Summer 2023

Recurrent Event Data Analysis with Mismeasured Covariates

Ravinath Alahakoon Mudiyanselage
Missouri University of Science and Technology

Department: Mathematics and Statistics

RECURRENT EVENT DATA ANALYSIS WITH
MISMEASURED COVARIATES

by

ALAHAKOON MUDIYANSELAGE
RAVINATH SASANKA BANDARA ALAHAKOON

A DISSERTATION

Presented to the Graduate Faculty of the

MISSOURI UNIVERSITY OF SCIENCE AND TECHNOLOGY

In Partial Fulfillment of the Requirements for the Degree

DOCTOR OF PHILOSOPHY

in

MATHEMATICS WITH STATISTICS EMPHASIS

2023

Approved by:

Dr. Akim Adekpedjou, Advisor
Dr. V. A. Samaranayake
Dr. Gayla Olbricht
Dr. Martin Bohner
Dr. Maciej Zawodniok

PUBLICATION DISSERTATION OPTION

This dissertation consists of the following two articles, formatted in the style used by the Missouri University of Science and Technology.

Paper I, Pages 22-63, is under review by Statistica Sinica.

Paper II, Pages 64-97, to be submitted.

ABSTRACT

Consider a study with n units wherein every unit is monitored for the occurrence of an event that can recur with random end of monitoring. At each recurrence, p concomitant variables associated to the event recurrence are recorded with q $(q \leq p)$ collected with errors. Of interest in this dissertation is the estimation of the regression parameters of event time regression models accounting for the covariates. To circumvent the problem of bias and consistency associated with model's parameter estimation in the presence of measurement errors, we propose inference for corrected estimating functions with well-behaved roots under additive measurement errors model. We consider two types of failure time regression models: one with additive effects and the other with multiplicative effects on the pure event history. We show that estimation is essentially unbiased under the corrected profile likelihood for recurrent events, in comparison to biased estimations under a likelihood function that ignores correction in both cases. We propose methods for obtaining estimators of error variance and discuss the property of the estimators. We further investigate the case of misspecified error models under the multiplicative regression model and show that the resulting estimators under misspecification converge to a value different from that of the true parameter–thereby providing a basis for bias assessment. In both cases, simulation studies indicate that the asymptotic properties of the regression parameters closely approximate its finite sample properties. We demonstrate the foregoing correction methods on an open source rhDNase dataset which was gathered in a clinical setting.

ACKNOWLEDGMENTS

My sincerest gratitude goes to my advisor, Dr. Akim Adekpedjou, for his unwavering guidance, motivation, and patience during my tenure at the Missouri University of Science and Technology's graduate school. Your support was instrumental in making this dissertation a reality. I am also deeply thankful to Dr. Gideon K. D. Zamba for his insightful suggestions and valuable contributions to my research work.

With the utmost respect and appreciation, I would like to extend my sincere gratitude to my esteemed committee members, Dr. V.A. Samaranayake, Dr. Gayla Olbricht, Dr. Martin Bohner, and Dr. Maciej Zawodniok, for serving in my committee. I would like to express my heartfelt gratitude to the Department of Mathematics and Statistics for their generous financial assistance.

I am truly grateful to the teachers who have impacted my life throughout my educational journey. I would like to extend my heartfelt thanks to all my teachers at Dharmaraja College Kandy, University of Peradeniya, and Missouri University of Science and Technology. Their guidance, wisdom, and mentorship have shaped me into the person I am today. Their unwavering commitment to my growth and success will always be remembered and deeply appreciated. Thank you for playing such a significant role in my academic and personal development.

My deepest gratitude goes to my family for their constant encouragement and unwavering support throughout my academic journey. I am especially thankful to my father, late mother, brother, and sister for inspiring me to pursue my doctoral degree in the United States. Their love and belief in me have been a source of strength and motivation. I am also incredibly grateful to my wife, Zahra, and our three children, Aleina, Shayaan, and Yara, for their love, understanding, and emotional support. I could not have achieved this milestone without their unwavering encouragement and understanding.

TABLE OF CONTENTS

Page

LIST OF ILLUSTRATIONS

LIST OF TABLES

Table Page

PAPER I

PAPER II

1. INTRODUCTION

In this section, we first discuss the mathematical prerequisites to facilitate the reading of this dissertation. Secondly, the basic concepts of recurrent events data analysis and measurement error are also given.

1.1. MATHEMATICAL PRELIMINARIES

The pioneering work by Aalen (1978) on the theory of counting processes has been the key to the development of statistical tools for analyzing data in reliability and survival analysis settings. A detailed discussion of these topics can be found in Andersen et al. (2012), Chung et al. (1990), and Fleming and Harrington (2011).

Let $(\Omega, \mathscr{F}, P)$ be a complete probability space and $T = [0, \tau] \subset \mathbb{R}$ be an interval of time.

Definition 1 *A filtration $\boldsymbol{F} = \{\mathscr{F}_t, t \in T\}$ on $(\Omega, \mathscr{F}, P)$ is an increasing family of σ-algebras, that is, $\forall t \leq s$, $\mathscr{F}_t \subseteq \mathscr{F}_s \subseteq \mathscr{F}$.*

Note here that in the case of a stochastic process, $\mathscr{F}_t$ could be taken to be all information generated by the process up to time t, and is called the natural history of the process. From now on, we will denote by $\mathbf{F}$ the natural filtration associated with the probability space $(\Omega, \mathscr{F}, P)$.

Definition 2 *A stochastic process $\boldsymbol{X} = \{X_t, t \geq 0\}$ is called cadlag if its simple paths $\{\boldsymbol{X}(t, w) : t \in T\}$ are right continuous with left hand limits for almost all w. Furthermore, the set of all cadlag functions is called the Skorohod space.*

Definition 3 *A counting process is a stochastic process $\{N(t) : t \geq 0\}$ adapted to a filtration $\boldsymbol{F}$ with $N(0) = 0$ and $N(t) < \infty$ almost surely (a.s), and whose paths are with probability one right-continuous, piecewise constant, and have only jump discontinuities, with jumps of size +1.*

Definition 4 *A stochastic process* $\boldsymbol{X} = \{X_t, t \geq 0\}$ *is:*

1. *Integrable if* $\sup_{t\in T} E(X(t)) < \infty$*,*

2. *Square integrable if* $\sup_{t\in T} E(X(t)^2) < \infty$*,*

3. *Bounded if there exists a finite constant* Γ *such that* $P\left\{\sup_{t\in T} |X(t)| < \Gamma\right\} = 1$.

Definition 5 *A collection* $\boldsymbol{M} = \{M_t, t \geq 0\}$ *is an* $\boldsymbol{F}$*-martingale if* $\boldsymbol{M}$ *is* $\boldsymbol{F}$*-adapted and satisfies:*

1. *Integrability:* $E(|M_t|) < \infty$ *for all* $t \in T$*,*

2. *Martingale property:* $E(M_t|\mathscr{F}_s) = M_s$ *a.s* $\forall s < t$.

We obtain a sub martingale if (2) in previous definition is replaced by $E(M_t|\mathscr{F}_s) \geq M_s$ $a.s$ $\forall s < t$. On the other hand, a super martingale is obtained by replacing (2) in previous definition by $E(M_t|\mathscr{F}_s) \leq M_s$ $a.s$ $\forall s < t$.

Definition 6 *(Local Martingale)*

1. *A stochastic process* $\boldsymbol{M} = \{M_t, t \geq 0\}$ *is a local martingale with respect to a filtration* $\boldsymbol{F}$ *if there exists a sequence* $\{\tau_n\}, n \in \mathbb{N}$*, of stopping times such that, for each* n*,* $M_n = \{\boldsymbol{M}(t \wedge \tau_n) : 0 \leq t < \infty\}$ *is an* $\boldsymbol{F}$*-martingale.*

2. *If* M_n *above is a martingale and is a square integrable process,* M_n *is called a square integrable martingale and* $\boldsymbol{M}$ *is called a local square integrable martingale.*

We now discuss the notion of a predictable process.

Definition 7 *The* σ*-algebra generated by all the sets of the form:*

1. $[0] \times A, A \in \mathscr{F}_0$ *and,*

2. $(a, b] \times A, 0 \leq a < b < \infty, A \in \mathscr{F}_a$,

is called the predictable σ-algebra for $\boldsymbol{F}$, where $\mathscr{F}_0$ is the information at time 0.

Lemma 1 *Let $\boldsymbol{F}$ be a filtration, and $\boldsymbol{X}$ a left-continuous real-valued process adapted to $\boldsymbol{F}$. Then $\boldsymbol{X}$ is predictable.*

Proposition 1 *Let $\boldsymbol{X}$ be an $\mathscr{F}_t$-predictable process. Then, for any $t > 0$, $X(t)$ is $\mathscr{F}_t$-measurable.*

We now discuss an important theorem that allows us to decompose a submartingale.

Theorem 1 (Doob-Meyer Decomposition) *Let $\boldsymbol{M} = \{M_t, t \geq 0\}$ be a right continuous, nonnegative submartingale with respect to the filtration $\boldsymbol{F}$. Then, there exists a right-continuous martingale $\mathscr{M}(t)$ and an increasing right-continuous predictable process $A(t)$ such that $M(t) = \mathscr{M}(t) + A(t)$ a.s.*

Note that, if M is a martingale with $E(M^2(t)) < \infty$ for $t > 0$, Jensen's inequality indicates that $M^2(t)$ is a submartingale.

Corollary 1 *Let $\boldsymbol{M}$ be a cadlag martingale with respect to $\boldsymbol{F}$. Then, there exists a unique increasing right-continuous predictable process denoted by $\langle \boldsymbol{M}, \boldsymbol{M} \rangle (t)$ called the predictable quadratic variation process of $\boldsymbol{M}$, such that $\langle \boldsymbol{M}, \boldsymbol{M} \rangle (0) = 0$ a.s, $E \langle \boldsymbol{M}, \boldsymbol{M} \rangle (t) < \infty$ for all t and $\left\{\boldsymbol{M}^2(t) - \langle \boldsymbol{M}, \boldsymbol{M} \rangle (t) : t \geq 0\right\}$ is a right continuous martingale.*

We now discuss about notion of stochastic integration. A detailed discussion can be found in Chung et al. (1990).

Theorem 2 *Suppose M is a finite variation local square integrable martingale, H a predictable process and $\int_0^t H^2 d\langle M \rangle$ locally integrable. Then, $\int_0^t H dM$ is a local square integrable martingale and its quadratic variation process is given by*

$$\left\langle \int H dM \right\rangle (t) = \int_0^t H^2 d\langle M \rangle.$$

The above theorem can be further generalized to a vector of martingales **M** and **M'** and matrices **H** and **K** of predictable processes. In that case, the predictable covariation process is given by

$$\left\langle \int \mathbf{H}d\mathbf{M}, \int \mathbf{K}d\mathbf{M'} \right\rangle = \int_0^t \mathbf{H}d\left\langle \mathbf{M}, \mathbf{M'} \right\rangle \mathbf{K'}.$$

where **A'** denotes the transpose of a matrix **A**.

Definition 8 *Suppose a filtration $\boldsymbol{F}$ on $(\Omega, \mathscr{F}, P)$ is given. A multivariate counting process $\boldsymbol{N} = (N_1, \ldots, N_k)$ is a vector of k $\boldsymbol{F}$- adapted cadlag processes for which:*

1. *$N_i = 0 \;\; \forall i = 1, 2, \ldots, k$,*
2. *There jumps are of size one and no two components can jump at the same time,*
3. *Their paths are nondecreasing and piecewise constant.*

Note that because the components of the counting process **N** are adapted, cadlag, locally bounded, and non-decreasing, they are local submartingales. So, by the Doob-Meyer decomposition, there exists a compensator of N_i, say Λ_i, which is referred to as the cumulative intensity process of the counting process.
The following proposition makes an important connection among counting processes, martingales and stochastic integration which is crucial in our study.

Proposition 2 *Let $\boldsymbol{N}$ be a multivariate counting process and let $\boldsymbol{\Lambda} = \int \lambda$ be its associated vector of compensator processes such that each component of $\boldsymbol{\Lambda}$ is absolutely continuous. Let $\boldsymbol{M} = \boldsymbol{N} - \boldsymbol{\Lambda}$ be the resulting vector of local martingales. If $\boldsymbol{H}$ is a vector of locally bounded and predictable processes, then $\int \boldsymbol{H}d\boldsymbol{M}$ are vectors of local square integrable martingales with a quadratic variation process given by*

$$\left\langle \int \boldsymbol{H}d\boldsymbol{M} \right\rangle = \int \boldsymbol{H} diag\left\{\lambda\right\} \boldsymbol{H'}ds,$$

where diag $\{\boldsymbol{\lambda}\}$ *is the diagonal matrix of associated intensity processes.*

The idea of constructing likelihood with counting process data was first introduced by Jacod (1975). Considering counting process data, one can write the likelihood function in a product integral form, which is a continuous version of the simple product Π.

Let $\Delta N_i(t) = N_i(t) - N_i(t-)$ be the jump process, and let our intensity process depend on some p-dimensional parameter θ. Then, the likelihood in $[0, t]$ can be written as

$$L(\theta, t) = \prod_{i=1}^{n} \prod_{v \in [0,t]} \left\{ \lambda_i(v, \theta)^{\Delta N_i(v)} \times (1 - \lambda_i(v, \theta))^{1 - \Delta N_i(v)} \right\}, \tag{1.1}$$

where $N_i(t)$ is the counting process for each individual i in the study and $\lambda_i(t, \theta)$ is the hazard rate at time t which is a function of θ for a parametric model. Simplifying (1.1) using Taylor expansion and noting that $1 - \lambda_i(v, \theta)dv \approx \exp(-\lambda_i(v, \theta))dv$, we obtain

$$L(\theta, t) \propto \prod_{i=1}^{n} \left[\prod_{v \in [0,t]} \left\{ \lambda_i(v, \theta)^{\Delta N_i(v)} \right\} \times \exp\left\{ - \int_0^t \lambda_i(v, \theta) dv \right\} \right]. \tag{1.2}$$

Next, by taking the logarithm of (1.2), we obtain the log-likelihood process given by

$$l(\theta, t) = \sum_{i=1}^{n} \left\{ \int_0^t \log[\lambda_i(v, \theta)] dN_i(v) - \int_0^t \lambda_i(v, \theta) dv \right\}. \tag{1.3}$$

The score process $U_\theta(\theta, t)$ is obtained by taking the derivative of (1.3) with respect to θ.

$$\begin{aligned} U_\theta(\theta, t) &= \sum_{i=1}^{n} \left\{ \int_0^t \nabla \log[\lambda_i(v, \theta)] dN_i(v) - \int_0^t \nabla \lambda_i(v, \theta) dv \right\} \\ &= \sum_{i=1}^{n} \left\{ \int_0^t \nabla \log[\lambda_i(v, \theta)] dM_i(v) \right\}, \end{aligned}$$

where ∇ stands for the gradient operator.

We now provide a result which is key to obtaining asymptotic properties of the estimators.

1.2. RECURRENT EVENTS DATA ANALYSIS

Survival analysis is a statistical field that focuses on analyzing the time until the occurrence of a specific event (time-to-event data) of interest. For example in medical and engineering disciplines this time could be the time elapsed from the beginning of a particular treatment to the occurrence of another condition, such as death or component breakdown. Stochastic process formulation, counting processes, and martingale theory are the dominant tools used to handle these time-to-event data today.

However, it is important to note that in some cases, study subjects may experience the event of interest multiple times as time goes by, which is known as a recurrent event, and it occurs in various fields, including public health, biomedicine, engineering, economics, and geology. Examples of recurrent events in public health and biomedical studies include drug abuse of teenagers, hospitalization of chronically ill individuals, onset of depression, and recurrence of tumors. In engineering settings, recurrent events could be the failure of an electronic system, the breakdown of computer software, or the power outage of an electric grid. Recurrences of hurricanes, earthquakes, or volcano eruptions are examples in geology.

The statistical methods used for analyzing single-event data cannot be directly applied to recurrent event data. One reason why traditional methods cannot be directly applied to recurrent event data is that the occurrence of one event can affect the probability of subsequent events. For example, a hospitalization may increase the likelihood of another hospitalization, or a customer's first visit to a store may affect their likelihood of visiting again. This correlation between events violates the assumption of independence that underlies many traditional statistical methods. To handle recurrent event data, specialized methods have been developed, such as the counting process or the frailty models. These models take into account the correlation between events and allow for the analysis of recurrent event data. The counting process model is based on counting the number of events that occur in a specific time interval, and the frailty model incorporates the unobserved heterogeneity between individuals into the analysis. Various approaches have been proposed, such as doubly-indexed processes formulated by Gill (1981), Selvin (1988), and later extended by Peña (2001), which have become the dominant tools used to handle recurrent event data.

1.2.1. Survival Models. Survival models are a crucial tool in survival analysis. At the core of survival modeling is the concept of hazard functions, which describe the probability of an event occurring at a specific time, given that an individual or unit has survived up until that point. Hazard functions model the rate at which events occur over time and are essential for modeling the time-to-event data.

There are several approaches to modeling the hazard function, including parametric, semi-parametric, and non-parametric models. Parametric models assume a specific functional form for the hazard function, such as the exponential, Weibull, or log-normal distributions. Semi-parametric models, such as the Cox proportional hazards model, assume a baseline hazard function that is not specified and allow the effect of covariates to be modeled. Non-parametric models, such as the Kaplan-Meier estimator, do not assume any functional form for the hazard function and estimate it directly from the data.

Survival models allow researchers to investigate the effect of one or more covariates on the instantaneous risk of an event occurring, making them a powerful tool for analyzing survival data. These models can be used in both single-event and recurrent event settings. By understanding the different approaches to modeling the hazard function, researchers can build and interpret survival models, gaining valuable insights into the risk factors associated with an event of interest.

1.2.1.1. Hazard function. The hazard function, defined below, gives the probability of the subject i failing at the next instant, given that the subject has survived up to time t;

$$\lambda(t) = \lim_{\Delta t \to 0} \frac{P\{t \leq T_i \leq t + \Delta t | T_i \geq t\}}{\Delta t}.$$

The hazard function fully specifies the distribution of t so that it also determines the survivor and density functions.

1.2.1.2. Cox model. The Cox model (multiplicative hazard function) is defined as follows:

$$\lambda(t; \mathbf{x}_i(t)) = \lambda_0(t) \exp(\boldsymbol{\beta}' \mathbf{x}_i(t)), \tag{1.4}$$

where $\lambda_0(\cdot)$ is the baseline hazard function and $\boldsymbol{\beta}$ is the regression parameter vector. This model assumes that the effect of covariates on the baseline hazard is multiplicative. As noted in Cox (1975), a special property of this model is estimating regression parameters by obtaining the partial likelihood. Furthermore, this partial likelihood method allows us to investigate the covariate effects even when the baseline hazard function is unspecified.

1.2.1.3. Additive hazard model. The additive hazard model is given by

$$\lambda(t; \mathbf{x}_i(t)) = \lambda_0(t) + \boldsymbol{\beta}' \mathbf{x}_i(t),$$

where $\lambda_0(\cdot)$ is the baseline hazard function and $\boldsymbol{\beta}$ is the regression parameter vector. This survival model assumes that the covariates have an additive effect on the hazard. More details can be found in Lin and Ying (1994), and Cox and Oakes (2018).

1.2.1.4. Accelerated failure time model. Accelerated failure time model Cox (1972a) assumes

$$\lambda(t; \mathbf{x}_i(t))) = \exp(-\boldsymbol{\beta}' \mathbf{x}_i) \lambda_0[t e^{-\boldsymbol{\beta}' \mathbf{x}_i}],$$

where $\lambda_0(\cdot)$ is the baseline hazard function, $\boldsymbol{\beta}$ is the regression parameter vector and $\mathbf{x}$ is a time invariant covariate vector. This model also assumes that $Y = \log T$ and covariate $\mathbf{x}$ are related via the linear model $Y = \boldsymbol{\beta}'\mathbf{x} + e$ where e is an error variable. More details are provided in Kalbfleisch and Prentice (2011).

1.2.2. Recurrent Events. This subsection introduces the modeling and notations used for recurrent events. Consider a study in which n units are being monitored for the occurrences and reoccurrences of an event for a time period of $[0, \tau_i]$ for each unit i, where τ_is are independent and identically distributed (i.i.d.) random variables. For unit i, define $S_{i,j}$ as the calendar time at the j^{th} recurrence and $T_{i,j}$ be the time difference between $(j-1)$ and the j^{th} recurrences. The $T_{i,j}$s are referred to as gap times or interoccurrence times, and in the dissertation, they are assumed to be independent and identically distributed (i.i.d.) random variables with an absolutely continuous function denoted by $F(t) = P(T_{i,j} \leq t)$. Let $\mathbf{x}_i(s)$ be a p-dimensional time varying covariates for the i^{th} unit. If K_i is the number of recurrent events experienced by unit i, then the total observables is $O = \{O_1, ..., O_n\}$ where

$$O_i = \left\{K_i, \tau_i, T_{i,1}, ..., T_{i,K_i}, \tau_i - S_{i,K_i}, \mathbf{x}_i(S_{i,1}), ..., \mathbf{x}_i(S_{i,K_i})\right\}. \tag{1.5}$$

Recurrent event data is illustrated in Figure 1.1.

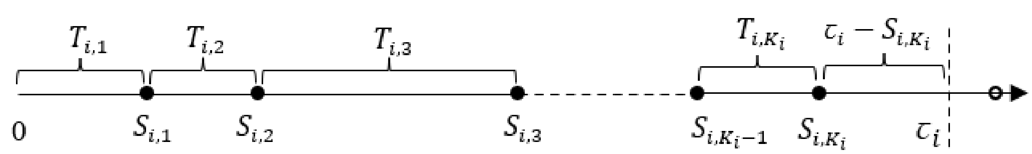

Figure 1.1. An illustration of recurrent event data.

1.2.3. Effective Age Process. The concept of the effective age process is a fundamental aspect of survival analysis that draws from reliability analysis, a field that deals with the repair and maintenance of systems and components. In this context, the term ”repair” refers to the process of restoring a damaged or failing system to a functional state. There are various models used to describe the effective age process in the literature, but two of the most commonly used models are the minimal repair model and the perfect repair model. The minimal repair model assumes that the repair can restore the system to the state it was in just before the failure occurred. In this case, the effective age process is simply the calendar time, or the amount of time that has passed since the system was put into use, denoted by $\varphi_i(s) = s$. In contrast, the perfect repair model assumes that a new, identical system can replace the failed system. This model assumes that the gaps between failures are independent and identically distributed, and the effective age process is defined as the time elapsed since the last failure; $\varphi_i(s) = s - S_{i,N_i^\dagger}(s-)$.

1.2.4. Stochastic Process Formulation. The observables in (1.5) can be expressed using stochastic processes. We define the following calendar time stochastic processes by,

$$N_i^\dagger = \{N_i^\dagger(s) : s \le \tau_i\} \quad \text{and} \quad Y_i^\dagger = \{Y_i^\dagger(s) : s \ge 0\},$$

where

$$N_i^\dagger(s) = \sum_{j=1}^{\infty} I\{S_{i,j} \le s \wedge \tau_i\} \quad \text{and} \quad Y_i^\dagger(s) = I\{\tau_i \ge s\}.$$

The process $N_i^\dagger(s)$ counts the number of observed events over the calendar period $[0, s]$ experienced by subject i while the $Y_i^\dagger(s)$ process indicates if the subject is still at risk of experiencing the event calendar by time s. Let $\lambda_0(\cdot)$ be the baseline hazard function and $\mathcal{F} = \{\mathcal{F}_s : s \ge 0\}$ a natural filtration generated by $\{(N_i^\dagger(s), Y_i^\dagger(s)) : s \ge 0\}$. Hence, the compensator process of $N_i^\dagger$ is given by

$$A_i^\dagger(s|\boldsymbol{\beta}) = \int_0^s Y_i^\dagger(v)\lambda_i(v)dv,$$

where

1. for multiplicative hazards: $\lambda_i(\cdot) = \lambda_0[\varphi_i(\cdot)] \exp[\boldsymbol{\beta}'\mathbf{x}_i(\cdot)]$ and

2. for additive hazards: $\lambda_i(\cdot) = \lambda_0[\varphi_i(\cdot)] + \boldsymbol{\beta}'\mathbf{x}_i(\cdot)$,

with $\varphi_i(\cdot)$ being the effective age process and $\boldsymbol{\beta}$ is a p-dimensional regression parameter. In this dissertation, we define the effective age process as $\varphi_i(s) = s - S_{i,N_i^\dagger}(s-)$, which is the time elapsed since the last event. The process $\varphi_i(s)$ is also called backward recurrence time. The martingale process $M_i^\dagger(s|\boldsymbol{\beta}) = N_i^\dagger(s) - A_i^\dagger(s|\boldsymbol{\beta})$ is a local square integrable martingale with regard to filtration $\mathcal{F}$.

1.2.5. Doubly Indexed Processes. Double indexed processes are an important concept in survival analysis that have gained attention due to their ability to address issues with martingales and renewal processes. The motivation for using double indexed processes stems from the breakdown of martingales, which are mathematical models that describe the expected value of a variable in a system over time. When a martingale breaks down, it can lead to a renewal process, where the system effectively ”resets” after each event. This can make it difficult to accurately model the behavior of the system over time, especially in cases where there are multiple factors that influence the likelihood of events occurring. Double indexed processes offer a solution to this problem by allowing for multiple variables to be incorporated into the model. By including two indices in the process, it becomes possible to account for both time-dependent and event-dependent variables, which can help to provide a more complete and accurate picture of the system being studied. Overall, the use of double indexed processes represents an important development in the field of survival analysis, enabling researchers to better understand and model complex systems where traditional martingale and renewal approaches may fall short.

For $i = 1, 2, ..., n$, define the doubly indexed processes $R_i(v,t)$, $N_i(s,t)$, $A_i(s,t|\boldsymbol{\beta})$ and $M_i(s,t|\boldsymbol{\beta})$ by

$$R_i(v,t) = I\{\varphi_i(s) \le t\},$$

$$N_i(s,t) = \int_0^s R_i(v,t) N_i^\dagger(dv),$$

$$A_i(s,t|\boldsymbol{\beta}) = \int_0^s R_i(v,t) A_i^\dagger(dv|\boldsymbol{\beta}), \tag{1.6}$$

$$M_i(s,t|\boldsymbol{\beta}) = \int_0^s R_i(v,t)M_i^{\dagger}(dv|\boldsymbol{\beta}) = N_i(s,t) - A_i(s,t|\boldsymbol{\beta}).$$

At calendar time v, $R_i(v,t)$ indicates if at the most t time units have elapsed since the last event occurrence. $N_i(s,t)$ counts number of recurrences experienced by the unit i at calendar time s whose effective age is at the most gap time t. $A_i(s,t|\boldsymbol{\beta})$ is the compensator process of $N_i(s,t)$ and $M_i(s,t|\boldsymbol{\beta})$ is a zero-mean square integrable martingale for fixed t since $M_i^{\dagger}(s|\boldsymbol{\beta})$ is a martingale and the process $R_i(v,t)$ is predictable.

1.2.6. Modeling of Recurrent Events Data with Cox Model. This subsection is devoted to recurrent events with multiplicative covariates effects. With a view towards estimating $\Lambda_0(t)$, an alternative form of (1.6) is needed due to the fact that (1.6) is not in multiplicative form. Therefore, from Pena et al. (2000), and Adekpedjou and Stocker (2015), the alternative form of (1.6) is defined by

$$A_i(s,t|\boldsymbol{\beta}) = \int_0^t Y_i(s,w|\boldsymbol{\beta})\lambda_0(w)dw,$$

where generalized at-risk process $Y_i(s,t|\boldsymbol{\beta})$, is given by

$$\begin{aligned} Y_i(s,t|\boldsymbol{\beta}) &= \sum_{j=1}^{N_i^{\dagger}((s\wedge\tau_i)-)} I(T_{ij} \geq t)\exp[\boldsymbol{\beta}'\mathbf{x}_i(\varphi_{i,j-1}^{-1}(t)] \\ &\quad + I\left((s\wedge\tau_i) - S_{i,N_i^{\dagger}((s\wedge\tau_i)-)} \geq t\right)\exp[\boldsymbol{\beta}'\mathbf{x}_i(\varphi_{i,N_i^{\dagger}((s\wedge\tau_i)-)}^{-1}(t))]. \end{aligned}$$

Moreover, the cumulative hazard function $\Lambda_0(t)$ under the Cox model is

$$\hat{\Lambda}_0(s,t|\boldsymbol{\beta}) = \int_0^t \frac{J(s,w|\boldsymbol{\beta})}{Y(s,w|\boldsymbol{\beta})}N(s,dw), \tag{1.7}$$

where $Y(s,w|\boldsymbol{\beta}) = \sum_{i=1}^{n} Y_i(s,w|\boldsymbol{\beta})$ and $J(s,w|\boldsymbol{\beta}) = I\{Y(s,w|\boldsymbol{\beta}) > 0\}$. More details on deriving $\hat{\Lambda}_0(s,t|\boldsymbol{\beta})$ can be found in Pena et al. (2001). Also, using (1.7), the product limit estimator for the baseline survivor function is

$$\hat{\bar{F}}_0(s,t|\boldsymbol{\beta}) = \prod_{w=0}^{t} \left\{1 - \hat{\Lambda}_0(s,dw|\boldsymbol{\beta})\right\}. \tag{1.8}$$

Note that the estimator $\hat{\boldsymbol{\beta}}$ is needed in order to estimate $\Lambda_0(s,t|\boldsymbol{\beta})$ and $\hat{\bar{F}}_0(s,t|\boldsymbol{\beta})$ using (1.7) and (1.8) respectively. We derive the full likelihood process following Jacod (1975) as follows:

$$\begin{aligned}
L_{full}(\Lambda_0,\boldsymbol{\beta},s) &= \prod_{i=1}^{n}\prod_{w=0}^{s}\left[dA_i^{\dagger}(w,\boldsymbol{\beta})\right]^{\Delta N_i^{\dagger}(w)} \times \left[1 - dA_i^{\dagger}(w,\boldsymbol{\beta})\right]^{1-\Delta N_i^{\dagger}(w)} \\
&= \prod_{i=1}^{n}\prod_{w=0}^{s}\left[Y_i^{\dagger}(w)\lambda_0[\varphi_i(w)]\exp(\boldsymbol{\beta}'\mathbf{x}_i(w))dw\right]^{\Delta N_i^{\dagger}(w)} \\
&\quad \times \left[1 - Y_i^{\dagger}(w)\lambda_0[\varphi_i(w)]\exp(\boldsymbol{\beta}'\mathbf{x}_i(w))dw\right]^{1-\Delta N_i^{\dagger}(w)} \\
&= \left\{\prod_{i=1}^{n}\prod_{w=0}^{s}\left[Y_i^{\dagger}(w)\lambda_0[\varphi_i(w)]\exp(\boldsymbol{\beta}'\mathbf{x}_i(w))dw\right]^{\Delta N_i^{\dagger}(w)}\right\} \\
&\quad \times \left\{\exp\left[-\sum_{i=1}^{n}\int_0^s Y_i^{\dagger}(w)\lambda_0[\varphi_i(w)]\exp(\boldsymbol{\beta}'\mathbf{x}_i(w))dw\right]\right\}.
\end{aligned} \tag{1.9}$$

From Adekpedjou and Stocker (2015), substituting $\hat{\Lambda}_0(s,t|\boldsymbol{\beta})$ for $\Lambda_0(w)$ in(1.9) and after simplifying, we get profile likelihood as

$$L_p(\boldsymbol{\beta}) = \prod_{i=1}^{n}\prod_{j=1}^{N_i^{\dagger}((s\wedge\tau_i)-)}\left\{\frac{\exp(\boldsymbol{\beta}'\mathbf{x}_i(S_{i,j}))}{Y(s,\varphi_i(S_{i,j})|\boldsymbol{\beta})}\right\}^{\Delta N_i^{\dagger}(S_{i,j})}.$$

Note that, the argument for the exponential part of (1.9) can be written in the following form (Peña et al. (2007a))):

$$\sum_{i=1}^{n}\int_0^s Y_i^{\dagger}(w)\lambda_0[\varphi_i(w)]\exp(\boldsymbol{\beta}'\mathbf{x}_i(w))dw = \int_0^{\infty} Y(s,w)\Lambda_0(dw).$$

From (1.7), it follows that $\int_0^\infty Y(s,w|\boldsymbol{\beta})\hat{\Lambda}_0(s,dw|\boldsymbol{\beta}) = \sum_{i=1}^n N_i(s,\infty)$, and this is independent of $\boldsymbol{\beta}$. As a result, the exponential part in (1.9) does not contribute to the profile likelihood of $\boldsymbol{\beta}$ and hence will be discarded. Next, by taking the logarithm of $L_p(\boldsymbol{\beta})$, we obtain log profile likelihood process as follows:

$$l_p(\boldsymbol{\beta}) = \sum_{i=1}^n \int_0^s \left[\boldsymbol{\beta}'\mathbf{x}_i(w) - \log[Y(s,\varphi_i(w)|\boldsymbol{\beta})]\right] N_i^\dagger(dw). \tag{1.10}$$

Estimating equation for profile maximum likelihood estimator of $\boldsymbol{\beta}$ can be obtained by equating the gradient of (1.10) to $\mathbf{0}$ as below.

$$\sum_{i=1}^n \int_0^s \left[\mathbf{x}_i(w) - \frac{\frac{\partial}{\partial \beta} Y(s,\varphi_i(w)|\boldsymbol{\beta})}{Y(s,\varphi_i(w)|\boldsymbol{\beta})}\right] N_i^\dagger(dw) = \mathbf{0}.$$

Numerical methods such as the Newton–Raphson algorithm or the Nelder–Mead simplex algorithm should be used to estimate $\boldsymbol{\beta}$, since a closed form expression for $\boldsymbol{\beta}$ is not obtainable.

1.2.7. Modeling of Recurrent Events Data with Additive Hazard Model. This subsection is devoted to recurrent events with additive covariates effects. From Stocker and Adekpedjou (2020), an equivalent expression for $A_i(s,t)$ in (1.6) is obtained where the argument of $\lambda_0(\cdot)$ is no longer the effective age process so that $\Lambda_0(t)$ can be obtained.

$$A_i(s,t) = \int_0^t Y_i(s,w)\left\{\lambda_0(w) + \boldsymbol{\beta}'\mathbf{x}_i(\varphi_i^{-1}(w))\right\} dw,$$

where the generalized at-risk process $Y_i(s,w)$ is given by

$$Y_i(s,w) = \sum_{j=1}^{N_i^\dagger((s\wedge\tau_i)-)} I(T_{ij} \geq w) + I\left((s\wedge\tau_i) - S_{i,N_i^\dagger((s\wedge\tau_i)-)} \geq w\right).$$

Also, by Stocker and Adekpedjou (2020), the cumulative hazard function $\Lambda_0(t)$ under the additive hazard function can be written as

$$\hat{\Lambda}_0(s,t|\boldsymbol{\beta}) = \int_0^t \frac{\sum_{i=1}^n \left\{N_i(s,dw) - Y_i(s,w)\boldsymbol{\beta}'\mathbf{x}_i(\varphi_i^{-1}(w))dw\right\}}{\sum_{i=1}^n Y_i(s,w)}.$$

A score function can be constructed using similar arguments as in Lin and Ying (1995). Details are provided in Stocker and Adekpedjou (2020). This score function is given by

$$U(\boldsymbol{\beta};\mathbf{x},s,t) = \sum_{i=1}^n \int_0^t [\mathbf{x}_i(\varphi_i^{-1}(w)) - \bar{\mathbf{x}}(\varphi^{-1}(w))][N_i(s,dw) - Y_i(s,w)\boldsymbol{\beta}^T\mathbf{x}_i(\varphi_i^{-1}(w))dw],$$

where

$$\bar{\mathbf{x}}(\varphi^{-1}(t)) = \frac{\sum_{j=1}^n Y_j(s,t)\mathbf{x}_j(\varphi_j^{-1}(t))}{\sum_{j=1}^n Y_j(s,t)}.$$

Solving $U(\boldsymbol{\beta};\mathbf{x},s,t) = \mathbf{0}$, we obtain

$$\begin{aligned}\hat{\boldsymbol{\beta}} &= \left\{\sum_{i=1}^n \int_0^t Y_i(s,w)[\mathbf{x}_i(\varphi_i^{-1}(w)) - \bar{\mathbf{x}}(\varphi^{-1}(w))]^{\otimes 2}dw\right\}^{-1} \\ &\quad \times\left\{\sum_{i=1}^n \int_0^t [\mathbf{x}_i(\varphi_i^{-1}(w)) - \bar{\mathbf{x}}(\varphi^{-1}(w))]N_i(s,dw)\right\}.\end{aligned}$$

1.3. MEASUREMENT ERROR

When dealing with measurement errors that affect covariates, it becomes crucial to understand the relationship between the error-contaminated covariates and their true (error-free) versions. This is necessary in order to make accurate inferences.

1.3.1. Measurement Error Models. In this section, we introduce commonly used error models in the literature. A comprehensive overview can be found in Carroll et al. (2006).

1.3.1.1. Classical additive error model. The classical additive error model has the form

$$x_i = z_i + \epsilon_i, \tag{1.11}$$

where ϵ_i are assumed to be independent and identically distributed with a mean of zero and a positive definite variance-covariance matrix $\boldsymbol{\Xi}$. It is also assumed that the ϵ_i are independent of the true covariates z_i. The multivariate normal distribution in the literature is a common choice for ϵ_i because of its reasonable assumptions, compatibility with other models and wide applicability.

1.3.1.2. Berkson model. The Berkson error model is given by

$$z_i = x_i + \epsilon_i, \tag{1.12}$$

where ϵ_i are assumed to be independent and identically distributed with a mean of zero and a positive definite variance covariance matrix $\boldsymbol{\Xi}$. It is also assumed that the ϵ_i are independent of the observed covariates x_i. Multivariate normal distribution is a common choice for ϵ_i in the literature.

The difference between the classical additive error model and the Berkson model is how they view the association between z and x. The Berkson model treats x as the independent variable and z as the dependent variable, while the classical additive error model treats z as the independent variable and x as the dependent variable.

1.3.1.3. Multiplicative model. The multiplicative error model is given by

$$x_i = z_i \epsilon_i, \tag{1.13}$$

where ϵ_i are assumed to be independent and identically distributed with a mean of zero and a positive definite variance-covariance matrix $\boldsymbol{\Xi}$. It is also assumed that the ϵ_i are independent of the true covariates z_i.

1.3.1.4. Latent variable model. The latent variable error model is a combination of both the classical error model and the Berkson model. Therefore, it is more flexible in handling measurement errors. This model uses a latent variable denoted by w_i to make the connection between x_i and z_i, and is defined by

$$x_i = w_i + \epsilon_{i,A} \tag{1.14}$$

and

$$z_i = w_i + \epsilon_{i,B}, \tag{1.15}$$

where $\epsilon_{i,A}$ and $\epsilon_{i,B}$ both have mean zero, and their corresponding error covariance variance matrices are Ξ_A and Ξ_B respectively. It is also assumed that $\epsilon_{i,A}$, $\epsilon_{i,A}$ and w_i are mutually independent.

The classical additive error model is the most popular in modeling survival data subject to covariate measurement error. More details on this error model and its applications can be found in Carroll et al. (2006).

1.3.1.5. Repeated measurements. Sometimes, replicate surrogate measurements of z_i may be available, say w_i. In particular, suppose z_i is measured $m_i(> 1)$ times repeatedly. Then for $l = 1, ..., m_i$, the classical additive error model becomes

$$w_{il} = z_{il} + \epsilon_{il},$$

where ϵ_{il} are zero mean i.i.d. random variables with a positive definite variance-covariance matrix Σ. For the classical additive errors, it is possible to estimate Σ as follows using these replications:

$$\hat{\Sigma} = \frac{\sum_{i=1}^{n} \sum_{l=1}^{m_i} (w_{il} - \bar{w}_{i.})^{\otimes 2}}{\sum_{i=1}^{n} (m_i - 1)},$$

where $\bar{w}_{i.} = \sum_{l=1}^{m_i} w_{il}/m_i$.

1.3.1.6. Validation subsample. A valid subsample usually contains measurements for both true (z_i) and surrogate (w_i) covariates. Furthermore, validation data can be categorized into internal and external groups depending on the response measurements' availability. In particular, an internal validation data set contains response measurements, whereas an external validation data set does not. Internal validation data can directly examine measurement error structure, often giving accurate estimators and inferences. In contrast, external validation data can be used to assess the measurement error model. More details on validation data can be found in Carroll et al. (2006).

1.3.1.7. Instrumental data. Sometimes, a second measurement of z_i, say $\tilde{z}_i$, may be available, measured using another mechanism. This variable, $\tilde{z}_i$, is often called an instrumental variable and is correlated with z_i albeit with a weaker relationship than w_i to z_i. The availability of so-called instrumental data can be useful in measurement error analysis. More details on instrumental data can be found in Carroll et al. (2006).

1.4. CONSEQUENCES OF IGNORING COVARIATE MEASUREMENT ERROR ON PARAMETER ESTIMATION

We performed a simulation study to demonstrate how parameter estimation is affected by error-contaminated covariates. We used the Cox model with a single covariate given by $\lambda(t; z_i) = \lambda_0(t)\exp(z_i\beta)$ where true covariate $z_i \sim N(0, 1)$. In this simulation study, we considered two scenarios below:

Scenario 1: We used classical additive error model $x_i = z_i + \epsilon_i$ where $\epsilon_i \sim N(0, \sigma^2)$. For each σ, we generated error-contaminated covariate x_i, replaced z_i in the Cox model by x_i, and estimated β. We replicated this β estimating procedure 200 times and plotted the average of those estimates labeled error-prone in Figure 1.2.

Scenario 2: For each σ, we used true covariate z_i in the Cox model and estimated β. We replicated this β estimating procedure 200 times and plotted the average of those estimates labeled error-free in Figure 1.2. We can see from the Figure 1.2 that the error-free parameter

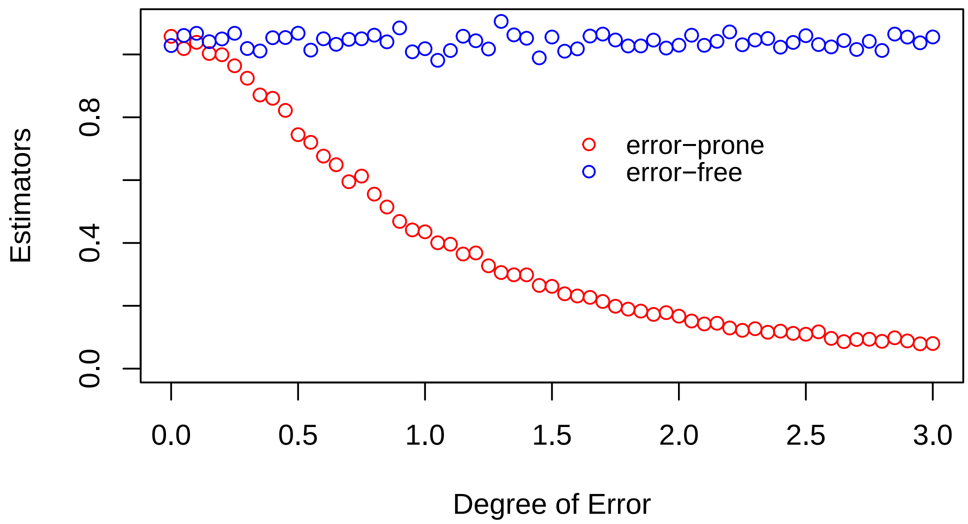

Figure 1.2. Impact of covariate measurement error on parameter estimation.

estimates are consistent around the true parameter value, which is 1. In contrast, error-prone parameter estimates are significantly biased. Moreover, the bias increases significantly as the magnitude of the error increases. Hence, ignoring covariate measurement error can lead to biased parameter estimates. Therefore, better estimating functions need to be developed in order to estimate β.

1.5. EXISTING METHODOLOGY ON RECURRENT EVENTS DATA WITH MEASUREMENT ERROR

Even though many methods have been developed in the survival data setting for single events with error-contaminated covariates, a little work has been done with regard to recurrent events with error-contaminated covariates. In this section, we review the literature on existing methods for analyzing recurrent events data subject to covariate measurement error.

Turnbull (1997) considered a mixed effects poisson regression model for recurrent event data with error-contaminated covariates. This author proposed adjustments for usual maximum likelihood estimators that are obtained from neglecting covariate measurement error.

Jiang et al. (1999a) investigated inference methods for discrete-time events in the presence of covariate measurement error. In particular, they used semi-parametric Poisson and mixed poisson process regression while accounting for possible random effects and covariate measurement error.

Yi and Lawless (2012) developed inferential methods that account for covariate measurement error. Particularly, their work included counting processes consisting of multiplicative intensity functions and mixed Poisson models. They discussed inference methods based on likelihood which led to obtaining estimation equations.

Yu et al. (2018) proposed non-parametric methods taking covariate measurement error into account in multivariate recurrent event data under informative censoring. However, their research was limited to time-independent covariates. Moreover, their approach did not require the Poisson-type assumption for recurrent event process and any distributional assumption for frailty or covariate measurement error.

In addition to the aforementioned existing methodology on recurrent events data with measurement error, there is some work in the literature related to measurement error in the field of survival analysis. Veierød and Laake (2001) and Guo and Li (2002) explored covariate measurement error effects on Poisson regression and misclassification. Zeger and Edelstein (1989) studied the Poisson regression model with error-contaminated covariates and used a likelihood method to correct the measurement error effects. Fung and Krewski (1999) investigated SIMEX and regression calibration algorithms empirically for Poisson regression with replicates of error-prone covariate measurements. Kim (2007) worked on a mean model for the event count data and used kernel estimates to obtain a correction method in the presence of categorical error-prone covariates while assuming a validation subsample is available. These studies did not investigate the asymptotic properties of the derived estimators. However, they provided simulation study results to assess their proposed methods' performance.

1.6. SPECIFIC OBJECTIVES OF THE DISSERTATION

In this dissertation, we develop statistical methods to analyze recurrent event data with mismeasured covariates. We consider two types of intensity functions namely the multiplicative and additive. While the first part of this dissertation is devoted to the development of statistical methods based on the multiplicative intensity model, the second part will concentrate on the additive intensity model. The aims of this dissertation are:

- Derive regression parameter estimators of the intensity models based on corrected scores, and obtain their asymptotic properties.
- Derive an estimator for cumulative baseline hazard function under the multiplicative regression model, and obtain its asymptotic properties.
- Investigate the effects of misspecified error models under the multiplicative regression model, and assess bias.
- Derive an estimator for error variance.
- Run simulation study to validate the theoretical results.
- Apply the results to real recurrent events data.

PAPER

I. ESTIMATION AND MODEL MISSPECIFICATION FOR RECURRENT EVENT DATA WITH COVARIATES UNDER MEASUREMENT ERRORS

Ravinath Alahakoon
Department of Mathematics & Statistics
Missouri University of Science and Technology
Rolla, Missouri 65409–0050
Email: ra8k7@mst.edu

Akim Adekpedjou
Department of Mathematics & Statistics
Missouri University of Science and Technology
Rolla, Missouri 65409–0050
Email: akima@mst.edu

K. D. Zamba
Department of Biostatistics
The University of Iowa
Iowa City, Iowa 52242
Email: gideon-zamba@uiowa.edu

ABSTRACT

For subject i, we monitor an event that can occur multiple times over a random observation window $[0, \tau_i)$. At each recurrence, p concomitant variables, $\mathbf{x}_i$, associated with the event recurrence are recorded–a subset ($q \leq p$) of which is measured with errors. To circumvent the problem of bias and consistency associated with parameter estimation in the presence of measurement errors, we propose inference for corrected estimating equations with well-behaved roots under the additive measurement errors model. We show that estimation is unbiased under the corrected profile likelihood for recurrent events compared to biased estimations under a likelihood function that ignores correction. We propose methods for obtaining estimators of error variance and discuss the properties of the estimators. We further investigate the case of misspecified error models and show that the resulting estimators under misspecification converge to a value different from that of the true parameter–thereby providing a basis for bias assessment. Finally, we demonstrate the preceding correction methods on an open-source rhDNase dataset gathered in a clinical setting.

Keywords: Recurrent events; Covariate measurement errors; Model misspecification; Bootstrap; Kullback-Leibler divergence; Corrected score.

1. INTRODUCTION

A recurrent event process is a process that repeatedly generates events serially. It is encountered in many fields, such as biomedical science, epidemiology, social science, reliability, and actuarial science, to name a few. The literature on methods and models that address various scientific questions about recurrent event processes is well-known. Regardless of the problem of interest, analyses of recurrent events can be broadly classified as either gap-time analysis or time-to-event analysis. Under these two paradigms, the focus can be put on estimating the intensity function $\lambda(\cdot)$, the survivor function $\bar{F}(\cdot)$, the mean rate function $\mu(\cdot)$, or other functionals of these unknowns in the presence of possible time-dependent covariates. Because covariates play an essential role in better understanding time to failure, incorporating them in modeling has always been beneficial. Denote $\mathbf{x}$ a p-dimensional vector of possibly time varying covariates, $\boldsymbol{\beta}$ a p-dimensional regressor, and $\lambda_0(\cdot)$ a baseline hazard function. For a vector $\mathbf{a}$, we call $\mathbf{a}'$ its transpose. Models that account for covariates include the multiplicative intensity model $\lambda(s) = \lambda_0(s)\exp(\boldsymbol{\beta}'\mathbf{x}(s))$ following Cox (1972), the additive model $\lambda(s) = \lambda_0(s)+\boldsymbol{\beta}'\mathbf{x}(s)$, the mean rate model $\mu(s) = \mu_0(s)\exp(\boldsymbol{\beta}'\mathbf{x}(s))$, the additive-multiplicative model $\lambda(s) = g(\boldsymbol{\beta}'\mathbf{x}(s)) + \lambda_0(s)h(\boldsymbol{\beta}'\mathbf{x}(s))$, and the accelerated failure time model $\lambda(t) = \mathbf{x}'\boldsymbol{\beta} + w$ as few examples. With specific choices of $g(\cdot)$ and $h(\cdot)$ in the additive-multiplicative model, one can easily retrieve the Cox and the additive model. The choice of a model depends on the area of application, the area of research interest, and its goodness of fit. What has been abundant in the literature is model specifications in which all covariates are assumed to be perfectly measured during data collection or study design. However, the notion that all covariates are perfectly measured is far-fetched in most research fields. A simple fact well-known to medicine is that blood pressure measured in healthcare facilities, for example, suffers from the psychological effect of subjects being in a doctor's office even if the instrument used appears to be well-calibrated. The same is true for measuring household lead levels, an error-prone process usually influenced by environmental factors, including air quality, dust movement, and soil quality. Similarly, measuring nutrient intake has been a long-documented error-prone process with measurement errors that can negatively impact health. Covariates measurement errors can be associated with the mechanism by which the measurements are taken, the situations under which they are measured, environmental factors, human errors, missingness, or with many other lurking factors.

Ignoring the errors in modeling and inference could lead to biased parameter estimates, distorted inference, and inaccurate conclusions. There has been extensive work on the topic with classical, censored, truncated, and uncensored data, resulting in numerous correction methods in models that include measurement errors. The correction methods are either parametric or nonparametric, wherein the error term acts on the true value of the covariates in an additive, multiplicative, or some other fashion. The parametric handling of the problem assumes a distribution of the error terms. In contrast, the nonparametric relax that assumption and uses replicate surrogates and instrumental variable for error correction. Regardless of the approaches and data type, the following references are noteworthy: Prentice (1982), Stefanski (1985), Armstrong (1990), Nakamura (1990), Nakamura (1992), Hu et al. (1998), Hu and Lin (2002), Song and Huang (2005), Yan and Yi (2015), Huque et al.

(2016), and Alexeeff et al. (2016). A comprehensive review of methods for measurement errors can be found in Carroll et al. (1995). Textbooks dealing with measurement errors include Fuller (1987), Carroll et al. (1995), and most recently Yi (2017).

Though the literature has been abundant with single events about measurement error models and correction methods, there needs to be more literature on recurrent events. Turnbull (1997), under a normal assumption for the errors, proposed a moment-based method for correcting a naive estimator. In contrast, Hu and Lin (2004) corrected a partial score function under a symmetric distribution for the errors. A few manuscripts on the topic with recurrent events have appeared in the last decade. Yi and Lawless (2012) using the correction from Nakamura (1990) and the simulation extrapolation (SIMEX) of Stefanski and Cook (1995), presented methods for modeling time to events which account for measurement errors under a broad class of models for hazard. In their approach, parameters were estimated using likelihood-based tools and estimating equations. More recently, some authors focused on the measurement errors problem with recurrent events while simultaneously dealing with informative censoring. For instance, Yu et al. (2016) developed a regression calibration and moment-corrected approach to adjust for measurement errors while accounting for informative censoring. Yu et al. (2016) modeled time to event and incorporated informative censoring using a shared frailty model. Yu et al. (2018), on the other hand, modeled informative censoring using a shared frailty as in Yu et al. (2016), but relaxed the distributional assumption on the errors and proposed a general nonparametric missing at random model to account for the errors. To the best of our knowledge, the most recent manuscript dealing with recurrent events is that of Chen and Yi (2021). They investigated another aspect of failure time data, namely left truncation. Since right censored and left truncated data are prevalent in practice, Chen and Yi (2021) developed models for simultaneously handling both data features while modeling measurement errors using moment correction.

To correct the bias induced by measurement errors in estimating model parameters with failure time data, researchers have relied on the so-called induced hazard rates, which are defined as the conditional hazard given the observed covariates and the events history. The idea is to construct an unbiased score function, called *corrected score*, upon which estimation and inference are based rather than naive score, *i.e.*, the score ignoring measurement errors, using an error-prone covariate, which tends to yield spurious results of no practical value. The main contributors to this idea in single event settings are Prentice (1982), Nakamura (1990), and Nakamura (1992). Many authors have shown that the corrected score is not the gradient of a corrected likelihood, leading Nakamura (1992) to introduce the concept of *approximately corrected partial likelihood*. Later, Augustin (2004) justified that the corrected score proposed in Nakamura (1992) and Kong and Gu (1999) are exact and that their corresponding estimators are consistent. In light of the recent clarifications and based on the results in Augustin (2004), we take the approach of corrected partial likelihood and consider a gap-time modeling of the intensity function with recurrent events when one or more covariates are measured with errors. We operate under the classical additive measurement errors model known to have broad applicability in scientific research. We make the blanket assumption that the errors fluctuate around the covariates. Other general measurement error models such as regression calibration in Wang et al. (1997), Yu and Nan (2010), and Chapter 4 of Carroll et al. (1995) could also be used. In our current development, we do not impose any distributional assumption on the errors other than their variance-covariance matrix being time-independent and possessing a consistent estimate. When properly standardized, we propose a corrected partial likelihood score process with root consistent estimators that follow a multivariate normal large sample distribution. Our results generalize those of Kong and Gu (1999) in two ways: (1) they are applied to recurrent event data by focusing on their stochastic feature, and in our case, (2) the errors do not follow any particular distribution. Kong and Gu (1999) deals with single event models with assumed normally distributed errors. Moreover, we augment our work by deriving a corrected baseline cumulative hazard with recurrent events and its asymptotic properties. We add a discussion on misspecified error models and develop properties of the estimators under error misspecification. We show, in that case, that the estimator converges to a value different from the true parameter, thereby allowing an assessment of bias and its magnitude.

This part of the dissertation proceeds as follows: Section 2 states the stochastic setting for our recurrent event model and defines some key concepts that guide our theoretical development along with preliminary results. Section 3 proposes the corrected estimator, including the corrected estimating equations, and develops their key asymptotic properties. We follow Section 4 with a simulation study that validates the theoretical properties and illustrates our findings in an applied setting. Section 5 proposes an estimator for the corrected baseline hazard and its asymptotic properties. In Section 6, we focus on the issue of the error model's misspecification and asymptotic bias. We finally conclude our comprehensive work with a conclusion, discussion, and recommendations section.

2. PRELIMINARIES

2.1. DYNAMIC MODELING AND OBSERVABLES

Consider n units that are monitored for an event that can recur up to a random time τ_i for each unit i. For unit i, let $S_{i,j}$ be the time of occurrence of the j^{th} event, and $T_{i,j}$ the gap between the $(j-1)$ and the j^{th} occurrence. For every unit i, a p-dimensional time-varying covariates $\mathbf{x}_i$ is recorded. We assume, for $j = 1, ..., K_i$, that the $T_{i,j}$s have a distribution function $F_i(\cdot)$, and a hazard function $\lambda_i(t) \equiv \lambda(t \mid \mathbf{x})$ that is a function of the covariates; and K_i being the total number of events per unit. For unit i, the observables at the censoring time τ_i are

$$\boldsymbol{O}_i = (K_i, \tau_i, T_{i,1}, \ldots, T_{i,K_i}, \tau_i - S_{i,K_i}, \mathbf{x}_{i,1}(s_1), ..., \mathbf{x}_{i,K_i}(s_{K_i})), \tag{1}$$

which define an aggregate vector $\boldsymbol{O} = (\boldsymbol{O}_1, ..., \boldsymbol{O}_n)$ over the observed sample of size n. In what follows, s represents calendar time, whereas t represents gap time. The relevant counting processes from the data on which estimation will be conducted are $N_i^\dagger = \{N_i^\dagger(s) : s \leq \tau_i\}$, $Y_i^\dagger = \{Y_i^\dagger(s) : s \geq 0\}$, where $N_i^\dagger(s) = \sum_{j=1}^{\infty} I\{S_{i,j} \leq s \wedge \tau_i\}$, and $Y_i^\dagger(s) = I\{\tau_i \geq s\}$. The process $N_i^\dagger(s)$ determines, for subject i, the event occurrences up to time s whereas the $Y_i^\dagger(s)$ process determines if the unit is at-risk for future recurrences. We write the effective age of the unit i at time s as $\varphi_i(s) = s - S_{i,N_i^\dagger(s^-)}$. Observe that $0 \leq \varphi(s) \leq s$, and is viewed as a process that keeps track of the time elapsed since the last occurrence of an event. In φ, we not only track the time elapsed but also verify whether the time being tracked is a true event time versus another time with the hopes of a future event. Following Cox (1972), we provide a link between the effective age process, the regressors, and the multiplicative intensity process by $\lambda(s) = \lambda_0(\varphi_i(s)) \exp(\boldsymbol{\beta}'\mathbf{x}_i(s))$ where $\boldsymbol{\beta} \in \mathcal{R}^p$ is a set of regressor parameters. The utility of the effective age rendering of the Cox model is that it allows one to explicitly model the effect of an intervention or treatment performed just after an event occurrence, an adaptation that is more in harmony with settings where treatments are administered during an observation window such as in the health science; see Beutner et al. (2020) . From the theory of stochastic integration, the compensator process of $N_i^\dagger(s)$ is $A_i^\dagger(s)$ given by $A_i^\dagger(s|\boldsymbol{\beta}) = \int_0^s Y_i^\dagger(v)\lambda_0[\varphi_i(v)] \exp[\boldsymbol{\beta}'\mathbf{x}_i(v)]dv$. Due to the randomness of the argument $\varphi_i(v)$ in $A_i^\dagger(s|\boldsymbol{\beta})$, one usually works with the doubly-indexed processes $N_i(s,t)$ and $A_i(s,t|\boldsymbol{\beta})$ that are functions of both the calendar and gap times in order to handle the random argument in the integrand. The $N_i(s,t)$ and $A_i(s,t|\boldsymbol{\beta})$ processes are the number of events that occurred by calendar time s for unit i whose effective age is at most gap time t and its associated doubly-indexed compensator respectively. Hence, for fixed t, $M_i(s,t|\boldsymbol{\beta}) = N_i(s,t) - A_i(s,t|\boldsymbol{\beta})$ is a zero-mean square integrable martingale. More notations and details on stochastic processes formulation in this section can be found in Pena et al. (2001), Adekpedjou and Stocker (2015), Zamba and Adekpedjou (2019), and references therein for interested readers.

2.2. MEASUREMENT ERRORS NOTATIONS

Under the additive model for errors, for unit i, at calendar time s, let $\mathbf{x}_i(s)$ be the p-dimensional covariates, possibly $q \leq p$ are measured with errors. If $\boldsymbol{\epsilon}_i(s) = (\epsilon_{i1}(s), ..., \epsilon_{ip}(s))$ is the p-dimensional errors on $\mathbf{x}_i(s) = (x_{i1}(s), ..., x_{ip}(s))$, then

$$\mathbf{x}_i(s) = \mathbf{z}_i(s) + \boldsymbol{\epsilon}_i(s),$$

where the $\mathbf{z}_i(s)$ are the true and unobserved covariates. We do not impose any distributional assumption on the $\boldsymbol{\epsilon}_i(s)$, other than having a zero mean and a variance-covariance matrix $\boldsymbol{\Xi}$ that is time-independent. We assume the existence of a consistent estimator $\hat{\boldsymbol{\Xi}}$ of $\boldsymbol{\Xi}$ which can be derived using validation methods or replicates (cf. Carroll et al. (1995), Chapter 4). It is to be noted that some authors have assumed the errors to have a multivariate normal distribution with error variance obtained using validation data and sample variance formulas; cf. section 2.1 of Yi and Lawless (2012). See also section 4.7 of Carroll et al. (1995) on ways to derive estimators for the error variance.

We now introduce some notation in the sequel to be used throughout this manuscript. All random entities are defined on a complete probability space $(\Omega, \mathcal{F}, \mathbf{P})$. The space $D[0, t]$ denotes the cadlag functions on $[0, t]$ equipped with the supremum norm $\|\cdot\|_\infty$; and all asymptotic results are taken as $n \longrightarrow \infty$. The notations $\overset{d}{\longrightarrow}$, $\overset{as}{\longrightarrow}$, and $\overset{p}{\longrightarrow}$ respectively denote convergence in distribution, almost sure convergence, and convergence in probability. For a vector $\mathbf{a} = (a_1, \ldots, a_p)$, $\mathbf{a}'$ is its transpose; and if $\mathbf{b}$ is also a p-dimensional vector, then $\mathbf{a} \otimes \mathbf{b}$ is the $p \times p$ matrix $\mathbf{ab}'$ with $(k, l)^{th}$ element $a_k b_l$. In addition, $\mathbf{a}^{\otimes 2} = \mathbf{a} \otimes \mathbf{a}$, and for the vector $\mathbf{a}$, $\|\mathbf{a}\| = \sup_k |a_k|$. Finally, we define the gradient operator $\nabla_{\boldsymbol{\beta}}$ by the vector of partial derivatives $\nabla_{\boldsymbol{\beta}} = \frac{\partial}{\partial \boldsymbol{\beta}} \equiv (\partial/\partial\beta_l, l = 1, 2, \ldots, p)'$ if $\boldsymbol{\beta}$ a p-dimensional vector.

3. CORRECTED SCORES AND ESTIMATORS

3.1. PRELIMINARY

Some preliminary results are contained in Adekpedjou and Stocker (2015); consequently, we will not provide another complete exposition here. With $s^\star = \max_{1\le i\le n} \tau_i$, the generalized at risk process is defined as

$$\begin{aligned} Y_i(s, \mathbf{x}(t)|\boldsymbol{\beta}) &= \left[\sum_{j=1}^{N_i^\dagger((s\wedge\tau_i)-)} I(T_{i,j} \ge t) + I((s \wedge \tau_i) - S_{i,N_i^\dagger((s\wedge\tau_i)-)} \ge t) \right] \cdot \exp(\boldsymbol{\beta}'\mathbf{x}_i(t)) \\ &:= h_i(s,t) \cdot \exp(\boldsymbol{\beta}'\mathbf{x}_i(t)). \end{aligned}$$

We write $S^{(0)}(s, \mathbf{x}(t)|\boldsymbol{\beta}) := n^{-1} \sum_{i=1}^n Y_i(s, \mathbf{x}(t)|\boldsymbol{\beta})$, and its k^{th} order derivative with respect to $\boldsymbol{\beta}$ written as

$$\begin{aligned} S^{(k)}(s, \mathbf{x}(t)|\boldsymbol{\beta}) &= \frac{1}{n} \boldsymbol{\nabla}_{\boldsymbol{\beta}}^{(k)} \sum_{i=1}^n h_i(s,t) \cdot \exp(\boldsymbol{\beta}'\mathbf{x}_i(t)) \\ &= \frac{1}{n} \sum_{i=1}^n h_i(s,t) \cdot \mathbf{x}_i(t)^{\otimes k} \exp(\boldsymbol{\beta}'\mathbf{x}_i(t)). \end{aligned}$$

For $k = 0, 1, 2$, write

$$E(s, \mathbf{x}(t)|\boldsymbol{\beta}) = S^{(1)}(s, \mathbf{x}(t)|\boldsymbol{\beta}) / S^{(0)}(s, \mathbf{x}(t)|\boldsymbol{\beta})$$

and

$$V(s, \mathbf{x}(t)|\boldsymbol{\beta}) = [S^{(2)}(s, \mathbf{x}(t)|\boldsymbol{\beta}) / S^{(0)}(s, \mathbf{x}(t)|\boldsymbol{\beta})] - E(s, \mathbf{x}(t)|\boldsymbol{\beta})^{\otimes 2}.$$

The partial log-likelihood is written as

$$l_p(\boldsymbol{\beta}, \mathbf{x}(s^\star)) = \sum_{i=1}^n \int_0^{s^\star} \left[\boldsymbol{\beta}'\mathbf{x}_i(v) - \log[Y_i(s^\star, \varphi_i(v)|\boldsymbol{\beta})] \right] N_i^\dagger(dv), \tag{2}$$

where the integral is obtained over the calendar time process. Because $\varphi_i(v)$ is a random variable, a change of variable $\varphi_i(v) = w$ in the integrand leads to a likelihood profile $l_P(\boldsymbol{\beta}, \mathbf{x}(s^\star))$ and an associated score process $\mathbf{U}(\boldsymbol{\beta}, \mathbf{x}(s^\star)) = \nabla_{\boldsymbol{\beta}} l_P(\boldsymbol{\beta}, \mathbf{x}(s^\star))$ given by

$$\mathbf{U}(\boldsymbol{\beta}, \mathbf{x}(t^\star)) = \sum_{i=1}^{n} \int_0^{t^\star} [\mathbf{x}_i(\varphi_i^{-1}(w)) - E(s^\star, w)] N_i(s^\star, dw), \tag{3}$$

where $t^\star = \max_{i,j} T_{i,j}$ is the maximum gap time. The estimator of $\boldsymbol{\beta}$, $\hat{\boldsymbol{\beta}}$ say, is the solution to $\mathbf{U}(\boldsymbol{\beta}, \mathbf{x}(t^\star)) = \mathbf{0}$, with predictors assumed error-free. Note that $\mathbf{x}_i(s)$ are not the true covariates; consequently, the solution $\hat{\boldsymbol{\beta}}$ are biased by virtue of the biasedness of $\mathbf{U}(\boldsymbol{\beta}, \mathbf{x}(t^\star))$. As $\mathbf{U}(\hat{\boldsymbol{\beta}}, \mathbf{x}(t^\star))$ does not equal zero in the presence of measurement errors, the corresponding likelihood and score, which are functions of $S^{(k)}(s^\star, \mathbf{x}(t)|\boldsymbol{\beta})$, $k = 0, 1$ are also biased and cannot portray a reasonable estimation mechanism. We provide a corrected expression for $\mathbf{U}(\boldsymbol{\beta}, \mathbf{x}(t^\star))$, which is inextricably linked to corrections of $S^{(0)}(s^\star, \mathbf{x}(t)|\boldsymbol{\beta})$ and $S^{(1)}(s^\star, \mathbf{x}(t)|\boldsymbol{\beta})$. We denote $E_\epsilon(\cdot|\boldsymbol{O})$, the conditional expectation under the distribution of $\boldsymbol{\epsilon}$ with respect to the observables $\boldsymbol{O}$. In the spirit of corrected likelihoods, we seek estimating functions $\breve{\mathbf{U}}(\boldsymbol{\beta}, \mathbf{x}(t))$, expressed in terms of the observed data and satisfying

$$E_\epsilon\left(\breve{\mathbf{U}}(\boldsymbol{\beta}, \mathbf{x}(t)|\boldsymbol{O})\right) = \mathbf{U}(\boldsymbol{\beta}, \mathbf{z}(t)). \tag{4}$$

This will be handled through the use of the first-order approximation of the ratio of expectations. We also operate under some regularity conditions, which are detailed in the technical appendix section. Under these regularity conditions, $\boldsymbol{\epsilon}_i$'s are independent zero-mean with time-independent covariance structure; all order moment generating functions of the error distribution exist; the integrated hazard is finite; the covariates are uniformly bounded and cannot escape to infinity; and uniform continuity of $S^{(k)}(s^\star, \mathbf{x}(t)|\boldsymbol{\beta})$ and their expectations is guaranteed.

We now introduce, the regularity conditions required for the proofs and to establish large sample properties.

Regularity Conditions A:

I. The $\epsilon_i(t)$s are independent with mean zero and independent of $\boldsymbol{O}$.

II. Var$(\boldsymbol{\epsilon}_i(t)) = \boldsymbol{\Xi}$ and is time independent.

III. The moment-generating function at all orders exists.

Assumptions I, II, and III are regular assumptions imposed on measurement errors. Assumption I is a trivial assumption in that errors, in general, should not deviate much from the true value and should average to zero. In addition, errors in one unit do not indicate errors on the next one, hence the independence between errors and observables. Assumption II is given to simplify the calculation of the variance-covariance matrix of the large sample distribution of the corrected error properly standardized. As for III, it is the usual assumption on the moment-generating function. It is needed for the large sample properties of the corrected score since the asymptotic properties require the in probability limit of the estimator of the moment-generating function.

Regularity Conditions B:

The regularity conditions I, II, III, IV on page 6 of Adekpedjou and Stocker (2015) with $\mathbf{x}_i$ replaced by $\mathbf{z}_i$ are in force, namely

I. $\int_0^{t^\star} \lambda_0(w)dw < \infty$.

II. $(\mathbf{X}_i(s)\colon s \le \tau_i)$ is uniformly bounded for $s \in [0, \tau_i]$ for all i.

III. For $k = 0, 1, 2$ there exists functions $s^{(k)}(s^\star, t|\boldsymbol{\beta}) = E(S^{(k)}(s^\star, t|\boldsymbol{\beta}))$ and a neighborhood $\mathcal{B}$ of $\boldsymbol{\beta}_0$ such that the functions are continuous functions of $\boldsymbol{\beta} \in \mathcal{B}$ uniformly in $t \in [0, t^\star]$ and bounded on $[0, t^\star] \times \mathcal{B}$. Furthermore,

$$\sup_{\substack{t\in[0,t^\star] \\ \boldsymbol{\beta}\in\mathcal{B}}} \left\| S^{(k)}(s, t|\boldsymbol{\beta}) - s^{(k)}(s, t|\boldsymbol{\beta}) \right\| \xrightarrow{p} 0$$

and

$$s^{(1)}(s^\star, t|\boldsymbol{\beta}) = \nabla_{\boldsymbol{\beta}} s^{(0)}(s^\star, t|\boldsymbol{\beta}) \quad \text{and} \quad s^{(2)}(s^\star, t|\boldsymbol{\beta}) = \nabla_{\boldsymbol{\beta}\boldsymbol{\beta}^t} s^{(0)}(s^\star, t|\boldsymbol{\beta}).$$

IV. Let $\mathcal{B}$, $s^{(0)}(s^\star, t|\boldsymbol{\beta})$, $s^{(1)}(s^\star, t|\boldsymbol{\beta}), s^{(2)}(s^\star, t|\boldsymbol{\beta})$ be as in Condition III. For all $\boldsymbol{\beta} \in \mathcal{B}$ and $t \in [0, t^\star]$ define

$$e(s^\star, t|\boldsymbol{\beta}) = \frac{s^{(1)}(s^\star, t|\boldsymbol{\beta})}{s^{(0)}(s^\star, t|\boldsymbol{\beta})} \quad \text{and} \quad v(s^\star, t|\boldsymbol{\beta}) = \frac{s^{(2)}(s^\star, t|\boldsymbol{\beta})}{s^{(0)}(s^\star, t|\boldsymbol{\beta})} - e(s^\star, t|\boldsymbol{\beta})^{\otimes 2}.$$

Assume that $s^{(0)}(s^\star, t|\boldsymbol{\beta})$ is bounded away from 0 on $[0, t^\star] \times \mathcal{B}$; there exists a positive-definite matrix $\boldsymbol{\Sigma}_1(\boldsymbol{\beta}_0; s^\star, t)$ such that

$$\sup_{t\in[0,t^\star]} \left\| \frac{1}{n} \sum_{i=1}^{n} \int_0^t \left[\mathbf{X}_i(\varphi_i^{-1}(w)) - E(s^\star, w|\boldsymbol{\beta}_0)\right]^{\otimes 2} S^{(0)}(s^\star, w|\boldsymbol{\beta}_0)\lambda_0(w)dw - \boldsymbol{\Sigma}_1(\boldsymbol{\beta}_0; s^\star, t) \right\|$$

converges in probability to 0; and the matrix

$$\boldsymbol{\Sigma}(\boldsymbol{\beta}_0; s^\star, t^\star) = \int_0^{t^\star} v(s^\star, w|\boldsymbol{\beta}_0) s^{(0)}(s^\star, w|\boldsymbol{\beta}_0)\lambda_0(w)dw$$

is positive-definite. In Regularity condition B, the $E(\cdot,\cdot), V(\cdot,\cdot)$, and $S(\cdot,\cdot)$ functions are the regular functions that arise in the analysis of failure time data. The $E(\cdot,\cdot)$ and $V(\cdot,\cdot)$ are the expectation and variance, respectively, of the covariates when they are truly observed, whereas the $S^{(k)}(\cdot,\cdot)$, $k = 0, 1, 2$ are normalizing constants.

Proposition 3 *The following proposition holds:*

$$E(s^\star, \mathbf{x}(t)|\boldsymbol{\beta}) = \frac{S^{(1)}(s^\star, \mathbf{z}(t))}{S^{(0)}(s^\star, \mathbf{z}(t))} + \boldsymbol{\nabla}_{\boldsymbol{\beta}} \ln(\phi(\boldsymbol{\beta})).$$

Proof: Under regularity condition I,

$$\begin{aligned} E_{\boldsymbol{\epsilon}}(S^{(0)}(s^{\star},\mathbf{x}(t)|\boldsymbol{O})) &= E_{\boldsymbol{\epsilon}}\left(\frac{1}{n}\sum_{i=1}^{n} h_i(s,t)\cdot\exp[\boldsymbol{\beta}'(\mathbf{z}_i(t)+\boldsymbol{\epsilon}_i(t))]\right) \\ &= E_{\boldsymbol{\epsilon}}\left(\exp(\boldsymbol{\beta}'\boldsymbol{\epsilon}_i(t))\right)\cdot E_{\boldsymbol{\epsilon}}\left(\frac{1}{n}\sum_{i=1}^{n} h_i(s,t)\exp(\boldsymbol{\beta}'\mathbf{z}_i(t))\right) \\ &= \phi(\boldsymbol{\beta})\cdot S^{(0)}(s^{\star},\mathbf{z}(t))+o_p(1). \end{aligned}$$

Likewise, using the fact that $\boldsymbol{\nabla}_{\boldsymbol{\beta}}\phi_i(\boldsymbol{\beta}) = E(\boldsymbol{\epsilon}_i(t)\exp(\boldsymbol{\beta}'\boldsymbol{\epsilon}_i(t))$, we obtain

$$E_{\boldsymbol{\epsilon}}(S^{(1)}(s^{\star},\mathbf{x}(t)|\boldsymbol{O}) = \phi(\boldsymbol{\beta})\cdot S^{(1)}(s^{\star},\mathbf{z}(t)) + \boldsymbol{\nabla}_{\boldsymbol{\beta}}\phi(\boldsymbol{\beta})S^{(0)}(s^{\star},\mathbf{z}(t)).$$

Hence, applying a first order approximation of $E_{\boldsymbol{\epsilon}}(S^{(k)}(s^{*},\mathbf{x}(t)|\boldsymbol{O})$ for $k=0,1$ at $\left[E_{\boldsymbol{\epsilon}}(S^{(0)}(s^{\star},\mathbf{x}(t)|\boldsymbol{O}), E_{\boldsymbol{\epsilon}}(S^{(1)}(s^{\star},\mathbf{x}(t)|\boldsymbol{O})))\right]$, we obtain

$$\begin{aligned} E_{\boldsymbol{\epsilon}}\left(E(s^{\star},\mathbf{x}(t)|O)\right) &= \frac{E_{\boldsymbol{\epsilon}}\left(S^{(1)}(s^{\star},\mathbf{x}(t))|\boldsymbol{O}\right)}{E_{\boldsymbol{\epsilon}}\left(S^{(0)}(s^{\star},\mathbf{x}(t)|\boldsymbol{O})\right)} \\ &= \frac{\phi(\boldsymbol{\beta})\cdot S^{(1)}(s^{\star},\mathbf{z}(t)) + \boldsymbol{\nabla}_{\boldsymbol{\beta}}\phi(\boldsymbol{\beta})S^{(0)}(s^{\star},\mathbf{z}(t))}{\phi(\boldsymbol{\beta})\cdot S^{(0)}(s^{\star},\mathbf{z}(t))} \\ &= \frac{S^{(1)}(s^{\star},\mathbf{z}(t))}{S^{(0)}(s^{\star},\mathbf{z}(t))} + \boldsymbol{\nabla}_{\boldsymbol{\beta}}[\ln(\phi(\boldsymbol{\beta}))]. \end{aligned}$$

So, the corrected score to be used for unbiased estimating equations based on the observables is

$$\breve{\mathbf{U}}(\boldsymbol{\beta};s^{\star},t) = \sum_{i=1}^{n}\int_0^t [\mathbf{x}_i(\varphi_i^{-1}(w)) - E(s^{\star},\mathbf{x}(w)) + \boldsymbol{\nabla}_{\boldsymbol{\beta}}\ln(\phi(\boldsymbol{\beta}))]N_i(s^{\star},dw). \tag{5}$$

The solution $\breve{\boldsymbol{\beta}}$ to the estimating equations $\breve{\mathbf{U}}(\boldsymbol{\beta};s^{\star},t) = \mathbf{0}$, is the corrected maximum profile likelihood estimator. In a like manner, one can also derive the corrected information matrix as

$$\mathcal{I}_n(\breve{\boldsymbol{\beta}};s^{\star},t) = -\boldsymbol{\nabla}_{\boldsymbol{\beta}}[\breve{\mathbf{U}}(\boldsymbol{\beta};s^{\star},t)]\Big|_{\breve{\boldsymbol{\beta}}}.$$

Note that when $\boldsymbol{\epsilon}(t) \equiv \mathbf{0}$ in (5), one recovers the no measurement errors covariates as in Adekpedjou and Stocker (2015). Similar to the argument about exact corrected scores and root consistency as those put forth in Augustin (2004), Nakamura (1992), and Kong and Gu (1999), we show that the results hold for recurrent events as well.

3.2. LARGE SAMPLE PROPERTIES OF CORRECTED VALUES

In order to establish the convergence in distribution of the standardized version of $\breve{\boldsymbol{\beta}}$, we would need the large sample properties of the scaled corrected score $\sqrt{n}^{-1}\breve{\mathbf{U}}(\boldsymbol{\beta}; s^{\star}, t)$. Assume $\boldsymbol{\beta}_0$ is the true value of $\boldsymbol{\beta}$, and $\mathbf{z}_i(t)$ are the true covariates free of measurement errors; then, the process

$$M_i(s,t) = N_i(s,t) - \int_0^t Y_i(s, \mathbf{z}(w)|\boldsymbol{\beta}_0)\lambda_0(w)dw$$

is, for fixed t, a zero-mean martingale with respect to the calendar time filtration $\mathfrak{F} = \{\mathfrak{F}_s : s \geq 0\}$, with $\mathfrak{F}_s$ the σ-field generated by $\{[(N_i^{\dagger}(s), Y_i^{\dagger}(s+)) : s \geq 0], i = 1, 2, \ldots, n\}$. The first result in this section pertains to the consistency of the corrected maximum likelihood $\breve{\boldsymbol{\beta}}$.

Theorem 3 *As $n \to \infty$, the sequence of solutions $\breve{\boldsymbol{\beta}}_n = (\breve{\beta}_{1n}, ..., \breve{\beta}_{np})$ to the sequence of equations $\breve{\boldsymbol{U}}_n(\boldsymbol{\beta}; s^{\star}, t) = \mathbf{0}$, $n = 1, 2, ...$ is consistent.*

Proof: The proof is based on a simplified functional uniform convergence argument for $M-$estimators as used in theorem 5.9 of van der Vaart (1998), by showing two conditions needed to establish consistency, namely convergence in probability (i) and separability of the root (ii). To show condition (i), observe that

$$\begin{aligned}\frac{1}{n}\breve{\mathbf{U}}(\boldsymbol{\beta}; s^\star) &= \frac{1}{n}\sum_{i=1}^{n}\int_0^{s^\star}[\mathbf{x}_i(w) - E(s^\star, \varphi_i(w); \mathbf{x}(w)) + \boldsymbol{\nabla}_{\boldsymbol{\beta}}\ln(\phi(\boldsymbol{\beta}))]M_i(dw|\boldsymbol{\beta}_0) \quad (6)\\ &\quad +\frac{1}{n}\sum_{i=1}^{n}\int_0^{s^\star}[\mathbf{x}_i(w) - E(s^\star, \varphi_i(w); \mathbf{x}(w)) + \boldsymbol{\nabla}_{\boldsymbol{\beta}}\ln(\phi(\boldsymbol{\beta}))]A_i^\dagger(dw|\boldsymbol{\beta}_0)\\ &= T_1 + T_2.\end{aligned}$$

We now use the time-transformed processes. Since $M_i(s^\star, dw|\boldsymbol{\beta}_0)$ is a zero mean martingale for fixed t, transforming the first term in (6), we see that it is $o_p(1)$. Next, replacing $\mathbf{x}_i(w)$ by $\mathbf{z}_i(w)+\boldsymbol{\epsilon}_i(w)$ and using the fact that $E(s^\star, \mathbf{x}(t)) = E(s^\star, \mathbf{z}(t))+\boldsymbol{\nabla}_{\boldsymbol{\beta}}[\ln(\phi(\boldsymbol{\beta}))]$, the second term becomes

$$\begin{aligned}T_2 &= \frac{1}{n}\sum_{i=1}^{n}\int_0^{t^\star}[\mathbf{z}_i(\varphi_i^{-1}(w)) - E(s^\star, \mathbf{z}(w))]A_i^\dagger(s, dw|\boldsymbol{\beta}_0)\\ &\quad +\frac{1}{n}\sum_{i=1}^{n}\int_0^{t^\star}\boldsymbol{\epsilon}_i(\varphi_i^{-1}(w))A_i^\dagger(s, dw|\boldsymbol{\beta}_0)\\ &\quad -\int_0^{s^\star}E(s^\star, \mathbf{z}(w))S^{(0)}(s^\star, \mathbf{z}(w))\lambda_0(w)dw\\ &\xrightarrow{p} \int_0^{t^\star}[s^{(1)}(s^\star, \mathbf{z}(w)|\boldsymbol{\beta}) - e(s^\star, \mathbf{z}(w))s^{(0)}(s^\star, \mathbf{z}(w)|\boldsymbol{\beta}_0)]\lambda_0(w)dw.\end{aligned}$$

Hence Condition (i) of Theorem 5.9 of van der Vaart (1998) has been established. For Condition (ii), note that the corrected score $\breve{\mathbf{U}}(\boldsymbol{\beta}; s^\star, t)$ is a continuous function and is $\mathbf{0}$ at $\boldsymbol{\beta}_0$. The partial derivative of the in-probability limit of T_2 is

$$-\int_0^{t^\star} s^{(0)}(s^\star, w|\boldsymbol{\beta}_0)\left[\frac{s^{(2)}(s^\star, w|\boldsymbol{\beta}_0)}{s^{(1)}(s^\star, w|\boldsymbol{\beta}_0)} - \left(\frac{s^{(1)}(s^\star, w|\boldsymbol{\beta})}{s^{(0)}(s^\star, w|\boldsymbol{\beta})}\right)^{\otimes 2}\right]\lambda_0(w)dw, \qquad (7)$$

which is negative definite at $\boldsymbol{\beta} = \boldsymbol{\beta}_0$. Hence $\breve{\boldsymbol{\beta}}$ is a global maximum, thus concluding the consistency of $\boldsymbol{\beta}$. $\|$

The moment-generating function of the $\boldsymbol{\epsilon}_i(s)$, $\phi_i(\boldsymbol{\beta}) = E(e^{\boldsymbol{\beta}'\boldsymbol{\epsilon}_i(s)})$ plays a key role in the asymptotic properties of corrected estimators, the corrected score, and corrected Fisher information. In the next few lines, we discuss its large sample property. For every i, note that $\boldsymbol{\beta}'\boldsymbol{\epsilon}_i(s) = \sum_{j=1}^{p} \beta_j \epsilon_{ij}(s)$. The function $\phi_i(\boldsymbol{\beta})$ is assumed to exist for $\boldsymbol{\beta}$ in the neighborhood of zero and is twice differentiable with respect $\boldsymbol{\beta}$. If a consistent estimator $\hat{\boldsymbol{\beta}}$ of $\boldsymbol{\beta}$ exists, the following lemma asserts the consistency of an estimator of $\phi(\boldsymbol{\beta})$ namely $\hat{\phi}(\boldsymbol{\beta}) := \phi(\hat{\boldsymbol{\beta}})$, the empirical moment generating function.

Lemma 2 *Let $\hat{\boldsymbol{\beta}}$ be the corrected and consistent maximum likelihood estimator of $\boldsymbol{\beta}$. The empirical moment generating function of $\phi(\boldsymbol{\beta})$ based on $\{\phi_i(\boldsymbol{\beta}) : i = 1, ..., n\}$ is defined by*

$$\hat{\phi}(\boldsymbol{\beta}) = \phi(\hat{\boldsymbol{\beta}}) = \frac{1}{n}\sum_{i=1}^{n} \phi_i(\hat{\boldsymbol{\beta}}),$$

and is a consistent estimator of $\phi(\boldsymbol{\beta})$.

Proof: The result follows from the functional continuous mapping theorem, the uniform law of large numbers, and the consistency of $\hat{\boldsymbol{\beta}}$.

As indicated earlier, to establish the distribution limit of the properly standardized corrected maximum likelihood estimator, we will first establish the asymptotic properties of the corrected score $\sqrt{n}^{-1}\breve{\mathbf{U}}(\boldsymbol{\beta}; s^{\star}, t)$.

Theorem 4 *As $n \to \infty$, the process $\left\{\frac{1}{\sqrt{n}}\breve{\boldsymbol{U}}(\boldsymbol{\beta}; s^{\star}, t) : t \in [0, t^{\star}]\right\}$ converges weakly on $D[0, t^{\star}]$ to a zero-mean Gaussian process $\boldsymbol{U}^{\infty}(\boldsymbol{\beta}_0, s^{\star}, t)$ with covariance matrix given by $\Psi(s, t_1 \wedge t_2 | \boldsymbol{\beta}_0))$ that can be estimated by $\breve{\Psi}(s, t_1 \wedge t_2 | \breve{\boldsymbol{\beta}})) = E([V_i(s, t_1 \wedge t_2)^{\otimes 2}])$, with $V_i(s, t)$ given in the proof below.*

Proof: $\sqrt{n}^{-1}\breve{\mathbf{U}}(\boldsymbol{\beta}; s^{\star}, t)$ can be written as Proof: Let $A = \sqrt{n}^{-1}\breve{\mathbf{U}}(\boldsymbol{\beta}; s^{\star}, t)$ which can be written as

$$\begin{aligned} A &= \frac{1}{\sqrt{n}}\sum_{i=1}^{n}\int_0^t \left\{\mathbf{x}_i(\varphi_i^{-1}(v)) - E(s^{\star}, \mathbf{x}(v)) + \nabla_{\boldsymbol{\beta}} \ln[\phi(\boldsymbol{\beta}_0)]\right\} M_i(s^{\star}, dv|\boldsymbol{\beta}_0) \\ &+\frac{1}{\sqrt{n}}\sum_{i=1}^{n}\int_0^t h_i(s^{\star}, v)\exp(\boldsymbol{\beta}_0'\mathbf{x}_i(v))e^{-\boldsymbol{\epsilon}_i'(v)\boldsymbol{\beta}_0}\lambda_0(v)\phi(\boldsymbol{\beta}_0)\phi^{-1}(\boldsymbol{\beta}_0)e^{-\boldsymbol{\epsilon}_i(v)\boldsymbol{\beta}_0}dv. \end{aligned}$$

For large n, it can be seen that A is equal to

$$\begin{aligned} A &= \frac{1}{\sqrt{n}}\sum_{i=1}^{n}\int_0^t \left\{\mathbf{x}_i(\varphi_i^{-1}(v)) - E(s^{\star}, \mathbf{x}(v)) + \nabla_{\boldsymbol{\beta}}[\ln\phi(\boldsymbol{\beta}_0)]\right\} M_i(s^{\star}, dv|\boldsymbol{\beta}_0) \\ &+\frac{1}{\sqrt{n}}\phi^{-1}(\boldsymbol{\beta}_0)\sum_{i=1}^{n}\int_0^t h_i(s^{\star}, v)\exp(\boldsymbol{\beta}_0'\mathbf{x}_i(v))e^{-\boldsymbol{\epsilon}'(v)\boldsymbol{\beta}_0}\lambda_0(v)dv. \end{aligned} \tag{8}$$

The display in (8) can be viewed as the sum of n independent $V_i(s,t)$ processes with $V_i(s,t)$ given by

$$\begin{aligned} V_i(s,t) &= \frac{1}{\sqrt{n}}\int_0^t \left\{\mathbf{x}_i(\varphi^{-1}(v)) - E(s, \mathbf{z}(v)) + \nabla_{\boldsymbol{\beta}}[\ln\phi(\boldsymbol{\beta}_0)]\right\} M_i(s, dv|\boldsymbol{\beta}_0) \\ &+\phi^{-1}(\boldsymbol{\beta}_0)\int_0^t h_i(s, v)\exp(\boldsymbol{\beta}_0'\mathbf{x}_i(v))e^{-\boldsymbol{\epsilon}'(v)\boldsymbol{\beta}_0}\lambda_0(v)dv. \end{aligned}$$

We can then apply the functional central limit theorem to find its large sample properties. To apply the functional central limit theorem, we have to show conditions (a) to (e) of Theorem 11.16 of Kosorok (2008) about the manageability of the process and the existence of an in-probability limit of the variance-covariance matrix of the standardized process. Numerous manuscripts have shown these conditions dealing with recurrent event modeling when the covariates are not mismeasured. The addition of the constant term $\nabla_{\boldsymbol{\beta}}[\ln\phi(\boldsymbol{\beta}_0)]$ does not prevent those conditions from holding since $M_i(s, dv)$ is a zero-mean martingale.

Therefore, by Theorem 11.16 of Kosorok (2008) the process

$$\left\{\sqrt{n}^{-1}\breve{\mathbf{U}}(\boldsymbol{\beta}; s^\star, t) : t \in [0, t^\star]\right\}$$

converges to a tight $vU^\infty(\boldsymbol{\beta}_0; s^\star, t)$ with a variance covariance matrix $\Psi(s, t_1 \wedge t_2|\boldsymbol{\beta}_0))$ that can be estimated by $\breve{\Psi}(s, t_1 \wedge t_2|\breve{\boldsymbol{\beta}}))$.

With the limiting distribution of $\sqrt{n}^{-1}\breve{\mathbf{U}}_n(\boldsymbol{\beta}; s^\star, t)$ established, we can now proceed with that of the properly standardized corrected maximum likelihood estimator. To that end, we need a corrected Fisher information matrix and its in probability limit. Recall that the Fisher information matrix is given by

$$\mathcal{I}_n(\breve{\boldsymbol{\beta}}; s^\star, t) = -\frac{1}{n}\boldsymbol{\nabla}_{\boldsymbol{\beta}'}\breve{\mathbf{U}}(\breve{\boldsymbol{\beta}}; s^\star, t).$$

We have

$$\begin{aligned}
\mathcal{I}_n(\boldsymbol{\beta};^\star, t) &= -\frac{1}{n}\nabla_{\boldsymbol{\beta}'}\breve{\mathbf{U}}(\boldsymbol{\beta}; s, t) \\
&= -\frac{1}{n}\int_0^t \left\{\nabla_{\boldsymbol{\beta}'}[\nabla_{\boldsymbol{\beta}} \ln \phi(\boldsymbol{\beta})] - \frac{S^{(2)}(s^\star, t|\boldsymbol{\beta})}{S^{(0)}(s^\star, t|\boldsymbol{\beta})} + E^{\otimes 2}(s^\star, t|\boldsymbol{\beta})\right\} \sum_{i=1}^n N_i(s^\star, dw).
\end{aligned}$$

To obtain the corrected Fisher information, we need to find the corrected expression of the second and third terms in the integrand. The third corrected expression has been provided in Proposition 1. Recall that the variance covariance matrix of $\boldsymbol{\epsilon}_i$ is assumed to exist and is independent of t, that is $\boldsymbol{\Xi} = E(\boldsymbol{\epsilon}_i^{\otimes 2})$, $i = 1, ..., n$ is independent of t. The next proposition pertains to the corrected expression of the second term, namely $S^{(2)}(s^\star, \boldsymbol{x}(t)|\boldsymbol{\beta})[S^{(0)}(s^\star, \boldsymbol{x}(t)|\boldsymbol{\beta})]^{-1}$.

Proposition 4 *The corrected expression of* $S^{(2)}(s^\star, \mathbf{x}(t)|\boldsymbol{\beta})[S^{(0)}(s^\star, \mathbf{x}(t)|\boldsymbol{\beta})]^{-1}$ *can be approximated by*

$$\begin{aligned}\frac{S^{(2)}(s^\star, \mathbf{x}(t)|\boldsymbol{\beta})}{S^{(0)}(s^\star, \mathbf{x}(t)|\boldsymbol{\beta})} &\approx \frac{S^{(2)}(s^\star, \mathbf{z}(t)|\boldsymbol{\beta})}{S^{(0)}(s^\star, \mathbf{z}(t)|\boldsymbol{\beta})} + [E(s^\star, \mathbf{z}(t)|\boldsymbol{\beta})(\boldsymbol{\beta}\boldsymbol{\Xi})]' \\ &+ [E(s^\star, \mathbf{z}(t)|\boldsymbol{\beta})(\boldsymbol{\beta}\boldsymbol{\Xi})] + \boldsymbol{\Xi} + [(\boldsymbol{\Xi}\boldsymbol{\beta})(\boldsymbol{\Xi}\boldsymbol{\beta})'].\end{aligned}$$

Proof: Replacing $\mathbf{x}$ by $\mathbf{z} + \boldsymbol{\epsilon}$, $S^{(2)}(s, \mathbf{x}(t)|\boldsymbol{\beta})$ can be written as

$$\begin{aligned}S^{(2)}(s^\star, \mathbf{x}(t)|\boldsymbol{\beta}) &= \frac{1}{n}\sum_{i=1}^{n} h_i(s,t)(\mathbf{z}_i + \boldsymbol{\epsilon}_i)^{\otimes 2}\exp(\boldsymbol{\beta}'(\mathbf{z}_i + \boldsymbol{\epsilon}_i)) \\ &= \frac{1}{n}\sum_{i=1}^{n} h_i(s,t)\mathbf{z}_i^{\otimes 2}\exp(\boldsymbol{\beta}'(\mathbf{z}_i + \boldsymbol{\epsilon}_i)) + \frac{1}{n}\sum_{i=1}^{n} h_i(s,t)\boldsymbol{\epsilon}_i^{\otimes 2}\exp(\boldsymbol{\beta}'(\mathbf{z}_i + \boldsymbol{\epsilon}_i)) \\ &+ \frac{1}{n}\sum_{i=1}^{n} h_i(s,t)\mathbf{z}_i\boldsymbol{\epsilon}_i'\exp(\boldsymbol{\beta}'(\mathbf{z}_i + \boldsymbol{\epsilon}_i)) + \frac{1}{n}\sum_{i=1}^{n} h_i(s,t)\mathbf{z}_i'\boldsymbol{\epsilon}_i\exp(\boldsymbol{\beta}'(\mathbf{z}_i + \boldsymbol{\epsilon}_i)) \\ &= A_1 + A_2 + A_3 + A_4.\end{aligned}$$

A_1 can be written as

$$A_1 = \phi(\beta)\frac{1}{n}\sum_{i=1}^{n} h_i(s,t)\mathbf{z}_i^{\otimes 2}\exp(\boldsymbol{\beta}'\mathbf{z}_i) + \frac{1}{n}\sum_{i=1}^{n} h_i(s,t)\mathbf{z}_i^{\otimes 2}\exp(\boldsymbol{\beta}'\mathbf{z}_i)\{\exp(\boldsymbol{\beta}'\boldsymbol{\epsilon}_i)\phi^{-1}(\boldsymbol{\beta}) - 1\}.$$

So that

$$E\,(A_1) = \phi(\boldsymbol{\beta})s^{(2)}(s^\star, \mathbf{z}) + o_p(1).$$

Likewise,

$$\begin{aligned}A_2 &= \phi(\boldsymbol{\beta})\frac{1}{n}\sum_{i=1}^{n} h_i(s,t)\exp(\boldsymbol{\beta}'\mathbf{z}_i)[\boldsymbol{\Xi} + (\boldsymbol{\Xi}\boldsymbol{\beta})(\boldsymbol{\Xi}\boldsymbol{\beta})'] \\ &+ \phi(\boldsymbol{\beta})\frac{1}{n}\sum_{i=1}^{n} h_i(s,t)\exp(\boldsymbol{\beta}'\mathbf{z}_i)\left[\sum_{i=1}^{n}\boldsymbol{\epsilon}_i^{\otimes 2}\right] \\ &+ \frac{1}{n}\left[\left[\sum_{i=1}^{n}\boldsymbol{\epsilon}_i^{\otimes 2}\right]\boldsymbol{\beta}\right]\left[\sum_{i=1}^{n}\boldsymbol{\epsilon}_i^{\otimes 2}\right]'\left[e^{\boldsymbol{\beta}'\boldsymbol{\epsilon}_i}\phi(\boldsymbol{\beta}) - 1\right].\end{aligned}$$

Hence,

$$E(A_2) = \phi(\boldsymbol{\beta})S^{(0)}(s^\star, \mathbf{z}(t))[\boldsymbol{\Xi} + (\boldsymbol{\Xi}\boldsymbol{\beta})(\boldsymbol{\Xi}\boldsymbol{\beta})'] + o_p(1).$$

The terms A_3 and A_4 can also be corrected in the same way, leading to

$$E(A_3) = \phi(\boldsymbol{\beta})[S^{(1)}(s^\star, \mathbf{z}(t))(\boldsymbol{\Xi}\boldsymbol{\beta})' + (\boldsymbol{\Xi}\boldsymbol{\beta})s^{(1)}(s^\star, \mathbf{z}(t))'] + o_p(1).$$

Because of symmetry between A_3 and A_4, the expressions are the same; and the corrected expression of $S^{(2)}(s^\star, \mathbf{x}(t)|\boldsymbol{\beta})[S^{(0)}(s^\star, \mathbf{x}(t)|\boldsymbol{\beta})]^{-1}$ follows. $\|$ In light of the proposition, the corrected Fisher information matrix is given by

$$\mathcal{I}_n(\boldsymbol{\beta}; s^\star, t^\star) = -\frac{1}{n}\int_0^{t\star}\left[E(s^\star, \mathbf{z}(w)|\boldsymbol{\beta})^{\otimes 2} - \frac{S^{(2)}(s^\star, \mathbf{z}(w)|\boldsymbol{\beta})}{S^{(0)}(s^\star, \mathbf{z}(w)|\boldsymbol{\beta})}\right]\sum_{i=1}^{n} N_i(s^\star, dw).$$

It can be seen that the corrected Fisher information based on the true values of the covariates $\mathbf{z}$ no longer contains any information about the error terms. This concurs with the results in Kong and Gu (1999) and those in Augustin (2004). Moreover, the Fisher information matrix can therefore be estimated by $\frac{1}{n}\breve{\mathbf{I}}(\breve{\boldsymbol{\beta}}; \mathbf{x}(t), s^\star)$ which is a consistent estimator of $\frac{1}{n}\breve{\mathbf{I}}(\breve{\boldsymbol{\beta}}; \mathbf{z}(t), s^\star)$ by virtue of the consistency of $\breve{\boldsymbol{\beta}}$ and the corrected expression. The final result in this section regards the asymptotic property of the properly standardized corrected maximum partial likelihood estimators.

Theorem 5 *As $n \to \infty$, $\sqrt{n}(\breve{\boldsymbol{\beta}}_n - \boldsymbol{\beta}_0) \xrightarrow{d} N_p(\mathbf{0}, \boldsymbol{\Sigma}^{-1}(\boldsymbol{\beta}_0))$, where $\boldsymbol{\Sigma}^{-1}(\boldsymbol{\beta}_0; s^\star, t^\star)$ can be consistently estimated by*

$$\breve{\boldsymbol{\Sigma}}^{-1} = \left([\mathcal{I}_n(\breve{\boldsymbol{\beta}}; \mathbf{x}(t), s^\star)]^{-1}\right)'\breve{\Psi}(\breve{\boldsymbol{\beta}}; s, t)[\mathcal{I}_n(\breve{\boldsymbol{\beta}}; \mathbf{x}(t), s^\star)]^{-1}.$$

The proof follows from the usual Taylor expansion of the corrected score around $\boldsymbol{\beta}_0$, the limiting distribution of $\sqrt{n}^{-1}\breve{\mathbf{U}}_n(\boldsymbol{\beta}; s, t)$, and the consistency of the corrected Fisher information matrix. The limiting variance is finally obtained using multivariate distribution theory.$\|$

3.3. MEASUREMENT ERRORS VARIANCE ESTIMATION

Estimating the measurement error variance is useful in addressing whether a simplified approach that ignores it would be acceptable. The size of the error variance is also a factor associated with bias in parameter estimates. In order to address this estimation, it is common to proceed by bootstrapping. Two approaches are usually taken to achieve this goal; the school of thoughts in Hjort (1985) through bootstrapping the empirical distribution and those of Efron (1981), Efron and Tibshirani (1986) through the bootstrap resampling of the observables. If the empirical distribution is the basis for bootstrapping, then given $K_i = l_i$, one can obtain the empirical function of the $\mathbf{x}_i$ as

$$F_n(w|\mathbf{x}) = \frac{1}{\sum_{i=1}^{n} l_i} \sum_{i=1}^{n} \sum_{j=1}^{l_i} \mathrm{I}\{\mathbf{x}_i(j) \leq w\}$$

and obtain draws from the empirical. One could also deal directly with the observables, target the within-subject observations as replicates, and derive an estimated variance-covariance structure $\hat{\boldsymbol{\Xi}}$ by restricting to subjects with multiple events and obtain

$$\hat{\boldsymbol{\Xi}} = \frac{\sum_{i=1}^{n} I\{N_i^{\dagger}((s \wedge \tau_i)-) > 0\} \sum_{j=1}^{N_i^{\dagger}((s\wedge\tau_i)-)} (\mathbf{x}_{i,j}(s_{i,j}) - \bar{\mathbf{x}}_{i.})^{\otimes 2}}{\sum_{i=1}^{n} [N_i^{\dagger}((s \wedge \tau_i)-) - 1]},$$

$$\bar{\mathbf{x}}_{i.} = \frac{\sum_{j=1}^{N_i^{\dagger}((s\wedge\tau_i)-)} \mathbf{x}_{i,j}(s_{i,j})}{N_i^{\dagger}((s \wedge \tau_i)-)},$$

and resample this estimate for smoothing purposes. Regardless of the approach taken, since the empirical distribution function is a consistent estimator of the true distribution function, the individual components of $\hat{\boldsymbol{\Xi}}$ converge almost surely to their true values.

Theorem 6 *The components of $\hat{\sigma}^2_{qr}$ of $\hat{\boldsymbol{\Xi}}$ satisfy $\hat{\sigma}^2_{qr} \xrightarrow{a.s.} \sigma^2_{qr}$ under the bootstrap probability, that is, the coordinate-wise bootstrap estimator converges in probability to its counterpart in the true variance of the errors.*

Proof: This can be proved using the concept of statistical functionals. Define the Mallow metric $d_1(\cdot,\cdot)$ on L_1. Let X be an observable with probability measure P for which $E(X) < \infty$, likewise for Y with probability measure Q. Define the metric $d_1(\cdot,\cdot)$ between P and Q by $d_1(P,Q) = \inf E(|X - Y|)$. By Lemma 8.1 of Bickel and Freedman (1981), $d_1(\cdot,\cdot)$ is a metric. The empirical distribution function $F_n(w|\mathbf{x})$ converges in distribution to the true distribution function of the errors $F(t)$. By Lemma 8.3 of Bickel and Freedman (1981), $d_1(F_n(t), F(t)) \to 0$ as $n \to \infty$. Given $K_i = l_i$, define the functionals

$$\int \mathbf{x} d\bar{F}_n = \frac{1}{n}\sum_{i=1}^{n}\sum_{j=1}^{K_i}\mathbf{x}_i(j), \quad \text{and} \quad \int \mathbf{x}^2 d\bar{F}_n = \frac{1}{n}\sum_{i=1}^{n}\sum_{j=1}^{K_i}\mathbf{x}_i^2(j).$$

Assume F_n and F are in L_1 and L_2, then by Bickel and Freedman (1981), $d_1(F_n, F) \to 0$ implies $\int \|\mathbf{x}\| d\bar{F}_n \to E(\|\mathbf{x}\|)$ and $\int \|\mathbf{x}\|^2 d\bar{F}_n \to E(\|\mathbf{x}\|^2)$. Then $\hat{\sigma}^2_{qr} \xrightarrow{a.s.} \sigma^2_{qr}$ in bootstrap probability, and consequently $\hat{\boldsymbol{\Xi}} \xrightarrow{a.s.} \boldsymbol{\Xi}$.

4. SIMULATION AND APPLICATION

A simulation study was performed using the R Studio software package to investigate the performance of proposed corrected regression parameter estimators. The specific objectives of this study were: (i) to examine the effect of sample size (n) on the distributional properties of $\hat{\boldsymbol{\beta}}_C$; (ii) to examine the bias and variance of $\hat{\boldsymbol{\beta}}_C$.

Survival Times: We generate survival times t by

$$t = \theta_2 \left[-(\log U)\exp(-\boldsymbol{\beta}'\mathbf{x})\right]^{\frac{1}{\theta_1}},$$

where

t = survival time

θ_1 = shape parameter of the Weibull distribution

θ_2 = scale parameter of the Weibull distribution

$\boldsymbol{\beta}$ = p dimensional regression parameter vector

$\mathbf{x}$ = p dimensional covariates vector

U = randomly generated value from Uniform$(0, 1)$

Next, we show how we obtained this equation. First, note that

$$\lambda(u) = \frac{f(u)}{1 - F(u)} = -\frac{d}{du}\log[1 - F(u)],$$

and

$$\Lambda(t) = \int_0^t \lambda(u)du = -\log[1 - F(t)]. \tag{9}$$

For a Weibull distribution with shape parameter (θ_1) and scale parameter(θ_2), the cumulative distribution function is given by

$$F(t) = 1 - \exp\left[-\left(\frac{t}{\theta_2}\right)^{\theta_1}\right]. \tag{10}$$

Hence by (9) and (10), we get the expression of baseline cumulative hazard function for a Weibull distribution by

$$\Lambda_0(t) = \left(\frac{t}{\theta_2}\right)^{\theta_1}. \tag{11}$$

Moreover, for proportional hazard function,

$$\lambda(u) = \lambda_0(u)\exp(\boldsymbol{\beta}'\mathbf{x}). \tag{12}$$

Therefore,

$$-\log[1-F(t)] = \Lambda(t) = \int_0^t \lambda(u)du = \int_0^t [\lambda_0(u)\exp(\boldsymbol{\beta}'\mathbf{x})]du = \Lambda_0(t)\exp(\boldsymbol{\beta}'\mathbf{x}) \tag{13}$$

We also know that,

$$1 - F(t) \sim \text{uniform}(0, 1). \tag{14}$$

Hence by (11), (12) , (13) and (14) we obtain,

$$-\log U = \left(\frac{t}{\theta_2}\right)^{\theta_1} \exp(\boldsymbol{\beta}'\mathbf{x}).$$

By solving for t, we get

$$t = \theta_2\left[-(\log U)\exp(-\boldsymbol{\beta}'\mathbf{x})\right]^{\frac{1}{\theta_1}} \tag{15}$$

True Parameter Values: In our study, we set $\theta_1 = 1$, $\theta_2 = 1.5$, $\boldsymbol{\beta} = \{-1, 1\}$, $X_1 \sim N(0, 1)$, and $X_2 \sim Bernoulli(0.5)$.

Censoring: We generate censoring times from an exponential distribution randomly. $C_i \sim exp(\theta)$. In our study, we set $\theta = 0.75$. Next, we calculate times T_i by $T_i = min(t_i, C_i)$. We will also create an indicator variable $\delta_i = I(t_i \geq C_i)$ to indicate if T_i is a survival time or a censoring time.

Error Contaminated Variables: We add a gaussian noise with variance σ^2 to X_1 to create error contaminated version of X_1, say $\tilde{X}_1$.

By this point, we have the knowledge to generate an observation tuple,

$$O_i = \left\{X_{1i}, \tilde{X}_{1i}, X_{2i}, T_i, \delta_i\right\}.$$

Recurrent Event Data: To generate the recurrent events data, we perform the following steps. First, we create a database with 1 million observation tuples, say D. Next, we split this database into two different sub databases based on the value of δ. Let us call the sub-database with $\delta = 0$ as D_c and the sub-database with $\delta = 1$ as D_{nc}. Suppose we need to generate recurrent events data for $n \in \{40, 80, 100, 200\}$ subjects. To do that, we determine how many recurrent events are experienced by each subject j, say K_j, by randomly selecting a number from $\{0, 1, 2, 3, 4, 5, 6\}$. After that, we randomly select K_j number of observations from D_{nc} sub-database followed by one observation tuple from D_c to mimic the recurrent events observed by the jth subject. Once we have generated the recurrent events data for each subject of the study, we use the entire dataset to find naive regression parameter estimates by using `coxph` function in `surviavl` package in R. We create 100 data sets and find the mean of regression parameter estimates, which we call the naive estimates $\hat{\beta_1}$ and $\hat{\beta_2}$. We also find the standard deviation of the regression parameter estimates. Similarly, using our proposed equation, for each of the generated data set, we find the corrected regression parameter estimates. Finally, we obtain the mean of the corrected regression parameter estimates, which are denoted by $\hat{\beta}_{1c}$ and $\hat{\beta}_{2c}$. We also find and the standard deviation of these corrected estimates. We change the value of σ to take the values $\{0.01, 0.05, 0.1, 0.2, 0.3, 0.4, 0.5\}$, to observe the ability of our proposed estimator to handle the error. We refer to the model which ignores the measurement errors as a naive fit and to the one which incorporates the measurement error in estimation via corrected score as the corrected fit. In the second part of our simulation study, We kept $\beta_2 = 1$ and error variance $\sigma_2 = 0.1^2$ fixed and varied $\beta_1 \in (-1, -.8, -.6, -.4, -.2, -.1, .1, .2, .4, .6, .8, 1)$, to observe how naive and corrected estimators behave.

4.1. SIMULATION RESULTS

The simulation results are summarized in Figure 1 and Tables 1-4. The pictorial representation in Figure 1 encapsulates the use of estimation based on the corrected partial score rather than the naive approach in the parameter estimation for the predictor recorded with measurement error. While in the naive case, bias increases with the magnitude of the effect of the predictor on the hazard of recurrence, we see that the corrected partial likelihood has created an estimation scheme with minimal, if not nonexistent, bias. This observation holds irrespective of n. Overall, standard error trivially decreases with n. The parameter estimate for the predictor without measurement error has been estimated with minimal bias in both the partial and the corrected partial scores. We measure the accuracy of estimating the parameter without measurement error by using the mean square error. We call `MSE(naive)` for the estimation based on the partial likelihood, and `MSE(corrected)` the one based on the corrected partial likelihood. Expectedly, these values are negligible and asymptotically similar. Correction is expected to have a minimal and negligible effect on this covariate. The results align with published literature in single events where the corrected likelihood outperforms the naive approach.

4.2. APPLICATION: RHDNASE DATA

This is a pulmonary exacerbation study that has appeared in many publications, including Fuchs et al. (1994), Cook and Lawless (2007), and Yi (2017) to name a few. This double-blind, randomized clinical trial aims to assess the effect of rhDNase, a recombinant deoxyribonuclease I enzyme, versus placebo, on respiratory exacerbations among patients with cystic fibrosis. Six hundred and forty-five patients were recruited for the trial, and each patient was followed up for about 169 days. We use a modified version of these data as an illustrative example. During the study, an argument was made concerning forced expiratory volume (`FEV`) measurements being recorded with measurement errors because two measure-

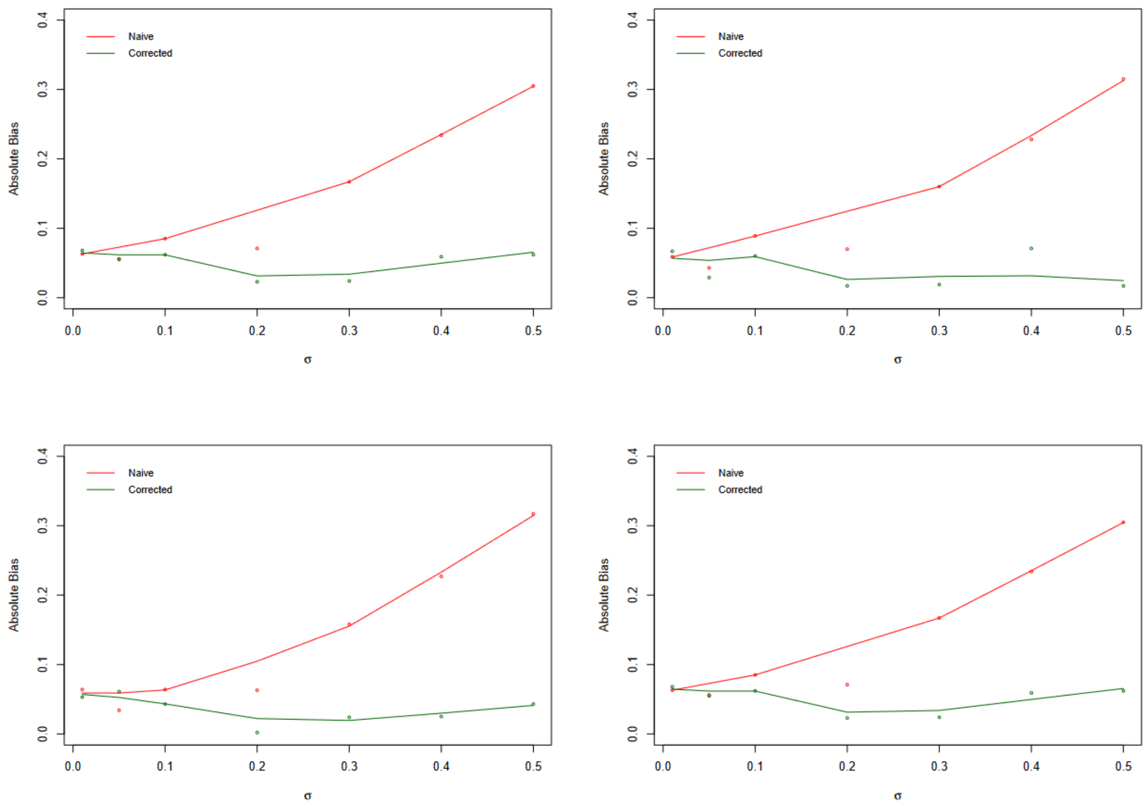

Figure 1. Bias assessment comparing the naive approach to the corrected.

ments taken a few minutes apart differed. For this illustrative example, we chose the seed value `12321` in `R` to create time-varying `FEV` values, for each subject, at each occurrence of respiratory exacerbation, using the subject-specific ordered measurements `FEV1, FEV2` according to a `Uniform(FEV1, FEV2)`. The dataset with time-varying `FEV` is posted on our web domain (`https://www.myweb.uiowa.edu/gzamba/`) for a reproducibility check on our proposed methods. Next, we add Gaussian noise to the time-varying `FEV` measurements, using the seed `121` in `R`, to create a covariate with measurement error. We use $\sigma = 0$ as a reference value, which would be the basis for our comparison; degeneracy under which both the corrected and uncorrected methods should yield similar results. We increase the error variance so that the resultant `FEV` value is still within the range of clinically acceptable and realistic pulmonary function values. By increasing the measurement error variance,

Table 1. Weibull Intensity: $\bar{F}(t;\boldsymbol{\theta}) = e^{-(\theta_2 t)^{\theta_1}}; \theta_2 = 1; \theta_1 = 1.5; \beta_2 = 1, n = 40, 80.$

					$\mathcal{N}(0, .1)$				
n	β_1	$\hat{\beta}_1$	$se(\hat{\beta}_1)$	$\hat{\beta}_{1c}$	$se(\hat{\beta}_{1c})$	$\hat{\beta}_2$	$se(\hat{\beta}_2)$	$\hat{\beta}_{2c}$	$se(\hat{\beta}_{2c})$
	-1.0	-0.839	0.095	-0.980	0.143	0.978	0.238	1.036	0.240
	-0.8	-0.678	0.109	-0.836	0.158	1.039	0.214	1.112	0.217
	-0.4	-0.365	0.117	-0.415	0.106	1.054	0.238	1.049	0.240
	-0.2	-0.177	0.117	-0.230	0.128	1.062	0.249	1.040	0.247
40	-0.1	-0.078	0.080	-0.081	0.120	1.002	0.248	1.040	0.230
	0.1	0.096	0.085	0.093	0.124	1.052	0.238	1.063	0.193
	0.2	0.170	0.079	0.183	0.125	1.019	0.260	1.028	0.207
	0.4	0.347	0.101	0.411	0.134	1.013	0.208	1.024	0.256
	0.8	0.709	0.138	0.773	0.150	1.067	0.260	1.079	0.244
	1.0	0.859	0.116	1.021	0.169	0.968	0.196	1.111	0.270
	-1.0	-0.840	0.079	-0.947	0.093	0.991	0.153	1.008	0.170
	-0.8	-0.687	0.079	-0.801	0.088	1.018	0.166	1.052	0.187
	-0.4	-0.342	0.064	-0.395	0.084	1.005	0.171	1.030	0.156
	-0.2	-0.187	0.062	-0.187	0.079	1.035	0.165	1.031	0.158
80	-0.1	-0.083	0.068	-0.105	0.059	1.019	0.164	0.976	0.154
	0.1	0.075	0.062	0.090	0.078	1.026	0.177	1.051	0.131
	0.2	0.195	0.062	0.187	0.056	1.035	0.128	1.024	0.160
	0.4	0.338	0.064	0.385	0.089	1.006	0.139	0.980	0.139
	0.8	0.665	0.075	0.784	0.093	0.972	0.156	1.046	0.146
	1.0	0.831	0.070	0.959	0.100	1.002	0.137	1.029	0.158

we verify in this illustrative example how the naive approach would progressively lead us to spurious inference about an FEV effect and how this inconvenience has been repaired through correction. We recognize that our method is not one-size-fits-all. The last row of table 4 highlights instants beyond which estimation crumbles down for the corrected partial likelihood—a situation partly addressed by Kong and Gu (1999) that failure in estimation for the corrected partial score tends to occur when $| \beta_1\sigma | \geq 0.8$; though in this example estimation appears to reasonably hold up to $| \beta_1\sigma | \leq 1$. The breakdown observed in the corrected score coefficient estimates is due to the divergence of the iteration process.

Table 2. Weibull Intensity: $\bar{F}(t;\boldsymbol{\theta}) = e^{-(\theta_2 t)^{\theta_1}}$; $\theta_2 = 1$; $\theta_1 = 1.5$; $\beta_2 = 1$, $n = 100, 200$.

n	β_1	$\hat{\beta}_1$	$se(\hat{\beta}_1)$	$\hat{\beta}_{1c}$	$\mathcal{N}(0,.1)$ $se(\hat{\beta}_{1c})$	$\hat{\beta}_2$	$se(\hat{\beta}_2)$	$\hat{\beta}_{2c}$	$se(\hat{\beta}_{2c})$
	-1.0	-0.826	0.069	-0.984	0.100	0.960	0.132	1.058	0.140
	-0.8	-0.674	0.060	-0.807	0.086	0.984	0.127	1.017	0.127
	-0.4	-0.354	0.063	-0.388	0.066	1.040	0.151	1.007	0.119
	-0.2	-0.177	0.060	-0.199	0.068	1.031	0.146	1.006	0.135
	-0.1	-0.086	0.064	-0.101	0.058	0.993	0.126	1.033	0.137
100	0.1	0.093	0.059	0.103	0.056	1.028	0.142	0.988	0.131
	0.2	0.183	0.063	0.211	0.072	1.057	0.140	1.029	0.115
	0.4	0.367	0.063	0.399	0.070	1.016	0.144	1.017	0.115
	0.8	0.685	0.070	0.784	0.083	1.011	0.141	1.026	0.166
	1.0	0.839	0.068	0.965	0.094	0.989	0.141	1.010	0.149
	-1.0	-0.821	0.039	-0.962	0.065	1.018	0.089	1.004	0.106
	-0.8	-0.668	0.053	-0.763	0.051	0.969	0.097	1.026	0.103
	-0.4	-0.338	0.039	-0.401	0.051	1.011	0.112	1.022	0.103
	-0.2	-0.182	0.046	-0.210	0.049	1.028	0.105	1.014	0.093
	-0.1	-0.092	0.045	-0.109	0.044	1.015	0.081	1.012	0.093
200	0.1	0.079	0.035	0.101	0.043	1.013	0.081	1.048	0.109
	0.2	0.177	0.042	0.208	0.047	1.016	0.099	1.022	0.089
	0.4	0.348	0.046	0.404	0.040	1.004	0.095	1.033	0.100
	0.8	0.678	0.043	0.758	0.050	0.976	0.107	1.001	0.099
	1.0	0.832	0.057	0.968	0.067	1.015	0.101	1.036	0.111

Table 3. MSE($\widehat{\beta}_2$): Error for the covariate with no measurement error.

n	40	80	100	200
`MSE(naive)`	0.0016	0.0005	0.0009	0.0004
`MSE(corrected)`	0.0041	0.0011	0.0007	0.0006

Table 4. Weibull Intensity: $\bar{F}(t;\boldsymbol{\theta}) = e^{-(\theta_2 t)^{\theta_1}}; \theta_2 = 1; \theta_1 = 1.5; (\beta_1, \beta_2) = (-1, 1)$.

					$\mathcal{N}(0,\sigma)$				
n	σ	$\hat{\beta}_1$	$se(\hat{\beta}_1)$	$\hat{\beta}_{1c}$	$se(\hat{\beta}_{1c})$	$\hat{\beta}_2$	$se(\hat{\beta}_2)$	$\hat{\beta}_{2c}$	$se(\hat{\beta}_{2c})$
	0.5	-0.704	0.085	-1.024	0.228	0.988	0.231	1.046	0.295
	0.4	-0.737	0.079	-0.980	0.193	1.012	0.224	1.129	0.265
	0.3	-0.848	0.094	-0.965	0.155	0.960	0.231	1.042	0.256
40	0.2	-0.909	0.107	-0.979	0.135	0.979	0.237	1.055	0.243
	0.1	-0.926	0.099	-0.911	0.116	1.029	0.227	1.036	0.193
	0.05	-0.947	0.122	-0.976	0.120	1.018	0.211	1.099	0.234
	0.01	-0.938	0.144	-0.928	0.134	0.956	0.177	1.044	0.227
	0.5	-0.685	0.071	-0.983	0.144	0.916	0.155	1.055	0.181
	0.4	-0.772	0.086	-0.929	0.086	0.966	0.152	1.049	0.162
	0.3	-0.840	0.077	-0.981	0.115	0.960	0.126	1.003	0.130
80	0.2	-0.930	0.091	-0.983	0.089	1.005	0.128	1.025	0.150
	0.1	-0.911	0.066	-0.940	0.082	0.989	0.135	1.056	0.138
	0.05	-0.957	0.081	-0.971	0.080	1.068	0.124	1.014	0.157
	0.01	-0.941	0.089	-0.933	0.071	1.005	0.132	1.001	0.161
	0.5	-0.683	0.068	-0.957	0.139	0.956	0.133	1.015	0.173
	0.4	-0.773	0.056	-0.975	0.108	1.026	0.143	1.118	0.173
	0.3	-0.842	0.075	-0.976	0.090	0.971	0.127	1.058	0.133
100	0.2	-0.937	0.074	-1.002	0.091	1.011	0.135	0.996	0.136
	0.1	-0.936	0.063	-0.957	0.078	0.997	0.137	1.042	0.124
	0.05	-0.966	0.066	-0.939	0.075	1.016	0.129	1.029	0.133
	0.01	-0.936	0.068	-0.947	0.071	1.011	0.154	1.024	0.120
	0.5	-0.695	0.045	-0.938	0.065	0.950	0.084	1.022	0.104
	0.4	-0.766	0.042	-0.941	0.059	1.000	0.081	1.095	0.110
	0.3	-0.833	0.052	-0.976	0.057	0.995	0.084	1.021	0.085
200	0.2	-0.929	0.051	-0.977	0.059	1.004	0.108	1.029	0.097
	0.1	-0.915	0.044	-0.938	0.047	0.978	0.076	1.004	0.094
	0.05	-0.944	0.055	-0.945	0.054	1.027	0.098	1.037	0.073
	0.01	-0.937	0.051	-0.932	0.049	1.013	0.109	1.010	0.087

Table 5. Regression Parameter Estimates.

σ	$\hat{\beta}_1$	$\hat{\beta}_2$	$\hat{\beta}_{1c}$	$\hat{\beta}_{2c}$
0.000	-1.7025	-0.2734	-1.7004	-0.2729
0.050	-1.6201	-0.2695	-1.6884	-0.2678
0.100	-1.4338	-0.2665	-1.6763	-0.2644
0.150	-1.2153	-0.2646	-1.6612	-0.2603
0.200	-1.0102	-0.2636	-1.6446	-0.2566
0.300	-0.6944	-0.2633	-1.6125	-0.2467
0.400	-0.4910	-0.2640	-1.5873	-0.2370
0.500	-0.3610	-0.2649	-1.5774	-0.2240
0.600	-0.2752	-0.2657	-1.6094	-0.2438
0.700	-0.2166	-0.2665	421.8588	86.7149

5. CORRECTED BASELINE HAZARD AND PROPERTIES

The baseline hazard plays a vital role in the Cox model. Whether time-varying or not, covariates act multiplicatively on the baseline hazard and give a real-time failure rate at time t. The Breslow cumulative baseline hazard was developed based on error-free covariates. So, any errors in the covariates measurement hinder the interpretation of the baseline hazard. It has been shown in the lifetime literature that the baseline hazard estimator based on the true covariates is consistent and converges to Gaussian processes when properly standardized. However, these large sample properties do not hold when the covariates are error-prone. As a consequence, a corrected baseline hazard needs to be developed. Some work has been done for the right censored data under the assumption of normally distributed errors, while large sample properties are derived, such as in Kong and Gu (1999). Below we propose a corrected baseline hazard for recurrent events with covariates under measurement errors. Our corrected estimator generalizes that in Augustin (2004) and Kong and Gu (1999) to recurrent events and is not restricted to normally distributed errors. The cumulative baseline hazard estimator based on error-free covariates under the Cox model was derived in Adekpedjou and Stocker (2015), Peña et al. (2007), and given by

$$\hat{\Lambda}_0(s^\star, \mathbf{x}(t)|\hat{\boldsymbol{\beta}}) = \int_0^t \frac{N(s^\star, dw)}{Y(s^\star, w|\hat{\boldsymbol{\beta}})},$$

where $\hat{\boldsymbol{\beta}}$ is the estimate based on error-free covariates. We seek a corrected estimator of $\Lambda_0(t)$, say $\breve{\Lambda}_0(s, t; \mathbf{z}|\breve{\boldsymbol{\beta}})$ that satisfies

$$E_{\boldsymbol{\epsilon}}(\hat{\Lambda}_0(s^\star, \mathbf{x}(t)|\hat{\boldsymbol{\beta}})) = \breve{\Lambda}_0(s^\star, \mathbf{z}(t)|\breve{\boldsymbol{\beta}}).$$

To that end, we assume expectation and integral formula are exchangeable, and we observe that from Taylor expansion of the ratio

$$\begin{aligned} E_{\boldsymbol{\epsilon}}\left(\frac{1}{S^{(0)}(s^\star, \mathbf{x}(t)|\boldsymbol{\beta})}\right) &= \frac{1}{E_{\boldsymbol{\epsilon}}(S^{(0)}(s^\star, \mathbf{x}(t)|\boldsymbol{\beta}))} \\ &= \frac{1}{\phi(\boldsymbol{\beta})S^{(0)}(s^\star, \mathbf{z}(t)|\boldsymbol{\beta})} + o_p(1). \end{aligned}$$

Then, it is not difficult to see that a corrected baseline cumulative hazard estimator for $\Lambda_0(t)$ based on the observables with covariates $\{\mathbf{x}_i(t) : i = 1, ..., n\}$ is given by

$$\breve{\Lambda}_0(s^\star, t|\breve{\boldsymbol{\beta}}) = \phi(\breve{\boldsymbol{\beta}}) \int_0^t \frac{N(s^\star, dw)}{S^{(0)}(s^\star, \mathbf{x}(t)|\breve{\boldsymbol{\beta}}))}, \tag{16}$$

where $\breve{\boldsymbol{\beta}}$ is the corrected estimate of $\boldsymbol{\beta}$. If the errors are assumed to be normally distributed, then we obtain a similarly corrected baseline hazard as those in Kong and Gu (1999). The corrected cumulative hazard in (16) is also similar in form to the one proposed by Augustin (2004) in the discrete failure time case. The consistency of $\breve{\Lambda}_0(s^\star, t|\breve{\boldsymbol{\beta}})$ is the consequence of the consistency of $\phi(\breve{\boldsymbol{\beta}})$ and $\hat{\Lambda}_0(s^\star, \mathbf{x}(t)|\hat{\boldsymbol{\beta}})$; from Theorem 2 of Adekpedjou and Stocker (2015).

Theorem 7 *As $n \to \infty$, the process $\{\sqrt{n}(\breve{\Lambda}_0(s^\star, t|\breve{\boldsymbol{\beta}}) - \Lambda_0(t))\}$ converges to a zero mean Gaussian process with variance function given by $\Gamma(s, t_1 \wedge t_2)$.*

Proof: Write

$$\begin{aligned}\sqrt{n}(\breve{\Lambda}_0(s^\star, t|\breve{\boldsymbol{\beta}}) - \Lambda_0(t)) &= \sqrt{n}(\breve{\Lambda}_0(s^\star, t|\breve{\boldsymbol{\beta}}) - \breve{\Lambda}_0(s^\star, t|\boldsymbol{\beta}_0)) \\ &\quad + \sqrt{n}(\breve{\Lambda}_0(s^\star, t|\boldsymbol{\beta}_0) - \Lambda_0(t)) \\ &= C_1 + C_2.\end{aligned}$$

Consider C_1. Taylor expanding C_1 yields

$$\begin{aligned} C_1 &= (\breve{\boldsymbol{\beta}} - \boldsymbol{\beta}_0)\left[\nabla_{\boldsymbol{\beta}}\phi(\boldsymbol{\beta}^\star)\int_0^t \frac{\sum_{i=1}^n N_i(s^\star, du)}{S^{(0)}(s^\star, u|\boldsymbol{\beta}^\star)}\right] \\ &\quad -(\breve{\boldsymbol{\beta}} - \boldsymbol{\beta}_0)\nabla_{\boldsymbol{\beta}}\left[\phi(\boldsymbol{\beta}^\star)\int_0^t \frac{E(\boldsymbol{\beta}^\star, \mathbf{x}(u))}{S^{(0)}(s^\star, \mathbf{x}(u)|\boldsymbol{\beta}^\star)}\sum_{i=1}^n N_i(s^\star, du)\right] \\ &= -(\breve{\boldsymbol{\beta}} - \boldsymbol{\beta}_0)H(\boldsymbol{\beta}^\star),\end{aligned}$$

where $\boldsymbol{\beta}^\star$ lies in the line segment between $\breve{\boldsymbol{\beta}}$ and $\boldsymbol{\beta}_0$ and $\boldsymbol{\beta}^\star \xrightarrow{p} \boldsymbol{\beta}_0$ as $n \to \infty$. Next, using the corrected expression of $S^{(0)}(s^\star, \mathbf{x}(t)|\boldsymbol{\beta})$ and $S^{(1)}(s^\star, \mathbf{x}(t)|\boldsymbol{\beta})$), it can be shown that, as $n \to \infty$, $H(\boldsymbol{\beta}^\star)$ converges to $-\int_0^t e(\boldsymbol{\beta}_0, u)\lambda_0(u)du$. Noting that $S^{(0)}(s^\star, \mathbf{x}(t)|\boldsymbol{\beta}_0) = \phi(\boldsymbol{\beta}_0)S^{(0)}(s^\star, \mathbf{z}(t)|\boldsymbol{\beta}_0)$, we have

$$\begin{aligned} C_2 &= \breve{\Lambda}_0(s^\star, t|\breve{\boldsymbol{\beta}}) - \Lambda_0(t) \\ &= \int_0^t \frac{dM(s, du)}{S^{(0)}(s^\star, \mathbf{z}(u))} + o_p(1).\end{aligned}$$

Hence, for large n

$$
\begin{aligned}
\sqrt{n}(\breve{\Lambda}_0(s^\star, t|\breve{\boldsymbol{\beta}}) - \Lambda_0(t)) &= \sqrt{n}H(\boldsymbol{\beta}^\star)(\breve{\boldsymbol{\beta}} - \boldsymbol{\beta}_0) \\
&\quad + \sqrt{n}\int_0^t \frac{\sum_{i=1}^n M_i(s, du)}{S^{(0)}(s^\star, \mathbf{z}(u))} \\
&= \frac{1}{\sqrt{n}}\int_0^t \frac{dM(s, du)}{S^{(0)}(s^\star, \mathbf{z}(u))} \\
&\quad + \frac{1}{\sqrt{n}} H(\boldsymbol{\beta}_0)\left[\frac{1}{n} I_n(\boldsymbol{\beta}_0; s^\star, t)\right]^{-1} \mathbf{U}_n(\boldsymbol{\beta}_0, s^\star, \mathbf{x}(t)) \\
&= \sum_{i=1}^n W_i(\boldsymbol{\beta}_0; s^\star, t).
\end{aligned}
$$

Therefore, $\sqrt{n}(\breve{\Lambda}_0(s^\star, t|\breve{\boldsymbol{\beta}}) - \Lambda_0(t))$ can be viewed as the sum of n independent components, each of which is a p-vector. This is the sum of n independent random process whose finite-dimensional distribution converges to those of $W^\infty(s, t)$. By the Functional Central Limit theorem, the sum converges weakly to a zero-mean Gaussian process with covariance matrix $\Gamma(s, t_1 \wedge t_2)$ that can be estimated consistently by $\breve{\Gamma}(s, t_1 \wedge t_2) = E(W_i(s, t_1; \breve{\boldsymbol{\beta}})W_i(s, t_2; \breve{\boldsymbol{\beta}}))$. ||

6. MISSPECIFIED ERRORS MODEL

A fundamental assumption underlying classical results on the properties of estimators obtained from the score process, namely the maximum likelihood estimators, is that the distribution or model determining the behavior under investigation is well specified. For example, thus far in this manuscript, we have assumed that the measurement errors have the additive form $\mathbf{x} = \mathbf{z} + \boldsymbol{\epsilon}$. However, the additive model for errors may be erroneous, and the real form may be another unknown model. Perhaps the Berkson model $\mathbf{z} = \mathbf{x} + \boldsymbol{\epsilon}$, or a mixture of the Berkson and additive, namely $\mathbf{x} = \mathbf{u} + \boldsymbol{\epsilon}$, where $\mathbf{u}$ is a latent variable having mean μ_u and a variance covariance matrix Σ_u. The true error model can also be a multiplicative or a transformed additive model. In many situations, knowing how the errors act on the covariates is hard. Suppose the measurement errors model is not correctly specified. In that case, it is natural to ask what would have resulted in properties such as consistency and the convergence in law of the properly standardized corrected partial likelihood estimator $\breve{\boldsymbol{\beta}}$. Does this estimator still converge asymptotically to some limit? If so, does this limit have any practical value? Does the asymptotic normality of the properly standardized $\breve{\boldsymbol{\beta}}$ still hold under the misspecified errors model? These questions can be answered using the theory of misspecified models. White and Domowitz (1984) provides a unified framework on the properties of maximum likelihood estimators obtained under various types of misspecification in different contexts.

6.1. PROPERTIES OF THE CORRECTED ESTIMATOR UNDER MISSPECIFIED ERRORS

To set the stage for this subsection, let $l_P(\boldsymbol{\beta}; s^\star, t)$ be the log-profile likelihood process under the working errors model, and $\boldsymbol{\beta}_n^\star$, the solution to $l_P(\boldsymbol{\beta}; s^\star, t) = \mathbf{0}$, if one exists. Suppose our true measurement errors model is defined by some unknown function $\boldsymbol{\kappa}(\mathbf{x}, \boldsymbol{\epsilon})$. Further, let $\tilde{l}_P(\boldsymbol{\beta}; s^\star, t)$ be the corrected likelihood under the true measurement errors $\boldsymbol{\kappa}(\mathbf{x}, \boldsymbol{\epsilon})$. If $\boldsymbol{\beta}$ is in a compact subset $\mathcal{B}$ of $\Re^p$, the solution $\tilde{\boldsymbol{\beta}}_n$ to $\tilde{l}(\boldsymbol{\beta}; s^\star, t) = \mathbf{0}$ is defined by

$$\tilde{\boldsymbol{\beta}}_n = \text{argmax}_{\boldsymbol{\beta} \in \mathcal{B}} \tilde{l}_P(\boldsymbol{\beta}; s^\star, t).$$

Likewise,

$$\boldsymbol{\beta}_n^\star = \text{argmax}_{\boldsymbol{\beta} \in \mathcal{B}} l_P(\boldsymbol{\beta}; s^\star, t)$$

is the sequence of the solution under the working likelihood. The first question is, does a sequence of solutions exist under the working model assumption? The answer is yes, given in the following lemma 2 of Jennrich (1969).

Lemma 3 *If $\mathcal{B}$ is a compact subset $\Re^p$. Given the data $\boldsymbol{O}$, there exists a sequence $\boldsymbol{\beta}_n^\star$ maximizing $l(\boldsymbol{\beta}; s^\star, t)$, that is*

$$\boldsymbol{\beta}_n^\star = \text{argmax}_{\boldsymbol{\beta}\in\mathcal{B}} l_P(\boldsymbol{\beta}; s^\star, t).$$

Following White and Domowitz (1984), to describe the discrepancy between the sequence of estimators $\{\tilde{\boldsymbol{\beta}}_n : n = 1, 2, 3...\}$ and $\{\breve{\boldsymbol{\beta}}_n : n = 1, 2, 3...\}$, we use the Kullback-Leibler (Kullback and Leibler (1951)) information criterion (KLI). The KLI gives an idea about how far apart the two solutions are and is given by

$$\text{KLI}(\boldsymbol{O}; \boldsymbol{\beta}) = E_{\boldsymbol{\kappa}(\mathbf{x},\boldsymbol{\epsilon})}\left[\frac{l_P(\boldsymbol{\beta}; s^\star, t)}{\tilde{l}_P(\boldsymbol{\beta}; s^\star, t)}\right].$$

The following theorem is on the large sample behavior of the sequence $\breve{\boldsymbol{\beta}}_n$, the solution obtained using the working likelihood.

Theorem 8 *As $n \longrightarrow \infty$, we have:*
(a) $\boldsymbol{\beta}_n^\star \xrightarrow{p} \boldsymbol{\beta}^\star \neq \boldsymbol{\beta}_0$
(b) $\sqrt{n}(\boldsymbol{\beta}_n^\star - \boldsymbol{\beta}^\star) \xrightarrow{d} N_p(\mathbf{0}, \boldsymbol{\Sigma}^{-1}(\boldsymbol{\beta}^\star))$, *where* $\boldsymbol{\Sigma}^{-1}(\boldsymbol{\beta}^\star))$ *is given by*

$$\boldsymbol{\Sigma}^{-1}(\boldsymbol{\beta}^\star) = \left([\mathcal{I}(\boldsymbol{\beta}^\star; s^\star, t)]^{-1}\right)'\Psi(\boldsymbol{\beta}^\star; s^\star, t)[\mathcal{I}(\boldsymbol{\beta}^\star; s^\star, t)]^{-1},$$

and can be consistently estimated with the observables data and expectation for the Fisher information taken under $\boldsymbol{\kappa}(\boldsymbol{x}, \boldsymbol{\epsilon})$ *the true error model,* $\Psi(\cdot; s^\star, t)$ *is the same as in 4.*

Remark 1 *The previous theorem gives us two important results when the model for errors is misspecified. First, the root of the corrected score under the misspecified model is still consistent- albeit it does not converge to the true value of* $\boldsymbol{\beta}$*, and we still have convergence to a multivariate normal distribution when the misspecified estimator is standardized. The other significant result is the fact that we can quantify the magnitude of the inconsistency,* *if any, by looking at $\boldsymbol{\beta}_n^\star - \boldsymbol{\beta}_0$. To obtain $\boldsymbol{\beta}_n^\star$, one needs to numerically solve the equation*

$$\begin{aligned}\tilde{\boldsymbol{U}}(\boldsymbol{\beta}; s^\star, t) &= \sum_{i=1}^{n} \int_0^t [\mathbf{x}_i(\varphi_i^{-1}(w)) - E_{\boldsymbol{\kappa}(\mathbf{x},\boldsymbol{\epsilon})}(s^\star, \mathbf{x}(w))] N_i(s^\star, dw) \\ &= \sum_{i=1}^{n} \int_0^t \left[\mathbf{x}_i(\varphi_i^{-1}(w)) - \frac{E_{\boldsymbol{\kappa}(\mathbf{x},\boldsymbol{\epsilon})}(S^{(1)}(s^\star, \mathbf{x}(w)|\boldsymbol{O}))}{E_{\boldsymbol{\kappa}(\mathbf{x},\boldsymbol{\epsilon})}(S^{(0)}(s^\star, \mathbf{x}(w)|\boldsymbol{O}))}\right] N_i(s^\star, dw) \\ &= \mathbf{0}.\end{aligned} \tag{17}$$

Proof: (a) We apply Theorem 1.14 page 28 of Yi and Lawless (2012). We have two tasks: (i) show that $l(\boldsymbol{\beta}; s^\star, t) = \ln[L(\boldsymbol{\beta}; s^\star, t)]$ is bounded by an integrable function concerning the distribution of the true errors and (ii) KLI$(\boldsymbol{O}; \boldsymbol{\beta})$ has a unique minimum at $\boldsymbol{\beta}_n^\star$ in $\mathcal{B}$, a compact subset of $\Re^p$. We begin with (i). Observe that the corrected score under the true errors model is

$$\tilde{\mathbf{U}}(\boldsymbol{\beta}; s^\star, t) := \sum_{i=1}^{n} \int_0^t \left[\mathbf{x}_i(\varphi_i^{-1}(w)) - \frac{E_{\boldsymbol{\kappa}(\mathbf{x},\boldsymbol{\epsilon})}(S^{(1)}(s^\star, w; \mathbf{x}|\boldsymbol{O}))}{E_{\boldsymbol{\kappa}(\mathbf{x},\boldsymbol{\epsilon})}(S^{(0)}(s^\star, w; \mathbf{x}|\boldsymbol{O}))}\right] N_i(s^\star, dw).$$

The concept of manageability can be used to show that an integrable function bounds it. We assume without loss of generality that $\mathbf{x}_{ik}$ is positive. Because x_{ik} has a total variation bounded by a constant, and $N_i(s, t)$ is a sum of indicator functions that are increasing functions for each i, we conclude that each of the p-components of $\tilde{l}_P(\boldsymbol{\beta}; s^\star, t)$ has pseudo-dimension not exceeding 10 (cf. Pollard (1990)) since the integrand have pseudo-dimension of at most 1. Therefore, $\tilde{l}_P(\boldsymbol{\beta}; s^\star, t)$ is manageable, hence bounded by an integrable function with respect to the true probability measure of the errors, which also guarantees that KLI$(\boldsymbol{O}; \boldsymbol{\beta})$ is well defined. To prove part (ii), let $l(\boldsymbol{\beta}; s^\star, t)$ be the true likelihood under the true errors model, and $\tilde{l}(\boldsymbol{\beta}; s^\star, t)$ be the working likelihood, that is one under the misspecified model. Define the random function $\mathcal{D}(\boldsymbol{\beta}; s^\star, t)$ in $\mathcal{B}$ by

$$\mathcal{D}(\boldsymbol{\beta}; s^\star, t) = \frac{1}{n} E_{\boldsymbol{\kappa}(\mathbf{x},\boldsymbol{\epsilon})}[l(\boldsymbol{\beta}_0; s^\star, t) - \tilde{l}(\boldsymbol{\beta}; s^\star, t)] \tag{18}$$

$$= \frac{1}{n} E_{\boldsymbol{\kappa}(\mathbf{x},\boldsymbol{\epsilon})}\left[\log\left(\frac{L(\boldsymbol{\beta}_0; s^\star, t)}{\tilde{L}(\boldsymbol{\beta}; s^\star, t)}\right)\right]. \tag{19}$$

Also, define the deterministic function

$$d(\boldsymbol{\beta}; s^\star, t) = \int_0^{t^\star} \left[(\tilde{\boldsymbol{\beta}}_0 - \boldsymbol{\beta})' \mathbf{x}(\varphi^{-1}(w)) - \ln \left[\frac{s^{(0)}(s^\star, w|\boldsymbol{\beta})}{s^{(0)}(s^\star, w|\boldsymbol{\beta}_0)} \right] \right] \lambda_0(w) dw. \tag{20}$$

These functions are twice-differentiable concerning $\boldsymbol{\beta} \in \mathcal{B}$. Observe that $\boldsymbol{\beta}_n^\star$ is the natural estimator of the working likelihood. Furthermore, $\tilde{\boldsymbol{\beta}}$ is the parameter value that minimizes KLI$(\boldsymbol{O}; \boldsymbol{\beta})$ since the numerator in KLI is a function of the true likelihood, and the denominator is the working likelihood with expectation taking with respect to the true measurement errors model distribution. Assume that, for $m = 0, 1, 2$

$$\sup_{(\boldsymbol{\beta}, t) \in \mathcal{B} \times [0, \tau]} \| S_i^{(m)}(\boldsymbol{\beta}; s, t) - E_{\boldsymbol{\kappa}(\mathbf{x}, \boldsymbol{\epsilon})} \left(S_i^{(m)}(\boldsymbol{\beta}; s, t) \right) \| \xrightarrow{p} 0,$$

with $E_{\boldsymbol{\kappa}(\mathbf{x}, \boldsymbol{\epsilon})} \left(S_i^{(m)}(\boldsymbol{\beta}; s, t) \right) = s^{\star(m)}(\boldsymbol{\beta}; s, t)$. Furthermore, observe that $\tilde{\boldsymbol{\beta}}_n$ is minimizer of $\boldsymbol{\beta} \mapsto D(s^\star, \boldsymbol{\beta})$. We also observe that

$$\sup_{(\boldsymbol{\beta}, t) \in \mathcal{B} \times [0, \tau]} \| \mathcal{D}(\boldsymbol{\beta}; s^\star, t) - d(\boldsymbol{\beta}; s^\star, t) \| \xrightarrow{p} 0,$$

as $n \to \infty$. Easy calculations show that the first derivative of $\mathcal{D}(\boldsymbol{\beta}; s^\star, t)$ is the score process, whose derivative is the Fisher information matrix with limit coinciding with the Hessian of $d(\boldsymbol{\beta}; s^\star, t)$. The Hessian of $d(\boldsymbol{\beta}; s^\star, t)$ turns out to be positive definite, which means that $\boldsymbol{\beta} \mapsto D(\boldsymbol{\beta}; s^\star, t)$ is strictly convex for $\boldsymbol{\beta} \in \mathcal{B}$. We take $\boldsymbol{\beta}_n^\star$ to be the minimizer of $D(s^\star, \boldsymbol{\beta})$ inside $\mathcal{B}$, and $\boldsymbol{\beta}^\star$ to be the unique minimizer of $d(\boldsymbol{\beta}; s^\star, t)$ in $\mathcal{B}$. Therefore, by Theorem 2.2 of White and Domowitz (1984), we obtain the convergence in probability of the sequence $\{\boldsymbol{\beta}_n^\star : n = 1, 2, 3, ...\}$ to a value $\boldsymbol{\beta}^\star$ minimizing the KLI$(\boldsymbol{O}; \boldsymbol{\beta})$. The proof of Part (b)follows the convergence to the multivariate normal distribution of the properly standardized corrected maximum likelihood when the error terms are correctly specified.

7. CONCLUSION AND DISCUSSION

Through our exposition, we have shown, in the context of recurrent events, that correcting the profile likelihood and the estimating equations is a wise course of action in the presence of suspected measurement errors in the Cox model covariates. We have further provided a mechanism for correcting the baseline integrated hazard in the presence of measurement errors–as it is a crucial ingredient for inference in reliability and recurrence. We operated under the additive measurement errors model and have shown that the corrected estimators proposed are consistent and, when properly standardized, converge to Gaussian processes. We further utilized the Kullback-Leibler information criterion and misspecification theorems from Domowitz and White (1982) to show that when the errors model is misspecified by taking a form that differs from the additive, bias can be easily quantified and that the new estimators obtained under this scenario converge when properly standardized, although not to the true parameter values. The latter indicates the size of the bias. Finally, we provided simulation results supporting bias and estimation and successfully applied our findings to open-source rhDNase data. We hope that practitioners will be convinced not to settle down for modeling without prior thoughts given to measurement errors but instead choose the course of wisdom by pertinently questioning the mechanism that generates the data at hand and scrutinizing whether bias has crept into the process. Furthermore, we provided a theoretical means for estimating the magnitude of the errors. We plan to disseminate the current work by software that will be published on the comprehensive R archive network and adaptable to practitioners with applied inclinations.

REFERENCES

Adekpedjou, A. and Stocker, R. (2015). A general class of semiparametric models for recurrent event data. *J. Statist. Plann. Inference*, 156:48–63.

Alexeeff, S. E., Carroll, R. J., and Coull, B. (2016). Spatial measurement error and correction by spatial SIMEX in linear regression models when using predicted air pollution exposures. *Biostatistics*, 17(2):377–389.

Armstrong, B. G. (1990). The effects of measurement errors on relatwe risk regressions. *American Journal of Epidemiology*, 132 (6):1176–1184.

Augustin, T. (2004). An exact corrected log-likelihood function for Cox's proportional hazards model under measurement error and some extensions. *Scand. J. Statist.*, 31(1):43–50.

Beutner, E., Bordes, L., and Doyen, L. (2020). Consistent semiparametric estimators for recurrent event times models with application to virtual age models. *Bernoulli*, 26(1):557–586.

Bickel, P. J. and Freedman, D. A. (1981). Some asymptotic theory for the bootstrap. *Ann. Statist.*, 9(6):1196–1217.

Carroll, R. J., Ruppert, D., and Stefanski, L. A. (1995). *Measurement error in nonlinear models*, volume 63 of *Monographs on Statistics and Applied Probability*. Chapman & Hall, London.

Chen, L.-P. and Yi, G. Y. (2021). Semiparametric methods for left-truncated and right-censored survival data with covariate measurement error. *Ann. Inst. Statist. Math.*, 73(3):481–517.

Cox, D. R. (1972). Regression models and life-tables. *Journal of the Royal Statistical Society: Series B (Methodological)*, 34(2):187–202.

Domowitz, I. and White, H. (1982). Misspecified models with dependent observations. *Journal of Econometrics*, 20(1):35–58.

Efron, B. (1981). Censored data and the bootstrap. *J. Amer. Statist. Assoc.*, 76(374):312–319.

Efron, B. and Tibshirani, R. (1986). Bootstrap methods for standard errors, confidence intervals, and other measures of statistical accuracy (with discussion). *Statistical Science*, 1(1):54–77.

Fuller, W. A. (1987). *Measurement error models*. Wiley Series in Probability and Mathematical Statistics: Probability and Mathematical Statistics. John Wiley & Sons, Inc., New York.

Hjort, N. L. (1985). Bootstrapping cox's regression model. *Technical Report 241, Department of Statistics, Stanford University*.

Hu, C. and Lin, D. Y. (2002). Cox regression with covariate measurement error. *Scand. J. Statist.*, 29(4):637–655.

Hu, C. and Lin, D. Y. (2004). Semiparametric failure time regression with replicates of mismeasured covariates. *J. Amer. Statist. Assoc.*, 99(465):105–118.

Hu, P., Tsiatis, A. A., and Davidian, M. (1998). Estimating the parameters in the Cox model when covariate variables are measured with error. *Biometrics*, 54(4):1407–1419.

Huque, M. H., Bondell, H. D., Carroll, R. J., and Ryan, L. M. (2016). Spatial regression with covariate measurement error: a semiparametric approach. *Biometrics*, 72(3):678–686.

Jennrich, R. I. (1969). Asymptotic properties of non-linear least squares estimators. *Ann. Math. Statist.*, 40:633–643.

Kong, F. H. and Gu, M. (1999). Consistent estimation in Cox proportional hazards model with covariate measurement errors. *Statist. Sinica*, 9(4):953–969.

Kosorok, M. R. (2008). *Introduction to empirical processes and semiparametric inference*. Springer Series in Statistics. Springer, New York.

Kullback, S. and Leibler, R. A. (1951). On information and sufficiency. *Ann. Math. Statistics*, 22:79–86.

Nakamura, T. (1990). Corrected score function for errors-in-variables models: methodology and application to generalized linear models. *Biometrika*, 77(1):127–137.

Nakamura, T. (1992). Proportional hazards model with covariates subject to measurement error. *Biometrics*, 48(3):829–838.

Peña, E. A., Slate, E. H., and Gonzalez, J. R. (2007). Semiparametric inference for a general class of models for recurrent events. *Journal of Statistical Planning and Inference*, 137:1727–1747.

Pena, E. A., Strawderman, R. L., and Hollander, M. (2001). Nonparametric estimation with recurrent event data. *Journal of the American Statistical Association*, 96(456):1299–1315.

Pollard, D. (1990). *Empirical processes: theory and applications*, volume 2 of *NSF-CBMS Regional Conference Series in Probability and Statistics*. Institute of Mathematical Statistics, Hayward, CA; American Statistical Association, Alexandria, VA.

Prentice, R. L. (1982). Covariate measurement errors and parameter estimation in a failure time regression model. *Biometrika*, 69(2):331–342.

Song, X. and Huang, Y. (2005). On corrected score approach for proportional hazards model with covariate measurement error. *Biometrics*, 61(3):702–714.

Stefanski, L. A. (1985). The effects of measurement error on parameter estimation. *Biometrika*, 72(3):583–592.

Stefanski, L. A. and Cook, J. R. (1995). Simulation-extrapolation: the measurement error jackknife. *J. Amer. Statist. Assoc.*, 90(432):1247–1256.

Turnbull, B., J. W. . C. L. (1997). Regression models for recurrent event data: Parametric random effects models with measurement error. *Stat in Med.*, 16:853–864.

van der Vaart, A. W. (1998). *Asymptotic statistics*, volume 3 of *Cambridge Series in Statistical and Probabilistic Mathematics*. Cambridge University Press, Cambridge.

Wang, C. Y., Hsu, L., Feng, Z. D., and Prentice, R. L. (1997). Regression calibration in failure time regression. *Biometrics*, 53(1):131–145.

White, H. and Domowitz, I. (1984). Nonlinear regression with dependent observations. *Econometrica*, 52(1):pp. 143–162.

Yan, Y. and Yi, G. Y. (2015). A corrected profile likelihood method for survival data with covariate measurement error under the Cox model. *Canad. J. Statist.*, 43(3):454–480.

Yi, G. Y. (2017). *Statistical analysis with measurement error or misclassification*. Springer Series in Statistics. Springer, New York. Strategy, method and application, With a foreword by Raymond J. Carroll.

Yi, G. Y. and Lawless, J. F. (2012). Likelihood-based and marginal inference methods for recurrent event data with covariate measurement error. *Canad. J. Statist.*, 40(3):530–549.

Yu, H., Cheng, Y.-J., and Wang, C.-Y. (2016). Semiparametric regression estimation for recurrent event data with errors in covariates under informative censoring. *Int. J. Biostat.*, 12(2):20160001, 18.

Yu, H., Cheng, Y.-J., and Wang, C.-Y. (2018). Methods for multivariate recurrent event data with measurement error and informative censoring. *Biometrics*, 74(3):966–976.

Yu, M. and Nan, B. (2010). Regression calibration in semiparametric accelerated failure time models. *Biometrics*, 66(2):405–414.

Zamba, K. D. and Adekpedjou, A. (2019). A khmaladze-transformed test of fit with ml estimation in the presence of recurrent events. *Sequential Analysis*, 38(3):318–341.

II. ADDITIVE SEMIPARAMETRIC MODEL FOR RECURRENT EVENT DATA SUBJECT TO COVARIATES MEASUREMENT ERROR

Ravinath Alahakoon
Department of Mathematics & Statistics
Missouri University of Science and Technology
Rolla, Missouri 65409–0050
Email: ra8k7@mst.edu

Akim Adekpedjou
Department of Mathematics & Statistics
Missouri University of Science and Technology
Rolla, Missouri 65409–0050
Email: akima@mst.edu

ABSTRACT

Consider that units are monitored for multiple occurrences of an event. At each occurrence, covariates contributing to those event times are recorded, with at least one of the covariates subject to measurement error. It is well known that error-contaminated covariates can distort inference and create biased and inconsistent estimators if not properly accounted for. In this manuscript, we propose a corrected score function under the assumption of the classical additive errors model while letting the effect of covariates be additive. Simulation studies show that the proposed estimators approximate the true values of regression parameters well. Finally, the proposed methods are applied to the rhDNase dataset.

Keywords: Recurrent events; Covariates measurement error; Corrected score; Additive model; Counting processes

1. INTRODUCTION

In survival analysis, the goal is to model the time until an event of interest occurs, such as the failure of a mechanical component or the death of a patient. The time-to-event data is often right-censored, meaning that some subjects do not experience the event by the end of the study or are lost to follow-up before the event occurs. The Cox proportional hazard model is a widely used method in survival analysis, which assumes that the hazard function at any time t is proportional to a set of covariates or predictors as below:

$$\lambda(t|\mathbf{x}) = \lambda_0(t)\exp(\boldsymbol{\beta}^T\mathbf{x}),$$

where $\lambda_0(t)$ is the baseline hazard function, which is independent of the covariates, and $\boldsymbol{\beta}$ are the regression coefficients of the predictors $\mathbf{x}$. To estimate the coefficients $\boldsymbol{\beta}$, we use the partial likelihood method that allows us to handle right-censored data and obtain consistent estimates of the coefficients. In part 1 of this dissertation, we used the Cox proportional hazards model to analyze the relationship between a set of predictors and the recurrent event time data which is in fact a type of time-to-failure data where each subject may experience more than one event during the study period.

Nevertheless, in this second part of the dissertation, we will use the additive hazards model below:

$$\lambda(t|\mathbf{x}) = \lambda_0(t) + \boldsymbol{\beta}^T\mathbf{x}.$$

The choice of the appropriate model depends on the specific research question and the underlying assumptions of the data. While the multiplicative intensity model is widely used in survival analysis, it may not always be appropriate. In some cases, the additive hazard model may be a more suitable alternative. One such situation is when the hazards are non-proportional, meaning that the effect of covariates on the hazard function may change over time. In this case, an additive model may be more flexible and better able to capture the changing relationship between the covariates and the outcome variable. For example, in cancer studies, the hazard of recurrence may be higher in the first few years after treatment and then decrease over time. In such cases, the additive hazard model may be more appropriate. Another situation is when there are significant interaction effects between covariates, which may not be captured by the multiplicative hazard model. In contrast, the additive hazard model can more easily account for such effects. For instance, in a study of heart disease, the effect of obesity on the hazard of a heart attack may be different in men and women. The additive hazard model can better capture such interaction effects. Non-linear effects may also pose a challenge for the multiplicative hazard model. For instance, in a study of smoking and lung cancer, the effect of smoking on the hazard of lung cancer may be non-linear. In such cases, the additive hazard model may be more appropriate. Sometimes, when the effect of a covariate changes over time, the multiplicative hazard model may not be appropriate. For instance, in a study of HIV/AIDS, the effect of CD4 count on the hazard of death may change as the disease progresses. In such cases, the additive hazard model may be more appropriate.

Next we give a brief introduction to recurrent events and measurement error. Recurrent event processes are processes that generate events repeatedly over time, and they arise in a variety of fields such as biomedical science, epidemiology, social science, reliability, and actuarial science. To better understand these processes, analysts often need to estimate unknown functions like the intensity function, the survivor function, or the mean rate function, taking into account possible time-dependent covariates. In many research fields, it is unrealistic to assume that all covariates are perfectly measured, and ignoring measurement errors can lead to biased estimates and inaccurate conclusions. Therefore, numerous correction methods have been developed to account for measurement errors in models, including censored, truncated, and uncensored data. These methods can be parametric or nonparametric, using various techniques such as replicate surrogates, instrumental variables, or moment-based approaches. While there is abundant literature on measurement error models for single events, there is a need for more research on these models in the context of recurrent events. A few recent studies have addressed this issue, including a moment-based method and a partial score function correction for symmetric errors. Additionally, some authors have developed methods for modeling recurrent events that account for measurement errors under a broad class of hazard models and informative censoring.

We now discuss some of the important references. Turnbull (1997) considered a mixed effects Poission regression model for recurrent event data with error-contaminated covariates. They proposed adjustments for usual maximum likelihood estimators that are obtained from neglecting covariate measurement error as the method of accounting for error-prone covariates. Jiang et al. (1999) investigated inference methods for discrete-time events in the presence of covariate measurement error. In particular, they used semi-parametric Poisson and mixed Poission process regression while accounting for possible random effects and covariate measurement error. Yi and Lawless (2012) developed inferential methods that account for covariate measurement error. Particularly, their work included counting processes consisting of multiplicative intensity functions and mixed Poisson models. They discussed inference methods based on likelihood, producing estimation equations. Yu et al. (2018) proposed non-parametric methods for correcting covariate measurement error in multivariate recurrent event data under informative censoring. However, their research was limited to time-independent covariates. Moreover, their approach did not require the Poisson-type assumption for recurrent event process and any distributional assumption for frailty or covariate measurement error.

In addition, there is some work in the literature related to our area of concentration in this manuscript, mainly measurement error. Veierød and Laake (2001) and Guo and Li (2002) explored covariate measurement error effects on Poisson regression and misclassification. Zeger and Edelstein (1989) studied the Poisson regression model with error-contaminated covariates and used a likelihood method to correct the measurement error effects. Fung and Krewski (1999) investigated SIMEX and regression calibration algorithms empirically for Poisson regression with replicates of error-prone covariate measurements. Kim (2007) produced a mean model for the event count data and used kernel estimates to obtain a correction method in the presence of categorical error-prone covariates while assuming a validation subsample is available. These studies did not investigate the asymptotic properties of the derived estimators. However, they provided simulation study results to assess their proposed methods' performance.

In this dissertation, we assume that the variance-covariance matrix of error variables is time-independent while no distributional properties are imposed on the errors. We propose a corrected score process considering additive intensity function. We consider the case where the errors are modeled using the classical additive measurement errors model, which may be the case in many real-world scenarios.

This part of the dissertation proceeds as follows: In Section 2, the model and the notation are introduced. In Section 3, the measurement error model and the assumptions are provided. Section 4 consists of a mathematical setup for our proposed corrected score. In Section 5, consistency of the proposed corrected regression parameter vector is established. Section 6 reports the simulation study to investigate finite sample properties. In Section 7, the rhNDase dataset is analyzed to illustrate our model. Section 9 lists concluding remarks of this research.

2. THE MODEL AND NOTATION

The notation employed in this section is identical to that used in part 1. However, there is a key distinction between the two: while a multiplicative hazard model was utilized in the first section of the dissertation, an additive hazard model will be employed in this one. Specifically:

$$\lambda_i(v) = \lambda_0(\varphi_i(v)) + \boldsymbol{\beta}^T \mathbf{x}_i(v).$$

3. MEASUREMENT ERROR

As in the first part of this dissertation, we consider the additive error model. To proceed in this section, we will follow the same approach by correcting the score function under the additive error model.

4. CORRECTED SCORE

Following Stocker and Adekpedjou (2020) if $\mathbf{x}_i(s)$ were the true covariates, then the score process is given by

$$U(\boldsymbol{\beta};\mathbf{x},s,t)=\sum_{i=1}^{n}\int_0^t[\mathbf{x}_i(\varphi_i^{-1}(w))-\bar{\mathbf{x}}(\varphi^{-1}(w))][N_i(s,dw)-Y_i(s,w)\boldsymbol{\beta}^T\mathbf{x}_i(\varphi_i^{-1}(w))dw],$$

where

$$\bar{\mathbf{x}}(\varphi^{-1}(t))=\frac{\sum_{j=1}^{n}Y_j(s,t)\mathbf{x}_j(\varphi_j^{-1}(t))}{\sum_{j=1}^{n}Y_j(s,t)}.$$

Unfortunately, the $\mathbf{x}_i(s)$ are not the true covariates, and as a consequence, the solution $\hat{\boldsymbol{\beta}}$ is biased and cannot be used. If ignored, error-contaminated covariates can lead to biased parameter estimates and inaccurate conclusions. Hence, regardless of the model used, the estimation functions need to be corrected to obtain accurate estimators. Intending to find a remedy for this issue, next we will derive a corrected score to estimate regression parameters. As before, we seek,

$$E_\epsilon[\mathbf{U}^*(\boldsymbol{\beta};\mathbf{x},s,t)]=\mathbf{U}(\boldsymbol{\beta};\mathbf{z},s,t). \tag{1}$$

The following propositions are needed in the sequel. All expectations are taken with respect to the error distribution denoted by $E_\epsilon(\cdot)$.

Proposition 5 $E_\epsilon\left\{x_i(\varphi_i^{-1}(w))\right\}=z_i(\varphi_i^{-1}(w))$.

Proof:

$$\begin{aligned} E_\epsilon\left\{\mathbf{x}_i(\varphi_i^{-1}(w))\right\} &= E_\epsilon\left\{\mathbf{z}_i(\varphi_i^{-1}(w)) + \boldsymbol{\epsilon}_i(\varphi_i^{-1}(w))\right\} \\ &= \mathbf{z}_i(\varphi_i^{-1}(w)). \ \| \end{aligned} \tag{2}$$

Proposition 6 $E_\epsilon\left\{\bar{\boldsymbol{x}}(\varphi^{-1}(w))\right\} = \bar{z}(\varphi^{-1}(w))$.

Proof:

$$\begin{aligned} E_\epsilon\left\{\bar{\mathbf{x}}(\varphi^{-1}(w))\right\} &= E_\epsilon\left\{\frac{\sum_{j=1}^n Y_j(s,w)\mathbf{x}_j(\varphi_j^{-1}(w))}{\sum_{j=1}^n Y_j(s,w)}\right\} \\ &= \frac{\sum_{j=1}^n Y_j(s,w)E_\epsilon[\mathbf{x}_j(\varphi_j^{-1}(w))]}{\sum_{j=1}^n Y_j(s,t)} \\ &= \frac{\sum_{j=1}^n Y_j(s,w)E_\epsilon[\mathbf{z}_j(\varphi_j^{-1}(w)) + \boldsymbol{\epsilon}_j(\varphi_j^{-1}(w))]}{\sum_{j=1}^n Y_j(s,t)} \\ &= \frac{\sum_{j=1}^n Y_j(s,w)\mathbf{z}_j(\varphi_j^{-1}(w))}{\sum_{j=1}^n Y_j(s,w)} \\ &= \bar{\mathbf{z}}(\varphi^{-1}(w)). \ \| \end{aligned}$$

Proposition 7 $E_\epsilon[\boldsymbol{x}_i^{\otimes 2}(\varphi_i^{-1}(w))] = \boldsymbol{z}_i^{\otimes 2}(\varphi_i^{-1}(w)) + \boldsymbol{\Xi}$.

Proof:

$E_\epsilon[\mathbf{x}_i^{\otimes 2}(\varphi_i^{-1}(w))]$

$$
\begin{aligned}
&= E_\epsilon[\mathbf{x}_i(\varphi_i^{-1}(w))\mathbf{x}_i^T(\varphi_i^{-1}(w))] \\
&= E_\epsilon \begin{bmatrix} x_{i,1}^2(\varphi_i^{-1}(w)) & \dots & x_{i,1}(\varphi_i^{-1}(w))x_{i,p}(\varphi_i^{-1}(w)) \\ x_{i,1}(\varphi_i^{-1}(w))x_{i,2}(\varphi_i^{-1}(w)) & \dots & x_{i,2}(\varphi_i^{-1}(w))x_{i,p}(\varphi_i^{-1}(w)) \\ \vdots & \ddots & \vdots \\ x_{i,1}(\varphi_i^{-1}(w))x_{i,p}(\varphi_i^{-1}(w)) & \dots & x_{i,p}^2(\varphi_i^{-1}(w)) \end{bmatrix} \\
&= \begin{bmatrix} E_\epsilon[x_{i,1}^2(\varphi_i^{-1}(w))] & \dots & E_\epsilon[x_{i,1}(\varphi_i^{-1}(w))x_{i,p}(\varphi_i^{-1}(w))] \\ E_\epsilon[x_{i,1}(\varphi_i^{-1}(w))x_{i,2}(\varphi_i^{-1}(w))] & \dots & E_\epsilon[x_{i,2}(\varphi_i^{-1}(w))x_{i,p}(\varphi_i^{-1}(w))] \\ \vdots & \ddots & \vdots \\ E_\epsilon[x_{i,1}(\varphi_i^{-1}(w))x_{i,p}(\varphi_i^{-1}(w))] & \dots & E_\epsilon[x_{i,p}^2(\varphi_i^{-1}(w))] \end{bmatrix}. \qquad (3)
\end{aligned}
$$

Diagonal elements $E_\epsilon[x_{i,m}^2(\varphi_i^{-1}(w))]$ for $m = 1, 2, ..., p$ are found as below:

$$
\begin{aligned}
E_\epsilon[x_{i,m}^2(\varphi_i^{-1}(w))] &= E_\epsilon[(z_{i,m}(\varphi_i^{-1}(w)) + \epsilon_{i,m}(\varphi_i^{-1}(w)))^2] \\
&= E_\epsilon[z_{i,m}^2(\varphi_i^{-1}(w))] + E_\epsilon[2z_{i,m}(\varphi_i^{-1}(w))\epsilon_{i,m}(\varphi_i^{-1}(w))] \\
&\quad + E_\epsilon[\epsilon_{i,m}^2(\varphi_i^{-1}(w))] \\
&= z_{i,m}^2(\varphi_i^{-1}(w)) + 2z_{i,m}(\varphi_i^{-1}(w))E_\epsilon[\epsilon_{i,m}(\varphi_i^{-1}(w))] \\
&\quad + E_\epsilon[\epsilon_{i,m}^2(\varphi_i^{-1}(w))] \\
&= z_{i,m}^2(\varphi_i^{-1}(w)) + \sigma_m^2. \qquad (4)
\end{aligned}
$$

Off diagonal elements $B = E_\epsilon[x_{i,q}(\varphi_i^{-1}(w))x_{i,m}(\varphi_i^{-1}(w))]$ for $q = 1, ..., p$ and $m = 1, 2, ..., p$ and $q \neq m$ are found as below:

$$\begin{aligned} B &= E_\epsilon[(z_{i,q}(\varphi_i^{-1}(w)) + \epsilon_{i,q}(\varphi_i^{-1}(w)))(z_{i,m}(\varphi_i^{-1}(w)) + \epsilon_{i,m}(\varphi_i^{-1}(w)))] \\ &= E_\epsilon[z_{i,q}(\varphi_i^{-1}(w))z_{i,m}(\varphi_i^{-1}(w))] \\ &\quad +z_{i,q}(\varphi_i^{-1}(w))E_\epsilon[\epsilon_{i,m}(\varphi_i^{-1}(w))] \\ &\quad +z_{i,m}(\varphi_i^{-1}(w))E_\epsilon[\epsilon_{i,q}(\varphi_i^{-1}(w))] \\ &\quad +E_\epsilon[\epsilon_{i,m}(\varphi_i^{-1}(w))]E_\epsilon[\epsilon_{i,q}(\varphi_i^{-1}(w))] \\ &= z_{i,q}(\varphi_i^{-1}(w))z_{i,m}(\varphi_i^{-1}(w)). \end{aligned} \tag{5}$$

From (4) and (5), we obtain

$$\begin{aligned} (3) &= \begin{bmatrix} z_{i,1}^2(\varphi_i^{-1}(w)) & \dots & z_{i,1}(\varphi_i^{-1}(w))z_{i,p}(\varphi_i^{-1}(w)) \\ z_{i,1}(\varphi_i^{-1}(w))z_{i,2}(\varphi_i^{-1}(w)) & \dots & z_{i,2}(\varphi_i^{-1}(w))z_{i,p}(\varphi_i^{-1}(w)) \\ \vdots & \ddots & \vdots \\ z_{i,1}(\varphi_i^{-1}(w))z_{i,p}(\varphi_i^{-1}(w)) & \dots & z_{i,p}^2(\varphi_i^{-1}(w)) \end{bmatrix} \\ &\quad + \begin{bmatrix} \sigma_1^2 & 0 & \dots & 0 \\ 0 & \sigma_2^2 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \sigma_m^2 \end{bmatrix} \\ &= \mathbf{z}_i^{\otimes 2}(\varphi_i^{-1}(w)) + \mathbf{\Xi}. \ \| \end{aligned}$$

Proposition 8 $E_\epsilon[\boldsymbol{x}_j(\varphi_j^{-1}(w))\boldsymbol{x}_i^T(\varphi_i^{-1}(w))] = z_j(\varphi_j^{-1}(w))\boldsymbol{z}_i^T(\varphi_i^{-1}(w))$.

Proof:

$E_\epsilon[\mathbf{x}_j(\varphi_j^{-1}(w))\mathbf{x}_i^T(\varphi_i^{-1}(w))]$

$$
\begin{aligned}
&= E_\epsilon \begin{bmatrix} x_{j,1}(\varphi_i^{-1}(w))x_{i,1}(\varphi_i^{-1}(w)) & \dots & x_{j,1}(\varphi_i^{-1}(w))x_{i,p}(\varphi_i^{-1}(w)) \\ x_{j,2}(\varphi_i^{-1}(w))x_{i,1}(\varphi_i^{-1}(w)) & \dots & x_{j,2}(\varphi_i^{-1}(w))x_{i,p}(\varphi_i^{-1}(w)) \\ \vdots & \ddots & \vdots \\ x_{j,p}(\varphi_i^{-1}(w))x_{i,1}(\varphi_i^{-1}(w)) & \dots & x_{j,p}(\varphi_i^{-1}(w))x_{i,p}(\varphi_i^{-1}(w)) \end{bmatrix} \\
&= \begin{bmatrix} E_\epsilon[x_{j,1}(\varphi_i^{-1}(w))x_{i,1}(\varphi_i^{-1}(w))] & \dots & E_\epsilon[x_{j,1}(\varphi_i^{-1}(w))x_{i,p}(\varphi_i^{-1}(w))] \\ E_\epsilon[x_{j,2}(\varphi_i^{-1}(w))x_{i,1}(\varphi_i^{-1}(w))] & \dots & E_\epsilon[x_{j,2}(\varphi_i^{-1}(w))x_{i,p}(\varphi_i^{-1}(w))] \\ \vdots & \ddots & \vdots \\ E_\epsilon[x_{j,p}(\varphi_i^{-1}(w))x_{i,1}(\varphi_i^{-1}(w))] & \dots & E_\epsilon[x_{j,p}(\varphi_i^{-1}(w))x_{i,p}(\varphi_i^{-1}(w))] \end{bmatrix}. \qquad (6)
\end{aligned}
$$

Element $C = E_\epsilon[x_{j,u}(\varphi_i^{-1}(w))x_{i,v}(\varphi_i^{-1}(w))]$ for $u = 1, 2, ..., p$ and $v = 1, 2, ..., p$ of the above matrix (6) is fond as follows:

$$
\begin{aligned}
C &= E_\epsilon[(z_{j,u}(\varphi_i^{-1}(w)) + \epsilon_{j,u}(\varphi_i^{-1}(w)))(z_{i,v}(\varphi_i^{-1}(w)) + \epsilon_{i,v}(\varphi_i^{-1}(w)))] \\
&= E_\epsilon[z_{j,u}(\varphi_i^{-1}(w))z_{i,v}(\varphi_i^{-1}(w))] \\
&\quad + z_{j,u}(\varphi_i^{-1}(w))E_\epsilon[\epsilon_{i,v}(\varphi_i^{-1}(w))] \\
&\quad + z_{i,v}(\varphi_i^{-1}(w))E_\epsilon[\epsilon_{j,u}(\varphi_i^{-1}(w))] \\
&\quad + E_\epsilon[\epsilon_{i,v}(\varphi_i^{-1}(w))]E_\epsilon[\epsilon_{j,u}(\varphi_i^{-1}(w))] \\
&= z_{j,u}(\varphi_i^{-1}(w))z_{i,v}(\varphi_i^{-1}(w)).
\end{aligned}
$$

Hence, (6) can be written as

$$
\begin{aligned}
(6) &= \begin{bmatrix} z_{j,1}(\varphi_i^{-1}(w))z_{i,1}(\varphi_i^{-1}(w)) & \dots & z_{j,1}(\varphi_i^{-1}(w))z_{i,p}(\varphi_i^{-1}(w)) \\ z_{j,2}(\varphi_i^{-1}(w))z_{i,1}(\varphi_i^{-1}(w)) & \dots & z_{j,2}(\varphi_i^{-1}(w))z_{i,p}(\varphi_i^{-1}(w)) \\ \vdots & \ddots & \vdots \\ z_{j,p}(\varphi_i^{-1}(w))z_{i,1}(\varphi_i^{-1}(w)) & \dots & z_{j,p}(\varphi_i^{-1}(w))z_{i,p}(\varphi_i^{-1}(w)) \end{bmatrix} \\
&= \mathbf{z}_j(\varphi_j^{-1}(w))\mathbf{z}_i^T(\varphi_i^{-1}(w)). \quad \|
\end{aligned}
$$

Proposition 9 $E_\epsilon\left\{\bar{\boldsymbol{x}}(\varphi^{-1}(w))\boldsymbol{x}_i^T(\varphi_i^{-1}(w))\right\} = \bar{\boldsymbol{z}}(\varphi^{-1}(w))\boldsymbol{z}_i^T(\varphi_i^{-1}(w)) + \frac{Y_i(s,w)\boldsymbol{\Xi}}{\sum_{j=1}^n Y_j(s,w)}$.

Proof:

$$
\begin{aligned}
&E_\epsilon\left\{\bar{\mathbf{x}}(\varphi^{-1}(w))\mathbf{x}_i^T(\varphi_i^{-1}(w))\right\} \\
&\quad = E_\epsilon\left\{\left[\frac{\sum_{j=1}^n Y_j(s,w)\mathbf{x}_j(\varphi_j^{-1}(w))}{\sum_{j=1}^n Y_j(s,w)}\right]\mathbf{x}_i^T(\varphi_i^{-1}(w))\right\} \\
&\quad = E_\epsilon\left\{\frac{Y_i(s,w)\mathbf{x}_i(\varphi_i^{-1}(w))\mathbf{x}_i^T(\varphi_i^{-1}(w))}{\sum_{j=1}^n Y_j(s,w)}\right\} \\
&\qquad + E_\epsilon\left\{\frac{\sum_{i\neq j} Y_j(s,w)\mathbf{x}_j(\varphi_j^{-1}(w))\mathbf{x}_i^T(\varphi_i^{-1}(w))}{\sum_{j=1}^n Y_j(s,w)}\right\} \\
&\quad = \frac{Y_i(s,w)E_\epsilon[\mathbf{x}_i(\varphi_i^{-1}(w))^{\otimes 2}]}{\sum_{j=1}^n Y_j(s,w)} \\
&\qquad + \frac{\sum_{i\neq j} Y_j(s,w)E_\epsilon[\mathbf{x}_j(\varphi_j^{-1}(w))\mathbf{x}_i^T(\varphi_i^{-1}(w))]}{\sum_{j=1}^n Y_j(s,w)} \\
&\quad = \frac{Y_i(s,w)[\mathbf{z}_i^{\otimes 2}(\varphi_i^{-1}(w)) + \boldsymbol{\Xi}]}{\sum_{j=1}^n Y_j(s,w)} \\
&\qquad + \frac{\sum_{i\neq j} Y_j(s,w)[\mathbf{z}_j(\varphi_j^{-1}(w))\mathbf{z}_i^T(\varphi_i^{-1}(w))]}{\sum_{j=1}^n Y_j(s,w)} \\
&\quad = \left[\frac{\sum_{j=1}^n Y_j(s,w)\mathbf{z}_j(\varphi_j^{-1}(w))}{\sum_{j=1}^n Y_j(s,w)}\right]\mathbf{z}_i^T(\varphi_i^{-1}(w)) + \frac{Y_i(s,w)\boldsymbol{\Xi}}{\sum_{j=1}^n Y_j(s,w)} \\
&\quad = \bar{\mathbf{z}}(\varphi^{-1}(w))\mathbf{z}_i^T(\varphi_i^{-1}(w)) + \frac{Y_i(s,w)\boldsymbol{\Xi}}{\sum_{j=1}^n Y_j(s,w)}. \quad \|
\end{aligned}
$$

Proposition 10

$$
\begin{aligned}
&\sum_{i=1}^n \int_0^t Y_i(s,w)[\boldsymbol{x}_i(\varphi_i^{-1}(w)) - \bar{\boldsymbol{x}}(\varphi^{-1}(w))]\boldsymbol{x}_i^T(\varphi_i^{-1}(w)dw \\
&\qquad = \sum_{i=1}^n \int_0^t Y_i(s,w)[\boldsymbol{x}_i(\varphi_i^{-1}(w)) - \bar{\boldsymbol{x}}(\varphi^{-1}(w))]^{\otimes 2}dw.
\end{aligned}
$$

Proof:

$\sum_{i=1}^n \int_0^t Y_i(s,w)[\mathbf{x}_i(\varphi_i^{-1}(w)) - \bar{\mathbf{x}}(\varphi^{-1}(w))]^{\otimes 2}dw$

$$
\begin{aligned}
&= \sum_{i=1}^{n} \int_0^t Y_i(s,w)[\mathbf{x}_i(\varphi_i^{-1}(w)) - \bar{\mathbf{x}}(\varphi^{-1}(w))][\mathbf{x}_i(\varphi_i^{-1}(w)) - \bar{\mathbf{x}}(\varphi^{-1}(w))]^T dw \\
&= \sum_{i=1}^{n} \int_0^t Y_i(s,w)[\mathbf{x}_i(\varphi_i^{-1}(w)) - \bar{\mathbf{x}}(\varphi^{-1}(w))]\mathbf{x}_i^T(\varphi_i^{-1}(w)dw \\
&\quad - \sum_{i=1}^{n} \int_0^t Y_i(s,w)[\mathbf{x}_i(\varphi_i^{-1}(w)) - \bar{\mathbf{x}}(\varphi^{-1}(w))]\bar{\mathbf{x}}^T(\varphi^{-1}(w))dw \\
&= \sum_{i=1}^{n} \int_0^t Y_i(s,w)[\mathbf{x}_i(\varphi_i^{-1}(w)) - \bar{\mathbf{x}}(\varphi^{-1}(w))]\mathbf{x}_i^T(\varphi_i^{-1}(w)dw \\
&\quad - \sum_{i=1}^{n} \int_0^t Y_i(s,w)\left[\mathbf{x}_i(\varphi_i^{-1}(w)) - \frac{\sum_{j=1}^n Y_j(s,w)\mathbf{x}_j(\varphi_j^{-1}(w))}{\sum_{j=1}^n Y_j(s,w)}\right]\bar{\mathbf{x}}^T(\varphi^{-1}(w))dw \\
&= \sum_{i=1}^{n} \int_0^t Y_i(s,w)[\mathbf{x}_i(\varphi_i^{-1}(w)) - \bar{\mathbf{x}}(\varphi^{-1}(w))]\mathbf{x}_i^T(\varphi_i^{-1}(w)dw \\
&\quad - \int_0^t \left[\sum_{i=1}^{n} Y_i(s,w)\mathbf{x}_i(\varphi_i^{-1}(w)) \right. \\
&\quad \left. - \sum_{i=1}^{n} Y_i(s,w) \cdot \frac{\sum_{j=1}^n Y_j(s,w)\mathbf{x}_j(\varphi_j^{-1}(w))}{\sum_{j=1}^n Y_j(s,w)}\right]\bar{\mathbf{x}}^T(\varphi^{-1}(w))dw \\
&= \sum_{i=1}^{n} \int_0^t Y_i(s,w)[\mathbf{x}_i(\varphi_i^{-1}(w)) - \bar{\mathbf{x}}(\varphi^{-1}(w))]\mathbf{x}_i^T(\varphi_i^{-1}(w)dw. \ \|
\end{aligned}
$$

We examine the expected value of $\mathbf{U}(\boldsymbol{\beta};\mathbf{x},s,t)$, where covariates are subject to classical additive errors. Applying propositions 5, 6, 7, and 9 we obtain

$$\begin{aligned}
E_\epsilon\left\{\mathbf{U}(\boldsymbol{\beta};\mathbf{x},s,t)\right\} &= E_\epsilon\left\{\sum_{i=1}^{n}\int_0^t[\mathbf{x}_i(\varphi_i^{-1}(w))-\bar{\mathbf{x}}(\varphi^{-1}(w))]\right.\\
&\quad\left.\times\,[N_i(s,dw)-Y_i(s,w)\boldsymbol{\beta}^T\mathbf{x}_i(\varphi_i^{-1}(w))dw]\right\}\\
&= \sum_{i=1}^{n}\int_0^t[E_\epsilon\left\{\mathbf{x}_i(\varphi_i^{-1}(w))\right\}-E_\epsilon\left\{\bar{\mathbf{x}}(\varphi^{-1}(w))\right\}]N_i(s,dw)\\
&\quad-\sum_{i=1}^{n}\int_0^t Y_i(s,w)E_\epsilon\left\{[\boldsymbol{\beta}^T\mathbf{x}_i(\varphi_i^{-1}(w))]\mathbf{x}_i(\varphi_i^{-1}(w))\right\}dw\\
&\quad+\sum_{i=1}^{n}\int_0^t Y_i(s,w)E_\epsilon\left\{[\boldsymbol{\beta}^T\mathbf{x}_i(\varphi_i^{-1}(w))]\bar{\mathbf{x}}(\varphi^{-1}(w))\right\}dw.\\
&= \sum_{i=1}^{n}\int_0^t[\mathbf{z}_i(\varphi_i^{-1}(w))-\bar{\mathbf{z}}(\varphi^{-1}(w))]N_i(s,dw)\\
&\quad-\sum_{i=1}^{n}\int_0^t Y_i(s,w)[\mathbf{z}_i^{\otimes 2}(\varphi_i^{-1}(w))+\boldsymbol{\Xi}]\boldsymbol{\beta}dw\\
&\quad+\sum_{i=1}^{n}\int_0^t Y_i(s,w)\left\{\bar{\mathbf{z}}(\varphi^{-1}(w))\mathbf{z}_i^T(\varphi_i^{-1}(w))\right.\\
&\quad\left.+\frac{Y_i(s,w)\boldsymbol{\Xi}}{\sum_{j=1}^{n}Y_j(s,w)}\right\}\boldsymbol{\beta}dw\\
&= \sum_{i=1}^{n}\int_0^t[\mathbf{z}_i(\varphi_i^{-1}(w))-\bar{\mathbf{z}}(\varphi^{-1}(w))]N_i(s,dw)\\
&\quad-\sum_{i=1}^{n}\int_0^t Y_i(s,w)[\boldsymbol{\beta}^T\mathbf{z}_i(\varphi_i^{-1}(w))\mathbf{z}_i(\varphi_i^{-1}(w))]dw\\
&\quad+\sum_{i=1}^{n}\int_0^t Y_i(s,w)\left\{[\boldsymbol{\beta}^T\mathbf{z}_i(\varphi_i^{-1}(w))]\bar{\mathbf{z}}(\varphi^{-1}(w))\right\}dw\\
&\quad-\sum_{i=1}^{n}\int_0^t\boldsymbol{\Xi}\boldsymbol{\beta}\left[1-\frac{Y_i(s,w)}{\sum_{j=1}^{n}Y_j(s,w)}\right]Y_i(s,w)dw\\
E_\epsilon\left\{\mathbf{U}(\boldsymbol{\beta};\mathbf{x},s,t)\right\} &= \mathbf{U}(\boldsymbol{\beta};\mathbf{z},s,t)-\sum_{i=1}^{n}\int_0^t\boldsymbol{\Xi}\boldsymbol{\beta}\left[1-\frac{Y_i(s,w)}{\sum_{j=1}^{n}Y_j(s,w)}\right]Y_i(s,w)dw. \quad (7)
\end{aligned}$$

Note that rearranging the terms in (7) yields

$$E_\epsilon \left\{ \mathbf{U}(\boldsymbol{\beta}; \mathbf{x}, s, t) + \sum_{i=1}^{n} \int_0^t \boldsymbol{\Xi\beta} \left[1 - \frac{Y_i(s,w)}{\sum_{j=1}^{n} Y_j(s,w)} \right] Y_i(s,w) dw \right\} = U(\boldsymbol{\beta}; \mathbf{z}, s, t).$$

From the foregoing derivation, the function $\mathbf{U}^*(\boldsymbol{\beta}; \mathbf{x}, s, t)$ is given by

$$\mathbf{U}^*(\boldsymbol{\beta}; \mathbf{x}, s, t) = \mathbf{U}(\boldsymbol{\beta}; \mathbf{x}, s, t) + \sum_{i=1}^{n} \int_0^t \boldsymbol{\Xi\beta} \left[1 - \frac{Y_i(s,w)}{\sum_{j=1}^{n} Y_j(s,w)} \right] Y_i(s,w) dw.$$

Since $E[\mathbf{U}(\boldsymbol{\beta}; \mathbf{z}, s, t)] = \mathbf{0}$, the corrected score function $\mathbf{U}^*(\boldsymbol{\beta}; \mathbf{x}, s, t)$ is an unbiased estimating function. Setting $\mathbf{U}^*(\boldsymbol{\beta}; \mathbf{x}, s, t)$ to $\mathbf{0}$, we obtain

$$\sum_{i=1}^{n} \int_0^t [\mathbf{x}_i(\varphi_i^{-1}(w)) - \bar{\mathbf{x}}(\varphi^{-1}(w))] N_i(s, dw) \quad = \quad \{A + B\} \boldsymbol{\beta},$$

where $A = \sum_{i=1}^{n} \int_0^t Y_i(s,w)[\mathbf{x}_i(\varphi_i^{-1}(w)) - \bar{\mathbf{x}}(\varphi^{-1}(w))]\mathbf{x}_i^T(\varphi_i^{-1}(w) dw$,
and $B = \sum_{i=1}^{n} \int_0^t \boldsymbol{\Xi} \left[1 - \frac{Y_i(s,w)}{\sum_{j=1}^{n} Y_j(s,w)} \right] Y_i(s,w) dw$. This gives us the following:

$$\begin{aligned} \hat{\boldsymbol{\beta}}_C \quad = \quad & \left\{ \sum_{i=1}^{n} \int_0^t Y_i(s,w)[\mathbf{x}_i(\varphi_i^{-1}(w)) - \bar{\mathbf{x}}(\varphi^{-1}(w))]\mathbf{x}_i^T(\varphi_i^{-1}(w) dw \right. \\ & \left. + \sum_{i=1}^{n} \int_0^t \boldsymbol{\Xi} \left[1 - \frac{Y_i(s,w)}{\sum_{j=1}^{n} Y_j(s,w)} \right] Y_i(s,w) dw \right\}^{-1} \\ & \times \left\{ \sum_{i=1}^{n} \int_0^t [\mathbf{x}_i(\varphi_i^{-1}(w)) - \bar{\mathbf{x}}(\varphi^{-1}(w))] N_i(s, dw) \right\}. \end{aligned}$$

By proposition 10, the corrected estimators, $\hat{\boldsymbol{\beta}}_C$ is the following form:

$$\begin{aligned} \hat{\boldsymbol{\beta}}_C \quad = \quad & \left\{ \sum_{i=1}^{n} \int_0^t Y_i(s,w)[\mathbf{x}_i(\varphi_i^{-1}(w)) - \bar{\mathbf{x}}(\varphi^{-1}(w))]^{\otimes 2} dw \right. \\ & \left. + \sum_{i=1}^{n} \int_0^t \boldsymbol{\Xi} \left[1 - \frac{Y_i(s,w)}{\sum_{j=1}^{n} Y_j(s,w)} \right] Y_i(s,w) dw \right\}^{-1} \\ & \times \left\{ \sum_{i=1}^{n} \int_0^t [\mathbf{x}_i(\varphi_i^{-1}(w)) - \bar{\mathbf{x}}(\varphi^{-1}(w))] N_i(s, dw) \right\}. \end{aligned}$$

5. LARGE SAMPLE PROPERTIES

Consistency of the estimators play a major role in estimation. This section is devoted to the consistency of corrected regression parameter estimators.

5.1. REGULARITY CONDITIONS

I. $N_i(s,t)$ is bounded for all i.

II. $P\{Y_i(s,t) = 1\} > 0$ for all i.

III. For $i = 1, ..., n$; $\sup_{t\in\tau} ||E[\mathbf{z}_i^{\otimes 2}(\varphi_i^{-1}(t))]|| < \infty$.

IV. The processes $\mathbf{z}_i(\varphi^{-1}(t))$ are of bounded total variation for all i and $t \in \tau$.

V. For $\alpha_i(s,t)$ being any of the functions $Y_i(s,t)$, $N_i(s,t)$, and $\mathbf{z}_i(\varphi^{-1}(t))$ or any function that can be expressed as a summation or a multiplication of the functions $Y_i(s,t)$, $N_i(s,t)$, and $\mathbf{z}_i(\varphi^{-1}(t))$, we have

(a)

$$\alpha(s,t) = \lim_{n\to\infty} \frac{1}{n} \sum_{i=1}^{n} E[\alpha_i(s,t)] < \infty.$$

(b) There exist envelopes $F_i(s^\star, t)$; $t \in \tau$ such that

$$\sum_{i=1}^{n} \frac{E(F_i^{\otimes 2}(s^\star, t))}{i^2} < \infty.$$

VI. For $i = 1, ..., n$; $||E[\boldsymbol{\epsilon}_i^{\otimes 2}(\varphi_i^{-1}(t))]|| < \infty$.

VII. For $i = 1, ..., n$; there exists non singular matrices defined as

$$\int_0^t E\left\{Y_i(s,w)\left[\mathbf{z}_i(\varphi_i^{-1}(w)) - \frac{E[Y_i(s,w)\mathbf{z}_i(\varphi_i^{-1}(w))]}{E[Y_i(s,w)]}\right]^{\otimes 2}\right\} dw.$$

Condition I. ensures that the number of events experienced by each subject is finite. Condition II. requires that all subjects in the study have a chance of being observed, which in turn guarantees that $E[Y_i(s, w)]$ is bounded away from zero. This is important for the proofs where $E[Y_i(s, w)]$ term appears in the denominator. Condition III. assumes bounded variation for the covariate process. Conditions IV. and V. are used to establish manageability which is needed to prove almost sure convergence of some quantities that appear in the proofs. Condition VI. controls the variability of the measurement error vectors, by making sure they have finite variance. Condition VII. is required to establish the consistency of the corrected regression parameter vector.

Lemma 4 *There exist finite functions*

$$\alpha(s, t) = \lim_{n\to\infty} \frac{1}{n} \sum_{i=1}^{n} E[\alpha_i(s, t)],$$

where $\alpha_i(s, t)$ can be expressed as a summation or a multiplication of the functions $Y_i(s, t)$, $N_i(s, t)$, and $z_i(\varphi^{-1}(t))$, such that

$$\sup_{t\in\tau} \left| \frac{1}{n} \sum_{i=1}^{n} \alpha_i(s, t) - \alpha(s, t) \right| \overset{a.s.}{\to} 0. \tag{8}$$

Proof:

$Y_i(s, t)$ is a monotonically increasing function. Therefore, according to the lemma A.2 from Bilias et al. (1997), it is manageable. $N_i(s, t)$ is also manageable, since $N_i(s, t)$ satisfies regularity condition I.. Additionally, $\mathbf{z}_i(\varphi^{-1}(w))$ is also manageable by regularity condition IV.. The preservation result for summations and products of manageable processes hence establishes the fact that any function which can be expressed as a summation or a multiplication of the functions $Y_i(s, t)$, $N_i(s, t)$ and $\mathbf{z}_i(\varphi^{-1}(w))$ is also manageable. Finally, by using regularity condition V. and applying the Uniform Strong Law of Large Numbers from Pollard (1990), we can establish (8). $\|$

5.2. CONSISTENCY OF $\hat{\boldsymbol{\beta}}_c$

Stocker and Adekpedjou (2020) showed that in the absence of measurement errors, the estimator $\hat{\boldsymbol{\beta}}_n$ can be estimated by

$$\hat{\boldsymbol{\beta}}_n = M_n^{-1} W_n,$$

where

$$M_n = \frac{1}{n}\sum_{i=1}^{n}\int_0^t Y_i(s,w)[\mathbf{z}_i(\varphi_i^{-1}(w)) - \bar{\mathbf{z}}(\varphi^{-1}(w))]^{\otimes 2} dw,$$

and

$$W_n = \frac{1}{n}\sum_{i=1}^{n}\int_0^t [\mathbf{z}_i(\varphi_i^{-1}(w)) - \bar{\mathbf{z}}(\varphi^{-1}(w))] N_i(s,dw).$$

They also showed that $\hat{\boldsymbol{\beta}}_n \overset{a.s.}{\to} \boldsymbol{\beta}_0$ as $n \to \infty$. To establish the consistency of the corrected regression parameter vector, we utilize this finding along with several lemmas that will be derived in this sequel.

Lemma 5 *As $n \to \infty$, $\bar{z}(\varphi^{-1}(w)) \overset{a.s.}{\to} \boldsymbol{e}(\varphi^{-1}(w))$, where*

$$\boldsymbol{e}(\varphi^{-1}(w)) = \frac{E[Y_i(s,w)\boldsymbol{z}_i(\varphi^{-1}(w))]}{E[Y_i(s,w)]}.$$

Proof: By lemma (4), we have $\frac{1}{n}\sum_{i=1}^{n} Y_i(s,w)\mathbf{z}_i(\varphi^{-1}(w)) \overset{a.s.}{\to} E[Y_i(s,w)\mathbf{z}_i(\varphi^{-1}(w))]$ uniformly in w. Since $Y_i(s,w)$ is a monotonically increasing function, according to the lemma A.2 from Bilias et al. (1997) it is manageable. So, under regularity condition (V.) applying the Strong Uniform Law of Large Numbers we get, $\frac{1}{n}\sum_{i=1}^{n} Y_i(s,w) \overset{a.s.}{\to} E[Y_i(s,w)]$ uniformly in w. Finally by regularity condition II., and the Strong Uniform Law of Large Numbers, we obtain $\frac{\sum_{j=1}^{n} Y_j(s,w)\mathbf{z}_j(\varphi_j^{-1}(w))}{\sum_{j=1}^{n} Y_j(s,w)} \overset{a.s.}{\to} \mathbf{e}(\varphi^{-1}(w))$. $\|$

Lemma 6 *As $n \to \infty$, $M_n \overset{a.s.}{\to} B_1$, where*

$$B_1 = E\left[\int_0^t Y_i(s,w)[\boldsymbol{z}_i(\varphi_i^{-1}(w)) - \boldsymbol{e}(\varphi^{-1}(w))]^{\otimes 2} dw\right].$$

Proof:

$$
\begin{aligned}
M_n &= \frac{1}{n}\sum_{i=1}^{n}\int_0^t Y_i(s,w)[\mathbf{z}_i(\varphi_i^{-1}(w)) - \bar{\mathbf{z}}(\varphi^{-1}(w))]^{\otimes 2}dw \\
&= \frac{1}{n}\sum_{i=1}^{n}\int_0^t Y_i(s,w)[\mathbf{z}_i(\varphi_i^{-1}(w)) - \mathbf{e}(\varphi_i^{-1}(w)) + \mathbf{e}(\varphi_i^{-1}(w)) - \bar{\mathbf{z}}(\varphi^{-1}(w))]^{\otimes 2}dw \\
&= \frac{1}{n}\sum_{i=1}^{n}\int_0^t Y_i(s,w)[\mathbf{z}_i(\varphi_i^{-1}(w)) - \mathbf{e}(\varphi_i^{-1}(w))]^{\otimes 2}dw \\
&\quad + \frac{1}{n}\sum_{i=1}^{n}\int_0^t Y_i(s,w)[\mathbf{e}(\varphi_i^{-1}(w)) - \bar{\mathbf{z}}(\varphi^{-1}(w))]^{\otimes 2}dw \\
&\quad + \frac{1}{n}\sum_{i=1}^{n}\int_0^t 2Y_i(s,w)[\mathbf{z}_i(\varphi_i^{-1}(w)) - \mathbf{e}(\varphi_i^{-1}(w))][\mathbf{e}(\varphi_i^{-1}(w)) - \bar{\mathbf{z}}(\varphi^{-1}(w))]^T dw \\
&= \frac{1}{n}\sum_{i=1}^{n}\int_0^t Y_i(s,w)[\mathbf{z}_i(\varphi_i^{-1}(w)) - \mathbf{e}(\varphi_i^{-1}(w))]^{\otimes 2}dw \\
&\quad + \frac{1}{n}\sum_{i=1}^{n}\int_0^t Y_i(s,w)[-\mathbf{e}(\varphi^{-1}(w))^{\otimes 2} + 2\mathbf{z}_i(\varphi_i^{-1}(w))\mathbf{e}^T(\varphi^{-1}(w)) \\
&\quad + \bar{\mathbf{z}}(\varphi^{-1}(w))^{\otimes 2} - 2\bar{\mathbf{z}}(\varphi^{-1}(w))\mathbf{z}_i^T(\varphi_i^{-1}(w))]dw \\
&= \frac{1}{n}\sum_{i=1}^{n}\int_0^t Y_i(s,w)[\mathbf{z}_i(\varphi_i^{-1}(w)) - \mathbf{e}(\varphi^{-1}(w))]^{\otimes 2}dw \\
&\quad + \frac{1}{n}\sum_{i=1}^{n}\int_0^t Y_i(s,w)[-\mathbf{e}(\varphi^{-1}(w))^{\otimes 2} + 2\mathbf{z}_i(\varphi_i^{-1}(w))\mathbf{e}^T(\varphi^{-1}(w)) \\
&\quad + \mathbf{e}(\varphi^{-1}(w))^{\otimes 2} - 2\mathbf{e}(\varphi^{-1}(w))\mathbf{z}_i^T(\varphi_i^{-1}(w))]dw + 0_{a.s.}(1) \quad \text{(applying lemma 5)} \\
&= \frac{1}{n}\sum_{i=1}^{n}\int_0^t Y_i(s,w)[\mathbf{z}_i(\varphi_i^{-1}(w)) - \mathbf{e}(\varphi^{-1}(w))]^{\otimes 2}dw + 0_{a.s.}(1) \\
&= E\left[\int_0^t Y_i(s,w)[\mathbf{z}_i(\varphi_i^{-1}(w)) - \mathbf{e}(\varphi^{-1}(w))]^{\otimes 2}dw\right] + 0_{a.s.}(1) \\
&= B_1 + 0_{a.s.}(1). \quad \|
\end{aligned}
$$

Lemma 7 *As $n \to \infty$, $W_n \overset{a.s.}{\to} C_1$, where*

$$
C_1 = E\left[\int_0^t [z_i(\varphi_i^{-1}(w)) - \boldsymbol{e}(\varphi^{-1}(w))]N_i(s,dw)\right].
$$

Proof: Under regularity conditions I. and V., applying the Strong Uniform Law of Large Numbers, we get $\frac{1}{n}\sum_{i=1}^{n} N_i(s,w) \overset{a.s.}{\to} E[N_i(s,w)]$ uniformly in w. By lemma (4) , as $n \to \infty$, $n^{-1}\sum_{i=1}^{n}\int_0^t \mathbf{z}_i(\varphi^{-1}(w))N_i(s,dw) \overset{a.s.}{\to} \int_0^t E[\mathbf{z}_i(\varphi^{-1}(w))N_i(s,dw)]$. Therefore, we obtain

$$\begin{aligned}
W_n &= \frac{1}{n}\sum_{i=1}^{n}\int_0^t [\mathbf{z}_i(\varphi_i^{-1}(w)) - \bar{\mathbf{z}}(\varphi^{-1}(w))]N_i(s,dw) \\
&= \frac{1}{n}\sum_{i=1}^{n}\int_0^t [\mathbf{z}_i(\varphi_i^{-1}(w))]N_i(s,dw) - [\bar{\mathbf{z}}(\varphi^{-1}(w))]\int_0^t \frac{1}{n}\sum_{i=1}^{n} N_i(s,dw) \\
&= \int_0^t E[\mathbf{z}_i(\varphi^{-1}(w))N_i(s,dw)] - \int_0^t E[\mathbf{e}(\varphi^{-1}(w))N_i(s,dw)] + 0_{a.s.}(1) \\
&= E\left[\int_0^t [\mathbf{z}_i(\varphi_i^{-1}(w)) - \mathbf{e}(\varphi^{-1}(w))]N_i(s,dw)\right] + 0_{a.s.}(1) \\
&= C_1 + 0_{a.s.}(1). \ \|
\end{aligned}$$

Hence by lemma 6, lemma 7, and Stocker and Adekpedjou (2020), we get that

$$B_1^{-1}C_1 = \boldsymbol{\beta}_0. \tag{9}$$

As previously proven, the corrected regression parameter vector can be expressed as follows:

$$\hat{\boldsymbol{\beta}}_C = L_n^{-1}K_n, \tag{10}$$

where

$$\begin{aligned}
L_n &= \frac{1}{n}\left\{\sum_{i=1}^{n}\int_0^t Y_i(s,w)[\mathbf{x}_i(\varphi_i^{-1}(w)) - \bar{\mathbf{x}}(\varphi^{-1}(w))]^{\otimes 2}dw \right. \\
&\quad \left. - \sum_{i=1}^{n}\int_0^t \boldsymbol{\Xi}\left[1 - \frac{Y_i(s,w)}{\sum_{j=1}^{n} Y_j(s,w)}\right] Y_i(s,w)dw\right\},
\end{aligned}$$

and

$$K_n \;=\; \frac{1}{n}\left\{\sum_{i=1}^{n}\int_0^t [\mathbf{x}_i(\varphi_i^{-1}(w)) - \bar{\mathbf{x}}(\varphi^{-1}(w))]N_i(s,dw)\right\}.$$

Lemma 8 *As $n \to \infty$, $L_n \overset{a.s.}{\to} B_1$, where*

$$L_n = Q_1 - Q_2,$$

$$Q_1 = \frac{1}{n}\sum_{i=1}^{n}\int_0^t Y_i(s,w)[\boldsymbol{x}_i(\varphi_i^{-1}(w)) - \bar{\boldsymbol{x}}(\varphi^{-1}(w))]^{\otimes 2}dw,$$

and

$$Q_2 = \frac{1}{n}\left[1 - \frac{1}{\sum_{j=1}^{n} Y_j(s,w)}\right]\sum_{i=1}^{n} Y_i(s,w)\boldsymbol{\Xi}dw.$$

Proof: First, we investigate the Q_1 term.

$$\begin{aligned}
Q_1 &= \frac{1}{n}\sum_{i=1}^{n}\int_0^t Y_i(s,w)[\mathbf{x}_i(\varphi_i^{-1}(w)) - \bar{\mathbf{x}}(\varphi^{-1}(w))]^{\otimes 2}dw \\
&= \frac{1}{n}\sum_{i=1}^{n}\int_0^t Y_i(s,w)[\mathbf{x}_i(\varphi_i^{-1}(w)) - \mathbf{e}(\varphi^{-1}(w))) + \mathbf{e}(\varphi^{-1}(w))) - \bar{\mathbf{x}}(\varphi^{-1}(w))]^{\otimes 2}dw \\
&= \frac{1}{n}\sum_{i=1}^{n}\int_0^t Y_i(s,w)[\mathbf{x}_i(\varphi_i^{-1}(w)) - \mathbf{e}(\varphi^{-1}(w)))]^{\otimes 2}dw \\
&\quad + \frac{1}{n}\sum_{i=1}^{n}\int_0^t Y_i(s,w)[\mathbf{e}(\varphi^{-1}(w))) - \bar{\mathbf{x}}(\varphi^{-1}(w))]^{\otimes 2}dw \\
&\quad + \frac{1}{n}\sum_{i=1}^{n}\int_0^t 2Y_i(s,w)[\mathbf{x}_i(\varphi_i^{-1}(w)) - \mathbf{e}(\varphi^{-1}(w)))][\mathbf{e}(\varphi^{-1}(w))) - \bar{\mathbf{x}}(\varphi^{-1}(w))]^T dw \\
&= \frac{1}{n}\sum_{i=1}^{n}\int_0^t Y_i(s,w)[\mathbf{x}_i(\varphi_i^{-1}(w)) - \mathbf{e}(\varphi^{-1}(w)))]^{\otimes 2}dw \\
&\quad + \int_0^t \left\{-\frac{1}{n}\sum_{i=1}^{n} Y_i(s,w)\mathbf{e}^{\otimes 2}(\varphi_i^{-1}(w)) - \frac{1}{n}\frac{\left\{\sum_{j=1}^{n} Y_j(s,w)\mathbf{x}_i(\varphi_i^{-1}(w))\right\}^{\otimes 2}}{\sum_{j=1}^{n} Y_j(s,w)}\right. \\
&\quad \left. + \frac{2}{n}\sum_{i=1}^{n} Y_i(s,w)\mathbf{x}_i(\varphi_i^{-1}(w))\mathbf{e}^T(\varphi^{-1}(w)))\right\} dw.
\end{aligned}$$

Now, we examine the second term (say G_n) of the Q_1 term above by applying the Strong Uniform Law of Large Numbers individually to each term.

$$\begin{aligned}
G_n &= \int_0^t \left\{ -\frac{1}{n}\sum_{i=1}^n Y_i(s,w)\mathbf{e}^{\otimes 2}(\varphi_i^{-1}(w)) - \frac{1}{n}\frac{\left\{\sum_{j=1}^n Y_j(s,w)\mathbf{x}_i(\varphi_i^{-1}(w))\right\}^{\otimes 2}}{\sum_{j=1}^n Y_j(s,w)} \right. \\
&\quad \left. + \frac{2}{n}\sum_{i=1}^n Y_i(s,w)\mathbf{x}_i(\varphi_i^{-1}(w))\mathbf{e}^T(\varphi_i^{-1}(w)) \right\} dw \\
&= \int_0^t \left\{ -E[Y_i(s,w)]\mathbf{e}^{\otimes 2}(\varphi_i^{-1}(w)) - \frac{\left\{E[Y_j(s,w)\mathbf{x}_i(\varphi_i^{-1}(w))]\right\}^{\otimes 2}}{E[Y_j(s,w)]} \right. \\
&\quad \left. + 2E[Y_i(s,w)\mathbf{x}_i(\varphi_i^{-1}(w))]\mathbf{e}^T(\varphi_i^{-1}(w)) \right\} dw + 0_{a.s.}(1) \\
&= \int_0^t \left\{ -E[Y_i(s,w)]\mathbf{e}^{\otimes 2}(\varphi_i^{-1}(w)) - \frac{\left\{E[Y_j(s,w)\mathbf{z}_i(\varphi_i^{-1}(w))]\right\}^{\otimes 2}}{E[Y_j(s,w)]} \right. \\
&\quad \left. + 2E[Y_i(s,w)\mathbf{z}_i(\varphi_i^{-1}(w))]\mathbf{e}^T(\varphi_i^{-1}(w)) \right\} dw + 0_{a.s.}(1) \\
&= \int_0^t \left\{ -E[Y_i(s,w)]\mathbf{e}^{\otimes 2}(\varphi_i^{-1}(w)) - E[Y_i(s,w)]\mathbf{e}^{\otimes 2}(\varphi_i^{-1}(w)) \right. \\
&\quad \left. +2E[Y_i(s,w)]\mathbf{e}^{\otimes 2}(\varphi_i^{-1}(w)) \right\} dw + 0_{a.s.}(1) \\
&= 0_{a.s.}(1).
\end{aligned}$$

Then, we continue applying the Strong Uniform Law of Large Numbers to the remaining terms of Q_1 as follows:

$$
\begin{aligned}
Q_1 &= \frac{1}{n}\sum_{i=1}^{n}\int_0^t Y_i(s,w)[\mathbf{x}_i(\varphi_i^{-1}(w)) - \mathbf{e}(\varphi_i^{-1}(w))]^{\otimes 2}dw + 0_{a.s.}(1) \\
&= \frac{1}{n}\sum_{i=1}^{n}\int_0^t Y_i(s,w)[\mathbf{z}_i(\varphi_i^{-1}(w)) - \mathbf{e}(\varphi_i^{-1}(w)) + \boldsymbol{\epsilon}_i(\varphi_i^{-1}(w))]^{\otimes 2}dw + 0_{a.s.}(1) \\
&= \frac{1}{n}\sum_{i=1}^{n}\int_0^t Y_i(s,w)[\mathbf{z}_i(\varphi_i^{-1}(w)) - \mathbf{e}(\varphi_i^{-1}(w))]^{\otimes 2}dw \\
&\quad + \frac{1}{n}\sum_{i=1}^{n}\int_0^t Y_i(s,w)\boldsymbol{\epsilon}_i(\varphi_i^{-1}(w))]^{\otimes 2}dw \\
&\quad + \int_0^t 2E[Y_i(s,w)(\mathbf{z}_i(\varphi_i^{-1}(w)) - \mathbf{e}(\varphi_i^{-1}(w)))]E[\boldsymbol{\epsilon}_i^T(\varphi_i^{-1}(w))]dw + 0_{a.s.}(1) \\
&= \frac{1}{n}\sum_{i=1}^{n}\int_0^t Y_i(s,w)[\mathbf{z}_i(\varphi_i^{-1}(w)) - \mathbf{e}(\varphi_i^{-1}(w))]^{\otimes 2}dw \\
&\quad + \frac{1}{n}\sum_{i=1}^{n}\int_0^t Y_i(s,w)\boldsymbol{\epsilon}_i(\varphi_i^{-1}(w))]^{\otimes 2}dw + 0_{a.s.}(1) \\
&= \frac{1}{n}\sum_{i=1}^{n}\int_0^t Y_i(s,w)[\mathbf{z}_i(\varphi_i^{-1}(w)) - \mathbf{e}(\varphi_i^{-1}(w))]^{\otimes 2}dw \\
&\quad + \int_0^t E[Y_i(s,w)]E[\boldsymbol{\epsilon}_i^{\otimes 2}(\varphi_i^{-1}(w))]dw + 0_{a.s.}(1) \\
&= \frac{1}{n}\sum_{i=1}^{n}\int_0^t Y_i(s,w)[\mathbf{z}_i(\varphi_i^{-1}(w)) - \mathbf{e}(\varphi_i^{-1}(w))]^{\otimes 2}dw \\
&\quad + \int_0^t E[Y_i(s,w)]\boldsymbol{\Xi}dw + 0_{a.s.}(1).
\end{aligned}
$$

Now, we examine the second term Q_2.

$$
\begin{aligned}
Q_2 &= \frac{1}{n}\left[1-\frac{1}{\sum_{j=1}^{n} Y_j(s,w)}\right]\sum_{i=1}^{n} Y_i(s,w)\Xi dw \\
&= \frac{1}{n}\int_0^t \sum_{i=1}^{n}\{Y_i(s,w)\Xi\}\, dw - \frac{1}{n}\Xi\int_0^t \frac{n^{-1}\sum_{i=1}^{n} Y_i(s,w)}{n^{-1}\sum_{i=1}^{n} Y_i(s,w)} dw \\
&= \frac{1}{n}\int_0^t \sum_{i=1}^{n}\{Y_i(s,w)\Xi\}\, dw - \frac{1}{n}\Xi t + 0_{a.s.}(1) \\
&= \frac{1}{n}\int_0^t \sum_{i=1}^{n}\{Y_i(s,w)\Xi\}\, dw + 0_{a.s.}(1) \\
&= \int_0^t E\{Y_i(s,w)\Xi\}\, dw + 0_{a.s.}(1).
\end{aligned}
$$

As a result, $Q_1 - Q_2$ gives

$$
\begin{aligned}
L_n = Q_1 - Q_2 &= \frac{1}{n}\sum_{i=1}^{n}\int_0^t Y_i(s,w)[\mathbf{z}_i(\varphi_i^{-1}(w)) - e(\varphi_i^{-1}(w))]^{\otimes 2} dw \\
&\quad + \int_0^t E[Y_i(s,w)]\Xi dw - \int_0^t E[Y_i(s,w)]\Xi dw + 0_{a.s.}(1) \\
&= E\left[\int_0^t Y_i(s,w)[\mathbf{z}_i(\varphi_i^{-1}(w)) - \mathbf{e}(\varphi_i^{-1}(w))]^{\otimes 2} dw\right] \\
&= B_1 + 0_{a.s.}(1). \ \|
\end{aligned}
$$

As shown in the previous proof, the inverse matrix L_n in $\hat{\boldsymbol{\beta}}_C$ converges almost surely to a positive definite matrix, subject to mild regularity conditions. Hence, the estimator $\hat{\boldsymbol{\beta}}_C$ is devoid of any singularity and instability problems.

Lemma 9 *As* $n \to \infty$, $\bar{\boldsymbol{x}}(\varphi^{-1}(w)) \overset{a.s.}{\to} \boldsymbol{e}(\varphi^{-1}(w))$.

Proof: By regularity condition III. and VI. $\mathbf{x}_i(\varphi^{-1}(w))$ is manageable. Since $Y_i(s,w)$ is a monotonically increasing function, according to the lemma A.2 from Bilias et al. (1997), it is manageable. The preservation result for summations and products of manageable processes hence establishes the fact that $\frac{1}{n}\sum_{i=1}^{n} Y_i(s,w)\mathbf{x}_i(\varphi^{-1}(w))$ is manageable. Under regularity condition V., applying the Strong Uniform Law of Large Numbers, we obtain

$\frac{1}{n}\sum_{i=1}^{n} Y_i(s,w)\mathbf{x}_i(\varphi^{-1}(w)) \stackrel{a.s.}{\rightarrow} E[Y_i(s,w)\mathbf{z}_i(\varphi^{-1}(w))]$ uniformly in w. We already showed that , $\frac{1}{n}\sum_{i=1}^{n} Y_i(s,w) \stackrel{a.s.}{\rightarrow} E[Y_i(s,w)]$ uniformly in w. Then by regularity condition II. and the Strong Uniform Law of Large Numbers, we obtain $\frac{\sum_{j=1}^{n} Y_j(s,w)\mathbf{x}_j(\varphi_j^{-1}(w))}{\sum_{j=1}^{n} Y_j(s,w)} \stackrel{a.s.}{\rightarrow} \mathbf{e}(\varphi^{-1}(w))$. ||

Lemma 10 *As $n \rightarrow \infty$, $K_n \stackrel{a.s.}{\rightarrow} C_1$, where*

$$C_1 = E\left[\int_0^t [z_i(\varphi_i^{-1}(w)) - \boldsymbol{e}(\varphi^{-1}(w))]N_i(s,dw)\right].$$

Proof: We already showed $\frac{1}{n}\sum_{i=1}^{n} N_i(s,w) \stackrel{a.s.}{\rightarrow} E[N_i(s,w)]$ uniformly in w. Also $\mathbf{x}_i(\varphi^{-1}(w)$ is manageable by regularity condition III.. The preservation result for summations and products of manageable processes hence establishes the fact that $n^{-1}\sum_{i=1}^{n}\int_0^t \mathbf{x}_i(\varphi^{-1}(w))N_i(s,dw)$ is manageable. So, under regularity condition V. , by applying the Strong Uniform Law of Large Numbers, we obtain, as $n \rightarrow \infty$, $n^{-1}\sum_{i=1}^{n}\int_0^t \mathbf{x}_i(\varphi^{-1}(w))N_i(s,dw) \stackrel{a.s.}{\rightarrow} \int_0^t E[\mathbf{z}_i(\varphi^{-1}(w))N_i(s,dw)]$. Finally we get

$$\begin{aligned}
K_n &= \frac{1}{n}\sum_{i=1}^{n}\int_0^t [\mathbf{x}_i(\varphi_i^{-1}(w))]N_i(s,dw) - [\bar{\mathbf{x}}(\varphi^{-1}(w))]\int_0^t \frac{1}{n}\sum_{i=1}^{n} N_i(s,dw) \\
&= \int_0^t E[\mathbf{z}_i(\varphi^{-1}(w))N_i(s,dw)] - \int_0^t E[\mathbf{e}(\varphi^{-1}(w))N_i(s,dw)] + 0_{a.s.}(1) \\
&= E\left[\int_0^t [\mathbf{z}_i(\varphi_i^{-1}(w)) - \mathbf{e}(\varphi^{-1}(w))]N_i(s,dw)\right] + 0_{a.s.}(1) \\
&= C_1 + 0_{a.s.}(1). \quad ||
\end{aligned}$$

We state the consistency of $\hat{\boldsymbol{\beta}}_C$ as theorem 9 below.

Theorem 9 *Under regular conditions, as $n \rightarrow \infty$, $\hat{\boldsymbol{\beta}}_C \stackrel{a.s.}{\rightarrow} \boldsymbol{\beta}_0$.*

Proof: By equation (10), lemma 8 and lemma 10, we obtain that

$$\hat{\boldsymbol{\beta}}_C \stackrel{a.s.}{\rightarrow} B_1^{-1}C_1. \tag{11}$$

By equation (9) since we know that $B_1^{-1}C_1 = \boldsymbol{\beta}_0$, we can rewrite (11) as below

$$\hat{\boldsymbol{\beta}}_C \xrightarrow{a.s.} \boldsymbol{\beta}_0. \quad \|$$

6. SIMULATION STUDIES

This section gives a description on the simulation design and discusses the simulation results obtained.

6.1. SIMULATION DESIGN

A simulation study was performed using the R Studio software package to investigate the performance of proposed corrected regression parameter estimators. The specific objectives of this study were: (i) to examine the effect of sample size (n) on the distributional properties of $\hat{\boldsymbol{\beta}}_C$; (ii) to examine the bias and variance of $\hat{\boldsymbol{\beta}}_C$.

Survival Times: We generate survival times t, by solving the equation below:

$$\left(\frac{t}{\theta_2}\right)^{\theta_1} + \boldsymbol{\beta}'\mathbf{x}t + \log U = 0,$$

where

t = survival time

θ_1 = shape parameter of the Weibull distribution

θ_2 = scale parameter of the Weibull distribution

$\boldsymbol{\beta}$ = p dimensional regression parameter vector

$\mathbf{x}$ = p dimensional covariates vector

U = randomly generated value from Uniform$(0, 1)$

Next, we show how we obtained this equation. First, note that

$$\lambda(u) = \frac{f(u)}{1 - F(u)} = -\frac{d}{du}\log[1 - F(u)],$$

and

$$\Lambda(t) = \int_0^t \lambda(u)du = -\log[1 - F(t)]. \tag{12}$$

For a Weibull distribution with shape parameter (θ_1) and scale parameter(θ_2), the cumulative distribution function is given by

$$F(t) = 1 - \exp\left[-\left(\frac{t}{\theta_2}\right)^{\theta_1}\right]. \tag{13}$$

Hence by (12) and (13), we get the expression of baseline cumulative hazard function for a Weibull distribution by

$$\Lambda_0(t) = \left(\frac{t}{\theta_2}\right)^{\theta_1}. \tag{14}$$

Moreover, for additive hazard function,

$$\lambda(u) = \lambda_0(u) + \boldsymbol{\beta}'\mathbf{x}. \tag{15}$$

Therefore,

$$-\log[1 - F(t)] = \Lambda(t) = \int_0^t \lambda(u)du = \int_0^t [\lambda_0(u) + \boldsymbol{\beta}'\mathbf{x}]du = \Lambda_0(t) + \boldsymbol{\beta}'\mathbf{x}t \tag{16}$$

We also know that,

$$1 - F(t) \sim \text{uniform}(0, 1) \tag{17}$$

Hence by (14), (15) , (16) and (17) we obtain,

$$-\log U = \left(\frac{t}{\theta_2}\right)^{\theta_1} + \boldsymbol{\beta}'\mathbf{x}t$$

We use `NLRoot` package in R to find the roots of this equation.

True Parameter Values: In our study, we set $\theta_1 = 1$, $\theta_2 \in \{0.9, 2\}$, $\boldsymbol{\beta} = \{-1, 1\}$, $X_1 \sim uniform(-1, 1)$, and $X_2 \sim Bernoulli(0.5)$.

Censoring: We generate censoring times from an exponential distribution randomly. $C_i \sim exp(\theta)$. In our study, we set $\theta = 0.8$. Next, we calculate times T_i by $T_i = min(t_i, C_i)$. We will also create an indicator variable $\delta_i = I(t_i \geq C_i)$ to indicate if T_i is a survival time or a censoring time.

Error Contaminated Variables: We add a gaussian noise with variance σ^2 to X_1 to create error contaminated version of X_1, say $\tilde{X_1}$.

By this point, we have the knowledge to generate an observation tuple,

$$O_i = \left\{X_{1i}, \tilde{X}_{1i}, X_{2i}, T_i, \delta_i\right\}.$$

Recurrent Event Data: To generate the recurrent events data, we perform the following steps. First, we create a database with 1 million observation tuples, say D. Next, we split this database into two different sub databases based on the value of δ. Let us call the sub-database with $\delta = 0$ as D_c and the sub-database with $\delta = 1$ as D_{nc}. Suppose we need to generate recurrent events data for $n \in \{30, 50, 80\}$ subjects. To do that, we determine how many recurrent events are experienced by each subject j, say K_j, by randomly selecting a number from $\{0, 1, 2, 3, 4, 5, 6\}$. After that, we randomly select K_j number of observations from D_{nc} sub-database followed by one observation tuple from D_c to mimic the recurrent events observed by the jth subject. Once we have generated the recurrent events data for each subject of the study, we use the entire dataset to find naive regression parameter estimates by using `addhazard` or `timereg` packages in R. We create 100 data sets and find these mean of the regresson parameter estimates, which we call the naive estimates $\hat{\beta_1}$ and $\hat{\beta_2}$. We also find the standard deviation of these regression parameter estimates. Similarly, using our proposed equation, for each of the generated data set, we find the corrected regression parameter estimates. Finally, we obtain the mean of the corrected regression parameter estimates, which are denoted by $\hat{\beta}_{1c}$ and $\hat{\beta}_{2c}$. We also find and the standard deviation of these corrected regression parameter estimates. We change the value of σ to take the values $\{0.01, 0.05, 0.1, 0.15, 0.2, 0.25, 0.3, 0.35\}$, to observe the ability of our proposed estimator to handle the error. We refer to the model which ignores the measurement errors as a naive fit and to the one which incorporates the measurement error in estimation via corrected score as the corrected fit.

6.2. SIMULATION RESULTS

Tables 1 and 2 summarize the simulation results. Figures 1 and 2 show the effectiveness of the proposed corrected regression parameter estimates versus the naive regression parameter estimates for increasing and decreasing hazards respectively. Irrespective of the sample size n, as the magnitude of the error variance increases naive estimators tend to be more biased, whereas the corrected estimators remain steady with minimal or no bias. It is also observed that the standard errors decrease with sample size n.

7. APPLICATION

Fuchs et al. (1994) reported a pulmonary exacerbation study which has later been used by many other authors in literature. It was a double-blind randomized multicenter clinical trial designed to evaluate the effect of rhDNase, a recombinant deoxyribonuclease I enzyme, versus placebo on the occurrence of respiratory exacerbations among patients with cystic fibrosis. Six hundred and forty five patients participated in this trial and each patient was followed up for approximately 169 days. Occurrences of all exacerbations were recorded for everyone in this trial. Due to the measurement error, two measurements of forced expiratory volume (FEV) reflecting lung capacity which were taken a few minutes apart were different for each patient. We used a modified version of this dataset to illustrate our proposed methods. We created time varying FEV values for each subject by a Uniform(FEV1-

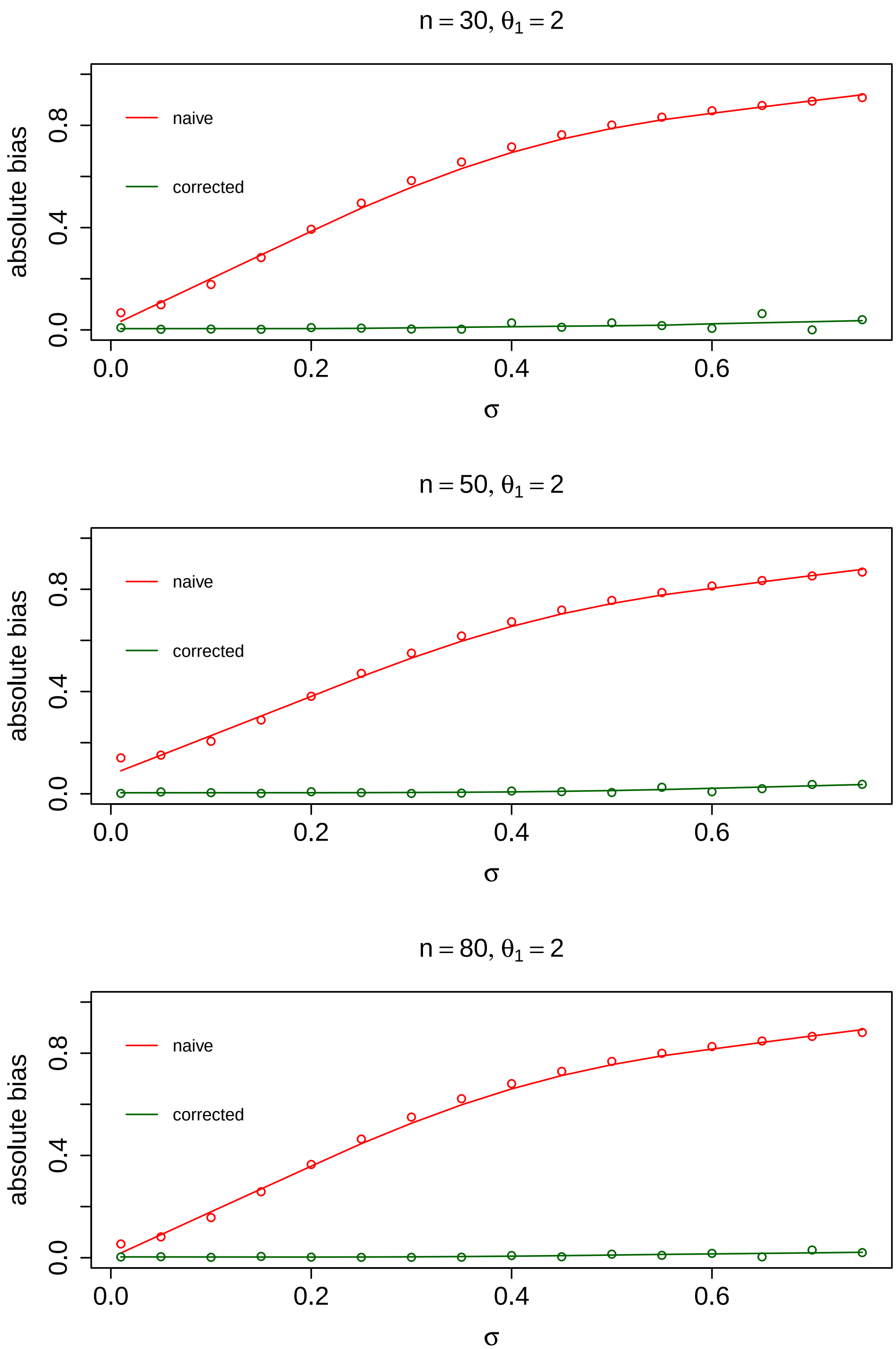

Figure 1. Bias assessment comparing the naive estimation approach to the corrected for an increasing hazard.

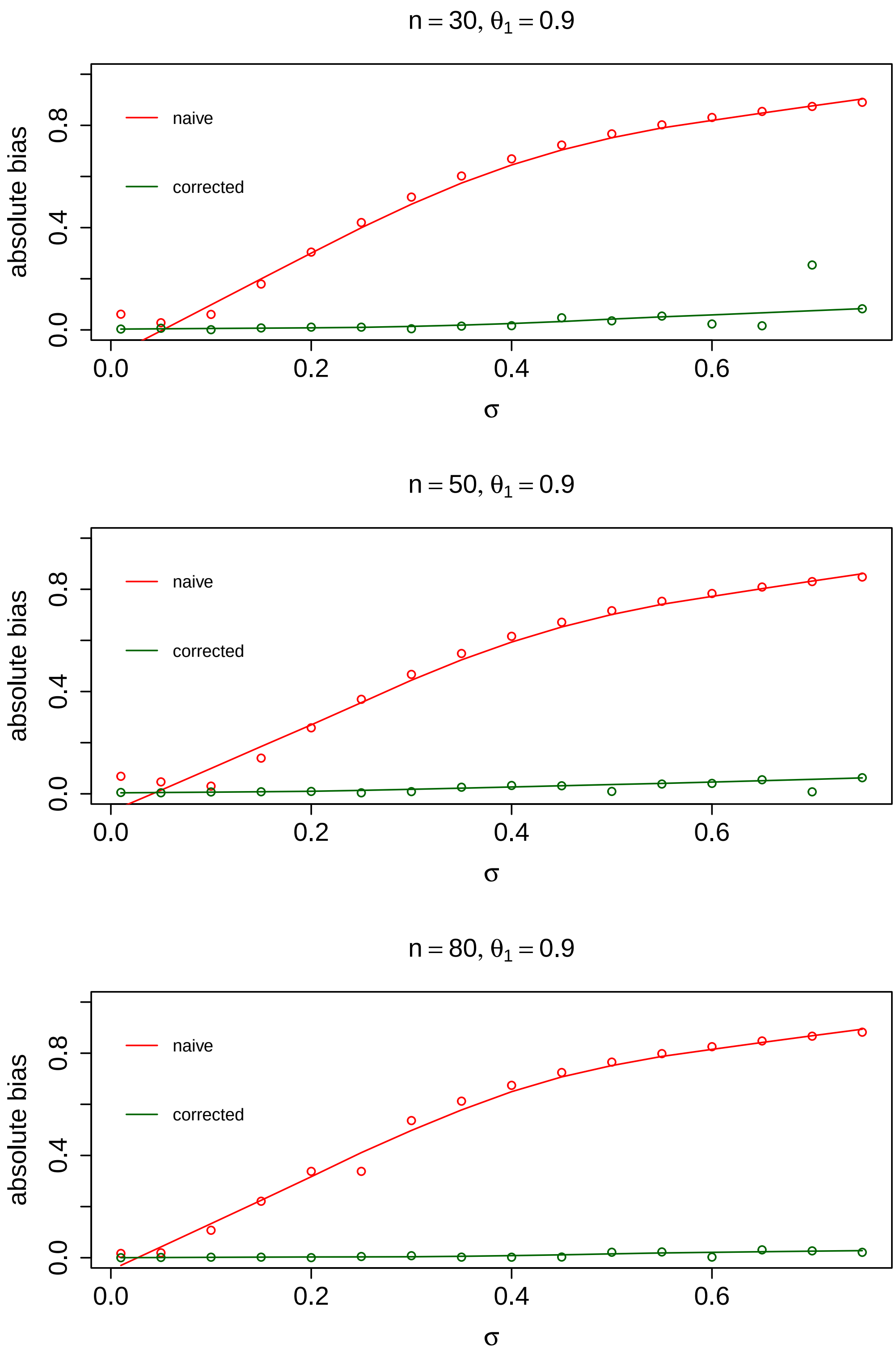

Figure 2. Bias assessment comparing the naive estimation approach to the corrected for a decreasing hazard.

Table 1. Weibull Intensity: $\bar{F}(t;\boldsymbol{\theta}) = e^{-(\theta_2 t)^{\theta_1}}; \theta_2 = 1; \theta_1 = 2; (\beta_1, \beta_2) = (-1, 1)$.

n	σ	$\hat{\beta}_1$	$se(\hat{\beta}_1)$	$\hat{\beta}_{1c}$	$\mathcal{N}(0,\sigma)$ $se(\hat{\beta}_{1c})$	$\hat{\beta}_2$	$se(\hat{\beta}_2)$	$\hat{\beta}_{2c}$	$se(\hat{\beta}_{2c})$
	0.35	-0.3434	0.3460	-1.0027	0.7873	0.9592	0.3495	0.9937	0.2686
	0.3	-0.4161	0.3847	-0.9967	0.6370	0.9597	0.3491	0.9983	0.2546
	0.25	-0.5041	0.4300	-1.0068	0.5447	0.9607	0.3487	0.9994	0.2388
30	0.2	-0.6065	0.4808	-1.0089	0.5045	0.9623	0.3485	1.0040	0.2273
	0.15	-0.7172	0.5326	-0.9977	0.4681	0.9646	0.3487	1.0032	0.2203
	0.1	-0.8225	0.5767	-1.0031	0.4263	0.9677	0.3495	1.0041	0.2158
	0.05	-0.9018	0.6029	-1.0024	0.3911	0.9711	0.3506	0.9957	0.2137
	0.01	-0.9330	0.6071	-0.9913	0.3681	0.9737	0.3515	0.9979	0.2125
	0.35	-0.3827	0.2824	-1.0027	0.5833	0.9527	0.2411	1.0015	0.1732
	0.3	-0.4498	0.3089	-0.9985	0.4784	0.9540	0.2421	0.9996	0.1662
	0.25	-0.5290	0.3381	-1.0044	0.3992	0.9556	0.2430	1.0008	0.1591
50	0.2	-0.6185	0.3686	-0.9919	0.3483	0.9575	0.2438	1.0019	0.1537
	0.15	-0.7116	0.3968	-1.0017	0.3154	0.9596	0.2440	1.0034	0.1482
	0.1	-0.7946	0.4171	-0.9955	0.2900	0.9617	0.2433	0.9972	0.1448
	0.05	-0.8487	0.4238	-1.0073	0.2781	0.9634	0.2414	0.9990	0.1410
	0.01	-0.8594	0.4188	-0.9985	0.2734	0.9642	0.2388	1.0004	0.1386
	0.35	-0.3780	0.1967	-0.9979	0.3879	0.9347	0.1689	1.0022	0.1364
	0.3	-0.4499	0.2110	-0.9984	0.3093	0.9357	0.1690	0.9981	0.1314
	0.25	-0.5360	0.2258	-0.9984	0.2530	0.9369	0.1690	1.0023	0.1313
80	0.2	-0.6352	0.2402	-1.0022	0.2207	0.9383	0.1691	1.0004	0.1294
	0.15	-0.7419	0.2533	-1.0050	0.1994	0.9399	0.1690	0.9988	0.1282
	0.1	-0.8431	0.2644	-0.9985	0.1865	0.9415	0.1686	0.9975	0.1274
	0.05	-0.9185	0.2739	-0.9961	0.1865	0.9427	0.1679	1.0022	0.1280
	0.01	-0.9462	0.2801	-0.9972	0.1851	0.9431	0.1672	1.0013	0.1272

SD,FEV2+SD), where FEV1 and FEV2 are two different FEV measurements of each subject and SD is the standard deviation of all FEV1 and FEV2 values of all subjects together. Next, we added a Gaussian noise to the time-varying FEV measurements, to create a covariate with measurement error. We used $\sigma = 0$ as a reference value, which is the

Table 2. Weibull Intensity: $\bar{F}(t;\boldsymbol{\theta}) = e^{-(\theta_2 t)^{\theta_1}}; \theta_2 = 1; \theta_1 = 0.9; (\beta_1, \beta_2) = (-1, 1)$.

n	σ	$\hat{\beta}_1$	$se(\hat{\beta}_1)$	$\hat{\beta}_{1c}$	$\mathcal{N}(0,\sigma)$ $se(\hat{\beta}_{1c})$	$\hat{\beta}_2$	$se(\hat{\beta}_2)$	$\hat{\beta}_{2c}$	$se(\hat{\beta}_{2c})$
	0.35	-0.3981	0.3914	-0.9853	0.9904	1.0364	0.4388	1.0056	0.3548
	0.3	-0.4806	0.4372	-0.9956	0.7197	1.0343	0.4376	1.0070	0.3208
	0.25	-0.5801	0.4900	-0.9893	0.5613	1.0322	0.4362	0.9994	0.2789
30	0.2	-0.6957	0.5475	-1.0108	0.4670	1.0301	0.4349	0.9981	0.2662
	0.15	-0.8208	0.6030	-0.9924	0.3866	1.0284	0.4340	0.9954	0.2535
	0.1	-0.9395	0.6445	-0.9992	0.3408	1.0277	0.4339	0.9952	0.2466
	0.05	-1.0280	0.6593	-1.0064	0.3259	1.0286	0.4347	0.9969	0.2441
	0.01	-1.0614	0.6503	-0.9968	0.3320	1.0308	0.4355	0.9994	0.2445
	0.35	-0.4512	0.3163	-1.0258	0.7053	1.0442	0.3304	1.0056	0.2624
	0.3	-0.5330	0.3473	-1.0085	0.5820	1.0434	0.3324	0.9965	0.2337
	0.25	-0.6304	0.3817	-0.9961	0.4877	1.0425	0.3342	0.9951	0.2182
50	0.2	-0.7420	0.4181	-1.0093	0.4213	1.0416	0.3355	1.0046	0.2097
	0.15	-0.8605	0.4525	-1.0080	0.3732	1.0407	0.3357	1.0048	0.2006
	0.1	-0.9700	0.4776	-1.0072	0.3317	1.0402	0.3343	0.9964	0.1934
	0.05	-1.0466	0.4864	-1.0039	0.2963	1.0402	0.3306	0.9984	0.1885
	0.01	-1.0685	0.4806	-1.0052	0.2807	1.0409	0.3260	1.0008	0.1858
	0.35	-0.3877	0.2125	-0.9977	0.4251	0.9929	0.2223	0.9995	0.1543
	0.3	-0.4636	0.2299	-0.9922	0.3407	0.9917	0.2226	0.9979	0.1450
	0.25	-0.5553	0.2481	-0.9953	0.2890	0.9903	0.2229	1.0024	0.1396
80	0.2	-0.6623	0.2661	-0.9997	0.2469	0.9886	0.2232	0.9972	0.1293
	0.15	-0.7791	0.2830	-1.0022	0.2178	0.9867	0.2236	1.0003	0.1269
	0.1	-0.8926	0.2991	-0.9980	0.1991	0.9846	0.2239	0.9968	0.1244
	0.05	-0.9808	0.3151	-1.0011	0.1903	0.9829	0.2238	1.0020	0.1284
	0.01	-1.0167	0.3271	-1.0000	0.1907	0.9819	0.2233	1.0007	0.1259

basis for our comparison. By increasing the measurement error variance, we demonstrate how the naive approach would lead us to incorrect inference about an FEV effect, and how this has been corrected through the use of our proposed method.

Table 3. Regression Parameter Estimates.

σ	$\hat{\beta}_1(\times 10^{-}2)$	$\hat{\beta}_2(\times 10^{-}2)$	$\hat{\beta}_{1c}(\times 10^{-}2)$	$\hat{\beta}_{2c}(\times 10^{-}2)$
0.0	-0.5605	-0.0944	-0.5605	-0.0944
0.1	-0.4648	-0.0941	-0.5348	-0.0927
0.15	-0.3881	-0.0949	-0.5166	-0.0920
0.2	-0.3154	-0.0959	-0.4957	-0.0915
0.25	-0.2537	-0.0970	-0.4727	-0.0911
0.3	-0.2044	-0.0979	-0.4482	-0.0908
0.35	-0.1657	-0.0987	-0.4229	-0.0907
0.4	-0.1356	-0.0993	-0.3974	-0.0907

8. CONCLUDING REMARKS

We have proposed a corrected score function when covariates effects are additive and measured with errors. Without the corrections, the estimates in the model are biased and inconsistent. Our results demonstrate that the corrected regression parameter estimators are consistent. All these results are based on the assumption of the additive error model.

REFERENCES

Bilias, Y., Gu, M., and Ying, Z. (1997). Towards a general asymptotic theory for cox model with staggered entry. *The Annals of Statistics*, 25(2):662–682.

Fuchs, H. J., Borowitz, D. S., Christiansen, D. H., Morris, E. M., Nash, M. L., Ramsey, B. W., Rosenstein, B. J., Smith, A. L., and Wohl, M. E. (1994). Effect of aerosolized recombinant human dnase on exacerbations of respiratory symptoms and on pulmonary function in patients with cystic fibrosis. *New England Journal of Medicine*, 331(10):637–642.

Fung, K. Y. and Krewski, D. (1999). On measurement error adjustment methods in poisson regression. *Environmetrics: The official journal of the International Environmetrics Society*, 10(2):213–224.

Guo, J. Q. and Li, T. (2002). Poisson regression models with errors-in-variables: implication and treatment. *Journal of statistical planning and inference*, 104(2):391–401.

Jiang, W., Turnbull, B. W., and Clark, L. C. (1999). Semiparametric regression models for repeated events with random effects and measurement error. *J. Amer. Statist. Assoc.*, 94(445):111–124.

Kim, Y.-J. (2007). Analysis of panel count data with measurement errors in the covariates. *Journal of Statistical computation and Simulation*, 77(2):109–117.

Lin, D. Y., Wei, L.-J., Yang, I., and Ying, Z. (2000). Semiparametric regression for the mean and rate functions of recurrent events. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, 62(4):711–730.

Pollard, D. (1990). *Empirical processes: theory and applications*, volume 2 of *NSF-CBMS Regional Conference Series in Probability and Statistics*. Institute of Mathematical Statistics, Hayward, CA; American Statistical Association, Alexandria, VA.

Sellke, T. (1988). Weak convergence of the Aalen estimator for a censored renewal process. In *Statistical decision theory and related topics, IV, Vol. 2 (West Lafayette, Ind., 1986)*, pages 183–194. Springer, New York.

Stocker, R. and Adekpedjou, A. (2020). A general class of additive semiparametric models for recurrent event data. *Journal of Statistical Planning and Inference*, 205:231–244.

Turnbull, B., J. W. . C. L. (1997). Regression models for recurrent event data: Parametric random effects models with measurement error. *Stat in Med.*, 16:853–864.

Veierød, M. B. and Laake, P. (2001). Exposure misclassification: bias in category specific poisson regression coefficients. *Statistics in medicine*, 20(5):771–784.

Yi, G. Y. and Lawless, J. F. (2012). Likelihood-based and marginal inference methods for recurrent event data with covariate measurement error. *Canad. J. Statist.*, 40(3):530–549.

Yu, H., Cheng, Y.-J., and Wang, C.-Y. (2018). Methods for multivariate recurrent event data with measurement error and informative censoring. *Biometrics*, 74(3):966–976.

Zeger, S. L. and Edelstein, S. L. (1989). Poisson regression with a surrogate x an analysis of vitamin a and indonesian children's mortality. *Journal of the Royal Statistical Society: Series C (Applied Statistics)*, 38(2):309–318.

SECTION

2. CONCLUSION

The initial section of this dissertation outlines a novel approach to obtaining a corrected partial score for recurrent events with one or more covariates that have been measured with errors. This involves using the Cox model to derive corrected regression parameters and cumulative hazards function, as well as discussing a method for estimating measurement error variance and examining its properties. Furthermore, we present the asymptotic properties of the proposed estimators and demonstrate their efficacy through numerical studies. The correction methods proposed are then applied to the rhDNase data. The second part of this dissertation presents a corrected score based on the additive hazard function for recurrent events with error-contaminated covariates. We derive regression parameter estimators and demonstrate their accuracy in estimating their true values through numerical studies. Both parts of this dissertation are based on the assumption that the errors are modeled using the classical additive measurement errors model, which is applicable to numerous real-life scenarios.

REFERENCES

Aalen, O. (1978). Nonparametric inference for a family of counting processes. *The Annals of Statistics*, pages 701–726.

Adekpedjou, A. and Stocker, R. (2015). A general class of semiparametric models for recurrent event data. *J. Statist. Plann. Inference*, 156:48–63.

Alexeeff, S. E., Carroll, R. J., and Coull, B. (2016). Spatial measurement error and correction by spatial SIMEX in linear regression models when using predicted air pollution exposures. *Biostatistics*, 17(2):377–389.

Andersen, P. K., Borgan, O., Gill, R. D., and Keiding, N. (2012). *Statistical models based on counting processes*. Springer Science & Business Media.

Armstrong, B. G. (1990). The effects of measurement errors on relatwe risk regressions. *American Journal of Epidemiology*, 132 (6):1176–1184.

Augustin, T. (2004). An exact corrected log-likelihood function for Cox's proportional hazards model under measurement error and some extensions. *Scand. J. Statist.*, 31(1):43–50.

Beutner, E., Bordes, L., and Doyen, L. (2020). Consistent semiparametric estimators for recurrent event times models with application to virtual age models. *Bernoulli*, 26(1):557–586.

Bickel, P. J. and Freedman, D. A. (1981). Some asymptotic theory for the bootstrap. *Ann. Statist.*, 9(6):1196–1217.

Bilias, Y., Gu, M., and Ying, Z. (1997). Towards a general asymptotic theory for cox model with staggered entry. *The Annals of Statistics*, 25(2):662–682.

Carroll, R. J., Ruppert, D., and Stefanski, L. A. (1995). *Measurement error in nonlinear models*, volume 63 of *Monographs on Statistics and Applied Probability*. Chapman & Hall, London.

Carroll, R. J., Ruppert, D., Stefanski, L. A., and Crainiceanu, C. M. (2006). *Measurement error in nonlinear models: a modern perspective*. Chapman and Hall/CRC.

Chen, L.-P. and Yi, G. Y. (2021). Semiparametric methods for left-truncated and right-censored survival data with covariate measurement error. *Ann. Inst. Statist. Math.*, 73(3):481–517.

Chung, K. L., Williams, R. J., and Williams, R. (1990). *Introduction to stochastic integration*, volume 2. Springer.

Cox, D. (1972a). Regression models and life tables (with discussion). *J. Royal Statist. Soc., B*, 34:187–220.

Cox, D. R. (1972b). Regression models and life-tables. *Journal of the Royal Statistical Society: Series B (Methodological)*, 34(2):187–202.

Cox, D. R. (1975). Partial likelihood. *Biometrika*, 62(2):269–276.

Cox, D. R. and Oakes, D. (2018). *Analysis of survival data*. Chapman and Hall/CRC.

Domowitz, I. and White, H. (1982). Misspecified models with dependent observations. *Journal of Econometrics*, 20(1):35–58.

Efron, B. (1981). Censored data and the bootstrap. *J. Amer. Statist. Assoc.*, 76(374):312–319.

Efron, B. and Tibshirani, R. (1986). Bootstrap methods for standard errors, confidence intervals, and other measures of statistical accuracy (with discussion). *Statistical Science*, 1(1):54–77.

Fleming, T. R. and Harrington, D. P. (2011). *Counting processes and survival analysis*. John Wiley & Sons.

Fuchs, H. J., Borowitz, D. S., Christiansen, D. H., Morris, E. M., Nash, M. L., Ramsey, B. W., Rosenstein, B. J., Smith, A. L., and Wohl, M. E. (1994). Effect of aerosolized recombinant human dnase on exacerbations of respiratory symptoms and on pulmonary function in patients with cystic fibrosis. *New England Journal of Medicine*, 331(10):637–642.

Fuller, W. A. (1987). *Measurement error models*. Wiley Series in Probability and Mathematical Statistics: Probability and Mathematical Statistics. John Wiley & Sons, Inc., New York.

Fung, K. Y. and Krewski, D. (1999). On measurement error adjustment methods in poisson regression. *Environmetrics: The official journal of the International Environmetrics Society*, 10(2):213–224.

Guo, J. Q. and Li, T. (2002). Poisson regression models with errors-in-variables: implication and treatment. *Journal of statistical planning and inference*, 104(2):391–401.

Hjort, N. L. (1985). Bootstrapping cox's regression model. *Technical Report 241, Department of Statistics, Stanford University*.

Hu, C. and Lin, D. Y. (2002). Cox regression with covariate measurement error. *Scand. J. Statist.*, 29(4):637–655.

Hu, C. and Lin, D. Y. (2004). Semiparametric failure time regression with replicates of mismeasured covariates. *J. Amer. Statist. Assoc.*, 99(465):105–118.

Hu, P., Tsiatis, A. A., and Davidian, M. (1998). Estimating the parameters in the Cox model when covariate variables are measured with error. *Biometrics*, 54(4):1407–1419.

Huque, M. H., Bondell, H. D., Carroll, R. J., and Ryan, L. M. (2016). Spatial regression with covariate measurement error: a semiparametric approach. *Biometrics*, 72(3):678–686.

Jacod, J. (1975). Multivariate point processes: predictable projection, radon-nikodym derivatives, representation of martingales. *Zeitschrift für Wahrscheinlichkeitstheorie und verwandte Gebiete*, 31(3):235–253.

Jennrich, R. I. (1969). Asymptotic properties of non-linear least squares estimators. *Ann. Math. Statist.*, 40:633–643.

Jiang, W., Turnbull, B. W., and Clark, L. C. (1999a). Semiparametric regression models for repeated events with random effects and measurement error. *Journal of the American Statistical Association*, 94(445):111–124.

Jiang, W., Turnbull, B. W., and Clark, L. C. (1999b). Semiparametric regression models for repeated events with random effects and measurement error. *J. Amer. Statist. Assoc.*, 94(445):111–124.

Kalbfleisch, J. D. and Prentice, R. L. (2011). *The statistical analysis of failure time data*. John Wiley & Sons.

Kim, Y.-J. (2007). Analysis of panel count data with measurement errors in the covariates. *Journal of Statistical computation and Simulation*, 77(2):109–117.

Kong, F. H. and Gu, M. (1999). Consistent estimation in Cox proportional hazards model with covariate measurement errors. *Statist. Sinica*, 9(4):953–969.

Kosorok, M. R. (2008). *Introduction to empirical processes and semiparametric inference*. Springer Series in Statistics. Springer, New York.

Kullback, S. and Leibler, R. A. (1951). On information and sufficiency. *Ann. Math. Statistics*, 22:79–86.

Lin, D. and Ying, Z. (1995). Semiparametric analysis of general additive-multiplicative hazard models for counting processes. *The annals of Statistics*, pages 1712–1734.

Lin, D. Y. and Ying, Z. (1994). Semiparametric analysis of the additive risk model. *Biometrika*, 81(1):61–71.

Nakamura, T. (1990). Corrected score function for errors-in-variables models: methodology and application to generalized linear models. *Biometrika*, 77(1):127–137.

Nakamura, T. (1992). Proportional hazards model with covariates subject to measurement error. *Biometrics*, 48(3):829–838.

Peña, E. A., Slate, E. H., and González, J. R. (2007a). Semiparametric inference for a general class of models for recurrent events. *Journal of Statistical Planning and Inference*, 137(6):1727–1747.

Peña, E. A., Slate, E. H., and Gonzalez, J. R. (2007b). Semiparametric inference for a general class of models for recurrent events. *Journal of Statistical Planning and Inference*, 137:1727–1747.

Pena, E. A., Strawderman, R. L., and Hollander, M. (2000). A weak convergence result relevant in recurrent and renewal models. *Recent Advances in Reliability Theory: Methodology, Practice, and Inference*, pages 493–514.

Pena, E. A., Strawderman, R. L., and Hollander, M. (2001). Nonparametric estimation with recurrent event data. *Journal of the American Statistical Association*, 96(456):1299–1315.

Pollard, D. (1990). *Empirical processes: theory and applications*, volume 2 of *NSF-CBMS Regional Conference Series in Probability and Statistics*. Institute of Mathematical Statistics, Hayward, CA; American Statistical Association, Alexandria, VA.

Prentice, R. L. (1982). Covariate measurement errors and parameter estimation in a failure time regression model. *Biometrika*, 69(2):331–342.

Song, X. and Huang, Y. (2005). On corrected score approach for proportional hazards model with covariate measurement error. *Biometrics*, 61(3):702–714.

Stefanski, L. A. (1985). The effects of measurement error on parameter estimation. *Biometrika*, 72(3):583–592.

Stefanski, L. A. and Cook, J. R. (1995). Simulation-extrapolation: the measurement error jackknife. *J. Amer. Statist. Assoc.*, 90(432):1247–1256.

Stocker, R. and Adekpedjou, A. (2020). A general class of additive semiparametric models for recurrent event data. *Journal of Statistical Planning and Inference*, 205:231–244.

Turnbull, B., J. W. . C. L. (1997). Regression models for recurrent event data: Parametric random effects models with measurement error. *Stat in Med.*, 16:853–864.

van der Vaart, A. W. (1998). *Asymptotic statistics*, volume 3 of *Cambridge Series in Statistical and Probabilistic Mathematics*. Cambridge University Press, Cambridge.

Veierød, M. B. and Laake, P. (2001). Exposure misclassification: bias in category specific poisson regression coefficients. *Statistics in medicine*, 20(5):771–784.

Wang, C. Y., Hsu, L., Feng, Z. D., and Prentice, R. L. (1997). Regression calibration in failure time regression. *Biometrics*, 53(1):131–145.

White, H. and Domowitz, I. (1984). Nonlinear regression with dependent observations. *Econometrica*, 52(1):pp. 143–162.

Yan, Y. and Yi, G. Y. (2015). A corrected profile likelihood method for survival data with covariate measurement error under the Cox model. *Canad. J. Statist.*, 43(3):454–480.

Yi, G. Y. (2017). *Statistical analysis with measurement error or misclassification*. Springer Series in Statistics. Springer, New York. Strategy, method and application, With a foreword by Raymond J. Carroll.

Yi, G. Y. and Lawless, J. F. (2012). Likelihood-based and marginal inference methods for recurrent event data with covariate measurement error. *Canad. J. Statist.*, 40(3):530–549.

Yu, H., Cheng, Y.-J., and Wang, C.-Y. (2016). Semiparametric regression estimation for recurrent event data with errors in covariates under informative censoring. *Int. J. Biostat.*, 12(2):20160001, 18.

Yu, H., Cheng, Y.-J., and Wang, C.-Y. (2018). Methods for multivariate recurrent event data with measurement error and informative censoring. *Biometrics*, 74(3):966–976.

Yu, M. and Nan, B. (2010). Regression calibration in semiparametric accelerated failure time models. *Biometrics*, 66(2):405–414.

Zamba, K. D. and Adekpedjou, A. (2019). A khmaladze-transformed test of fit with ml estimation in the presence of recurrent events. *Sequential Analysis*, 38(3):318–341.

Zeger, S. L. and Edelstein, S. L. (1989). Poisson regression with a surrogate x an analysis of vitamin a and indonesian children's mortality. *Journal of the Royal Statistical Society: Series C (Applied Statistics)*, 38(2):309–318.

VITA

Alahakoon Mudiyanselage, Ravinath Sasanka Bandara Alahakoon attended Dharmaraja Colleage, Kandy, Sri Lanka. In 2014, he received his bachelor's degree specializing in Statistics from University of Peradeniya, Sri Lanka. In May 2020, Ravinath received a Master of Science in Applied Mathematics with Statistics emphasis from Missouri University of Science and Technology. He successfully defended his dissertation in May 2023 and received his Doctor of Philosophy in Applied Mathematics with Statistics emphasis from Missouri University of Science and Technology in July 2023.