beta}(t; \gamma_0) - \beta_0(t)\}]\]
\[= \int_{t_1}^{t_2} W_i(t)n^{-1/2} \sum_{i=1}^{n} R_i \{ Z_i(t) - \tilde{Z}_x(t)X_i(t) \} X_i^T(t)dO_i(t) \{ \tilde{\beta}(t; \gamma_0) - \beta_0(t) \} \]

\[= \int_{t_1}^{t_2} W_i(t)n^{-1/2} \sum_{i=1}^{n} R_i \{ Z_i(t) - z_x(t)X_i(t) \} X_i^T(t)dO_i(t) \{ \tilde{\beta}(t; \gamma_0) - \beta_0(t) \} \]

\[= \int_{t_1}^{t_2} d \left( n^{-1/2} \sum_{i=1}^{n} \int_{t_1}^{t} R_i \{ Z_i(s) - z_x(s)X_i(s) \} X_i^T(s)dO_i(s) \right) W_i(t) \{ \tilde{\beta}(t; \gamma_0) - \beta_0(t) \} \]

Under the condition (I) and by Lemma 1 of Sun & Wu (2005), both

\[n^{-1/2} \sum_{i=1}^{n} \int_{t_1}^{t} R_i \{ Z_i(s) - z_x(s)X_i(s) \} X_i^T(s)dO_i(s)\]

and

\[n^{-1/2} \sum_{i=1}^{n} \int_{t_1}^{t} R_i X_i(s)X_i^T(s)dO_i(s)\]

converge weakly to vectors of mean zero Gaussian processes with continuous sample paths respectively. And from the early derivation, \(W_i(t) \{ \tilde{\beta}(t; \gamma_0) - \beta_0(t) \} \) and \(\tilde{Z}_x(t) - z_x(t)\) are of bounded variations and both converge to zero in probability uniformly in \(t\). Hence by Lemma A.1 of Lin & Ying (2001), the first term converges to zero in probability.

As the second term of \(\eta_2\)

\[n^{-1/2} \sum_{i=1}^{n} \int_{t_1}^{t_2} \left[ (1 - R_i) \tilde{E}_s \{ W_i(t) \{ Z_i(t) - \tilde{Z}_x(t)X_i(t) \} X_i^T(t)dO_i(t) \mid \mathcal{X} \} \right] \{ \tilde{\beta}(t; \gamma_0) - \beta_0(t) \} \]

\[= n^{-1/2} \sum_{i=1}^{n} \int_{t_1}^{t_2} \left[ (1 - R_i) \tilde{E}_s \{ W_i(t) \{ Z_i(t) - z_x(t)X_i(t) \} X_i^T(t)dO_i(t) \mid \mathcal{X} \} \right] \{ \tilde{\beta}(t; \gamma_0) - \beta_0(t) \} \]
\[-n^{-1/2} \sum_{i=1}^{n} \int_{t_1}^{t_2} \{(1 - R_i) \hat{E}_s \{W_i(t) \{ \tilde{Z}_x(t) - z_x(t) \} X_i(t) X^T_i(t) dO_i(t) \mid X \} \{\tilde{\beta}(t; \gamma_0) - \beta_0(t)\} \}
\]

\[= n^{-1/2} \sum_{i=1}^{n} \int_{t_1}^{t_2} \{(1 - R_i) \hat{E}_s \{W_i(t) \{ Z_i(t) - z_x(t)X_i(t) \} X^T_i(t) dO_i(t) \mid X_i \} \{\tilde{\beta}(t; \gamma_0) - \beta_0(t)\} \}
\]

\[-\hat{E}_s \{n^{-1/2} \sum_{i=1}^{n} \int_{t_1}^{t_2} W_i(t)(1 - R_i) \{ \tilde{Z}_x(t) - z_x(t) \} X_i(t) X^T_i(t) dO_i(t) \{\tilde{\beta}(t; \gamma_0) - \beta_0(t)\} \mid X \}
\]

\[= n^{-1/2} \sum_{i=1}^{n} \int_{t_1}^{t_2} \{(1 - R_i) \hat{E}_s \{W_i(t) \{ Z_i(t) - z_x(t)X_i(t) \} X^T_i(t) dO_i(t) \mid D_i, R_i = 0\} \}
\]

\[\{\tilde{\beta}(t; \gamma_0) - \beta_0(t)\}
\]

\[-\hat{E}_s \{\int_{t_1}^{t_2} W_i(t) \{ \tilde{Z}_x(t) - z_x(t) \} d\left(n^{-1/2} \sum_{i=1}^{n} \int_{t_1}^{t} (1 - R_i) X_i(u) X^T_i(u) dO_i(u) \right) \{\tilde{\beta}(t; \gamma_0) - \beta_0(t)\} \mid X \},
\]

also by Lemma 1 of Sun & Wu (2005) \(n^{-1/2} \sum_{i=1}^{n} \int_{t_1}^{t} (1 - R_i) X_i(u) X^T_i(u) dO_i(u)\) converges weakly to a vector of mean zero Gaussian processes with continuous sample paths. Then from the early derivation, \(W_i(t) \{\tilde{\beta}(t; \gamma_0) - \beta_0(t)\}\) is of bounded variations and converges to zero in probability uniformly in \(t\). Hence by Lemma A.1 of Lin & Ying (2001),

\[\int_{t_1}^{t_2} W_i(t) \{ \tilde{Z}_x(t) - z_x(t) \} d\left(n^{-1/2} \sum_{i=1}^{n} \int_{t_1}^{t} (1 - R_i) X_i(u) X^T_i(u) dO_i(u) \right) \xrightarrow{P} 0.
\]

Also using the similar argument in Lemma A.2.1, the second term of \(\eta_2\) equals to

\[n^{-1/2} \sum_{i=1}^{n} \int_{t_1}^{t_2} \{(1 - R_i) E_s \{W_i(t) \{ Z_i(t) - z_x(t)X_i(t) \} X^T_i(t) dO_i(t) \mid D_i, R_i = 0\} \}
\]

\[\{\tilde{\beta}(t; \gamma_0) - \beta_0(t)\} + o_p(1)
\]

\[= \int_{t_1}^{t_2} \left[n^{-1/2} \sum_{i=1}^{n} (1 - R_i) \{ Z_i(t) - z_x(t)X_i(t) \} X^T_i(t) E_s \{dO_i(t) \mid D_i, R_i = 0\} W_i(t) \right]
\]

\[\{\tilde{\beta}(t; \gamma_0) - \beta_0(t)\} + o_p(1)
\]

\[= \int_{t_1}^{t_2} \left[d\left(n^{-1/2} \sum_{i=1}^{n} \int_{t_1}^{t} (1 - R_i) \{ Z_i(u) - z_x(u)X_i(u) \} X^T_i(u) E_s \{dO_i(u) \mid D_i, R_i = 0\} \right) \right]
\]
\[ R_i = 0 \) \( W_i(t) \{ \tilde{\beta}(t; \gamma_0) - \beta_0(t) \} \] + o_p(1)

Now apply Lemma 1 of Sun & Wu (2005) again.

\[
n^{-1/2} \sum_{i=1}^{n} \int_{t_i}^{t} (1 - R_i) \{ Z_i(u) - z_x(u)X_i(u) \} X_i^T(u)E_s \{ dO_i(u) \mid D_i, R_i = 0 \}
\]

converges weakly to a vector of mean zero Gaussian processes with continuous sample paths. Also from the early derivation, \( W_i(t) \{ \tilde{\beta}(t; \gamma_0) - \beta_0(t) \} \) is of bounded variations and converges to zero in probability uniformly in \( t \). Hence by Lemma A.1 of Lin & Ying (2001) the second term of \( \eta_2 \overset{p}{\longrightarrow} 0 \). Then \( \eta = \eta_1 + \eta_2 \overset{p}{\longrightarrow} 0 \). Thus \( n^{-1/2} U(\gamma_0) \) equals

\[
n^{-1/2} \sum_{i=1}^{n} \int_{t_i}^{t_2} \ll W_i(t) \{ Z_i(t) - \tilde{Z}_x(t)X_i(t) \} \{ dM_i(t) - \tilde{\beta}_0^T(t)X_i(t) dO_i(t) \} \gg_R.
\]

Since

\[
dM_i(t) - \tilde{\beta}_0^T(t)X_i(t) dO_i(t)
= \left[ Y_i(t) - \gamma_0^T Z_i(t) \right] dN_i^e(t) - \xi_i(t) \alpha_i(t) X_i^T(t) dB(t) - \beta_0^T(t)X_i(t) dN_i^c(t)
+ \beta_0^T(t)X_i(t)\xi_i(t)\alpha_i(t) dt
= \left[ Y_i(t) - \gamma_0^T Z_i(t) - \beta_0^T(t)X_i(t) \right] dN_i^e(t) - \xi_i(t) \alpha_i(t) X_i^T(t) \beta_0(t) d(t)
+ \beta_0^T(t)X_i(t)\xi_i(t)\alpha_i(t) dt
= \epsilon_i(t) dN_i^c(t),
\]

\[
n^{-1/2} U(\gamma_0) = n^{-1/2} \sum_{i=1}^{n} \int_{t_i}^{t_2} \ll W_i(t) \{ Z_i(t) - \tilde{Z}_x(t)X_i(t) \} \epsilon_i(t) dN_i^c(t) \gg_R
= n^{-1/2} \sum_{i=1}^{n} \int_{t_i}^{t_2} R_i W(t) \{ Z_i(t) - \tilde{Z}_x(t)X_i(t) \} \epsilon_i(t) dN_i^c(t)\ (A.11)
+ n^{-1/2} \sum_{i=1}^{n} \int_{t_i}^{t_2} (1 - R_i) \tilde{E}_s \{ W_i(t) \{ Z_i(t) - \tilde{Z}_x(t)X_i(t) \} \epsilon_i(t) dN_i^c(t) \mid X \}.\ (A.12)
\]
\begin{align}
&= n^{-1/2} \sum_{i=1}^{n} \int_{t_1}^{t_2} R_i W(t) \{Z_i(t) - z_x(t)X_i(t)\} \epsilon_i(t) dN^c_i(t) \\
&- n^{-1/2} \sum_{i=1}^{n} \int_{t_1}^{t_2} R_i W(t) \{\tilde{Z}_x(t) - z_x(t)\} X_i(t) \epsilon_i(t) dN^c_i(t) \\
&= n^{-1/2} \sum_{i=1}^{n} \int_{t_1}^{t_2} R_i W(t) \{Z_i(t) - z_x(t)X_i(t)\} \epsilon_i(t) dN^c_i(t) \\
&- \int_{t_1}^{t_2} W_i(t) \{\tilde{Z}_x(t) - z_x(t)\} n^{-1/2} \sum_{i=1}^{n} R_i X_i(t) [dM_i(t) - \beta_0^T(t)X_i(t)dO_i(t)] \\
&= n^{-1/2} \sum_{i=1}^{n} \int_{t_1}^{t_2} R_i W(t) \{Z_i(t) - z_x(t)X_i(t)\} \epsilon_i(t) dN^c_i(t) \\
&- \int_{t_1}^{t_2} W_i(t) \{\tilde{Z}_x(t) - z_x(t)\} \left( n^{-1/2} \sum_{i=1}^{n} \int_{t_1}^{t} R_i X_i(u) dM_i(u) \right) \\
&+ \int_{t_1}^{t_2} W_i(t) \{\tilde{Z}_x(t) - z_x(t)\} \left( n^{-1/2} \sum_{i=1}^{n} \int_{t_1}^{t} R_i X_i(t) X_i^T(t) dO_i(t) \right) \beta_0(t) \\
&= n^{-1/2} \sum_{i=1}^{n} \int_{t_1}^{t_2} R_i W(t) \{Z_i(t) - z_x(t)X_i(t)\} \epsilon_i(t) dN^c_i(t) + o_p(1).
\end{align}

The last equality holds because of the joint weak convergence of

\[
\left( n^{-1/2} \sum_{i=1}^{n} \int_{t_1}^{t} R_i X_i(u) dM_i(u), \ n^{-1/2} \sum_{i=1}^{n} \int_{t_1}^{t} R_i X_i(u) X_i^T(u) dO_i(u) \right)
\]

by Lemma 1 of Sun & Wu (2005), the consistency of \(W_i(t) \{\tilde{Z}_x(t) - z_x(t)\}\) and Lemma A.1 of Lin & Ying (2001).

\begin{align}
&= n^{-1/2} \sum_{i=1}^{n} \int_{t_1}^{t_2} (1 - R_i) \tilde{E}_s \{W_i(t) \{Z_i(t) - \tilde{Z}_x(t)X_i(t)\} \epsilon_i(t) dN^c_i(t) | \mathcal{X} \} \\
&= n^{-1/2} \sum_{i=1}^{n} \int_{t_1}^{t_2} (1 - R_i) \tilde{E}_s \{W_i(t) \{Z_i(t) - z_x(t)X_i(t)\} \epsilon_i(t) dN^c_i(t) | \mathcal{X} \}
\end{align}
by Lemma 1 of Sun & Wu (2005), the consistency of
The last equality holds also because of the weak convergence of
A.1 of Lin & Ying (2001). Similarly the
\[ \sum_{i=1}^{n} \int_{t_1}^{t_2} (1 - R_i) \hat{E}_s \{ W_i(t) \{ \tilde{Z}_x(t) - z_x(t) \} X_i(t) \} dM_i(t) \]
\[ - \beta_0^T(t) X_i(t) dO_i(t) \mid \mathcal{X} \}
\[ = \sum_{i=1}^{n} \int_{t_1}^{t_2} (1 - R_i) \hat{E}_s \{ W_i(t) \{ Z_i(t) - z_x(t) X_i(t) \} \epsilon_i(t) \} dN_i^c(t) \mid \mathcal{X} \}
\[ = \sum_{i=1}^{n} \int_{t_1}^{t_2} (1 - R_i) \hat{E}_s \{ W_i(t) \{ Z_i(t) - z_x(t) X_i(t) \} \epsilon_i(t) \} dN_i^c(t) \mid \mathcal{D}_i, \]
\[ R_i = 0 \} + o_p(1). \]

The last equality holds also because of the weak convergence of
\[ n^{-1/2} \sum_{i=1}^{n} \int_{t_1}^{t} (1 - R_i) X_i(u) dM_i(u) \] and
\[ n^{-1/2} \sum_{i=1}^{n} \int_{t_1}^{t} (1 - R_i) X_i(u) X_i^T(u) dO_i(u) \]
by Lemma 1 of Sun & Wu (2005), the consistency of \( W_i(t) \{ \tilde{Z}_x(t) - z_x(t) \} \) and Lemma A.1 of Lin & Ying (2001). Similarly the \( W_i(t) \) can be replaced by its limit \( w(t) \). Then

\[ n^{-1/2} \sum_{i=1}^{n} \int_{t_1}^{t_2} (1 - R_i) E_s \{ w(t) \{ Z_i(t) - z_x(t) X_i(t) \} \epsilon_i(t) \} dN_i^c(t) \mid \mathcal{D}_i, \]
\[ R_i = 0 \} \]
\[ + n^{-1/2} \sum_{i=1}^{n} \int_{t_1}^{t_2} (1 - R_i) \hat{E}_s \{ w(t) \{ Z_i(t) - z_x(t) X_i(t) \} \epsilon_i(t) \} dN_i^c(t) \mid \mathcal{D}_i, \]
\[ R_i = 0 \} \]
\[ - n^{-1/2} \sum_{i=1}^{n} \int_{t_1}^{t_2} (1 - R_i) E_s \{ w(t) \{ Z_i(t) - z_x(t) X_i(t) \} \epsilon_i(t) \} dN_i^c(t) \mid \mathcal{D}_i, \]
\[ R_i = 0 \} + o_p(1) \]
\[ = n^{-1/2} \sum_{i=1}^{n} \int_{t_1}^{t_2} (1 - R_i) E_s \{ w(t) \{ Z_i(t) - z_x(t) X_i(t) \} \epsilon_i(t) \} dN_i^c(t) \mid \mathcal{D}_i, \]
\[ R_i = 0 \} \]
\begin{align*}
&+n^{-1/2} \sum_{i=1}^{n} (1 - R_i) \int_0^L \sum_{j=1}^{n_i} I(t_1 \leq s + T_{ij} \leq t_2) w(s + T_{ij}) \{Z_{ij} - z_x(s + T_{ij})X_{ij}\} \epsilon_i(s + T_{ij}) I(C_i \geq T_{ij}) \left[ \frac{d\hat{F}_s(s)}{F_s(V_i)} - \frac{dF_s(s)}{F_s(V_i)} \right] \\
&+o_p(1)
\end{align*}

Referring to the argument in Lemma A.2.4, (A.13) has the following decomposition.

\begin{align*}
(A.13) &= n^{-1/2} \int_0^\infty \int_0^L \int_{t_1}^{t_2} \left\{ I(x < (L - (V_i))) \right\} \frac{dM^R(x)}{y^R(x)} dF_s(s) \frac{dN^*_x(v-s)}{F_s(V_i)} \\
&+n^{-1/2} \int_0^L \int_{0}^{(L-x)^-} E \left\{ \int_{t_1}^{t_2} (1 - R_i) w(v)(Z_i(v) - z_x(v)X_i(v)) \epsilon_i(v) \frac{dN^*_x(v-s)}{F_s(V_i)} \right\} \\
&+n^{-1/2} \int_0^L E \left\{ \int_{t_1}^{t_2} (1 - R_i) w(v)(Z_i(v) - z_x(v)X_i(v)) \epsilon_i(v) \frac{dN^*_x(v-s)}{F_s(V_i)} \right\} \\
&= n^{-1/2} \int_0^\infty \int_0^L E \left\{ \int_{t_1}^{t_2} (1 - R_i) w(v)(Z_i(v) - z_x(v)X_i(v)) \epsilon_i(v) \frac{dN^*_x(v-s)}{F_s(V_i)} \right\} \\
&+n^{-1/2} \int_0^L \int_{0}^{(L-x)^-} E \left\{ \int_{t_1}^{t_2} (1 - R_i) w(v)(Z_i(v) - z_x(v)X_i(v)) \epsilon_i(v) \frac{dN^*_x(v-s)}{F_s(V_i)} \right\} \\
&+n^{-1/2} \int_0^L E \left\{ \int_{t_1}^{t_2} (1 - R_i) w(v)(Z_i(v) - z_x(v)X_i(v)) \epsilon_i(v) \frac{dN^*_x(v-s)}{F_s(V_i)} \right\} \\
&= n^{-1/2} \int_0^\infty \int_0^L E \left\{ \int_{t_1}^{t_2} (1 - R_i) w(v)(Z_i(v) - z_x(v)X_i(v)) E[\epsilon_i(v) | X_i(\cdot), Z_i(\cdot), N_i(\cdot), S_i, V_i, C_i] \frac{dN^*_x(v-s)}{F_s(V_i)} \right\} dF_s(s)
\end{align*}
\[
\begin{align*}
\frac{dM^R(x)}{y^R(x)} & + n^{-1/2} \int_0^L \int_0^{(L-x)-} E \left\{ \int_{t_1}^{t_2} (1 - R_s) w(v)(Z_i(v) - z_x(v)X_i(v))E[\epsilon_i(v) X_i(\cdot), Z_i(\cdot), N_i(\cdot), S_i, V_i, C_i] \frac{dN_i^s(v - s)}{F_s(V_i)} \right\} dF_s(s) F^R_s(x) \\
+ n^{-1/2} \int_0^L E \left\{ \int_{t_1}^{t_2} (1 - R_s) w(v)(Z_i(v) - z_x(v)X_i(v))E[\epsilon_i(v) X_i(\cdot), Z_i(\cdot), N_i(\cdot), S_i, V_i, C_i] \frac{dN_i^s(v - s)}{F_s(V_i)} \right\} \frac{dM^R(L - s) - y^R(L - s)}{y^R(x)} + o_p(1).
\end{align*}
\]

Under the assumption that \(E\{Y_i(t)|X_i(\cdot), Z_i(\cdot), N_i(\cdot), S_i, V_i, C_i\} = E\{Y_i(t)|X_i(\cdot), Z_i(\cdot)\}\),

\[
E[\epsilon_i(v) | X_i(\cdot), Z_i(\cdot), N_i(\cdot), S_i, V_i, C_i] = E[\epsilon_i(v) | X_i(\cdot), Z_i(\cdot)] = 0.
\]

Then (A.13) = \(0 + o_p(1) \xrightarrow{P} 0.\) Hence

\[
n^{-1/2}U(\gamma_0) = n^{-1/2} \sum_{i=1}^{n} \int_{t_1}^{t_2} R_i w(t)\{Z_i(t) - z_x(t)X_i(t)\}\epsilon_i(t)dN_i^c(t) \\
+ n^{-1/2} \sum_{i=1}^{n} \int_{t_1}^{t_2} (1 - R_i) E_s\{w(t)\{Z_i(t) - z_x(t)X_i(t)\}\epsilon_i(t)dN_i^c(t) | D_i, R_i = 0\} + o_p(1)
\]

\[
= n^{-1/2} \sum_{i=1}^{n} \int_{t_1}^{t_2} w(t)\{Z_i(t) - z_x(t)X_i(t)\}\epsilon_i(t)R_i dN_i^c(t) \\
+ n^{-1/2} \sum_{i=1}^{n} \int_{t_1}^{t_2} w(t)\{Z_i(t) - z_x(t)X_i(t)\}\epsilon_i(t)E_s\{(1 - R_i)dN_i^c(t) | D_i, R_i = 0\} + o_p(1)
\]

\[
= n^{-1/2} \sum_{i=1}^{n} \int_{t_1}^{t_2} w(t)\{Z_i(t) - z_x(t)X_i(t)\}\epsilon_i(t)[R_i dN_i^c(t) \\
+ E_s\{(1 - R_i)dN_i^c(t) | D_i, R_i = 0\}] + o_p(1).
\]

Applying theorem 5.21 (van der Vaart, 1998) to the score function, (A.8) becomes

\[
n^{1/2}(\hat{\gamma} - \gamma_0) = D^{-1}[n^{-1/2}U(\gamma_0)] + o_p(1).
\]

Hence \(n^{1/2}(\hat{\gamma} - \gamma_0) \xrightarrow{D} \mathcal{N}(0, D^{-1}VD^{-1}).\) \(\Box\)
Proof of Theorem 3.4

By the definitions, we have

\[
\bar{\beta}(t; \gamma_0) - \beta^*(t) = \bar{Y}_x^T(t) - \bar{Z}_x^T(t)\gamma_0 + [\bar{g}_x^T(t) - \bar{z}_x^T(t)\gamma_0]
\]

\[
= (\bar{\bar{E}}_{xx}(t))^{-1}\bar{E}_{xy}(t) - (\bar{\bar{E}}_{xx}(t))^{-1}\bar{E}_{xz}(t)\gamma_0 - (\bar{e}_{xx}(t))^{-1}\bar{e}_{xy}(t)
\]

\[
+ (\bar{e}_{xx}(t))^{-1}\bar{e}_{xz}(t)\gamma_0
\]

\[
= (\bar{\bar{E}}_{xx}(t))^{-1}[(\bar{E}_{xy}(t) - \bar{e}_{xy}(t)) - (\bar{E}_{xz}(t) - \bar{e}_{xz}(t))\gamma_0]
\]

\[
- (\bar{e}_{xx}(t))^{-1}[\bar{E}_{xx}(t) - \bar{e}_{xx}(t)](\bar{E}_{xx}(t))^{-1}[\bar{e}_{xy}(t) - \bar{e}_{xz}(t)\gamma_0]
\]

\[
= (e_{xx}(t))^{-1}[(\bar{E}_{xy}(t) - \bar{e}_{xy}(t)) - (\bar{E}_{xz}(t) - \bar{e}_{xz}(t))\gamma_0]
\]

\[
- (e_{xx}(t))^{-1}[\bar{E}_{xx}(t) - \bar{e}_{xx}(t)](e_{xx}(t))^{-1}[e_{xy}(t) - e_{xz}(t)\gamma_0] + o_p(1).
\]

The last equality holds by Slutsky’s theorem. Then

\[
\bar{\beta}(t; \gamma_0) - \beta^*(t)
\]

\[
= (e_{xx}(t))^{-1}[(\bar{E}_{xy}(t) - \bar{e}_{xy}(t)) - (\bar{E}_{xz}(t) - \bar{e}_{xz}(t))\gamma_0]
\]

\[
- (e_{xx}(t))^{-1}[(\bar{E}_{xy}(t) - \bar{e}_{xy}(t)) - (\bar{E}_{xz}(t) - \bar{e}_{xz}(t))\gamma_0]
\]

\[
+ o_p(1)
\]

\[
= (e_{xx}(t))^{-1}\left( n^{-1} \sum_{i=1}^{n} R_i \int_{0}^{\tau} K_h(u - t)X_i(u)|Y_i(u) - Z_i^T(u)\gamma_0
\]

\[
- X_i^T(u)\beta_0(u)]dN_i^e(u)
\]

\[
+ n^{-1} \sum_{i=1}^{n} (1 - R_i)\tilde{E}_a \left\{ \int_{0}^{\tau} K_h(u - t)X_i(u)|Y_i(u) - Z_i^T(u)\gamma_0
\]

\[
- X_i^T(u)\beta_0(u)]dN_i^e(u) \mid \mathcal{X} \right\}
\]

\[
- \int_{0}^{\tau} K_h(u - t)E\{\xi_i(u)\alpha_i(u)X_i(u)|Y_i(u) - Z_i^T(u)\gamma_0
\]

\[
\right).\]
\[
\begin{aligned}
-X_i^T(u)\beta_0(u)]du \\
-(e_{xx}(t))^{-1} \left( n^{-1} \sum_{i=1}^{n} R_i \int_0^t K_h(u-t)X_i(u)X_i^T(u)[\beta_0(t) - \beta_0(u)]dN_i^c(u) \\
+n^{-1} \sum_{i=1}^{n} (1 - R_i) \tilde{E}_a \left\{ \int_0^t K_h(u-t)X_i(u)[\beta_0(t) - \beta_0(u)]dN_i^c(u) | \mathcal{X} \right\} \\
- \int_0^t K_h(u-t)E\{\xi(u)\alpha_i(u)X_i(u)X_i^T(u)\}[\beta_0(t) - \beta_0(u)]du \right) + o_p(1) \\
= (e_{xx}(t))^{-1} \left( \int_0^t K_h(u-t)d \left[ n^{-1} \sum_{i=1}^{n} \int_0^u R_i X_i(w)X_i^T(w) dN_i^c(w) \right][\beta_0(t) - \beta_0(u)] \\
+n^{-1} \sum_{i=1}^{n} (1 - R_i) \tilde{E}_a \left\{ \int_0^t K_h(u-t)d \left[ n^{-1} \sum_{i=1}^{n} \int_0^u (1 - R_i) X_i(w)X_i^T(w) dN_i^c(w) \right] \\
[\beta_0(t) - \beta_0(u)] | \mathcal{X} \right\} \\
- \int_0^t K_h(u-t)E\{\xi(u)\alpha_i(u)X_i(u)X_i^T(u)\}[\beta_0(t) - \beta_0(u)]du \right) + o_p(1)
\end{aligned}
\]

We know that
\[
\int_0^t K_h(u-t)E\{\xi(u)\alpha_i(u)X_i(u)\epsilon_i(u)\}du \\
= \int_0^t K_h(u-t)E\{E[\xi(u)\alpha_i(u)X_i(u)\epsilon_i(u) | X_i(\cdot), Z_i(\cdot)]\}du \\
= \int_0^t K_h(u-t)E\{\xi(u)\alpha_i(u)X_i(u)E[\epsilon_i(u) | X_i(\cdot), Z_i(\cdot)]\}du = 0.
\]

Therefore,
\[
(nh)^{1/2}(\tilde{\beta}(t; \gamma_0) - \beta^* (t))
\]
\[
\begin{align*}
&= (nh)^{1/2}(e_{xx}(t))^{-1}(n^{-1} \sum_{i=1}^{n} \ll \int_{0}^{T} K_h(u - t)X_i(u)\epsilon_i(u) dN^c_i(u) \gg R)

&= (e_{xx}(t))^{-1}\left(\int_{0}^{T} h^{1/2}K_h(u - t)d\left[n^{-1/2} \sum_{i=1}^{n} \int_{0}^{u} R_iX_i(w)X_i^T(w)dN^c_i(w)\right] [\beta_0(t) - \beta_0(u)]\right)

&= \int_{-1}^{1} h^{1/2}K(x)d\left[n^{-1/2} \sum_{i=1}^{n} \int_{0}^{x+th} R_iX_i(w)X_i^T(w)dN^c_i(w)\right] \frac{\beta_0(t) - \beta_0(t + xh)}{h} \\
&= -\int_{-1}^{1} h^{1/2}K(x)d\left[n^{-1/2} \sum_{i=1}^{n} \int_{0}^{x+th} R_iX_i(w)X_i^T(w)dN^c_i(w)\right] [x\beta'_0(t) + O(h)] \\
&\xrightarrow{P} 0
\end{align*}
\]

since \(n^{-1/2} \sum_{i=1}^{n} \int_{0}^{x+th} R_iX_i(w)X_i^T(w)dN^c_i(w)\) converges weakly as \(h \to 0\) and \(n \to \infty\).

Similarly,
\[
\int_{0}^{T} h^{1/2}K_h(u - t)d\left[n^{-1/2} \sum_{i=1}^{n} \int_{0}^{u} (1 - R_i)X_i(w)X_i^T(w)dN^c_i(w)\right] [\beta_0(t) - \beta_0(u)] \xrightarrow{P} 0
\]
as \(h \to 0\) and \(n \to \infty\). And
\[
\begin{align*}
&= \int_{-1}^{1} (nh)^{1/2}K(x)e_{xx}(t + xh)[\beta_0(t) - \beta_0(t + xh)]dx \\
&= -\int_{-1}^{1} (nh)^{1/2}K(x)e_{xx}(t + xh)[xh\beta'_0(t) + (1/2)x^2h^2\beta''_0(t) + o(h^2)]dx
\end{align*}
\]
\[
\int_{-1}^{1} \left( (nh)^{1/2}K(x)[e_{xx}(t) + xhe'_{xx}(t) + (1/2)x^2h^2e''_{xx}(t) + o(h^2)]xh\beta'_0(t) \\
+ (1/2)x^2h^2\beta''_0(t) + o(h^2) \right) dx \\
= -(nh)^{1/2} \int_{-1}^{1} K(x)[e_{xx}(t) + x^2h^2e''_{xx}(t)\beta'_0(t) + (1/2)x^2h^2e_{xx}(t)\beta''_0(t) \\
+ o(h^2)] dx \\
= -(nh)^{1/2} \int_{-1}^{1} xK(x)dx e_{xx}(t)\beta'_0(t) - (nh^5)^{1/2}[e'_{xx}(t)\beta'_0(t) \\
+ (1/2)e_{xx}(t)\beta''_0(t)] \int_{-1}^{1} x^2K(x)dx + o_p((nh^5)^{1/2}) \\
= -0 - (nh^5)^{1/2}[e'_{xx}(t)\beta'_0(t) + (1/2)e_{xx}(t)\beta''_0(t)] \int_{-1}^{1} x^2K(x)dx + o_p((nh^5)^{1/2}) \\
= -(nh^5)^{1/2}[e'_{xx}(t)\beta'_0(t) + (1/2)e_{xx}(t)\beta''_0(t)] \int_{-1}^{1} x^2K(x)dx + o_p((nh^5)^{1/2}) \\
as nh^5 = O(1). Thus
\]

\[
(nh)^{1/2} \left( \beta(t; \gamma_0) - \beta^*(t) + h^2(e_{xx}(t))^{-1}[e'_{xx}(t)\beta'_0(t) \\
+ (1/2)e_{xx}(t)\beta''_0(t)] \int_{-1}^{1} x^2K(x)dx \right) \\
= (nh)^{1/2}(e_{xx}(t))^{-1} \left( n^{-1} \sum_{i=1}^{n} \ll \int_{0}^{T} K_h(u - t)X_i(u)\epsilon_i(u)dN^c_i(u) \gg R \right) \quad (A.14) \\
= (nh)^{1/2}(e_{xx}(t))^{-1} \left( n^{-1} \sum_{i=1}^{n} R_i \int_{0}^{T} K_h(u - t)X_i(u)\epsilon_i(u)dN^c_i(u) \\
+ n^{-1} \sum_{i=1}^{n} (1 - R_i) \hat{E}_i \left\{ \int_{0}^{T} K_h(u - t)X_i(u)\epsilon_i(u)dN^c_i(u) \mid \mathcal{X}_i \right\} \right) \\
= (nh)^{1/2}(e_{xx}(t))^{-1} \left( n^{-1} \sum_{i=1}^{n} R_i \int_{0}^{T} K_h(u - t)X_i(u)\epsilon_i(u)dN^c_i(u) \\
+ n^{-1} \sum_{i=1}^{n} (1 - R_i) \hat{E}_i \left\{ \int_{0}^{T} K_h(u - t)X_i(u)\epsilon_i(u)dN^c_i(u) \mid \mathcal{X}_i \right\} \right) \\
= (nh)^{1/2}(e_{xx}(t))^{-1} \left( n^{-1} \sum_{i=1}^{n} R_i \int_{0}^{T} K_h(u - t)X_i(u)\epsilon_i(u)dN^c_i(u) \\
+ n^{-1} \sum_{i=1}^{n} (1 - R_i) \hat{E}_i \left\{ \int_{0}^{T} K_h(u - t)X_i(u)\epsilon_i(u)dN^c_i(u) \mid \mathcal{D}_i, R_i = 0 \right\} \right) \\
= (nh)^{1/2}(e_{xx}(t))^{-1} \left( n^{-1} \sum_{i=1}^{n} R_i \int_{0}^{T} K_h(u - t)X_i(u)\epsilon_i(u)dN^c_i(u) \\
+ n^{-1} \sum_{i=1}^{n} (1 - R_i) \hat{E}_i \left\{ \int_{0}^{T} K_h(u - t)X_i(u)\epsilon_i(u)dN^c_i(u) \mid \mathcal{D}_i \right\} \right)
\]
\[+n^{-1} \sum_{i=1}^{n} (1 - R_i) E_{k} \left\{ \int_{0}^{T} K_h(u-t)X_i(u)\epsilon_i(u)dN_i^c(u) \mid \mathcal{D}_i, R_i = 0 \right\} \]

\[+n^{-1} \sum_{i=1}^{n} (1 - R_i) \hat{E}_{k} \left\{ \int_{0}^{T} K_h(u-t)X_i(u)\epsilon_i(u)dN_i^c(u) \mid \mathcal{D}_i, R_i = 0 \right\} \]

\[-n^{-1} \sum_{i=1}^{n} (1 - R_i) E_{k} \left\{ \int_{0}^{T} K_h(u-t)X_i(u)\epsilon_i(u)dN_i^c(u) \mid \mathcal{D}_i, R_i = 0 \right\} \]

\[= (nh)^{1/2}(e_{xx}(t))^{-1} \left\{ n^{-1} \sum_{i=1}^{n} R_i \int_{0}^{T} K_h(u-t)X_i(u)\epsilon_i(u)dN_i^c(u) \right\} \]

\[+n^{-1} \sum_{i=1}^{n} (1 - R_i) E_{k} \left\{ \int_{0}^{T} K_h(u-t)X_i(u)\epsilon_i(u)dN_i^c(u) \mid \mathcal{D}_i, R_i = 0 \right\} \]

\[+h^{1/2}(e_{xx}(t))^{-1} \left\{ n^{-1/2} \left[ \int_{0}^{\infty} \int_{0}^{L} E \left( (1 - R_i)X_i(u)\epsilon_i(u)\alpha_i^*(s-u) \mid F_s(V_i) \right) F_s(s)^{DMR(x)} y^{R(x)} \right. \right. \]

\[+ \int_{0}^{L} \int_{0}^{(L-x)-} E \left( (1 - R_i)X_i(u)\epsilon_i(u)\alpha_i^*(s-u) \mid F_s(V_i) \right) F_s(s)^{DMR((L-x)-)} y^{R((L-x)-)} \right] \]

\[+o_p(1) + O_p(n^{-1/2}h^2) \right) \]

\[= (nh)^{1/2}(e_{xx}(t))^{-1} \left\{ n^{-1} \sum_{i=1}^{n} R_i \int_{0}^{T} K_h(u-t)X_i(u)\epsilon_i(u)dN_i^c(u) \right\} \]

\[+n^{-1} \sum_{i=1}^{n} (1 - R_i) E_{k} \left\{ \int_{0}^{T} K_h(u-t)X_i(u)\epsilon_i(u)dN_i^c(u) \mid \mathcal{D}_i, R_i = 0 \right\} \]

\[+O(h^{1/2}) + o_p(h^{1/2}) + O_p(n^{-1/2}h^{5/2}) \]

which for each fixed time point \( t \), converges in distribution to a multivariate distribution with mean 0 and covariance matrix \( \mu_0 \Sigma(t) \) by Lindeberg-Feller theorem.

We derive the asymptotic covariance matrix in the following way.

\[\text{cov} \left\{ (nh)^{1/2} \left( \hat{\beta}(t; \gamma_0) - \beta^*(t) + h^2(e_{xx}(t))^{-1}[e'_{xx}(t)] \beta_0'(t) \right. \right. \]

\[+ (1/2)e_{xx}(t) \beta_0''(t) \right\} \int_{-1}^{1} x^2 K(x) dx \right] \]

\[= \text{cov} \left\{ (nh)^{1/2}(e_{xx}(t))^{-1} \left\{ n^{-1} \sum_{i=1}^{n} R_i \int_{0}^{T} K_h(u-t)X_i(u)\epsilon_i(u)dN_i^c(u) \right\} \right. \]

\[+ O(h^{1/2}) + o_p(h^{1/2}) + O_p(n^{-1/2}h^{5/2}) \]
By the Doob-Meyer decomposition of $N_i$, Note that all the subjects are i.i.d. and that $R_i$ is an indicator,

\[
+ n^{-1} \sum_{i=1}^{n} (1 - R_i) E_s \left\{ \int_0^\tau K_h(u - t) X_i(u) \epsilon_i(u) dN_i^c(u) \mid D_i, R_i = 0 \right\}
\]

\[
= n^{-1} h(e_{xx}(t))^{-1} \text{cov} \left[ \left( \sum_{i=1}^{n} R_i \int_0^\tau K_h(u - t) X_i(u) \epsilon_i(u) dN_i^c(u) \right. \right.
\]

\[
+ \sum_{i=1}^{n} (1 - R_i) E_s \left\{ \int_0^\tau K_h(u - t) X_i(u) \epsilon_i(u) dN_i^c(u) \mid D_i, R_i = 0 \right\} \right] (e_{xx}(t))^{-1}
\]

By the Doob-Meyer decomposition of $N_i$, $N_i^c(t) = \int_0^t Y_i^c(s) \alpha_i^c(s) ds + M_i^c(t)$. Let

\[
Y_i^c(t) = \sum_{j=1}^{n_i} I(T_{ij} \geq t).
\]

So

\[
h(e_{xx}(t))^{-1} \text{cov} \left[ \int_0^\tau K_h(u - t) R_i X_i(u) \epsilon_i(u) dN_i^c(u) \right] (e_{xx}(t))^{-1}
\]

\[
= h(e_{xx}(t))^{-1} \text{cov} \left[ \int_0^\tau K_h(u - t) R_i X_i(u) \epsilon_i(u) dM_i^c(u) \right] (e_{xx}(t))^{-1}
\]

\[
+ 2h(e_{xx}(t))^{-1} \text{cov} \left[ \int_0^\tau K_h(u - t) R_i X_i(u) \epsilon_i(u) dM_i^c(u),
\right.
\]

\[
\int_0^\tau K_h(u - t) R_i X_i(u) \epsilon_i(u) Y_i^c(u) \alpha_i^c(u) du \right] (e_{xx}(t))^{-1}
\]

\[
+ h(e_{xx}(t))^{-1} \text{cov} \left[ \int_0^\tau K_h(u - t) R_i X_i(u) \epsilon_i(u) Y_i^c(u) \alpha_i^c(u) du \right] (e_{xx}(t))^{-1}.
\]

$R_i$, $X_i(t)$ and $\epsilon_i(t)$ are $\mathcal{F}_t^c$-predictable. This leads the first term above to

\[
h(e_{xx}(t))^{-1} \text{cov} \left[ \int_0^\tau K_h(u - t) R_i X_i(u) \epsilon_i(u) dM_i^c(u) \right] (e_{xx}(t))^{-1}
\]
\begin{align*}
&= h(e_{xx}(t))^{-1}E \left( \int_0^T K_h^2(u-t) R_i^2 X_i(u) X_i^T(u) \epsilon_i^2(u) d < M > \right) (e_{xx}(t))^{-1} \\
&= h(e_{xx}(t))^{-1}E \left( \int_0^T K_h^2(u-t) R_i^2 X_i(u) X_i^T(u) \epsilon_i^2(u) Y_i^c(u) \alpha_i^c(u) du \right) (e_{xx}(t))^{-1} \\
&= (e_{xx}(t))^{-1}E \left( \int_{-1}^1 K^2(x) R_i^2 X_i(t + xh) X_i^T(t + xh) \epsilon_i^2(t + xh) Y_i^c(t + xh) \alpha_i^c(t + xh) dx \right) (e_{xx}(t))^{-1} \\
&= (e_{xx}(t))^{-1}E \left[ R_i^2 X_i(t) X_i^T(t) \epsilon_i^2(t) Y_i^c(t) \alpha_i^c(t) \int_{-1}^1 K^2(x) dx + O(h^2) \right] (e_{xx}(t))^{-1} \\
&= \mu_0 (e_{xx}(t))^{-1}E[ R_i^2 X_i(t) X_i^T(t) \epsilon_i^2(t) Y_i^c(t) \alpha_i^c(t) ] (e_{xx}(t))^{-1} + O(h^2). \\
\end{align*}

And

\begin{align*}
&= h(e_{xx}(t))^{-1} \text{cov} \left( \int_0^T K_h(u-t) R_i X_i(u) \epsilon_i(u) Y_i^c(u) \alpha_i^c(u) du \right) (e_{xx}(t))^{-1} \\
&= h(e_{xx}(t))^{-1} \text{cov} \left( \int_{-1}^1 K(x) R_i X_i(t + xh) \epsilon_i(t + xh) Y_i^c(t + xh) \alpha_i^c(t + xh) dx \right) (e_{xx}(t))^{-1} \\
&= h(e_{xx}(t))^{-1} \text{cov}[ R_i X_i(t) \epsilon_i(t) Y_i^c(t) \alpha_i^c(t) + O(h^2) ] (e_{xx}(t))^{-1} = O(h).
\end{align*}

Since

\begin{align*}
E \left( \int_0^T K_h(u-t) R_i X_i(u) \epsilon_i(u) dM_i^c(u) \right) &= E \left[ E \left( \int_0^T K_h(u-t) R_i X_i(u) \epsilon_i(u) dM_i^c(u) \bigg| X_i(\cdot), Z_i(\cdot), N_i(\cdot), S_i, V_i, C_i \right) \right] \\
&= E \left( \int_0^T K_h(u-t) R_i X_i(u) E(\epsilon_i(u)|X_i(\cdot), Z_i(\cdot), N_i(\cdot), S_i, V_i, C_i) dM_i^c(u) \right) \\
&= E \left( \int_0^T K_h(u-t) R_i X_i(u) E(\epsilon_i(u)|X_i(\cdot), Z_i(\cdot)) dM_i^c(u) \right) \bigg| X_i(\cdot), Z_i(\cdot), N_i(\cdot), S_i, V_i, C_i \bigg] = 0
\end{align*}

and

\begin{align*}
E \left( \int_0^T K_h(u-t) R_i X_i(u) \epsilon_i(u) Y_i^c(u) \alpha_i^c(u) du \right) &= E \left[ E \left( \int_0^T K_h(u-t) R_i X_i(u) \epsilon_i(u) Y_i^c(u) \alpha_i^c(u) du \bigg| X_i(\cdot), Z_i(\cdot), N_i(\cdot), S_i, V_i, C_i \right) \right] \\
&= E \left( \int_0^T K_h(u-t) R_i X_i(u) \epsilon_i(u) Y_i^c(u) \alpha_i^c(u) du \bigg| X_i(\cdot), Z_i(\cdot), N_i(\cdot), S_i, V_i, C_i \right) \bigg]\end{align*}
Then by the Cauchy-Schwarz inequality,

\[ h \leq hE \left( \int_0^\tau K_h(u-t)R_iX_i(u)E(\varepsilon_i(u)|X_i(\cdot), Z_i(\cdot), N_i(\cdot), S_i, V_i, C_i)Y_i^c(u)\alpha_i^c(u)du \right) \]

we have

\[
cov\left( \int_0^\tau K_h(u-t)R_iX_i(u)\varepsilon_i(u)\alpha_i^c(u)du \right)
= E \left( \int_0^\tau K_h(u-t)R_iX_i(u)\varepsilon_i(u)\alpha_i^c(u)du \right)^2,
\]

\[
cov\left( \int_0^\tau K_h(u-t)R_iX_i(u)\varepsilon_i(u)Y_i^c(u)\alpha_i^c(u)du \right)
= E \left( \int_0^\tau K_h(u-t)R_iX_i(u)\varepsilon_i(u)Y_i^c(u)\alpha_i^c(u)du \right)^2
\]

\[
cov\left( \int_0^\tau K_h(u-t)R_iX_i(u)\varepsilon_i(u)dM_i^c(u), \int_0^\tau K_h(u-t)R_iX_i(u)\varepsilon_i(u)Y_i^c(u)\alpha_i^c(u)du \right)
= E \left[ \left( \int_0^\tau K_h(u-t)R_iX_i(u)\varepsilon_i(u)dM_i^c(u) \right)^T \left( \int_0^\tau K_h(u-t)R_iX_i(u)\varepsilon_i(u)Y_i^c(u)\alpha_i^c(u)du \right) \right].
\]

Then by the Cauchy-Schwarz inequality,

\[ h \cdot \text{cov}(\int_0^\tau K_h(u-t)R_iX_i(u)\varepsilon_i(u)dM_i^c(u), \int_0^\tau K_h(u-t)R_iX_i(u)\varepsilon_i(u)Y_i^c(u)\alpha_i^c(u)du) \]

\[ = hE \left[ \left( \int_0^\tau K_h(u-t)R_iX_i(u)\varepsilon_i(u)dM_i^c(u) \right)^T \right] \left[ E \left( \int_0^\tau K_h(u-t)R_iX_i(u)\varepsilon_i(u)Y_i^c(u)\alpha_i^c(u)du \right)^2 \right]^{1/2} \]
\[
\begin{align*}
= h \left\{ \left[ \cov \left( \int_0^t K_h(u-t)R_iX_i(u)\epsilon_i(u)dM_i^c(u) \right) \right] + \left[ \cov \left( \int_0^t K_h(u-t)R_iX_i(u)\epsilon_i(u)dM_i^c(u) \right) \right] \right\}^{1/2} \\
= \left\{ \left[ \h cov \left( \int_0^t K_h(u-t)R_iX_i(u)\epsilon_i(u)dM_i^c(u) \right) \right] + \left[ \h cov \left( \int_0^t K_h(u-t)R_iX_i(u)\epsilon_i(u)dM_i^c(u) \right) \right] \right\}^{1/2} \\
= \left\{ \left[ \mu_0(e_{xx}(t))^-1E[R_i^2X_i(t)X_i^T(t)e_i^2(t)Y_i^c(t)\alpha_i^c(t)](e_{xx}(t))^-1 + O(h^2) \right][O(h)] \right\}^{1/2} \\
= O(h)
\end{align*}
\]

Hence
\[
\begin{align*}
h(e_{xx}(t))^-1 \cov \left( \int_0^t K_h(u-t)R_iX_i(u)\epsilon_i(u)dN_i^c(u) \right)(e_{xx}(t))^-1 \\
= \mu_0(e_{xx}(t))^-1E[R_i^2X_i(t)X_i^T(t)e_i^2(t)Y_i^c(t)\alpha_i^c(t)](e_{xx}(t))^-1 + O(h^2) + O(h).
\end{align*}
\]

Note that
\[
\begin{align*}
\tilde{e}_{xy}(t) &= \int_0^t K_h(s-t)e_{xy}(s)ds = \int_{t-h}^{t+h} h^{-1}K(s/t)h^{-1}e_{xy}(s)ds \\
&= \int_{-1}^{1} K(x)e_{xy}(t + xh)dx \\
&= \int_{-1}^{1} K(x)(e_{xy}(t) + hxe_{xy}'(t) + (1/2)h^2x^2e_{xy}''(t) + o(h^2))dx \\
&= e_{xy}(t) \int_{-1}^{1} K(x)dx + he_{xy}'(t) \int_{-1}^{1} xK(x)dx + (1/2)h^2e_{xy}''(t) \int_{-1}^{1} x^2K(x)dx \\
&\quad + o(h^2) \\
&= e_{xy}(t) + (1/2)h^2e_{xy}''(t) \int_{-1}^{1} x^2K(x)dx + o(h^2).
\end{align*}
\]

Similar results hold for \( \tilde{e}_{xx}(t) \) and \( \tilde{e}_{xz}(t) \). Let \( \mu_2 = \int_{-1}^{1} x^2K(x)dx \). So by the long division of functions
\[
\begin{align*}
\tilde{y}_x^T(t) &= (\tilde{e}_{xx}(t))^{-1}\tilde{e}_{xy}(t) \\
&= (e_{xx}(t) + (1/2)\mu_2h^2e_{xx}''(t) + o(h^2))^{-1}(e_{xy}(t) + (1/2)\mu_2h^2e_{xy}''(t) + o(h^2))
\end{align*}
\]
\[
\begin{align*}
&= y_x^T(t) + (1/2)\mu_2 h^2(e_{xx}(t))^{-1}[e''_{xy}(t) - e''_{xx}(t)(e_{xx}(t))^{-1}e_{xy}(t)] + o(h^2).
\end{align*}
\]

Also
\[
\tilde{z}_x^T(t) = z_x^T(t) + (1/2)\mu_2 h^2(e_{xx}(t))^{-1}[e''_{xx}(t) - e''_{xx}(t)(e_{xx}(t))^{-1}e_{xx}(t)] + o(h^2).
\]

Then
\[
\beta^*(t) = \bar{y}^T_x(t) - \tilde{z}_x^T(t)\gamma_0
\]
\[
= y_x^T(t) - z_x^T(t)\gamma_0 + (1/2)\mu_2 h^2(e_{xx}(t))^{-1}[e''_{xy}(t) - e''_{xx}(t)(e_{xx}(t))^{-1}e_{xy}(t)]
- (1/2)\mu_2 h^2(e_{xx}(t))^{-1}[e''_{xx}(t) - e''_{xx}(t)(e_{xx}(t))^{-1}e_{xx}(t)]\gamma_0 + o(h^2)
\]
\[
= y_x^T(t) - z_x^T(t)\gamma_0 + (1/2)\mu_2 h^2(e_{xx}(t))^{-1}[e''_{xy}(t) - e''_{xx}(t)y_x^T(t)]
- (1/2)\mu_2 h^2(e_{xx}(t))^{-1}[e''_{xx}(t) - e''_{xx}(t)\tilde{z}_x^T(t)]\gamma_0 + o(h^2)
\]
\[
= \beta_0(t) + (1/2)\mu_2 h^2(e_{xx}(t))^{-1}[e''_{xy}(t) - e''_{xx}(t)\gamma_0 - e''_{xx}(t)\beta(t)] + o(h^2).
\]
\[(A.15)\]

So
\[
(nh)^{1/2}(\tilde{\beta}(t) - \beta^*(t))
\]
\[
= (nh)^{1/2}(\tilde{\beta}(t; \tilde{\gamma}) - \beta^*(t))
\]
\[
= (nh)^{1/2}(\tilde{\beta}(t; \gamma_0) - \beta^*(t)) + (nh)^{1/2}(\tilde{\gamma} - \gamma_0)\frac{\partial \tilde{\beta}(t; \gamma_0)}{\partial \gamma} + O_p(n^{-1/2}h^{1/2})
\]
\[
= (nh)^{1/2}(\tilde{\beta}(t; \gamma_0) - \beta^*(t)) + O(h^{1/2}) + O_p(n^{-1/2}h^{1/2})
\]

since \(n^{1/2}(\tilde{\gamma} - \gamma_0) \xrightarrow{D} \mathcal{N}(0, D^{-1}V D^{-1})\) and
\[
\frac{\partial \tilde{\beta}(t; \gamma_0)}{\partial \gamma} = -\tilde{Z}_x(t) \xrightarrow{P} - z_x(t).
\]

Therefore,
\[
(nh)^{1/2}(\tilde{\beta}(t) - \beta_0(t) - \beta_{Bias}(t)) \xrightarrow{D} \mathcal{N}(0, \mu_0 \Sigma(t)),
\]
as \(n \to \infty, h \to 0, nh \to \infty, nh^5 = O(1)\). \(\square\)