9-6-2022 3:00 PM

Regression-based Methods for Dynamic Treatment Regimes with Mismeasured Covariates or Misclassified Response

Dan Liu, *The University of Western Ontario*

Supervisor: He, Wenqing, *The University of Western Ontario*

A thesis submitted in partial fulfillment of the requirements for the Doctor of Philosophy degree in Statistics and Actuarial Sciences

Abstract

The statistical study of dynamic treatment regimes (DTRs) focuses on estimating sequential treatment decision rules tailored to patient-level information across multiple stages of intervention. Regression-based methods in DTR have been studied in the literature with a critical assumption that all the observed variables are precisely measured. However, this assumption is often violated in many applications. One example is the STAR*D study, in which the patient's depressive score is subject to measurement error. In this thesis, we explore problems in the context of DTR with measurement error or misclassification considered in the observed data.

The first project deals with covariate measurement error in Q-learning with continuous outcomes. The true covariate is not observable, but its replicate measurements are available in each stage. We propose a modified Q-learning algorithm with regression calibration to handle the measurement error. Given the replicate measurements, the proposed method obtains and uses the estimates of the unobserved true covariate in each stage of Q-learning.

The second project explores covariate measurement error in dynamic weighted survival modeling (DWSurv), a regression-based method dealing with survival outcomes in DTR. Internal validation data are assumed to be available with true covariates only observed in a subset of the data. Two correction methods are proposed to eliminate the effect of mismeasured covariate by obtaining the estimates of the missing true covariate in each stage of DWSurv. The consistency of the proposed estimator is established.

The third project examines Q-learning with binary outcomes being subject to misclassification. We investigate the outcome misclassification effect for internal validation data and develop a correction method to adjust for the effect in Q-learning. A probability relationship is established between the true outcome and the misclassified outcome. The estimation procedure in Q-learning is modified by including the derived probability relationship in the proposed method.

Extensive simulation studies are conducted to assess the performance of the proposed methods and to compare them with the naive method. Real data are analyzed for illustration of the proposed methods. The results showcase the importance of incorporating the errors in DTR and the competency of the proposed methods in obtaining the optimal DTR.

Keywords: Precision medicine, dynamic treatment regimes, regression-based methods, Q-learning, dynamic weighted survival modeling, measurement error, misclassification, regression calibration, *k*-nearest neighbors, weighted least squares, maximum likelihood estimation

Summary for Lay Audience

Precision medicine is a new approach that recommends individualized treatment to a patient by taking the patient's information into account. It differs from the traditional 'one-size-fits-all' clinical strategy, which ignores the patient's heterogeneity in response to the treatment. Dynamic treatment regimes (DTRs) realize this process by providing sequential treatment decisions. However, in practice, a patient's information that is used to infer a treatment decision often contains error-corrupted covariates or misclassified outcomes, which can be viewed as incorrect records of the patient's characteristics or mislabeled clinical outcomes of the patient. The contaminated information may misrepresent the health status of the patients and further lead to inaccurate treatment decision-making. In this thesis, three situations are investigated in the context of DTR with error-corrupted covariates or misclassified outcomes.

The first study focuses on the problem of error-corrupted covariates in a DTR method with continuous outcomes, provided that the true covariate is not observed, but only its repeated measurements are available. The regression calibration method is employed to correct the error by using a new variable for the error-corrupted covariates, which are obtained from the available repeated measurements in the data.

The second study deals with the error-corrupted covariates in a survival-based DTR, given that the true covariate is partially observed in the data. Two correction methods are developed to correct the error-corrupted covariates. The proposed methods create estimates for the unobserved true covariate using the available error-corrupted covariate and use the created values for modeling.

The third study addresses the misclassified outcome problem in a DTR method with binary outcomes, assuming that the true outcome is only observed in a subset of data. A likelihood-based approach is proposed, which incorporates the relationship between the true outcome and misclassified outcome, through which the outcome misclassification can be corrected.

For each topic, simulation studies have demonstrated significant improvements in error correction and treatment decision-making. Real data applications have also shown the importance of including the errors in the DTR context.

Acknowledgments

First, I would like to express my gratitude to my advisor Dr. Wenqing He. This dissertation would not have been possible without his support and guidance. From Dr. He, I have learned to be attentive to details and have scientific thinking. He allowed me to have flexibility in the process of doing research and provided me with insightful advice when I was directionless. I would also like to thank Dr. Hao Yu, Dr. Hyukjun Gweon, Dr. Daniel Lizotte, and Dr. Michael Wallace for reviewing my thesis and providing invaluable comments.

Secondly, I want to thank Dr. Grace Y. Yi and all the members of the GW-DSRG (Grace-Wenqing Data Science Research Group). Dr. Yi's deep knowledge, passion, and rigorous attitude toward research have influenced and encouraged me throughout my Ph.D. study. The discussions in the regular group meeting have enriched my knowledge in many research areas. Through the team, I also came to befriend many talented and diligent peers Jingyu Cui, Yasin Khadem Charvadeh, Chengqian Xian, Kexin Luo, Li-Pang Chen, Qihuang Zhang, Junhan Fang, etc. I would like to thank Chi-Kuang Yeh for his friendship over this journey. I would also like to acknowledge my friend Dr. Sarah Zheng, who generously shared the MIMIC-III data.

Personally, I want to express my deep gratitude to my parents for their unconditional love and sacrifice. They have encouraged and supported me unwaveringly in my academic journey. I would not be who I am today without my parents. I would also like to thank my husband for his dedicated love and encouragement through this journey. He has been very supportive and provides help when I need him the most.

I would like to thank my brothers and sisters in Christ for their prayers and love in carrying me over the years. Last but not least, I give my thanksgiving and all the glory to the everlasting God, my heavenly Father, for His love, grace, and faithfulness.

Contents

List of Figures

List of Tables

Chapter 1

Introduction

It is a long history that clinicians have been using a 'one-size-fits-all' strategy to treat patients with chronic diseases over a multi-stage period. However, in practice, this treatment strategy faces many challenges. As the disease and patients' health conditions are constantly changing over time, they may not respond to the treatments that used to be effective. It implies a new treatment should be adopted. The financial concerns for both patients and hospitals may be incurred when a costly but ineffective treatment has been constantly applied over multiple stages. Moreover, the best treatment regarded at a given stage may yet lead to a suboptimal clinical outcome in the long run (Chakraborty, 2013; Chakraborty and Murphy, 2014).

To tackle these challenges, the study of precision medicine has begun to arise, with the objective of searching for optimal dynamic treatment regimes (DTRs) for patients. A dynamic treatment regime is a sequence of treatment decision rules, one per stage of intervention, recommended to a patient by taking the individual's characteristics and treatment history into account. An optimal DTR is a sequence of treatment decision rules that yields optimal treatments, with which the long-term clinical outcome is optimized. In recent years, there have been a variety of case studies associated with the estimation of optimal DTR. However, the observed data in studies are often assumed to be measured error-free, which may be violated in practice. The following examples show that variables with measurement error often exist.

The first example is the Sequenced Treatment Alternatives to Relieve Depression (STAR*D) study, designed as a multisite, multistage randomized controlled trial. The STAR*D study aimed to evaluate the effect of treatments for patients who suffered from a major depressive

disorder (Rush et al., 2003,0). The entire study possessed four levels, and in each level, one or a combination of treatments was assigned to the patients. The severity of depressive disorder was measured by the Quick Inventory of Depressive Symptomatology (QIDS) score. If a patient who received treatment did not meet the requirement of remission (QIDS ≤ 5) at the end of the level, this patient would have entered the next level with a different treatment assigned. The QIDS score is used to derive the sequential optimal treatment rules for each patient, but in the trial, the QIDS score was reported by both clinicians and patients. Due to unavoidable human errors, these reported scores may be subject to measurement error. Thus, the estimated optimal DTR may be problematic.

The second example comes from the Medical Information Mart for Intensive Care-III (MIMIC-III) Clinical Database, comprising large-scale observational admission data collected at Beth Israel Deaconess Medical Center from 2001 to 2012 (Johnson et al., 2016,0). The MIMIC-III dataset was used to study the association between the use of transthoracic echocardiography (TTEC) and mortality, conditional on the intensive care unit (ICU) patients' characteristics and lab test results (Feng et al., 2018; Chen et al., 2021). However, in the MIMIC-III data, variables such as positive end-expiratory pressure are significantly associated with TTEC but suffer considerable missingness. Without accounting for such missingness, the conclusions may be misleading.

Another example is the National Health and Nutrition Examination Survey Data I Epidemiologic Follow-up Study (NHEFS), a national longitudinal study. It aimed to investigate the relationships between clinical, nutritional, and behavioral factors and subsequent morbidity, mortality, and operational factors with hospital utilization. The cohort NHEFS dataset contains cigarette smokers with baseline variables collected from 1971 to 1975. They were later followed up through personal interviews in 1982, in which their smoking status was collected. Apparently, the answers from the cigarette smokers without confirmation are subject to misclassification as they may not report the truth. A similar case is also found in a smoking cessation program, which examined the effectiveness of a perioperative smoking cessation intervention (Lee et al., 2013). The patients reported their status of quitting smoking, with lab tests for confirmation. It has been discovered that 11 out of the 146 patients misreported their smoking status in this program.

Although these four examples were initiated with different objectives, they share a common fact that the variables in the collected data might be contaminated with either measurement error or misclassification. As the primary goal of DTR is to identify sequential optimal treatment rules, the resulting optimal DTR may be altered if the error in the observed data is taken into account. Such considerations motivate us to investigate the measurement error and misclassification effects in the DTR context and develop correction methods to eliminate the effects.

1.1 Dynamic Treatment Regimes

Identifying an optimal DTR depends on the statistical approaches, given the structure of available data and research questions. The common approaches for estimating the optimal DTR can be classified into two categories, regression-based and classification-based methods.

Regression-based methods also referred as indirect methods. They model and estimate the conditionally expected outcome to yield an optimal DTR that maximizes the expected outcome. The classical regression-based methods that are widely studied in the literature include Q-learning (Watkins, 1989; Chakraborty and Murphy, 2014), G-estimation (Robins, 2004), A-learning (Murphy, 2003; Schulte et al., 2014), and regret-regression (Henderson et al., 2010). However, these methods assumed continuous outcomes in their approaches. Moodie et al. (2014) made an attempt to extend the Q-learning to binary outcomes and count outcomes. Ghosh and Chakraborty (2018) proposed a likelihood-based approach to estimate and compare two embedded DTR with ordinal outcomes in a two-stage sequential multiple assignment randomized trials. The Bayesian approach was proposed to estimate the optimal embedded DTR with binary outcomes by Artman et al. (2020). The estimation of DTR with survival outcomes is considered with the accelerated failure time (AFT) model in the Q-learning framework (Goldberg and Kosorok, 2012; Huang and Ning, 2012; Huang et al., 2014). Although statistical methods in the G-estimation framework have been proposed (Robins and Greenland, 1994; Hernán and Robins, 2020), they are not widely used because of the complexity in theory and implementation due to the nature of G-estimation. Murray et al. (2018) and Klausch et al. (2018) also developed Bayesian-based approaches to estimate the optimal treatment regimes.

Regression-based methods enjoy the advantages of being built on regression models and

easily implemented. However, they suffer from a strict assumption of the correct specification of the outcome model to yield consistent estimates of parameters (Chakraborty, 2013). To overcome their weaknesses, Wallace and Moodie (2015) proposed a dynamic weighted ordinary least squares (dWOLS), a doubly robust estimation method that integrates the implementation simplicity of Q-learning and double robustness property of G-estimation. Simoneau et al. (2020b) further developed dynamic weighted survival modeling (DWSurv), a doubly robust regression-based method to deal with DTR with survival outcomes. Xiao et al. (2019) examined loss-based robust regression estimators to accommodate baseline function misspecification and skewed, heterogeneous, heavy-tailed errors or outliers.

In contrast, instead of requiring a specification of the outcome model beforehand, classification-based methods, also referred to as direct methods or value search methods, directly estimate the marginal mean outcome of a regime and identify an optimal DTR that maximizes the estimated value over all possible DTRs (Laber et al., 2014). Some popular classification-based methods utilize the inverse probability weighting (IPW) method to estimate the marginal mean of outcome in DTR, but they are sensitive to the misspecification of the propensity score (Robins, 2000; van der Laan, 2006; van der Laan and Petersen, 2007). Zhang et al. (2012) and Zhang et al. (2013) proposed a doubly robust method by introducing an augmented IPW estimator. Marginal structural mean models were studied to construct DTRs (Robins et al., 2008; Orellana et al., 2010).

Machine learning techniques are also introduced to make a class prediction to find the optimal DTR across stages. For instance, Laber and Zhao (2015) introduced decision trees as a new estimation method to obtain an optimal regime, and the intuitive value-based classification meaning makes the resulting DTR more interpretable. Zhao et al. (2012) and Zhao et al. (2015a) proposed outcome weighted learning (OWL) by borrowing the idea of the support vector machine to redefine the DTR problem into a weighted classification problem. Zhou et al. (2017) extended the OWL into a more generalized version, residual weighted learning (RWL), to include variable selection and different classes of the outcome. Liu et al. (2018) developed an augmented outcome weighted learning (AOL) that combines OWL and regression models to estimate an optimal DTR. Fu et al. (2019) modified the loss function of the OWL to be bounded and proposed a robust outcome weighted learning (ROWL), by which more stable

optimal treatment rules were produced. For censored data, Zhao et al. (2015b) proposed a doubly robust estimator for expected survival time and utilized outcome-weighted learning to estimate sequential optimal treatment rules. Methods have also been developed using survival probability as the outcome of interest (Bai et al., 2017; Jiang et al., 2017; Xue et al., 2022).

In this thesis, we will focus on the study of regression-based methods with the covariate subject to measurement error or binary outcome with misclassification. Q-learning with continuous outcomes, Q-learning with binary outcomes, and dynamic weighted survival modeling are explored.

1.1.1 Notations and Concepts

Before describing the methodology for DTR, we introduce some basic notations in the DTR framework. Based on the outcome type, the data for DTR can be categorized into uncensored data and censored data.

DTR with Uncensored Data

Let the uppercase letters represent random variables and lower-class letters represent the realization of the random variables. A DTR data trajectory across a maximum of J stages follows $\{\boldsymbol{X_1}, A_1, Y_1, \boldsymbol{X_2}, A_2, Y_2, ..., \boldsymbol{X_J}, A_J, Y_J\}$, where $\boldsymbol{X_1}$ is the baseline covariate vector, measured at the beginning of stage 1 before initial treatment is applied and $\boldsymbol{X_j}$ represents the updated information about the patient, collected at the beginning of stage j (j = 2, ..., J). A_j denotes a binary treatment action taken at the beginning of stage j, where $A_j = 1$ if the patient received a treatment, and $A_j = 0$ otherwise. A patient's history $\boldsymbol{H_j}$ with values taken as $\boldsymbol{h_j}$ is defined as the accumulative information collected up to j^{th} stage before making treatment decision A_j: $\boldsymbol{H_j} = (\boldsymbol{X_1}, A_1, \boldsymbol{X_2}, ..., \boldsymbol{X_j})$. In the data trajectory, Y_j is the outcome observed at the end of stage j, as a reward subsequent to the treatment A_j. The outcome Y_j can be of any type, such as continuous outcome, survival time or discrete-valued outcome. In some circumstances, only a single terminal outcome Y is observed at the end of the last stage. In a two-stage setting, it can be viewed as a special case that $Y_1 \equiv 0$ and $Y_2 = Y$. A DTR $\boldsymbol{a}$ is defined as a sequence of treatment decision rules such that $\boldsymbol{a} = \{a_1, a_2, ..., a_J\}$, where $a_j = a_j(\boldsymbol{h_j})$ is the treatment assigned

at stage j. An optimal DTR denoted as $\boldsymbol{a^{opt}}$ is a sequence of treatment rules that maximizes the conditional mean outcome Y (or mean sum of Y_j's), where $\boldsymbol{a^{opt}} = \{a_1^{opt}, a_2^{opt}, ..., a_J^{opt}\}$ and $a_j^{opt} = a_j^{opt}(\boldsymbol{h_j})$ is the optimal treatment at stage j.

In DTR with observational studies, an important concept is the treatment model, which is defined as the probability of assigning treatment a_j conditional on patients' history $\boldsymbol{h_j}$, denoted as $\pi(\boldsymbol{h_j}) = P(A_j = a_j|\boldsymbol{H_j} = \boldsymbol{h_j})$. The treatment model is often used in the statistical approaches to remove the confounding treatment effects on parameter estimation so that unbiased estimates of the treatment effect can be obtained (Austin, 2011; Moodie et al., 2012; Chakraborty, 2013; Tsiatis, 2019).

DTR with Censored Data

For censored data, DTR with survival outcomes follows a data trajectory across a maximum of J stages $\{\eta_1, \boldsymbol{X_1}, A_1, Y_1, \eta_2, \boldsymbol{X_2}, A_2, Y_2, ..., \eta_J, \boldsymbol{X_J}, A_J, Y_J, \Delta\}$, where η_j is an indicator of whether the individuals entered the j^{th} stage for treatment ($\eta_j = 1$) or not ($\eta_j = 0$). X_j is the covariates collected at the beginning of stage j. $A_j \in \{1, 0\}$ is the binary treatment received at stage j. Let T_j and T be the stage-j survival time and the overall survival time across all the J stages with $T = \sum_{j=1}^{J} \eta_j T_j$. C_j is the stage-j censoring time with C being the sum of the censoring times $C = \sum_{j=1}^{J} \eta_j C_j$. Let Δ be a censoring indicator such that $\Delta = \mathbb{1}(T \leq C)$ and δ be the realization of Δ. The observed outcome Y_j is defined as the minimum of the survival time and censoring time at stage j, $Y_j = \min(T_j, C_j)$. The history $\boldsymbol{H_j}$ with values taken as $\boldsymbol{h_j}$ is a collection of all the covariates and the treatments prior to the time of making the j^{th} treatment decision A_j: $\boldsymbol{H_j} = (\boldsymbol{X_1}, A_1, \boldsymbol{X_2}, ..., \boldsymbol{X_j})$. Then, we can obtain a sequence of treatment decision rules up to J stage $\boldsymbol{a} = \{a_1, a_2, ..., a_J\}$, where $a_j = a_j(\boldsymbol{h_j})$ is the treatment assigned at stage j. An optimal DTR $\boldsymbol{a^{opt}} = \{a_1^{opt}, a_2^{opt}, ..., a_J^{opt}\}$, which is a sequence of treatment rules that maximizes the expected overall survival time T, where $a_j^{opt} = a_j^{opt}(\boldsymbol{h_j})$ is the optimal treatment received at stage j.

In DTR with survival data, a treatment model is defined as the probability of receiving treatment a_j conditional on a history of those who entered the j^{th} stage, denoted as $\pi(\boldsymbol{h_j}) = P(A_j = a_j|\boldsymbol{H_j} = \boldsymbol{h_j}, \eta_j = 1)$. A censoring model is also defined for those who entered the j^{th} stage.

It models the probability of experiencing the event of interest conditional on patients' history and treatment, denoted as $g(a_j, \boldsymbol{h_j}) = P(\Delta = 1|\boldsymbol{H_j} = \boldsymbol{h_j}, A_j = a_j, \eta_j = 1)$.

Basic Assumptions

The methodology in DTR is established on the following two assumptions (Chakraborty, 2013):

(A1) *Stable unit treatment value*: an individual's outcome is not influenced by other individuals' treatment allocation.

(A2) *No unmeasured confounders*: for any possible treatment rule, treatment A_j received in the j^{th} stage is independent of any future (potential) covariate or outcome, $\{X_{j+1}, Y_{j+1}, ..., X_J, Y_J\}$, conditional on the history $\boldsymbol{H_j}$.

The first assumption ensures that the patients in the study are independent of each other regarding the treatment effects. The second assumption allows for no future covariate or outcome to affect the current treatment decision-making.

1.1.2 Q-learning with Continuous Outcomes

Q-learning originates from reinforcement learning and has become one of the most popular regression-based methods to estimate an optimal DTR (Watkins, 1989; Chakraborty and Murphy, 2014). The Q-learning is modeled by stage-specific Q-functions, which measure the expected future reward conditional on the history of a patient's characteristics and treatment action. For a two-stage DTR, the Q-functions are defined as

$$Q_2(\boldsymbol{H_2}, A_2) = E[Y|\boldsymbol{H_2}, A_2],$$
$$Q_1(\boldsymbol{H_1}, A_1) = E[\max_{a_2} Q_2(\boldsymbol{H_2}, a_2)|\boldsymbol{H_1}, A_1].$$

Since Q-functions are usually unknown, they need to be estimated from the data using a backward recursive procedure (Chakraborty, 2013). At stage 2, the Q-function $Q_2(\boldsymbol{H_2}, A_2)$ is the expected terminal outcome Y conditional on the history $\boldsymbol{H_2}$ and treatment A_2. Having worked backward recursively, the first stage $Q_1(\boldsymbol{H_1}, A_1)$ is modeled with a pseudo-outcome

$\widetilde{Y}_1$ constructed as $\max_{a_2} Q_2(\boldsymbol{H_2}, a_2)$, which would be the future reward had the patients received the second stage optimal treatment a_2^{opt}. By using $\widetilde{Y}_1$, it allows the first stage treatment effect comparison to be reasonable.

To estimate the Q-functions, a common statistical approach is to parameterize $\mathrm{Q}_j(\boldsymbol{H_j}, A_j)$ at stage j via regression models

$$Q_j(\boldsymbol{H_j}, A_j; \boldsymbol{\beta_j}, \boldsymbol{\psi_j}) = f(\boldsymbol{H_{j0}}; \boldsymbol{\beta_j}) + g(\boldsymbol{H_{j1}}, A_j; \boldsymbol{\psi_j}), \tag{1.1}$$

where the treatment-free component $\mathrm{f}(\boldsymbol{H_{j0}}; \boldsymbol{\beta_j})$ is a function of $\boldsymbol{H_{j0}}$, a subset of history vector $\boldsymbol{H_j}$ without regard to A_j, and the blip component $\mathrm{g}(\boldsymbol{H_{j1}}, A_j; \boldsymbol{\psi_j})$ is a function of A_j and $\boldsymbol{H_{j1}}$, a different subset of history vector $\boldsymbol{H_j}$. The covariates collected in $\boldsymbol{H_{j1}}$ are called tailoring variables. The functions $\mathrm{f}(\cdot)$ and $\mathrm{g}(\cdot)$ can be specified in any form, such as splines, neural network and regression trees (Chakraborty, 2013). The simplest case might be modeling the Q-functions linearly as

$$Q_j(\boldsymbol{H_j}, A_j; \boldsymbol{\beta_j}, \boldsymbol{\psi_j}) = \boldsymbol{\beta_j^T H_{j0}} + A_j\left(\boldsymbol{\psi_j^T H_{j1}}\right). \tag{1.2}$$

In the set of parameters $(\boldsymbol{\beta_j}, \boldsymbol{\psi_j})$ in the linear setting (1.2), we are generally interested in estimating the blip parameter $\boldsymbol{\psi_j}$ since it contains both the effect of treatment A_j and the interaction between treatment and tailoring variables in $\boldsymbol{H_{j1}}$, by which the optimal DTR is determined. $\hat{a}_j^{opt}$, the estimated optimal treatment at stage j, is the treatment that maximizes $\mathrm{Q}_j(\boldsymbol{h_j}, a_j; \hat{\boldsymbol{\beta}}_{\boldsymbol{j}}, \hat{\boldsymbol{\psi}}_{\boldsymbol{j}})$. That is,

$$\hat{a}_j^{opt} = \arg\max_{a_j} Q_j(\boldsymbol{h_j}, a_j; \hat{\boldsymbol{\beta}}_{\boldsymbol{j}}, \hat{\boldsymbol{\psi}}_{\boldsymbol{j}}).$$

Given $\mathrm{A}_j \in \{1, 0\}$ in the linear Q-function (1.2), $\hat{a}_j^{opt}$ is further inferred to be $\hat{a}_j^{opt} = \mathbb{1}\left(\hat{\boldsymbol{\psi}}_{\boldsymbol{j}}^{\boldsymbol{T}} \boldsymbol{h_{j1}} > 0\right)$, where $\mathbb{1}(\cdot)$ is the indicator function. It implies that $\hat{a}_j^{opt} = 1$ if $\hat{\boldsymbol{\psi}}_{\boldsymbol{j}}^{\boldsymbol{T}} \boldsymbol{h_{j1}} > 0$, and $\hat{a}_j^{opt} = 0$, otherwise. Then, following this new expression for the estimated optimal treatment, the pseudo-outcome $\widetilde{Y}$ can be further written as

$$\widetilde{Y}_1 = \max_{a_2} Q_2(\boldsymbol{H_2}, a_2; \hat{\boldsymbol{\beta}}_{\boldsymbol{2}}, \hat{\boldsymbol{\psi}}_{\boldsymbol{2}}) = \hat{\boldsymbol{\beta}}_{\boldsymbol{2}}^{\boldsymbol{T}} \boldsymbol{H_{20}} + (\hat{\boldsymbol{\psi}}_{\boldsymbol{2}}^{\boldsymbol{T}} \boldsymbol{H_{21}}) \mathbb{1}\left(\hat{\boldsymbol{\psi}}_{\boldsymbol{2}}^{\boldsymbol{T}} \boldsymbol{H_{21}} > 0\right). \tag{1.3}$$

In (1.3), the term $\max_{a_2} Q_2(\boldsymbol{H_2}, a_2; \hat{\boldsymbol{\beta}}_{\boldsymbol{2}}, \hat{\boldsymbol{\psi}}_{\boldsymbol{2}})$ is expanded into a combination of treatment-free component and blip component, assuming that all the patients had received the optimal treatment at stage 2. This pseudo-outcome (1.3) can then be used for the first stage estimation.

A two-stage linear Q-learning algorithm is summarized in the following steps:

1. Parameterize the stage 2 Q-function

$$\mathrm{Q}_2(\boldsymbol{H_2}, A_2; \boldsymbol{\beta_2}, \boldsymbol{\psi_2}) = E[Y|\boldsymbol{H_2}, A_2] = \boldsymbol{\beta}_{\boldsymbol{2}}^{\boldsymbol{T}} \boldsymbol{H_{20}} + A_2\left(\boldsymbol{\psi}_{\boldsymbol{2}}^{\boldsymbol{T}} \boldsymbol{H_{21}}\right).$$

2. Apply ordinary least squares (OLS) procedure to obtain the stage 2 estimator $(\hat{\boldsymbol{\beta}}_{\boldsymbol{2}}, \hat{\boldsymbol{\psi}}_{\boldsymbol{2}})$

$$(\hat{\boldsymbol{\beta}}_{\boldsymbol{2}}, \hat{\boldsymbol{\psi}}_{\boldsymbol{2}}) = \underset{(\boldsymbol{\beta_2}, \boldsymbol{\psi_2})}{\arg\min} \frac{1}{n} \sum_{i=1}^{n} \left(Y_i - Q_2(\boldsymbol{H_2}, A_2; \boldsymbol{\beta_2}, \boldsymbol{\psi_2})\right)^2.$$

3. Derive the stage 2 optimal treatment as $\hat{a}_2^{opt} = \mathbb{1}\left(\hat{\boldsymbol{\psi}}_{\boldsymbol{2}}^{\boldsymbol{T}} \boldsymbol{h_{21}} > 0\right)$.

4. Construct the pseudo-outcome for estimation at stage 1

$$\widetilde{Y}_1 = \hat{\boldsymbol{\beta}}_{\boldsymbol{2}}^{\boldsymbol{T}} \boldsymbol{H_{20}} + (\hat{\boldsymbol{\psi}}_{\boldsymbol{2}}^{\boldsymbol{T}} \boldsymbol{H_{21}}) \mathbb{1}\left(\hat{\boldsymbol{\psi}}_{\boldsymbol{2}}^{\boldsymbol{T}} \boldsymbol{H_{21}} > 0\right).$$

5. Parameterize the stage 1 Q-function

$$\mathrm{Q}_1(\boldsymbol{H_1}, A_1; \boldsymbol{\beta_1}, \boldsymbol{\psi_1}) = \boldsymbol{\beta}_{\boldsymbol{1}}^{\boldsymbol{T}} \boldsymbol{H_{10}} + A_1\left(\boldsymbol{\psi}_{\boldsymbol{1}}^{\boldsymbol{T}} \boldsymbol{H_{11}}\right).$$

6. Apply OLS procedure to obtain the stage 1 estimator $(\hat{\boldsymbol{\beta}}_{\boldsymbol{1}}, \hat{\boldsymbol{\psi}}_{\boldsymbol{1}})$

$$(\hat{\boldsymbol{\beta}}_{\boldsymbol{1}}, \hat{\boldsymbol{\psi}}_{\boldsymbol{1}}) = \underset{(\boldsymbol{\beta_1}, \boldsymbol{\psi_1})}{\arg\min} \frac{1}{n} \sum_{i=1}^{n} \left(\widetilde{Y}_{i1} - Q_1(\boldsymbol{H_1}, A_1; \boldsymbol{\beta_1}, \boldsymbol{\psi_1})\right)^2.$$

7. Derive the stage 1 optimal treatment as $\hat{a}_1^{opt} = \mathbb{1}\left(\hat{\boldsymbol{\psi}}_{\boldsymbol{1}}^{\boldsymbol{T}} \boldsymbol{h_{11}} > 0\right)$.

Q-learning enjoys the advantage of simplicity in implementation. Following the procedures above, the regression parameters ($\boldsymbol{\beta}_j$, $\boldsymbol{\psi}_j$) are consistently estimated using the ordinary least squares method (Chakraborty, 2013). However, the validity of Q-learning requires a correct specification of the outcome model (1.2). If the outcome model is misspecified, it results in inconsistent estimates of parameters (Chakraborty, 2013).

1.1.3 Q-learning with Binary Outcomes

When the outcome of interest is binary, Moodie et al. (2014) presented a modified Q-function using the inverse-logit function at stage j

$$Q_j(\boldsymbol{H_j}, A_j; \boldsymbol{\beta_j}, \boldsymbol{\psi_j}) = \text{expit}\Big(\boldsymbol{\beta_j^T H_{j0}} + A_j(\boldsymbol{\psi_j^T H_{j1}})\Big),$$

where expit(x) = 1/(1 + exp(-x)), and $Q_j(\boldsymbol{H_j}, A_j)$ is bounded by [0, 1]. Then, the two-stage Q-functions are followed by

$$\begin{aligned} Q_2(\boldsymbol{H_2}, A_2; \boldsymbol{\beta_2}, \boldsymbol{\psi_2}) &= E[Y|\boldsymbol{H_2}, A_2] = \text{expit}\Big(\boldsymbol{\beta_2^T H_{20}} + A_2(\boldsymbol{\psi_2^T H_{21}})\Big), \\ Q_1(\boldsymbol{H_1}, A_1; \boldsymbol{\beta_1}, \boldsymbol{\psi_1}) &= \text{expit}\Big(\boldsymbol{\beta_1^T H_{10}} + A_1(\boldsymbol{\psi_1^T H_{11}})\Big). \end{aligned} \tag{1.4}$$

In Q-learning with binary response, the pseudo-outcome $\widetilde{Y}_1$ is constructed as the logit of $\text{Q}_2(\boldsymbol{H_2}, a_2; \hat{\boldsymbol{\beta}}_{\boldsymbol{2}}, \hat{\boldsymbol{\psi}}_{\boldsymbol{2}})$

$$\widetilde{Y}_1 = \max_{a_2} \text{logit} Q_2(\boldsymbol{H_2}, a_2; \hat{\boldsymbol{\beta}}_{\boldsymbol{2}}, \hat{\boldsymbol{\psi}}_{\boldsymbol{2}}), \tag{1.5}$$

where (1.5) is essentially the logit of predicted probability had the patients received the second stage optimal treatment. In this way, $\widetilde{Y}_1$ is projected to values in the real line for the stage-1 estimation, which is performed using the OLS in a model for the logit of $Q_1(\boldsymbol{H_{i1}}, A_{i1}; \boldsymbol{\beta_1}, \boldsymbol{\psi_1})$.

Once the stage-j estimator $(\hat{\boldsymbol{\beta}}_{\boldsymbol{j}}, \hat{\boldsymbol{\psi}}_{\boldsymbol{j}})$ is obtained, the estimated optimal treatment $\hat{a}_j^{opt}$ can be obtained by either directly maximizing $Q_j(\boldsymbol{h_j}, a_j; \hat{\boldsymbol{\beta}}_{\boldsymbol{j}}, \hat{\boldsymbol{\psi}}_{\boldsymbol{j}})$ or only maximizing the blip component $a_j(\hat{\boldsymbol{\psi}}_{\boldsymbol{j}}^{\boldsymbol{T}} \boldsymbol{h_{j1}})$, as the inverse-logit function is strictly increasing.

A two-stage linear Q-learning algorithm with binary outcomes is summarized in the following steps:

1. Parameterize the stage 2 Q-function

$$\text{Q}_2(\boldsymbol{H_2}, A_2; \boldsymbol{\beta_2}, \boldsymbol{\psi_2}) = \text{expit}\Big(\boldsymbol{\beta_2^T H_{20}} + A_2(\boldsymbol{\psi_2^T H_{21}})\Big).$$

2. Apply logistic regression to obtain the stage 2 estimator $(\hat{\boldsymbol{\beta}}_{\boldsymbol{2}}, \hat{\boldsymbol{\psi}}_{\boldsymbol{2}})$ of the conditional mean model for Y, $\text{Q}_2(\boldsymbol{H_2}, A_2; \boldsymbol{\beta_2}, \boldsymbol{\psi_2})$.

3. Derive the optimal treatment as $\hat{a}_2^{opt} = \arg\max_{a_2} Q_2(\boldsymbol{h_2}, a_2; \hat{\boldsymbol{\beta}}_{\boldsymbol{2}}, \hat{\boldsymbol{\psi}}_{\boldsymbol{2}})$.

4. Construct the pseudo-outcome for estimation at stage 1

$$\widetilde{Y}_1 = \max_{a_2} \text{logit} Q_2(\boldsymbol{H_2}, a_2; \hat{\boldsymbol{\beta}}_{\boldsymbol{2}}, \hat{\boldsymbol{\psi}}_{\boldsymbol{2}}).$$

5. Apply OLS regression to obtain the stage 1 estimator $(\hat{\boldsymbol{\beta}}_{\boldsymbol{1}}, \hat{\boldsymbol{\psi}}_{\boldsymbol{1}})$

$$(\hat{\boldsymbol{\beta}}_{\boldsymbol{1}}, \hat{\boldsymbol{\psi}}_{\boldsymbol{1}}) = \arg\min_{(\boldsymbol{\beta_1}, \boldsymbol{\psi_1})} \tfrac{1}{n} \textstyle\sum_{i=1}^{n} \Big(\widetilde{Y}_{i1} - Q_1(\boldsymbol{H_{i1}}, A_{i1}; \boldsymbol{\beta_1}, \boldsymbol{\psi_1})\Big)^2.$$

6. Derive the optimal treatment as $\hat{a}_1^{opt} = \arg\max_{a_1} Q_1(\boldsymbol{h_1}, a_1; \hat{\boldsymbol{\beta}}_1, \hat{\boldsymbol{\psi}}_1)$.

This modified Q-learning algorithm distinguishes itself from ordinary Q-learning in some aspects. Firstly, instead of modeling the Q-functions linearly, Moodie et al. (2014) used the inverse-logit function to model $\mathrm{Q}_j(\boldsymbol{H_j}, A_j)$, which is the probability of success at stage j. Another difference lies in the construction of the pseudo-outcome. In the ordinary Q-learning, $\widetilde{Y}_1$ is constructed as an estimate of the expected outcome of all the patients who were optimally treated at the second stage. However, in this method, $\widetilde{Y}_1$ is the logit of predicted probability had the patients received the second stage optimal treatment. This modification allows the pseudo-outcome to be transformed from the probability to the values in the real line.

1.1.4 Dynamic Weighted Survival Modeling

When it comes to the DTR with survival outcomes, the estimation of optimal DTR is challenging because of the censoring. Censoring occurs when the patients withdraw from the study or are lost of follow-up during the study period. A multi-stage treatment period complicates the estimation as patients may experience the event of interest before the end of any stage.

To deal with the censored outcome in DTR, Simoneau et al. (2020b) proposed the dynamic weighted survival modeling (DWSurv), a doubly robust method, to estimate an optimal DTR with survival times being subject to right-censoring. To ensure the feasibility of DTR with survival outcomes, Simoneau et al. (2020b) made two more assumptions in addition to the assumptions (A1) and (A2) described in (1.1.1):

(A3) *Coarsening at random*: at the beginning of each stage, the probability of censoring onward is independent of future outcomes, given accrued information.

(A4) *Positivity*: at any j^{th} stage, $\mathrm{P}(A_j = a_j|\boldsymbol{H_j}, \eta_j = 1) > 0$ for all treatment options a_j and $\mathrm{P}(\Delta = 1|\boldsymbol{H_j}, A_j, \eta_j = 1) > 0$.

In a two-stage setting, the DWSurv method models the logarithm of survival times at stage 2 and stage 1 based on the accelerated failure time (AFT) models

$$\begin{aligned} \log T_2 &= f(\boldsymbol{h_{20}}; \boldsymbol{\beta_2}) + g(\boldsymbol{h_{21}}, a_2; \boldsymbol{\psi_2}) + \epsilon_2, \\ \log \widetilde{T} &= f(\boldsymbol{h_{10}}; \boldsymbol{\beta_1}) + g(\boldsymbol{h_{11}}, a_1; \boldsymbol{\psi_1}) + \epsilon_1, \end{aligned} \tag{1.6}$$

where the error term ϵ_j is independent and identically distributed with mean zero, and $\widetilde{T}$ is the overall pseudo-survival time had all the patients who entered the second stage received the optimal treatment a_2^{opt}. f($\boldsymbol{h_{j0}};\boldsymbol{\beta_j}$) and g($\boldsymbol{h_{j1}}, a_j;\boldsymbol{\psi_j}$) are the treatment-free component and blip component, respectively, with functions f(·) and g(·) specified in any form. The simplest case is to consider log-survival times in linear form

$$\begin{aligned} \log T_2 &= \boldsymbol{\beta_2^T h_{20}} + a_2(\boldsymbol{\psi_2^T h_{21}}) + \epsilon_2, \\ \log \widetilde{T} &= \boldsymbol{\beta_1^T h_{10}} + a_1(\boldsymbol{\psi_1^T h_{11}}) + \epsilon_1. \end{aligned} \tag{1.7}$$

The pseudo-survival time $\widetilde{T}$ is constructed as

$$\widetilde{T} = T_1 + \eta_2 T_2 exp\Big\{\boldsymbol{\psi_2^T h_{21}}[a_2^{opt} - a_2]\Big\}. \tag{1.8}$$

Intuitively, (1.8) reflects three possible situations. If patients did not enter the second stage, $\widetilde{T}$ is equal to T_1, the survival time at the first stage. If patients entered the second stage and received the optimal treatment a_2^{opt}, $\widetilde{T}$ is equal to the observed overall survival time $T = T_1 + T_2$. If patients entered the second stage but did not receive the optimal treatment a_2^{opt}, $\widetilde{T}$ is larger than T due to the non-zero term $\exp\Big\{\boldsymbol{\psi_2^T h_{21}}[a_2^{opt} - a_2]\Big\}$.

The DWSurv method is designed for observational studies, where confounders may exist. Thus, Simoneau et al. (2020b) introduced weights in the algorithm so that by including carefully chosen weights, any possible confounding effect on estimating the parameters could be eliminated. The balancing property (1.9) is provided to find the appropriate weights

$$g(1, \boldsymbol{h_j})\pi(\boldsymbol{h_j})w_j(1, 1, \boldsymbol{h_j}) = [1 - g(0, \boldsymbol{h_j})][1 - \pi(\boldsymbol{h_j})]w_j(0, 0, \boldsymbol{h_j}), \tag{1.9}$$

where $\pi(\boldsymbol{h_j}) = P(A_j = 1|\boldsymbol{H_j = h_j}, \eta_j = 1)$ is the treatment model, and $g(a_j, \boldsymbol{h_j}) = P(\Delta = 1|\boldsymbol{H_j = h_j}, A_j = a_j, \eta_j = 1)$ is the censoring model. Simoneau et al. (2020b) demonstrates that the DWSurv method yields consistent blip parameter estimates if the weights satisfy (1.9). There is a broad class of weight choices that satisfy (1.9), and the use of weight (1.10) is recommended

$$w_j(\delta, a_j, \boldsymbol{h_j}) = \frac{|a_j - P(A_j = 1|\boldsymbol{h_j}, \eta_j = 1)|}{P(\Delta = \delta|\boldsymbol{h_j}, a_j, \eta_j = 1)}. \tag{1.10}$$

A two-stage linear DWSurv algorithm is summarized in the following steps:

1. Propose parametric models for the probability of treatment $\mathrm{P}(A_2 = 1|\boldsymbol{h_2}, \eta_2 = 1; \boldsymbol{\alpha_2})$ and the probability of censoring $\mathrm{P}(\Delta = 0|\boldsymbol{h_2}, a_2, \eta_2 = 1; \boldsymbol{\lambda_2})$ at stage 2 and find the estimated weight $\hat{w}_2$ from (1.10).

2. Assume a linear AFT model for the logarithm of survival time at stage 2 $\log T_2 = \boldsymbol{\beta_2^T h_{20}} + a_2\left(\boldsymbol{\psi_2^T h_{21}}\right) + \epsilon_2$ and obtain the estimator $(\hat{\boldsymbol{\beta}}_\mathbf{2}, \hat{\boldsymbol{\psi}}_\mathbf{2})$ by solving

$$\mathrm{U}_2(\boldsymbol{\beta_2}, \boldsymbol{\psi_2}) = \sum_{i=1}^n \delta_i \eta_{i2} \hat{w}_{i2} \begin{pmatrix} \boldsymbol{h_{i20}} \\ a_{i2}\boldsymbol{h_{i21}} \end{pmatrix} \left(\log T_{i2} - \boldsymbol{\beta_2^T h_{i20}} - a_{i2}\boldsymbol{\psi_2^T h_{i21}}\right) = 0.$$

3. Derive the stage 2 optimal treatment as $\hat{a}_2^{opt} = \mathbb{1}\left(\hat{\boldsymbol{\psi}}_\mathbf{2}^\mathbf{T}\boldsymbol{h_{21}} > 0\right)$.

4. Construct the pseudo-survival time for estimation at stage 1

$$\widetilde{T} = T_1 + \eta_2 T_2 exp\left\{\hat{\boldsymbol{\psi}}_\mathbf{2}^\mathbf{T}\boldsymbol{h_{21}}[\hat{a}_2^{opt} - a_2]\right\}.$$

5. Propose parametric models for the probability of treatment $\mathrm{P}(A_1 = 1|\boldsymbol{h_1}, \eta_1 = 1; \boldsymbol{\alpha_1})$ and the probability of censoring $\mathrm{P}(\Delta = 0|\boldsymbol{h_1}, a_1, \eta_1 = 1; \boldsymbol{\lambda_1})$ and find the estimated weight $\hat{w}_1$ from (1.10).

6. Assume a linear AFT model for the pseudo-survival time at stage 1 $\log\widetilde{T} = \boldsymbol{\beta_1^T h_{10}} + a_1\left(\boldsymbol{\psi_1^T h_{11}}\right) + \epsilon_1$ and obtain the estimator $(\hat{\boldsymbol{\beta}}_\mathbf{1}, \hat{\boldsymbol{\psi}}_\mathbf{1})$ by solving

$$\mathrm{U}_1(\boldsymbol{\beta_1}, \boldsymbol{\psi_1}) = \sum_{i=1}^n \delta_i \eta_{i1} \hat{w}_{i1} \begin{pmatrix} \boldsymbol{h_{i10}} \\ a_{i1}\boldsymbol{h_{i11}} \end{pmatrix} \left(\log\widetilde{T}_i - \boldsymbol{\beta_1^T h_{i10}} - a_{i1}\boldsymbol{\psi_1^T h_{i11}}\right) = 0.$$

7. Derive the stage 1 optimal treatment as $\hat{a}_1^{opt} = \mathbb{1}\left(\hat{\boldsymbol{\psi}}_\mathbf{1}^\mathbf{T}\boldsymbol{h_{11}} > 0\right)$.

Like Q-learning, DWSurv is implemented backward recursively from the last stage to the first stage. Moreover, DWSurv is a doubly robust method, which means that the consistency of the estimators remains if either the treatment-free model or weight model (treatment model and censoring model) is correctly specified. Thus, this double robustness property allows for model misspecification to some extent.

Simoneau et al. (2020b) further developed a formula to estimate the asymptotic variance of $(\hat{\boldsymbol{\beta}}, \hat{\boldsymbol{\psi}})$ in a single-stage, which is given by

$$Var(\hat{\boldsymbol{\beta}}, \hat{\boldsymbol{\psi}}) = E\left[\left\{E\left[\frac{\partial}{\partial(\boldsymbol{\beta}, \boldsymbol{\psi})} U_{\text{adj}}(\boldsymbol{\beta}, \boldsymbol{\psi})\right]^{-1} U_{\text{adj}}(\boldsymbol{\beta}, \boldsymbol{\psi})\right\}^{\otimes 2}\right], \tag{1.11}$$

where $\mathrm{E}[U_{\mathrm{adj}}(\boldsymbol{\beta},\boldsymbol{\psi})^{\otimes 2}] = \mathrm{E}[U_{\mathrm{adj}}(\boldsymbol{\beta},\boldsymbol{\psi})U_{\mathrm{adj}}(\boldsymbol{\beta},\boldsymbol{\psi})^T]$. $U_{\mathrm{adj}}(\boldsymbol{\beta},\boldsymbol{\psi})$, the estimating equation adjusted for the plug-in estimates of the nuisance parameters $\boldsymbol{\alpha}$ and $\boldsymbol{\lambda}$, is expressed as

$$U_{\mathrm{adj}}(\boldsymbol{\beta},\boldsymbol{\psi}) \approx U(\boldsymbol{\beta},\boldsymbol{\psi}) - E\left[\frac{\partial}{\partial \boldsymbol{\alpha}} U(\boldsymbol{\beta},\boldsymbol{\psi})\right] E\left[\frac{\partial}{\partial \boldsymbol{\alpha}} \boldsymbol{s}_{\alpha}\right]^{-1} \boldsymbol{s}_{\alpha} - E\left[\frac{\partial}{\partial \boldsymbol{\lambda}} U(\boldsymbol{\beta},\boldsymbol{\psi})\right] E\left[\frac{\partial}{\partial \boldsymbol{\lambda}} \boldsymbol{s}_{\lambda}\right]^{-1} \boldsymbol{s}_{\lambda},$$

where $\boldsymbol{s}_{\alpha}$ and $\boldsymbol{s}_{\lambda}$ are the score functions of the treatment and censoring models. With two or more stages, (1.11) applies in the last stage but with an additional term added to the estimating equations in the previous stages. Simoneau et al. (2020a) recommended the use of asymptotic variance, which takes much less computation time than the bootstrap approaches.

1.2 Measurement Error and Misclassification

1.2.1 Measurement Error in Covariates

Measurement error models are the statistical models that reveal the underlying mechanism of measurement error. It describes the relationship between the observed variable and the true variables. This section focuses on the review of measurement error in covariates.

For i = 1, ..., n, let $\boldsymbol{X_i}$ be an error-prone covariate, $\boldsymbol{Z_i}$ be an error-free covariate and $\boldsymbol{W_i}$ be a surrogate, mismeasured measurement of $\boldsymbol{X_i}$. We introduce three commonly used measurement error models:

Classical Additive Model

$$\boldsymbol{W_i} = \boldsymbol{X_i} + \boldsymbol{e_i}, \tag{1.12}$$

where the error term $\boldsymbol{e_i}$ is independent of $\boldsymbol{X_i}$ with mean $\mathbf{0}$ and covariance $\boldsymbol{\Sigma_{ee}}$.

Berkson Model

$$\boldsymbol{X_i} = \boldsymbol{W_i} + \boldsymbol{e_i},$$

where $\boldsymbol{e_i}$ is independent of $\boldsymbol{W_i}$ with mean $\mathbf{0}$ and covariance $\boldsymbol{\Sigma_{ee}}$.

Multiplicative Model

$$\boldsymbol{W_i} = \boldsymbol{X_i e_i},$$

where $\boldsymbol{e_i}$ is independent of $\boldsymbol{X_i}$ with mean $\mathbf{1}$.

An important concept in measurement error models is the non-differential error, which means the error term $\boldsymbol{e}$ contains no extra information about the outcome Y. Otherwise, the error is differential with respect to Y. The classical additive model is the most popular and widely used model among the three measurement error models (Carroll et al., 2006; Yi, 2017).

1.2.2 Misclassification in Response

When an error-prone variable is discrete, it is often described as a misclassification problem. We consider here a binary response Y that is subject to misclassification. Instead of fully observing the true response Y, a surrogate Y^* is observed as a mismeasured version of Y. Similar to the measurement error model, the misclassification modeling process is characterized by a set of misclassification probabilities ($\gamma_{10}(\boldsymbol{X})$, $\gamma_{01}(\boldsymbol{X})$), also called misclassification rates, to associate Y^* with Y such that

$$\gamma_{10}(\boldsymbol{X}) = P(Y^* = 1|Y = 0, \boldsymbol{X}),\ \gamma_{01}(\boldsymbol{X}) = P(Y^* = 0|Y = 1, \boldsymbol{X}). \tag{1.13}$$

In (1.13), the error in Y is differential in that Y^* is dependent on the covariate $\boldsymbol{X}$, conditional on Y. Otherwise, the error is non-differential if Y^* is independent of $\boldsymbol{X}$ conditional on Y.

1.2.3 Methods for Measurement Error in Covariates

There has been substantial research in the measurement error literature for correcting the bias caused by the measurement error in parameter estimation. It's worth noting that the choice of measurement error models and error correction methods largely depends on the nature of the research question and the structure of the available data. In this thesis, we concentrate on reviewing error correction methods for the classical additive model based on the data structure.

Methods for Validation Data

In practice, clinicians are sometimes only able to collect a small subset of data, in which all the variables $\{\boldsymbol{X_i}, \boldsymbol{W_i}, \boldsymbol{Z_i}, Y_i\}$ are observed, while the majority of data only have $(\boldsymbol{W_i}, \boldsymbol{Z_i}, \mathrm{Y}_i)$ observed. Such data are called (internal) validation data and the data structure is viewed as

$$\{\boldsymbol{X_i}, \boldsymbol{W_i}, \boldsymbol{Z_i}, Y_i\} \quad \text{if i} \in V,$$
$$\{\boldsymbol{W_i}, \boldsymbol{Z_i}, Y_i\} \quad \text{if i} \in \overline{V},$$

where the first group of data is the validation data denoted as V and the second group of data is main study data, also called non-validation data, denoted as $\overline{V}$. The main difference between these two groups is whether or not the true covariate $\boldsymbol{X_i}$ is available. It has been well documented that the covariate measurement error results in biased estimation of parameters without any corrections (Carroll et al., 2006; Yi, 2017). Under this class of data structure, the availability of a few $\boldsymbol{X_i}$ in V motivates researchers to develop methods to find the estimates $\hat{\boldsymbol{X}}_{\boldsymbol{i}}$ of the unobserved $\boldsymbol{X_i}$ in $\overline{V}$ by making use of the surrogates in the data. In the case of internal validation data, it is also regarded as a missing data problem (Cole et al., 2006).

Regression calibration (RC) is a classical approach to find such estimates $\hat{\boldsymbol{X}}_{\boldsymbol{i}}$ with two steps (Carroll et al., 2006). In the first step, a linear model is assumed for $\boldsymbol{X}$ and $(\boldsymbol{W}, \boldsymbol{Z})$, and the regression coefficients of the linear model are estimated using the OLS method based on the small set of validation data. In the second step, given the values of $(\boldsymbol{W_i}, \boldsymbol{Z_i})$ in $\overline{V}$, the estimates $\hat{\boldsymbol{X}}_{\boldsymbol{i}}$ for the missing $\boldsymbol{X_i}$ are predicted using the regression coefficients obtained in the previous step. This approach enjoys high popularity due to its simplicity in theory and implementation. However, the RC method is known to yield unbiased estimates of parameters in linear models and only approximately unbiased in nonlinear models (Carroll et al., 2006; Yi, 2017).

To overcome the limitation of the RC method in nonlinear models, Freedman et al. (2004) proposed a moment reconstruction (MR) method to create 'adjusted values' that have the same first and second moments as the unobserved true covariates. The MR method yields not only consistent estimators in linear models but also in nonlinear models. Thomas et al. (2011) introduced a moment-adjusted imputation (MAI) method, extending the MR method to higher-order moment matching. The MAI method is more advantageous than MR for non-normal

distributed covariates. Similar to MR and MAI, a multiple imputation (MI) method was proposed as another imputation method that can handle differential measurement error (Cole et al., 2006). The performances of the RC, MR, and MI methods were compared in linear and logistic regressions, and the MR and MI outperformed in terms of bias reduction but tended to be less efficient (Freedman et al., 2008).

In the survival data context, Jin et al. (2019) proposed a weighted least squares (WLS) method for the accelerated failure time (AFT) model when the true covariate $\boldsymbol{X_i}$ is only available in the validation data. Instead of having a continuous surrogate, a categorical auxiliary variable $\boldsymbol{M}$ is observed. Then, the structure of the data follows

$$\{\boldsymbol{X_i}, \boldsymbol{M_i}, \boldsymbol{Z}_i, Y_i, \Delta_i\} \quad \text{if i} \in V,$$
$$\{\boldsymbol{M_i}, \boldsymbol{Z}_i, Y_i, \Delta_i\} \quad \text{if i} \in \overline{V}.$$

For $\text{i} \in \overline{V}$, the $\boldsymbol{X_i}$ is estimated by

$$\hat{\boldsymbol{X}}_{\boldsymbol{i}} = \frac{\sum_{j \in V} \mathbb{1}(\boldsymbol{M_j} = \boldsymbol{M_i}) \boldsymbol{X_j}}{\sum_{j \in V} \mathbb{1}(\boldsymbol{M_j} = \boldsymbol{M_i})}. \tag{1.14}$$

This method has a few advantages. It does not need to specify the measurement error model. The consistency and asymptotic normality of the estimator were also established. However, one limitation of this method is that the variable $\boldsymbol{M}$ must be discrete to produce the estimates $\hat{\boldsymbol{X}}_{\boldsymbol{i}}$. For any continuous surrogate, Jin et al. (2019) suggested discretizing it into a categorical variable first and then applying (1.14) to obtain the estimates $\hat{\boldsymbol{X}}_{\boldsymbol{i}}$.

Machine learning techniques have also been applied in recent years to handle the measurement error for the use of survival data. Zhou and Wang (2000) introduced kernel smoothing to the Cox model and explored asymptotic properties for the estimators. Similar work on the Cox model can also be found in Hu and Lin (2002), Liu et al. (2009), Fan and Wang (2009) and Liu et al. (2010). To study the impact on the AFT model, Granville and Fan (2012) applied kernel smoothing to impute the $\hat{\boldsymbol{X}}_{\boldsymbol{i}}$ in the main study data. Granville and Fan (2014) further utilized local polynomial approximation to obtain a Buckley-James estimator of the AFT model.

Methods for Replicate Data

The data with k multiple surrogates or measurements $\boldsymbol{W} = (\boldsymbol{W_1}, \boldsymbol{W_2}, ..., \boldsymbol{W_k})$ for the unobserved $\boldsymbol{X}$ are called replicate data. For any patient i (i = 1, ..., n), the data structure is

$$\{\boldsymbol{W_{i1}}, \boldsymbol{W_{i2}}, ..., \boldsymbol{W_{ik_i}}, \boldsymbol{Z_i}, Y_i\},$$

where k_i is the number of replicate surrogates for the i^{th} subject. Let $\overline{\boldsymbol{W}}_{\boldsymbol{i}}$ be the mean value of $(\boldsymbol{W_{i1}}, \boldsymbol{W_{i2}}, ..., \boldsymbol{W_{ik_i}})$. Simply using the replicate surrogates in the analysis may still yield biased estimators. Thus, the question of interest is to search for a best linear approximation to $\boldsymbol{X_i}$ given $(\overline{\boldsymbol{W}}_{\boldsymbol{i}}, \boldsymbol{Z_i})$, that is, $E[\boldsymbol{X_i}|\overline{\boldsymbol{W}}_{\boldsymbol{i}}, \boldsymbol{Z_i}]$ for $\boldsymbol{X_i}$ as a linear function of $\overline{\boldsymbol{W}}_{\boldsymbol{i}}$ and $\boldsymbol{Z_i}$ (Carroll and Stefanski, 1990; Gleser, 1990).

Regression calibration is an approach to provide the best linear approximation $\hat{\boldsymbol{X}}_{\boldsymbol{i}}$, which is given by (Carroll et al., 2006)

$$\hat{\boldsymbol{X}}_{\boldsymbol{i}} = \hat{\mu}_w + \begin{bmatrix} \hat{\Sigma}_{xx}, & \hat{\Sigma}_{xz} \end{bmatrix} \begin{bmatrix} \hat{\Sigma}_{xx} + \hat{\Sigma}_{ee}/k_i & \hat{\Sigma}_{xz} \\ \hat{\Sigma}_{xz}^T & \hat{\Sigma}_{zz} \end{bmatrix}^{-1} \begin{pmatrix} \overline{\boldsymbol{W}}_{\boldsymbol{i}} - \hat{\mu}_w \\ \boldsymbol{Z_i} - \hat{\mu}_z \end{pmatrix}, \tag{1.15}$$

where

$$\begin{aligned}
\overline{\boldsymbol{W}}_{\boldsymbol{i}} &= \frac{1}{k_i} \sum_{j=1}^{k_i} \boldsymbol{W_{ij}}, \\
\hat{\mu}_x = \hat{\mu}_w &= \sum_{i=1}^{n} k_i \overline{\boldsymbol{W}}_{\boldsymbol{i}} \Big/ \sum_{i=1}^{n} k_i, \\
\hat{\mu}_z &= \overline{\boldsymbol{Z}}, \\
\nu &= \sum_{i=1}^{n} k_i - \sum_{i=1}^{n} k_i^2 \Big/ \sum_{i=1}^{n} k_i, \\
\hat{\Sigma}_{xx} &= \left[\left\{ \sum_{i=1}^{n} k_i (\overline{\boldsymbol{W}}_{\boldsymbol{i}} - \hat{\mu}_w)(\overline{\boldsymbol{W}}_{\boldsymbol{i}} - \hat{\mu}_w)^T \right\} - (n-1)\hat{\Sigma}_{ee} \right] \Big/ \nu, \\
\hat{\Sigma}_{xz} &= \sum_{i=1}^{n} k_i (\overline{\boldsymbol{W}}_{\boldsymbol{i}} - \hat{\mu}_w)(\boldsymbol{Z_i} - \overline{\boldsymbol{Z}})^T \Big/ \nu, \\
\hat{\Sigma}_{zz} &= (n-1)^{-1} \sum_{i=1}^{n} (\boldsymbol{Z_i} - \overline{\boldsymbol{Z}})(\boldsymbol{Z_i} - \overline{\boldsymbol{Z}})^T, \\
\hat{\Sigma}_{ee} &= \sum_{i=1}^{n} \sum_{j=1}^{k_i} (\boldsymbol{W_{ij}} - \overline{\boldsymbol{W}}_{\boldsymbol{i}})(\boldsymbol{W_{ij}} - \overline{\boldsymbol{W}}_{\boldsymbol{i}})^T \Big/ \sum_{i=1}^{n} (k_i - 1).
\end{aligned}$$

More recently, several other approaches have been developed as alternative methods for replicate data. Bartlett et al. (2009) proposed an efficient likelihood-based method for the linear and logistic outcome models applicable to replicate data. Keogh and White (2014) described an approach using the idea of MI in a study with repeated measurements. Muff et al. (2015) extended the integrated nested Laplace approximations method to correct classical measurement error in exposure when a replicate study is available. Gray (2018) extended MI approach suitable for the use with replicate data.

SIMEX-based Methods

Another class of methods directly deals with the naive estimators to correct the bias. Stefanski and Cook (1995) proposed Simulation-Extrapolation (SIMEX), a simulation-based approach to adjust for the covariate measurement error effect. The key idea is to first model the trend of estimators with different strengths of measurement errors through simulation. Then given the trend, the estimates are extrapolated back to the situation without measurement error. The main advantage of the SIMEX method is that it requires no specification of the distribution of true covariate, which makes SIMEX robust.

There are several applications and extensions of the SIMEX method in the literature. The R package `simex` implements the SIMEX method with a range of extrapolation functions provided (Lederer and Küchenhoff, 2006). Ronning and Rosemann (2008) took into account the correlation of error terms and proposed generalized SIMEX to accommodate the correlation. In the context of survival data, He et al. (2007) applied the SIMEX method to the AFT model when true covariates are error-prone and not restricted to a specific distribution. He et al. (2012) also developed an easy-to-implement R package `simexaft` for the use of the SIMEX method in the AFT model. Yi et al. (2015b) extended the SIMEX method to accommodate the effect of missingness in response and measurement error in covariates. Yi et al. (2015a) generalized the usual SIMEX method to treat measurement error and misclassification in covariates simultaneously. Zhang and Yi (2019) further developed an R package `augSIMEX` for the use of the generalized SIMEX method proposed in Yi et al. (2015a).

1.2.4 Methods for Misclassification in Response

The misclassification problem has been increasingly discussed in the literature. Statistical analysis with misclassified responses may result in severely biased estimators. Its negative impact is likely to be greater than the covariate measurement error because the misclassification can alter the structure of the response model (Zhu and Wu, 2004; Carroll et al., 2006; Yi, 2017). Neuhaus (1999) and Yi (2017) pointed out that ignoring the misclassification in the response during the analysis process is equivalent to modeling the data with a misspecified link function. Moreover, it may also incur a loss of efficiency of the estimators (Neuhaus, 1999).

Classical methods dealing with misclassification in the response can be found in the literature. If the distribution of the binary data can be specified, Hausman et al. (1998) and Neuhaus (1999) derived a relationship of models for observed surrogate Y^* and true response Y as

$$P(Y^* = 1|\boldsymbol{X}) = \gamma_{10} + (1 - \gamma_{10} - \gamma_{01})P(Y = 1|\boldsymbol{X}). \tag{1.16}$$

By equation (1.16), the maximum likelihood estimation (MLE) method can be used for the estimation of parameters by maximizing the log-likelihood of the data (1.17)

$$l = \sum_{i=1}^{n} P(Y_i^* = y_i^*|\boldsymbol{X_i} = \boldsymbol{x_i}) = \sum_{i=1}^{n} \sum_{y_i=0}^{1} P(Y_i^* = y_i^*|Y_i = y_i, \boldsymbol{X_i} = \boldsymbol{x_i})P(Y_i = y_i|\boldsymbol{X_i} = \boldsymbol{x_i}). \tag{1.17}$$

Neuhaus (2002) extended the likelihood method to clustered and longitudinal binary data with responses being subject to misclassification. Lyles and Lin (2010) utilized the direct MLE method to handle the outcome misclassification and proposed a predictive value weighting approach to correct the covariate misclassification. Lyles et al. (2011) further illustrated the likelihood-based method with the use of internal validation data in case-control studies to address the outcome misclassification.

Since the likelihood-based methods could be computationally intensive, the mean score method has been proposed as an alternative choice (Pepe, 1992; Pepe et al., 1994). Yi (2017) further elaborated it to be a semiparametric approach. In addition to these methods, an EM algorithm was introduced by Magder and Hughes (1997) to handle the misclassified outcome, which is also applicable to the differential misclassification. Edwards et al. (2013) developed a

multiple imputation approach when validation data are available. Bayesian methods have also been established using prior distributions to correct the misclassified binary response (Prescott and Garthwaite, 2002,0; Daniel Paulino et al., 2003; Gerlach and Stamey, 2007).

Machine learning approaches have received increasing attention over the years to accommodate the misclassification effect. Xu et al. (2006) proposed a robust support vector machine (SVM) to account for the misclassification in the response using a robust loss function. Yang et al. (2007) developed a weighted SVM to improve the outlier sensitivity problem in standard SVM. Random forest was demonstrated to be the most robust classifier compared with the other ten classifiers in noisy data with a misclassified response (Folleco et al., 2008).

The effect of misclassification in response in other applications has also been explored. Mwalili et al. (2008) described an approach to correct the misclassification in a zero-inflated negative binomial regression model. Chen et al. (2014) introduced a marginal method for longitudinal ordinal data with misclassification in both response and covariates. Shu and Yi (2019a) studied the misclassified outcome with missingness in causal inference and proposed methods to correct misclassification and missingness effects simultaneously. An R package `ipwErrorY` was developed by Shu and Yi (2019b) for the corrected estimation of average treatment effect in causal inference with a misclassified response. In genetics studies, Zhang and Yi (2020) explored bivariate mixed responses with measurement error and misclassification, and used the likelihood-based methods to correct the measurement error and misclassification effects simultaneously. Zhang and Yi (2021) further proposed estimating equation approaches to deal with measurement error and misclassification in bivariate responses.

1.3 Objectives and Organizations

Although there is extensive literature on dynamic treatment regimes and measurement error/misclassification, to the best of our knowledge, the study of the measurement error or misclassification effect on dynamic treatment regimes is scarce. Most existing methods in DTR literature ignore the fact that the patient's characteristics and the outcome may be contaminated with measurement error or misclassification. Consequently, the estimation of optimal DTR that relies on a collection of error-prone variables may be severely biased if those errors

are not addressed. Spicker and Wallace (2020) demonstrated the substantial impact of measurement error on dynamic weighted ordinary least squares. In this thesis, we extend to study and accommodate the effect of covariate measurement error and outcome misclassification in the contexts of DTR approaches, including Q-learning with continuous outcomes, Q-learning with binary outcomes, and dynamic weighted survival modeling.

The remainder of this thesis is organized as follows. In Chapter 2, we consider Q-learning with continuous outcomes, in which the covariates are considered mismeasured with repeated measurements. The regression calibration method is employed to correct the measurement error in Q-learning. In Chapter 3, we consider the situation of DTR with survival outcomes based on the DWSurv method for internal validation data with covariates being contaminated. Two correction methods, the k-nearest neighbors method and the weighted least squares method, are developed to eliminate the effect of error-prone covariates. In Chapter 4, we consider Q-learning with misclassified binary outcomes and internal validation data. The maximum likelihood estimation method is proposed to accommodate the misclassification effect in Q-learning. A summary of findings and future work is presented in Chapter 5.

Chapter 2

Dynamic Treatment Regimes with Measurement Error in Covariates: a Q-learning Approach

2.1 Introduction

In this chapter, we study the effect of covariate measurement error on Q-learning, a DTR method with continuous outcomes. The existing research work in the study of Q-learning assumes that the collected covariates are free from measurement error. However, this assumption is commonly violated in clinical practice. To date, it remains unclear whether and how much the covariate measurement error plays a role in affecting the performance of Q-learning. This chapter aims to study the effect of measurement error in covariates on Q-learning. Specifically, the impact of covariate measurement error in Q-learning will be examined, and regression calibration will be explored to adjust for the measurement error effect.

The remainder of this chapter is organized as follows. In Section 2.2, we describe the Q-learning with mismeasured covariates and the use of the regression calibration method in Q-learning to correct the covariate measurement error. Simulation studies are carried out to examine the performance of the RC method in Section 2.3. In Section 2.4, we apply the proposed method to the STAR*D study. The conclusions are summarized in Section 2.5.

2.2 Methodology

2.2.1 Notations and Model Framework

We restrict the notations and framework set-up to DTR with two decision points. The data trajectory follows $\{\boldsymbol{X_1}, \boldsymbol{Z_1}, A_1, \boldsymbol{X_2}, \boldsymbol{Z_2}, A_2, Y\}$, where $\boldsymbol{X_j}$ and $\boldsymbol{Z_j}$ are error-prone covariate vector and error-free covariate vector (j = 1, 2). We consider a situation where the true covariate $\boldsymbol{X_j}$ is not observable at stage j. Instead, there are up to k_j unbiased replicate surrogates observed for $\boldsymbol{W_j} = (\boldsymbol{W_{j1}}, ..., \boldsymbol{W_{jk_j}})$, where $\boldsymbol{W_{jl}}$ (l = 1, ..., k_j) denotes a surrogate or mismeasured version of $\boldsymbol{X_j}$. The classical additive model is assumed to describe the relationship of $\boldsymbol{W_{jl}}$ and $\boldsymbol{X_j}$, that is $\boldsymbol{W_{jl}} = \boldsymbol{X_j} + \boldsymbol{e_{jl}}$, where the $\boldsymbol{e_{jl}}$ follow a normal distribution with mean $\mathbf{0}$ and covariance $\boldsymbol{\Sigma_{ee}}$ and are independent of each other and of all other variables. The binary treatment $\mathrm{A}_j \in \{1, 0\}$ is assigned at stage j. Y is a continuous outcome observed at the end of the second stage.

In the presence of measurement error, the true covariate $\boldsymbol{X_j}$ is absent but only the replicate surrogates $\boldsymbol{W_j}$ are observed at stage j. Then, the data trajectory is replaced by

$$\{\boldsymbol{W_1}, \boldsymbol{Z_1}, A_1, \boldsymbol{W_2}, \boldsymbol{Z_2}, A_2, Y\}.$$

In this case, the naive histories are formed as $\boldsymbol{H_1^n} = (\boldsymbol{W_1}, \boldsymbol{Z_1})$ and $\boldsymbol{H_2^n} = (\boldsymbol{W_1}, \boldsymbol{Z_1}, A_1, \boldsymbol{W_2}, \boldsymbol{Z_2})$. As a result, the Q-functions that use the naive histories are called naive Q-functions, which contain the replicate surrogates only rather than the true covariates. Then, the naive Q-functions are given by

$$\begin{aligned} Q_2(\boldsymbol{H_2^n}, A_2; \boldsymbol{\beta_2^n}, \boldsymbol{\psi_2^n}) &= f(\boldsymbol{W_1}, \boldsymbol{Z_1}, A_1, \boldsymbol{W_2}, \boldsymbol{Z_2}; \boldsymbol{\beta_2^n}) + g(\boldsymbol{W_1}, \boldsymbol{Z_1}, A_1, \boldsymbol{W_2}, \boldsymbol{Z_2}, A_2; \boldsymbol{\psi_2^n}), \\ Q_1(\boldsymbol{H_1^n}, A_1; \boldsymbol{\beta_1^n}, \boldsymbol{\psi_1^n}) &= f(\boldsymbol{W_1}, \boldsymbol{Z_1}; \boldsymbol{\beta_1^n}) + g(\boldsymbol{W_1}, \boldsymbol{Z_1}, A_1; \boldsymbol{\psi_1^n}). \end{aligned}$$

Using the naive histories $\boldsymbol{H_1^n}$ and $\boldsymbol{H_2^n}$, the naive Q-functions can be further summarized as

$$\begin{aligned} Q_2(\boldsymbol{H_2^n}, A_2; \boldsymbol{\beta_2^n}, \boldsymbol{\psi_2^n}) &= f(\boldsymbol{H_{20}^n}; \boldsymbol{\beta_2^n}) + g(\boldsymbol{H_{21}^n}, A_2; \boldsymbol{\psi_2^n}), \\ Q_1(\boldsymbol{H_1^n}, A_1; \boldsymbol{\beta_1^n}, \boldsymbol{\psi_1^n}) &= f(\boldsymbol{H_{10}^n}; \boldsymbol{\beta_1^n}) + g(\boldsymbol{H_{11}^n}, A_1; \boldsymbol{\psi_1^n}). \end{aligned} \tag{2.1}$$

If the functions f(·) and g(·) are modeled linearly, then the naive Q-function at stage j is

given by

$$Q_j(\boldsymbol{H}_j^n, A_j; \boldsymbol{\beta}_j^n, \boldsymbol{\psi}_j^n) = \boldsymbol{\beta}_j^{nT} \boldsymbol{H}_{j0}^n + A_j\left(\boldsymbol{\psi}_j^{nT} \boldsymbol{H}_{j1}^n\right). \tag{2.2}$$

The naive Q-functions (2.1) and (2.2) are different from (1.1) and (1.2) in the sense that the original history is replaced with the naive history. By applying the ordinary least squares (OLS), the naive estimator $(\hat{\boldsymbol{\beta}}_j^n, \hat{\boldsymbol{\psi}}_j^n)$ can be obtained. According to Carroll et al. (2006) and Yi (2017), it is reasonable to believe that the naive estimator $(\hat{\boldsymbol{\beta}}_j^n, \hat{\boldsymbol{\psi}}_j^n)$ may be biased from $(\boldsymbol{\beta}_j, \boldsymbol{\psi}_j)$. Let the blip parameter $\boldsymbol{\psi} = (\boldsymbol{\psi_2}, \boldsymbol{\psi_1})$, which is the parameter of primary interest for estimation. Then the naive blip estimator $\hat{\boldsymbol{\psi}}^n = (\hat{\boldsymbol{\psi}}_2^n, \hat{\boldsymbol{\psi}}_1^n)$ may be biased from $\boldsymbol{\psi}$. Consequently, we are motivated to assess the degree of biases in the parameter estimation and search for a good approximation $\hat{\boldsymbol{X}}_j$ to $\boldsymbol{X}_j$ using the available replicate surrogates in the data.

2.2.2 Regression Calibration

Prentice (1982) pioneered the regression calibration method to address covariate measurement error in a survival data context. It has now become a widely used error correction method, which can tackle the measurement error problems for both validation data and replicate data (Carroll et al., 2006). In this chapter, we focus on the study with replicate data. The key idea of regression calibration is to find the estimates $\hat{\boldsymbol{X}}$ of $\boldsymbol{X}$ using the available replicate surrogates and proceed with the analysis using the estimates $\hat{\boldsymbol{X}}$ so that the bias caused by the measurement error is reduced.

For any stage j, we can obtain the RC estimates $\hat{\boldsymbol{X}}_j$ using the replicate surrogates $\boldsymbol{W}_j$ according to the formula (1.15). Then, by replacing the unobserved $\boldsymbol{X}_j$ with the corrected estimates $\hat{\boldsymbol{X}}_j$, the data trajectory is updated to be

$$\{\hat{\boldsymbol{X}}_\mathbf{1}, \boldsymbol{Z_1}, A_1, \hat{\boldsymbol{X}}_\mathbf{2}, \boldsymbol{Z_2}, A_2, Y\}.$$

The corresponding Q-functions using the corrected estimates $\hat{\boldsymbol{X}}_j$ are followed by

$$\begin{aligned}
Q_2(\boldsymbol{H}_\mathbf{2}^{rc}, A_2; \boldsymbol{\beta}_\mathbf{2}^{rc}, \boldsymbol{\psi}_\mathbf{2}^{rc}) &= f(\hat{\boldsymbol{X}}_\mathbf{1}, \boldsymbol{Z_1}, A_1, \hat{\boldsymbol{X}}_\mathbf{2}, \boldsymbol{Z_2}; \boldsymbol{\beta}_\mathbf{2}^{rc}) + g(\hat{\boldsymbol{X}}_\mathbf{1}, \boldsymbol{Z_1}, A_1, \hat{\boldsymbol{X}}_\mathbf{2}, \boldsymbol{Z_2}, A_2; \boldsymbol{\psi}_\mathbf{2}^{rc}), \\
Q_1(\boldsymbol{H}_\mathbf{1}^{rc}, A_1; \boldsymbol{\beta}_\mathbf{1}^{rc}, \boldsymbol{\psi}_\mathbf{1}^{rc}) &= f(\hat{\boldsymbol{X}}_\mathbf{1}, \boldsymbol{Z_1}; \boldsymbol{\beta}_\mathbf{1}^{rc}) + g(\hat{\boldsymbol{X}}_\mathbf{1}, \boldsymbol{Z_1}, A_1; \boldsymbol{\psi}_\mathbf{1}^{rc}).
\end{aligned}$$

From the new data trajectory, we can obtain the RC histories in a form $\boldsymbol{H}_\mathbf{1}^{rc} = (\hat{\boldsymbol{X}}_\mathbf{1}, \boldsymbol{Z_1})$ and

$\boldsymbol{H}_{\boldsymbol{2}}^{rc} = (\hat{\boldsymbol{X}}_{\boldsymbol{1}}, \boldsymbol{Z}_{\boldsymbol{1}}, A_1, \hat{\boldsymbol{X}}_{\boldsymbol{2}}, \boldsymbol{Z}_{\boldsymbol{2}})$. The Q-functions based on the RC histories are given by

$$\begin{aligned} Q_2(\boldsymbol{H}_{\boldsymbol{2}}^{rc}, A_2; \boldsymbol{\beta}_{\boldsymbol{2}}^{rc}, \boldsymbol{\psi}_{\boldsymbol{2}}^{rc}) &= f(\boldsymbol{H}_{\boldsymbol{20}}^{rc}; \boldsymbol{\beta}_{\boldsymbol{2}}^{rc}) + g(\boldsymbol{H}_{\boldsymbol{21}}^{rc}, A_2; \boldsymbol{\psi}_{\boldsymbol{2}}^{rc}), \\ Q_1(\boldsymbol{H}_{\boldsymbol{1}}^{rc}, A_1; \boldsymbol{\beta}_{\boldsymbol{1}}^{rc}, \boldsymbol{\psi}_{\boldsymbol{1}}^{rc}) &= f(\boldsymbol{H}_{\boldsymbol{10}}^{rc}; \boldsymbol{\beta}_{\boldsymbol{1}}^{rc}) + g(\boldsymbol{H}_{\boldsymbol{11}}^{rc}, A_1; \boldsymbol{\psi}_{\boldsymbol{1}}^{rc}). \end{aligned} \tag{2.3}$$

If each Q-function in (2.3) is modeled linearly, then it can be expressed as

$$Q_j(\boldsymbol{H}_{\boldsymbol{j}}^{rc}, A_j; \boldsymbol{\beta}_{\boldsymbol{j}}^{rc}, \boldsymbol{\psi}_{\boldsymbol{j}}^{rc}) = \boldsymbol{\beta}_{\boldsymbol{j}}^{rcT} \boldsymbol{H}_{\boldsymbol{j0}}^{rc} + A_j\left(\boldsymbol{\psi}_{\boldsymbol{j}}^{rcT} \boldsymbol{H}_{\boldsymbol{j1}}^{rc}\right). \tag{2.4}$$

The modified Q-functions (2.3) and (2.4) are formalized based on the RC histories, which consist of the corrected estimates for the error-prone covariates and other variables. Then, the estimator $(\hat{\boldsymbol{\beta}}_{\boldsymbol{j}}^{rc}, \hat{\boldsymbol{\psi}}_{\boldsymbol{j}}^{rc})$ obtained from the Q-functions (2.3) is the RC estimator. It is discussed that the RC method yields consistent estimators in linear models but is approximately consistent in nonlinear models (Carroll et al., 2006; Yi, 2017). Thus, if the Q-function is in a form of (2.4), $(\hat{\boldsymbol{\beta}}_{\boldsymbol{j}}^{rc}, \hat{\boldsymbol{\psi}}_{\boldsymbol{j}}^{rc})$ is a consistent estimator of $(\boldsymbol{\beta}_j, \boldsymbol{\psi}_j)$ and $\hat{\boldsymbol{\psi}}^{rc} = (\hat{\boldsymbol{\psi}}_{\boldsymbol{2}}^{rc}, \hat{\boldsymbol{\psi}}_{\boldsymbol{1}}^{rc})$ consistently estimates the blip parameter $\boldsymbol{\psi}$. However, if the Q-function is in a nonlinear form, regression calibration can still provide a considerable bias reduction in the parameter estimation in Q-learning.

Modified Q-learning Algorithm with Regression Calibration:

The following modified Q-learning algorithm with regression calibration details the estimation procedure:

1. Parameterize the stage 2 Q-function

$$\mathrm{Q}_2(\boldsymbol{H}_{\boldsymbol{2}}^{rc}, A_2; \boldsymbol{\beta}_{\boldsymbol{2}}^{rc}, \boldsymbol{\psi}_{\boldsymbol{2}}^{rc}) = \boldsymbol{\beta}_{\boldsymbol{2}}^{rcT} \boldsymbol{H}_{\boldsymbol{20}}^{rc} + A_2\left(\boldsymbol{\psi}_{\boldsymbol{2}}^{rcT} \boldsymbol{H}_{\boldsymbol{21}}^{rc}\right).$$

2. Apply OLS procedure to obtain the stage 2 estimator $(\hat{\boldsymbol{\beta}}_{\boldsymbol{2}}^{rc}, \hat{\boldsymbol{\psi}}_{\boldsymbol{2}}^{rc})$

$$(\hat{\boldsymbol{\beta}}_{\boldsymbol{2}}^{rc}, \hat{\boldsymbol{\psi}}_{\boldsymbol{2}}^{rc}) = \underset{(\boldsymbol{\beta}_{\boldsymbol{2}}^{rc}, \boldsymbol{\psi}_{\boldsymbol{2}}^{rc})}{\arg\min} \tfrac{1}{n} \textstyle\sum_{i=1}^{n} \left(Y_i - Q_2(\boldsymbol{H}_{\boldsymbol{i2}}^{rc}, A_{i2}; \boldsymbol{\beta}_{\boldsymbol{2}}^{rc}, \boldsymbol{\psi}_{\boldsymbol{2}}^{rc})\right)^2.$$

3. Derive the stage 2 optimal treatment as $\hat{a}_2^{opt} = \mathbb{1}\left(\hat{\boldsymbol{\psi}}_{\boldsymbol{2}}^{rcT} \boldsymbol{h}_{\boldsymbol{21}}^{rc} > 0\right)$.
4. Construct the pseudo-outcome for estimation at stage 1

$$\widetilde{Y}_1 = \hat{\boldsymbol{\beta}}_{\boldsymbol{2}}^{rcT} \boldsymbol{H}_{\boldsymbol{20}}^{rc} + \left(\hat{\boldsymbol{\psi}}_{\boldsymbol{2}}^{rcT} \boldsymbol{H}_{\boldsymbol{21}}^{rc}\right) \mathbb{1}\left(\hat{\boldsymbol{\psi}}_{\boldsymbol{2}}^{rcT} \boldsymbol{H}_{\boldsymbol{21}}^{rc} > 0\right).$$

5. Parameterize the stage 1 Q-function

$$\mathrm{Q}_1(\boldsymbol{H}_{\mathbf{1}}^{\boldsymbol{rc}}, A_1; \boldsymbol{\beta}_{\mathbf{1}}^{\boldsymbol{rc}}, \boldsymbol{\psi}_{\mathbf{1}}^{\boldsymbol{rc}}) = \boldsymbol{\beta}_{\mathbf{1}}^{\boldsymbol{rcT}} \boldsymbol{H}_{\mathbf{10}}^{\boldsymbol{rc}} + A_1\left(\boldsymbol{\psi}_{\mathbf{1}}^{\boldsymbol{rcT}} \boldsymbol{H}_{\mathbf{11}}^{\boldsymbol{rc}}\right).$$

6. Apply OLS procedure to obtain the stage 1 estimator $(\hat{\boldsymbol{\beta}}_{\mathbf{1}}^{\boldsymbol{rc}}, \hat{\boldsymbol{\psi}}_{\mathbf{1}}^{\boldsymbol{rc}})$

$$(\hat{\boldsymbol{\beta}}_{\mathbf{1}}^{\boldsymbol{rc}}, \hat{\boldsymbol{\psi}}_{\mathbf{1}}^{\boldsymbol{rc}}) = \underset{(\boldsymbol{\beta}_{\mathbf{1}}^{\boldsymbol{rc}}, \boldsymbol{\psi}_{\mathbf{1}}^{\boldsymbol{rc}})}{\arg\min} \tfrac{1}{n} \textstyle\sum_{i=1}^{n} \left(\widetilde{Y}_{i1} - Q_1(\boldsymbol{H}_{\boldsymbol{i1}}^{\boldsymbol{rc}}, A_{i1}; \boldsymbol{\beta}_{\mathbf{1}}^{\boldsymbol{rc}}, \boldsymbol{\psi}_{\mathbf{1}}^{\boldsymbol{rc}})\right)^2.$$

7. Derive the stage 1 optimal treatment as $\hat{a}_1^{opt} = \mathbb{1}\left(\hat{\boldsymbol{\psi}}_{\mathbf{1}}^{\boldsymbol{rcT}} \boldsymbol{h}_{\mathbf{11}}^{\boldsymbol{rc}} > 0\right)$.

This modified Q-learning algorithm distinguishes itself from the original Q-learning in Section (1.1.2) in the sense that the history used for the parameter estimation is only an approximation of the true underlying history.

2.3 Simulation Studies

In this section, we conduct a series of simulation studies to assess the measurement error effect on estimating the parameters and predicting the optimal treatment decision rules and optimal value function in Q-learning. The performance of the regression calibration method is examined and compared with the naive method in one-stage and two-stage Q-learning.

2.3.1 One-Stage Estimation

We begin with one-stage parameter estimation in Q-learning. Let X and Z be the error-prone and error-free covariates, respectively, which are generated from N(1, 1). Instead of observing X, two replicate surrogates W_1, W_2 are observed as mismeasured version for X, modeled by $\mathrm{W}_l = \mathrm{X} + \mathrm{e}_l$ ($l = 1, 2$), where $\mathrm{e}_l \sim \mathrm{N}(0, \sigma^2)$. $\overline{W}$ is an average value of W_1 and W_2, given by $\overline{W} = (W_1 + W_2)/2$. Treatment $\mathrm{A} \in \{1, 0\}$ is generated from a Bernoulli distribution with probability $\mathrm{P}(\mathrm{A} = 1) = 1/2$. The outcome Y is generated by $\mathrm{Y} = 0.5 + \beta_z \mathrm{Z} + \beta_x \mathrm{X} + \mathrm{A}(\psi_{10} + \psi_{11}\mathrm{X}) + \epsilon$, where $(\boldsymbol{\beta}, \boldsymbol{\psi}) = (\beta_z, \beta_x, \psi_{10}, \psi_{11}) = (0.5, 1, 0.5, 1)$ and $\epsilon \sim \mathrm{N}(0, 1)$, independent of each other and all the other variables.

Four estimators are considered and compared in each round of 500 simulations: (1) true estimator $\hat{\boldsymbol{\psi}}^{\boldsymbol{t}}$ obtained using the true covariate X, (2) naive estimator $\hat{\boldsymbol{\psi}}^{\boldsymbol{n}}$ obtained using a single surrogate W_1, (3) naive estimator $\hat{\boldsymbol{\psi}}^{\boldsymbol{nb}}$ obtained using the averaged surrogate $\overline{W}$, (4) RC

estimator $\hat{\boldsymbol{\psi}}^{rc}$ obtained using the RC estimates $\hat{X}$. Analyses are conducted under two different sample sizes of n = 500 and n = 2000. In each setting, the bias, empirical standard error (SE), root mean square error (RMSE) and 95% coverage probability (CP%) of the estimators are calculated using the standard bootstrap. The measurement error mechanism is assumed with $\sigma \in \{0.2, 0.5, 0.8\}$, which reflects a small, moderate or large measurement error on the true covariate X. Numerical results of n = 500 and n = 2000 are reported in Table 2.1 and Table 2.2, respectively. The estimates of $(\boldsymbol{\beta}, \boldsymbol{\psi})$ under $\sigma = 0.8$ are visualized in Figure 2.1 for n = 500 and Figure 2.2 for n = 2000, respectively.

Tables 2.1 and 2.2 show that ignoring the covariate measurement error leads to biased results with noticeable biases, and the coverage probabilities are below the nominal level of 95%. As the degree of measurement error increases, the biases are more severe. In contrast, the RC estimator presents a satisfactory performance in correcting for the effect with small biases and coverage probabilities around 95%. Its performance also seems robust against the various magnitude of measurement error. Moreover, we also see that the sample size affects the performance of the methods. As the sample size becomes larger, the associated variability decreases in all the scenarios.

2.3.2 Two-Stage Estimation in Linear Case

This simulation study aims to investigate the effect of measurement error on the parameter estimation in DTR with two decision points. Let $\mathrm{X}_j \sim \mathrm{N}(1, 1)$ and $\mathrm{Z}_j \sim \mathrm{N}(0.5, 1)$ be the error-prone and error-free covariates at stage j (j = 1, 2), respectively. A treatment $\mathrm{A}_j \in \{1, 0\}$ is assigned with probability $\mathrm{P}(\mathrm{A}_j = 1) = 1/2$. In practice, the number of replicate surrogates may vary from person to person. To mimic this situation, we consider a scenario with 3 replicate surrogates W_{j1}, W_{j2}, W_{j3}, generated by $W_{jl} = X_j + \mathrm{e}_{jl}$ (l = 1, 2, 3), where $\mathrm{e}_{jl} \sim \mathrm{N}(0, \sigma_j^2)$. The degree of measurement error at stage j is reflected by σ_j, which is assumed to be known or estimated from a pilot study. Each patient is assumed to possess at least W_{j1} and W_{j2} as primary proxies while W_{j3} may not be available. The degree of missingness in W_{j3} is set to be 80%. Let $\overline{W}_j$ be an average value of W_{j1}, W_{j2} and W_{j3}, $\overline{W}_j = (W_{j1} + W_{j2} + W_{j3})/3$. The outcome is modeled linearly in the treatment-free component as $\mathrm{Y} = \mathrm{X}_1 + \mathrm{Z}_1 + \mathrm{X}_2 + \mathrm{Z}_2 +$

$A_1(\psi_{10} + \psi_{11}X_1) + A_2(\psi_{20} + \psi_{21}X_2) + \epsilon$, where $\boldsymbol{\psi} = (\psi_{20}, \psi_{21}, \psi_{10}, \psi_{11}) = (0.5, -1, 0.5, -1)$ and $\epsilon \sim N(0, 1)$, independent of each other and of all other variables. In each round of 500 simulations, a dataset with the size of 2000 patients is generated.

Four estimators are compared in each stage to evaluate the performance of the RC method: (1) true estimator $\hat{\boldsymbol{\psi}}^{\boldsymbol{t}}$ obtained using the true covariate X_j, (2) naive estimator $\hat{\boldsymbol{\psi}}^{\boldsymbol{n}}$ obtained using a single surrogate W_{j1}, (3) naive estimator $\hat{\boldsymbol{\psi}}^{\boldsymbol{nb}}$ obtained using the averaged surrogate $\overline{W}_j$, (4) RC estimator $\hat{\boldsymbol{\psi}}^{\boldsymbol{rc}}$ obtained using the RC estimates $\hat{X}_j$. The degree of measurement error σ_j is specified as 0.2, 0.5 and 0.8. Results for the bias, SE, RMSE and CP% of $\hat{\boldsymbol{\psi}}$ computed using the standard bootstrap are reported in Table 2.3. Figures 2.3, 2.4 and 2.5 provide the visualized parameter estimates under $\sigma_2 = 0.8$.

Similar to the findings in one-stage estimation, both naive blip estimators $\hat{\boldsymbol{\psi}}^{\boldsymbol{n}}$ and $\hat{\boldsymbol{\psi}}^{\boldsymbol{nb}}$ are biased due to the ignorance of the covariate measurement error. As the degree of measurement error increases, the biases of the naive estimators exacerbate. On the contrary, the RC estimator $\hat{\boldsymbol{\psi}}^{\boldsymbol{rc}}$ yields small biases, and the coverage probabilities are close to the nominal level of 95%. Moreover, the performance of the RC estimator is also shown to be robust against the different magnitude of measurement error across the two stages.

2.3.3 Two-Stage Estimation in Nonlinear Case

In this section, we explore the measurement error effect on the estimation of blip parameters and optimal DTR in a nonlinear outcome model with two decision points. The data generation mechanism is the same with the one in (2.3.2) except that the outcome model is given by $Y = f(X_1) + Z_1 + f(X_2) + Z_2 + A_1(\psi_{10} + \psi_{11}X_1) + A_2(\psi_{20} + \psi_{21}X_2) + \epsilon$, where $\boldsymbol{\psi} = (\psi_{20}, \psi_{21}, \psi_{10}, \psi_{11}) = (0.5, -1, 0.5, -1)$ and $\epsilon \sim N(0, 1)$, independent of everything else. In the outcome model, three nonlinear functions are considered for X_j: (1) $f(X_j) = X_j + X_j^3$ (cubic), (2) $f(X_j) = X_j + e^{X_j}$ (exponential), (3) $f(X_j) = X_j + \sin(X_j^2) + \cos(X_j^2)$ (complex).

We continue the analysis with four estimators $\hat{\boldsymbol{\psi}}^{\boldsymbol{t}}$, $\hat{\boldsymbol{\psi}}^{\boldsymbol{n}}$, $\hat{\boldsymbol{\psi}}^{\boldsymbol{nb}}$ and $\hat{\boldsymbol{\psi}}^{\boldsymbol{rc}}$. The measurement error σ_j is chosen from a range of {0.2, 0.5, 0.8}. Table 2.4 displays the results for the bias, SE, RMSE and CP% of $\hat{\boldsymbol{\psi}}$ over various measurement errors in each nonlinear case. The blip estimates for three nonlinear examples under $\sigma = (\sigma_2, \sigma_1) = (0.8, 0.8)$ are visualized in Figures

2.6, 2.7 and 2.8. We repeat the set of simulations 500 times.

In general, the results in Table 2.4 reveal a larger measurement error effect with bigger biases and standard errors in the nonlinear case than those in the linear case. In comparison, the RC method remains effective and robust, though it produces slightly less reduced biases in this setting compared with that in (2.3.2), especially for the estimation of $\boldsymbol{\psi_1}$. Moreover, in the three scenarios, the models containing the cubic function and exponential function in the treatment-free component are sensitive to measurement error.

2.3.4 Predicted Optimal DTR

In this section, we explore the effect of measurement error on the predicted optimal treatment decision rules by evaluating the proportion of optimally treated patients across two stages. In Carroll et al. (2006), there is an argument about the necessity of modeling measurement error in a predictive setting. However, considering the importance of correctly identifying and recommending the optimal treatments to the patients, it's worth looking into the role of measurement error in predicting the optimal DTR in Q-learning.

The analysis follows the simulation design (2.3.2) and is done with the training data of 2000 patients and test data of 5000 patients. We first use the training data to produce three estimators ($\hat{\boldsymbol{\psi}}^{\boldsymbol{n}}$, $\hat{\boldsymbol{\psi}}^{\boldsymbol{nb}}$, $\hat{\boldsymbol{\psi}}^{\boldsymbol{rc}}$) with a single surrogate W_{j1}, averaged surrogate $\overline{W}_j$, and RC estimates $\hat{X}_j$, respectively. Then, we use the test data to find the prediction accuracy of optimal DTR, which is measured by the proportion of the patients whose optimal treatments are correctly identified in the test data at stage 2 and/or stage 1.

In each stage, six scenarios are considered to predict the optimal DTR using (1) naive estimator $\hat{\boldsymbol{\psi}}^{\boldsymbol{n}}$ and true covariate X_j (nt), (2) naive estimator $\hat{\boldsymbol{\psi}}^{\boldsymbol{nb}}$ and true covariate X_j (nbt), (3) RC estimator $\hat{\boldsymbol{\psi}}^{\boldsymbol{rc}}$ and true covariate X_j (ct), (4) naive estimator $\hat{\boldsymbol{\psi}}^{\boldsymbol{n}}$ and a single surrogate W_{j1} (nn), (5) naive estimator $\hat{\boldsymbol{\psi}}^{\boldsymbol{nb}}$ and averaged surrogate $\overline{W}_j$ (nbnb), (6) RC estimator $\hat{\boldsymbol{\psi}}^{\boldsymbol{rc}}$ and RC estimates $\hat{X}_j$ (cc), respectively. The first three scenarios aim to examine the measurement error effect on the prediction accuracy using the true covariates in the test data, while the last three evaluate the measurement error effect using the surrogates and corrected estimates in the test data. A total of 500 runs are simulated for each scenario. Numerical results are summarized in

Table 2.5. The prediction accuracy results under (σ_2, σ_1) = (0.8, 0.8) are shown in Figure 2.9.

Table 2.5 shows that the existence of measurement error leads to a remarkable degradation of the prediction accuracy of optimal DTR, and it achieves the lowest prediction accuracy when a single surrogate is used. However, the RC method outperforms the naive method and significantly improves the prediction accuracy in all the scenarios. In the last two scenarios (nbnb and cc), the RC method yields similar accuracy results to the naive method, indicating that the worst scenario of using the RC method is comparable to that of using the naive method.

2.3.5 Predicted Optimal Value Function

In this study, we evaluate the measurement error effect on the predicted optimal value function, which is the expected outcome under the optimal treatment regimes. The data generation mechanism follows (2.3.4), and we continue with the three estimators ($\hat{\boldsymbol{\psi}}^{\boldsymbol{n}}$, $\hat{\boldsymbol{\psi}}^{\boldsymbol{nb}}$, $\hat{\boldsymbol{\psi}}^{\boldsymbol{rc}}$) obtained from the training data. We use the test data to predict the value functions under the (1) true optimal DTR (opt), (2) optimal DTR estimated using $\hat{\boldsymbol{\psi}}^{\boldsymbol{n}}$ and X_j (nt), (3) optimal DTR estimated using $\hat{\boldsymbol{\psi}}^{\boldsymbol{nb}}$ and X_j (nbt), (4) optimal DTR estimated using $\hat{\boldsymbol{\psi}}^{\boldsymbol{rc}}$ and X_j (ct), (5) optimal DTR estimated using $\hat{\boldsymbol{\psi}}^{\boldsymbol{n}}$ and W_{j1} (nn), (6) optimal DTR estimated using $\hat{\boldsymbol{\psi}}^{\boldsymbol{nb}}$ and $\overline{W}_j$ (nbnb), (7) optimal DTR estimated using $\hat{\boldsymbol{\psi}}^{\boldsymbol{rc}}$ and $\hat{X}_j$ (cc). Simulations are repeated 500 times. For each scenario, the average value function is computed and reported in Table 2.6, along with its standard deviations. Figure 2.10 also displayed the predicted optimal value function under (σ_2, σ_1) = (0.8, 0.8).

In Table 2.6, we see that the measurement error effect is pronounced in terms of value function estimation under the optimal DTR. By comparison, the naive method generally yields lower value function estimates, and the optimal value function achieves the lowest value with a single surrogate being used, as expected. The RC method, however, improves the estimated optimal value function, even comparable to the true optimal value function when the true covariate is used.

2.4 Application to STAR*D Study

To illustrate the proposed correction method, we analyze the data arising from the Sequenced Treatment Alternatives to Relieve Depression study (Rush et al., 2003,0). The STAR*D study was designed as a multisite, multistage randomized controlled trial. It aimed to evaluate the effect of treatments for patients who suffered from major depressive disorder. The severity of depressive disorder was measured by the Quick Inventory of Depressive Symptomatology score, which was assessed by both patients (QIDS-S) and clinicians (QIDS-C). The entire study possessed four levels, in which one or a combination of treatments was assigned to the patients. At level 1, all of the patients were prescribed citalopram (CIT). At the end of level 1, if patients had QIDS ≤ 5, they achieved remission and were removed from the study but those who otherwise entered level 2. They were again randomized into one of the seven treatment options: either switching from CIT to one of four other treatment options (venlafaxine[VEN], sertraline[SER], bupropion[BUP], and cognitive therapy[CT]) or augmenting CIT with one of three treatments (BUP, CT and buspirone[BUS]). Then, at the end of level 2, they were again assessed with the QIDS score, and those who failed to achieve remission (QIDS ≤ 5) entered level 3. In level 3, they were randomized to receive either one of two new treatments (lithium[Li] or thyroid hormone[THY]) or one of two augmented treatment options (mirtazapine[MIRT], nortriptyline[NTP]). The QIDS score for remission was evaluated at the end of level 3.

In the literature, depression is found to be significantly associated with functional impairment (Greer et al., 2010). Patients with major depressive disorder were shown to have considerable deficits in the physical and social functioning (Lin et al., 2014; Trivedi et al., 2013). IsHak et al. (2016) analyzed the STAR*D data and pointed out the importance of developing individualized treatments for patients with a major depressive disorder to improve their long-term functioning. The perceived functional impairment is measured at each level of the STAR*D study by the Work and Social Adjustment Scale (WSAS) score, which reflects the functioning aspects of the work, home management, social activities, private activities, and relationships with others.

We follow the criteria in the literature (Chakraborty, 2013; Chakraborty et al., 2013; Wal-

lace et al., 2019) to select the data, where the two-stage DTR is considered by combining level 2 and level 2A as the first stage and treating level 3 as the second stage. The stage j treatment A_j is coded based on whether the treatment involves selective serotonin reuptake inhibitor ($A_j = 1$) or not ($A_j = 0$). Three tailoring variables are considered, Q_j: the QIDS-C score measured at the beginning of each level j, S_j: the QIDS slope, the change in QIDS-C divided by the time in the previous level, and P_j: the patients' preference indicating whether they wished to switch previous treatment ($P_j = 1$), to augment previous treatment or have no preference ($P_j = 0$). The outcome of interest is defined as the negative WSAS score across two stages

$$Y = R_1 \cdot Y_1 + (1 - R_1) \cdot \tfrac{1}{2}(Y_1 + Y_2),$$

where Y_1 and Y_2 are the negative WSAS scores observed at the end of stage 1 and stage 2, and R_1 is an indicator of whether the patients achieved remission ($R_1 = 1$) or not ($R_1 = 0$) at the end of stage 1. The selected data contain 1438 patients at stage 1, of whom 377 patients have entered the stage 2.

The previous analyses of the STAR*D data often assume that the QIDS-C score is error-free, which is usually not the case in practice. Spicker and Wallace (2020) studied the measurement error effect on sequential optimal treatment rules, assuming that the true QIDS score was unknown and both the QIDS-C score and QIDS-S score were considered as the repeated measurements of the true underlying QIDS score. In this work, we are interested in estimating the optimal treatment decision rules using Q-learning that maximize the negative WASA score, provided that the QIDS score is subject to measurement error. We compare three estimators, including two naive estimators using the QIDS-C score or QIDS-S score as the tailoring variable and the RC estimator using the corrected estimates computed by (QIDS-C, QIDS-S). The analysis results of the parameter estimates, bootstrap standard error, and 95% confidence interval are summarized in Table 2.7.

In Table 2.7, the parameter estimates of each stage vary remarkably between the naive method and the RC method, leading to different optimal treatment decision rules. More importantly, the results show that the significance of the tailoring variable differs between these two methods. The patients' preference to switch treatment and QIDS score have significant treatment effects in the interaction with the second stage treatment when the QIDS-S score is used.

However, by using the RC estimates, no significant term is observed across the two stages. It emphasizes that the measurement error effect is not negligible in an error-prone setting since it is possible to alter the estimation of optimal treatment decision rules and the significance of the tailoring variable.

2.5 Conclusion

This study aims to build a bridge between Q-learning with continuous outcomes and covariate measurement error where there exist replicated measurements for the error-prone covariates. It is demonstrated in both simulation studies and data analysis that ignoring measurement error in covariates will lead to severely biased results. To adjust for the measurement error effect, we apply the regression calibration method in Q-learning and present a modified Q-learning algorithm. On average, the RC method shows superior performance over the naive method in all the scenarios in terms of bias-reduction and coverage probability, especially in the linear Q-learning setting. Moreover, the RC method is generally robust against the magnitude of measurement error.

Another important topic discussed in the study is evaluating the performance of the proposed method from a predictive perspective. We predict the future optimal treatment decision rules by finding the proportion of patients whose optimal treatments are correctly identified across two stages. It turns out that using a single mismeasured covariate leads to the worst performance among all methods. In contrast, the RC method improves the prediction accuracy even when the degree of measurement error is high. Moreover, we also compare the naive method and RC method in terms of value function estimation. The optimal value function estimated from the naive method is generally lower, but the RC method enhances the optimal value function comparable to the true optimal value function.

Table 2.1: One-stage estimates of blip parameters (ψ_{10}, ψ_{11}) (n = 500)

		ψ_{10}				ψ_{11}			
σ	$\hat{\boldsymbol{\psi}}$	Bias	SE	RMSE	CP%	Bias	SE	RMSE	CP%
	$\hat{\boldsymbol{\psi}}^{t}$	0.002	0.133	0.133	96.8	-0.002	0.095	0.095	95.2
0.2	$\hat{\boldsymbol{\psi}}^{n}$	0.043	0.137	0.143	94.4	-0.043	0.096	0.106	93.6
	$\hat{\boldsymbol{\psi}}^{nb}$	0.022	0.134	0.136	95.4	-0.023	0.095	0.098	95.0
	$\hat{\boldsymbol{\psi}}^{rc}$	0.002	0.137	0.137	96.2	-0.003	0.097	0.097	95.8
0.5	$\hat{\boldsymbol{\psi}}^{n}$	0.209	0.149	0.257	70.4	-0.208	0.100	0.231	46.8
	$\hat{\boldsymbol{\psi}}^{nb}$	0.118	0.143	0.185	86.4	-0.117	0.098	0.153	74.8
	$\hat{\boldsymbol{\psi}}^{rc}$	0.008	0.151	0.152	94.4	-0.007	0.110	0.110	95.6
0.8	$\hat{\boldsymbol{\psi}}^{n}$	0.406	0.160	0.436	29.6	-0.403	0.099	0.414	2.4
	$\hat{\boldsymbol{\psi}}^{nb}$	0.254	0.152	0.296	59.8	-0.254	0.100	0.273	27.6
	$\hat{\boldsymbol{\psi}}^{rc}$	0.012	0.176	0.176	94.8	-0.012	0.132	0.133	94.8

Table 2.2: One-stage estimates of blip parameters (ψ_{10}, ψ_{11}) (n = 2000)

		ψ_{10}				ψ_{11}			
σ	$\hat{\boldsymbol{\psi}}$	Bias	SE	RMSE	CP%	Bias	SE	RMSE	CP%
	$\hat{\boldsymbol{\psi}}^{t}$	-0.002	0.066	0.066	95.4	0.004	0.047	0.047	96.6
0.2	$\hat{\boldsymbol{\psi}}^{n}$	0.035	0.068	0.077	93.4	-0.034	0.048	0.059	88.6
	$\hat{\boldsymbol{\psi}}^{nb}$	0.016	0.068	0.070	94.6	-0.015	0.048	0.05	95.4
	$\hat{\boldsymbol{\psi}}^{rc}$	-0.003	0.068	0.068	94.8	0.004	0.048	0.049	96.4
0.5	$\hat{\boldsymbol{\psi}}^{n}$	0.203	0.075	0.216	21.0	-0.200	0.050	0.206	2.2
	$\hat{\boldsymbol{\psi}}^{nb}$	0.115	0.072	0.135	65.8	-0.112	0.049	0.122	38.6
	$\hat{\boldsymbol{\psi}}^{rc}$	0.004	0.076	0.076	94.0	-0.001	0.055	0.055	95.6
0.8	$\hat{\boldsymbol{\psi}}^{n}$	0.388	0.080	0.396	0.2	-0.389	0.049	0.392	0.0
	$\hat{\boldsymbol{\psi}}^{nb}$	0.241	0.076	0.252	11.8	-0.241	0.050	0.246	0.0
	$\hat{\boldsymbol{\psi}}^{rc}$	-0.002	0.088	0.088	94.4	0.002	0.066	0.066	93.0

Table 2.3: Two-stage estimates of blip parameters ($\psi_{20}, \psi_{21}, \psi_{10}, \psi_{11}$) in linear case

			ψ_{20}				ψ_{21}				ψ_{10}				ψ_{11}			
σ_2	σ_1	$\hat{\psi}$	Bias	SE	RMSE	CP%	Bias	SE	RMSE	CP%	Bias	SE	RMSE	CP%	Bias	SE	RMSE	CP%
		$\hat{\psi}^t$	0.000	0.072	0.072	97.2	0.001	0.052	0.052	97.8	-0.007	0.105	0.105	96.2	0.003	0.075	0.075	96.0
0.2	0.2	$\hat{\psi}^n$	-0.039	0.072	0.082	94.2	0.039	0.051	0.064	90.0	-0.045	0.104	0.114	91.8	0.041	0.073	0.084	93.0
		$\hat{\psi}^{nb}$	-0.018	0.072	0.074	96.4	0.019	0.051	0.055	95.8	-0.025	0.105	0.108	94.8	0.022	0.074	0.077	96.2
		$\hat{\psi}^{rc}$	0.000	0.072	0.072	97.0	0.000	0.052	0.052	97.8	-0.007	0.106	0.106	95.8	0.004	0.075	0.076	96.2
	0.5	$\hat{\psi}^n$	-0.039	0.073	0.083	94.4	0.040	0.052	0.065	90.6	-0.206	0.101	0.229	45.8	0.202	0.067	0.213	12.0
		$\hat{\psi}^{nb}$	-0.018	0.073	0.075	95.8	0.019	0.052	0.055	95.6	-0.111	0.103	0.151	79.8	0.108	0.071	0.129	68.4
		$\hat{\psi}^{rc}$	0.000	0.073	0.073	96.4	0.000	0.053	0.053	97.8	-0.007	0.109	0.109	95.6	0.004	0.079	0.079	96.4
	0.8	$\hat{\psi}^n$	-0.040	0.075	0.085	93.4	0.040	0.054	0.067	90.8	-0.396	0.096	0.407	0.8	0.392	0.059	0.396	0.0
		$\hat{\psi}^{nb}$	-0.018	0.074	0.077	95.6	0.019	0.053	0.056	96.0	-0.236	0.100	0.257	33.0	0.233	0.066	0.242	5.8
		$\hat{\psi}^{rc}$	0.000	0.075	0.075	96.4	0.000	0.054	0.054	97.6	-0.007	0.114	0.115	95.2	0.004	0.086	0.086	95.8
0.5	0.2	$\hat{\psi}^n$	-0.203	0.070	0.214	17.2	0.201	0.047	0.206	0.4	-0.037	0.105	0.111	94.0	0.038	0.074	0.083	93.4
		$\hat{\psi}^{nb}$	-0.106	0.071	0.128	68.2	0.104	0.049	0.116	45.0	-0.016	0.105	0.107	95.6	0.017	0.075	0.077	95.6
		$\hat{\psi}^{rc}$	-0.002	0.075	0.075	97.0	0.000	0.055	0.055	96.4	0.002	0.106	0.106	95.4	-0.001	0.076	0.076	95.4
	0.5	$\hat{\psi}^n$	-0.203	0.072	0.216	18.2	0.201	0.048	0.207	0.4	-0.199	0.101	0.223	50.6	0.199	0.068	0.211	15.6
		$\hat{\psi}^{nb}$	-0.107	0.072	0.128	68.4	0.105	0.050	0.116	45.8	-0.102	0.103	0.145	83.0	0.103	0.072	0.126	70.6
		$\hat{\psi}^{rc}$	-0.002	0.076	0.076	96.6	0.000	0.056	0.056	96.6	0.002	0.109	0.109	95.4	-0.001	0.080	0.080	96.2
	0.8	$\hat{\psi}^n$	-0.204	0.074	0.217	20.0	0.201	0.050	0.207	0.6	-0.389	0.096	0.401	1.6	0.390	0.059	0.394	0.0
		$\hat{\psi}^{nb}$	-0.107	0.073	0.130	68.4	0.105	0.051	0.117	46.4	-0.228	0.100	0.249	36.8	0.229	0.067	0.238	7.6
		$\hat{\psi}^{rc}$	-0.002	0.077	0.077	96.4	0.000	0.057	0.057	96.6	0.003	0.115	0.115	95.4	-0.002	0.086	0.086	95.8
0.8	0.2	$\hat{\psi}^n$	-0.389	0.067	0.394	0.0	0.391	0.042	0.393	0.0	-0.041	0.105	0.113	92.8	0.039	0.074	0.083	93.0
		$\hat{\psi}^{nb}$	-0.228	0.069	0.238	7.0	0.229	0.046	0.234	0.0	-0.021	0.106	0.108	95.2	0.019	0.075	0.077	95.4
		$\hat{\psi}^{rc}$	0.002	0.079	0.079	97.0	-0.001	0.060	0.060	96.6	-0.003	0.106	0.106	95.8	0.001	0.076	0.076	95.0
	0.5	$\hat{\psi}^n$	-0.389	0.069	0.395	0.0	0.390	0.043	0.393	0.0	-0.202	0.102	0.226	49.0	0.199	0.068	0.211	14.8
		$\hat{\psi}^{nb}$	-0.229	0.070	0.239	7.4	0.229	0.047	0.234	0.0	-0.108	0.104	0.149	84.6	0.106	0.072	0.128	70.0
		$\hat{\psi}^{rc}$	0.002	0.080	0.080	97.0	-0.001	0.061	0.061	96.4	-0.003	0.109	0.109	95.0	0.001	0.080	0.080	95.8
	0.8	$\hat{\psi}^n$	-0.390	0.071	0.396	0.0	0.390	0.044	0.393	0.0	-0.392	0.097	0.403	1.2	0.389	0.060	0.393	0.0
		$\hat{\psi}^{nb}$	-0.229	0.072	0.240	9.0	0.229	0.048	0.234	0.0	-0.233	0.101	0.254	36.0	0.231	0.067	0.241	7.0
		$\hat{\psi}^{rc}$	0.001	0.082	0.082	97.0	-0.001	0.062	0.062	96.2	-0.002	0.115	0.115	94.2	0.001	0.086	0.086	96.0

Table 2.4: Two-stage estimates of blip parameters (ψ_{20}, ψ_{21}, ψ_{10}, ψ_{11}) in nonlinear case

			ψ_{20}				ψ_{21}				ψ_{10}				ψ_{11}			
Scenario	(σ_2, σ_1)	$\hat{\psi}$	Bias	SE	RMSE	CP%	Bias	SE	RMSE	CP%	Bias	SE	RMSE	CP%	Bias	SE	RMSE	CP%
Cub		$\hat{\psi}^t$	-0.003	0.068	0.068	96.2	0.001	0.049	0.049	96.4	0.021	0.538	0.538	95.0	-0.002	0.381	0.381	93.6
	(0.2, 0.2)	$\hat{\psi}^n$	-0.050	0.164	0.172	93.2	0.043	0.175	0.180	94.0	-0.013	0.542	0.542	95.4	0.033	0.403	0.404	94.6
		$\hat{\psi}^{nb}$	-0.027	0.125	0.128	94.4	0.019	0.128	0.130	94.2	0.005	0.539	0.539	95.6	0.016	0.393	0.393	94.6
		$\hat{\psi}^{rc}$	-0.009	0.125	0.125	95.8	0.000	0.130	0.130	94.4	0.023	0.545	0.546	95.8	-0.002	0.400	0.400	95.4
	(0.5, 0.5)	$\hat{\psi}^n$	-0.214	0.316	0.382	89.4	0.197	0.334	0.387	89.0	-0.197	0.538	0.572	94.0	0.226	0.445	0.499	91.2
		$\hat{\psi}^{nb}$	-0.117	0.246	0.272	93.4	0.105	0.267	0.287	92.8	-0.098	0.544	0.553	95.6	0.128	0.429	0.447	94.2
		$\hat{\psi}^{rc}$	-0.010	0.263	0.263	95.2	-0.001	0.296	0.296	93.2	0.005	0.574	0.574	95.2	0.025	0.475	0.475	95.4
	(0.8, 0.8)	$\hat{\psi}^n$	-0.377	0.389	0.542	84.0	0.368	0.372	0.524	83.8	-0.401	0.522	0.659	87.4	0.388	0.432	0.581	83.2
		$\hat{\psi}^{nb}$	-0.234	0.333	0.407	89.6	0.216	0.347	0.409	90.4	-0.257	0.540	0.598	94.2	0.249	0.448	0.513	89.4
		$\hat{\psi}^{rc}$	0.007	0.394	0.394	94.8	-0.022	0.444	0.444	94.8	-0.029	0.625	0.626	96.4	0.022	0.575	0.575	95.4
Exp		$\hat{\psi}^t$	-0.003	0.068	0.068	96.0	0.001	0.049	0.049	96.6	0.016	0.420	0.420	95.8	-0.004	0.298	0.298	94.0
	(0.2, 0.2)	$\hat{\psi}^n$	-0.047	0.138	0.146	92.6	0.043	0.149	0.155	92.8	-0.018	0.424	0.425	95.4	0.031	0.319	0.321	95.0
		$\hat{\psi}^{nb}$	-0.025	0.108	0.111	95.0	0.018	0.110	0.112	93.8	0.000	0.422	0.422	95.2	0.013	0.310	0.310	94.2
		$\hat{\psi}^{rc}$	-0.006	0.109	0.109	96.6	0.000	0.112	0.112	93.8	0.019	0.426	0.426	95.2	-0.006	0.315	0.315	94.0
	(0.5, 0.5)	$\hat{\psi}^n$	-0.210	0.253	0.329	84.8	0.202	0.277	0.342	85.6	-0.195	0.423	0.466	92.2	0.215	0.359	0.418	90.6
		$\hat{\psi}^{nb}$	-0.115	0.200	0.230	90.6	0.109	0.223	0.248	90.6	-0.096	0.429	0.439	94.8	0.115	0.344	0.362	94.2
		$\hat{\psi}^{rc}$	-0.009	0.216	0.217	94.6	0.003	0.247	0.247	93.4	0.008	0.453	0.453	95.0	0.013	0.380	0.380	94.6
	(0.8, 0.8)	$\hat{\psi}^n$	-0.382	0.302	0.487	74.8	0.373	0.302	0.479	73.6	-0.398	0.407	0.569	84.0	0.386	0.349	0.520	75.4
		$\hat{\psi}^{nb}$	-0.228	0.264	0.349	86.8	0.213	0.285	0.356	87.6	-0.250	0.424	0.492	91.6	0.241	0.362	0.435	87.6
		$\hat{\psi}^{rc}$	0.009	0.321	0.321	95.8	-0.021	0.367	0.367	94.8	-0.021	0.496	0.497	96.6	0.014	0.466	0.466	95.2
Com		$\hat{\psi}^t$	-0.004	0.070	0.07	97.0	0.001	0.051	0.051	97.2	0.000	0.103	0.103	96.2	0.002	0.074	0.074	95.4
	(0.2, 0.2)	$\hat{\psi}^n$	-0.044	0.076	0.088	91.6	0.041	0.060	0.072	89.8	-0.036	0.104	0.110	94.8	0.037	0.078	0.086	92.6
		$\hat{\psi}^{nb}$	-0.022	0.073	0.077	94.4	0.019	0.056	0.059	94.0	-0.017	0.104	0.105	95.8	0.019	0.076	0.079	95.0
		$\hat{\psi}^{rc}$	-0.004	0.074	0.074	96.0	0.001	0.057	0.057	96.6	0.001	0.105	0.105	96.6	0.001	0.078	0.078	95.6
	(0.5, 0.5)	$\hat{\psi}^n$	-0.200	0.083	0.216	33.4	0.201	0.062	0.210	12.4	-0.207	0.104	0.232	48.2	0.204	0.075	0.217	21.0
		$\hat{\psi}^{nb}$	-0.103	0.081	0.131	74.6	0.104	0.064	0.122	62.8	-0.113	0.104	0.153	84.0	0.109	0.078	0.134	71.6
		$\hat{\psi}^{rc}$	0.001	0.085	0.085	95.4	-0.001	0.069	0.069	94.8	-0.009	0.110	0.110	96.0	0.006	0.086	0.086	95.4
	(0.8, 0.8)	$\hat{\psi}^n$	-0.391	0.083	0.400	0.2	0.388	0.056	0.392	0.0	-0.396	0.101	0.409	2.8	0.393	0.066	0.399	0.0
		$\hat{\psi}^{nb}$	-0.231	0.083	0.245	21.4	0.228	0.062	0.236	3.0	-0.235	0.104	0.257	38.4	0.232	0.074	0.243	12.0
		$\hat{\psi}^{rc}$	0.001	0.095	0.095	94.4	-0.005	0.079	0.079	95.6	-0.005	0.118	0.118	95.6	0.003	0.095	0.095	95.2

Cub: cubic, exp: exponential, com: complex

Table 2.5: Prediction accuracy of optimal DTR (%)

		Stage 2						Stage 1						Stage 2 & Stage 1					
σ_2	σ_1	nt	nbt	ct	nn	nbnb	cc	nt	nbt	ct	nn	nbnb	cc	nt	nbt	ct	nn	nbnb	cc
0.2	0.2	98.3	98.4	98.5	94.3	95.9	95.9	97.5	97.6	97.7	94.0	95.5	95.5	95.9	96.1	96.2	88.6	91.6	91.6
	0.5	98.3	98.4	98.5	94.2	95.9	95.9	95.2	96.8	97.7	86.9	90.5	90.5	93.6	95.3	96.1	81.9	86.8	86.8
	0.8	98.2	98.4	98.4	94.2	95.9	95.9	89.6	94.5	97.6	81.3	85.9	85.9	88.0	93.0	96.0	76.6	82.4	82.4
0.5	0.2	95.6	97.5	98.4	87.0	90.7	90.7	97.6	97.6	97.7	94.0	95.5	95.5	93.3	95.2	96.1	81.8	86.6	86.6
	0.5	95.6	97.5	98.4	87.0	90.7	90.7	95.4	97.0	97.6	86.9	90.5	90.5	91.2	94.6	96.1	75.6	82.1	82.1
	0.8	95.5	97.5	98.4	87.0	90.7	90.7	89.9	94.8	97.6	81.3	85.9	85.9	85.9	92.4	96.0	70.7	77.9	77.9
0.8	0.2	89.9	95.0	98.5	81.4	86.1	86.1	97.6	97.6	97.7	94.0	95.5	95.5	87.7	92.8	96.2	76.5	82.2	82.2
	0.5	89.8	95.0	98.5	81.4	86.1	86.1	95.3	96.9	97.6	86.8	90.5	90.5	85.7	92.0	96.2	70.7	77.9	77.9
	0.8	89.8	94.9	98.4	81.4	86.0	86.0	89.7	94.6	97.6	81.2	85.9	85.9	80.6	89.8	96.0	66.1	73.9	73.9

Table 2.6: Predicted optimal value function (standard deviations)

σ_2	σ_1	opt	nt	nbt	ct	nn	nbnb	cc
0.2	0.2	3.393 (0.028)	3.366 (0.050)	3.380 (0.050)	3.393 (0.051)	3.380 (0.051)	3.387 (0.051)	3.387 (0.051)
	0.5	3.393 (0.028)	3.312 (0.049)	3.351 (0.050)	3.393 (0.052)	3.350 (0.051)	3.371 (0.051)	3.371 (0.051)
	0.8	3.393 (0.028)	3.255 (0.049)	3.310 (0.050)	3.393 (0.055)	3.313 (0.052)	3.348 (0.051)	3.348 (0.053)
0.5	0.2	3.392 (0.029)	3.313 (0.052)	3.352 (0.052)	3.395 (0.053)	3.351 (0.053)	3.373 (0.052)	3.373 (0.053)
	0.5	3.392 (0.029)	3.258 (0.051)	3.322 (0.052)	3.395 (0.055)	3.321 (0.054)	3.357 (0.053)	3.357 (0.054)
	0.8	3.392 (0.029)	3.200 (0.051)	3.281 (0.051)	3.395 (0.057)	3.284 (0.054)	3.334 (0.053)	3.334 (0.055)
0.8	0.2	3.396 (0.028)	3.256 (0.049)	3.311 (0.048)	3.396 (0.049)	3.315 (0.049)	3.351 (0.048)	3.351 (0.049)
	0.5	3.396 (0.028)	3.202 (0.049)	3.282 (0.047)	3.396 (0.050)	3.286 (0.051)	3.335 (0.048)	3.336 (0.050)
	0.8	3.396 (0.028)	3.145 (0.050)	3.241 (0.047)	3.396 (0.052)	3.250 (0.052)	3.312 (0.049)	3.312 (0.051)

Table 2.7: Analysis results of the STAR*D data for the blip parameters

	Clinicians			Patients			Corrected		
Variables	Est	SE	95%CI	Est	SE	95%CI	Est	SE	95%CI
A_2	-0.330	0.455	(-1.221, 0.561)	-0.436	0.367	(-1.155, 0.284)	-0.351	0.396	(-1.128, 0.426)
A_2P_2	0.243	0.436	(-0.611, 1.098)	0.620	0.252	(0.127, 1.113)	0.404	0.271	(-0.128, 0.935)
A_2S_2	-0.205	1.260	(-2.674, 2.264)	-1.011	0.979	(-2.931, 0.908)	-0.539	1.316	(-3.119, 2.041)
A_2Q_2	0.441	0.765	(-1.060, 1.941)	0.936	0.459	(0.036, 1.836)	0.702	0.644	(-0.561, 1.965)
A_1	-0.194	0.159	(-0.506, 0.117)	-0.084	0.167	(-0.411, 0.244)	-0.150	0.157	(-0.458, 0.159)
A_1P_1	0.186	0.183	(-0.173, 0.545)	0.089	0.187	(-0.278, 0.457)	0.148	0.178	(-0.201, 0.497)
A_1S_1	0.062	0.091	(-0.116, 0.241)	0.069	0.093	(-0.114, 0.252)	0.090	0.091	(-0.088, 0.268)
A_1Q_1	0.048	0.066	(-0.081, 0.178)	-0.064	0.084	(-0.229, 0.101)	-0.014	0.070	(-0.152, 0.124)

Est: estimates, SE: standard error, CI: confidence interval

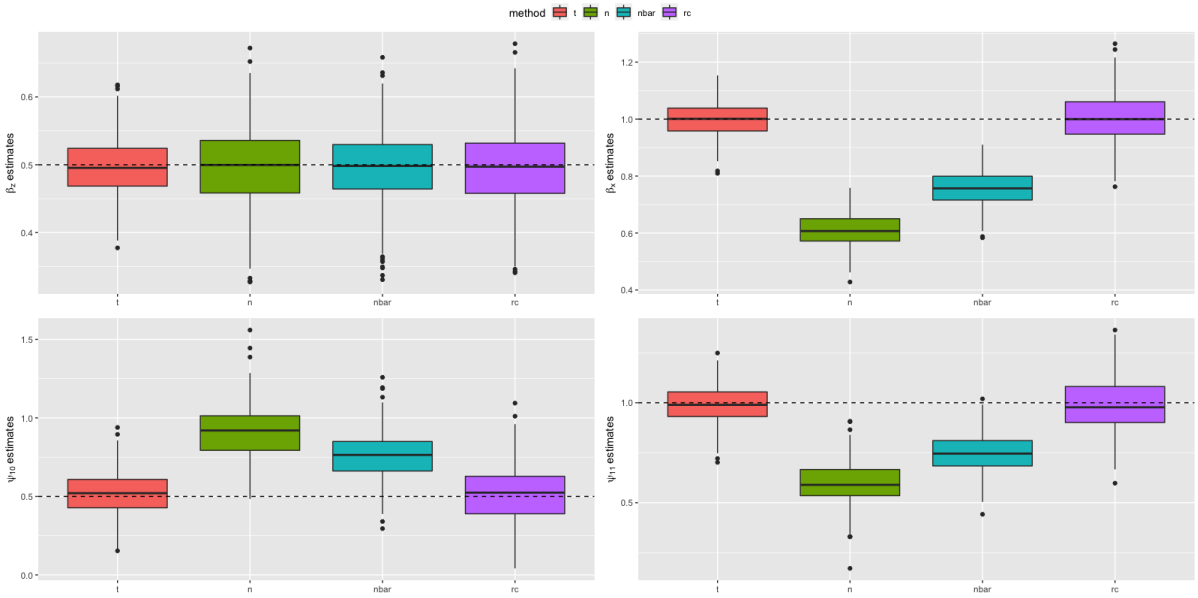

Figure 2.1: One-stage estimates of $(\beta_z, \beta_x, \psi_{10}, \psi_{11})$ with $\sigma = 0.8$ (n = 500)

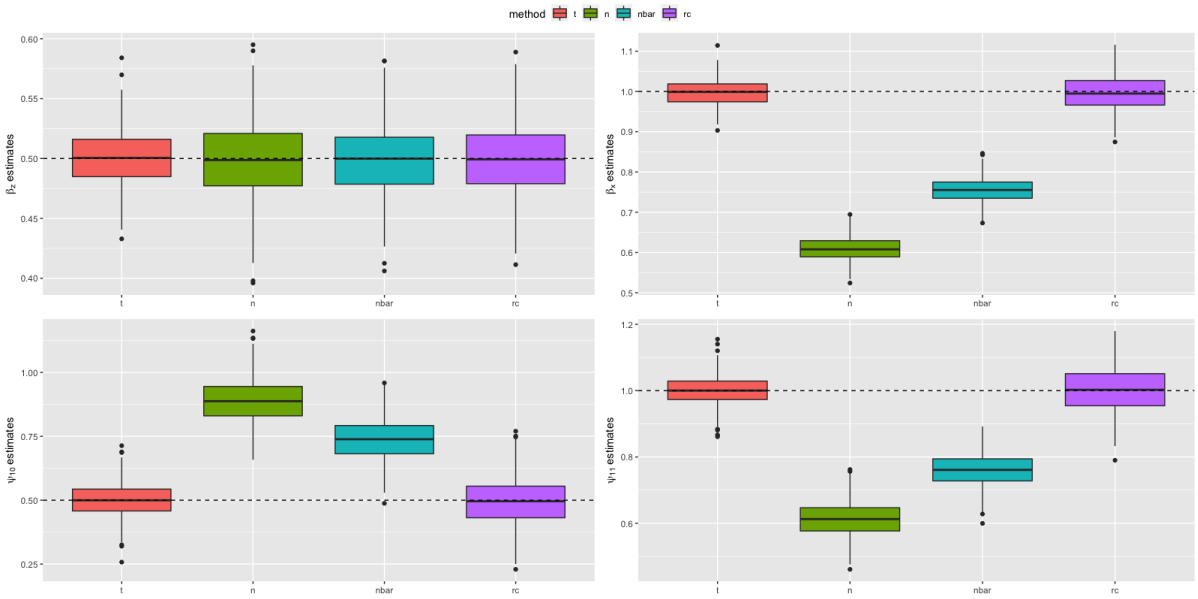

Figure 2.2: One-stage estimates of $(\beta_z, \beta_x, \psi_{10}, \psi_{11})$ with $\sigma = 0.8$ (n = 2000)

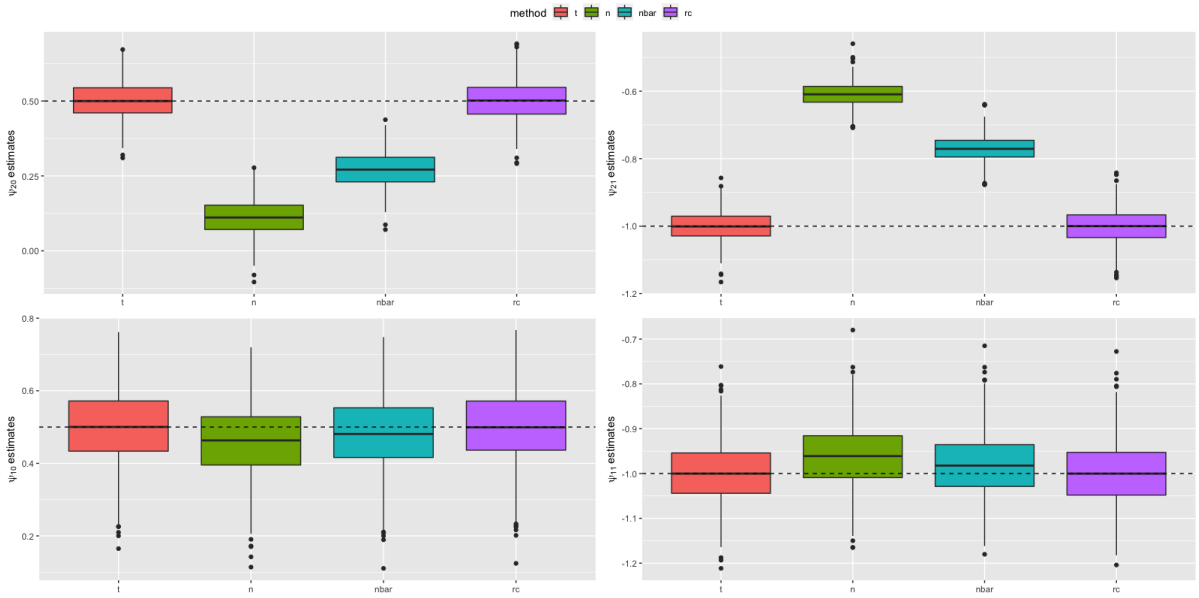

Figure 2.3: Two-stage estimates of $(\psi_{20}, \psi_{21}, \psi_{10}, \psi_{11})$ with $(\sigma_2, \sigma_1) = (0.8, 0.2)$

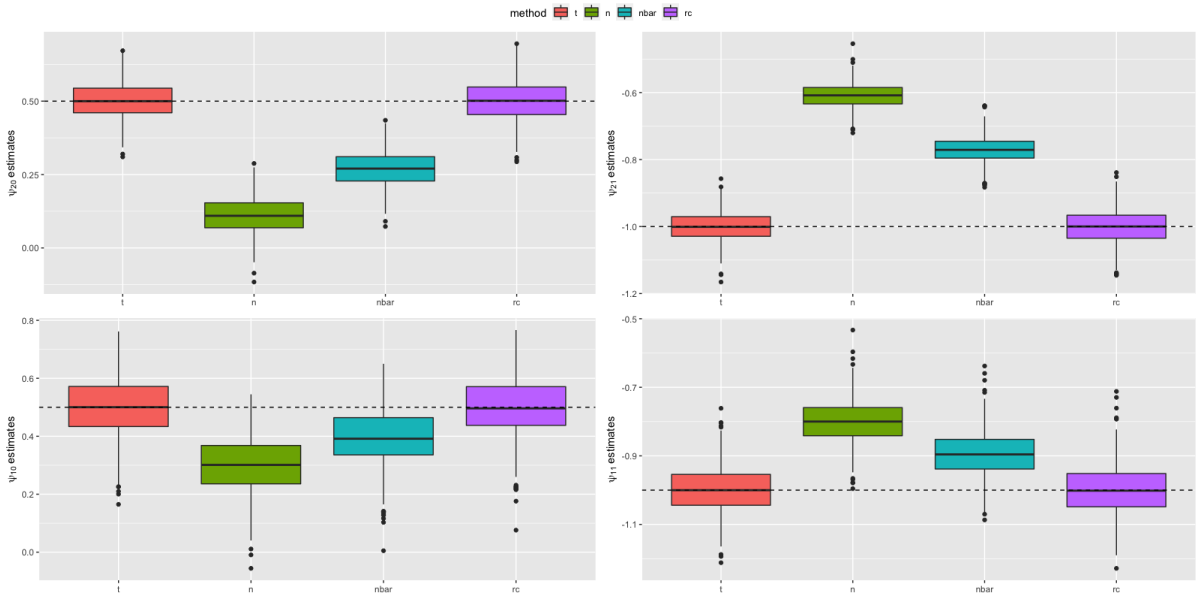

Figure 2.4: Two-stage estimates of $(\psi_{20}, \psi_{21}, \psi_{10}, \psi_{11})$ with $(\sigma_2, \sigma_1) = (0.8, 0.5)$

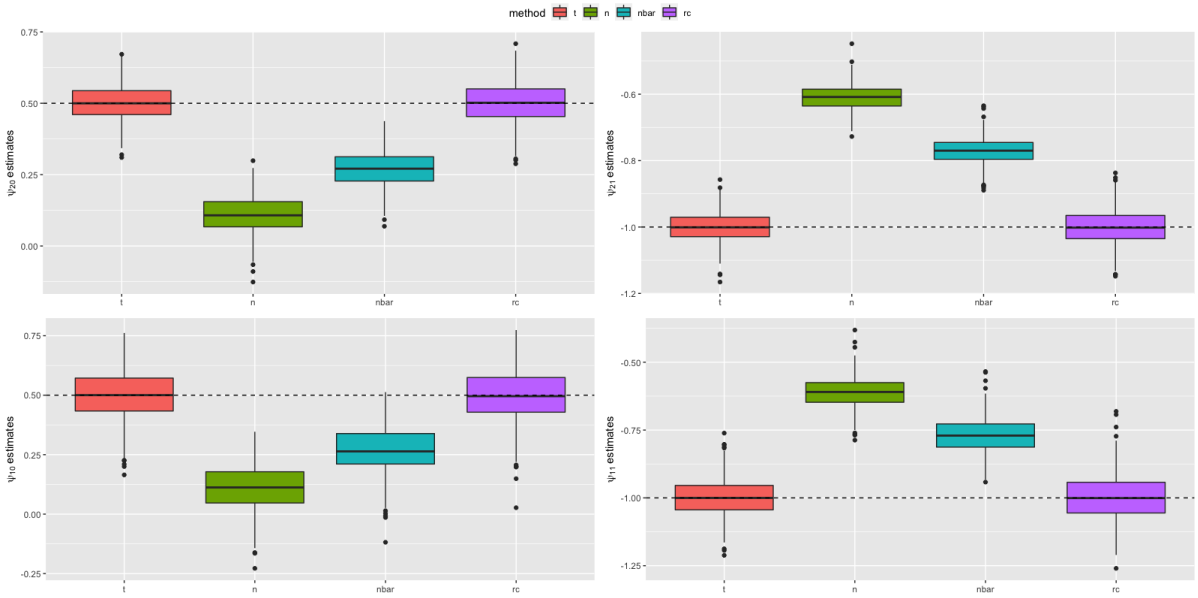

Figure 2.5: Two-stage estimates of $(\psi_{20}, \psi_{21}, \psi_{10}, \psi_{11})$ with $(\sigma_2, \sigma_1) = (0.8, 0.8)$

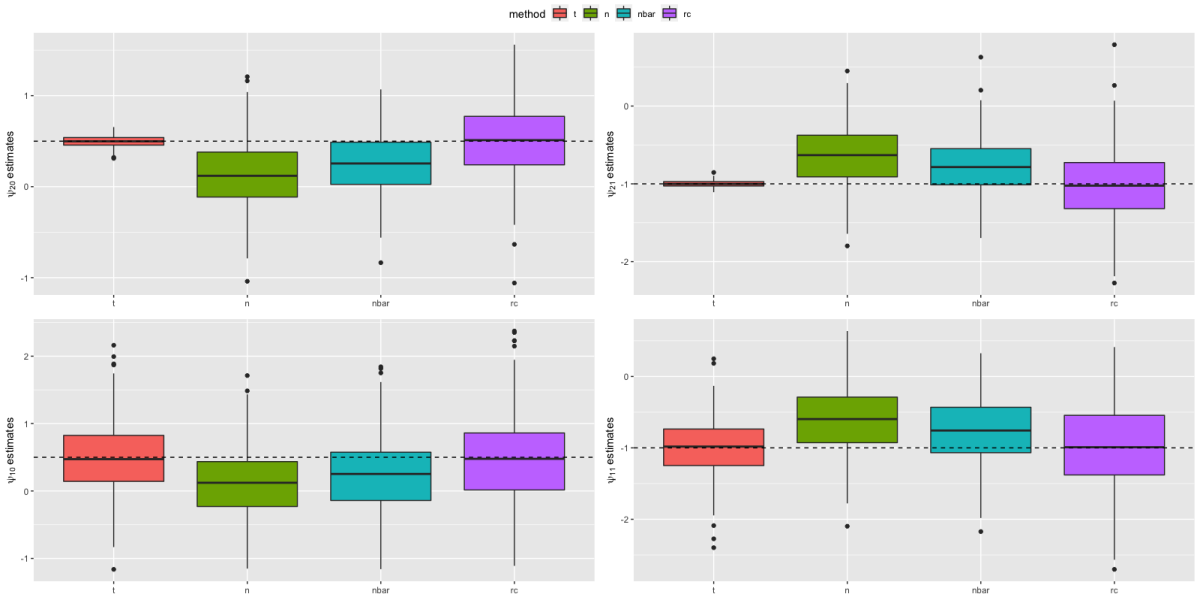

Figure 2.6: Two-stage estimates of $(\psi_{20}, \psi_{21}, \psi_{10}, \psi_{11})$ with cubic term and $(\sigma_2, \sigma_1) = (0.8, 0.8)$

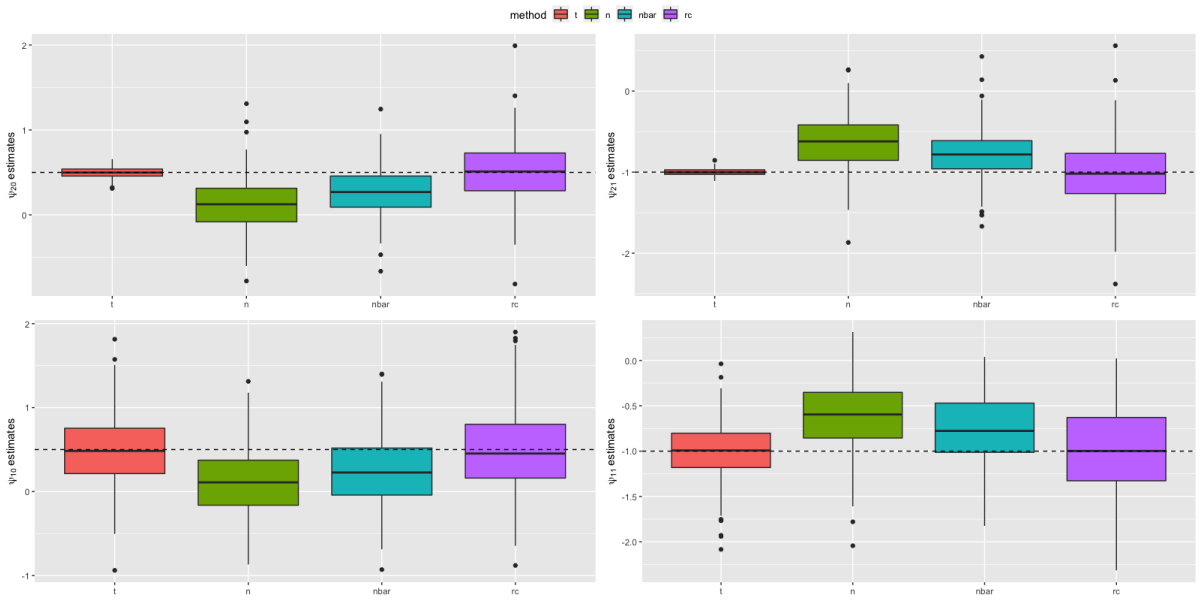

Figure 2.7: Two-stage estimates of (ψ_{20}, ψ_{21}, ψ_{10}, ψ_{11}) with exponential term and (σ_2, σ_1) = (0.8, 0.8)

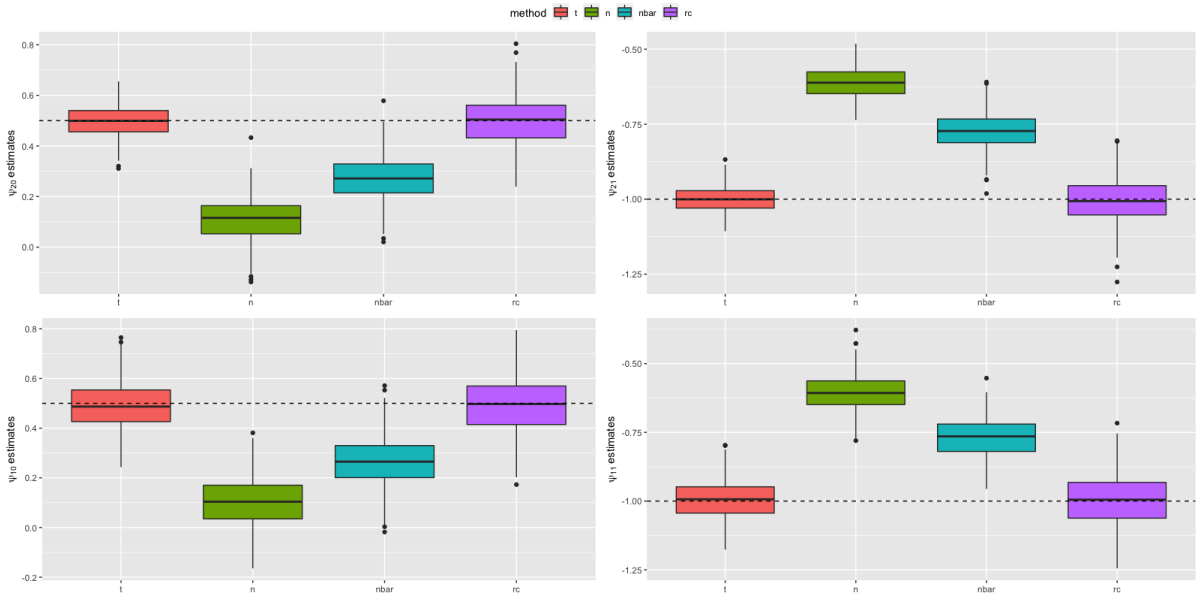

Figure 2.8: Two-stage estimates of (ψ_{20}, ψ_{21}, ψ_{10}, ψ_{11}) with complex term and (σ_2, σ_1) = (0.8, 0.8)

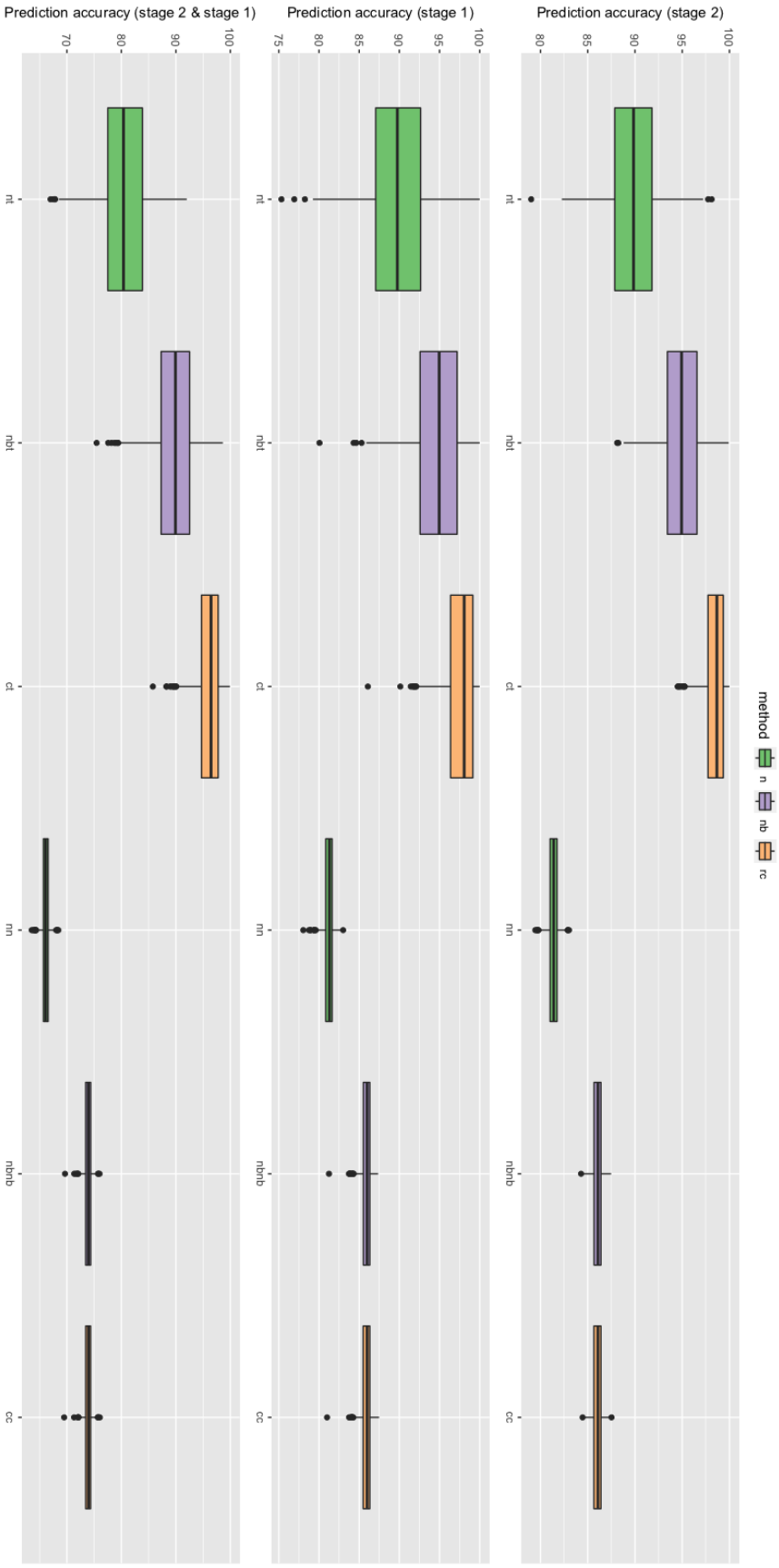

Figure 2.9: Prediction accuracy of optimal DTR with $(\sigma_2, \sigma_1) = (0.8, 0.8)$

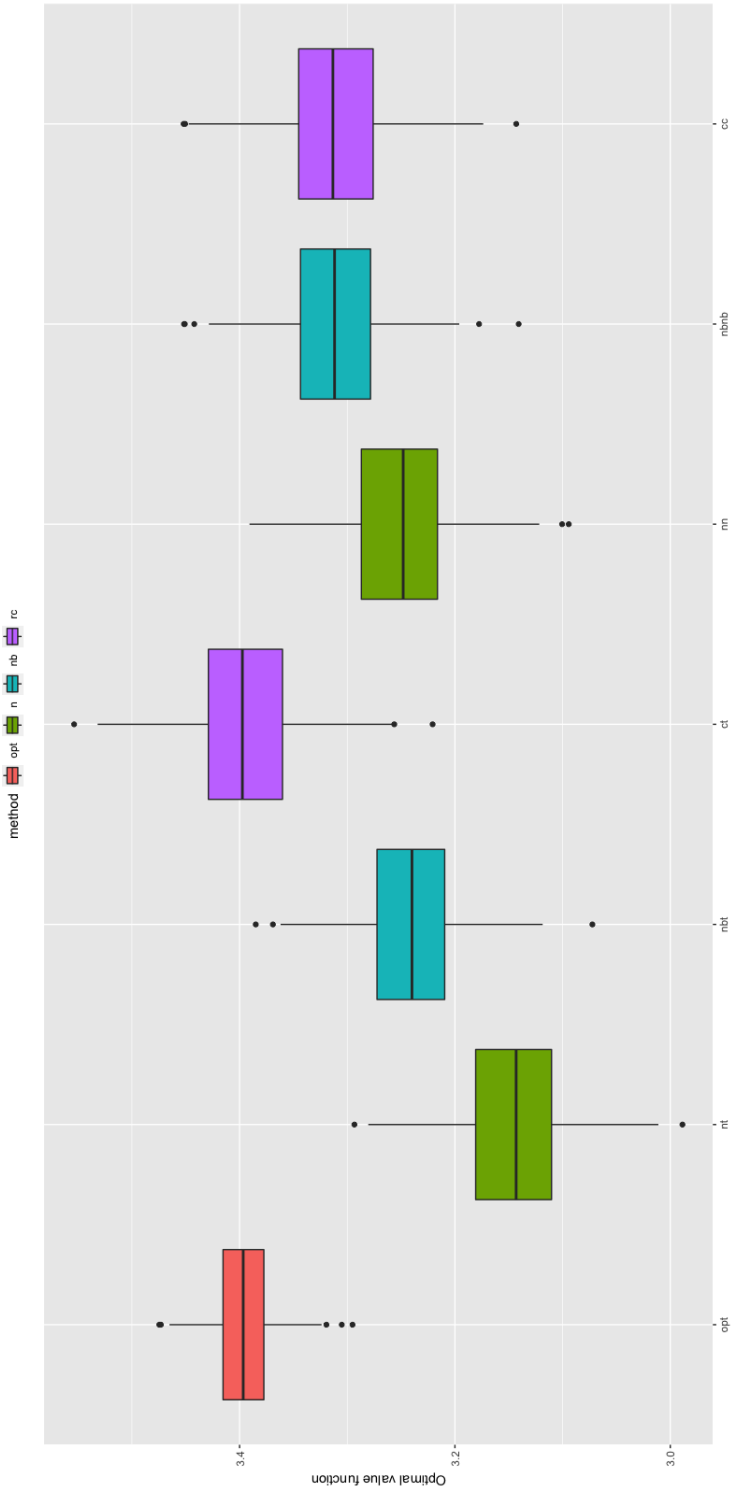

Figure 2.10: Predicted optimal value function with $(\sigma_2, \sigma_1) = (0.8, 0.8)$

Chapter 3

Dynamic Treatment Regimes with Survival Response and Covariate Measurement Error

3.1 Introduction

In the last chapter, we explore the effect of measurement error in covariates on DTR with continuous outcomes. This chapter examines the covariate measurement error effect on DTR with survival outcomes. In particular, the error-prone covariates are incorporated in the dynamic weighted survival modeling proposed by Simoneau et al. (2020b), given that internal validation data are available in each stage. The validation data contain both the true observations and continuous auxiliary variables/surrogates for the covariates. Two correction methods are proposed in DWSurv to handle the mismeasured and incomplete covariates. The first method is the k-nearest neighbors (kNN) method, which directly deals with the continuous surrogates to eliminate the measurement error effect in DWSurv. The second method is an extension of the weighted least squares method developed by Jin et al. (2019). The extended version requires the transformation of a continuous surrogate into a discrete variable for the use of the WLS method in DWSurv.

The remainder of the chapter is organized as follows. Section 3.2 describes the basic no-

tations and the framework of the *k*NN and WLS methods in DWSurv. Simulation studies are conducted in Section 3.3 to assess the performance of the proposed methods in DWSurv. In Section 3.4, the proposed correction methods are applied to the ICU data from the MIMIC-III database. Concluding remarks are summarized in Section 3.5.

3.2 Methodology

3.2.1 Notations and Model Framework

For simplicity, the notations and framework set-up are restricted to two decision points in DTR. Let X_j and $\boldsymbol{Z_j}$ be the error-prone covariate and error-free covariate vector at stage j (j = 1, 2). W_j denotes an auxiliary covariate, a surrogate to X_j with a classical additive relationship.

We consider a situation with the data trajectory (η_{i1}, X_{i1}, W_{i1}, $\boldsymbol{Z_{i1}}$, A_{i1}, Y_{i1}, η_{i2}, X_{i2}, W_{i2}, $\boldsymbol{Z_{i2}}$, A_{i2}, Y_{i2}, Δ_i), where the true covariate X_j is only observed in a subset of the data, but W_j is fully observed. In other words, at any stage j, the data with $\eta_j = 1$ are partitioned into validation data V and main study data $\overline{V}$

$$\begin{aligned} &\{X_{ij}, W_{ij}, \boldsymbol{Z_{ij}}, A_{ij}, Y_{ij}, \Delta_i\} \quad \text{if i} \in V, \\ &\{W_{ij}, \boldsymbol{Z_{ij}}, A_{ij}, Y_{ij}, \Delta_i\} \quad \text{if i} \in \overline{V}. \end{aligned}$$

Given the data structure, let $\bar{X}_{ij} = \mathrm{E}[X_{ij}|W_{ij}, \boldsymbol{Z_{ij}}, A_{ij}]$ in place of the unobserved X_{ij} in the AFT models. Define X^*_{ij} as a variable such that

$$X^*_{ij} = \tau_{ij} X_{ij} + (1 - \tau_{ij})\bar{X}_{ij},$$

where τ_{ij} is an indicator to denote whether the patient i is in the validation at stage j ($\tau_{ij} = 1$) or not ($\tau_{ij} = 0$).

If measurement error is ignored, by replacing X_j with W_j, we obtain naive histories as $\boldsymbol{H^n_1} = (W_1, \boldsymbol{Z_1})$, $\boldsymbol{H^n_2} = (W_1, \boldsymbol{Z_1}, A_1, W_2, \boldsymbol{Z_2})$. Then, the naive AFT models based on the naive histories in DWSurv are given by

$$\begin{aligned} \log T_2 &= \boldsymbol{\beta^{nT}_2 h^n_{20}} + a_2(\boldsymbol{\psi^{nT}_2 h^n_{21}}) + \epsilon_2, \\ \log \widetilde{T} &= \boldsymbol{\beta^{nT}_1 h^n_{10}} + a_1(\boldsymbol{\psi^{nT}_1 h^n_{11}}) + \epsilon_1. \end{aligned} \tag{3.1}$$

He et al. (2007) demonstrated that the naive estimator is biased in the AFT models with covariate measurement error. Thus, it is reasonable to believe that applying the naive AFT models (3.1) into the DWSurv algorithm may yield a biased blip estimator $\hat{\boldsymbol{\psi}}_j^{\boldsymbol{n}}$. Then, the naive blip estimators $\hat{\boldsymbol{\psi}}^{\boldsymbol{n}} = (\hat{\boldsymbol{\psi}}_{\mathbf{2}}^{\boldsymbol{n}}, \hat{\boldsymbol{\psi}}_{\mathbf{1}}^{\boldsymbol{n}})$ may be biased from $\boldsymbol{\psi} = (\boldsymbol{\psi_2}, \boldsymbol{\psi_1})$, which can further affect the estimation of the optimal DTR. Such concerns motivate us to explore effective approaches to find the estimates for the missing X_j in $\overline{V}$ and replace X_j with the new substitute values in the AFT models, and in turn, to adjust for the estimation of the optimal DTR.

3.2.2 *k*-Nearest Neighbors Method

k-nearest neighbors (*k*NN) is a non-parametric statistical learning method. It is known as a lazy learning method because it assumes no distribution of the data specified (Aha, 1997). *k*NN method can be used for both regression problems and classification problems. The key idea of the *k*NN method is to find the nearest *k* neighbors in the training data for the test objects in the testing data. Once the *k* neighbors are identified, the value or label for the object in the testing data can be determined (Biau and Devroye, 2015). In the case of unobservable values of X_{ij} in the main study in DWSurv, we may borrow this idea to develop a method to find the estimates from the validation data in place of X_{ij} in the main study in DWSurv.

Recall that the data structure of DWSurv at stage j follows

$$\{X_{ij}, W_{ij}, \boldsymbol{Z_{ij}}, A_{ij}, Y_{ij}, \Delta_i\} \quad \text{if i} \in V,$$
$$\{W_{ij}, \boldsymbol{Z_{ij}}, A_{ij}, Y_{ij}, \Delta_i\} \quad \text{if i} \in \overline{V}.$$

The true covariate X_{ij} is only observed in the validation data V but is unobservable in the main study data $\overline{V}$. Let $\hat{X}_{ij}^k$ be the estimates of the unobserved X_{ij} in $\overline{V}$. We denote by $\mathrm{W}_{ij}^{(s)}$ the s^{th} nearest neighbor of $\mathrm{W}_{ij} \in \overline{V}$ among $\mathrm{W}_{ij} \in V$, s = 1, ..., k, and $\mathrm{X}_{ij}^{(s)}$ the corresponding data point among $\mathrm{X}_{ij} \in V$. We obtain $\hat{X}_{ij}^k$ by first locating the nearest k data points $\mathrm{W}_{ij}^{(s)} \in V$ around $\mathrm{W}_{ij} \in \overline{V}$ and averaging the associated $\mathrm{X}_{ij}^{(s)} \in V$. Mathematically, $\hat{X}_{ij}^k$ in $\overline{V}$ is given by

$$\hat{X}_{ij}^k = \frac{1}{k} \sum_{s=1}^{k} \mathbb{1}(W_{ij}^{(s)} \text{ is near } W_{ij})\, X_{ij}^{(s)}, \tag{3.2}$$

where k is a given positive integer. Let $\widetilde{X}_{ij}^{k}$ be defined as a random variable at stage j such that

$$\widetilde{X}_{ij}^{k} = \tau_{ij} X_{ij} + (1 - \tau_{ij}) \hat{X}_{ij}^{k}. \tag{3.3}$$

Using $\widetilde{X}_{j}^{k}$ obtained from (3.3) leads to kNN histories $\boldsymbol{H}_{\mathbf{1}}^{\boldsymbol{k}} = (\widetilde{X}_{1}^{k}, \boldsymbol{Z}_1)$ and $\boldsymbol{H}_{\mathbf{2}}^{\boldsymbol{k}} = (\widetilde{X}_{1}^{k}, \boldsymbol{Z}_1, A_1, \widetilde{X}_{2}^{k}, \boldsymbol{Z}_2)$. Thus, the AFT models based on the kNN histories at two stages are given by

$$\log T_2 = \boldsymbol{\beta}_{\mathbf{2}}^{\boldsymbol{kT}} \boldsymbol{h}_{\mathbf{20}}^{\boldsymbol{k}} + a_2\left(\boldsymbol{\psi}_{\mathbf{2}}^{\boldsymbol{kT}} \boldsymbol{h}_{\mathbf{21}}^{\boldsymbol{k}}\right) + \epsilon_2,$$
$$\log \widetilde{T} = \boldsymbol{\beta}_{\mathbf{1}}^{\boldsymbol{kT}} \boldsymbol{h}_{\mathbf{10}}^{\boldsymbol{k}} + a_1\left(\boldsymbol{\psi}_{\mathbf{1}}^{\boldsymbol{kT}} \boldsymbol{h}_{\mathbf{11}}^{\boldsymbol{k}}\right) + \epsilon_1.$$

Distance Measure

The performance of the kNN method is determined by two factors, the distance measure and the choice of k (Biau and Devroye, 2015; Zhang, 2016). Distance measure calculates the relative distance between two data points based on a choice of the distance functions, including Euclidean, Manhattan, Minkowsky, Chebychev, and Chi-square distances, etc. It has been studied in the literature that the distance measure has a significant influence on the performance of the kNN algorithm (Alkasassbeh et al., 2015; Chomboon et al., 2015; Lopes and Ribeiro, 2015; Mulak and Talhar, 2015; Hu et al., 2016). In this study, we use the Euclidean distance to measure the distance. Since W_{ij} is assumed to be a scalar, the resulting distance measure is the absolute value of the difference between two points.

Choice of *k*

Tuning parameter k denotes the number of nearest neighbors to be selected. Similar to the distance measure, the choice of k is empirical. It is often chosen by specific criteria defined in the studies. Researchers have made efforts into developing methods to select the optimal k in various situations (Sun and Huang, 2010; Gou et al., 2011; Cheng et al., 2014; Hassanat et al., 2014; Zhang et al., 2018; Azadkia, 2019). Devroye et al. (1994) showed the consistency of the kNN estimates if k was chosen under the conditions: $\lim_{n_v \to \infty} k = \infty$ and $\lim_{n_v \to \infty} k/n_v = 0$, where n_v is the number of observations in the validation data. Such conditions cover a family

of k that depends on n_v. Practically, the rule of thumb for choosing the k is $\sqrt{n_v}$ (Hassanat et al., 2014; Nadkarni, 2016). In this study, we adopt this idea to define a more general k that satisfies the conditions in Devroye et al. (1994) and use the f-fold cross-validation, a data-driven approach to select the optimal k.

In detail, let k be defined as $k = n_v^{\alpha}$, where $\alpha \in (0, 1)$ is a real number. For $0 < \alpha_{min} < \alpha_{max} < 1$, a sequence of finite values for k ($\lceil n_v^{\alpha_{min}} \rceil = k_{min} < k < k_{max} = \lceil n_v^{\alpha_{max}} \rceil$) is selected, and the validation data are randomly split into a number of f subsamples. For each k, we select (f-1) subsamples and apply the proposed kNN method to obtain the predicted values $\hat{X}_{ij}^{f}(k)$ in the remaining f^{th} subsample denoted as V^f. Then, we calculate the mean squared error (MSE) of $\hat{X}_{ij}^{f}(k)$ by using the X_{ij}^{f} in V^f. This process is repeated f times with each of the subsamples used only once. Then, the mean of fold-based MSE is obtained for that k. When all the values in the sequence are evaluated, the optimal k denoted as k^{opt} can be determined as the one that has the lowest value of MSE

$$k^{opt} = \underset{k_{min}<k<k_{max}}{\arg\min}\, \mathrm{MSE}(k), \tag{3.4}$$

where MSE(k) is the mean of $\frac{1}{n_{vf}} \sum_{i \in V^f} (\hat{X}_{ij}^{f}(k) - X_{ij}^{f})^2$, and n_{vf} is the number of observations in V^f. This adaptive approach is easy to implement and efficient to produce the optimal k.

Based on the distance measure and the choice of optimal k discussed above, we can establish the theoretical property of the proposed kNN method. The detailed conditions and proof are given in the Appendix in Section 3.6.

Theorem 3.2.1 *Under the conditions (C1) - (C3) in the Appendix, Section 3.6, the proposed kNN method yields consistent estimates of $\boldsymbol{\psi}$ in DWSurv.*

Modified Dynamic Weighted Survival Modeling Algorithm I:

Provided the chosen distance measure and empirical choice of k, the modified dynamic weighted survival modeling algorithm I consists of the following steps:

1. Propose parametric models for the probability of treatment $\mathrm{P}(A_2 = 1|\boldsymbol{h}_2^{\boldsymbol{k}}, \eta_2 = 1)$ and the probability of censoring $\mathrm{P}(\Delta = 0|\boldsymbol{h}_2^{\boldsymbol{k}}, a_2, \eta_2 = 1)$ and find the estimated weight $\hat{w}_2$ from (1.10).

2. Assume a linear AFT model for the logarithm of survival time at stage 2 $\log T_2 =$

$\boldsymbol{\beta}_{\mathbf{2}}^{\boldsymbol{kT}}\boldsymbol{h}_{\mathbf{20}}^{\boldsymbol{k}} + a_2\left(\boldsymbol{\psi}_{\mathbf{2}}^{\boldsymbol{kT}}\boldsymbol{h}_{\mathbf{21}}^{\boldsymbol{k}}\right) + \epsilon_2$ and obtain the estimator $(\hat{\boldsymbol{\beta}}_{\mathbf{2}}^{\boldsymbol{k}}, \hat{\boldsymbol{\psi}}_{\mathbf{2}}^{\boldsymbol{k}})$ by solving

$$\mathrm{U}_2(\boldsymbol{\beta}_{\mathbf{2}}^{\boldsymbol{k}}, \boldsymbol{\psi}_{\mathbf{2}}^{\boldsymbol{k}}) = \sum_{i=1}^{n} \delta_i \eta_{i2} \hat{w}_{i2} \begin{pmatrix} \boldsymbol{h}_{\boldsymbol{i20}}^{\boldsymbol{k}} \\ a_{i2}\boldsymbol{h}_{\boldsymbol{i21}}^{\boldsymbol{k}} \end{pmatrix} \left(\log T_{i2} - \boldsymbol{\beta}_{\mathbf{2}}^{\boldsymbol{kT}}\boldsymbol{h}_{\boldsymbol{i20}}^{\boldsymbol{k}} - a_{i2}\boldsymbol{\psi}_{\mathbf{2}}^{\boldsymbol{kT}}\boldsymbol{h}_{\boldsymbol{i21}}^{\boldsymbol{k}}\right) = 0.$$

3. Derive the stage 2 optimal treatment as $\hat{a}_2^{opt} = \mathbb{1}\left(\hat{\boldsymbol{\psi}}_{\mathbf{2}}^{\boldsymbol{kT}}\boldsymbol{h}_{\mathbf{21}}^{\boldsymbol{k}} > 0\right)$.

4. Construct the pseudo-survival time for estimation at stage 1

$$\widetilde{T} = T_1 + \eta_2 T_2 \exp\left\{\hat{\boldsymbol{\psi}}_{\mathbf{2}}^{\boldsymbol{kT}}\boldsymbol{h}_{\mathbf{21}}^{\boldsymbol{k}}[\hat{a}_2^{opt} - a_2]\right\}.$$

5. Propose parametric models for the probability of treatment $\mathrm{P}(A_1 = 1|\boldsymbol{h}_{\mathbf{1}}^{\boldsymbol{k}}, \eta_1 = 1)$ and the probability of censoring $\mathrm{P}(\Delta = 0|\boldsymbol{h}_{\mathbf{1}}^{\boldsymbol{k}}, a_1, \eta_1 = 1)$ and find the estimated weight $\hat{w}_1$ from (1.10).

6. Assume a linear AFT model for the counterfactual logarithm of survival time at stage 1 $\log\widetilde{T} = \boldsymbol{\beta}_{\mathbf{1}}^{\boldsymbol{kT}}\boldsymbol{h}_{\mathbf{10}}^{\boldsymbol{k}} + a_1\left(\boldsymbol{\psi}_{\mathbf{1}}^{\boldsymbol{kT}}\boldsymbol{h}_{\mathbf{11}}^{\boldsymbol{k}}\right) + \epsilon_1$ and obtain the estimator $(\hat{\boldsymbol{\beta}}_{\mathbf{1}}^{\boldsymbol{k}}, \hat{\boldsymbol{\psi}}_{\mathbf{1}}^{\boldsymbol{k}})$ by solving

$$\mathrm{U}_1(\boldsymbol{\beta}_{\mathbf{1}}^{\boldsymbol{k}}, \boldsymbol{\psi}_{\mathbf{1}}^{\boldsymbol{k}}) = \sum_{i=1}^{n} \delta_i \eta_{i1} \hat{w}_{i1} \begin{pmatrix} \boldsymbol{h}_{\boldsymbol{i10}}^{\boldsymbol{k}} \\ a_{i1}\boldsymbol{h}_{\boldsymbol{i11}}^{\boldsymbol{k}} \end{pmatrix} \left(\log\widetilde{T}_i - \boldsymbol{\beta}_{\mathbf{1}}^{\boldsymbol{kT}}\boldsymbol{h}_{\boldsymbol{i10}}^{\boldsymbol{k}} - a_{i1}\boldsymbol{\psi}_{\mathbf{1}}^{\boldsymbol{kT}}\boldsymbol{h}_{\boldsymbol{i11}}^{\boldsymbol{k}}\right) = 0.$$

7. Derive the stage 1 optimal treatment as $\hat{a}_1^{opt} = \mathbb{1}\left(\hat{\boldsymbol{\psi}}_{\mathbf{1}}^{\boldsymbol{kT}}\boldsymbol{h}_{\mathbf{11}}^{\boldsymbol{k}} > 0\right)$.

3.2.3 Weighted Least Squares Method

In this section, we extend the weighted least squares (WLS) method of Jin et al. (2019) to a case with continuous surrogates in a multi-stage setting. Let M_j be an auxiliary variable at stage j, which takes the value of {1, 2, 3, 4}, according to the interval of $(-\infty, \mathrm{Q}_1]$, $(\mathrm{Q}_1, \mathrm{Q}_2]$, $(\mathrm{Q}_2, \mathrm{Q}_3]$, $(\mathrm{Q}_3, +\infty)$ that W_j lies in, where Q_1, Q_2 and Q_3 denote the quartiles of W_j. Then, we have a new data structure

$$\{X_{ij}, W_{ij}, M_{ij}, \boldsymbol{Z_{ij}}, A_{ij}, Y_{ij}, \Delta_i\} \quad \text{if i} \in V,$$
$$\{W_{ij}, M_{ij}, \boldsymbol{Z_{ij}}, A_{ij}, Y_{ij}, \Delta_i\} \quad \text{if i} \in \overline{V}.$$

As M_j is a discretized variable obtained from W_j, we can use M_j to assist us in finding the estimates $\hat{X}_{ij}^{w}$ of the unobserved X_{ij} in $\overline{V}$. For any $\mathrm{i} \in \overline{V}$, we modify the equation (1.14) to be applicable for estimating the X_{ij} at stage j, which is given by

$$\hat{X}_{ij}^{w} = \frac{\sum_{s\in V} \mathbb{1}(M_{sj} = M_{ij}) X_{sj}}{\sum_{s\in V} \mathbb{1}(M_{sj} = M_{ij})}. \tag{3.5}$$

With (3.5), we can obtain a substitute $\widetilde{X}_{ij}^{w}$ for X_{ij} at stage j such that

$$\widetilde{X}_{ij}^{w} = \tau_{ij} X_{ij} + (1 - \tau_{ij})\hat{X}_{ij}^{w}. \tag{3.6}$$

Then, using $\widetilde{X}_{j}^{w}$ obtained from (3.6) yields WLS histories $\boldsymbol{H}_{\mathbf{1}}^{\boldsymbol{w}} = (\widetilde{X}_{1}^{w}, \boldsymbol{Z}_1)$ and $\boldsymbol{H}_{\mathbf{2}}^{\boldsymbol{w}} = (\widetilde{X}_{1}^{w}, \boldsymbol{Z}_1, A_1, \widetilde{X}_{2}^{w}, \boldsymbol{Z}_2)$. Hence, the AFT models based on the WLS histories are given by

$$\log T_2 = \boldsymbol{\beta}_{\mathbf{2}}^{\boldsymbol{wT}} \boldsymbol{h}_{\mathbf{20}}^{\boldsymbol{w}} + a_2\left(\boldsymbol{\psi}_{\mathbf{2}}^{\boldsymbol{wT}} \boldsymbol{h}_{\mathbf{21}}^{\boldsymbol{w}}\right) + \epsilon_2,$$
$$\log \widetilde{T} = \boldsymbol{\beta}_{\mathbf{1}}^{\boldsymbol{wT}} \boldsymbol{h}_{\mathbf{10}}^{\boldsymbol{w}} + a_1\left(\boldsymbol{\psi}_{\mathbf{1}}^{\boldsymbol{wT}} \boldsymbol{h}_{\mathbf{11}}^{\boldsymbol{w}}\right) + \epsilon_1.$$

The intuition behind the equation (3.5) is that we can find the estimates for the unobserved X_{ij} in $\overline{V}$ by directly searching for the matched $M_{ij} \in V$ and $M_{ij} \in \overline{V}$ at stage j and averaging the corresponding true covariate X_{ij} in V. This method is similar to the kNN method, except that a categorical variable M_j is used for calculation, which is obtained from the surrogate W_j. As a result, we can use this method as an alternative to the kNN method.

Modified Dynamic Weighted Survival Modeling Algorithm II:

Provided the WLS equations in (3.5) and (3.6), the modified dynamic weighted survival modeling algorithm II consists of the following steps:

1. Propose parametric models for the probability of treatment $P(A_2 = 1|\boldsymbol{h}_{\mathbf{2}}^{\boldsymbol{w}}, \eta_2 = 1)$ and the probability of censoring $P(\Delta = 0|\boldsymbol{h}_{\mathbf{2}}^{\boldsymbol{w}}, a_2, \eta_2 = 1)$ and find the estimated weight $\hat{w}_2$ from (1.10).

2. Assume a linear AFT model for the logarithm of survival time at stage 2 $\log T_2 = \boldsymbol{\beta}_{\mathbf{2}}^{\boldsymbol{wT}} \boldsymbol{h}_{\mathbf{20}}^{\boldsymbol{w}} + a_2\left(\boldsymbol{\psi}_{\mathbf{2}}^{\boldsymbol{wT}} \boldsymbol{h}_{\mathbf{21}}^{\boldsymbol{w}}\right) + \epsilon_2$ and obtain the estimator $(\hat{\boldsymbol{\beta}}_{\mathbf{2}}^{\boldsymbol{w}}, \hat{\boldsymbol{\psi}}_{\mathbf{2}}^{\boldsymbol{w}})$ by solving

$$U_2(\boldsymbol{\beta}_{\mathbf{2}}^{\boldsymbol{w}}, \boldsymbol{\psi}_{\mathbf{2}}^{\boldsymbol{w}}) = \sum_{i=1}^{n} \delta_i \eta_{i2} \hat{w}_{i2} \begin{pmatrix} \boldsymbol{h}_{\boldsymbol{i20}}^{\boldsymbol{w}} \\ a_{i2} \boldsymbol{h}_{\boldsymbol{i21}}^{\boldsymbol{w}} \end{pmatrix} \left(\log T_{i2} - \boldsymbol{\beta}_{\mathbf{2}}^{\boldsymbol{wT}} \boldsymbol{h}_{\boldsymbol{i20}}^{\boldsymbol{w}} - a_{i2} \boldsymbol{\psi}_{\mathbf{2}}^{\boldsymbol{wT}} \boldsymbol{h}_{\boldsymbol{i21}}^{\boldsymbol{w}}\right) = 0.$$

3. Derive the stage 2 optimal treatment as $\hat{a}_2^{opt} = \mathbb{1}\left(\hat{\boldsymbol{\psi}}_{\mathbf{2}}^{\boldsymbol{wT}} \boldsymbol{h}_{\mathbf{21}}^{\boldsymbol{w}} > 0\right)$.

4. Construct the pseudo-survival time for estimation at stage 1

$$\widetilde{T} = T_1 + \eta_2 T_2 \exp\left\{\hat{\boldsymbol{\psi}}_{\mathbf{2}}^{\boldsymbol{wT}} \boldsymbol{h}_{\mathbf{21}}^{\boldsymbol{w}} [\hat{a}_2^{opt} - a_2]\right\}.$$

5. Propose parametric models for the probability of treatment $\mathrm{P}(A_1 = 1|\boldsymbol{h}_{\mathbf{1}}^{\boldsymbol{w}}, \eta_1 = 1)$ and the probability of censoring $\mathrm{P}(\Delta = 0|\boldsymbol{h}_{\mathbf{1}}^{\boldsymbol{w}}, a_1, \eta_1 = 1)$ and find the estimated weight $\hat{w}_1$ from (1.10).

6. Assume a linear AFT model for the counterfactual logarithm of survival time at stage 1 $\log\widetilde{T} = \boldsymbol{\beta}_{\mathbf{1}}^{\boldsymbol{wT}}\boldsymbol{h}_{\mathbf{10}}^{\boldsymbol{w}} + a_1\left(\boldsymbol{\psi}_{\mathbf{1}}^{\boldsymbol{wT}}\boldsymbol{h}_{\mathbf{11}}^{\boldsymbol{w}}\right) + \epsilon_1$ and obtain the estimator $(\hat{\boldsymbol{\beta}}_{\mathbf{1}}^{\boldsymbol{w}}, \hat{\boldsymbol{\psi}}_{\mathbf{1}}^{\boldsymbol{w}})$ by solving

$$\mathrm{U}_1(\boldsymbol{\beta}_{\mathbf{1}}^{\boldsymbol{w}}, \boldsymbol{\psi}_{\mathbf{1}}^{\boldsymbol{w}}) = \sum_{i=1}^{n} \delta_i\eta_{i1}\hat{w}_{i1}\begin{pmatrix} \boldsymbol{h}_{\boldsymbol{i10}}^{\boldsymbol{w}} \\ a_{i1}\boldsymbol{h}_{\boldsymbol{i11}}^{\boldsymbol{w}} \end{pmatrix}\left(\log\widetilde{T}_i - \boldsymbol{\beta}_{\mathbf{1}}^{\boldsymbol{wT}}\boldsymbol{h}_{\boldsymbol{i10}}^{\boldsymbol{w}} - a_{i1}\boldsymbol{\psi}_{\mathbf{1}}^{\boldsymbol{wT}}\boldsymbol{h}_{\boldsymbol{i11}}^{\boldsymbol{w}}\right) = 0.$$

7. Derive the stage 1 optimal treatment as $\hat{a}_1^{opt} = \mathbb{1}\left(\hat{\boldsymbol{\psi}}_{\mathbf{1}}^{\boldsymbol{wT}}\boldsymbol{h}_{\mathbf{11}}^{\boldsymbol{w}} > 0\right)$.

3.3 Simulation Studies

In this section, extensive simulation studies are conducted to assess the impact of measurement error in covariates on the estimation of parameters, prediction of optimal DTR and optimal overall survival time. The performance of the proposed *k*NN method and WLS method is evaluated in one-stage and multi-stage settings. The validity of the double robustness property in DWSurv is also examined in the presence of measurement error.

3.3.1 One-Stage Estimation

Let X be an error-prone covariate and Z be an error-free covariate, generated from a uniform distribution U(0.1, 2). W is a mismeasured version for X, with the relationship W = X + e, where e ~ N(0, σ^2) and σ reflects the degree of measurement error. The treatment A $\in \{1, 0\}$ is generated from a Bernoulli distribution with probability P(A = 1) = expit(0.5X + 0.5Z - 1), where expit(x) = 1/(1 + exp(-x)) is the inverse-logit function. Censoring time C is generated from an exponential distribution with a rate of 1/300. The log-survival time is generated as

$$\log T = 2 + \beta_z Z + \beta_x X + A(\psi_{10} + \psi_{11}X) + \epsilon,$$

where $(\boldsymbol{\beta}, \boldsymbol{\psi}) = (\beta_z, \beta_x, \psi_{10}, \psi_{11}) = (0.5, 0.5, 1, -1)$ and $\epsilon \sim \mathrm{N}(0, 0.5^2)$ is generated independent of all other variables. A dataset of 2000 patients is simulated 500 times. In each simulation, the dataset is randomly split into validation data and main study data with a validation ratio ρ, where the validation data contain $100\times\rho\%$ of the observations. A sequence of k is generated

with $\alpha \in \{0.1, ..., 0.9\}$ with a step size of 0.1. The primary interest is to evaluate the measurement error effect on the estimation of blip parameters, but its impact on the estimation of parameters for the main terms is also examined in the study.

Five estimators are considered to evaluate the performance of the proposed methods: (1) validation estimator ($\hat{\boldsymbol{\beta}}^{\boldsymbol{v}}$, $\hat{\boldsymbol{\psi}}^{\boldsymbol{v}}$) obtained using the validation data, (2) complete estimator ($\hat{\boldsymbol{\beta}}^{\boldsymbol{c}}$, $\hat{\boldsymbol{\psi}}^{\boldsymbol{c}}$) obtained using the complete data, (3) naive estimator ($\hat{\boldsymbol{\beta}}^{\boldsymbol{n}}$, $\hat{\boldsymbol{\psi}}^{\boldsymbol{n}}$) obtained using W, (4) kNN estimator ($\hat{\boldsymbol{\beta}}^{\boldsymbol{k}}$, $\hat{\boldsymbol{\psi}}^{\boldsymbol{k}}$) obtained using $\widetilde{X}^k$ according to the formula (3.3) with k^{opt} chosen from the 10-fold cross-validation and the selected sequence of k, (5) WLS estimator ($\hat{\boldsymbol{\beta}}^{\boldsymbol{w}}$, $\hat{\boldsymbol{\psi}}^{\boldsymbol{w}}$) obtained using $\widetilde{X}^w$ according to the formula (3.6). In simulations, the validation ratio ρ is set to be 0.5 and 0.7. The degree of measurement error $\sigma \in \{0.2, 0.5, 0.8\}$ is specified, and the degree of independent censoring is set as 30% and 70%. A summary of k^{opt} chosen in the simulations is provided in Table 3.1. Table 3.2 and Table 3.3 report the bias, asymptotic standard error (SE), root mean square error (RMSE) and 95% coverage probability (CP%) of $\hat{\boldsymbol{\psi}}$ under 30% and 70% independent censoring, respectively. Figure 3.1 and Figure 3.6 provide the visualized parameter estimates under $\rho = 0.7$.

In Table 3.1, we see that as ρ increases, the optimal k gets larger, indicating the relationship of k with the size of the validation data. Moreover, the optimal k is positively related to σ, reflecting that more data points are needed as the degree of measurement error increases. Tables 3.2 and 3.3 show that using the naive method causes considerable biases in the estimation of all the parameters except β_z due to the measurement error. The biases become larger as σ increases. There is little impact on the estimation of β_z, which corresponds to the error-free covariate Z. In comparison, the proposed kNN and WLS methods significantly reduce biases of parameter estimation and improve the coverage probabilities in all the scenarios. The size of the validation data and the censoring rate also affect the performance of the proposed methods. The empirical biases and standard errors of the proposed methods are reduced as ρ increases. But with more censored patients, all the methods experience larger variation as a result of the loss of information about the survival time.

3.3.2 Two-Stage Estimation

In this section, we explore the effect of covariate measurement error on the parameter estimation in a two-stage setting. The design of data generation mechanism in the multi-stage DTR with survival outcomes is difficult and lacks of realism (Simoneau et al., 2020a). Therefore, our two-stage simulation setting follows the idea of the data generation mechanism from Simoneau et al. (2020b). Let X_j and Z_j be the error-prone covariate and error-free covariate at stage j (j = 1, 2), which are generated from uniform distributions U(0.1, 1.29) at stage 1 and U(0.9, 2) at stage 2, respectively. The surrogate W_j is generated by the classical additive model $W_j = X_j + e_j$, where $e_j \sim N(0, \sigma_j^2)$. The treatment $A_j \in \{1, 0\}$ is assigned with $P(A_1 = 1) = \text{expit}(X_1 + Z_1 - 1)$ and $P(A_2 = 1) = \text{expit}(-X_2 - Z_2 + 2.8)$, respectively. The censoring time is generated from an exponential distribution with a rate of 1/300. Based on the AFT model, the observed survival time at stage 2 is given by

$$T_2 = \exp(4 + 1.1X_2 - 0.3X_2^2 - 0.1Z_2 - 0.1X_1 + A_2(\psi_{20} + \psi_{21}X_2) + \epsilon_2),$$

where $(\psi_{20}, \psi_{21}) = (-0.9, 0.6)$ and $\epsilon_2 \sim N(0, 0.3^2)$, independent of all other variables. The optimal survival time had all the patients received the second stage optimal treatment is

$$T_2^{opt} = \exp(\log T_2 + (A_2^{opt} - A_2)(\psi_{20} + \psi_{21}X_2)).$$

The overall survival time under the optimal treatment at stage 2 is generated from the AFT model as

$$\widetilde{T} = \exp(6.3 + 1.5X_1 - 0.8X_1^2 + 0.1Z_1 + A_1(\psi_{10} + \psi_{11}X_1) + \epsilon_1),$$

where $(\psi_{10}, \psi_{11}) = (0.8, -0.9)$ and $\epsilon_1 \sim N(0, 0.3^2)$, independent of all other variables. We assume that the total size of the simulated dataset is n = 2000, and 60% of the patients have entered stage 2. In each stage, 70% of the data are randomly selected as the validation data. The measurement error degree (σ_2, σ_1) is considered with $\sigma_j \in \{0.2, 0.5, 0.8\}$. The independent censoring rate is set to be 30% and 70%. A sequence of k is generated with $\alpha \in \{0.1, ..., 0.9\}$ with a step size of 0.1. Simulations are run 500 times.

In the two-stage DWSurv, we are interested in estimating the blip parameters $\boldsymbol{\psi} = (\psi_{20}, \psi_{21}, \psi_{10}, \psi_{11})$. Five blip estimators are compared to evaluate the performance of the proposed methods: (1) validation estimator $\hat{\boldsymbol{\psi}}^v$ obtained using X_j based on the validation data, (2) complete

estimator $\hat{\boldsymbol{\psi}}^{\boldsymbol{c}}$ obtained using X_j based on the complete data, (3) naive estimator $\hat{\boldsymbol{\psi}}^{\boldsymbol{n}}$ obtained by using W_j, (4) kNN estimator $\hat{\boldsymbol{\psi}}^{\boldsymbol{k}}$ obtained using $\widetilde{X}_j^k$ according to the formula (3.3) with k^{opt} chosen from the 10-fold cross-validation and the generated sequence of k, (5) WLS estimator $\hat{\boldsymbol{\psi}}^{\boldsymbol{w}}$ obtained using $\widetilde{X}_j^w$ according to the formula (3.6).

We would like to examine the validity of the double robustness property of the DWSurv method, whether it still holds in the presence of measurement error. Four scenarios are included in each stage: (1) both the weight (treatment and censoring models) and the treatment-free models are correctly specified, (2) the weight model is correctly specified, but the treatment-free model is misspecified, (3) the weight model is misspecified, but the treatment-free model is correctly specified, (4) both the weight and the treatment-free models are misspecified. The bias, SE, RMSE and CP% of each blip estimator are summarized in Table 3.4 to Table 3.9, accompanied by the parameter estimates of the four scenarios visualized in Figure 3.7 to Figure 3.12.

From the results, we see that the double robustness property fails with the naive estimator $\hat{\boldsymbol{\psi}}^{\boldsymbol{n}}$, even in the case of both the weight and the treatment-free models correctly specified (scenario 1). Moreover, the increase in the degree of measurement error exacerbates the biases of $\hat{\boldsymbol{\psi}}^{\boldsymbol{n}}$. In contrast, the proposed estimators $\hat{\boldsymbol{\psi}}^{\boldsymbol{k}}$ and $\hat{\boldsymbol{\psi}}^{\boldsymbol{w}}$ perform satisfactorily with fairly small biases and coverage probabilities are close to the nominal level under various combinations of censoring rate and (σ_2, σ_1). The double robustness property is substantially restored using the proposed estimators, when at least one of the weight model and treatment-free model is correctly specified (scenarios 1, 2 and 3). Moreover, a similar pattern is observed concerning the censoring rate on the proposed estimators. As the censoring rate is higher, the variability of the estimators increase due to more information lost about the survival time. In the last scenario (scenario 4) where both models are misspecified, all the methods yield severely biased results.

3.3.3 Prediction of Optimal DTR

The previous simulation studies examine the measurement error effect on parameter estimation. In this section, we investigate the measurement error effect on the accuracy of the predicted optimal treatment rules. The simulation design follows (3.3.2) but with training and test data

included, where the training data are used to fit models for parameter estimation, and the test data are used to predict the optimal DTR across two stages.

The training data and test data are generated with sizes of 2000 and 5000 patients, respectively. The training data of each stage are randomly split into validation data and main study data with a validation ratio $\rho \in \{0.5, 0.7\}$. Five estimators $\hat{\boldsymbol{\psi}}^{v}$, $\hat{\boldsymbol{\psi}}^{c}$, $\hat{\boldsymbol{\psi}}^{n}$, $\hat{\boldsymbol{\psi}}^{k}$ and $\hat{\boldsymbol{\psi}}^{w}$ are considered and estimated from the training data. We use the test data to find the prediction accuracy of the optimal DTR, which is measured by the proportion of patients whose optimal treatments are correctly predicted at stage 2 and/or stage 1. For j = 1, 2, the degree of measurement error σ_j is specified as 0.2, 0.5 and 0.8, and the rate of independent censoring is considered to be 30% and 70%. Simulations are repeated 500 times for each pair of ρ and (σ_2, σ_1), and the results are displayed in Table 3.10. The prediction accuracy of optimal DTR under $\rho = 0.7$ are visualized in Figure 3.13 and Figure 3.14.

The numerical results show that, in general, the prediction accuracy of optimal DTR is adversely affected by the covariate measurement error. The naive estimator $\hat{\boldsymbol{\psi}}^{n}$ leads to a remarkable degeneration in the prediction results in all the scenarios. The performance of $\hat{\boldsymbol{\psi}}^{n}$ becomes worse as the measurement error gets larger. In contrast, the proposed estimators $\hat{\boldsymbol{\psi}}^{k}$ and $\hat{\boldsymbol{\psi}}^{w}$ perform similarly in terms of the prediction accuracy of optimal DTR. Both proposed estimators significantly improved the prediction accuracy, which is even higher than the prediction accuracy obtained using the validation estimator $\hat{\boldsymbol{\psi}}^{v}$, suggesting a favorable choice of using the proposed methods to derive the sequential optimal treatment rules.

3.3.4 Prediction of the Expected Survival Time

In this section, we assess the prediction of the expected overall survival time under the optimal DTR in contaminated data with covariates being subject to measurement error. The data generation mechanism follows the setting in (3.3.3) with one training data to estimate the parameters and one test data to predict the expected overall survival time under the optimal DTR.

In the first step, five estimators $(\hat{\boldsymbol{\beta}}^{v}, \hat{\boldsymbol{\psi}}^{v})$, $(\hat{\boldsymbol{\beta}}^{c}, \hat{\boldsymbol{\psi}}^{c})$, $(\hat{\boldsymbol{\beta}}^{n}, \hat{\boldsymbol{\psi}}^{n})$, $(\hat{\boldsymbol{\beta}}^{k}, \hat{\boldsymbol{\psi}}^{k})$ and $(\hat{\boldsymbol{\beta}}^{w}, \hat{\boldsymbol{\psi}}^{w})$ are estimated from the training data. We use the test data to obtain five predicted optimal DTR based on the estimators and covariates. Then, for each scenario, the average value of the

overall log-survival time is computed under the estimated optimal DTR. The validation ratio ρ is set to be 0.5 and 0.7. The degree of measurement error $\sigma_j \in \{0.2, 0.5, 0.8\}$ is considered, and the censoring rate is set as 30% and 70%. Simulation is run 500 times. Numerical results for the mean optimal log-survival times, along with standard deviations, are summarized in Table 3.11.

The results in Table 3.11 show that the naive method tends to yield shorter optimal overall log-survival times. Moreover, the mean of the predicted optimal log-survival times using the naive method decreases as the measurement error increases. By comparison, the proposed kNN and WLS methods perform similarly and enhance the predicted optimal log-survival times in all the scenarios.

3.4 Application to MIMIC-III Data

We apply the proposed correction methods to a cohort of ICU patients with sepsis. The study data are taken from the MIMIC-III database, which contains the observational admission data collected at Beth Israel Deaconess Medical Center from 2001 to 2012 (Johnson et al., 2016,0). Feng et al. (2018) showed the significant association between the use of TTEC and improvement in 28-day mortality. However, Cook et al. (2002) suggested that the use of TTEC in all critically ill surgical patients was not cost-effective. Chen et al. (2021) revealed the heterogeneity in the treatment effects of TTEC and demonstrated the improvement in the 28-day survival rate by customizing the use of TTEC for ICU patients. In this work, we are interested in deriving the optimal treatment decision rules for the use of transthoracic echocardiography that maximize the overall survival time of ICU patients with sepsis.

We follow the same criteria in Feng et al. (2018) to select the cohort data, except that the information of patients whose second admission to ICU is also included. The outcome of interest is the survival time of the patients with sepsis, which is calculated as the difference between the death time and the first ICU admission time. The patients' characteristics at first and second admissions include age, gender, body mass index, simplified acute physiology score (SAPS), sequential organ failure assessment score (SOFA), Elixhauser comorbidity score, heart rate, lab test for cholesterol, positive end-expiratory pressure (PEEP) and the use of mechanical

ventilation (VENT) during the first 24 hours of ICU admission. We select VENT and PEEP as the tailoring variables as they are significantly associated with TTEC (Cook et al., 2002).

At the first ICU admission (stage 1), a total number of 6294 patients were admitted to ICU, and the treatment was initiated as a recommendation of TTEC ($A_1 = 1$) or not ($A_1 = 0$). About 10% of the patients experienced re-admission (stage 2) into ICU with continuing the use of TTEC ($A_2 = 1$) or dropping the use of TTEC ($A_2 = 0$). In the final cohort data, the variable PEEP was found largely underreported with about 35.7% and 37.2% missingness at stage 1 and stage 2, respectively. Removing such an amount of missing values in the analysis may lead to biased estimation. However, the heart rate, which is known to be positively associated with PEEP (Zhou et al., 2019), is completely observed. For j = 1, 2, we treat the cohort data with the observed PEEP_j as the validation data and estimate the unobserved PEEP_j, which forms the main study data, by using the heart rate as the auxiliary covariate.

In the analysis, we consider three blip estimators to construct the optimal treatment decision rules: (1) validation estimator obtained using the patients with observed PEEP_j only, (2) kNN estimator obtained using the imputed PEEP_j according to the formula (3.3), (3) WLS estimator obtained using the imputed PEEP_j according to the formula (3.6). The optimal k is selected by training the validation data using 10-fold cross-validation with $\alpha \in \{0.1, ..., 0.9\}$ with a step size of 0.5. Table 3.12 summarizes the estimation and associated inference results based on the validation method and proposed kNN and WLS methods.

The results in Table 3.12 show that the proposed kNN and WLS methods perform similarly, but the blip parameter estimates and the standard errors vary notably between the validation method and the proposed methods, leading to different optimal treatment decision rules. The variable VENT at stage 1 is statistically significant with respect to the treatment effect in all the methods, reflecting its significant association with the use of TTEC and the overall survival time of the patients. However, the significance of the coefficients for the treatments in two stages is shown to differ between methods. These results emphasize the impact of omitting the data with missing covariates is pronounced, which can result in different optimal treatment decision rules.

3.5 Conclusion

This chapter studies the effect of covariate measurement error on DWSurv with internal validation data provided. Two correction methods, the kNN method and the WLS method, are developed to adjust for the measurement error effect in DWSurv. The first kNN method directly uses the available surrogates to find the estimates for the unobserved true covariates in each stage of DWSurv. This method adopts the cross-validation method, a data-driven approach, with which the optimal number of nearest neighbors is identified. The second WLS method extends the use of the original WLS method to a multi-stage setting with continuous surrogates provided. In each stage, by transforming the continuous surrogate, the estimates of the unobserved covariates can be estimated using the WLS method. One advantage both methods share is that they do not require the specification of the relationship between true covariate and surrogate, making the modeling more robust.

Simulation studies demonstrate the satisfactory performance of the proposed methods in one-stage and multi-stage settings. On average, the kNN and WLS methods provide significant bias reduction in parameter estimation and substantial restoration of the double robustness property in the original DWSurv, even when the magnitude of measurement error is large. Moreover, the proposed methods show their superior performance in a predictive setting with higher prediction accuracy of optimal DTR and longer optimal survival times. However, the proposed methods experience larger standard errors with a higher censoring rate, resulting from losing information about the survival time. The proposed methods are applied to the MIMIC-III data as an illustration to estimate optimal treatment decision rules. The data analysis shows that the estimated optimal treatment decision rules can be altered if the unavailable covariate is not addressed.

Table 3.1: Summary of the optimal k from 10-fold cross-validation

% Censoring	ρ	σ	Min	Mean	Median	Max
30%	0.5	0.2	16	67	64	126
		0.5	32	101	126	126
		0.8	16	105	126	252
	0.7	0.2	19	88	78	160
		0.5	38	133	160	160
		0.8	38	135	160	160
70%	0.5	0.2	16	67	64	126
		0.5	32	101	126	126
		0.8	32	105	126	252
	0.7	0.2	38	92	78	160
		0.5	38	133	160	160
		0.8	38	139	160	329

Table 3.2: One-stage estimates of blip parameters (ψ_{10}, ψ_{11}) with 30% independent censoring

			ψ_{10}				ψ_{11}			
ρ	σ	$\hat{\psi}$	Bias	SE	RMSE	CP%	Bias	SE	RMSE	CP%
		$\hat{\psi}^c$	-0.002	0.059	0.059	93.6	0.001	0.050	0.050	93.4
0.5		$\hat{\psi}^v$	-0.002	0.084	0.084	96.0	0.002	0.071	0.071	94.6
	0.2	$\hat{\psi}^n$	-0.125	0.057	0.137	42.8	0.119	0.048	0.128	31.8
		$\hat{\psi}^k$	0.002	0.061	0.061	94.4	-0.002	0.052	0.052	94.8
		$\hat{\psi}^w$	-0.002	0.062	0.062	95.0	0.002	0.053	0.053	95.0
	0.5	$\hat{\psi}^n$	-0.479	0.050	0.482	0.0	0.457	0.039	0.458	0.0
		$\hat{\psi}^k$	0.007	0.067	0.068	94.2	-0.006	0.058	0.058	95.8
		$\hat{\psi}^w$	-0.004	0.067	0.068	95.2	0.005	0.058	0.058	96.2
	0.8	$\hat{\psi}^n$	-0.717	0.043	0.719	0.0	0.684	0.030	0.684	0.0
		$\hat{\psi}^k$	0.005	0.072	0.072	94.8	-0.004	0.063	0.063	95.6
		$\hat{\psi}^w$	-0.004	0.072	0.073	95.0	0.005	0.063	0.063	96.0
0.7		$\hat{\psi}^v$	-0.003	0.071	0.071	96.0	0.001	0.060	0.060	95.4
	0.2	$\hat{\psi}^n$	-0.125	0.057	0.137	43.4	0.117	0.048	0.126	32.8
		$\hat{\psi}^k$	0.000	0.060	0.060	95.8	-0.002	0.051	0.051	95.4
		$\hat{\psi}^w$	-0.003	0.061	0.061	95.8	0.001	0.052	0.052	95.4
	0.5	$\hat{\psi}^n$	-0.479	0.050	0.482	0.0	0.455	0.039	0.457	0.0
		$\hat{\psi}^k$	0.003	0.063	0.064	95.0	-0.004	0.054	0.055	94.8
		$\hat{\psi}^w$	-0.003	0.064	0.064	95.2	0.001	0.055	0.055	95.0
	0.8	$\hat{\psi}^n$	-0.718	0.043	0.719	0.0	0.683	0.030	0.683	0.0
		$\hat{\psi}^k$	0.001	0.066	0.066	95.2	-0.002	0.057	0.057	94.8
		$\hat{\psi}^w$	-0.004	0.066	0.066	94.8	0.002	0.057	0.057	95.8

Table 3.3: One-stage estimates of blip parameters (ψ_{10}, ψ_{11}) with 70% independent censoring

			ψ_{10}				ψ_{11}			
ρ	σ	$\hat{\psi}$	Bias	SE	RMSE	CP%	Bias	SE	RMSE	CP%
		$\hat{\psi}^c$	-0.002	0.091	0.091	96.2	0.003	0.077	0.077	95.8
0.5		$\hat{\psi}^v$	0.006	0.128	0.128	96.0	-0.004	0.108	0.108	97.0
	0.2	$\hat{\psi}^n$	-0.125	0.087	0.153	70.0	0.120	0.073	0.140	62.0
		$\hat{\psi}^k$	0.002	0.093	0.093	95.8	-0.001	0.079	0.079	95.2
		$\hat{\psi}^w$	-0.004	0.094	0.094	95.6	0.005	0.080	0.081	95.2
	0.5	$\hat{\psi}^n$	-0.479	0.076	0.485	0.0	0.457	0.059	0.461	0.0
		$\hat{\psi}^k$	0.006	0.103	0.103	95.6	-0.004	0.089	0.089	95.8
		$\hat{\psi}^w$	-0.004	0.103	0.103	95.4	0.005	0.089	0.089	94.6
	0.8	$\hat{\psi}^n$	-0.716	0.066	0.719	0.0	0.684	0.046	0.685	0.0
		$\hat{\psi}^k$	0.006	0.111	0.111	95.6	-0.003	0.097	0.097	96.0
		$\hat{\psi}^w$	-0.004	0.111	0.111	95.2	0.006	0.097	0.097	95.6
0.7		$\hat{\psi}^v$	0.004	0.108	0.108	94.2	0.000	0.091	0.091	92.2
	0.2	$\hat{\psi}^n$	-0.121	0.087	0.149	70.6	0.116	0.073	0.137	63.8
		$\hat{\psi}^k$	0.003	0.092	0.092	95.6	-0.003	0.078	0.078	94.8
		$\hat{\psi}^w$	0.001	0.093	0.093	95.8	0.000	0.079	0.079	95.4
	0.5	$\hat{\psi}^n$	-0.477	0.076	0.483	0.0	0.456	0.059	0.460	0.0
		$\hat{\psi}^k$	0.004	0.097	0.097	95.2	-0.004	0.083	0.083	93.8
		$\hat{\psi}^w$	-0.002	0.097	0.097	95.0	0.002	0.083	0.083	93.8
	0.8	$\hat{\psi}^n$	-0.716	0.066	0.719	0.0	0.685	0.046	0.686	0.0
		$\hat{\psi}^k$	0.002	0.101	0.101	95.2	-0.002	0.087	0.087	92.2
		$\hat{\psi}^w$	-0.002	0.101	0.101	96.0	0.002	0.087	0.087	92.8

Table 3.4: Two-stage estimates of blip parameters $(\psi_{20}, \psi_{21}, \psi_{10}, \psi_{11})$ with (σ_2, σ_1) = (0.2, 0.2) and 30% independent censoring

		ψ_{20}				ψ_{21}				ψ_{10}				ψ_{11}			
$\hat{\psi}$	Scenario	Bias	SE	RMSE	CP%	Bias	SE	RMSE	CP%	Bias	SE	RMSE	CP%	Bias	SE	RMSE	CP%
$\hat{\psi}^v$	1	0.000	0.143	0.143	95.6	0.002	0.097	0.097	95.6	-0.004	0.049	0.050	95.0	0.007	0.066	0.067	95.8
	2	0.000	0.144	0.144	95.4	0.001	0.098	0.098	96.0	-0.004	0.052	0.052	95.2	0.007	0.069	0.070	96.8
	3	0.001	0.141	0.141	95.6	0.001	0.096	0.096	96.0	-0.004	0.049	0.049	94.8	0.006	0.064	0.064	96.0
	4	-0.033	0.141	0.145	94.8	0.024	0.095	0.098	94.8	0.047	0.051	0.069	85.0	-0.067	0.067	0.095	83.0
$\hat{\psi}^c$	1	0.006	0.100	0.100	95.4	-0.003	0.068	0.068	95.4	-0.002	0.037	0.037	94.8	0.002	0.050	0.050	95.2
	2	0.005	0.101	0.101	95.6	-0.003	0.068	0.068	95.4	-0.002	0.039	0.039	94.2	0.002	0.053	0.053	95.8
	3	0.007	0.099	0.099	95.8	-0.004	0.067	0.067	96.4	-0.002	0.037	0.037	94.2	0.002	0.049	0.049	94.6
	4	-0.027	0.099	0.102	93.8	0.020	0.067	0.069	94.8	0.049	0.039	0.062	76.6	-0.072	0.051	0.088	72.0
$\hat{\psi}^n$	1	0.227	0.090	0.244	25.2	-0.166	0.060	0.177	19.8	-0.143	0.035	0.148	1.8	0.208	0.046	0.213	0.4
	2	0.230	0.090	0.247	24.8	-0.168	0.060	0.179	19.0	-0.148	0.036	0.153	1.6	0.215	0.047	0.220	0.4
	3	0.228	0.088	0.244	24.4	-0.167	0.059	0.177	17.4	-0.143	0.034	0.148	1.6	0.208	0.044	0.212	0.2
	4	0.217	0.088	0.234	29.0	-0.160	0.058	0.170	21.2	-0.123	0.035	0.128	5.8	0.178	0.045	0.184	3.2
$\hat{\psi}^k$	1	-0.007	0.106	0.106	95.0	0.002	0.072	0.072	95.2	0.004	0.039	0.039	92.8	-0.006	0.053	0.053	94.6
	2	-0.007	0.106	0.107	95.0	0.002	0.072	0.072	95.8	0.004	0.041	0.041	93.6	-0.006	0.055	0.055	94.8
	3	-0.006	0.105	0.105	94.8	0.002	0.071	0.071	95.4	0.004	0.039	0.039	93.2	-0.006	0.051	0.052	94.6
	4	-0.036	0.104	0.110	94.4	0.022	0.070	0.074	94.2	0.049	0.040	0.063	76.2	-0.071	0.053	0.089	74.6
$\hat{\psi}^w$	1	-0.003	0.106	0.107	95.0	-0.001	0.072	0.072	95.4	0.002	0.039	0.039	93.0	-0.003	0.053	0.053	94.6
	2	-0.004	0.107	0.107	95.0	0.000	0.072	0.072	95.4	0.003	0.041	0.041	94.8	-0.003	0.055	0.055	95.0
	3	-0.002	0.105	0.105	95.0	-0.001	0.071	0.071	95.8	0.002	0.039	0.039	93.8	-0.003	0.051	0.052	95.0
	4	-0.031	0.104	0.109	94.4	0.019	0.071	0.073	95.4	0.046	0.040	0.061	78.8	-0.066	0.053	0.085	76.4

Scenario 1: both models are correctly specified, Scenario 2: weight model is correctly specified, but treatment-free model is misspecified, Scenario 3: weight model is misspecified, but treatment-free model is correctly specified, Scenario 4: both models are misspecified

Table 3.5: Two-stage estimates of blip parameters $(\psi_{20}, \psi_{21}, \psi_{10}, \psi_{11})$ with (σ_2, σ_1) = (0.5, 0.5) and 30% independent censoring

		ψ_{20}				ψ_{21}				ψ_{10}				ψ_{11}			
$\hat{\psi}$	Scenario	Bias	SE	RMSE	CP%	Bias	SE	RMSE	CP%	Bias	SE	RMSE	CP%	Bias	SE	RMSE	CP%
$\hat{\psi}^v$	1	0.000	0.144	0.144	93.2	0.000	0.097	0.097	94.0	-0.001	0.049	0.049	94.8	0.003	0.066	0.066	94.0
	2	0.000	0.144	0.144	93.2	0.000	0.098	0.098	94.2	-0.002	0.052	0.052	95.8	0.004	0.069	0.069	95.4
	3	0.001	0.141	0.141	94.2	-0.001	0.095	0.095	94.2	-0.001	0.049	0.049	94.6	0.003	0.064	0.064	94.0
	4	-0.034	0.141	0.145	93.6	0.023	0.095	0.098	93.6	0.048	0.051	0.070	83.8	-0.069	0.067	0.096	80.8
$\hat{\psi}^c$	1	0.004	0.100	0.100	94.4	-0.003	0.068	0.068	94.2	0.000	0.037	0.037	93.0	0.001	0.050	0.050	93.2
	2	0.004	0.101	0.101	94.2	-0.003	0.068	0.068	93.8	0.000	0.039	0.039	93.6	0.001	0.052	0.052	94.2
	3	0.005	0.099	0.099	94.4	-0.003	0.067	0.067	94.8	0.000	0.037	0.037	94.2	0.001	0.049	0.049	94.2
	4	-0.029	0.099	0.103	94.0	0.020	0.066	0.069	94.0	0.050	0.039	0.063	76.0	-0.072	0.051	0.088	72.8
$\hat{\psi}^n$	1	0.577	0.063	0.581	0.0	-0.425	0.040	0.427	0.0	-0.411	0.028	0.411	0.0	0.596	0.031	0.597	0.0
	2	0.578	0.063	0.581	0.0	-0.425	0.040	0.427	0.0	-0.411	0.028	0.412	0.0	0.597	0.031	0.598	0.0
	3	0.577	0.062	0.581	0.0	-0.425	0.039	0.427	0.0	-0.410	0.027	0.411	0.0	0.596	0.030	0.597	0.0
	4	0.577	0.062	0.580	0.0	-0.425	0.039	0.426	0.0	-0.408	0.028	0.409	0.0	0.592	0.030	0.593	0.0
$\hat{\psi}^k$	1	-0.013	0.115	0.116	94.2	0.001	0.078	0.078	94.8	0.007	0.042	0.043	94.0	-0.008	0.057	0.058	94.0
	2	-0.013	0.115	0.116	94.0	0.001	0.078	0.078	94.8	0.006	0.044	0.044	94.2	-0.007	0.059	0.060	93.6
	3	-0.012	0.113	0.114	93.8	0.000	0.077	0.077	94.4	0.007	0.041	0.042	93.6	-0.008	0.055	0.056	93.4
	4	-0.041	0.112	0.120	93.6	0.020	0.076	0.079	94.4	0.049	0.043	0.065	79.6	-0.070	0.057	0.090	75.8
$\hat{\psi}^w$	1	-0.009	0.115	0.115	95.0	-0.002	0.078	0.078	93.8	0.004	0.042	0.042	93.0	-0.004	0.057	0.057	93.6
	2	-0.010	0.115	0.116	94.2	-0.002	0.078	0.078	94.2	0.004	0.044	0.044	94.2	-0.004	0.059	0.060	93.8
	3	-0.008	0.113	0.114	94.0	-0.003	0.077	0.077	94.8	0.004	0.042	0.042	93.4	-0.004	0.056	0.056	94.4
	4	-0.037	0.112	0.118	94.2	0.017	0.076	0.078	94.8	0.046	0.043	0.063	80.4	-0.066	0.057	0.087	78.4

Scenario 1: both models are correctly specified, Scenario 2: weight model is correctly specified, but treatment-free model is misspecified, Scenario 3: weight model is misspecified, but treatment-free model is correctly specified, Scenario 4: both models are misspecified

Table 3.6: Two-stage estimates of blip parameters $(\psi_{20}, \psi_{21}, \psi_{10}, \psi_{11})$ with (σ_2, σ_1) = (0.8, 0.8) and 30% independent censoring

		ψ_{20}				ψ_{21}				ψ_{10}				ψ_{11}			
$\hat{\psi}$	Scenario	Bias	SE	RMSE	CP%	Bias	SE	RMSE	CP%	Bias	SE	RMSE	CP%	Bias	SE	RMSE	CP%
$\hat{\psi}^v$	1	0.009	0.143	0.143	94.6	-0.008	0.097	0.097	93.8	-0.002	0.049	0.049	94.4	0.002	0.066	0.066	95.8
	2	0.009	0.143	0.144	94.4	-0.007	0.097	0.097	94.6	-0.002	0.051	0.052	94.4	0.002	0.069	0.069	94.8
	3	0.007	0.141	0.141	94.0	-0.007	0.095	0.095	93.6	-0.002	0.048	0.049	94.8	0.002	0.064	0.064	95.6
	4	-0.028	0.140	0.143	92.8	0.018	0.095	0.097	94.0	0.050	0.051	0.071	84.6	-0.073	0.067	0.099	80.8
$\hat{\psi}^c$	1	0.004	0.100	0.100	94.6	-0.003	0.068	0.068	94.4	0.000	0.037	0.037	95.6	0.000	0.050	0.050	95.6
	2	0.004	0.101	0.101	94.8	-0.002	0.068	0.068	94.4	0.001	0.039	0.039	96.4	-0.001	0.053	0.053	95.6
	3	0.003	0.099	0.099	94.0	-0.002	0.067	0.067	94.2	0.000	0.037	0.037	95.6	0.000	0.049	0.049	95.6
	4	-0.032	0.098	0.103	93.2	0.022	0.066	0.070	92.0	0.052	0.039	0.064	75.4	-0.075	0.051	0.091	68.4
$\hat{\psi}^n$	1	0.706	0.048	0.708	0.0	-0.519	0.028	0.520	0.0	-0.521	0.024	0.521	0.0	0.753	0.022	0.754	0.0
	2	0.706	0.048	0.708	0.0	-0.519	0.028	0.520	0.0	-0.521	0.024	0.522	0.0	0.754	0.022	0.754	0.0
	3	0.705	0.047	0.707	0.0	-0.519	0.028	0.520	0.0	-0.521	0.024	0.521	0.0	0.754	0.021	0.754	0.0
	4	0.706	0.047	0.707	0.0	-0.519	0.028	0.520	0.0	-0.520	0.024	0.521	0.0	0.753	0.021	0.753	0.0
$\hat{\psi}^k$	1	-0.011	0.117	0.118	94.6	-0.003	0.080	0.080	94.2	0.005	0.043	0.044	94.8	-0.005	0.059	0.059	95.0
	2	-0.010	0.118	0.118	94.6	-0.003	0.080	0.080	94.6	0.005	0.045	0.045	95.2	-0.006	0.061	0.061	94.0
	3	-0.011	0.116	0.116	95.0	-0.003	0.078	0.079	94.2	0.005	0.043	0.043	95.2	-0.005	0.057	0.057	95.2
	4	-0.039	0.115	0.121	93.4	0.016	0.078	0.080	94.0	0.049	0.044	0.066	81.0	-0.070	0.059	0.092	77.6
$\hat{\psi}^w$	1	-0.008	0.117	0.118	93.8	-0.005	0.080	0.080	94.6	0.003	0.043	0.043	94.2	-0.002	0.059	0.059	95.2
	2	-0.007	0.118	0.118	93.4	-0.005	0.080	0.080	94.4	0.003	0.045	0.045	95.2	-0.003	0.061	0.061	94.8
	3	-0.008	0.116	0.116	94.8	-0.005	0.078	0.079	94.0	0.003	0.043	0.043	94.2	-0.002	0.057	0.057	95.8
	4	-0.036	0.115	0.120	93.4	0.014	0.078	0.079	94.0	0.047	0.044	0.065	81.4	-0.068	0.059	0.090	78.4

Scenario 1: both models are correctly specified, Scenario 2: weight model is correctly specified, but treatment-free model is misspecified, Scenario 3: weight model is misspecified, but treatment-free model is correctly specified, Scenario 4: both models are misspecified

Table 3.7: Two-stage estimates of blip parameters $(\psi_{20}, \psi_{21}, \psi_{10}, \psi_{11})$ with $(\sigma_2, \sigma_1) = (0.2, 0.2)$ and 70% independent censoring

		ψ_{20}				ψ_{21}				ψ_{10}				ψ_{11}			
$\hat{\psi}$	Scenario	Bias	SE	RMSE	CP%	Bias	SE	RMSE	CP%	Bias	SE	RMSE	CP%	Bias	SE	RMSE	CP%
$\hat{\psi}^v$	1	0.005	0.216	0.216	94.2	-0.005	0.146	0.146	94.0	-0.007	0.076	0.076	95.4	0.007	0.101	0.101	95.8
	2	0.005	0.217	0.217	93.8	-0.005	0.147	0.147	94.8	-0.006	0.079	0.080	93.4	0.005	0.106	0.106	94.6
	3	0.004	0.214	0.214	93.4	-0.004	0.145	0.145	94.2	-0.006	0.074	0.075	95.4	0.005	0.098	0.098	95.8
	4	-0.031	0.214	0.216	94.2	0.020	0.144	0.146	93.8	0.044	0.078	0.090	91.6	-0.069	0.103	0.124	90.6
$\hat{\psi}^c$	1	0.003	0.152	0.152	94.2	-0.002	0.103	0.103	93.6	-0.006	0.058	0.058	95.4	0.005	0.077	0.077	95.6
	2	0.003	0.153	0.153	94.6	-0.002	0.103	0.103	94.2	-0.005	0.060	0.061	94.6	0.004	0.081	0.081	95.2
	3	0.000	0.150	0.150	93.8	0.000	0.101	0.101	93.0	-0.006	0.057	0.057	94.2	0.005	0.075	0.075	95.4
	4	-0.035	0.150	0.154	92.6	0.024	0.101	0.104	92.6	0.044	0.059	0.074	88.0	-0.069	0.078	0.104	85.4
$\hat{\psi}^n$	1	0.218	0.136	0.257	63.6	-0.162	0.090	0.185	57.4	-0.146	0.054	0.156	23.2	0.208	0.069	0.219	14.4
	2	0.222	0.136	0.261	63.6	-0.164	0.091	0.188	56.8	-0.151	0.055	0.161	22.2	0.214	0.071	0.226	14.0
	3	0.217	0.134	0.255	61.2	-0.161	0.090	0.184	56.2	-0.147	0.053	0.156	21.0	0.208	0.067	0.219	13.0
	4	0.207	0.133	0.246	63.6	-0.154	0.089	0.178	60.8	-0.127	0.054	0.138	33.6	0.179	0.068	0.192	25.4
$\hat{\psi}^k$	1	-0.009	0.160	0.161	94.4	0.003	0.108	0.108	94.8	0.000	0.060	0.060	95.4	-0.004	0.081	0.081	96.0
	2	-0.009	0.161	0.162	94.6	0.003	0.109	0.109	93.8	0.001	0.063	0.063	94.2	-0.006	0.084	0.084	95.8
	3	-0.012	0.159	0.159	94.4	0.005	0.107	0.107	94.4	0.000	0.059	0.059	94.0	-0.004	0.078	0.079	95.2
	4	-0.042	0.158	0.164	93.4	0.026	0.107	0.110	93.0	0.045	0.061	0.076	89.2	-0.069	0.081	0.107	86.4
$\hat{\psi}^w$	1	-0.004	0.161	0.161	94.6	-0.001	0.109	0.109	94.4	-0.002	0.060	0.060	94.8	-0.001	0.081	0.081	95.6
	2	-0.004	0.162	0.162	94.4	-0.001	0.109	0.109	94.0	-0.001	0.063	0.063	94.8	-0.003	0.084	0.084	96.8
	3	-0.007	0.159	0.160	94.0	0.001	0.108	0.108	94.4	-0.002	0.059	0.059	94.6	-0.001	0.079	0.079	95.6
	4	-0.036	0.159	0.163	93.4	0.022	0.107	0.110	93.8	0.042	0.061	0.074	90.2	-0.065	0.082	0.104	88.0

Scenario 1: both models are correctly specified, Scenario 2: weight model is correctly specified, but treatment-free model is misspecified, Scenario 3: weight model is misspecified, but treatment-free model is correctly specified, Scenario 4: both models are misspecified

Table 3.8: Two-stage estimates of blip parameters $(\psi_{20}, \psi_{21}, \psi_{10}, \psi_{11})$ with (σ_2, σ_1) = (0.5, 0.5) and 70% independent censoring

		ψ_{20}				ψ_{21}				ψ_{10}				ψ_{11}			
$\hat{\psi}$	Scenario	Bias	SE	RMSE	CP%	Bias	SE	RMSE	CP%	Bias	SE	RMSE	CP%	Bias	SE	RMSE	CP%
$\hat{\psi}^v$	1	-0.008	0.217	0.217	93.8	0.002	0.147	0.147	93.4	-0.001	0.075	0.075	93.2	-0.002	0.101	0.101	93.4
	2	-0.011	0.218	0.219	93.4	0.004	0.148	0.148	93.8	-0.002	0.079	0.079	93.2	0.000	0.107	0.107	93.2
	3	-0.006	0.215	0.215	93.2	0.000	0.145	0.145	93.2	0.000	0.074	0.074	92.6	-0.002	0.098	0.098	92.8
	4	-0.042	0.214	0.218	94.0	0.025	0.145	0.147	93.6	0.048	0.078	0.091	88.0	-0.073	0.103	0.126	87.4
$\hat{\psi}^c$	1	0.003	0.152	0.152	95.8	-0.004	0.103	0.103	95.0	-0.002	0.057	0.057	93.6	0.001	0.076	0.076	94.2
	2	0.002	0.153	0.153	95.6	-0.003	0.103	0.103	95.0	-0.003	0.060	0.060	93.4	0.002	0.081	0.081	93.8
	3	0.004	0.150	0.150	95.4	-0.004	0.102	0.102	94.2	-0.002	0.056	0.056	93.2	0.000	0.074	0.074	93.0
	4	-0.031	0.150	0.153	93.8	0.020	0.101	0.103	94.4	0.047	0.059	0.075	84.2	-0.072	0.078	0.106	84.6
$\hat{\psi}^n$	1	0.576	0.096	0.583	0.0	-0.426	0.061	0.430	0.0	-0.416	0.042	0.418	0.0	0.599	0.047	0.601	0.0
	2	0.576	0.096	0.584	0.2	-0.426	0.060	0.430	0.0	-0.416	0.043	0.419	0.0	0.600	0.047	0.602	0.0
	3	0.576	0.095	0.584	0.0	-0.426	0.060	0.430	0.0	-0.416	0.042	0.418	0.0	0.600	0.046	0.601	0.0
	4	0.576	0.094	0.583	0.0	-0.426	0.060	0.430	0.0	-0.413	0.042	0.415	0.0	0.595	0.046	0.597	0.0
$\hat{\psi}^k$	1	-0.008	0.175	0.175	95.8	-0.004	0.118	0.119	96.2	0.006	0.064	0.065	92.6	-0.010	0.087	0.088	94.6
	2	-0.010	0.175	0.176	95.6	-0.002	0.119	0.119	96.2	0.005	0.067	0.067	93.2	-0.009	0.091	0.092	94.6
	3	-0.009	0.173	0.173	96.4	-0.003	0.117	0.117	96.4	0.006	0.063	0.064	93.0	-0.011	0.085	0.086	93.4
	4	-0.036	0.171	0.175	94.8	0.016	0.116	0.117	94.0	0.048	0.065	0.081	88.0	-0.072	0.088	0.113	86.2
$\hat{\psi}^w$	1	-0.005	0.175	0.175	95.8	-0.006	0.119	0.119	95.6	0.004	0.065	0.065	94.2	-0.006	0.087	0.088	94.6
	2	-0.007	0.175	0.175	95.6	-0.005	0.119	0.119	95.4	0.003	0.067	0.067	93.4	-0.005	0.091	0.091	94.4
	3	-0.005	0.173	0.173	95.6	-0.006	0.117	0.117	96.2	0.004	0.063	0.063	92.8	-0.007	0.085	0.085	93.8
	4	-0.033	0.171	0.175	95.0	0.013	0.116	0.117	94.2	0.045	0.065	0.079	88.4	-0.067	0.088	0.111	87.2

Scenario 1: both models are correctly specified, Scenario 2: weight model is correctly specified, but treatment-free model is misspecified, Scenario 3: weight model is misspecified, but treatment-free model is correctly specified, Scenario 4: both models are misspecified

Table 3.9: Two-stage estimates of blip parameters $(\psi_{20}, \psi_{21}, \psi_{10}, \psi_{11})$ with (σ_2, σ_1) = (0.8, 0.8) and 70% independent censoring

		ψ_{20}				ψ_{21}				ψ_{10}				ψ_{11}			
$\hat{\psi}$	Scenario	Bias	SE	RMSE	CP%	Bias	SE	RMSE	CP%	Bias	SE	RMSE	CP%	Bias	SE	RMSE	CP%
$\hat{\psi}^v$	1	0.008	0.217	0.217	91.6	-0.008	0.147	0.147	91.4	0.004	0.076	0.076	95.6	-0.004	0.101	0.101	94.6
	2	0.007	0.218	0.218	92.8	-0.007	0.147	0.148	92.0	0.004	0.079	0.079	95.0	-0.003	0.106	0.106	95.2
	3	0.008	0.214	0.215	92.0	-0.008	0.145	0.145	91.8	0.005	0.074	0.074	95.8	-0.004	0.098	0.098	94.0
	4	-0.027	0.214	0.216	93.0	0.017	0.145	0.146	92.0	0.054	0.078	0.094	88.0	-0.076	0.103	0.128	88.4
$\hat{\psi}^c$	1	0.009	0.152	0.152	94.2	-0.007	0.103	0.103	94.6	0.002	0.057	0.057	94.8	0.001	0.076	0.076	95.4
	2	0.008	0.153	0.153	94.6	-0.007	0.104	0.104	94.6	0.001	0.060	0.060	93.8	0.001	0.080	0.080	95.2
	3	0.009	0.151	0.151	94.8	-0.007	0.102	0.102	95.0	0.002	0.056	0.056	95.0	0.000	0.074	0.074	95.0
	4	-0.025	0.150	0.152	95.0	0.016	0.101	0.103	94.6	0.052	0.059	0.079	83.2	-0.073	0.078	0.107	85.4
$\hat{\psi}^n$	1	0.703	0.073	0.707	0.0	-0.518	0.043	0.520	0.0	-0.517	0.037	0.518	0.0	0.752	0.034	0.752	0.0
	2	0.703	0.073	0.707	0.0	-0.518	0.043	0.520	0.0	-0.517	0.037	0.518	0.0	0.752	0.034	0.753	0.0
	3	0.703	0.072	0.707	0.0	-0.518	0.043	0.520	0.0	-0.517	0.037	0.518	0.0	0.752	0.032	0.752	0.0
	4	0.703	0.072	0.707	0.0	-0.518	0.043	0.520	0.0	-0.516	0.037	0.517	0.0	0.751	0.033	0.751	0.0
$\hat{\psi}^k$	1	-0.006	0.178	0.178	94.0	-0.007	0.121	0.121	93.4	0.010	0.066	0.067	94.4	-0.009	0.090	0.090	96.0
	2	-0.006	0.179	0.179	94.6	-0.007	0.122	0.122	93.8	0.009	0.069	0.069	94.6	-0.009	0.094	0.094	96.2
	3	-0.006	0.176	0.176	94.8	-0.007	0.120	0.120	94.0	0.010	0.065	0.066	93.6	-0.009	0.087	0.088	95.2
	4	-0.035	0.175	0.178	94.4	0.013	0.119	0.120	94.4	0.053	0.067	0.085	84.2	-0.072	0.090	0.115	87.8
$\hat{\psi}^w$	1	-0.004	0.178	0.178	94.6	-0.008	0.121	0.121	93.2	0.008	0.066	0.067	94.6	-0.006	0.090	0.090	95.8
	2	-0.004	0.179	0.179	94.4	-0.008	0.122	0.122	93.8	0.007	0.069	0.069	93.8	-0.006	0.094	0.094	95.6
	3	-0.004	0.176	0.176	94.8	-0.008	0.120	0.120	94.0	0.008	0.065	0.066	94.2	-0.006	0.087	0.088	95.8
	4	-0.032	0.175	0.178	94.4	0.011	0.119	0.120	94.6	0.051	0.067	0.084	84.8	-0.069	0.090	0.113	88.6

Scenario 1: both models are correctly specified, Scenario 2: weight model is correctly specified, but treatment-free model is misspecified, Scenario 3: weight model is misspecified, but treatment-free model is correctly specified, Scenario 4: both models are misspecified

Table 3.10: Prediction accuracy of optimal DTR (%)

			Stage 2					Stage 1					Stage 2 & Stage 1				
%Censoring	ρ	(σ_2, σ_1)	v	c	n	k	w	v	c	n	k	w	v	c	n	k	w
30%	0.5	(0.2, 0.2)	94.3	97.4	94.3	97.2	97.1	97.4	98.6	94.9	98.5	98.5	91.8	96.0	89.5	95.7	95.6
		(0.5, 0.5)	94.1	97.1	69.5	96.0	95.9	97.4	98.5	69.5	98.3	98.3	91.7	95.7	48.4	94.4	94.2
		(0.8, 0.8)	94.5	97.4	56.2	95.6	95.6	97.4	98.5	66.3	98.2	98.2	92.0	96.0	37.3	93.9	93.8
	0.7	(0.2, 0.2)	96.5	97.4	94.3	97.3	97.2	98.1	98.6	94.9	98.5	98.5	94.6	96.0	89.5	95.8	95.8
		(0.5, 0.5)	96.4	97.1	69.6	96.7	96.6	98.1	98.5	69.5	98.4	98.3	94.5	95.7	48.3	95.1	95.0
		(0.8, 0.8)	96.1	97.4	56.2	96.6	96.6	98.0	98.5	66.3	98.3	98.3	94.2	95.9	37.2	94.9	94.9
70%	0.5	(0.2, 0.2)	89.4	95.9	92.3	95.6	95.5	96.0	97.8	94.7	97.7	97.7	85.9	93.7	87.5	93.4	93.3
		(0.5, 0.5)	89.6	95.9	69.2	94.3	94.2	96.1	97.8	70.8	97.5	97.5	86.1	93.8	49.1	91.9	91.8
		(0.8, 0.8)	90.0	95.6	58.9	93.3	93.2	96.1	97.8	66.3	97.4	97.4	86.5	93.4	39.1	90.8	90.8
	0.7	(0.2, 0.2)	94.0	95.9	92.3	95.8	95.8	97.0	97.8	94.8	97.7	97.7	91.2	93.7	87.5	93.7	93.6
		(0.5, 0.5)	93.9	95.9	69.3	95.3	95.3	97.0	97.8	70.8	97.6	97.6	91.1	93.8	49.1	93.1	93.0
		(0.8, 0.8)	93.2	95.6	58.9	94.7	94.7	97.1	97.8	66.3	97.6	97.6	90.5	93.4	39.0	92.4	92.4

v: validation estimator, c: complete estimator, n: naive estimator, k: *k*NN estimator, w: WLS estimator

Table 3.11: Prediction of the expected optimal log-survival times (standard deviations)

% Censoring	ρ	(σ_2, σ_1)	v	c	n	k	w
30%	0.5	(0.2, 0.2)	7.166 (0.020)	7.166 (0.013)	7.143 (0.012)	7.154 (0.013)	7.152 (0.014)
		(0.5, 0.5)	7.166 (0.020)	7.166 (0.013)	7.094 (0.014)	7.136 (0.015)	7.135 (0.015)
		(0.8, 0.8)	7.167 (0.019)	7.167 (0.012)	7.085 (0.015)	7.131 (0.016)	7.130 (0.016)
	0.7	(0.2, 0.2)	7.166 (0.016)	7.166 (0.013)	7.143 (0.012)	7.158 (0.013)	7.157 (0.013)
		(0.5, 0.5)	7.166 (0.015)	7.166 (0.013)	7.094 (0.014)	7.146 (0.013)	7.145 (0.013)
		(0.8, 0.8)	7.167 (0.016)	7.167 (0.012)	7.085 (0.015)	7.142 (0.013)	7.142 (0.013)
70%	0.5	(0.2, 0.2)	7.168 (0.031)	7.166 (0.019)	7.144 (0.017)	7.154 (0.019)	7.152 (0.019)
		(0.5, 0.5)	7.169 (0.029)	7.168 (0.017)	7.095 (0.020)	7.137 (0.021)	7.136 (0.021)
		(0.8, 0.8)	7.169 (0.029)	7.168 (0.018)	7.088 (0.020)	7.134 (0.022)	7.133 (0.022)
	0.7	(0.2, 0.2)	7.166 (0.024)	7.166 (0.018)	7.144 (0.017)	7.159 (0.019)	7.157 (0.018)
		(0.5, 0.5)	7.168 (0.023)	7.168 (0.017)	7.095 (0.020)	7.148 (0.019)	7.147 (0.019)
		(0.8, 0.8)	7.169 (0.022)	7.168 (0.017)	7.088 (0.020)	7.144 (0.020)	7.144 (0.020)

v: validation estimator, c: complete estimator, n: naive estimator, k: *k*NN estimator, w: WLS estimator

Table 3.12: Analysis results of the MIMIC-III data for the two-stage estimation of blip parameters

	Validation			*k*NN			WLS		
Variables	Est	SE	95%CI	Est	SE	95%CI	Est	SE	95%CI
A_2	1.890	0.675	(0.567, 3.212)	0.735	0.409	(-0.067, 1.537)	0.726	0.410	(-0.079, 1.530)
A_2A_1	-0.703	0.564	(-1.809, 0.403)	-0.271	0.356	(-0.968, 0.427)	-0.263	0.358	(-0.964, 0.438)
A_2VENT_2	-0.862	0.258	(-1.369, -0.356)	-0.340	0.248	(-0.826, 0.145)	-0.333	0.248	(-0.819, 0.152)
A_2PEEP_2	0.473	0.476	(-0.460, 1.406)	0.118	0.259	(-0.389, 0.626)	0.143	0.260	(-0.367, 0.652)
A_1	0.448	0.099	(0.255, 0.642)	0.158	0.090	(-0.017, 0.334)	0.159	0.090	(-0.017, 0.334)
A_1VENT_1	0.241	0.052	(0.140, 0.342)	0.788	0.060	(0.670, 0.906)	0.789	0.060	(0.671, 0.907)
A_1PEEP_1	0.132	0.084	(-0.033, 0.297)	0.103	0.065	(-0.025, 0.231)	0.102	0.065	(-0.026, 0.229)

Est: estimates, SE: asymptotic standard error, CI: confidence interval

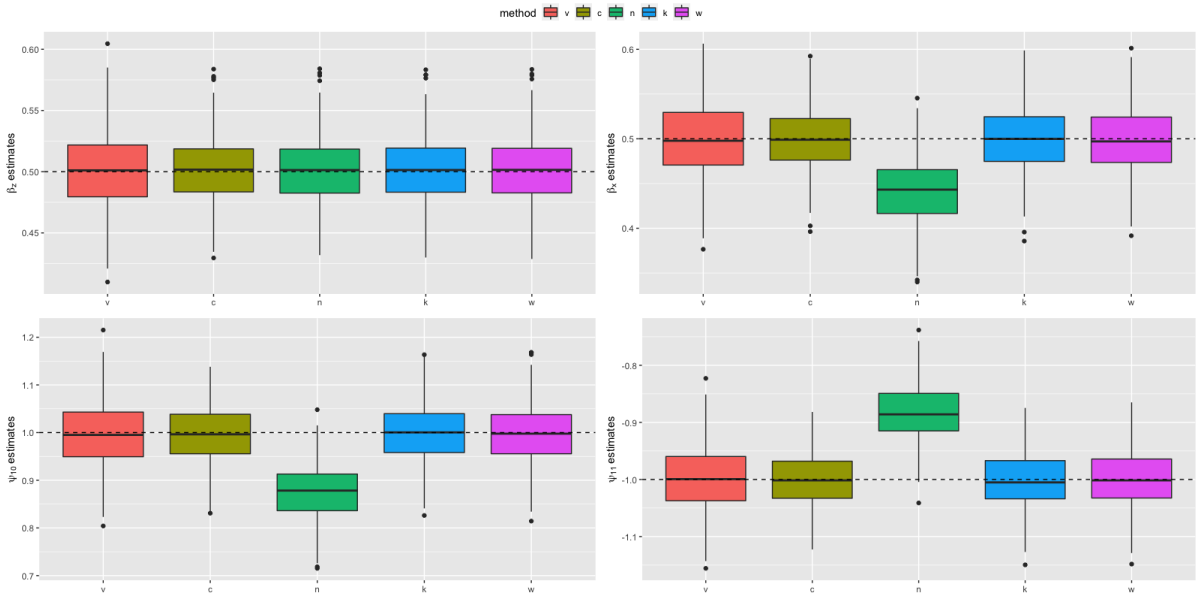

Figure 3.1: One-stage estimates of $(\beta_z, \beta_x, \psi_{10}, \psi_{11})$ with $\rho = 0.7$, $\sigma = 0.2$ and 30% independent censoring

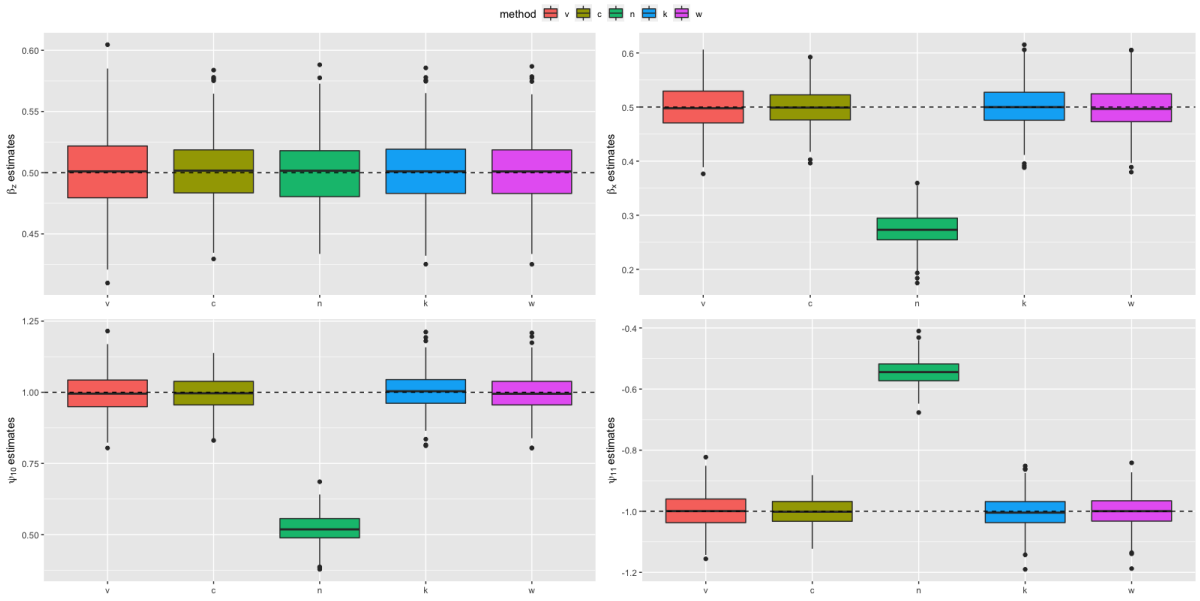

Figure 3.2: One-stage estimates of $(\beta_z, \beta_x, \psi_{10}, \psi_{11})$ with $\rho = 0.7$, $\sigma = 0.5$ and 30% independent censoring

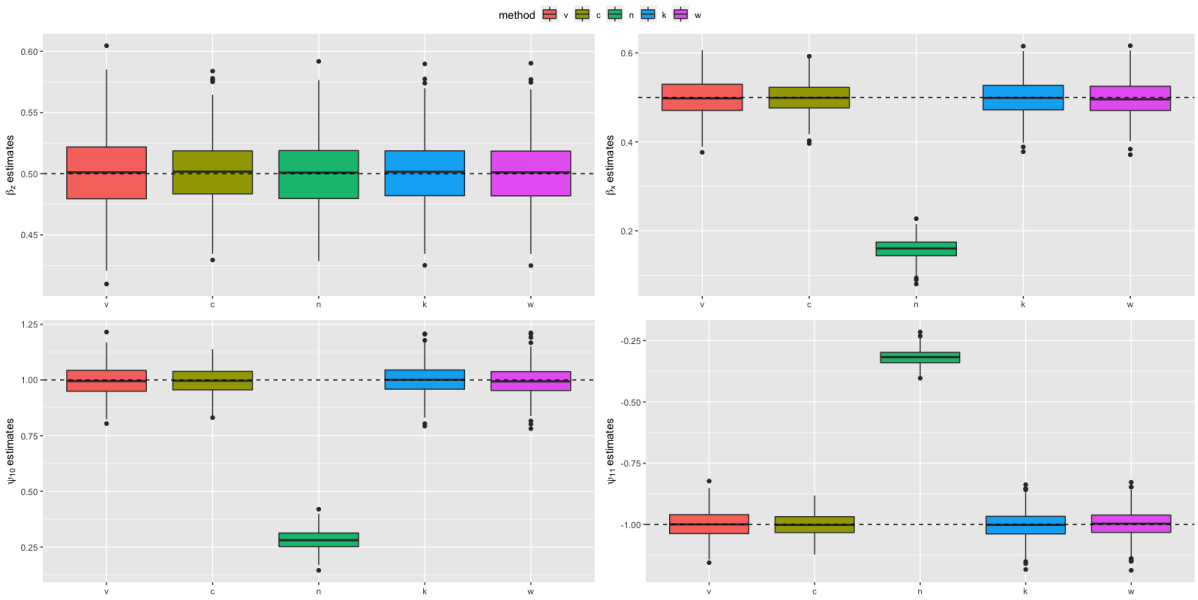

Figure 3.3: One-stage estimates of $(\beta_z, \beta_x, \psi_{10}, \psi_{11})$ with $\rho = 0.7$, $\sigma = 0.8$ and 30% independent censoring

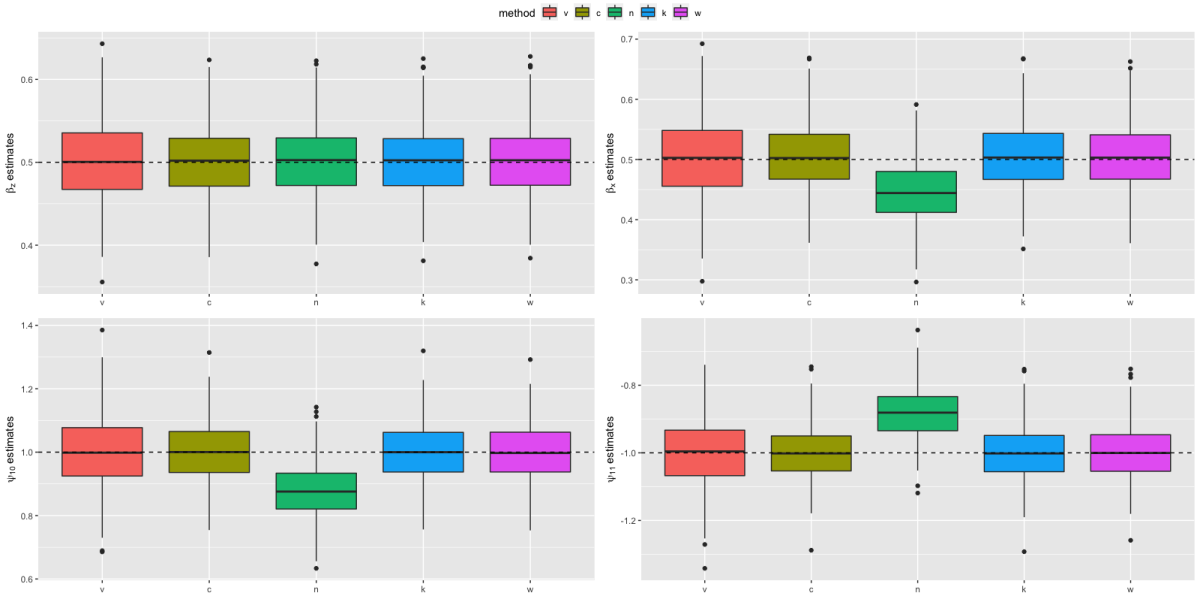

Figure 3.4: One-stage estimates of $(\beta_z, \beta_x, \psi_{10}, \psi_{11})$ with $\rho = 0.7$, $\sigma = 0.2$ and 70% independent censoring

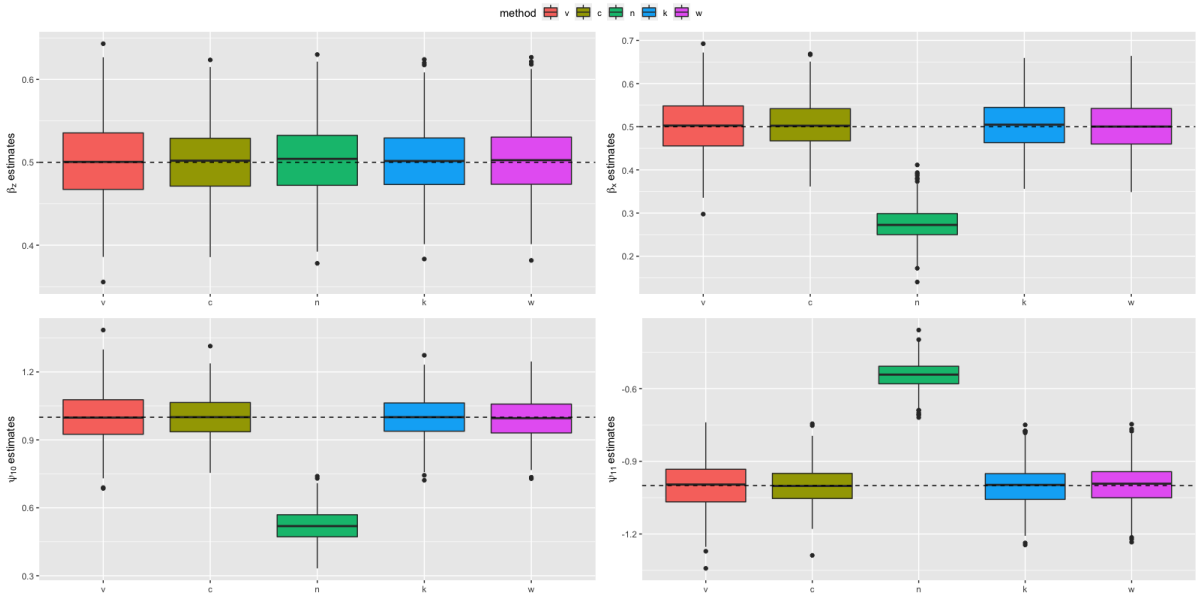

Figure 3.5: One-stage estimates of $(\beta_z, \beta_x, \psi_{10}, \psi_{11})$ with $\rho = 0.7$, $\sigma = 0.5$ and 70% independent censoring

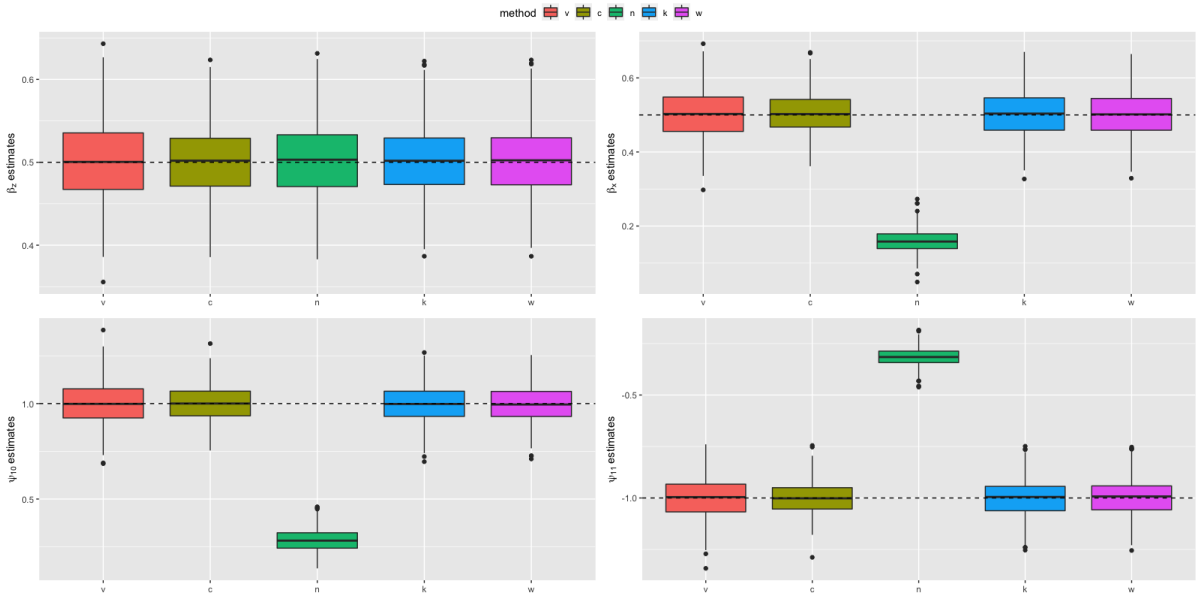

Figure 3.6: One-stage estimates of $(\beta_z, \beta_x, \psi_{10}, \psi_{11})$ with $\rho = 0.7$, $\sigma = 0.8$ and 70% independent censoring

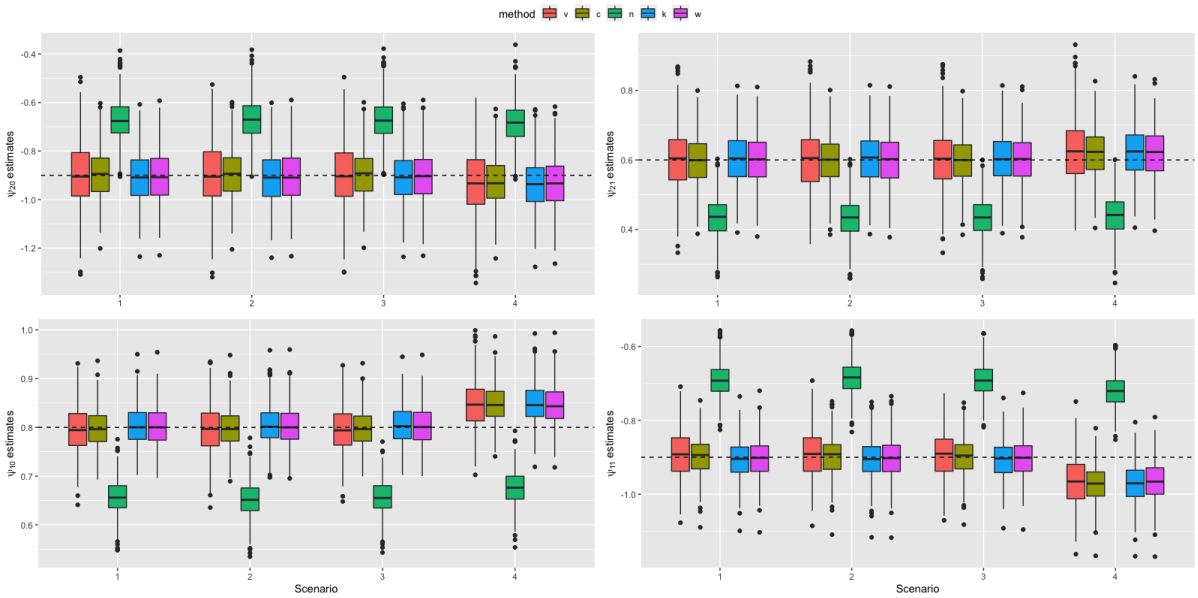

Figure 3.7: Two-stage estimates of (ψ_{20}, ψ_{21}, ψ_{10}, ψ_{11}) with $\rho = 0.7$, (σ_2, σ_1) = (0.2, 0.2) and 30% independent censoring under four scenarios: (1) both models correctly specified, (2) weight model correctly specified, but treatment-free model misspecified, (3) weight model misspecified, but treatment-free model correctly specified, (4) both models misspecified

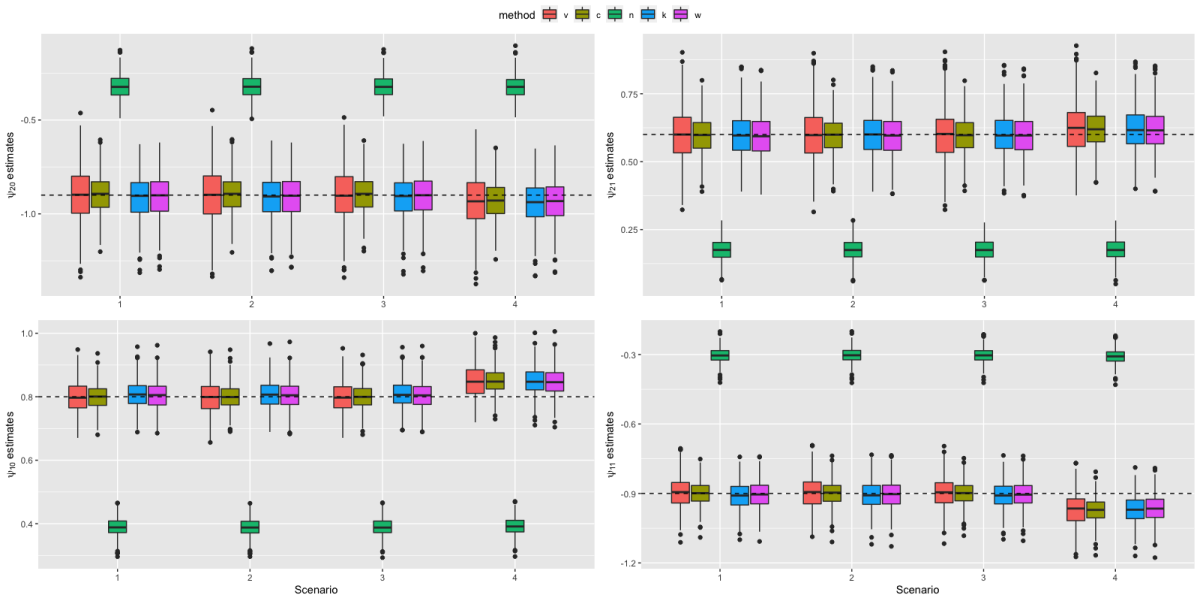

Figure 3.8: Two-stage estimates of (ψ_{20}, ψ_{21}, ψ_{10}, ψ_{11}) with $\rho = 0.7$, (σ_2, σ_1) = (0.5, 0.5) and 30% independent censoring under four scenarios: (1) both models correctly specified, (2) weight model correctly specified, but treatment-free model misspecified, (3) weight model misspecified, but treatment-free model correctly specified, (4) both models misspecified

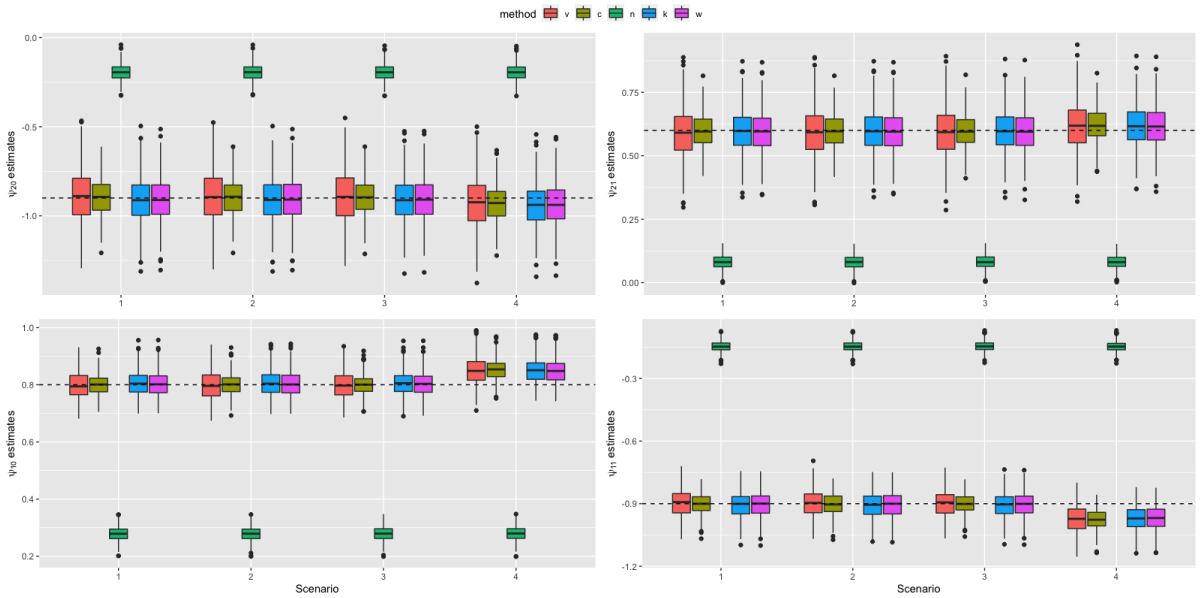

Figure 3.9: Two-stage estimates of (ψ_{20}, ψ_{21}, ψ_{10}, ψ_{11}) with $\rho = 0.7$, (σ_2, σ_1) = (0.8, 0.8) and 30% independent censoring under four scenarios: (1) both models correctly specified, (2) weight model correctly specified, but treatment-free model misspecified, (3) weight model misspecified, but treatment-free model correctly specified, (4) both models misspecified

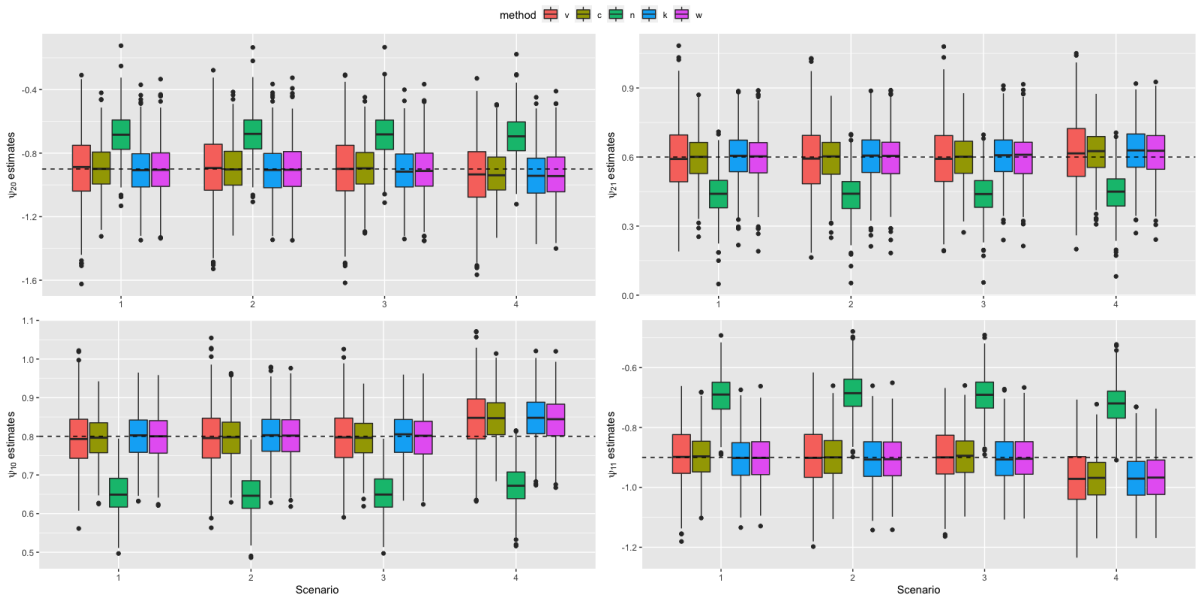

Figure 3.10: Two-stage estimates of (ψ_{20}, ψ_{21}, ψ_{10}, ψ_{11}) with $\rho = 0.7$, (σ_2, σ_1) = (0.2, 0.2) and 70% independent censoring under four scenarios: (1) both models correctly specified, (2) weight model correctly specified, but treatment-free model misspecified, (3) weight model misspecified, but treatment-free model correctly specified, (4) both models misspecified

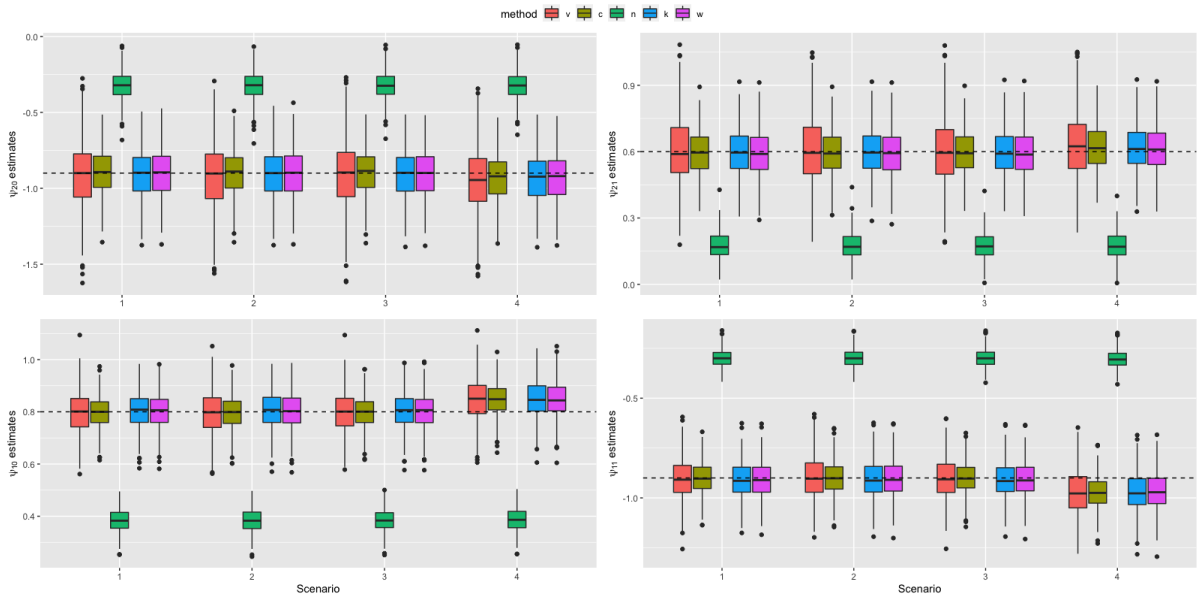

Figure 3.11: Two-stage estimates of $(\psi_{20}, \psi_{21}, \psi_{10}, \psi_{11})$ with $\rho = 0.7$, $(\sigma_2, \sigma_1) = (0.5, 0.5)$ and 70% independent censoring under four scenarios: (1) both models correctly specified, (2) weight model correctly specified, but treatment-free model misspecified, (3) weight model misspecified, but treatment-free model correctly specified, (4) both models misspecified

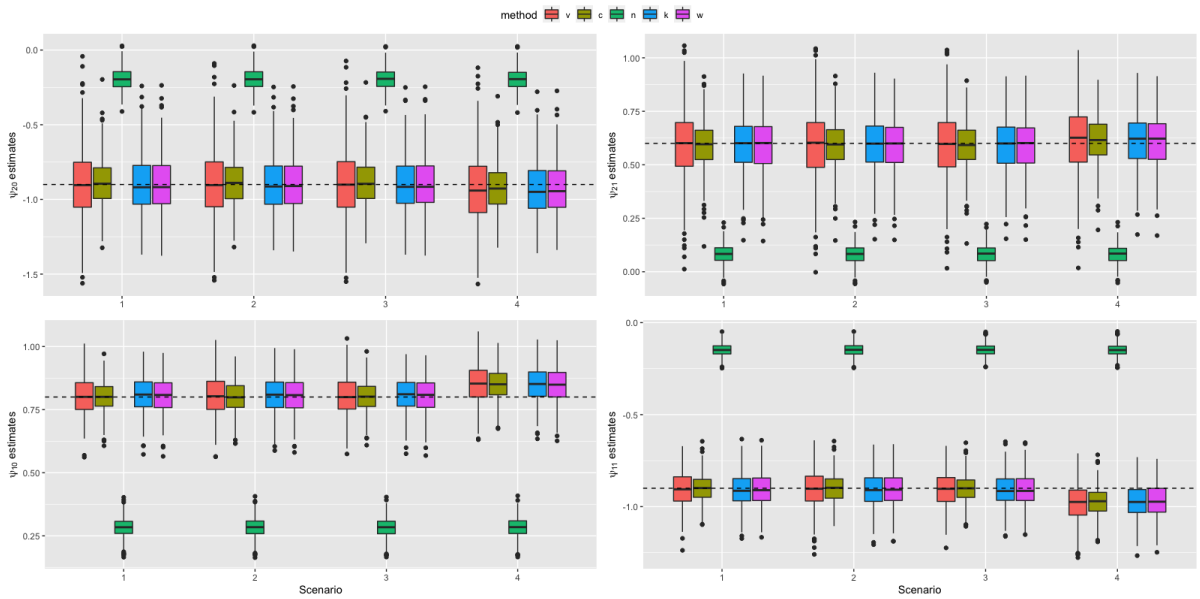

Figure 3.12: Two-stage estimates of $(\psi_{20}, \psi_{21}, \psi_{10}, \psi_{11})$ with $\rho = 0.7$, $(\sigma_2, \sigma_1) = (0.8, 0.8)$ and 70% independent censoring under four scenarios: (1) both models correctly specified, (2) weight model correctly specified, but treatment-free model misspecified, (3) weight model misspecified, but treatment-free model correctly specified, (4) both models misspecified

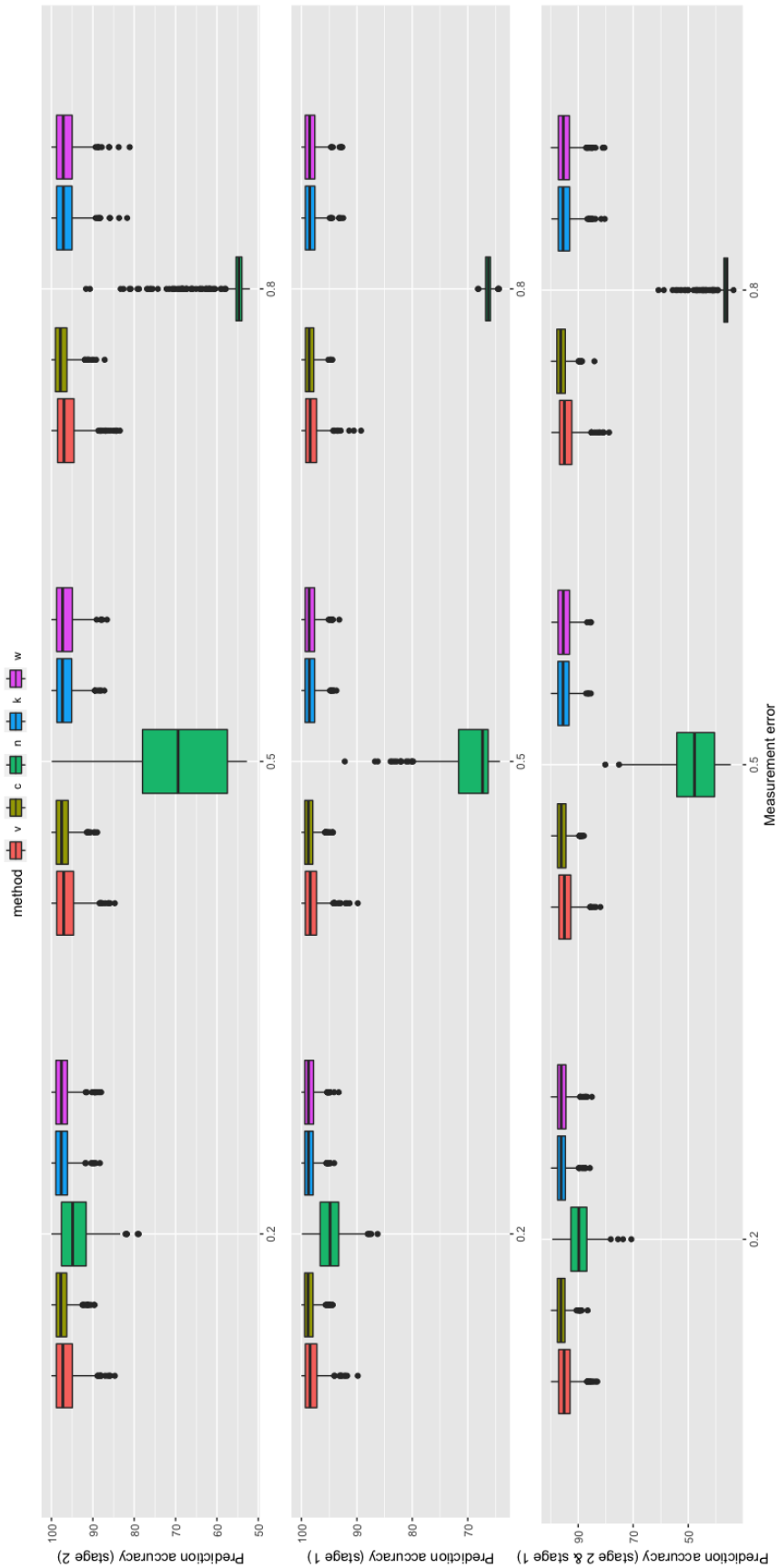

Figure 3.13: Prediction accuracy of optimal DTR with $\rho = 0.7$ and 30% independent censoring

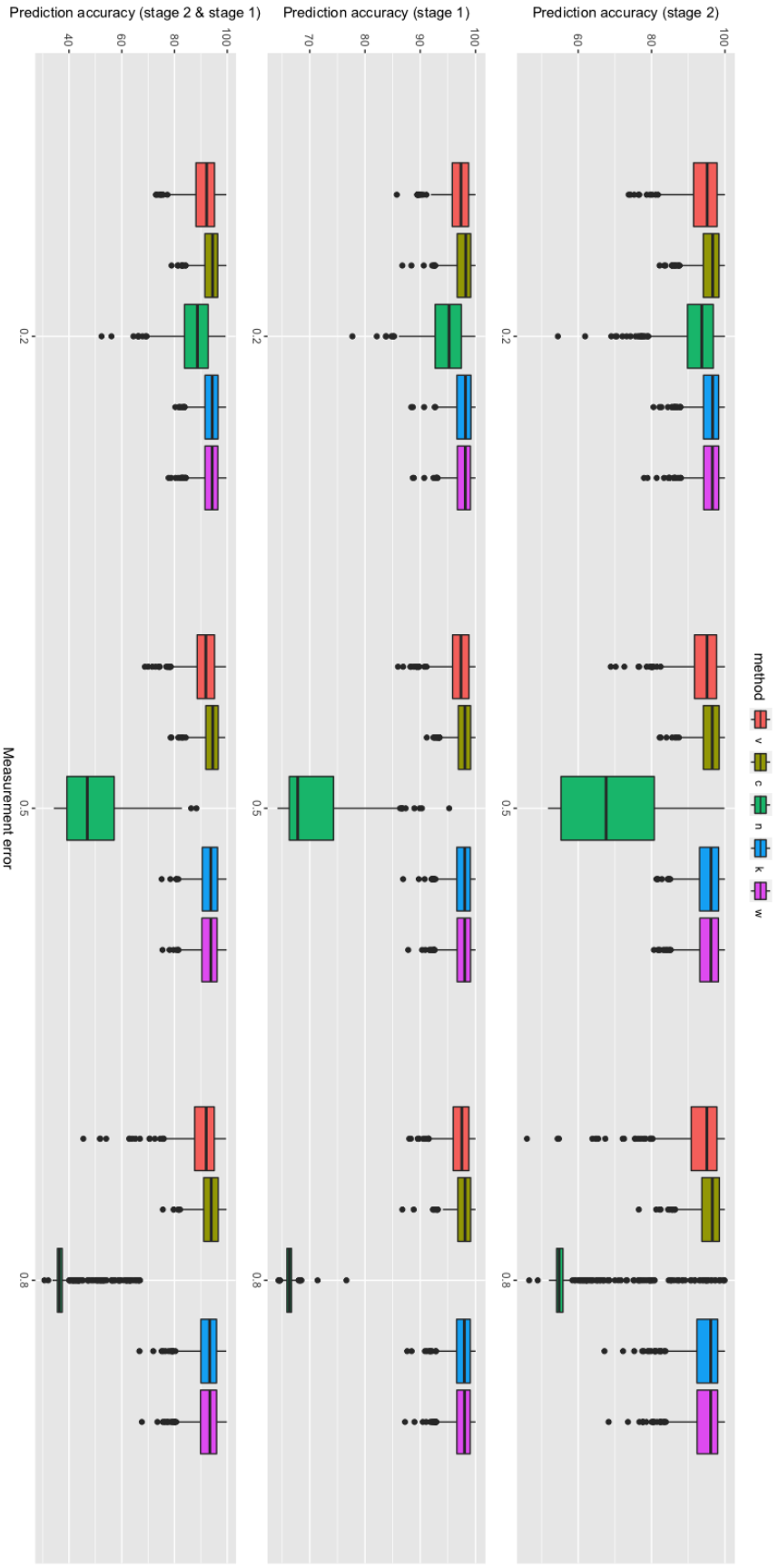

Figure 3.14: Prediction accuracy of optimal DTR with $\rho = 0.7$ and 70% independent censoring

3.6 Appendix

The proof of consistency in this section is based on a one-stage setting and it can be intuitively extended to multiple stages. The conditions for the property of consistency in DWSurv include:

(C1) Assumptions (A1) - (A4) in Sections (1.1.1) and (1.1.4) hold.

(C2) The weights satisfy the balancing property (1.9).

(C3) $E|X_i| < \infty$, and the choice of k satisfies that $\lim_{n_v \to \infty} k = \infty$ and $\lim_{n_v \to \infty} k/n_v = 0$.

The condition (C1) includes the assumptions that are necessary for the dynamic weighted survival modeling. The condition (C2) contains the requirement for choosing the weights in DWSurv, as described in Section (1.1.4). The condition (C3) assumes that the covariate is bounded, and the conditions the choice of k satisfies for the consistency of the kNN estimates.

Proof of Theorem 3.2.1:

For $\eta_1 = 1$, the original estimating equation is

$$U(\boldsymbol{\beta}, \boldsymbol{\psi}) = \sum_{i=1}^{n} \delta_i \hat{w}_i \begin{pmatrix} \boldsymbol{h_{i\beta}} \\ a_i \boldsymbol{h_{i\psi}} \end{pmatrix} \left(\log T_i - \boldsymbol{\beta^T h_{i\beta}} - a_i \boldsymbol{\psi^T h_{i\psi}} \right) = 0.$$

Let G(·) be defined as a function such that

$$G(X_i, \boldsymbol{Z_i}, A_i, T_i; \boldsymbol{\beta}, \boldsymbol{\psi}) = \begin{pmatrix} \boldsymbol{h_{i\beta}} \\ a_i \boldsymbol{h_{i\psi}} \end{pmatrix} \left(\log T_i - \boldsymbol{\beta^T h_{i\beta}} - a_i \boldsymbol{\psi^T h_{i\psi}} \right).$$

Then, the estimating equation can be further written as

$$U(\boldsymbol{\beta}, \boldsymbol{\psi}) = \sum_{i=1}^{n} \delta_i \hat{w}_i G(X_i, \boldsymbol{Z_i}, A_i, T_i; \boldsymbol{\beta}, \boldsymbol{\psi}) = 0.$$

Then, the estimating equation using the proposed kNN method is given by

$$\begin{aligned}
\hat{U}(\boldsymbol{\beta}, \boldsymbol{\psi}) &= \sum_{i=1}^{n} \delta_i \hat{w}_i \, G(\widetilde{X}_i^k, \boldsymbol{Z_i}, A_i, T_i; \boldsymbol{\beta}, \boldsymbol{\psi}) \\
&= \sum_{i \in V} \delta_i \hat{w}_i \, G(X_i, \boldsymbol{Z_i}, A_i, T_i; \boldsymbol{\beta}, \boldsymbol{\psi}) + \sum_{i \in \overline{V}} \delta_i \hat{w}_i \, G(\hat{X}_i^k, \boldsymbol{Z_i}, A_i, T_i; \boldsymbol{\beta}, \boldsymbol{\psi}) \\
&= \sum_{i=1}^{n} \delta_i \hat{w}_i \, G(X_i^*, \boldsymbol{Z_i}, A_i, T_i; \boldsymbol{\beta}, \boldsymbol{\psi}) + \sum_{i \in \overline{V}} \delta_i \hat{w}_i \left[G(\bar{X}_i^k, \boldsymbol{Z_i}, A_i, T_i; \boldsymbol{\beta}, \boldsymbol{\psi}) - G(\bar{X}_i, \boldsymbol{Z_i}, A_i, T_i; \boldsymbol{\beta}, \boldsymbol{\psi}) \right] \\
&= U(\boldsymbol{\beta}, \boldsymbol{\psi}) + \sum_{i \in \overline{V}} \delta_i \hat{w}_i \left[G(\hat{X}_i^k, \boldsymbol{Z_i}, A_i, T_i; \boldsymbol{\beta}, \boldsymbol{\psi}) - G(\bar{X}_i, \boldsymbol{Z_i}, A_i, T_i; \boldsymbol{\beta}, \boldsymbol{\psi}) \right].
\end{aligned}$$

By Cauchy-Schwarz inequality, for l = 1, 2, ..., L in $\overline{V}$,

$$\left\| \sum_{i\in\overline{V}} \delta_i \hat{w}_i \left[G(\hat{X}_i^k, \boldsymbol{Z_i}, A_i, T_i; \boldsymbol{\beta}, \boldsymbol{\psi}) - G(\bar{X}_i, \boldsymbol{Z_i}, A_i, T_i; \boldsymbol{\beta}, \boldsymbol{\psi}) \right] \right\| =$$

$$\left\| \sum_{l=1}^{L} \left[G(\hat{X}_l^k, \boldsymbol{Z_l}, A_l, T_l; \boldsymbol{\beta}, \boldsymbol{\psi}) - G(\bar{X}_l, \boldsymbol{Z_l}, A_l, T_l; \boldsymbol{\beta}, \boldsymbol{\psi}) \right] \sum_{i\in\overline{V}} \delta_i \hat{w}_i \mathbb{1}(W_l \text{ is near } W_i) \right\|$$

$$\leq \sum_{l=1}^{L} \left\| G(\hat{X}_l^k, \boldsymbol{Z_l}, A_l, T_l; \boldsymbol{\beta}, \boldsymbol{\psi}) - G(\bar{X}_l, \boldsymbol{Z_l}, A_l, T_l; \boldsymbol{\beta}, \boldsymbol{\psi}) \right\| \cdot \left\| \sum_{i\in\overline{V}} \delta_i \hat{w}_i \mathbb{1}(W_l \text{ is near } W_i) \right\|.$$

According to Devroye et al. (1994), for $E|X_i| < \infty$ and the choice of k satisfies $\lim_{n_v\to\infty} k = \infty$ and $\lim_{n_v\to\infty} k/n_v = 0$, we have $\left\|\hat{X}_l^k - \bar{X}_i\right\| \xrightarrow{P} 0 \;\; as\; n \to \infty$. By continuous mapping theorem, we obtain

$$\left\|G(\hat{X}_l^k, \boldsymbol{Z_l}, A_l, T_l; \boldsymbol{\beta}, \boldsymbol{\psi}) - G(\bar{X}_l, \boldsymbol{Z_l}, A_l, T_l; \boldsymbol{\beta}, \boldsymbol{\psi})\right\| \xrightarrow{P} 0 \quad as\; n \to \infty.$$

For some finite M, $\left\| \sum_{i\in\overline{V}} \delta_i \hat{w}_i \mathbb{1}(W_l \text{ is near } W_i) \right\| < M < \infty$. Therefore,

$$\left\| \sum_{l=1}^{L} \left[G(\hat{X}_l^k, \boldsymbol{Z_l}, A_l, T_l; \boldsymbol{\beta}, \boldsymbol{\psi}) - G(\bar{X}_l, \boldsymbol{Z_l}, A_l, T_l; \boldsymbol{\beta}, \boldsymbol{\psi}) \right] \sum_{i\in\overline{V}} \delta_i \hat{w}_i \mathbb{1}(W_l \text{ is near } W_i) \right\| \xrightarrow{P} 0 \quad as\; n \to \infty.$$

Equivalently,

$$\left\| \sum_{i\in\overline{V}} \delta_i \hat{w}_i \left[G(\hat{X}_i^k, \boldsymbol{Z_i}, A_i, T_i; \boldsymbol{\beta}, \boldsymbol{\psi}) - G(\bar{X}_i, \boldsymbol{Z_i}, A_i, T_i; \boldsymbol{\beta}, \boldsymbol{\psi}) \right] \right\| \xrightarrow{P} 0 \quad as\; n \to \infty.$$

Thus,

$$\hat{U}(\boldsymbol{\beta}, \boldsymbol{\psi}) = U(\boldsymbol{\beta}, \boldsymbol{\psi}) + o_p(1).$$

Since $(\hat{\boldsymbol{\beta}}^{\boldsymbol{k}}, \hat{\boldsymbol{\psi}}^{\boldsymbol{k}})$ is the solution to $\hat{U}(\boldsymbol{\beta}, \boldsymbol{\psi}) = 0$ and $(\hat{\boldsymbol{\beta}}, \hat{\boldsymbol{\psi}})$ is the solution to $U(\boldsymbol{\beta}, \boldsymbol{\psi}) = 0$, the blip estimator $\hat{\boldsymbol{\psi}}^{\boldsymbol{k}}$ converges to $\hat{\boldsymbol{\psi}}$ in probability. It is easy to prove that $\hat{\boldsymbol{\psi}}$ consistently estimates $\boldsymbol{\psi}$. Thus, $\hat{\boldsymbol{\psi}}^{\boldsymbol{k}}$ is a consistent estimator of $\boldsymbol{\psi}$.

Chapter 4

Q-learning with Misclassified Binary Outcomes

4.1 Introduction

In the previous chapters, we investigate the scenarios with continuous outcomes and survival outcomes in DTR, in which the covariates are subject to measurement error. In this chapter, we study the DTR with discrete-valued outcomes. In particular, a binary response in Q-learning (Moodie et al., 2014) is considered. We target our study on the misclassification effect on Q-learning in binary regression with internal validation data provided. The maximum likelihood estimation is proposed as an effective correction method to accommodate the misclassification effect in Q-learning.

The remainder of the chapter is organized as follows. In Section 4.2, the Q-learning model in binary regression is presented, and the misclassification process for the binary outcome is introduced. The correction method to account for the misclassification effect in Q-learning is described in Section 4.2. In Section 4.3, we conduct simulation studies to evaluate the performance of the proposed method in one-stage and multi-stage settings. Data analysis is conducted in Section 4.4 for illustration using two real data examples, the NHEFS data and smoking cessation data, followed by conclusions in the last section.

4.2 Methodology

4.2.1 Notations and Model Framework

The model framework is focused on a two-stage setting in Q-learning with a data trajectory $\{\boldsymbol{X_1}, A_1, \boldsymbol{X_2}, A_2, Y\}$, where $\boldsymbol{X}_j$ is a covariate vector precisely measured and collected prior to a treatment $A_j \in \{1, -1\}$ assigned at stage j (j = 1, 2). $Y \in \{1, 0\}$ is a binary outcome measured at the end of second stage.

In the applications, the outcome Y may be subject to misclassification, and let Y^* be a surrogate outcome, the actually observed version of Y. We focus on a situation where the study has both internal validation data V and main study data $\overline{V}$ available. That is,

$$\{\boldsymbol{X_{i1}}, A_{i1}, \boldsymbol{X_{i2}}, A_{i2}, Y_i, Y_i^*\} \quad \text{if i} \in V,$$
$$\{\boldsymbol{X_{i1}}, A_{i1}, \boldsymbol{X_{i2}}, A_{i2}, Y_i^*\} \quad \text{if i} \in \overline{V},$$

where the surrogate outcome Y^* is observed for all individuals, but the true outcome Y is only observed for individuals in the validation data V.

We consider the case of non-differential misclassification mechanism, where the probability of Y^* depends only on Y. Then, the misclassification probabilities $(\gamma_{10}, \gamma_{01})$ are defined as

$$\gamma_{10} = P(Y^* = 1|Y = 0), \ \gamma_{01} = P(Y^* = 0|Y = 1). \tag{4.1}$$

In order for the misclassification probabilities $(\gamma_{10}, \gamma_{01})$ and regression parameters $(\boldsymbol{\beta_j}, \boldsymbol{\psi_j})$ to be identifiable, one additional assumption is imposed to the Q-learning

(A5) *Monotonicity condition*: $\gamma_{10} + \gamma_{01} < 1$.

This assumption (A5) guarantees that $(\gamma_{10}, \gamma_{01}, \boldsymbol{\beta_j}, \boldsymbol{\psi_j})$ are identifiable if $\mathrm{E}[\boldsymbol{X_j}\boldsymbol{X_j^T}]$ exists and is non-singular for j = 1, 2 (Hausman et al., 1998). Otherwise, if $\gamma_{10} + \gamma_{01} \geq 1$, this set of misclassification probabilities are deemed problematic and Y^* is regarded not to be produced by chance (Neuhaus, 1999).

In Q-learning, when the outcome misclassification is ignored, and we proceed to fit a model by simply replacing Y with Y^* in (1.4), we obtain a naive model

$$Q_2(\boldsymbol{H_2}, A_2; \boldsymbol{\beta_2^n}, \boldsymbol{\psi_2^n}) = E[Y^*|\boldsymbol{H_2}, A_2] = P(Y^* = 1|\boldsymbol{H_2}, A_2) = \mathrm{expit}\Big(\boldsymbol{\beta_2^{nT}}\boldsymbol{H_{20}} + A_2\Big(\boldsymbol{\psi_2^{nT}}\boldsymbol{H_{21}}\Big)\Big). \tag{4.2}$$

It has been discussed in the literature that ignoring the misclassification in the response can result in attenuated covariate effects and a change in the model structure (Neuhaus, 1999; Carroll et al., 2006; Yi, 2017). Thus, using the naive model (4.2) in logistic regression yields a naive estimator $(\hat{\boldsymbol{\beta}}_2^n, \hat{\boldsymbol{\psi}}_2^n)$, which may be biased for $(\boldsymbol{\beta_2}, \boldsymbol{\psi_2})$. Furthermore, a biased naive estimator $(\hat{\boldsymbol{\beta}}_2^n, \hat{\boldsymbol{\psi}}_2^n)$ may affect the first stage parameter estimation and the determination of optimal DTR. Such potential issues motivate us to search for an effective approach to accommodate the outcome misclassification effect in Q-learning.

4.2.2 Maximum Likelihood Estimation Method

When the outcome is subject to misclassification, the original Q-learning in (1.1.3) needs modifications to produce consistent estimates of parameters. Carroll et al. (2006) and Yi (2017) described the general use of the maximum likelihood estimation method in the logistic regression model to deal with the outcome misclassification. We borrow their ideas and extend the MLE approach to Q-learning in the internal validation/main study data context.

The main idea of the MLE method in Q-learning is to derive likelihood functions for the validation data and main study data and then combine them for a total likelihood for parameter estimation. Given $(\boldsymbol{H_2}, A_2)$, we establish a relationship of the conditional probability of the surrogate outcome with the conditional probability of the true outcome as

$$
\begin{aligned}
P(Y^* = 1|\boldsymbol{H_2}, A_2) &= P(Y^* = 1, Y = 1|\boldsymbol{H_2}, A_2) + P(Y^* = 1, Y = 0|\boldsymbol{H_2}, A_2) \\
&= P(Y^* = 1|Y = 1, \boldsymbol{H_2}, A_2)P(Y = 1|\boldsymbol{H_2}, A_2) + P(Y^* = 1|Y = 0, \boldsymbol{H_2}, A_2)P(Y = 0|\boldsymbol{H_2}, A_2) \\
&= P(Y^* = 1|Y = 1)P(Y = 1|\boldsymbol{H_2}, A_2) + P(Y^* = 1|Y = 0)P(Y = 0|\boldsymbol{H_2}, A_2) \\
&= \Big[1 - P(Y^* = 0|Y = 1)\Big]P(Y = 1|\boldsymbol{H_2}, A_2) + P(Y^* = 1|Y = 0)\Big[1 - P(Y = 1|\boldsymbol{H_2}, A_2)\Big] \\
&= (1 - \gamma_{01})P(Y = 1|\boldsymbol{H_2}, A_2) + \gamma_{10}\Big[1 - P(Y = 1|\boldsymbol{H_2}, A_2)\Big] \\
&= \gamma_{10} + (1 - \gamma_{10} - \gamma_{01})P(Y = 1|\boldsymbol{H_2}, A_2).
\end{aligned}
\tag{4.3}
$$

Based on (4.3), we first derive the likelihood function for patients in the main study data $\overline{V}$, where only Y^* is observed. Thus, for any i^{th} patient ($i \in \overline{V}$), the corresponding likelihood L_i is

given by

$$
\begin{aligned}
L_i &= P(Y_i^* = y_i^* | \boldsymbol{H_{i2}}, A_{i2}) = P(Y_i^* = y_i^*, Y_i = 1 | \boldsymbol{H_{i2}}, A_{i2}) + P(Y_i^* = y_i^*, Y_i = 0 | \boldsymbol{H_{i2}}, A_{i2}) \\
&= P(Y_i^* = y_i^* | Y_i = 1, \boldsymbol{H_{i2}}, A_{i2}) P(Y_i = 1 | \boldsymbol{H_{i2}}, A_{i2}) + P(Y_i^* = y_i^* | Y_i = 0, \boldsymbol{H_{i2}}, A_{i2}) P(Y_i = 0 | \boldsymbol{H_{i2}}, A_{i2}) \\
&= P(Y_i^* = y_i^* | Y_i = 1) P(Y_i = 1 | \boldsymbol{H_{i2}}, A_{i2}) + P(Y_i^* = y_i^* | Y_i = 0) P(Y_i = 0 | \boldsymbol{H_{i2}}, A_{i2}).
\end{aligned}
$$

Then, the likelihood $L_{\overline{v}}$ is the product of the likelihoods across $n_{\overline{v}}$ patients from $\overline{V}$

$$
\begin{aligned}
L_{\overline{v}} &= \prod_{i=1}^{n_{\overline{v}}} L_i = \prod_{i=1}^{n_{\overline{v}}} \Big\{ P(Y_i^* = y_i^* | Y_i = 1) P(Y_i = 1 | \boldsymbol{H_{i2}}, A_{i2}) + P(Y_i^* = y_i^* | Y_i = 0) P(Y_i = 0 | \boldsymbol{H_{i2}}, A_{i2}) \Big\} \\
&= \prod_{i=1}^{n_{\overline{v}}} \Big\{ P(Y_i^* = 1 | Y_i = 1) P(Y_i = 1 | \boldsymbol{H_{i2}}, A_{i2}) + P(Y_i^* = 1 | Y_i = 0) P(Y_i = 0 | \boldsymbol{H_{i2}}, A_{i2}) \Big\} \times \\
&\quad \Big\{ P(Y_i^* = 0 | Y_i = 1) P(Y_i = 1 | \boldsymbol{H_{i2}}, A_{i2}) + P(Y_i^* = 0 | Y_i = 0) P(Y_i = 0 | \boldsymbol{H_{i2}}, A_{i2}) \Big\} \\
&= \prod_{i=1}^{n_{\overline{v}}} \Big\{ (1 - \gamma_{01}) P(Y_i = 1 | \boldsymbol{H_{i2}}, A_{i2}) + \gamma_{10} P(Y_i = 0 | \boldsymbol{H_{i2}}, A_{i2}) \Big\}^{y_i^*} \times \\
&\quad \Big\{ \gamma_{01} P(Y_i = 1 | \boldsymbol{H_{i2}}, A_{i2}) + (1 - \gamma_{10}) P(Y_i = 0 | \boldsymbol{H_{i2}}, A_{i2}) \Big\}^{1 - y_i^*} \\
&= \prod_{i=1}^{n_{\overline{v}}} \Big\{ (1 - \gamma_{01}) P(Y_i = 1 | \boldsymbol{H_{i2}}, A_{i2}) + \gamma_{10} \Big[1 - P(Y_i = 1 | \boldsymbol{H_{i2}}, A_{i2}) \Big] \Big\}^{y_i^*} \times \\
&\quad \Big\{ \gamma_{01} P(Y_i = 1 | \boldsymbol{H_{i2}}, A_{i2}) + (1 - \gamma_{10}) \Big[1 - P(Y_i = 1 | \boldsymbol{H_{i2}}, A_{i2}) \Big] \Big\}^{1 - y_i^*} \\
&= \prod_{i=1}^{n_{\overline{v}}} \Big\{ \gamma_{10} + (1 - \gamma_{10} - \gamma_{01}) P(Y_i = 1 | \boldsymbol{H_{i2}}, A_{i2}) \Big\}^{y_i^*} \times \Big\{ (1 - \gamma_{10}) - (1 - \gamma_{10} - \gamma_{01}) P(Y_i = 1 | \boldsymbol{H_{i2}}, A_{i2}) \Big\}^{1 - y_i^*}.
\end{aligned}
$$

For any i^{th} patient in the validation data ($i \in V$), the likelihood involves both Y_i^* and Y_i, which is expressed as

$$
\begin{aligned}
L_i &= P(Y_i^* = y_i^*, Y_i = y_i | \boldsymbol{H_{i2}}, A_{i2}) \\
&= P(Y_i^* = y_i^* | Y_i = y_i, \boldsymbol{H_{i2}}, A_{i2}) P(Y_i = y_i | \boldsymbol{H_{i2}}, A_{i2}) \\
&= P(Y_i^* = y_i^* | Y_i = y_i) P(Y_i = y_i | \boldsymbol{H_{i2}}, A_{i2}).
\end{aligned}
$$

Then, the likelihood L_v across n_v patients in V follows

$$
\begin{aligned}
L_v &= \prod_{i=1}^{n_v} L_i = \prod_{i=1}^{n_v} P(Y_i^* = y_i^* | Y_i = y_i) P(Y_i = y_i | \boldsymbol{H_{i2}}, A_{i2}) \\
&= \prod_{i=1}^{n_v} \Bigg\{ \Big[P(Y_i^* = 1 | Y_i = 1) P(Y_i = 1 | \boldsymbol{H_{i2}}, A_{i2}) \Big]^{y_i^*=1, y_i=1} \times \Big[P(Y_i^* = 1 | Y_i = 0) P(Y_i = 0 | \boldsymbol{H_{i2}}, A_{i2}) \Big]^{y_i^*=1, y_i=0} \times \\
&\quad \Big[P(Y_i^* = 0 | Y_i = 1) P(Y_i = 1 | \boldsymbol{H_{i2}}, A_{i2}) \Big]^{y_i^*=0, y_i=1} \times \Big[P(Y_i^* = 0 | Y_i = 0) P(Y_i = 0 | \boldsymbol{H_{i2}}, A_{i2}) \Big]^{y_i^*=0, y_i=0} \Bigg\} \\
&= \prod_{i=1}^{n_v} \Bigg\{ \Big[(1 - \gamma_{01}) P(Y_i = 1 | \boldsymbol{H_{i2}}, A_{i2}) \Big]^{y_i^* y_i} \times \Big[\gamma_{10} P(Y_i = 0 | \boldsymbol{H_{i2}}, A_{i2}) \Big]^{y_i^* (1 - y_i)} \times \\
&\quad \Big[\gamma_{01} P(Y_i = 1 | \boldsymbol{H_{i2}}, A_{i2}) \Big]^{(1 - y_i^*) y_i} \times \Big[(1 - \gamma_{10}) P(Y_i = 0 | \boldsymbol{H_{i2}}, A_{i2}) \Big]^{(1 - y_i^*)(1 - y_i)} \Bigg\} \\
&= \prod_{i=1}^{n_v} \Bigg\{ \Big[(1 - \gamma_{01}) P(Y_i = 1 | \boldsymbol{H_{i2}}, A_{i2}) \Big]^{y_i^* y_i} \times \Big[\gamma_{10} \Big(1 - P(Y_i = 1 | \boldsymbol{H_{i2}}, A_{i2}) \Big) \Big]^{y_i^* (1 - y_i)} \times \\
&\quad \Big[\gamma_{01} P(Y_i = 1 | \boldsymbol{H_{i2}}, A_{i2}) \Big]^{(1 - y_i^*) y_i} \times \Big[(1 - \gamma_{10}) \Big(1 - P(Y_i = 1 | \boldsymbol{H_{i2}}, A_{i2}) \Big) \Big]^{(1 - y_i^*)(1 - y_i)} \Bigg\}.
\end{aligned}
$$

Thus, the total likelihood function L across all the patients is given by

$$
\begin{aligned}
L = L_v \times L_{\overline{v}} &= \Bigg\{ \prod_{i=1}^{n_v} \Bigg\{ \Big[(1 - \gamma_{01}) P(Y_i = 1 | \boldsymbol{H_{i2}}, A_{i2}) \Big]^{y_i^* y_i} \times \Big[\gamma_{10} \Big(1 - P(Y_i = 1 | \boldsymbol{H_{i2}}, A_{i2}) \Big) \Big]^{y_i^* (1 - y_i)} \times \\
&\quad \Big[\gamma_{01} P(Y_i = 1 | \boldsymbol{H_{i2}}, A_{i2}) \Big]^{(1 - y_i^*) y_i} \times \Big[(1 - \gamma_{10}) \Big(1 - P(Y_i = 1 | \boldsymbol{H_{i2}}, A_{i2}) \Big) \Big]^{(1 - y_i^*)(1 - y_i)} \Bigg\} \Bigg\} \times \\
&\quad \Bigg\{ \prod_{i=1}^{n_{\overline{v}}} \Big\{ \gamma_{10} + (1 - \gamma_{10} - \gamma_{01}) P(Y_i = 1 | \boldsymbol{H_{i2}}, A_{i2}) \Big\}^{y_i^*} \times \Big\{ (1 - \gamma_{10}) - (1 - \gamma_{10} - \gamma_{01}) P(Y_i = 1 | \boldsymbol{H_{i2}}, A_{i2}) \Big\}^{1 - y_i^*} \Bigg\}.
\end{aligned}
\tag{4.4}
$$

From (4.4), we can obtain a total log-likelihood function that is to be maximized

$$
\begin{aligned}
logL = \sum_{i=1}^{n_v} \Bigg\{ & y_i^* y_i log \Big[(1 - \gamma_{01}) P(Y_i = 1 | \boldsymbol{H_{i2}}, A_{i2}) \Big] + y_i^* (1 - y_i) log \Big[\gamma_{10} \Big(1 - P(Y_i = 1 | \boldsymbol{H_{i2}}, A_{i2}) \Big) \Big] + \\
& (1 - y_i^*) y_i log \Big[\gamma_{01} P(Y_i = 1 | \boldsymbol{H_{i2}}, A_{i2}) \Big] + (1 - y_i^*)(1 - y_i) log \Big[(1 - \gamma_{10}) \Big(1 - P(Y_i = 1 | \boldsymbol{H_{i2}}, A_{i2}) \Big) \Big] \Bigg\} + \\
& \sum_{i=1}^{n_{\overline{v}}} \Big\{ y_i^* log \Big[\gamma_{10} + (1 - \gamma_{10} - \gamma_{01}) P(Y_i = 1 | \boldsymbol{H_{i2}}, A_{i2}) \Big] + (1 - y_i^*) log \Big[(1 - \gamma_{10}) - (1 - \gamma_{10} - \gamma_{01}) P(Y_i = 1 | \boldsymbol{H_{i2}}, A_{i2}) \Big] \Big\}
\end{aligned}
\tag{4.5}
$$

Let $\boldsymbol{\theta} = (\boldsymbol{\beta_2}, \boldsymbol{\psi_2}, \gamma_{10}, \gamma_{01})$. Maximizing $logL(\boldsymbol{\theta})$, the total log-likelihood function (4.5) with respect to $\boldsymbol{\theta}$, results in a MLE estimator $\hat{\boldsymbol{\theta}}^{\boldsymbol{mle}}$. It is equivalent to solving the score equation

$$\sum_{i=1}^{n} S_i(\boldsymbol{\theta}) = 0,$$

where $\mathrm{S}_i(\boldsymbol{\theta}) = \frac{\partial}{\partial \boldsymbol{\theta}} logL_i(\boldsymbol{\theta})$. Numerical methods such as Newton-Raphson method can be used to find the MLE estimator $\hat{\boldsymbol{\theta}}^{\boldsymbol{mle}}$.

Theorem 4.2.1 *Under the conditions (C1) - (C5) in the Appendix, Section 4.6, the MLE estimator* $\hat{\boldsymbol{\theta}}^{\boldsymbol{mle}}$ *is a consistent estimator of* $\boldsymbol{\theta}$*. That is,*

$$\hat{\boldsymbol{\theta}}^{\boldsymbol{mle}} \xrightarrow{p} \boldsymbol{\theta} \quad as\ n \to \infty.$$

As a result, the stage 2 estimator $(\hat{\boldsymbol{\beta}}_{\boldsymbol{2}}^{\boldsymbol{mle}}, \hat{\boldsymbol{\psi}}_{\boldsymbol{2}}^{\boldsymbol{mle}})$ is consistent. Then, with the consistent estimation of the pseudo-outcome, the stage 1 estimator $(\hat{\boldsymbol{\beta}}_{\boldsymbol{1}}^{\boldsymbol{mle}}, \hat{\boldsymbol{\psi}}_{\boldsymbol{1}}^{\boldsymbol{mle}})$ is consistent by using the ordinary least squares. Thus, the MLE method yields consistent estimates of blip parameter $\boldsymbol{\psi} = (\boldsymbol{\psi_2}, \boldsymbol{\psi_1})$ in Q-learning. The details are provided in the Appendix in Section 4.6.

Modified Q-learning Algorithm in Binary Regression:

The following two-stage Q-learning algorithm details the modified estimation procedures:

1. Parameterize the stage 2 Q-function

$$\mathrm{Q}_2(\boldsymbol{H_2}, A_2; \boldsymbol{\beta}_{\boldsymbol{2}}^{\boldsymbol{mle}}, \boldsymbol{\psi}_{\boldsymbol{2}}^{\boldsymbol{mle}}) = \mathrm{expit}\Big(\boldsymbol{\beta}_{\boldsymbol{2}}^{\boldsymbol{mleT}} \boldsymbol{H_{20}} + A_2(\boldsymbol{\psi}_{\boldsymbol{2}}^{\boldsymbol{mleT}} \boldsymbol{H_{21}})\Big).$$

2. Apply maximum likelihood estimation method to obtain the stage 2 estimator $(\hat{\boldsymbol{\beta}}_{\boldsymbol{2}}^{\boldsymbol{mle}}, \hat{\boldsymbol{\psi}}_{\boldsymbol{2}}^{\boldsymbol{mle}})$ by maximizing the log-likelihood function (4.5).

3. Derive the stage 2 optimal treatment as $\hat{a}_2^{opt} = \arg\max_{a_2} Q_2(\boldsymbol{h_2}, a_2; \hat{\boldsymbol{\beta}}_{\boldsymbol{2}}^{\boldsymbol{mle}}, \hat{\boldsymbol{\psi}}_{\boldsymbol{2}}^{\boldsymbol{mle}})$.

4. Construct the pseudo-outcome for estimation at stage 1

$$\widetilde{Y}_1 = \max_{a_2} \mathrm{logit} Q_2(\boldsymbol{H_2}, a_2; \hat{\boldsymbol{\beta}}_{\boldsymbol{2}}^{\boldsymbol{mle}}, \hat{\boldsymbol{\psi}}_{\boldsymbol{2}}^{\boldsymbol{mle}}).$$

5. Apply OLS regression to obtain the stage 1 estimator $(\hat{\boldsymbol{\beta}}_{\boldsymbol{1}}^{\boldsymbol{mle}}, \hat{\boldsymbol{\psi}}_{\boldsymbol{1}}^{\boldsymbol{mle}})$

$$(\hat{\boldsymbol{\beta}}_{\boldsymbol{1}}^{\boldsymbol{mle}}, \hat{\boldsymbol{\psi}}_{\boldsymbol{1}}^{\boldsymbol{mle}}) = \underset{(\boldsymbol{\beta}_{\boldsymbol{1}}^{\boldsymbol{mle}}, \boldsymbol{\psi}_{\boldsymbol{1}}^{\boldsymbol{mle}})}{\arg\min} \tfrac{1}{n} \textstyle\sum_{i=1}^{n} \Big(\widetilde{Y}_{i1} - Q_1(\boldsymbol{H_{i1}}, A_{i1}; \boldsymbol{\beta}_{\boldsymbol{1}}^{\boldsymbol{mle}}, \boldsymbol{\psi}_{\boldsymbol{1}}^{\boldsymbol{mle}})\Big)^2.$$

6. Derive the stage 1 optimal treatment as $\hat{a}_1^{opt} = \arg\max_{a_1} Q_1(\boldsymbol{h_1}, a_1; \hat{\boldsymbol{\beta}}_{\mathbf{1}}^{\boldsymbol{mle}}, \hat{\boldsymbol{\psi}}_{\mathbf{1}}^{\boldsymbol{mle}})$.

This modified Q-learning algorithm distinguishes itself from the original Q-learning algorithm in Section (1.1.3) in Step 2, which replaces the application of logistic regression with the maximum likelihood estimation method.

4.3 Simulation Studies

4.3.1 One-Stage Estimation

We begin with the one-stage estimation in Q-learning. Let X be a continuous covariate and Z be a binary covariate, where X $\sim$ N(1, 1) and Z $\in \{1, -1\}$ is generated with probability of 0.5. The treatment A $\in \{1, -1\}$ that depends on X is drawn from a Bernoulli distribution with probability P(A = 1) = expit(1 - X), where expit(x) = 1/(1 + exp(-x)). The true outcome Y is drawn from a Bernoulli distribution with probability expit($1 + \beta_z$Z + β_xX + A($\psi_{10} + \psi_{11}$X)), where ($\boldsymbol{\beta}$, $\boldsymbol{\psi}$) = ($\beta_z, \beta_x, \psi_{10}, \psi_{11}$) = (0.5, -1, 0.5, -0.5). Misclassified outcome Y* is simulated from a Bernoulli distribution based on the misclassification probabilities (γ_{10}, γ_{01}).

The generated dataset is randomly divided into validation data and main study data with a validation ratio ρ, where the validation data contain 100×ρ% of the observations. We consider three estimators to evaluate the performance of the proposed MLE method: (1) validation estimator ($\hat{\boldsymbol{\beta}}^{\boldsymbol{v}}$, $\hat{\boldsymbol{\psi}}^{\boldsymbol{v}}$) obtained using the validation data only, (2) naive estimator ($\hat{\boldsymbol{\beta}}^{\boldsymbol{n}}$, $\hat{\boldsymbol{\psi}}^{\boldsymbol{n}}$) obtained using the surrogate outcome Y*, (3) MLE estimator ($\hat{\boldsymbol{\beta}}^{\boldsymbol{mle}}$, $\hat{\boldsymbol{\psi}}^{\boldsymbol{mle}}$) obtained from the modified algorithm (4.2.2).

We compare results under two sample sizes of n = 500 and n = 2000. The validation ratio ρ is specified as 0.3 and 0.5. The set of (γ_{10}, γ_{01}) is considered to be (0.1, 0.1), (0.2, 0.2) and (0.3, 0.3), which can be estimated from the validation data. Simulations are repeated 500 times for each pair of ρ and (γ_{10}, γ_{01}). The bias, empirical standard error (SE), and root mean square error (RMSE) of $\hat{\boldsymbol{\psi}}$ are calculated and assessed. The percentile bootstrap confidence intervals are used with 200 bootstrap samples to derive the coverage probability (CP%) of 95% confidence intervals. Numerical results for n = 500 and n = 2000 under various ρ and (γ_{10}, γ_{01}) are provided in Table 4.1 and Table 4.2, respectively. The parameter estimates ($\hat{\boldsymbol{\beta}}$, $\hat{\boldsymbol{\psi}}$) under ρ

= 0.5 are visualized in Figure 4.1 to Figure 4.6.

Table 4.1 and Table 4.2 show that ignoring the outcome misclassification, the naive estimator $\hat{\boldsymbol{\psi}}^{\boldsymbol{n}}$ produces severely biased results. The results are more biased as the misclassification rate increases. On the contrary, the MLE estimator yields small biases and coverage probabilities close to the nominal level of 95%. Moreover, the proposed MLE method is numerically stable and robust against various ρ and $(\gamma_{10}, \gamma_{01})$. The sample size also plays an important role in the performance of methods. As ρ or n increases, the biases and variability of the estimators are reduced.

4.3.2 Two-Stage Estimation

In this section, we extend the study to evaluate the performance of the proposed methods with two decision points. For simplicity, we follow the same simulation design in Moodie et al. (2014), where the confounding variables are present.

A dataset with 2000 patients forms the data trajectory $(X_1, Z_1, A_1, X_2, Z_2, A_2, Y)$. For j = 1, 2, X_j is a continuous confounding covariate at stage j, where $X_1 \sim N(0, 1)$ and $X_2 \sim N(\eta_0 + \eta_1 X_1, 1)$ for $\eta_0 = -0.5, \eta_1 = 0.5$. The treatment $A_j \in \{1, -1\}$ is assigned depending on X_j with probability $P(A_j = 1) = \text{expit}(\zeta_0 + \zeta_1 X_j)$ for $\zeta_0 = -0.8$ and $\zeta_1 = 1.25$. Two binary covariates $Z_j \in \{1, -1\}$ are generated as $P(Z_1 = 1) = 0.5$ and $P(Z_2 = 1|Z_1, A_1) = \text{expit}(\delta_1 Z_1 + \delta_2 A_1)$. Given the data trajectory, the history at each stage is $\boldsymbol{H_1} = (X_1, Z_1)$ and $\boldsymbol{H_2} = (X_1, Z_1, A_1, X_2, Z_2)$. The outcome model is given by

$$E[Y|\boldsymbol{H_2}, A_2; \boldsymbol{\gamma}] = expit(\gamma_0 + \gamma_1 X_1 + \gamma_2 Z_1 + \gamma_3 A_1 + \gamma_4 Z_1 A_1 + \gamma_5 X_2 + \gamma_6 A_2 + \gamma_7 Z_2 A_2 + \gamma_8 A_1 A_2).$$

In this example, we consider a complete regular scenario that sets $\boldsymbol{\gamma}$ = (0, 1, 0, -0.5, 0, 1, 0.25, 0.5, 0.5) and (δ_1, δ_2) = (0.1, 0.1). While the stage 2 blip parameter can be easily identified as $\boldsymbol{\psi_2}$ = (ψ_6, ψ_7, ψ_8), the true values for the stage 1 blip parameter $\boldsymbol{\psi_1}$ = (ψ_{10}, ψ_{11}) need to be calculated based on the data-generating parameters. Moodie et al. (2014) derived a formula to quantify the true values for $\boldsymbol{\psi_1}$ in this setting, which are given by ψ_{10} = -0.3688, ψ_{11} = 0.0187. Once true outcome model is specified, the observed surrogate Y^* is generated from a Bernoulli distribution based on the misclassification model (4.1), where the misclassification probabilities $(\gamma_{10}, \gamma_{01})$ are set to be (0.1, 0.1), (0.2, 0.2) and (0.3, 0.3). Once the dataset is

generated, the validation data are randomly separated with a ratio $\rho \in \{0.3, 0.5\}$. We continue with the three estimators described in Section (4.3.1). A total of 500 simulations are run for each pair of ρ and $(\gamma_{10}, \gamma_{01})$. Numerical results for the bias, SE, RMSE and CP% of $\hat{\boldsymbol{\psi}} = (\hat{\boldsymbol{\psi}}_{\mathbf{2}}, \hat{\boldsymbol{\psi}}_{\mathbf{1}})$ using the percentile bootstrap with 200 bootstrap samples are reported in Table 4.3. The parameter estimates $\hat{\boldsymbol{\psi}}$ under $\rho = 0.5$ are presented in Figure 4.7 to Figure 4.9.

Similar to the one-stage setting, Table 4.3 shows that the naive estimator $\hat{\boldsymbol{\psi}}^{\boldsymbol{n}}$ leads to broadly biased results. However, the proposed estimator $\hat{\boldsymbol{\psi}}^{\boldsymbol{mle}}$ outperforms the naive estimator with small biases in all the scenarios, and the coverage probabilities of $\hat{\boldsymbol{\psi}}^{\boldsymbol{mle}}$ are close to 95%. It may result from the full log-likelihood function the MLE method relies on for estimation, which accurately describes the relationship with the true model. Moreover, the results show that for the set of first stage estimators, $\hat{\psi}_{10}$ is generally more vulnerable to bias compared with $\hat{\psi}_{11}$, which agrees with the findings in Chakraborty et al. (2010), Moodie et al. (2014) and Song et al. (2015).

4.3.3 Prediction of Optimal DTR

In this section, we explore the misclassification effect in a linear Q-learning response model from a prediction perspective. We are particularly interested in assessing the prediction accuracy of optimal DTR in a two-stage setting.

The simulation design follows (4.3.2), but the simulated data consist of training data with 2000 patients and test data with 5000 patients. The training data are randomly split into validation data and main study data with $\rho \in \{0.3, 0.5\}$, by which the misclassification probabilities and the regression parameters are estimated. To evaluate the performance of the proposed correction method in a predictive setting, we continue the previous three estimators $(\hat{\boldsymbol{\beta}}^{\boldsymbol{v}}, \hat{\boldsymbol{\psi}}^{\boldsymbol{v}})$, $(\hat{\boldsymbol{\beta}}^{\boldsymbol{n}}, \hat{\boldsymbol{\psi}}^{\boldsymbol{n}})$, $(\hat{\boldsymbol{\beta}}^{\boldsymbol{mle}}, \hat{\boldsymbol{\psi}}^{\boldsymbol{mle}})$. The test data are used to compute the prediction accuracy of optimal DTR, which is measured by the proportion of patients whose optimal treatments are correctly predicted at stage 2 and/or stage 1. Simulations are run 500 times. Table 4.4, along with Figure 4.10, summarizes the simulation results under a variety of ρ and $(\gamma_{10}, \gamma_{01})$.

Table 4.4 shows that the prediction accuracy of optimal DTR is adversely affected by misclassification. The naive estimator leads to a pronounced degeneration in the accuracy of pre-

dicted optimal DTR, and its performance is worsened as the misclassification rate increases. In comparison, the proposed MLE estimator considerably improves the prediction accuracy, especially when the optimal treatments in both stages are evaluated. The performance of the MLE estimator is also robust against the various magnitudes of ρ and $(\gamma_{10}, \gamma_{01})$. It substantially restores the precision to a level that is even superior to the validation estimator, suggesting a favorable choice to derive the sequential optimal treatment rules.

4.3.4 Prediction of the Outcome

In this study, we examine the performance of the proposed correction method in terms of the predicted error rates, sensitivity, and specificity of the outcome under the optimal DTR.

The predictive simulation setting follows (4.3.3) with three estimators $(\hat{\boldsymbol{\beta}}^{v}, \hat{\boldsymbol{\psi}}^{v})$, $(\hat{\boldsymbol{\beta}}^{n}, \hat{\boldsymbol{\psi}}^{n})$, $(\hat{\boldsymbol{\beta}}^{mle}, \hat{\boldsymbol{\psi}}^{mle})$ obtained from the training data. We use the test data to estimate the corresponding optimal DTR using each estimator and then calculate the (1) predicted error rates of the outcome, which is measured by the proportion of patients whose outcomes are incorrectly predicted under the estimated optimal DTR, (2) predicted sensitivity of the outcome, which is measured by the proportion of patients whose positive outcomes (Y = 1) are correctly predicted under the estimated optimal DTR, (3) predicted specificity of the outcome, which is measured by the proportion of patients whose negative outcomes (Y = 0) are correctly predicted under the estimated optimal DTR. For the training data, the validation ratio ρ is specified as 0.3 and 0.5, and the misclassification probabilities $(\gamma_{10}, \gamma_{01})$ are set to be (0.1, 0.1), (0.2, 0.2) and (0.3, 0.3). Simulations are repeated 500 times. Results are summarized in Table 4.5 and displayed in Figure 4.11 and Figure 4.12.

Table 4.5 shows that the naive method leads to the worst results in terms of the predicted error rates, sensitivity, and specificity of the outcome in most scenarios. Moreover, compared with the specificity, sensitivity is generally more sensitive to outcome misclassification. In contrast, the proposed MLE method produces the best performance with the lowest error rates and highest sensitivity and specificity results in all the scenarios. As ρ increases, the predicted error rates of the proposed method are lower, and the predicted sensitivity and specificity of the proposed method are higher.

4.4 Data Analysis

4.4.1 Application to NHEFS Data

In this example, we apply the proposed methods to the NHEFS data, which was collected by the National Center for Health Statistics and the National Institute on Aging in collaboration with other agencies of the Public Health Service. A detailed description of the NHEFS is available at https://wwwn.cdc.gov/nchs/nhanes/nhefs/. The NHEFS study aimed to investigate the relationships between clinical, nutritional, and behavioral factors assessed in the first National Health and Nutrition Examination Survey NHANES I and subsequent morbidity, mortality, and operational factors with hospital utilization. In this work, we are interested in estimating an optimal treatment decision rule using the cohort NHEFS dataset in Hernán and Robins (2020). The dataset consists of 1566 cigarette smokers aged 25-74 years, with a number of baseline variables collected from 1971 to 1975. They were followed up through personal interviews in 1982 and reported quitting smoking status, which is the outcome of interest in the analysis. We consider a binary indicator for regular exercise as the treatment variable, with A = 1 indicating those who had little or no exercise and A = -1 otherwise. The baseline variables to be included are age, gender, race, body mass index, systolic blood pressure (SBP), physical activity status, cholesterol, weight, diabetes, the number of years of smoking, and the number of cigarettes smoked each day (SmokeIntensity). Since the measured SBP is right-skewed in the dataset, we take the logarithmic transformation of SBP to be log(SBP-50) following Carroll et al. (2006). Diabetes and SmokeIntensity are shown to be significantly associated with the treatment variable from the treatment model. We regard these two variables as the tailoring variables to derive the optimal treatment decision rule. All the continuous variables are standardized in the analysis.

As described, the smoking status is reported by the patients and thus subject to misclassification. In the dataset, there is no information available to infer the degree of misclassification probabilities. Therefore, we specify a series of values for the misclassification probabilities and conduct sensitivity analyses to evaluate how the misclassification rate affects the estimated optimal treatment decision rule. In the work of Magder and Hughes (1997), it is discussed that the smokers who have really quit smoking are unlikely to report they are still smoking, while

those who have not are very likely to misreport their smoking cessation status. Magder and Hughes (1997) specified $\gamma_{10} = 10\%$, and Lee et al. (2013) provided an estimate for the misclassification rate to be $\gamma_{10} = 7.5\%$ in another smoking cessation study. Thus, we consider $\gamma_{01} = 0$ and $\gamma_{10} \in (5\%, 7.5\%, 10\%, 12.5\%)$ in our analysis. Table 4.6 summarizes the associated inference results, including the estimates, bootstrap standard error (SE), and 95% confidence intervals (CI) for the blip parameters obtained from the naive method and the proposed method.

In Table 4.6, the estimated optimal treatment decision rule from the naive method is $\hat{a}^{opt} =$ 1 if -0.148 + 0.130Diabetes + 0.075SmokeIntensity > 0, and $\hat{a}^{opt}$ = -1 otherwise. In general, the proposed MLE method produces slightly larger estimates than the naive method, leading to different optimal treatment decision rules. As γ_{10} increases, the blip parameter estimates and estimated SEs obtained from the proposed method become bigger. Moreover, the diabetes variable is shown to have a significant treatment effect using the naive method, but the MLE method displays different statistical significance for diabetes in all the scenarios. Therefore, it reveals that the misclassification effect is not negligible in an error-prone setting, which can alter the inference results, including the statistical significance, when the misclassification is taken into account in the analysis.

4.4.2 Application to Smoking Cessation Data

In the second example, we explore the misclassification effect by analyzing the smoking cessation data, which were collected at St. Joseph's Hospital (Lee et al., 2013). The smoking cessation study is a randomized controlled trial and aims to examine the effectiveness of a perioperative smoking cessation intervention with one decision point involved. We are interested in using the smoking cessation data to estimate an optimal treatment decision rule. In this trial, 168 patients were recruited and randomly assigned with the same probability to one of the two treatment groups, the intervention group (A = 1) or the control group (A = 0). The patients were followed up at the time of the 30-day postoperative phone call and self-reported their smoking cessation status, which is the outcome of interest with Y = 1 indicating smoking cessation.

In the study, the smoking cessation status reported by the smokers was examined with the

exhaled carbon monoxide (CO) levels (ppm), where an exhaled CO of $\leq$ 10 ppm confirmed smoking quitting (Lee et al., 2013). It has been found that out of 146 patients with exhaled CO greater than 10ppm, 11 patients misreported their smoking cessation status. We assume a non-differential misclassification mechanism in this analysis. Then, the misclassification probability can be estimated as $\gamma_{10} = 11/146 = 7.5\%$. For those who have already quit smoking, Magder and Hughes (1997) pointed out that they were highly likely to report that they have stopped smoking. Then, we assume that $\gamma_{01} = 0$. It should be noted that these $(\gamma_{10}, \gamma_{01})$ are just the estimates of misclassification probabilities, while the true misclassification probabilities are unknown. Thus, we specify a series of values for $\gamma_{10} \in (2.5\%, 5\%, 7.5\%, 8.5\%)$ and conduct sensitivity analyses to evaluate how the misclassification rate affects the optimal treatment decision rule estimation. The baseline variables in the analysis include age, gender, body mass index, diabetes status, hypertension, chronic obstructive pulmonary disease, cigarettes smoked per day, and the number of years of smoking. The hypertension variable was found statistically significant with respect to the treatment (Shu and Yi, 2019a). We consider hypertension (HTN) and the number of years of smoking (YrsSmoke) as the tailoring variables to derive the optimal treatment decision rule. All the continuous variables are standardized in the analysis. Table 4.7 summarizes the inference results obtained from the naive method and the proposed method.

The analysis results suggest that the misclassification effect is conspicuous. The naive method leads to an optimal decision rule, which is determined by the values of (1.363 - 0.696HTN - 0.189YrsSmoke). In comparison, the proposed MLE method yields notably larger parameter estimates and estimated standard errors than the naive method. As γ_{10} increases, we observe that the MLE estimator is sensitive to the change in the misclassification rate. One possible reason might be the limited size of the dataset. However, these results still reveal a non-negligible impact of outcome misclassification on the optimal treatment decision rule estimation for smoking cessation.

4.5 Conclusion

In this chapter, we discuss the outcome misclassification effect in Q-learning with binary outcomes in the context of internal validation/main study data design. The MLE method is pro-

posed to adjust for the misclassification effect in Q-learning. The correction method is established based on a relationship between two conditional probabilities of the true outcome and surrogate outcome. The likelihoods for both the validation data and main study data are derived and combined to create a total likelihood for parameter estimation in Q-learning. The proposed MLE method itself is straightforward, and under certain conditions, it yields consistent estimates of blip parameters in Q-learning.

We compare the proposed correction method with the naive method in both simulation studies and real data analysis. Ignoring the outcome misclassification leads to severely biased results in parameter estimation. By making use of the observed surrogate outcome and validation data, the proposed method provides satisfactory performance in simulation studies. We show that employing the proposed method in Q-learning considerably reduces the bias, improves the prediction accuracy of optimal DTR, predicted sensitivity and specificity of the outcome, and reduces the predicted error rates of the outcome. Moreover, the MLE method is numerically stable and robust against various magnitudes of validation ratio and misclassification rates. The proposed method is also applied to real data examples to estimate the optimal treatment decision rule. The data analysis suggests that the misclassification effect is not negligible in terms of parameter estimation and associated statistical significance.

Table 4.1: One-stage estimates of blip parameters (ψ_{10}, ψ_{11}) (n = 500)

			ψ_{10}				ψ_{11}			
ρ	$(\gamma_{10}, \gamma_{01})$	$\hat{\boldsymbol{\psi}}$	Bias	SE	RMSE	CP%	Bias	SE	RMSE	CP%
0.3		$\hat{\boldsymbol{\psi}}^{v}$	-0.005	0.722	0.722	93.6	-0.014	0.607	0.607	94.2
	(0.1, 0.1)	$\hat{\boldsymbol{\psi}}^{n}$	-0.175	0.161	0.238	85.2	0.175	0.124	0.214	77.2
		$\hat{\boldsymbol{\psi}}^{mle}$	0.016	0.224	0.225	95.4	-0.016	0.186	0.187	93.0
	(0.2, 0.2)	$\hat{\boldsymbol{\psi}}^{n}$	-0.285	0.152	0.323	65.8	0.288	0.114	0.310	46.4
		$\hat{\boldsymbol{\psi}}^{mle}$	0.007	0.267	0.267	94.8	-0.014	0.222	0.222	93.2
	(0.3, 0.3)	$\hat{\boldsymbol{\psi}}^{n}$	-0.362	0.148	0.391	48.4	0.366	0.108	0.382	24.0
		$\hat{\boldsymbol{\psi}}^{mle}$	0.009	0.307	0.307	92.6	-0.020	0.249	0.250	94.6
0.5		$\hat{\boldsymbol{\psi}}^{v}$	0.002	0.262	0.262	95.2	-0.014	0.210	0.210	94.4
	(0.1, 0.1)	$\hat{\boldsymbol{\psi}}^{n}$	-0.177	0.161	0.239	89.4	0.171	0.124	0.211	87.6
		$\hat{\boldsymbol{\psi}}^{mle}$	-0.003	0.204	0.204	94.2	-0.008	0.165	0.165	94.2
	(0.2, 0.2)	$\hat{\boldsymbol{\psi}}^{n}$	-0.285	0.152	0.323	75.8	0.288	0.114	0.310	63.8
		$\hat{\boldsymbol{\psi}}^{mle}$	0.001	0.226	0.226	95.0	-0.010	0.184	0.184	94.2
	(0.3, 0.3)	$\hat{\boldsymbol{\psi}}^{n}$	-0.367	0.147	0.395	58.4	0.370	0.107	0.385	37.8
		$\hat{\boldsymbol{\psi}}^{mle}$	0.005	0.245	0.245	94.8	-0.014	0.198	0.198	95.0

Table 4.2: One-stage estimates of blip parameters (ψ_{10}, ψ_{11}) (n = 2000)

			ψ_{10}				ψ_{11}			
ρ	$(\gamma_{10}, \gamma_{01})$	$\hat{\boldsymbol{\psi}}$	Bias	SE	RMSE	CP%	Bias	SE	RMSE	CP%
0.3		$\hat{\boldsymbol{\psi}}^{v}$	0.008	0.158	0.158	94.2	-0.007	0.126	0.126	95.0
	(0.1, 0.1)	$\hat{\boldsymbol{\psi}}^{n}$	-0.170	0.078	0.187	55.6	0.171	0.061	0.182	34.8
		$\hat{\boldsymbol{\psi}}^{mle}$	0.008	0.105	0.105	93.0	-0.008	0.086	0.086	93.2
	(0.2, 0.2)	$\hat{\boldsymbol{\psi}}^{n}$	-0.287	0.074	0.296	11.0	0.288	0.055	0.293	1.0
		$\hat{\boldsymbol{\psi}}^{mle}$	0.004	0.123	0.123	93.8	-0.007	0.100	0.100	93.4
	(0.3, 0.3)	$\hat{\boldsymbol{\psi}}^{n}$	-0.374	0.072	0.381	1.6	0.376	0.052	0.380	0.0
		$\hat{\boldsymbol{\psi}}^{mle}$	0.000	0.139	0.139	92.6	-0.005	0.113	0.113	93.8
0.5		$\hat{\boldsymbol{\psi}}^{v}$	0.001	0.121	0.121	94.6	-0.005	0.096	0.096	94.6
	(0.1, 0.1)	$\hat{\boldsymbol{\psi}}^{n}$	-0.176	0.078	0.193	63.8	0.177	0.060	0.187	47.8
		$\hat{\boldsymbol{\psi}}^{mle}$	0.004	0.097	0.097	93.4	-0.005	0.078	0.078	95.0
	(0.2, 0.2)	$\hat{\boldsymbol{\psi}}^{n}$	-0.287	0.074	0.296	24.2	0.288	0.055	0.293	5.0
		$\hat{\boldsymbol{\psi}}^{mle}$	0.002	0.106	0.106	94.2	-0.005	0.085	0.085	93.6
	(0.3, 0.3)	$\hat{\boldsymbol{\psi}}^{n}$	-0.369	0.072	0.376	4.8	0.374	0.052	0.378	0.4
		$\hat{\boldsymbol{\psi}}^{mle}$	0.002	0.114	0.114	95.8	-0.002	0.091	0.091	96.0

Table 4.3: Two-stage estimates of blip parameters ($\psi_{20}, \psi_{21}, \psi_{22}, \psi_{10}, \psi_{11}$)

			ψ_{20}				ψ_{21}				ψ_{22}				ψ_{10}				ψ_{11}			
ρ	$(\gamma_{10}, \gamma_{01})$	$\hat{\psi}$	Bias	SE	RMSE	CP%	Bias	SE	RMSE	CP%	Bias	SE	RMSE	CP%	Bias	SE	RMSE	CP%	Bias	SE	RMSE	CP%
0.3		$\hat{\psi}^v$	0.005	0.143	0.143	93.2	0.004	0.111	0.111	95.4	0.001	0.139	0.139	93.8	-0.007	0.053	0.053	93.6	-0.004	0.047	0.047	92.8
	(0.1, 0.1)	$\hat{\psi}^n$	-0.076	0.066	0.101	86.2	-0.169	0.052	0.177	25.4	-0.171	0.063	0.182	39.4	0.127	0.019	0.128	76.6	-0.012	0.017	0.021	94.0
		$\hat{\psi}^{mle}$	0.003	0.090	0.090	92.4	0.000	0.073	0.073	94.4	0.007	0.089	0.089	93.2	-0.001	0.029	0.029	92.8	-0.009	0.026	0.028	93.2
	(0.2, 0.2)	$\hat{\psi}^n$	-0.132	0.062	0.146	54.8	-0.280	0.049	0.284	0.0	-0.284	0.058	0.290	2.0	0.211	0.012	0.211	41.4	-0.011	0.011	0.016	94.0
		$\hat{\psi}^{mle}$	0.005	0.107	0.107	94.8	0.004	0.086	0.086	95.4	0.009	0.105	0.105	91.0	0.003	0.029	0.029	93.4	-0.001	0.026	0.026	94.6
	(0.3, 0.3)	$\hat{\psi}^n$	-0.174	0.061	0.184	34.4	-0.364	0.047	0.367	0.0	-0.368	0.055	0.372	0.0	0.271	0.008	0.271	11.6	-0.016	0.007	0.017	94.4
		$\hat{\psi}^{mle}$	0.003	0.124	0.124	93.8	0.004	0.098	0.098	95.4	0.002	0.120	0.120	94.6	-0.006	0.029	0.030	94.2	-0.004	0.026	0.026	94.4
0.5		$\hat{\psi}^v$	0.002	0.107	0.107	94.0	0.004	0.085	0.085	93.8	0.013	0.103	0.104	92.4	-0.003	0.041	0.041	93.8	0.004	0.037	0.037	94.6
	(0.1, 0.1)	$\hat{\psi}^n$	-0.077	0.066	0.101	86.0	-0.170	0.052	0.178	39.4	-0.171	0.063	0.182	52.8	0.127	0.019	0.128	81.2	-0.005	0.017	0.018	95.0
		$\hat{\psi}^{mle}$	0.001	0.085	0.085	93.8	0.003	0.068	0.068	94.6	0.008	0.082	0.082	93.8	-0.003	0.029	0.029	94.2	0.004	0.026	0.026	93.8
	(0.2, 0.2)	$\hat{\psi}^n$	-0.134	0.063	0.148	67.4	-0.282	0.048	0.286	3.6	-0.283	0.058	0.289	8.8	0.210	0.012	0.210	54.6	-0.008	0.011	0.014	93.8
		$\hat{\psi}^{mle}$	0.000	0.094	0.094	94.6	0.003	0.075	0.075	94.8	0.011	0.090	0.091	93.0	-0.003	0.029	0.029	94.4	0.005	0.026	0.026	95.0
	(0.3, 0.3)	$\hat{\psi}^n$	-0.176	0.061	0.186	46.4	-0.366	0.046	0.369	0.0	-0.365	0.055	0.369	0.2	0.279	0.008	0.279	25.6	-0.014	0.007	0.016	93.2
		$\hat{\psi}^{mle}$	0.000	0.100	0.100	93.8	0.003	0.080	0.080	93.4	0.007	0.097	0.097	95.6	-0.002	0.029	0.029	91.6	0.003	0.026	0.026	94.0

Table 4.4: Prediction accuracy of optimal DTR (%)

		Stage 2			Stage 1			Stage 2 & Stage 1		
ρ	$(\gamma_{10}, \gamma_{01})$	v	n	mle	v	n	mle	v	n	mle
0.3	(0.1, 0.1)	91.7	97.2	98.5	92.7	97.7	98.8	88.1	96.1	97.9
	(0.2, 0.2)	92.0	91.7	96.0	92.9	91.4	96.4	88.5	87.5	94.3
	(0.3, 0.3)	92.0	86.3	95.4	93.4	85.1	95.9	88.8	78.8	93.3
0.5	(0.1, 0.1)	95.6	97.3	98.9	97.1	98.1	99.5	94.2	96.3	98.6
	(0.2, 0.2)	96.3	91.2	98.0	97.3	92.7	98.8	94.9	87.5	97.4
	(0.3, 0.3)	96.4	84.8	97.1	97.3	86.5	98.0	95.1	77.9	96.2

v: validation estimator, n: naive estimator, mle: MLE estimator

Table 4.5: Predicted error rates, sensitivity, and specificity of the outcome (%)

		Error Rates			Sensitivity			Specificity		
ρ	$(\gamma_{10}, \gamma_{01})$	v	n	mle	v	n	mle	v	n	mle
0.3	(0.1, 0.1)	5.5	4.0	3.4	92.7	94.5	95.6	95.7	97.1	97.2
	(0.2, 0.2)	5.4	5.9	4.2	92.4	91.7	94.5	95.9	95.7	96.7
	(0.3, 0.3)	5.4	8.5	4.5	92.6	86.6	94.0	95.9	94.6	96.4
0.5	(0.1, 0.1)	4.0	3.8	3.0	95.0	94.9	96.1	96.7	97.0	97.5
	(0.2, 0.2)	3.9	5.9	3.4	95.0	92.2	95.8	96.9	95.3	97.1
	(0.3, 0.3)	4.0	8.9	3.7	94.9	86.1	95.2	96.8	94.3	97.0

v: validation estimator, n: naive estimator, mle: MLE estimator

Table 4.6: Sensitivity analysis results of the NHEFS data for the blip estimators

		A			A*Diabetes			A*SmokeIntensity		
Method	γ_{10}	Est	SE	95%CI	Est	SE	95%CI	Est	SE	95%CI
Naive		-0.148	0.091	(-0.300, 0.027)	0.130	0.066	(0.005, 0.257)	0.075	0.070	(-0.037, 0.236)
MLE	5%	-0.179	0.113	(-0.376, 0.034)	0.149	0.079	(-0.017, 0.304)	0.121	0.097	(-0.023, 0.343)
	7.5%	-0.201	0.129	(-0.433, 0.039)	0.160	0.089	(-0.025, 0.342)	0.148	0.114	(-0.011, 0.399)
	10%	-0.231	0.151	(-0.514, 0.047)	0.173	0.101	(-0.034, 0.374)	0.179	0.136	(-0.005, 0.470)
	12.5%	-0.270	0.186	(-0.618, 0.077)	0.187	0.117	(-0.054, 0.409)	0.213	0.174	(-0.048, 0.565)

Est: estimates, SE: standard error, CI: confidence interval

Table 4.7: Sensitivity analysis results of the smoking cessation data for the blip estimators

		A			A*HTN			A*YrsSmoke		
Method	γ_{10}	Est	SE	95%CI	Est	SE	95%CI	Est	SE	95%CI
Naive		1.363	0.895	(-0.391, 3.118)	-0.696	0.850	(-2.363, 0.970)	-0.189	0.896	(-1.945, 1.567)
MLE	2.5%	1.718	1.355	(-0.938, 4.374)	-0.826	1.235	(-3.247, 1.595)	-0.228	1.558	(-3.281, 2.826)
	5%	2.675	1.716	(-0.688, 6.039)	-1.116	1.540	(-4.134, 1.903)	-0.893	1.792	(-4.406, 2.620)
	7.5%	3.274	1.923	(-0.496, 7.043)	-1.338	1.725	(-4.720, 2.044)	-1.368	1.844	(-4.982, 2.246)
	8.5%	3.583	2.029	(-0.393, 7.560)	-1.418	1.773	(-4.893, 2.056)	-1.671	1.830	(-5.257, 1.915)

Est: estimates, SE: standard error, CI: confidence interval

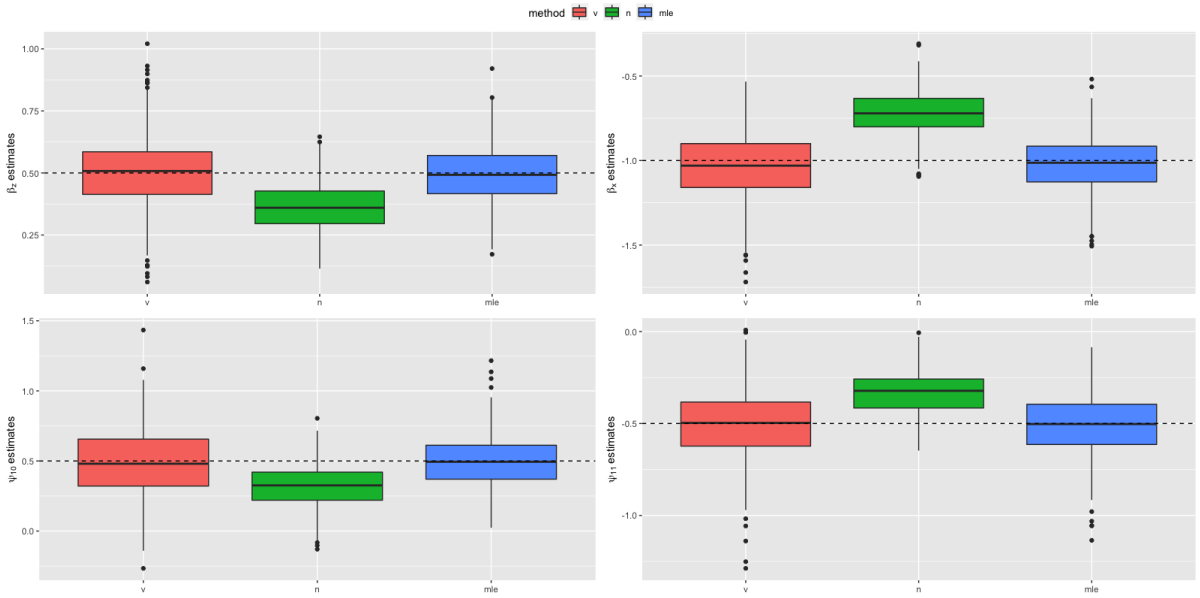

Figure 4.1: One-stage estimates of $(\beta_z, \beta_x, \psi_{10}, \psi_{11})$ for n = 500 with $\rho = 0.5$ and $(\gamma_{10}, \gamma_{01}) = (0.1, 0.1)$

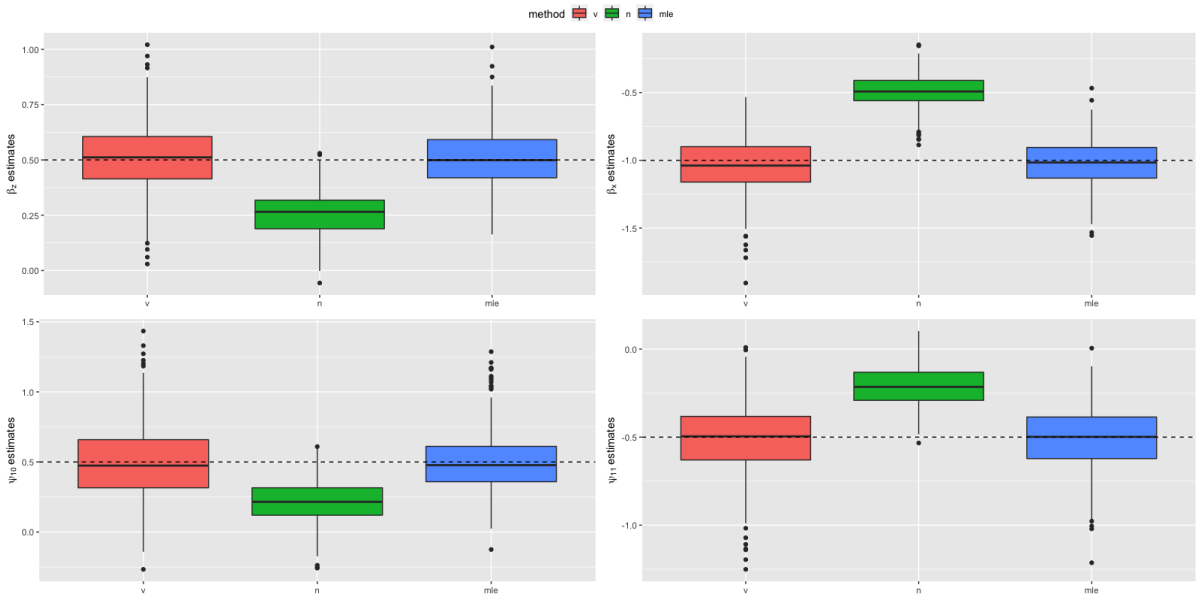

Figure 4.2: One-stage estimates of $(\beta_z, \beta_x, \psi_{10}, \psi_{11})$ for n = 500 with $\rho = 0.5$ and $(\gamma_{10}, \gamma_{01}) = (0.2, 0.2)$

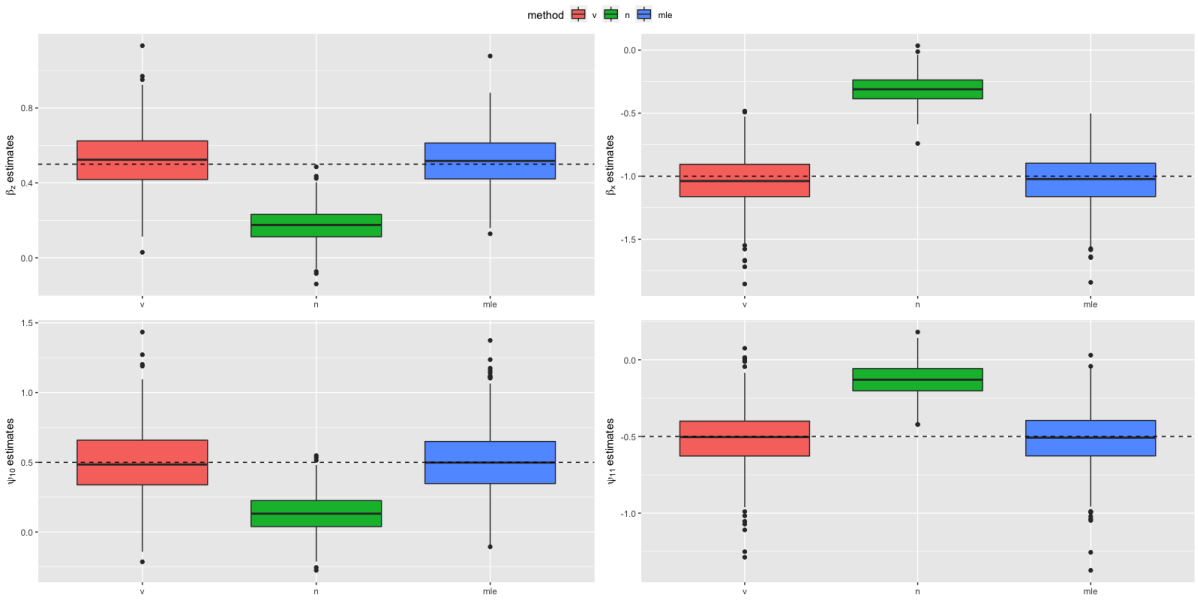

Figure 4.3: One-stage estimates of $(\beta_z, \beta_x, \psi_{10}, \psi_{11})$ for n = 500 with $\rho = 0.5$ and $(\gamma_{10}, \gamma_{01}) = (0.3, 0.3)$

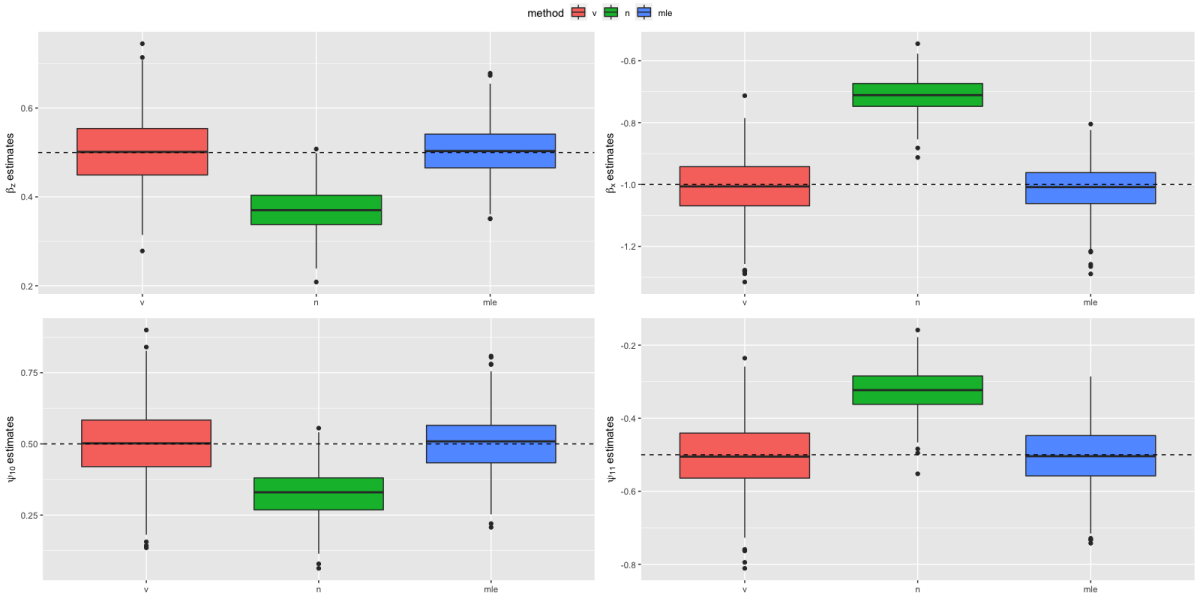

Figure 4.4: One-stage estimates of $(\beta_z, \beta_x, \psi_{10}, \psi_{11})$ for n = 2000 with $\rho = 0.5$ and $(\gamma_{10}, \gamma_{01}) = (0.1, 0.1)$

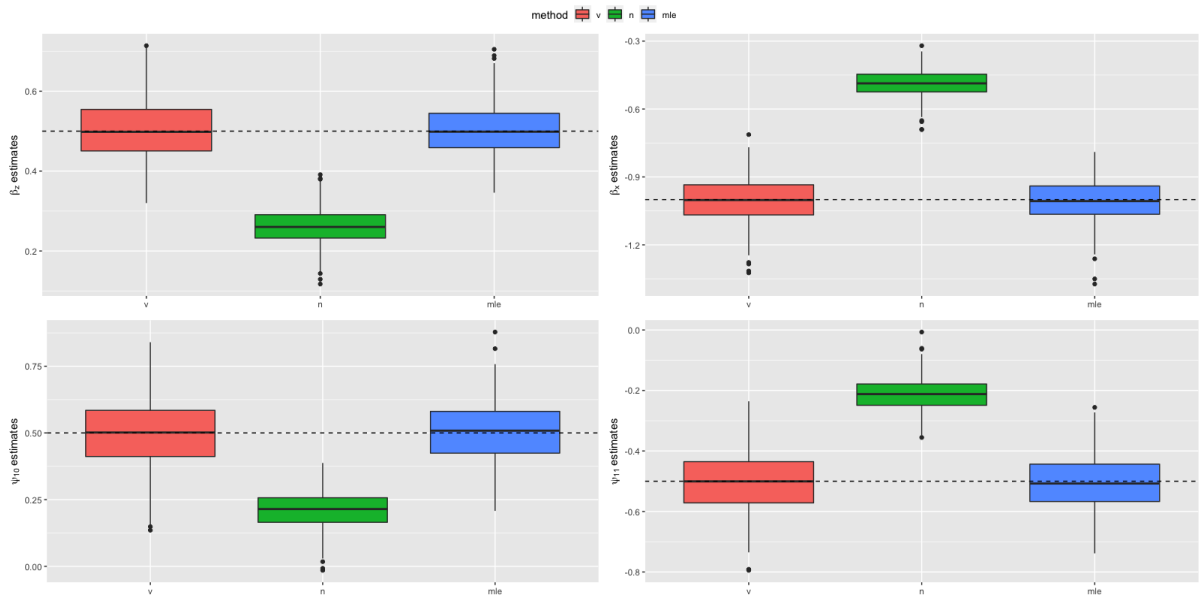

Figure 4.5: One-stage estimates of $(\beta_z, \beta_x, \psi_{10}, \psi_{11})$ for n = 2000 with $\rho = 0.5$ and $(\gamma_{10}, \gamma_{01}) = (0.2, 0.2)$

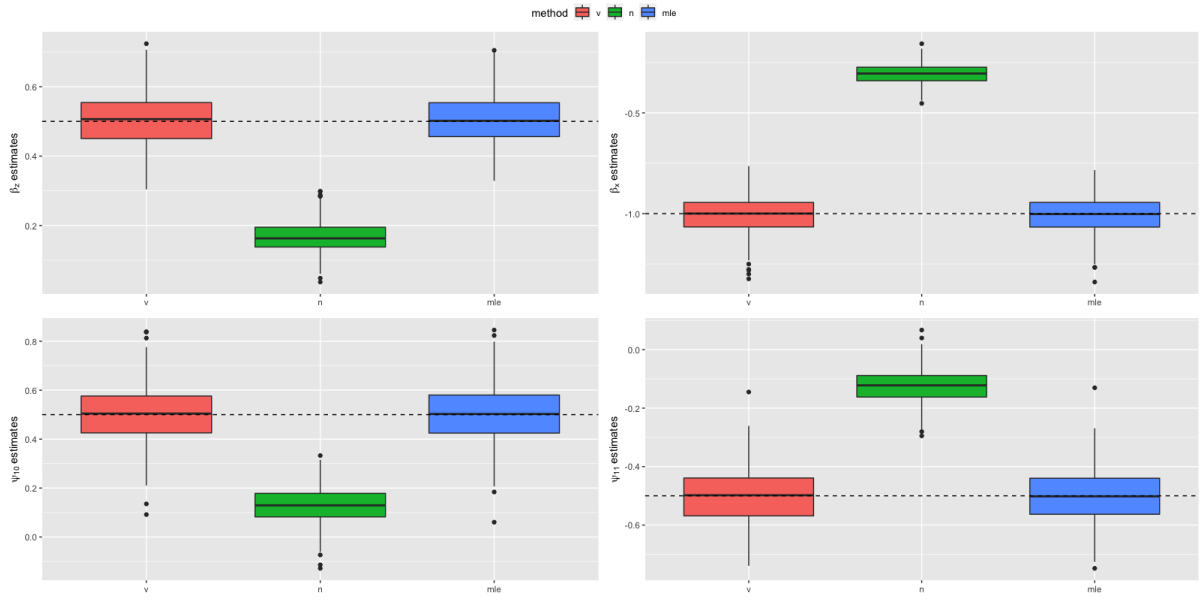

Figure 4.6: One-stage estimates of $(\beta_z, \beta_x, \psi_{10}, \psi_{11})$ for n = 2000 with $\rho = 0.5$ and $(\gamma_{10}, \gamma_{01}) = (0.3, 0.3)$

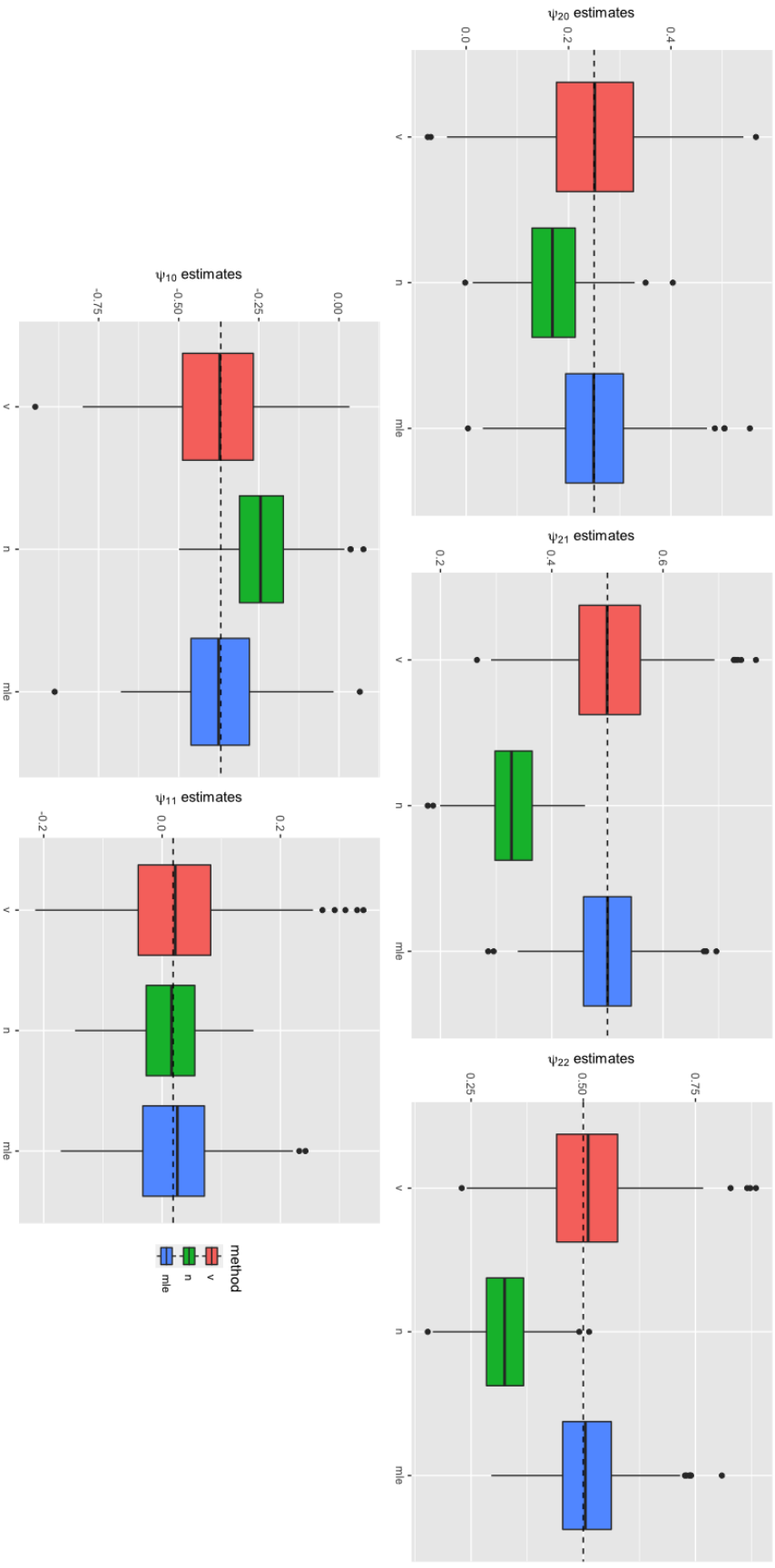

Figure 4.7: Two-stage estimates of $(\psi_{20}, \psi_{21}, \psi_{22}, \psi_{10}, \psi_{11})$ with $\rho = 0.5$ and $(\gamma_{10}, \gamma_{01}) = (0.1, 0.1)$

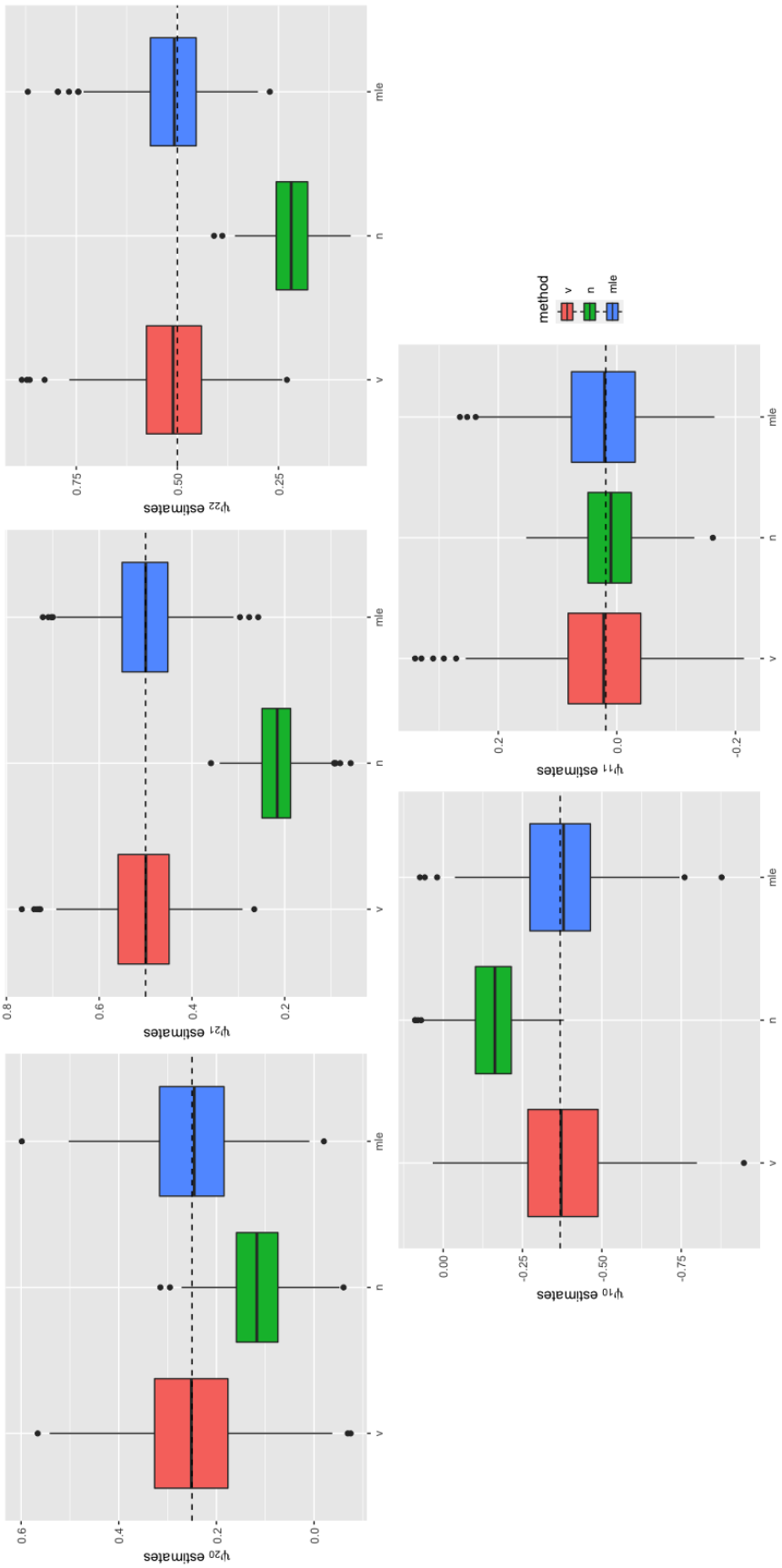

Figure 4.8: Two-stage estimates of $(\psi_{20}, \psi_{21}, \psi_{22}, \psi_{10}, \psi_{11})$ with $\rho = 0.5$ and $(\gamma_{10}, \gamma_{01}) = (0.2, 0.2)$

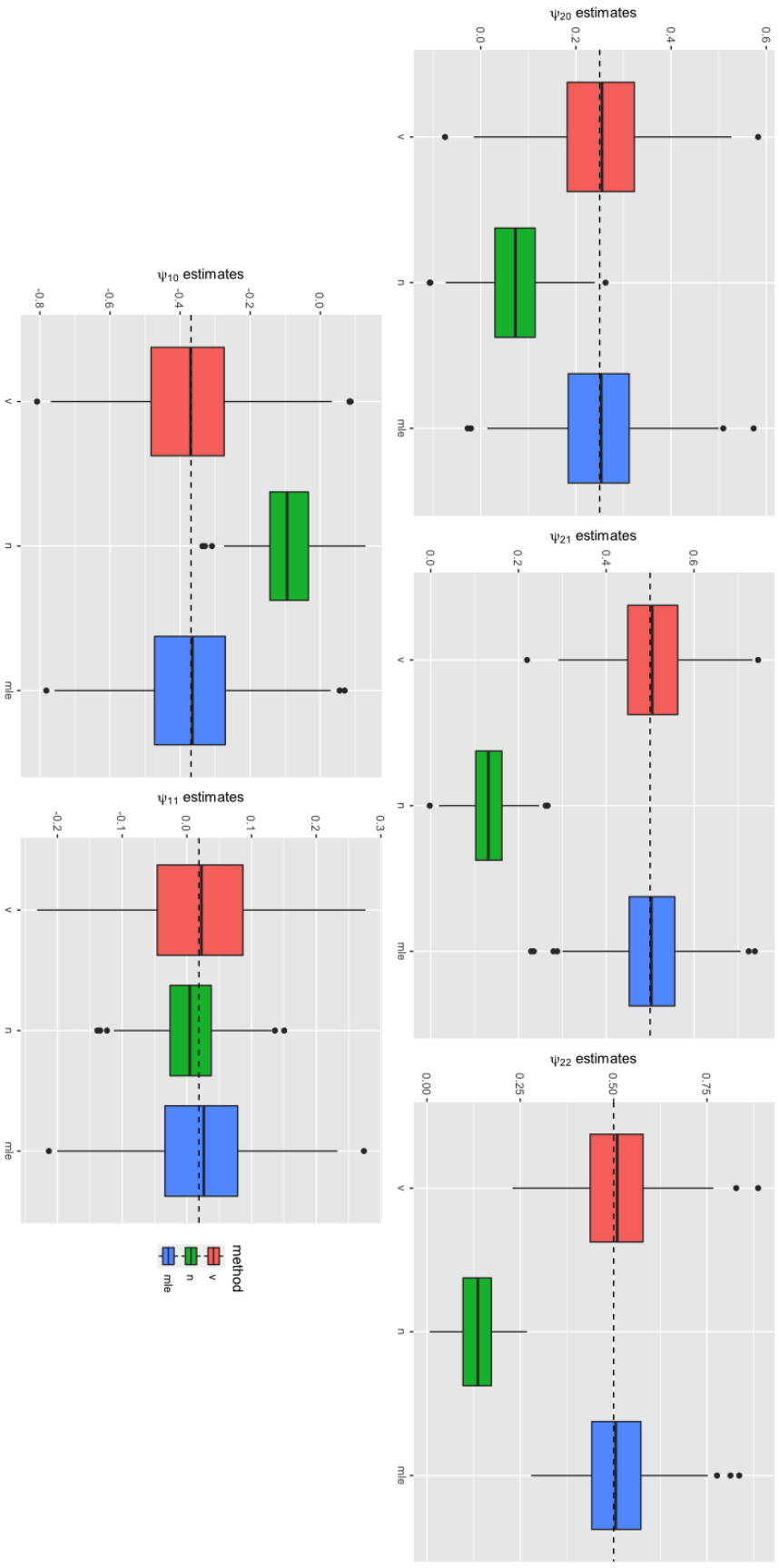

Figure 4.9: Two-stage estimates of $(\psi_{20}, \psi_{21}, \psi_{22}, \psi_{10}, \psi_{11})$ with $\rho = 0.5$ and $(\gamma_{10}, \gamma_{01}) = (0.3, 0.3)$

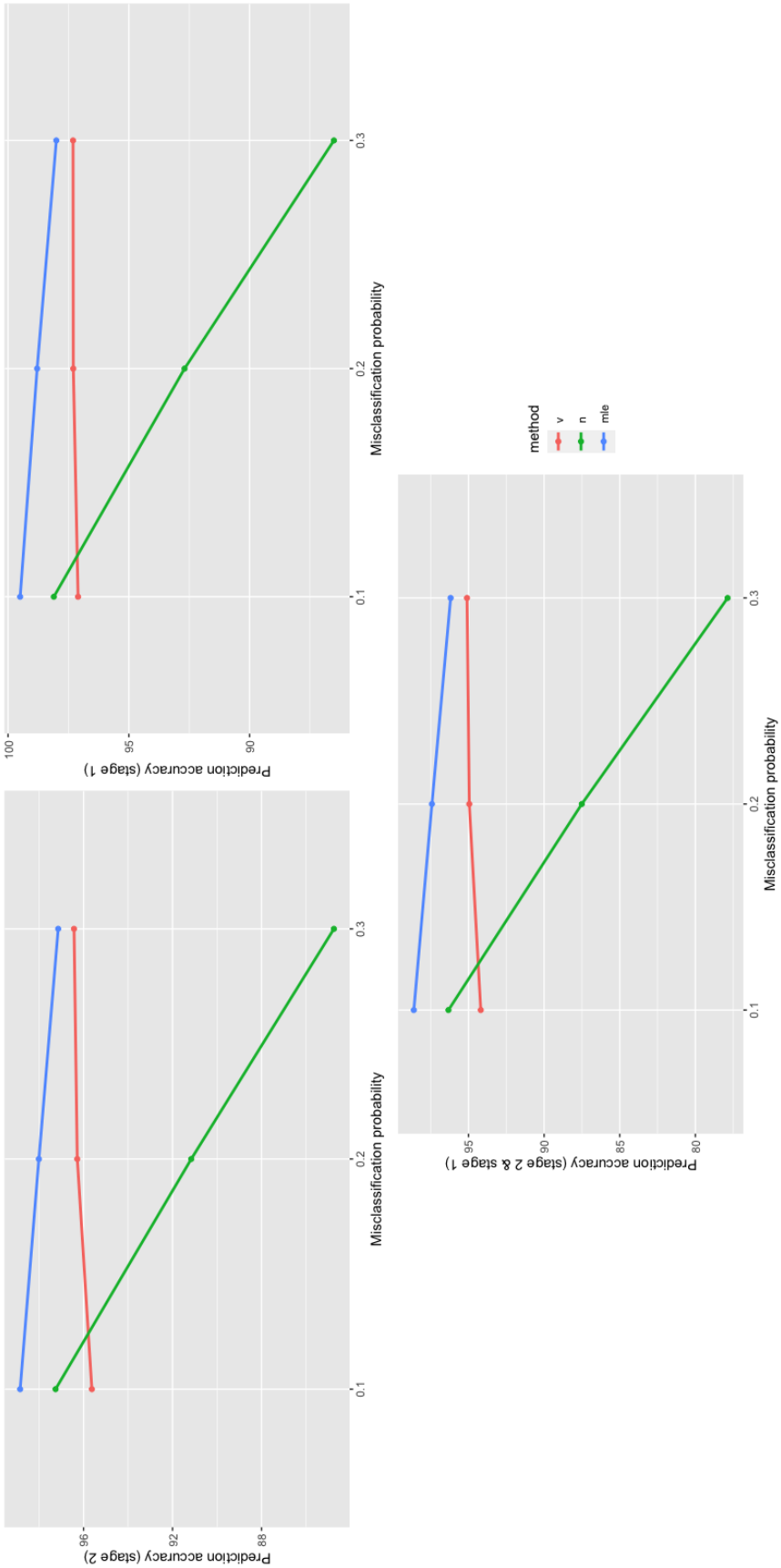

Figure 4.10: Prediction accuracy of optimal DTR with $\rho = 0.5$

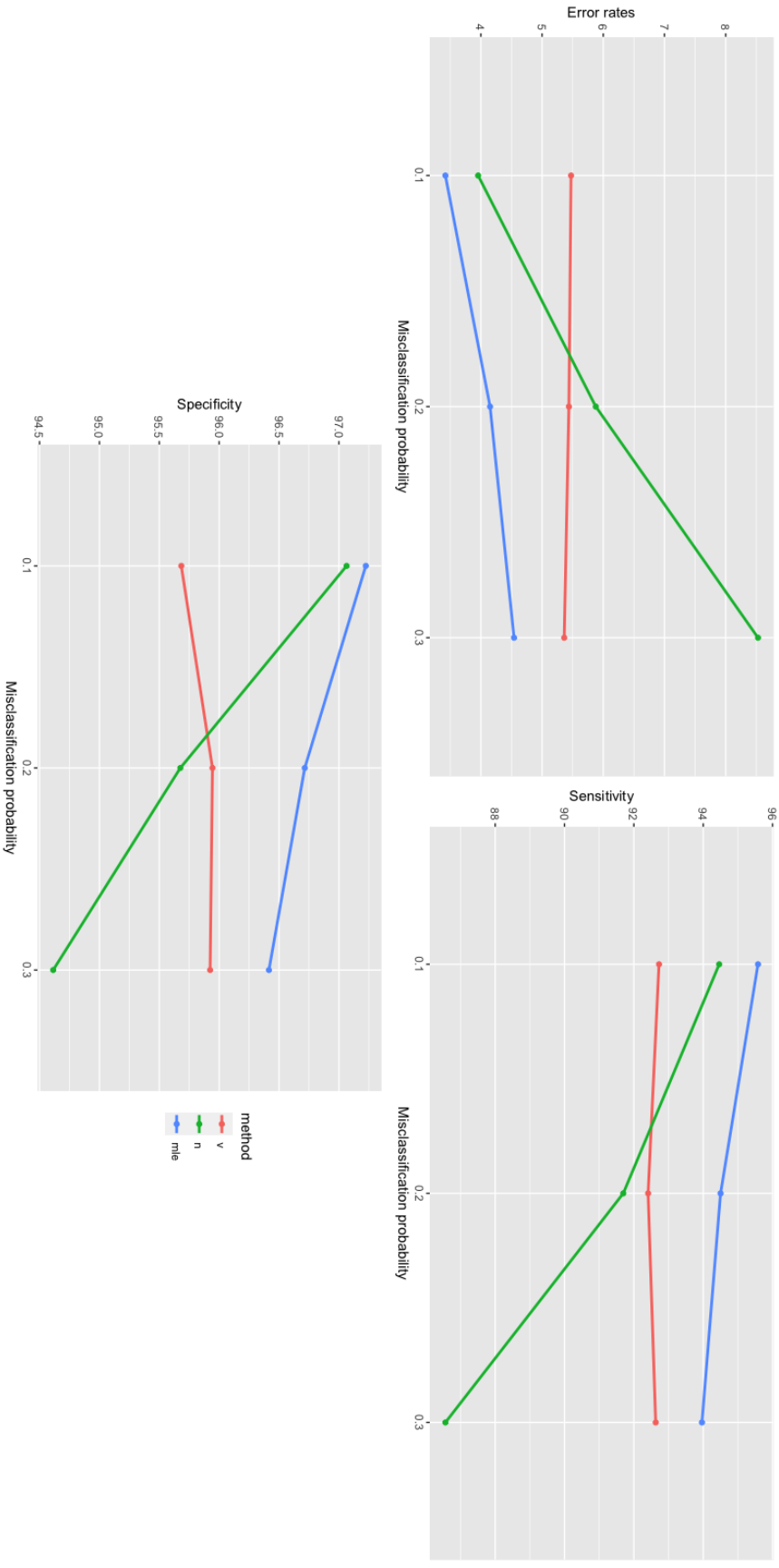

Figure 4.11: Predicted error rates, sensitivity, and specificity of the outcome with $\rho = 0.3$

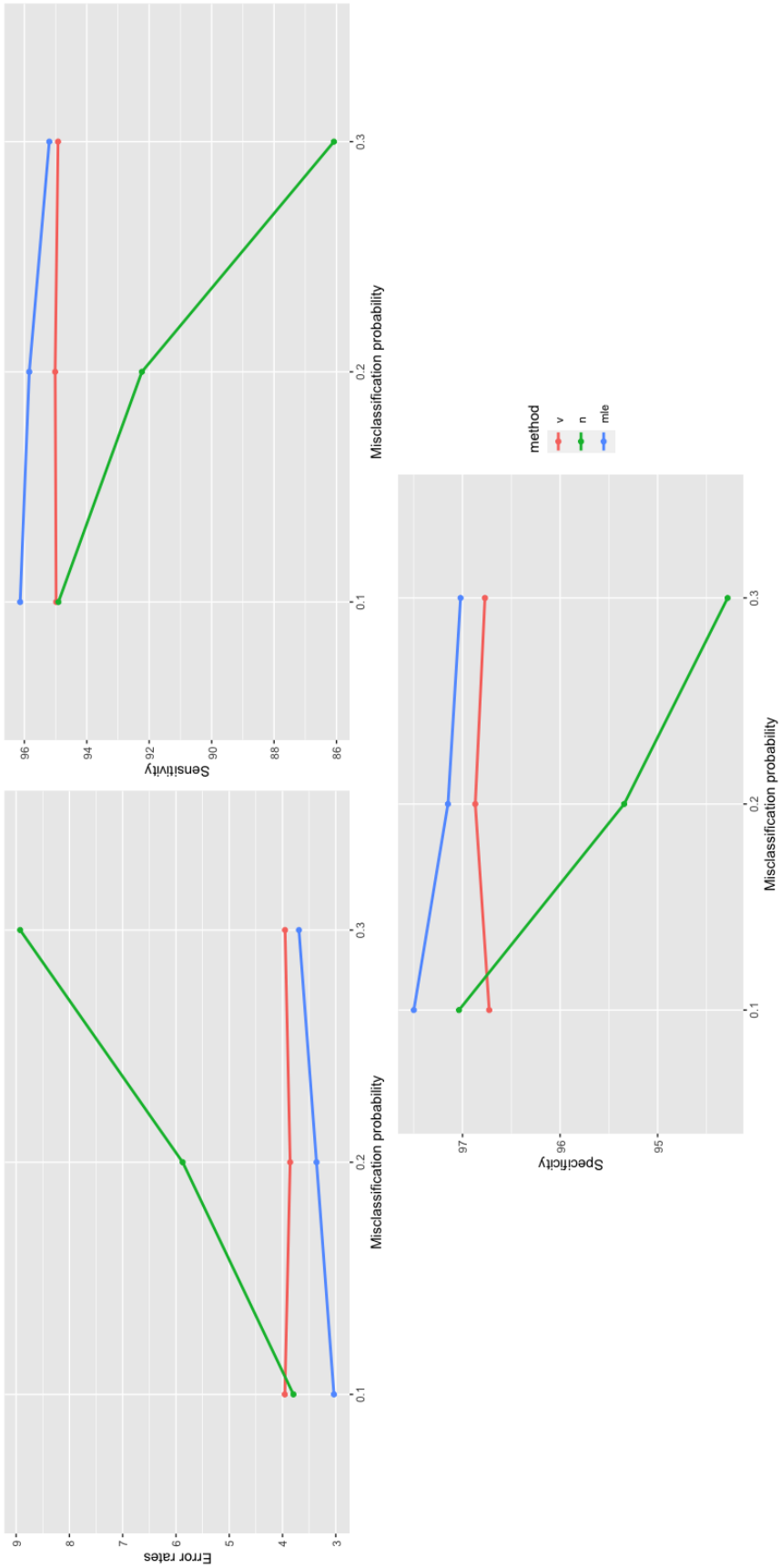

Figure 4.12: Predicted error rates, sensitivity, and specificity of the outcome with $\rho = 0.5$

4.6 Appendix

The proof of consistency in this section is based on a one-stage setting, and it can be intuitively extended to multiple stages.

Let $\boldsymbol{\theta} = (\boldsymbol{\beta}, \boldsymbol{\psi}, \gamma_{10}, \gamma_{01})$ and $\hat{\boldsymbol{\theta}}^{\boldsymbol{mle}}$ be the MLE estimator. The conditions for the property of consistency in Q-learning include:

(C1) Let Ω be the parameter space with finite dimension for $\boldsymbol{\theta}$. Ω is closed and compact. The true parameter value of $\boldsymbol{\theta}$ is interior to Ω.

(C2) The probability distributions with any two different values of $\boldsymbol{\theta}$ are distinct.

(C3) For an open subset ω of Ω that contains the true parameter value of $\boldsymbol{\theta}$, the first three derivatives of the log-likelihood $l(\boldsymbol{\theta})$ exist for $\boldsymbol{\theta} \in \omega$ almost surely. There exists a function M such that the n^{-1} times the absolute value of the the third derivative is bounded by M for $\boldsymbol{\theta} \in \omega$ and $E[M] < \infty$.

(C4) The information matrix $I(\boldsymbol{\theta})$ is finite and positive definite for $\boldsymbol{\theta} \in \omega$.

(C5) Assumptions (A1), (A2), (A5) in Sections (1.1.1) and (4.2.1) hold.

The conditions contain the regularity conditions (C1) - (C4) (Cox and Hinkley, 1979, p.281) and the assumptions that are necessary for Q-learning. The condition (C5) guarantees the identifiability of the parameter $\boldsymbol{\theta}$ in Q-learning to estimate a dynamic treatment regime. According to Pepe (1992), under the conditions (C1) - (C5), the MLE estimator $\hat{\boldsymbol{\theta}}^{\boldsymbol{mle}}$ that solves the equation $\frac{\partial}{\partial \boldsymbol{\theta}} \log L(\boldsymbol{\theta}) = 0$ satisifies

$$\hat{\boldsymbol{\theta}}^{\boldsymbol{mle}} \xrightarrow{p} \boldsymbol{\theta} \quad as\ n \to \infty,$$

where $L(\boldsymbol{\theta})$ is the likelihood stated in (4.4). Thus, $\hat{\boldsymbol{\psi}}^{\boldsymbol{mle}}$ is a consistent estimator of blip parameter $\boldsymbol{\psi}$.

Chapter 5

Summary and Future Work

The work presented in this thesis explores several statistical methods to address the issues in dynamic treatment regimes caused by covariate measurement error or outcome misclassification. Several regression-based methods in DTR with different types of outcomes are considered, and the effect of measurement error and misclassification on those methods is explored. Both simulation studies and data applications demonstrate the substantial impact of measurement error or misclassification on the analysis without errors corrected and the usefulness of the proposed correction methods to adjust for the effects.

Chapter 2 is motivated by the STAR*D study, in which the patients with a major depressive disorder were randomized at each level of study to one of the treatment options. The main objective of this study was to compare the effectiveness of different dynamic treatment regimes across multiple levels based on the QIDS score, which both clinicians and patients reported. In practice, the QIDS scores reported by patients and clinicians may be different from the true underlying QIDS score and, therefore, subject to measurement error. Q-learning is a widely used regression-based method to estimate optimal DTRs. This chapter explores the application of regression calibration in Q-learning to accommodate the effect of covariate measurement error with repeated measurements. Using the observed replicates, the RC estimates are created for the unobserved true covariates. Then, the patient's history is updated with the RC estimates, and a modified Q-learning algorithm is proposed to estimate the parameters and optimal DTR. Simulation studies demonstrate the significant improvements using the RC method in Q-learning in terms of bias reduction, the prediction accuracy of the optimal DTR,

and predicted optimal value function compared with the naive method in one-stage and multi-stage settings. Lastly, the proposed method is applied to the STAR*D data and compares its results with the naive method. The analysis results show that the statistical significance of the tailoring variable differs if the correction is made using the RC method, which suggests that the measurement error issue should not be ignored in an error-prone setting.

This work studies the measurement error effect with the classical additive model. It is of interest to explore other measurement error models such as Berkson and multiplicative models in Q-learning. Moreover, as regression calibration is known to perform well in linear models, other correction methods are worth exploring for highly nonlinear models.

In Chapter 3, the covariate measurement error in dynamic weighted survival modeling is studied. This DWSurv approach is practical but developed under the assumption that the covariates are free from mismeasurement. If this assumption is violated, it remains unclear what the impact would be on the estimation of parameters and optimal DTR. Therefore, in this chapter, we investigate the covariate measurement error effect on DWSurv for validation data and develop two correction methods, the k-nearest neighbors method and the weighted least squares method, to eliminate the effect. The proposed correction methods estimate the missing values of the true covariates using the mismeasured variables that are completely observable in the data. Both methods are easy to understand and fast to implement. The theoretical property of the kNN estimator is also established. Both simulation studies and data analysis showcase the competency of proposed methods in one-stage and multi-stage settings. The results show that using the proposed kNN and WLS methods leads to significant improvements in bias-reduction and restoration of the double robustness property in DWSurv. In the predictive scenarios, the proposed methods enhance the prediction accuracy of optimal DTR and the predicted optimal overall survival times. Lastly, the proposed methods are applied to the MIMIC-III data to estimate the optimal treatment decision rules. The analysis results reveal the significant impacts of discarding some data with missing covariates in the estimated optimal DTR.

There are a few possible directions for future work. First of all, the proposed kNN method uses Euclidean distance as the distance measure in this project. It is of interest to consider other distance functions, such as Manhattan, Minkowsky, Chebychev, Chi-square distances, etc., for the proposed method. Secondly, as k is defined in relation to the size of the validation data,

one may also explore different choices of k to improve the performance of the kNN method. Lastly, besides independent censoring, other types of censoring can be considered in the study, such as the censoring that depends on covariates.

In Chapter 4, Q-learning with binary outcomes is explored, with the outcome being subject to misclassification in the context of internal validation/main study data design. When the outcome misclassification is ignored, the estimation in Q-learning is severely biased. Therefore, the maximum likelihood estimation method is proposed to accommodate the misclassification effect in Q-learning. The proposed MLE method is established based on the relationship between two conditional probabilities of the true outcome and the outcome observed with error. Simulation studies are conducted to demonstrate the satisfactory performance of the proposed method in both one-stage and multi-stage settings. In particular, the MLE method is shown to be numerically stable and robust against various magnitudes of misclassification rates in the outcome model. Sensitivity analyses are also conducted using the NHEFS data and smoking cessation data to compare the optimal treatment decision rules estimated from the naive method and the proposed method. By incorporating the misclassification in the analysis, the estimated optimal treatment rules are shown to be different, and the statistical significance of the tailoring variable is also altered. It reveals a non-negligible impact of misclassification in the NHEFS and smoking cessation data.

In Chapter 4, a non-differential misclassification model is assumed for the proposed method, in which the dependence on the covariates and/or treatment is suppressed. Therefore, it is interesting to study other misclassification models, such as a differential one dependent on covariates and/or treatment. Secondly, one may also consider the misclassification problem with replicate data. In other words, instead of observing the true outcome in a small subset of data, the replicates of the outcome are observed. In such circumstances, it is necessary to explore other approaches to correct the misclassification in Q-learning.

In summary, the errors-in-variables problem in dynamic treatment regimes is a new and challenging topic. While the problems that have been studied in this thesis focus on measurement error and misclassification in a few popular DTR approaches based on the class of outcomes, many more complex situations remain unexplored. We anticipate the pursuit of exploring these in the future.

Bibliography

Aha, D. (1997). Special issue on lazy learning. *Artificial Intelligence Review*, 11:7–10.

Alkasassbeh, M., Altarawneh, G. A., and Hassanat, A. (2015). On enhancing the performance of nearest neighbour classifiers using hassanat distance metric. *arXiv preprint arXiv:1501.00687*.

Artman, W. J., Ertefaie, A., Lynch, K. G., and McKay, J. R. (2020). Bayesian set of best dynamic treatment regimes and sample size determination for smarts with binary outcomes. *arXiv preprint arXiv:2008.02341*.

Austin, P. C. (2011). An introduction to propensity score methods for reducing the effects of confounding in observational studies. *Multivariate Behavioral Research*, 46(3):399–424.

Azadkia, M. (2019). Optimal choice of k for k-nearest neighbor regression. *arXiv preprint arXiv:1909.05495*.

Bai, X., Tsiatis, A. A., Lu, W., and Song, R. (2017). Optimal treatment regimes for survival endpoints using a locally-efficient doubly-robust estimator from a classification perspective. *Lifetime Data Analysis*, 23(4):585–604.

Bartlett, J. W., De Stavola, B. L., and Frost, C. (2009). Linear mixed models for replication data to efficiently allow for covariate measurement error. *Statistics in Medicine*, 28(25):3158–3178.

Biau, G. and Devroye, L. (2015). *Lectures on the nearest neighbor method*. Springer, New York, NY, USA.

Carroll, R. J., Ruppert, D., Stefanski, L. A., and Crainiceanu, C. M. (2006). *Measurement error in nonlinear models: a modern perspective*. CRC press, Boca Raton, FL, USA.

Carroll, R. J. and Stefanski, L. A. (1990). Approximate quasi-likelihood estimation in models with surrogate predictors. *Journal of the American Statistical Association*, 85(411):652–663.

Chakraborty, B. (2013). *Statistical methods for dynamic treatment regimes*. Springer, New York, NY, USA.

Chakraborty, B., Laber, E. B., and Zhao, Y. (2013). Inference for optimal dynamic treatment regimes using an adaptive m-out-of-n bootstrap scheme. *Biometrics*, 69(3):714–723.

Chakraborty, B., Murphy, S., and Strecher, V. (2010). Inference for non-regular parameters in optimal dynamic treatment regimes. *Statistical Methods in Medical Research*, 19(3):317–343.

Chakraborty, B. and Murphy, S. A. (2014). Dynamic treatment regimes. *Annual Review of Statistics and Its Application*, 1:447–464.

Chen, R., Huling, J. D., Chen, G., and Yu, M. (2021). Robust sample weighting to facilitate individualized treatment rule learning for a target population. *arXiv preprint arXiv:2105.00581*.

Chen, Z., Yi, G. Y., and Wu, C. (2014). Marginal analysis of longitudinal ordinal data with misclassification in both response and covariates. *Biometrical Journal*, 56(1):69–85.

Cheng, D., Zhang, S., Deng, Z., Zhu, Y., and Zong, M. (2014). knn algorithm with data-driven k value. In *International Conference on Advanced Data Mining and Applications*, pages 499–512, Guilin, Guangxi, China. Springer.

Chomboon, K., Chujai, P., Teerarassamee, P., Kerdprasop, K., and Kerdprasop, N. (2015). An empirical study of distance metrics for k-nearest neighbor algorithm. In *Proceedings of the 3rd International Conference on Industrial Application Engineering*, pages 280–285, Kitakyushu, Japan. Institute of Industrial Applications Engineers.

Cole, S. R., Chu, H., and Greenland, S. (2006). Multiple-imputation for measurement-error correction. *International Journal of Epidemiology*, 35(4):1074–1081.

Cook, C. H., Praba, A. C., Beery, P. R., and Martin, L. C. (2002). Transthoracic echocardiography is not cost-effective in critically ill surgical patients. *Journal of Trauma and Acute Care Surgery*, 52(2):280–284.

Cox, D. R. and Hinkley, D. V. (1979). *Theoretical statistics*. CRC Press, Boca Raton, FL, USA.

Daniel Paulino, C., Soares, P., and Neuhaus, J. (2003). Binomial regression with misclassification. *Biometrics*, 59(3):670–675.

Devroye, L., Gyorfi, L., Krzyzak, A., and Lugosi, G. (1994). On the strong universal consistency of nearest neighbor regression function estimates. *The Annals of Statistics*, 22(3):1371–1385.

Edwards, J. K., Cole, S. R., Troester, M. A., and Richardson, D. B. (2013). Accounting for misclassified outcomes in binary regression models using multiple imputation with internal validation data. *American Journal of Epidemiology*, 177(9):904–912.

Fan, Z. and Wang, X.-F. (2009). Marginal hazards model for multivariate failure time data with auxiliary covariates. *Journal of Nonparametric Statistics*, 21(7):771–786.

Feng, M., McSparron, J. I., Kien, D. T., Stone, D. J., Roberts, D. H., Schwartzstein, R. M., Vieillard-Baron, A., and Celi, L. A. (2018). Transthoracic echocardiography and mortality in sepsis: analysis of the mimic-iii database. *Intensive Care Medicine*, 44(6):884–892.

Folleco, A., Khoshgoftaar, T. M., Van Hulse, J., and Bullard, L. (2008). Identifying learners robust to low quality data. In *2008 IEEE International Conference on Information Reuse and Integration*, pages 190–195, Las Vegas, NV, USA. IEEE.

Freedman, L. S., Fainberg, V., Kipnis, V., Midthune, D., and Carroll, R. J. (2004). A new method for dealing with measurement error in explanatory variables of regression models. *Biometrics*, 60(1):172–181.

Freedman, L. S., Midthune, D., Carroll, R. J., and Kipnis, V. (2008). A comparison of regression calibration, moment reconstruction and imputation for adjusting for covariate measurement error in regression. *Statistics in Medicine*, 27(25):5195–5216.

Fu, S., He, Q., Zhang, S., and Liu, Y. (2019). Robust outcome weighted learning for optimal individualized treatment rules. *Journal of Biopharmaceutical Statistics*, 29(4):606–624.

Gerlach, R. and Stamey, J. (2007). Bayesian model selection for logistic regression with misclassified outcomes. *Statistical Modelling*, 7(3):255–273.

Ghosh, P. and Chakraborty, B. (2018). Comparison of dynamic treatment regimes with an ordinal outcome. *arXiv preprint arXiv:1808.07287*.

Gleser, L. J. (1990). Improvements of the naive approach to estimation in nonlinear errors-in-variables regression models. *Contemporary Mathematics*, 112:99–114.

Goldberg, Y. and Kosorok, M. R. (2012). Q-learning with censored data. *Annals of Statistics*, 40(1):529–560.

Gou, J., Xiong, T., and Kuang, Y. (2011). A novel weighted voting for k-nearest neighbor rule. *Journal of Computers*, 6(5):833–840.

Granville, K. and Fan, Z. (2012). Accelerated failure time models with auxiliary covariates. *Journal of Biometrics and Biostatistics*, 3(6):1000152.

Granville, K. and Fan, Z. (2014). Buckley-james estimator of AFT models with auxiliary covariates. *PLOS ONE*, 9(8):e104817.

Gray, C. M. (2018). *Use of the Bayesian family of methods to correct for effects of exposure measurement error in polynomial regression models*. PhD thesis, London School of Hygiene & Tropical Medicine.

Greer, T. L., Kurian, B. T., and Trivedi, M. H. (2010). Defining and measuring functional. *CNS Drugs*, 24(4):267–284.

Hassanat, A. B., Abbadi, M. A., Altarawneh, G. A., and Alhasanat, A. A. (2014). Solving the problem of the k parameter in the knn classifier using an ensemble learning approach. *arXiv preprint arXiv:1409.0919*.

Hausman, J. A., Abrevaya, J., and Scott-Morton, F. M. (1998). Misclassification of the dependent variable in a discrete-response setting. *Journal of Econometrics*, 87(2):239–269.

He, W., Xiong, J., Yi, G. Y., et al. (2012). Simex r package for accelerated failure time models with covariate measurement error. *Journal of Statistical Software*, 46(1):1–14.

He, W., Yi, G. Y., and Xiong, J. (2007). Accelerated failure time models with covariates subject to measurement error. *Statistics in Medicine*, 26(26):4817–4832.

Henderson, R., Ansell, P., and Alshibani, D. (2010). Regret-regression for optimal dynamic treatment regimes. *Biometrics*, 66(4):1192–1201.

Hernán, M. A. and Robins, J. M. (2020). *Causal inference: what if*. Chapman & Hall/CRC, Boca Raton, FL, USA.

Hu, C. and Lin, D. (2002). Cox regression with covariate measurement error. *Scandinavian Journal of Statistics*, 29(4):637–655.

Hu, L.-Y., Huang, M.-W., Ke, S.-W., and Tsai, C.-F. (2016). The distance function effect on k-nearest neighbor classification for medical datasets. *SpringerPlus*, 5(1):1–9.

Huang, X. and Ning, J. (2012). Analysis of multi-stage treatments for recurrent diseases. *Statistics in Medicine*, 31(24):2805–2821.

Huang, X., Ning, J., and Wahed, A. S. (2014). Optimization of individualized dynamic treatment regimes for recurrent diseases. *Statistics in Medicine*, 33(14):2363–2378.

IsHak, W. W., James, D. M., Mirocha, J., Youssef, H., Tobia, G., Pi, S., Collison, K. L., and Cohen, R. M. (2016). Patient-reported functioning in major depressive disorder. *Therapeutic Advances in Chronic Disease*, 7(3):160–169.

Jiang, R., Lu, W., Song, R., and Davidian, M. (2017). On estimation of optimal treatment regimes for maximizing t-year survival probability. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, 79(4):1165–1185.

Jin, L. H., Liu, Y. Y., and Wu, L. (2019). Weighted least squares method for the accelerated failure time model with auxiliary covariates. *Acta Mathematica Sinica, English Series*, 35(7):1163–1178.

Johnson, A. E., Pollard, T. J., Shen, L., Li-Wei, H. L., Feng, M., Ghassemi, M., Moody, B., Szolovits, P., Celi, L. A., and Mark, R. G. (2016). Mimic-iii, a freely accessible critical care database. *Scientific Data*, 3(1):1–9.

Johnson, A. E., Stone, D. J., Celi, L. A., and Pollard, T. J. (2018). The mimic code repository: enabling reproducibility in critical care research. *Journal of the American Medical Informatics Association*, 25(1):32–39.

Keogh, R. H. and White, I. R. (2014). A toolkit for measurement error correction, with a focus on nutritional epidemiology. *Statistics in Medicine*, 33(12):2137–2155.

Klausch, T., van de Ven, P., van de Brug, T., van de Wiel, M. A., and Berkhof, J. (2018). Estimating bayesian optimal treatment regimes for dichotomous outcomes using observational data. *arXiv preprint arXiv:1809.06679*.

Laber, E. B., Lizotte, D. J., Qian, M., Pelham, W. E., and Murphy, S. A. (2014). Dynamic treatment regimes: Technical challenges and applications. *Electronic Journal of Statistics*, 8(1):1225–1272.

Laber, E. B. and Zhao, Y.-Q. (2015). Tree-based methods for individualized treatment regimes. *Biometrika*, 102(3):501–514.

Lederer, W. and Küchenhoff, H. (2006). A short introduction to the simex and mcsimex. *The Newsletter of the R Project*, 6(4):26–31.

Lee, S. M., Landry, J., Jones, P. M., Buhrmann, O., and Morley-Forster, P. (2013). The effectiveness of a perioperative smoking cessation program: a randomized clinical trial. *Anesthesia & Analgesia*, 117(3):605–613.

Lin, C.-H., Yen, Y.-C., Chen, M.-C., and Chen, C.-C. (2014). Depression and pain impair daily functioning and quality of life in patients with major depressive disorder. *Journal of Affective Disorders*, 166:173–178.

Liu, Y., Wang, Y., Kosorok, M. R., Zhao, Y., and Zeng, D. (2018). Augmented outcome-weighted learning for estimating optimal dynamic treatment regimens. *Statistics in Medicine*, 37(26):3776–3788.

Liu, Y., Wu, Y., and Zhou, H. (2010). Multivariate failure times regression with a continuous auxiliary covariate. *Journal of Multivariate Analysis*, 101(3):679–691.

Liu, Y., Zhou, H., and Cai, J. (2009). Estimated pseudopartial-likelihood method for correlated failure time data with auxiliary covariates. *Biometrics*, 65(4):1184–1193.

Lopes, N. and Ribeiro, B. (2015). On the impact of distance metrics in instance-based learning algorithms. In *Iberian Conference on Pattern Recognition and Image Analysis*, pages 48–56, Santiago de Compostela, Spain. Springer.

Lyles, R. H. and Lin, J. (2010). Sensitivity analysis for misclassification in logistic regression via likelihood methods and predictive value weighting. *Statistics in Medicine*, 29(22):2297–2309.

Lyles, R. H., Tang, L., Superak, H. M., King, C. C., Celentano, D. D., Lo, Y., and Sobel, J. D. (2011). Validation data-based adjustments for outcome misclassification in logistic regression: an illustration. *Epidemiology*, 22(4):589–597.

Magder, L. S. and Hughes, J. P. (1997). Logistic regression when the outcome is measured with uncertainty. *American Journal of Epidemiology*, 146(2):195–203.

Moodie, E. E., Chakraborty, B., and Kramer, M. S. (2012). Q-learning for estimating optimal dynamic treatment rules from observational data. *Canadian Journal of Statistics*, 40(4):629–645.

Moodie, E. E., Dean, N., and Sun, Y. R. (2014). Q-learning: Flexible learning about useful utilities. *Statistics in Biosciences*, 6(2):223–243.

Muff, S., Riebler, A., Held, L., Rue, H., and Saner, P. (2015). Bayesian analysis of measurement error models using integrated nested laplace approximations. *Journal of the Royal Statistical Society: Series C (Applied Statistics)*, 64(2):231–252.

Mulak, P. and Talhar, N. (2015). Analysis of distance measures using k-nearest neighbor algorithm on kdd dataset. *International Journal of Science and Research*, 4(7):2101–2104.

Murphy, S. A. (2003). Optimal dynamic treatment regimes. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, 65(2):331–355.

Murray, T. A., Yuan, Y., and Thall, P. F. (2018). A bayesian machine learning approach for optimizing dynamic treatment regimes. *Journal of the American Statistical Association*, 113(523):1255–1267.

Mwalili, S. M., Lesaffre, E., and Declerck, D. (2008). The zero-inflated negative binomial regression model with correction for misclassification: an example in caries research. *Statistical Methods in Medical Research*, 17(2):123–139.

Nadkarni, P. (2016). *Clinical research computing: A practitioner's handbook*. Academic Press, Cambridge, MA, USA.

Neuhaus, J. M. (1999). Bias and efficiency loss due to misclassified responses in binary regression. *Biometrika*, 86(4):843–855.

Neuhaus, J. M. (2002). Analysis of clustered and longitudinal binary data subject to response misclassification. *Biometrics*, 58(3):675–683.

Orellana, L., Rotnitzky, A., and Robins, J. M. (2010). Dynamic regime marginal structural mean models for estimation of optimal dynamic treatment regimes, part i: main content. *The International Journal of Biostatistics*, 6(2):Article 8.

Pepe, M. S. (1992). Inference using surrogate outcome data and a validation sample. *Biometrika*, 79(2):355–365.

Pepe, M. S., Reilly, M., and Fleming, T. R. (1994). Auxiliary outcome data and the mean score method. *Journal of Statistical Planning and Inference*, 42(1-2):137–160.

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

Prescott, G. J. and Garthwaite, P. H. (2002). A simple bayesian analysis of misclassified binary data with a validation substudy. *Biometrics*, 58(2):454–458.

Prescott, G. J. and Garthwaite, P. H. (2005). Bayesian analysis of misclassified binary data from a matched case–control study with a validation sub-study. *Statistics in Medicine*, 24(3):379–401.

Robins, J., Orellana, L., and Rotnitzky, A. (2008). Estimation and extrapolation of optimal treatment and testing strategies. *Statistics in Medicine*, 27(23):4678–4721.

Robins, J. M. (2000). Marginal structural models versus structural nested models as tools for causal inference. In *Statistical models in epidemiology, the environment, and clinical trials*, pages 95–133. Springer, New York, NY, USA.

Robins, J. M. (2004). Optimal structural nested models for optimal sequential decisions. In *Proceedings of the Second Seattle Symposium in Biostatistics*, pages 189–326, Seattle, WA, USA. Springer.

Robins, J. M. and Greenland, S. (1994). Adjusting for differential rates of prophylaxis therapy for pcp in high-versus low-dose azt treatment arms in an aids randomized trial. *Journal of the American Statistical Association*, 89(427):737–749.

Ronning, G. and Rosemann, M. (2008). Simex estimation in case of correlated measurement errors. *AStA Advances in Statistical Analysis*, 92(4):391–404.

Rush, A. J., Fava, M., Wisniewski, S. R., Lavori, P. W., Trivedi, M. H., Sackeim, H. A., Thase, M. E., Nierenberg, A. A., Quitkin, F. M., Kashner, T. M., et al. (2004). Sequenced treatment alternatives to relieve depression (STAR* D): rationale and design. *Controlled Clinical Trials*, 25(1):119–142.

Rush, A. J., Trivedi, M. H., Ibrahim, H. M., Carmody, T. J., Arnow, B., Klein, D. N., Markowitz, J. C., Ninan, P. T., Kornstein, S., Manber, R., et al. (2003). The 16-item quick in-

ventory of depressive symptomatology (qids), clinician rating (qids-c), and self-report (qids-sr): a psychometric evaluation in patients with chronic major depression. *Biological Psychiatry*, 54(5):573–583.

Schulte, P. J., Tsiatis, A. A., Laber, E. B., and Davidian, M. (2014). Q-and a-learning methods for estimating optimal dynamic treatment regimes. *Statistical Science*, 29(4):640–661.

Shu, D. and Yi, G. Y. (2019a). Causal inference with measurement error in outcomes: Bias analysis and estimation methods. *Statistical Methods in Medical Research*, 28(7):2049–2068.

Shu, D. and Yi, G. Y. (2019b). ipwerrory: An r package for estimation of average treatment effect with misclassified binary outcome. *R Journal*, 11(1):337–351.

Simoneau, G., Moodie, E. E., Nijjar, J. S., and Platt, R. W. (2020a). Finite sample variance estimation for optimal dynamic treatment regimes of survival outcomes. *Statistics in Medicine*, 39(29):4466–4479.

Simoneau, G., Moodie, E. E., Nijjar, J. S., Platt, R. W., Investigators, S. E. R. A. I. C., et al. (2020b). Estimating optimal dynamic treatment regimes with survival outcomes. *Journal of the American Statistical Association*, 115(531):1531–1539.

Song, R., Wang, W., Zeng, D., and Kosorok, M. R. (2015). Penalized q-learning for dynamic treatment regimens. *Statistica Sinica*, 25(3):901–920.

Spicker, D. and Wallace, M. P. (2020). Measurement error and precision medicine: Error-prone tailoring covariates in dynamic treatment regimes. *Statistics in Medicine*, 39(26):3732–3755.

Stefanski, L. A. and Cook, J. R. (1995). Simulation-extrapolation: the measurement error jackknife. *Journal of the American Statistical Association*, 90(432):1247–1256.

Sun, S. and Huang, R. (2010). An adaptive k-nearest neighbor algorithm. In *2010 Seventh International Conference on Fuzzy Systems and Knowledge Discovery*, volume 1, pages 91–94, Yantai, Shandong, China. IEEE.

Thomas, L., Stefanski, L., and Davidian, M. (2011). A moment-adjusted imputation method for measurement error models. *Biometrics*, 67(4):1461–1470.

Trivedi, M. H., Morris, D. W., Wisniewski, S. R., Lesser, I., Nierenberg, A. A., Daly, E., Kurian, B. T., Gaynes, B. N., Balasubramani, G., and Rush, A. J. (2013). Increase in work productivity of depressed individuals with improvement in depressive symptom severity. *American Journal of Psychiatry*, 170(6):633–641.

Tsiatis, A. A. (2019). *Dynamic Treatment Regimes: Statistical Methods for Precision Medicine*. CRC press, Boca Raton, FL, USA.

van der Laan, M. J. (2006). Causal effect models for intention to treat and realistic individualized treatment rules. *University of California, Berkeley Division of Biostatistics Working Paper Series. Working Paper 203.*

van der Laan, M. J. and Petersen, M. L. (2007). Causal effect models for realistic individualized treatment and intention to treat rules. *The International Journal of Biostatistics*, 3(1):Article 3.

Wallace, M. P. and Moodie, E. E. (2015). Doubly-robust dynamic treatment regimen estimation via weighted least squares. *Biometrics*, 71(3):636–644.

Wallace, M. P., Moodie, E. E., and Stephens, D. A. (2019). Model selection for g-estimation of dynamic treatment regimes. *Biometrics*, 75(4):1205–1215.

Watkins, C. J. C. H. (1989). *Learning from delayed rewards*. PhD thesis, King's College, Cambridge, UK.

Xiao, W., Zhang, H. H., and Lu, W. (2019). Robust regression for optimal individualized treatment rules. *Statistics in Medicine*, 38(11):2059–2073.

Xu, L., Crammer, K., Schuurmans, D., et al. (2006). Robust support vector machine training via convex outlier ablation. In *Association for the Advancement of Artificial Intelligence*, volume 6, pages 536–542, Boston, MA, USA. AAAI Press.

Xue, F., Zhang, Y., Zhou, W., Fu, H., and Qu, A. (2022). Multicategory angle-based learning for estimating optimal dynamic treatment regimes with censored data. *Journal of the American Statistical Association*, 117(539):1438–1451.

Yang, X., Song, Q., and Wang, Y. (2007). A weighted support vector machine for data classification. *International Journal of Pattern Recognition and Artificial Intelligence*, 21(5):961–976.

Yi, G. Y. (2017). *Statistical analysis with measurement error or misclassification: strategy, method and application*. Springer, New York, NY, USA.

Yi, G. Y., Ma, Y., Spiegelman, D., and Carroll, R. J. (2015a). Functional and structural methods with mixed measurement error and misclassification in covariates. *Journal of the American Statistical Association*, 110(510):681–696.

Yi, G. Y., Tan, X., and Li, R. (2015b). Variable selection and inference procedures for marginal analysis of longitudinal data with missing observations and covariate measurement error. *Canadian Journal of Statistics*, 43(4):498–518.

Zhang, B., Tsiatis, A. A., Laber, E. B., and Davidian, M. (2012). A robust method for estimating optimal treatment regimes. *Biometrics*, 68(4):1010–1018.

Zhang, B., Tsiatis, A. A., Laber, E. B., and Davidian, M. (2013). Robust estimation of optimal dynamic treatment regimes for sequential treatment decisions. *Biometrika*, 100(3):681–694.

Zhang, Q. and Yi, G. Y. (2019). R package for analysis of data with mixed measurement error and misclassification in covariates: augsimex. *Journal of Statistical Computation and Simulation*, 89(12):2293–2315.

Zhang, Q. and Yi, G. Y. (2020). Genetic association studies with bivariate mixed responses subject to measurement error and misclassification. *Statistics in Medicine*, 39(26):3700–3719.

Zhang, Q. and Yi, G. Y. (2021). Marginal analysis of bivariate mixed responses with measurement error and misclassification. *Statistical Methods in Medical Research*, 30(5):1155–1186.

Zhang, S., Cheng, D., Deng, Z., Zong, M., and Deng, X. (2018). A novel knn algorithm with data-driven k parameter computation. *Pattern Recognition Letters*, 109:44–54.

Zhang, Z. (2016). Introduction to machine learning: k-nearest neighbors. *Annals of Translational Medicine*, 4(11):218.

Zhao, Y., Zeng, D., Rush, A. J., and Kosorok, M. R. (2012). Estimating individualized treatment rules using outcome weighted learning. *Journal of the American Statistical Association*, 107(499):1106–1118.

Zhao, Y.-Q., Zeng, D., Laber, E. B., and Kosorok, M. R. (2015a). New statistical learning methods for estimating optimal dynamic treatment regimes. *Journal of the American Statistical Association*, 110(510):583–598.

Zhao, Y.-Q., Zeng, D., Laber, E. B., Song, R., Yuan, M., and Kosorok, M. R. (2015b). Doubly robust learning for estimating individualized treatment with censored data. *Biometrika*, 102(1):151–168.

Zhou, H. and Wang, C.-Y. (2000). Failure time regression with continuous covariates measured with error. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, 62(4):657–665.

Zhou, L., Cai, G., Xu, Z., Weng, Q., Ye, Q., and Chen, C. (2019). High positive end expiratory pressure levels affect hemodynamics in elderly patients with hypertension admitted to the intensive care unit: a prospective cohort study. *BMC Pulmonary Medicine*, 19(1):1–9.

Zhou, X., Mayer-Hamblett, N., Khan, U., and Kosorok, M. R. (2017). Residual weighted learning for estimating individualized treatment rules. *Journal of the American Statistical Association*, 112(517):169–187.

Zhu, X. and Wu, X. (2004). Class noise vs. attribute noise: A quantitative study. *Artificial Intelligence Review*, 22(3):177–210.

Curriculum Vitae

Name:	Dan Liu
Post-Secondary Education	Western University, *London, ON, Canada* Ph.D. in Statistics, 2018 - 2022
	Boston University, *Boston, MA, USA* M.S. in Management Studies, 2013 - 2015 M.S. in Systems Engineering, 2011 - 2013
	Wuhan University, *Wuhan, Hubei, China* B.Eng. in Electrical Engineering, 2007 - 2011
Honours and Awards:	Western Graduate Research Scholarship, 2018 - 2022 Western Graduate Teaching Assistantship, 2018 - 2022 Statistical Society of Canada Student Travel Award, 2021 Vector Institute Deep Learning and Reinforcement Learning Summer School Certificate, 2019
Related Work Experience:	Western University, 2018 - 2022 Teaching Assistant and Research Assistant

Publications:

1. Liu, D., Du, Y., Charvadeh, Y. K., Cui, J., Chen, L. P., Deng, G., ... & Yi, G. Y. (2020). A real time and interactive web-based platform for visualizing and analyzing COVID-19 in Canada. *International Journal of Statistics and Probability, 9*(5), 23-29.

2. Liu, D. and He. W. Dynamic Treatment Regimes with Measurement Error in Covariates: a Q-learning Approach. In preparation.

3. Liu, D. and He. W. Dynamic Treatment Regimes with Survival Response and Covariate

Measurement Error. In preparation.

4. Liu, D. and He. W. Q-learning with Misclassified Binary Outcomes. In preparation.

Conference Presentations:

1. Dan Liu & Wenqing He (2021). Effect of Measurement Error on Dynamic Weighted Survival Modeling. *The Annual General Meeting of the Statistical Society of Canada*, 06/2021.

2. Dan Liu & Wenqing He (2022). Q-learning with Misclassified Response in Binary Regression. *The Annual General Meeting of the Statistical Society of Canada*, 06/2022.