8-18-2023 1:30 PM

Optimizing Dynamic Treatment Regimes with Q-Learning: Complications due to Error-Prone Data and Applications to COVID-19 Data

Yasin Khadem Charvadeh, *Western University*

Supervisor: Yi, Grace Y., *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

Much of everyday life depends on making informed decisions. The process of decision-making depends heavily on gathering pertinent information which comes from quality data. To obtain such relevant information, one needs appropriate tools for collecting and analyzing data. Statistical methods are an essential and vital tool for collecting and turning data into useful information. The role of statistics in decision-making became more apparent with the emergence of COVID-19 as a global pandemic which drew significant attention to using data as a fundamental component of evidence-based decision-making. With the involvement of COVID-19 in many areas, such as medicine, epidemiology, and economy, many researchers from different branches of science conducted research on COVID-19-related problems from various angles. In this thesis, we employ statistical modeling and methods to examine COVID-19 data, and we develop new methods to address new issues that invalidate some standard methods.

The first project examines the clinical manifestations and epidemiological features of COVID-19 pertinent to the early pandemic stage. We employ semiparametric and nonparametric survival models as well as text mining and data visualization techniques to shed light on COVID-19 incubation and recovery times. Using a dataset from January 22, 2020 to March 29, 2020, we explore some of the risk factors associated with the severity of the disease. Our analysis shows that the median incubation time is about 5 days, and older people tend to have a longer incubation period. The median time for infected people to recover is about 20 days, and the recovery time is significantly associated with age but not gender.

The second project uses a data-driven approach to explore optimal preventive measures with the goal of lowering the COVID-19 case fatality rate. In this study, using a reinforcement learning algorithm to characterize the evolving situation and country-specific features, we study the effectiveness of country-level preventive actions. Our analytical results suggest that country-specific characteristics and the baseline information of COVID-19 determine optimal preventive policies. Furthermore, our study reveals that the factors significantly associated with the COVID-19 case fatality rate include the population proportion of elders ages 65 and over, gross domestic product per capita, obesity prevalence, substance use prevalence, population density, and health system quality.

The framework of the second project is cast under dynamic treatment regimes, which has

attracted lots of interest among statisticians and quantitative researchers. Dynamic treatment regimes require a special class of designs, called *sequential multiple assignment randomized trial* (SMART) designs, for the validity of the results derived from them. However, similar to randomized controlled designs, SMART designs are prone to the violation of the condition that all variables must be precisely measured. Misclassification in categorical variables as well as measurement error in continuous variables are among common problems in the applications. In the next two projects, we explore correction strategies to overcome the effects induced by such erroneous data.

The third project examines Q-learning with covariates subject to misclassification. We present two correction methods, namely regression calibration and corrected estimation equations, where we consider the main study/validation study setup. We assess the performances of the proposed methods by conducting extensive simulation studies as well as real data analysis. Numerical results confirm the satisfactory performance of the proposed methods in reducing or eliminating the bias induced by covariate misclassification in parameter estimates.

In the fourth project, we aim to expand upon the developments made in the third project by considering the scenario of compound outcome in conjunction with mixed measurement error and misclassification in covariates. The effectiveness of the correction strategies proposed in this study is evaluated through comprehensive simulation studies and real-world data analysis. The numerical findings indicate that the suggested methods exhibit satisfactory performance by reducing or removing the bias resulting from covariate mismeasurement.

Keywords: Case fatality rate, COVID-19 data, COVID-19 risk factors, dynamic treatment regimes, estimating function, incubation time, measurement error, misclassification, optimal preventive policy, precision medicine, Q-learning, regression calibration, regression models, recovery time, survival analysis, text mining, validation data.

Summary for Lay Audience

This thesis explores various statistical and reinforcement learning methods to gain insights into epidemiological characteristics and effective containment measures related to COVID-19. However, caution must be exercised when interpreting the results from such analyses due to the use of error-contaminated data. Consequently, as a complementary remedy, we develop procedures for addressing such complications.

In the first study, we use a dataset, dated from January 22, 2020 to March 29, 2020, to examine epidemiological characteristics of COVID-19. We use survival analysis techniques to quantify how the recovery time may be associated with age and gender. Using data visualization and text mining tools, we study incubation times, fatality rate, as well as most common symptoms. Based on our numerical results, the median incubation time is about 5 days, and the elders are more likely to have longer incubation periods. Furthermore, we find that the median recovery time for infected patients is about 20 days, and there is no gender difference in recovery times.

In the second study, we use data from 175 countries from January 13 of 2020 to March 9 of 2021, and investigate possible factors associated with the case fatality rate of COVID-19. The Q-learning algorithm is employed to assess optimal preventive policies adopted by individual countries to reduce their COVID-19 case fatality rates. The data analysis suggests that, in addition to addressing traditional risk factors, policymakers should tailor the strictness of preventive policies to country-specific characteristics and evolving situation to alleviate the risk of death from COVID-19.

The third study investigates the effects of misclassified covariates in developing dynamic treatment regimes with the Q-learning approach. We present two procedures to account for the bias induced by covariate misclassification. The satisfactory performance of these procedures is demonstrated through extensive simulation studies.

The fourth study deals with mixed measurement error and misclassification in covariates within the context of Q-learning with compound outcome. We demonstrate that the presence of such measurement inaccuracies can pose significant challenges to the accurate estimation process in Q-learning. To address this issue, we propose effective correction strategies that successfully alleviate the impact of mismeasurement.

Acknowledgments

I would like to express my deepest gratitude and appreciation to all those who have contributed to successfully completing my Ph.D. journey. Without their unwavering support, guidance, and encouragement, this accomplishment would not have been possible.

First and foremost, I extend my heartfelt appreciation to my esteemed supervisor, Dr. Grace Yun Yi, for her invaluable guidance, unwavering support, and insightful feedback throughout this research endeavor. Her expertise, patience, and dedication have been instrumental in shaping my academic and research pursuits. I am truly grateful for the knowledge and wisdom she has imparted to me, which has undoubtedly shaped the trajectory of my academic career.

I thank the thesis examiners, Dr. Bei Jiang, Dr. Hyukjun (Jay) Gweon, Dr. Simon Bonner, and Dr. Renfang Tian, for their feedback on my thesis.

I extend my sincere appreciation to my beloved parents, Taha Khadem Charvadeh and Elaheh Shahmari Ardehjani, for their unwavering support, love, and encouragement throughout my academic journey. Their constant belief in my abilities, sacrifices, and unwavering faith have been a constant source of inspiration for me. I am deeply grateful for their unconditional love and for instilling in me the values of perseverance and dedication.

I would also like to express my heartfelt gratitude to my brothers, Saeed and Vahid, and my lovely sister Sara. Their unwavering support, encouragement, and belief in my abilities have been a source of strength during challenging times. Their motivation and faith in my potential have continually inspired me to strive for excellence. I am truly grateful for their presence in my life and the invaluable support they have provided throughout my academic journey.

I would like to extend my gratitude to the faculty members, and members of the GW-DSRG (Grace-Wenqing Data Science Research Group), who have provided their guidance, support, and intellectual stimulation throughout my academic journey. Their insightful discussions, constructive feedback, and collaborations have been instrumental in shaping my research ideas and expanding my knowledge horizon.

Finally, I am deeply grateful to all my friends and well-wishers who have supported me with their encouragement, understanding, and positivity throughout this challenging yet rewarding journey. Their presence and camaraderie have made the research process more enjoyable and memorable.

Contents

List of Figures

List of Tables

Chapter 1

Introduction and Literature Review

SARS-CoV-2 (Lai et al., 2020) is a member of the coronaviruses family, which causes a transmittable infectious respiratory disease known as COVID-19. The novel virus was first reported in December 2019 in the city of Wuhan, China (Zhang et al., 2020). On March 11, 2020, the World Health Organization (WHO) upgraded the status of the COVID-19 outbreak from epidemic to a global pandemic, and now almost all countries have reported confirmed cases, with the USA and India having the highest numbers of confirmed cases (Worldometers, 2022). As of November 17, 2022, the WHO reported 633,263,617 confirmed cases with 6,594,491 deaths (WHO, 2022).

This research is motivated by data on public health, social and economic aspects of COVID-19. Analysis of different aspects of COVID-19-related issues requires subject-specific statistical methods. Therefore, the aim of this chapter is to provide brief descriptions of some statistical methods employed in the proceeding chapters.

1.1 Survival Analysis

Survival analysis is a collection of statistical methods primarily concerned with the modeling of survival time (lifetime) or the length of time until the occurrence of some specific event. Such data arise in many areas, including epidemiological, social, engineering, and reliability studies. There are basically three requirements for time-to-event data: (A) a well-defined origin of measure to the occurrence of the event of interest, (B) a scale for the measurement, and (C)

a precise definition of the occurrence of the event of interest (Lawless, 2003).

1.1.1 Notation and Notion

Let T be a non-negative random variable representing the time to a specific event occurrence associated with individuals in the population, and let the small letter t denote any realization for the variable T. Let $f(t)$ denote the probability density function of T and let $F(t)$ denote the cumulative distribution function, defined as

$$F(t) = P(T \leq t) = \int_0^t f(x)dx \quad \text{for} \quad t \geq 0.$$

The survivor function, denoted $S(t)$, gives the probability of an individual surviving beyond time t, given by

$$S(t) = 1 - F(t) = P(T \geq t).$$

The hazard function, denoted $h(t)$, gives the instantaneous rate of failure or death at time t, given that the individual has survived up to time t. That is, in notation

$$h(t) = \lim_{\Delta t \to 0} \frac{P(t \leq T < t + \Delta t \mid T \geq t)}{\Delta t} = \frac{f(t)}{S(t)},$$

where Δt denotes a small time magnitude. It is clear that the values of the hazard function range from zero to infinity.

Censoring

A common feature in survival analysis is censoring. Censoring is said to occur when there is some information about an individual event time, but the exact event time cannot be observed. Ignoring censored observations or treating them as if they were uncensored measurements can lead to a considerable degree of bias in results. Censoring can be due to termination of follow-up, death, withdrawal, migration, etc. There are different types of censoring, such as "right censoring", "left censoring", and "interval censoring" among which right censoring has received the most attention in the literature (Lawless, 2003).

Right censoring occurs when the survival time is "incomplete" at the right side of the follow-up period. Suppose that C is a censoring time for an individual (e.g., the end time of the study). If $T > C$, then we say the subject is right censored. Often, we introduce the censoring indicator, $\delta = I(T > C)$, to indicate whether the observed time for a study subject is censoring time or survival time, where $I(\cdot)$ is the indicator function.

A lifetime is said to be left-censored at time C if the exact event time is not observed but is known to be less than time C. For example, if the event of interest has already happened before the individual is enrolled in the study and the exact event time is unknown, then the enrollment time is the left censoring time (i.e., C). In the case of left censoring, the observed time is $t = \max(T, C)$.

Interval censoring occurs when the event of interest is known to occur within a known time interval of non-zero length. That is, the observed data consists of $(L, U]$ with the information that $L < T \leq U$. Interval censoring can be thought of as a generalization of left and right censoring. To be specific, interval censoring yields left-censoring if $L = 0$ and U is a known upper bound on the true survival time. On the other hand, interval censoring gives right-censoring if L is a known lower bound on the true survival time and $U = \infty$.

Truncation

Truncation is defined as a condition that excludes individuals from the study population if their event times are smaller or larger than certain values. There are two types of truncation, namely "left truncation" and "right truncation". Left truncation, also called delayed entry, occurs if individuals survive past a certain time u, say, the entry of the study. This means that individual event time $t > u$. In the presence of left truncation, the observed data for an individual consists of $\{u, t, \delta\}$, where $t > u$ and t is an event time or censoring time. On the other hand, time-to-event data are right-truncated if only individuals who have experienced the event of interest by a certain time can be included in the study.

Incubation Period

The incubation period is the time elapsed between exposure to a disease-causing organism and symptom onset (Wagner et al., 2006). Let S denote the symptom onset time, and let E_L and E_U denote the lower and upper time points for the exposure period, respectively, where the exposure period pertains to the duration within which an individual comes into contact with a pathogen. Define $t_L \triangleq S - E_U$ and $t_U \triangleq S - E_L$. Then the incubation time is between t_L and t_U, shown in Figure 1.1. A good understanding of the incubation period of an infectious disease is critical because it can provide information about when infected individuals will be symptomatic and, therefore, infectious. The incubation period also offers insights into the decision-making process around the control of infectious diseases (e.g., by determining a sensible quarantine time).

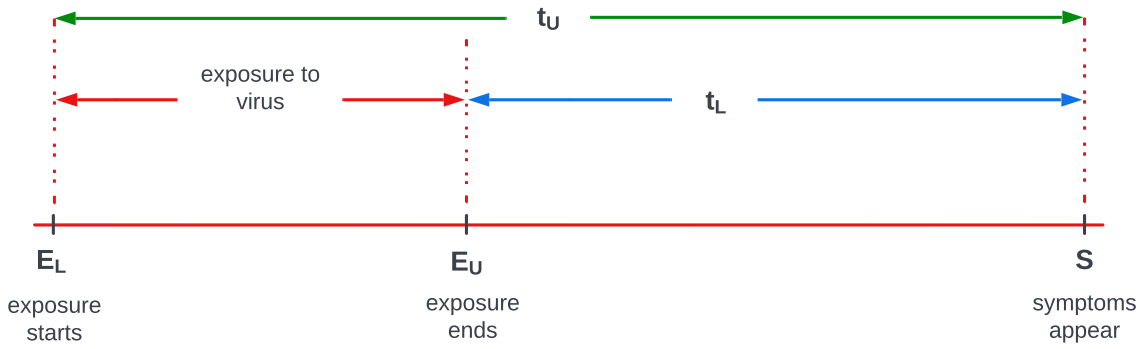

Figure 1.1: Visualization of the incubation period. E_L: Lower time point of exposure period; E_U: Upper time point of exposure period; S: Symptoms onset; $t_L = S - E_U$; $t_U = S - E_L$.

1.1.2 Useful Methods and Models

Kaplan-Meier Method

The Kaplan-Meier (KM) method is a nonparametric method used to estimate survival function from data that are censored or truncated. In what follows, we describe the procedure for

computing KM probabilities of survival, given survival time and censoring status. Let n represent the size of a random sample, and let i denote the index for a subject in the sample. Suppose we have the observed data $\Big\{\{t_i, \delta_i\} : i = 1, \cdots, n\Big\}$, where δ_i is the censoring indicator for subject i with $\delta_i = 1$ (or $\delta_i = 0$) indicating the observed value t_i to be the event (or censoring) time. We order distinct event times to be $t'_1 < t'_2 < \cdots < t'_k$, where $k \leq n$.

For $j = 1, \cdots, k$, let $d_j = \sum_{i=1}^{n} I(t_i = t'_j, \delta_i = 1)$ represent the number of events occurring at time t'_j, and let $n_j = \sum_{i=1}^{n} I(t_i \geq t'_j)$ denote the number of individuals at risk prior to time t'_j. Then, the KM estimator (Lawless, 2003), also known as the product limit estimator, of $S(t)$ is expressed as

$$\hat{S}(t) = \prod_{j:\ t'_j < t} \hat{P}\Big(T > t'_j \mid T \geq t'_j\Big) = \prod_{j:\ t'_j < t} \frac{n_j - d_j}{n_j}.$$

Cox Proportional Hazards Model

The Cox proportional hazards (PH) model (Cox, 1972) is a semi-parametric model used to study relationships between time-to-event outcomes and risk factors or exposures. The Cox PH regression model is usually written in terms of the hazard function:

$$h(t, X) = h_0(t) \exp\left\{\sum_{j=1}^{p} \beta_j X_j\right\}, \tag{1.1}$$

where $X \triangleq (X_1, \cdots, X_p)^{\mathrm{T}}$ is a p-dimensional vector of explanatory variables, $h_0(t)$ is the baseline hazard function, and the β_j are regression coefficients associated with X_j. Here the explanatory variables X_j's are time-independent. With time-dependent explanatory variables, the Cox PH model form can still be used and is called the extended Cox model.

The Cox PH model is popular because of its attractive characteristics. Estimation of the covariate effects β_j may be performed by leaving the baseline hazard function, $h_0(t)$, unspecified, and thus, gaining robust estimation results. To be specific, let $\beta = (\beta_1, \cdots, \beta_p)^{\mathrm{T}}$ denote the vector of parameters. Estimation of β can be performed by maximizing the partial likelihood (Lawless, 2003).

Suppose we have observations $\{t_i, \delta_i, X_i\}$ for $i = 1, \cdots, n$, and suppose there is no tie at each event time. The partial likelihood, denoted $\mathcal{L}(\beta)$, for the Cox PH model is given by

$$\mathcal{L}(\beta) = \prod_{i=1}^{n} \left\{ \frac{h_0(t_i) \exp\left(\beta^{\mathrm{T}} X_i\right)}{\sum_{j \in R_i} h_0(t_i) \exp\left(\beta^{\mathrm{T}} X_j\right)} \right\}^{\delta_i}$$
$$= \prod_{i=1}^{n} \left\{ \frac{\exp\left(\beta^{\mathrm{T}} X_i\right)}{\sum_{j \in R_i} \exp\left(\beta^{\mathrm{T}} X_j\right)} \right\}^{\delta_i},$$

where $R_i = \{l : T_l \geq t_i\}$ denotes the risk set.

Accelerated Failure Time Model

The accelerated failure time (AFT) model is a parametric regression model commonly used in survival analysis. The AFT model assumes that the effect of a covariate is to accelerate or decelerate the life course of a disease by some constant. Unlike its counterpart, the Cox PH model, it allows us to directly model the relationship between $\log T$ and covariates, given by

$$\log T = u + \beta^{\mathrm{T}} X + \epsilon, \tag{1.2}$$

where u is the intercept term, β is the vector of regression coefficients, X is a vector of covariates, and ϵ is the error term that is independent of X and has a given distribution.

The AFT model (1.2) describes a general class of models where the distribution of ϵ can assume different forms. The AFT model facilitates survival times directly and thus enables us to describe survival curves. The AFT model is transparent in the interpretation of survival curves that can be used to construct confidence intervals, either for survival times or parameters. Goodness-of-fit may be performed under the AFT model with the use of different forms of residuals (Balakrishnan et al., 2013).

Further, assuming zero mean for the error term in (1.2) but leaving its distribution unspecified, one may develop a semi-parametric AFT model (e.g., Jin et al., 2006; He and Yi, 2020).

1.2 Dynamic Treatment Regimes

1.2.1 A Literature Review

Precision medicine, also referred to as personalized medicine, is an emerging practice of medicine that helps produce targeted therapies for patients with chronic diseases and reduces the cost of healthcare by reducing the quantity, duration, frequency, or scope of unnecessary treatments. The primary motivation for this practice stems from the concerns that the previously widely regarded *acute inpatient care model* with "one-size-fits-all" or *static* treatment strategy, which does not take into account the heterogeneities and time-varying conditions of patients. Personalized medicine for chronic diseases, however, follows the concept of *chronic care model* (CCM) developed by Wagner et al. (2001) to improve chronic illness care programs. Some of the key features of CCM are "embedding evidence-based guidelines into daily clinical practice", "individualizing the treatment type, dosage and timing based on patient's ongoing conditions", and "using individual patient's case history to change therapeutical strategies in a sequential manner" (Wagner et al., 2001). The latter feature is known as *dynamic* treatment and highlights the heterogeneity of patients' responses to treatments.

In the context of personalized medicine, changing therapeutical interventions in a sequential manner requires personalizing interventions over multiple stages of a patient's recovery trajectory. The implementation of such practice can be done through dynamic treatment regimes (DTRs) (Murphy et al., 2001; Murphy, 2003; Robins, 2004). DTRs, aka adaptive interventions, have become increasingly important in medical studies. Casting the problem under a dynamic framework, the objective of a DTR is to optimize the long-term clinical outcome of an individual with his or her characteristics and medical history taken into account. DTRs allow individualization of treatments, and thus are mostly beneficial for patients with chronic illnesses who need long-term medication. In contrast to the acute inpatient care model, where clinicians recommend a one-size-fits-all treatment strategy based on a priori empirical population information, DTRs are an evidence-based strategy to implement subject-specific treatment or intervention. This highly customized strategy is regarded as a key element of the CCM (Wagner et al., 2001).

Basically, a DTR is a set of decision rules over a sequence of time-intervals or stages, and

each decision rule recommends an effective treatment for an interval based on a patient's individual characteristics and treatments histories (Murphy, 2003,0). DTRs facilitate therapeutic intervention that evolves over time, which is mainly dedicated by the patient's characteristics and other relevant information. They can be particularly useful for accounting for patients' heterogeneities and time-varying needs. In the development of the optimal DTR, the objective of each decision rule is not to optimize the outcome of a patient in any specific stage, but to optimize the patient's long-term outcome.

Many methods have been developed for estimating the optimal DTRs. These methods can be broadly classified as either *direct* or *indirect* estimation methods. Direct estimation methods, also known as policy search methods, establish the optimal sequence of treatment rules by optimizing the expected outcome over a set of candidate treatments. Popular methods under this category include *inverse probability of treatment weighting* (IPTW) approaches, *marginal structural models*, and *classification-based methods* (Chakraborty and Moodie, 2013). In contrast, indirect methods use approximate dynamic programming to describe models for the stage-specific conditional mean outcome, and then optimize the resulting approximate conditional mean outcome to find the optimal DTRs. Popular indirect methods are Q-learning (Murphy, 2005b; Chakraborty and Moodie, 2013), A-learning (Robins, 2004, pp.189-326), and G-estimation for structural nested models (Robins, 1997, pp.69-117).

1.2.2 Estimating Treatment Rules by Q-Learning

Q-learning, originating from the computer science community (Watkins, 1989), is a well-known method for constructing optimal DTRs. The implementation of the Q-learning method relies on the specification of a function called the Q-function. In the context of DTRs, Q-functions represent the expected cumulative rewards (e.g., clinical effectiveness or patient health) obtained by following specific treatment regimes (policy) in a sequential decision-making problem. The implementation of the Q-learning method is simple, which typically involves two steps: (1) estimate the stage-specific Q-functions, and (2) recommend the actions that optimize the estimated Q-functions (Qian et al., 2012, pp.127-148). These steps basically hinge on the specification of the Q-functions, which can be modeled parametrically,

semi-parametrically, or nonparametrically (Zhao et al., 2009).

In this section, we present the basic notation and necessary steps for constructing optimal treatment regimes using the Q-learning algorithm in a mismeasurement-free context. Consider the setting with K stages of decision rules, where K is a positive integer greater than 1.

For $k = 1, \cdots, K$, let A_k denote the binary treatment indicator received at the kth stage for a subject, taking value 0 or 1. Let X_k denote an error-prone binary covariate taking value 0 or 1, and let Z_k denote a vector of precisely measured covariates where X_k and Z_k are measured prior to the treatment receipt at the beginning of the kth stage. For $k = 1, \cdots, K$, let $\overline{X}_k = \left\{X_1, \cdots, X_k\right\}$, $\overline{Z}_k = \left\{Z_1, \cdots, Z_k\right\}$, and $\overline{A}_k = \left\{A_1, \cdots, A_k\right\}$.

For $k = 1, \cdots, K$, let Y_k denote the observed outcome at the end of stages k, which is assumed "the bigger the better" and is regarded as a function, say $g(\cdot)$, of the history of the treatment, $\overline{A}_{k-1}$, together with the current treatment A_k, and the history of the covariates, $\overline{X}_k$ and $\overline{Z}_k$, as well as the covariates X_{k+1} and Z_{k+1} at the next stage. That is,

$$Y_k = g(\overline{A}_k, \overline{X}_{k+1}, \overline{Z}_{k+1}) \tag{1.3}$$

for $k = 1, \cdots, K$ (Chakraborty and Moodie, 2013, p.35), where X_{K+1} and Z_{K+1} are null.

The objective is to find a sequence of optimal treatments by retrospectively maximizing the expected value of the outcome for each stage in a *backward* manner (i.e., from stage K back to stage 1), with the impacts of the history of treatments and the covariates incorporated. This idea is realized by the Q-learning method, which is rooted in the use of Q-functions defined for K stages (Chakraborty and Moodie, 2013, Section 3.4.1).

To be specific, set $Q_{K+1} = 0$. Define

$$Q_K(\overline{A}_K, \overline{X}_K, \overline{Z}_K) = E(Y_K \mid \overline{A}_K, \overline{X}_K, \overline{Z}_K), \tag{1.4}$$

and for stage k with $k = K - 1, \cdots, 1$, define the stage k Q-function recursively in a backward direction:

$$Q_k(\overline{A}_k, \overline{X}_k, \overline{Z}_k) = E\left\{Y_k + \max_{a_{k+1}} Q_{k+1}(\overline{A}_k, \overline{X}_{k+1}, \overline{Z}_{k+1}, a_{k+1}) \,\middle|\, \overline{A}_k, \overline{X}_k, \overline{Z}_k\right\}, \tag{1.5}$$

where $\overline{A}_{k+1}$ in $Q_{k+1}(\overline{A}_{k+1}, \overline{X}_{k+1}, \overline{Z}_{k+1})$ in the right-hand side of (1.5) is written as $\overline{A}_k \cup \{a_{k+1}\}$ so that the future treatment a_{k+1} for stage k can be evaluated to find the optimal one.

In other words, for $k = K-1, \cdots, 1$, let

$$\hat{Y}_k \triangleq Y_k + \max_{a_{k+1}} Q_{k+1}(\overline{A}_k, \overline{X}_{k+1}, \overline{Z}_{k+1}, a_{k+1}) \tag{1.6}$$

denote the stage k pseudo-outcome constructed in a way that assumes the subject is given the best treatment at the subsequent stage. The additive structure between Y_k and the maximized value of $Q_{k+1}(\overline{A}_k, \overline{X}_{k+1}, \overline{Z}_{k+1}, a_{k+1})$, related to the Bellman equation (Bellman, 2010), reflects the belief that the treatments over the multiple stages affect the outcomes cumulatively in an additive manner. Then the stage k Q-function is defined as

$$Q_k(\overline{A}_k, \overline{X}_k, \overline{Z}_k) = E\{\hat{Y}_k \mid \overline{A}_k, \overline{X}_k, \overline{Z}_k\}, \qquad \text{for} \qquad k = K-1, \cdots, 1, \tag{1.7}$$

together with (1.4).

To find optimal treatments, a natural approach is to use a backward procedure by examining the $Q_k(\overline{A}_k, \overline{X}_k, \overline{Z}_k)$ functions for $k = K, \cdots, 1$. In the case where the true Q-functions are known, optimal treatments, denoted d_k, are determined by:

$$d_k = \arg\max_{a_k} Q_k(\overline{A}_{k-1}, \overline{X}_k, \overline{Z}_k, a_k) \qquad \text{for} \qquad k = K, \cdots, 1. \tag{1.8}$$

When the true Q-functions are unknown, as is often the case in applications, we employ regression models, especially linear regression models, to delineate them as they represent conditional expectations. To be specific, for $k = K, \cdots, 1$, consider the linear regression model

$$Q_k(\overline{A}_k, \overline{X}_k, \overline{Z}_k) = \beta_k^{\mathrm{T}} H_{k0} + (\psi_k^{\mathrm{T}} H_{k1}) A_k, \tag{1.9}$$

where we separate the treatment effects from those of covariates and express $\overline{A}_{k-1} \cup \overline{X}_k \cup \overline{Z}_k$ as $\{H_{k0}, H_{k1}\}$, with H_{k0} representing the covariates that have a predictive effect on the outcome, and with H_{k1} standing for the covariates that interact with treatment; H_{k0} and H_{k1} may include a constant, or intercept, term, and they may include the same covariates. Here β_k and ψ_k are

the associated regression coefficients (Chakraborty and Moodie, 2013, p.40), and we write $\theta_k = (\beta_k^{\mathrm{T}}, \psi_k^{\mathrm{T}})^{\mathrm{T}}$ for $k = 1, \cdots, K$.

Suppose we have a random sample $\mathcal{D} \triangleq \left\{\{A_{ki}, X_{ki}, Z_{ki}, Y_{ki}\} : k = 1, \cdots, K;\ i = 1, \cdots, n\right\}$ with the $\{A_{ki}, X_{ki}, Z_{ki}, Y_{ki}\}$ for $i = 1, \cdots, n$ being independent and identically distributed (i.i.d.) following the same distribution of $\{A_k, X_k, Z_k, Y_k\}$. In the following development, index i is added to $\overline{X}_k$, $\overline{Z}_k$, and $\overline{A}_k$ in a similar manner.

Now we describe a procedure for finding optimal treatments, called the Q-learning algorithm, using the data in the random sample. The algorithm is basically a backward recursion process that aims to minimize the mean squared error between the (pseudo-) outcome and its conditional expectation at each stage, starting from the last stage backward to the first stage. At the last stage K, the minimization is essentially the least squares method which minimizes the squared difference between the observed outcome Y_K and its conditional expectation (1.4); for stage k with $k = K-1, \cdots, 1$, the least squares method is individually applied to conceptually minimize the squared difference between the pseudo-outcome $\hat{Y}_k$ in (1.6) and its conditional expectation (1.7).

To be specific, with $Q_{K+1} \triangleq 0$, we estimate the regression coefficients for each stage via the least squares approach:

$$\hat{\theta}_k = \arg\min_{\theta_k} \frac{1}{n} \sum_{i=1}^{n} \left[\hat{Y}_{ki} - Q_k(\overline{A}_{ki}, \overline{X}_{ki}, \overline{Z}_{ki}; \theta_k)\right]^2 \tag{1.10}$$

for $k = K, \cdots, 1$, where $\hat{Y}_{Ki} = Y_{Ki}$; for $k = K-1, \cdots, 1$,

$$\hat{Y}_{ki} = Y_{ki} + \max_{a_{k+1}} Q_{k+1}(\overline{A}_{ki}, \overline{X}_{(k+1)i}, \overline{Z}_{(k+1)i}, a_{k+1}; \hat{\theta}_{k+1}), \tag{1.11}$$

representing stage k pseudo-outcome for subject i; and $Q_k(\overline{A}_{ki}, \overline{X}_{ki}, \overline{Z}_{ki}; \theta_k)$ is determined by regression model such as (1.9).

Substituting the estimates of (1.10) into (1.8), we estimate the optimal treatment by

$$\hat{d}_k = \arg\max_{a_k} Q_k(\overline{A}_{k-1}, \overline{X}_k, \overline{Z}_k, a_k; \hat{\theta}_k) \qquad \text{for} \qquad k = K, \cdots, 1, \tag{1.12}$$

where $Q_k(\overline{A}_{k-1}, \overline{X}_k, \overline{Z}_k, a_k; \hat{\theta}_k)$ is determined by (1.9) with θ_k replaced by its estimate $\hat{\theta}_k$, deter-

mined by (1.10).

For $K = 2$, the implementation of (1.10) under model (1.9) can be carried out using the *R* package ***qLearn***. The *R* package, ***DTRreg***, initially developed for the G-estimation and dWOLS methods, can also be used for Q-learning with any K.

We conclude this section with comments on inference about the model parameters using the estimators derived from (1.10). While the least squares method is applied to all $k = K, \cdots, 1$ in (1.10), the interpretation of the resulting estimators differ for $k = K$ and $k < K$. When $k = K$, the resulting estimator $\hat{\theta}_K$ is the least squares estimator of θ_K, and thus inference about θ_K can be carried out in a usual way. For instance, coverage rates (CRs) of confidence intervals (CIs) for the Kth stage estimators in (1.10) can be calculated using either Wald-type (W-type) CIs or percentile bootstrap (PB) CIs. However, caution should be exercised when obtaining CIs for the estimators in stage k with $k < K$, as those estimators $\hat{\theta}_k$ with $k = K-1, \cdots, 1$ are not the least squares estimators of the θ_k due to the unobservable feature of pseudo-outcomes $\hat{Y}_{ki}$. For $k = K-1, \cdots, 1$, the pseudo-outcome, $\hat{Y}_{ki}$, may be a non-smooth function of $\hat{\theta}_{k+1}$, which creates the *non-regularity* or *weak non-regularity* issue to be discussed in Section 1.2.3.

1.2.3 Inference in the Presence of Non-regularity

In implementing (1.12), the non-regularity issue for estimators may arise when optimal treatments at subsequent stages are not unique for a portion of subjects in the population; this happens when the coefficient of the treatment is equal to zero with a positive probability. Furthermore, weak non-regularity occurs if the coefficient of the treatment assumes values near zero (Robins, 2004). Non-regularity or weak non-regularity often distorts usual inferential procedures that root in the asymptotic normal distribution. In applications, it is generally advised to examine analysis results for any severe non-regularity issue. A few approaches have been proposed to address the non-regularity or weak non-regularity problem. One strategy to circumvent non-regularity is to employ the soft-thresholding or hard-thresholding approach to modify the pseudo-outcome in the Q-learning algorithm to regularise the non-regular estimators (Chakraborty et al., 2010).

An alternative remedy is to report double bootstrap (DB) or m-out-of-n bootstrap CIs

(Chakraborty and Moodie, 2013, Chapter 8) when doing inference about the model parameters. DB CIs are of particular interest in cases where asymptotically pivotal statistics are not available (Nankervis, 2005). To be specific, let ϕ denote a parameter of interest, and let $\hat{\phi}$ denote its least squares estimate. For $0 < \alpha < 1$, we construct a $(1-\alpha)100\%$ DB CI for ϕ by running the following steps, with M_1 and M_2 being given positive integers:

1. Draw M_1 first-stage bootstrap samples from the original data. For $b = 1, \cdots, M_1$, use the bth bootstrap sample, denoted S_b, to obtain an estimate of ϕ, denoted $\hat{\phi}^{*b}$.

2. For $b = 1, \cdots, M_1$, draw M_2 second-stage bootstrap samples from the first-stage bootstrap sample S_b. For $m = 1, \cdots, M_2$, use the mth bootstrap sample, denoted S_{bm}, to obtain an estimate of ϕ denoted by $\hat{\phi}^{**bm}$.

3. For $b = 1, \cdots, M_1$, calculate $u^{*b} = \frac{1}{M_2} \sum_{m=1}^{M_2} \mathbb{1}\{\hat{\phi}^{**bm} \leq \hat{\phi}\}$, where $\mathbb{1}(\cdot)$ is the indicator function.

4. Let $u^*_{(\frac{\alpha}{2})}$ and $u^*_{(1-\frac{\alpha}{2})}$ denote the $(\frac{\alpha}{2})$-percentile and $(1-\frac{\alpha}{2})$-percentile of $\left\{u^{*b} : b = 1, \cdots, M_1\right\}$, respectively. Find the $u^*_{(\frac{\alpha}{2})}$-percentile and $u^*_{(1-\frac{\alpha}{2})}$-percentile of $\left\{\hat{\phi}^{*b} : b = 1, \cdots, M_1\right\}$, denoted $\hat{\phi}^*_{u^*_{(\frac{\alpha}{2})}}$ and $\hat{\phi}^*_{u^*_{(1-\frac{\alpha}{2})}}$, respectively. Then a $(1-\alpha)100\%$ DB CI for ϕ is given by $\left(\hat{\phi}^*_{u^*_{(\frac{\alpha}{2})}}, \hat{\phi}^*_{u^*_{(1-\frac{\alpha}{2})}}\right)$.

While larger M_1 and M_2 are expected to give better results, the choice of M_1 and M_2 is usually driven by the trade-off of the computational burden and accuracy of the results.

1.3 Optimization Procedures

1.3.1 Constrained Optimization

Optimization problems are generally categorized into two broad categories: convex and nonconvex optimization. Convex optimization is further categorized into linear and nonlinear programming, while nonconvex optimization is categorized into discrete and continuous optimization (e.g., Lin et al., 2012). The review in this section is kept general in nature. To start with, we present the following optimization problem:

$$\begin{aligned} \text{minimize} \quad & f(x) \\ \text{subject to} \quad & f_i(x) \le b_i, \quad i = 1, \cdots, m, \\ & h_j(x) = c_j, \quad j = 1, \cdots, p, \\ & x_{iL} \le x_i \le x_{iU} \quad i = 1, \cdots, n, \end{aligned} \tag{1.13}$$

where the vector $x = (x_1, \cdots, x_n)^{\mathrm{T}} \in \mathbb{R}^n$, with coordinate x_i bounded by given constants x_{iL} and x_{iU} for $i = 1, \cdots, n$, called the design variables; $f : \mathbb{R}^n \to \mathbb{R}$ is the objective (or goal) function; $f_i : \mathbb{R}^n \to \mathbb{R}$ is an inequality constraint function for $i = 1, \cdots, m$; $h_j : \mathbb{R}^n \to \mathbb{R}$ is an equality constraint function for $j = 1, \cdots, p$; b_i is a given constant for $i = 1, \cdots, m$; and c_j is a given constant for $j = 1, \cdots, p$.

Optimization techniques are used to find the optimal vector, denoted $x^\star$, that has the smallest objective value among all vectors that satisfy the constraints in (1.13). That is, for any z with $f_1(z) \le b_1, \cdots, f_m(z) \le b_m$ and $h_1(z) = c_1, \cdots, h_p(z) = c_p$, we have $f(z) \ge f(x^\star)$.

If the objective and constraint functions in the optimization problem (1.13) are all linear functions, the optimization problem (1.13) is called a linear programming, and otherwise, a nonlinear programming.

If the equality constraints are affine, and the objective function, as well as inequality constraint functions, are convex, i.e., they respectively satisfy the inequalities

$$f(\alpha x + \beta y) \le \alpha f(x) + \beta f(y)$$

and

$$f_i(\alpha x + \beta y) \le \alpha f_i(x) + \beta f_i(y) \quad \text{with} \quad i = 1, \cdots, m,$$

for all $x, y \in \mathbb{R}^n$ and all $\alpha, \beta \in \mathbb{R}$ with $\alpha + \beta = 1$, $\alpha \ge 0, \beta \ge 0$, the optimization problem (1.13) is called convex optimization problem.

It is noted that the variables x in (1.13) can also be of discrete or integer nature. Optimization problems with discrete variables are referred to as discrete optimization problems. Discrete optimization problems are generally approached using global optimization algorithms rather than local optimization algorithms. In what follows, we briefly review some local and global optimization algorithms.

1.3.2 Nature of Optimization Algorithms

Local Optimization Algorithms

Local optimization algorithms are mostly gradient-based, and they result in a locally optimal solution in the vicinity of which we cannot find other feasible solutions with better objective function values. In convex optimization problems, a locally optimal solution is also globally optimal. That is, for linear programming problems, quadratic programming problems with a positive or negative definite objective function, and nonlinear programming problems with convex or concave objective and constraint functions, a locally optimal solution is also globally optimal (Boyd and Vandenberghe, 2004).

Global Optimization Algorithms

An objective function may have multiple optima, and an optimization problem with such an objective function is referred to as a multimodal optimization problem. Some of the solutions to a multimodal optimization problem can be globally optimal solutions having identical objective function value, and some can be locally optimal solutions having different objective function value. Although one can still use local optimization algorithms in conjunction with a multi-start approach to deal with such problems, it is more practical to employ global optimization algorithms that focus on finding the best possible solution in the entire search space. Global optimization algorithms generally require more time, but they are more likely to result in more reliable solutions (Horst and Tuy, 2013).

1.3.3 Useful Optimization Algorithms

Newton Raphson Algorithm

The Newton Raphson algorithm (Avriel, 2003) is an iterative gradient-based optimization algorithm without constraints, i.e., no constraints in (1.13). The Newton Raphson algorithm

basically consists of two general steps:

- **Step 1**: Use the second-order Taylor series expansion to construct a quadratic approximation of the objective function around some initial design point x^0.
- **Step 2**: Adjust the design point to that which maximizes the quadratic approximation.

Iterate through these steps until the design point stabilizes. To clarify, consider the following second-order Taylor series expansion of the objective function $f(x)$

$$f(x) \approx f\left(x^0\right) + \nabla f\left(x^0\right)^{\mathrm{T}}\left(x - x^0\right) + \frac{1}{2}\left(x - x^0\right)^{\mathrm{T}} H\left(x^0\right)\left(x - x^0\right), \tag{1.14}$$

where $\nabla f\left(x^0\right)$ is the gradient, and $H\left(x^0\right)$ is the Hessian matrix. Differentiating (1.14) with respect to x and setting the result equal to zero results in the following update formula for the current design point:

$$x = x^0 - H\left(x^0\right)^{-1} \nabla f\left(x^0\right).$$

One of the downsides to the Newton Raphson algorithm is that calculating both the Hessian matrix and its inversion is computationally expensive, especially when dimensions get large. To overcome this issue, other optimization methods are available, as indicated by the following methods.

BFGS Method

The Broyden–Fletcher–Goldfarb–Shanno (BFGS) algorithm (Shanno, 1970) is a second-order optimization algorithm and is usually referred to as the most popular quasi-Newton algorithm. It is known as a quasi-Newton method since it approximates the inverse of the Hessian matrix using the gradient. This means that, unlike the Newton Raphson method, which requires the calculation of the inverse of the Hessian matrix, the BFGS method does not need the Hessian and its inverse to be available or calculated precisely for each step of the algorithm. The approximation to the inverse of the Hessian matrix is updated at each iteration using the first-order gradient information from that iteration.

Nelder-Mead Method

The Nelder-Mead method (Nelder and Mead, 1965), also known as the Simplex search method, is a heuristic multidimensional unconstrained optimization. The algorithm does not require the availability of the gradient and therefore is suitable for problems with non-smooth functions. It basically navigates through the search space stochastically to find the best optimal point it encounters during the search. The Nelder-Mead algorithm uses a geometrical shape called a simplex to search over the domain of the optimization problem. The algorithm starts with a randomly-generated simplex and evolves by moving and/or reshaping the simplex at every iteration until a desired bound is obtained to result in the most optimal objective value.

Genetic Algorithm

The genetic algorithm (De Jong, 1988) is a stochastic search that uses the concept of survival of the fittest to approximate the solutions for a problem. The algorithm is suitable for solving various optimization problems, including problems in which the objective function is nonlinear, discontinuous, or non-differentiable. The genetic algorithm starts with a population of all possible solutions, called "generation zero". The next step is to score the population by evaluating how good each solution in the generation zero population is. Based on the resulting scores, some solutions are selected for reproduction (also called parents). Crossover rules combine two parents to form new solutions for the next generation. This process continues until the score of the best solution stabilizes and does not change for many generations.

1.4 Estimation Equations Approach

Suppose $\{Y_1, \cdots, Y_n\}$ consists of i.i.d. random variables, and let $\{y_1, \cdots, y_n\}$ be their realized values. Suppose the i.i.d. random variables are drawn from the density $f(y;\theta)$, where the finite vector $\theta = (\theta_1, \cdots, \theta_p)^{\mathrm{T}} \subseteq \mathbb{R}^p$. When interest lies in learning about the unknown θ, one may make use of a parametric model with the full specification of the joint distribution $f(y;\theta)$. In such a case, inference is basically done using the *maximum likelihood* method,

which is a centerpiece of statistical inference (e.g., Lehmann and Casella, 2006). Given the data $\{y_1, \cdots, y_n\}$, the likelihood function is defined as

$$L(\theta) = \prod_{i=1}^{n} f(y_i; \theta).$$

With the given data, maximizing $L(\theta)$ with respect to θ gives the maximum likelihood estimator of θ.

Formulating a likelihood function requires correct specification of the distribution form, which may not always be possible in applications. To get around this problem, we can employ estimation function (Godambe, 1991) approaches as an alternative to estimate θ.

The estimating function approach starts with specifying a $p \times 1$ vector of functions which involves both the unknown parameter θ and y_i. Let $S(\theta; y_i)$ denote such a vector-valued function.

An estimate for θ, denoted $\hat{\theta}$, can be obtained by solving the equations

$$\sum_{i=1}^{n} S(\theta; y_i) = 0. \tag{1.15}$$

Here $S(\theta; y_i)$ is called an estimating function for θ, and (1.15) is referred to as estimating equations. To ensure the resulting estimator $\hat{\theta}$ to be consistent and asymptotically normal, we need to impose certain conditions on $S(\theta; y_i)$. Typically, $S(\theta; y_i)$ is required to satisfy

$$E\Big\{S(\theta; Y_i)\Big\} = 0 \tag{1.16}$$

together with other conditions. A function satisfying (1.16) is called an *unbiased estimating function*.

Define

$$I(\theta) = E\Big\{\frac{\partial S(\theta; Y_i)}{\partial \theta^{\mathrm{T}}}\Big\}, \quad J(\theta) = E\Big[S(\theta; Y_i)\{S(\theta; Y_i)\}^{\mathrm{T}}\Big], \tag{1.17}$$

and

$$\Sigma(\theta) = \Big\{I(\theta)\Big\}^{-1} J(\theta) \Big\{I(\theta)\Big\}^{-1\mathrm{T}},$$

where the expectations are taken with respect to $f(y_i; \theta)$ and the inverse matrices are assumed

to exist. Under the regularity conditions, $\sqrt{n}(\hat{\theta} - \theta)$ is asymptotically normal with mean zero and asymptotic covariance matrix $\Sigma(\theta)$ (Yi, 2017, Section 1.3).

The estimating equation method outlined here has been commonly used to conduct inference about model parameter θ. The usefulness of this approach is mainly pertinent to its nice feature of not requiring the full distribution of the associated variables, yet yielding consistent estimators which have asymptotic normal distributions as long as certain regularity conditions are satisfied. In applications, different sets of regularity conditions may be imposed when using the estimating equation method to handle different problems, and those regularity conditions are usually only sufficient but not necessary conditions to ensure consistency of $\hat{\theta}$ and asymptotic normality of $\sqrt{n}(\hat{\theta} - \theta)$.

For example, assuming θ is a scalar, the following regularity conditions were considered by Godambe (1960):

(i) $E\left[S\left(\theta; y_i\right)\right] = \int S(\theta; y_i) f(y_i; \theta) dy_i = 0$ for all $\theta \in \Theta$, where Θ is the parameter space that is open;

(ii) $\frac{dS(\theta; y_i)}{d\theta}$ exists for all $\theta \in \Theta$;

(iii) $\int S(\theta; y_i) f(y_i; \theta) dy_i$ is differentiable under the sign of integration,

(iv) $E\left[\frac{dS(\theta; y_i)}{d\theta}\right]^2 > 0$ for all $\theta \in \Theta$;

together with the assumptions for $f(y_i; \theta)$. More discussions of regularity conditions can be found in Yi (2017, Section 1.3) and the references therein.

1.5 Measurement Error and Misclassification

A variable is said to be subject to measurement error or misclassification when a difference exists between the true value of the measurement and the observed measurement. The term "measurement error" is mostly used for continuous error-prone variables, while the term "misclassification" is used for error-prone discrete variables. In this thesis, we use the term "mismeasurement" to refer to any setting where a measured quantity and its true value may be

different. Mismeasurement can be classified as "random error" and "systematic error". Random errors are regarded as a chance difference between a measured quantity and its true value. These types of errors are often described with a distribution, whereas systematic errors are referred to those that are featured by a constant, thus removable by calibrating the measuring instruments or procedures.

In the following sections, we briefly touch on some of the most important mismeasurement models and mechanisms. To this end, for $i = 1, \cdots, n$, let Y_i denote the error-free response variable, let X_i denote an error-prone explanatory variable whose imprecisely measured surrogate is denoted by X_i^*, and let Z_i denote an error-free explanatory variable.

1.5.1 Measurement Error and Misclassification Models

Measurement error and misclassification models are used to uncover the underlying relationship between the observed variables and true variables. We begin by describing one of the most widely used measurement error models, called the *classical additive error* model.

Classical Additive Error Model

A classical additive error model has the following form

$$X_i^* = X_i + e_i, \tag{1.18}$$

where the error term e_i is independent of the true variable X_i as well as Z_i, and has mean zero.

Berkson Model

The *Berkson model* is of the form

$$X_i = X_i^* + e_i,$$

where e_i is independent of X_i^* as well as Z_i, and has mean zero.

The classical additive error model and the Berkson model differ in the perspective of viewing the relationship between X_i^* and X_i, where one is treated as a dependent variable, and the other is regarded as an independent variable.

Misclassification Model

In settings where X_i is not continuous but discrete, the relationship between X_i and X_i^* can be studied in two ways through modeling: (1) the conditional probability $P(X_i^* = x_i^* \mid X_i = x_i, Z_i)$ or (2) the conditional probability $P(X_i = x_i \mid X_i^* = x_i^*, Z_i)$. The probabilities $P(X_i^* = x_i^* \mid X_i = x_i, Z_i)$ and $P(X_i = x_i \mid X_i^* = x_i^*, Z_i)$ are known as *misclassification probabilities* and *reclassification probabilities*, respectively (Yi, 2017, p.70).

1.5.2 Impacts of Covariate Measurement Error or Misclassification

Consider a random sample of n observations $\{y_i, x_i\}$ of i.i.d. random variables $\{Y_i, X_i\}$. Suppose the relationship between Y_i and X_i is delineated by a simple linear regression of the form

$$Y_i = \beta_0 + \beta_x X_i + \epsilon_i \qquad \text{for} \qquad i = 1, \cdots, n, \tag{1.19}$$

where β_0 and β_x are regression parameters, and ϵ_i is independent of X_i with mean zero and variance σ^2. Suppose that X_i is not observed, but its surrogate version X_i^* is available.

If applying the least squares method with X_i in (1.19) replaced by X_i^*, then the resultant estimator of β_x is given by:

$$\hat{\beta}_x^* = \frac{\sum_i^n (x_i^* - \bar{x}^*)(y_i - \bar{y})}{\sum_i^n (x_i^* - \bar{x}^*)^2},$$

and it can be shown that as $n \to \infty$,

$$\hat{\beta}_x^* \xrightarrow{p} \frac{Cov(X_i^*, Y_i)}{Var(X_i^*)}. \tag{1.20}$$

We now consider two cases. In the first case, we assume that X_i is continuous and linked with X_i^* by the model

$$X_i^* = X_i + e_i,$$

where error e_i has mean zero and variance σ_e^2, and is independent of X_i and ϵ_i. In this case, (1.20) becomes

$$\hat{\beta}_x^* \xrightarrow{p} \mathscr{W}\beta_x \qquad \text{as } n \to \infty,$$

where $\mathscr{W} = \frac{\sigma_x^2}{\sigma_x^2+\sigma_e^2}$, and σ_x^2 is the variance of X_i (Fuller, 1987; Yi, 2017, Section 2.2). Since $\mathscr{W} \leq 1$, the measurement error in X_i has caused an attenuation bias in using $\hat{\beta}_x^*$ to estimate β_x.

Now, we consider the second case where X_i is not a continuous random variable. Let X_i be a binary random variable taking 0 or 1, with $\pi = P(X_i = 1)$. Suppose that X_i^* and X_i are linked by the misclassification matrix

$$\Pi_j = \begin{pmatrix} \pi_{00} & \pi_{01} \\ \pi_{10} & \pi_{11} \end{pmatrix}, \tag{1.21}$$

where $\pi_{kl} = P(X_i^* = k \mid X_i = l)$ for $k = 0, 1$ and $l = 0, 1$, and X_i^* is assumed to be independent of Y_i given X_i.

To calculate $Cov(X_i^*, Y_i)$, we first obtain the expected value of X_i^*:

$$\begin{aligned} E(X_i^*) &= E\Big\{E(X_i^* \mid X_i)\Big\} \\ &= E\Big\{\pi_{11}X_i + (1-\pi_{00})(1-X_i)\Big\} \\ &= \pi_{11}\pi + (1-\pi_{00})(1-\pi). \end{aligned}$$

Let $p \triangleq \pi_{11}\pi + (1-\pi_{00})(1-\pi)$. By the conditional independence of X_i^* and Y_i, given X_i, we obtain the expected value of $X_i^*Y_i$:

$$\begin{aligned} E(X_i^*Y_i) =& E\Big\{E(X_i^*Y_i \mid X_i)\Big\} \\ =& E\Big\{E(X_i^* \mid X_i)E(Y_i \mid X_i)\Big\} \\ =& E\Big\{E(X_i^* \mid X_i)(\beta_0 + \beta_x X_i)\Big\} \\ =& E\Big[\{\pi_{11}X_i + (1-\pi_{00})(1-X_i)\}(\beta_0 + \beta_x X_i)\Big] \\ =& E\Big\{\pi_{11}\beta_0 X_i + (1-\pi_{00})(1-X_i)\beta_0\Big\} + E\Big\{\pi_{11}\beta_x X_i^2 + (1-\pi_{00})\beta_x(X_i - X_i^2)\Big\} \\ =& p\beta_0 + \pi\beta_x\pi_{11}, \end{aligned}$$

where the last step is due to $X_i^2 = X_i$.

Finally, we have

$$\begin{aligned}\frac{Cov(X_i^*, Y_i)}{Var(X_i^*)} &= \frac{E(X_i^* Y_i) - E(X_i^*)E(Y_i)}{Var(X_i^*)} \\ &= \frac{(p\beta_0 + \pi\beta_x\pi_{11}) - p(\beta_0 + \beta_x\pi)}{p(1-p)} \\ &= \frac{\pi(1-\pi)(\pi_{00} + \pi_{11} - 1)}{p(1-p)}\beta_x,\end{aligned}$$

where we use $Var(X_i^*) = p(1-p)$.

Shieh (2009, p.41) showed that $\frac{\pi(1-\pi)(\pi_{00}+\pi_{11}-1)}{p(1-p)} \leq 1$, and thus, (1.20) suggests that covariate misclassification has an attenuated effect on the estimation of covariate effect β_x.

1.5.3 Model Identifiability

Identifiability is a fundamental requirement in statistical modeling. It means that different parameter values give rise to different probability distributions. Conversely, if two sets of parameter values generate identical distributions of the data, the model is not identifiable. Let Θ denote parameter space, and let S_Y denote the sample space of data. A model, $\{F(y;\theta) : \theta \in \Theta\}$, is called *identifiable* if for all $\theta_0, \theta_1 \in \Theta$ and for all $y \in S_Y$:

$$F(y;\theta_0) = F(y;\theta_1) \text{ if and only if } \theta_0 = \theta_1,$$

where $F(y;\cdot)$ is a probability distribution (Guillaume et al., 2019).

Nonidentifiability occurs when a model is poorly specified or when the parameter space is "too large". For example, when the number of the model parameters exceeds the number of observations in the data, the model is nonidentifiable. In the presence of error-prone data, we usually need additional modeling of mismeasurement and/or covariate processes in addition to modeling the response process. This, in particular, results in the parameter space expansion, which typically generates model nonidentifiability issues. One strategy to overcome model nonidentifiability is to collect additional data to help characterize the mismeasurement process. Now, we describe two types of data sources that are used in the analysis of error-contaminated

data.

Validation Subsample

Consider the main study data $\left\{\left\{Y_i, X_i^*, Z_i\right\} : \ i \in \mathcal{M}\right\}$, where $\mathcal{M}$ denote the index set. An internal validation subsample contains true measurements of the variable X_i in addition to $\{Y_i, X_i^*, Z_i\}$ for $i \in \mathcal{V}$, where $\mathcal{V} \subset \mathcal{M}$, i.e., $\left\{\left\{Y_i, X_i, X_i^*, Z_i\right\} : \ i \in \mathcal{V}\right\}$ is available. In contrast an external validation subsample often contain $\left\{\left\{X_i, X_i^*, Z_i\right\} : \ i \in \mathcal{V}\right\}$ with $\mathcal{V} \cap \mathcal{M} = \emptyset$.

Repeated Measurements

In some settings, we observe repeated surrogate measurements for X_i. Such data are called replicate data and have the form

$$\left\{\left\{Y_i, X_{ij}^*, Z_i)\right\} : \ i = 1, \cdots, n; \ j = 1, \cdots, n_i\right\},$$

where n_i is the number of replicate surrogate measurements for the ith subject. Usually, one would make replicate measurements of X_i if there were good reasons to believe that the average of replicates is a better estimate of X_i than a single observation (Yi, 2017, Section 2.4).

1.5.4 Regression Calibration

Regression calibration (RC) (Carroll et al., 2006, Section 4.1) is a simple yet effective correction strategy applicable to almost any regression model with covariate mismeasurement. Unlike the naive analysis where X_i is replaced with its observed surrogate X_i^*, the primary idea behind the RC approach is to replace the unobserved variable X_i by the conditional expectation $E(X_i \mid X_i^*, Z_i)$, which can be estimated by regressing X_i on $\{X_i^*, Z_i\}$ if for example, validation data are available. The RC method bears relevance to the EM algorithm and is regarded as a special case of it, as discussed by Yi (2017, p.60).

Suppose the mean of X_i given (X_i^*, Z_i) can be described by

$$E(X_i \mid X_i^*, Z_i) = m_X(X_i^*, Z_i; \zeta),$$

where ζ indicates regression parameters. The RC algorithm consists of the following three steps:

Step 1. Estimate $m_X(X_i^*, Z_i; \zeta)$ by validation subsample.

Step 2. Replace the unobserved X_i by its estimate $m_X(X_i^*, Z_i; \hat{\zeta})$, and run the standard analysis to obtain point estimates of parameters of interest.

Step 3. Implement the bootstrap method to obtain the standard errors associated with the estimated parameters in step 2.

1.6 COVID-19 Data

Our research is motivated by the emerging COVID-19 data. Here we outline some features of COVID-19 that will be examined in this thesis.

1.6.1 COVID-19: Epidemiology

Estimating the incubation period is crucial for disease control. Having a sensible estimate of the median incubation time helps the government and healthcare sector decide on a rationale quarantine time. With the aim of determining a reasonable quarantine time, Khadem Charvadeh et al. (2022) examined some useful methods for modeling the distribution of the COVID-19 incubation time. Estimating recovery times for infected patients is of great importance for healthcare workers to effectively allocate the limited medical resources to cope with the COVID-19 crisis. Moreover, understanding the relationships of demographic factors, such as age and gender, with COVID-19 is essential as it helps healthcare professionals prioritize treatment of patients with different characteristics. The Canada COVID-19 website (Liu et al., 2020) examined and displayed some of the epidemiological characteristics of COVID-19 in Canada. While various efforts have been made to study the behavior of SARS-CoV-2 since the

outbreak of COVID-19, the understanding of COVID-19 has been constantly enhanced as more COVID-19 data become available. Extensive evidence-based studies from multiple angles are required to comprehensively unveil the clinical characteristics of COVID-19 by examining the data coming from different sources as the pandemic evolves.

1.6.2 COVID-19: Case Fatality Rate and Prevention Strategies

The case fatality rate (*CFR*), defined as the proportion of deaths among those individuals infected with the disease, is considered a key indicator for describing the severity of the disease. To reduce the detrimental impact of COVID-19, including lowering the *CFR*, various preventive measures such as facial coverings, social distancing, lockdowns, testing, and contact tracing were implemented by different countries. While clinical and epidemiological characteristics of patients contribute to the *CFR*, the stringency of responses and containment measures resulted in country-level variation in *CFR* (Liang et al., 2020). For instance, as of October 13, 2021, the *CFR* of COVID-19 in Yemen was reported to be as high as 18.95%, whereas in Singapore, it was reported to be 0.14% (Our World in Data, 2021).

Since the COVID-19 pandemic started, estimation of the *CFR* and identifying the associated factors have attracted extensive research. Some studies focused on exploring patient-level risk factors. For example, Zhou et al. (2020) conducted a retrospective cohort study of 191 adult COVID-19 patients and found that older age, higher sequential organ failure assessment scores, and elevated d-dimer at admission were risk factors for death. Chen et al. (2020) examined the clinical characteristics and symptoms of 799 COVID-19 patients and found that the median age of deceased patients was significantly older than recovered patients. Chronic hypertension and other cardiovascular comorbidities were found to be associated with deceased patients. Other studies explored the association of *CFR* with gender (Jin et al., 2020), obesity (Klang et al., 2020), diabetes (Guo et al., 2020), and kidney diseases (Cheng et al., 2020). Those studies indicated that, to some degree, the regional *CFR* may be partly explained by the factors such as age and health status, including the presence of having diseases such as coronary heart disease and hypertension.

However, national-level variation in *CFR* suggests the existence of other driving forces

behind fatality risk. This gap in COVID-19 fatality rate has been investigated in several studies to help national health organizations lay the groundwork for protecting high-risk patients. Some researchers investigated possible scenarios for the disparity in the risk of death from COVID-19 and suggested that the healthcare system's capacity, effectiveness of government policies, people's compliance, testing, and case detection capacities were effective in prevention and attenuation of COVID-19 severity (Liang et al., 2020; Ji et al., 2020; Chaudhry et al., 2020; Kayano and Nishiura, 2020). Nevertheless, to the best of our knowledge, there has been no research evaluating the consequences of different strictness of policies taken by different countries, which typically change over time.

In Chapter 3, we cast the problem into the Q-learning framework, which basically aims to determine optimal preventive policies using regression models. Q-learning is a useful reinforcement learning scheme to select optimal policies sequentially for an agent in a given environment. While it is commonly implemented to estimate the value of a single action that maximizes the expected cumulative reward, in Chapter 3, we apply the Q-learning approach to a setting where two actions need to be optimized. We use the Q-learning method and data from open-access databases to explore how country-specific preventive measures may possibly lower the *CFR*. Our study sheds light on the association among government actions, socioeconomic factors, and the *CFR*.

1.6.3 COVID-19: Health and Economic Costs

During the COVID-19 pandemic, many countries imposed unprecedented containment restrictions to curb the spread of the disease, and the strictness of such restrictions changed from country to country and from time to time, with some countries imposing long total lockdowns of economic and social activities and others trading off deaths against the economy. Those countries seeking to merely minimize the health costs of COVID-19 pandemic made use of epidemiology models developed for the spread of infectious diseases. Many new or refined versions of epidemiology models have been proposed during COVID-19. For example, Okhuese (2020) used differential equations and modified the SEIR model to evaluate the probability of reinfection in the recovered class and found that the rate of reinfection by the

recovered population will decline to zero over time as the virus is cleared clinically. Higazy (2020) modeled COVID-19 pandemic by fractional order SIDARTHE model. The proposed mathematical model predicts the evolution of COVID-19 pandemic and helps understand the impact of different preventive measures with different values of the fractional order. In a study conducted by Pribylová and Hajnova (2020), they proposed a generalization of the SEIR model by accommodating asymptomatic infectious cohort in understanding the epidemic dynamics of COVID-19, and demonstrated that the asymptomatic cohort plays a crucial role in the spread of the COVID-19. These models, however, have limited applicability in policy analysis for they have economic consequences for not taking the trade-off between health and economic outcomes into account. This suggests the need for developing models to minimize economic losses as well as health costs simultaneously.

Some researchers investigated the trade-off between the COVID-19 outbreak and economic activities. For example, Kano et al. (2021) proposed an abstract agent-based model of the COVID-19 outbreak that accounts for economic activities. Lasaulce et al. (2021) proposed a simple yet practical model to study the fundamental trade-off between economic and health aspects of the COVID-19 pandemic. However, these studies implicitly assume that the covariates are error-free. In practice, this assumption often does not hold, necessitating the need for correction strategies when dealing with data that involve misclassified discrete covariates and mismeasured continuous covariates. Therefore, the objective of Chapter 5 is to introduce appropriate correction strategies to address this issue within the framework of weighted dynamic programming.

1.7 Thesis Organizations

In this thesis, we employ survival analysis and reinforcement learning techniques to examine COVID-19 data. While the literature on DTRs and mismeasurement is vast, to the best of our knowledge, there has been little work on optimizing DTRs with misclassified covariates or mixed measurement error and misclassification in covariates. In this thesis, we present correction strategies to ameliorate the effects of covariate mismeasurement in developing DTRs with the Q-learning method.

The remainder of this thesis is organized as follows. In Chapter 2, we employ semiparametric and nonparametric survival models as well as text mining and data visualization techniques to examine the clinical manifestations and epidemiological features of COVID-19 in the early pandemic by studying a dataset from January 22, 2020 to March 29, 2020. The results of this chapter have been published by *The Journal of Data Science* (Khadem Charvadeh and Yi, 2020). Khadem Charvadeh led this project and conducted data analysis.

In Chapter 3, we use Q-learning for tailoring the strictness of preventive policies to country-specific characteristics and evolving situation to leverage the salutary effects of prevention strategies with an objective of lowering the *CFR*. The results of this chapter have been published online by *Statistics in Biosciences* (Khadem Charvadeh and Yi, 2023b).

In Chapter 4, we study how Q-learning may be affected by covariate misclassification, and we present correction strategies to reduce or eliminate the bias induced by covariate misclassification. The results of this chapter have been wrapped up as a paper and submitted for publication (Khadem Charvadeh and Yi, 2023a).

In Chapter 5, we study the impact of the mix misclassified discrete covariates and mismeasured continuous covariates in Q-learning with a compound outcome. The chapter also demonstrates the practical application of Q-learning with a compound outcome in analyzing the trade-off between health and economic costs associated with the COVID-19 pandemic, with specific attention given to the presence of mixed measurement error and misclassification in covariates.

Chapter 2

Data Visualization and Descriptive Analysis for Understanding Epidemiological Characteristics of COVID-19

2.1 Introduction

COVID-19 is a disease caused by the severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2). The disease was reported to spread in people in December 2019, and it was important to understand epidemiological features of COVID-19, including answering urgent questions such as: (1) What was the average time of symptom onset? (2) How long did it take for infected patients to recover? (3) Was there any age or gender difference in the recovery of infected patients? (4) What were the common symptoms of infected patients?

While each of these questions warranted in-depth research when more data about COVID-19 became available, in the beginning of the pandemic, we conducted a prompt exploratory analysis of the clinical manifestations and epidemiological features of COVID-19. The research reported in this chapter was conducted right after the COVID-19 pandemic was declared by the WHO. The goal was to provide a timely examination of available COVID-19

data to offer intuitive insights into further in-depth research. With the available Kaggle novel coronavirus dataset of 3397 patients dated from January 22, 2020 to March 29, 2020, we analyzed the data using semiparametric and nonparametric survival models as well as text mining and data visualization techniques.

The remainder of this chapter is organized as follows. In Section 2.2, we describe the data and examine different features of COVID-19 by data visualization. In Section 2.3, we employ survival analysis techniques to estimate the distribution of recovery times for infected patients. In Section 2.4, we estimate the average time of symptom onset. We conclude the manuscript with discussions in the last section.

2.2 Data Visualization

2.2.1 Data Description

In this study, we use the Kaggle novel coronavirus dataset from January 22, 2020 to March 29, 2020. The dataset, available as a Google spreadsheet at https://www.kaggle.com/datasets/sudalairajkumar/novel-corona-virus-2019-dataset, has been updated automatically every five minutes based on Johns Hopkins Center for System Science and Engineering (CSSE) data (https://github.com/CSSEGISandData/COVID-19). The dataset consists of measurements of 3397 people with the novel virus from 39 countries including those in Europe, Asia, and Africa. There are 14 variables representing the *summary, location, country, gender, age, symptom onset, hospital visit date, exposure start, exposure end, visiting Wuhan, from Wuhan, death, recovery status, and symptoms* of the infected cases. Using the information given in the *summary*, *exposure start*, *exposure end*, *symptom onset*, and *recovery status*, we further extract more specific information from the original dataset, including *infection source, travel history, time gap between exposure to symptom onset, and time gap between symptom onset to recovery*. A copy of the dataset is available at https://github.com/YasinKhc/Covid-19. Among 3397 patients, only 1449 of them have the information of age which ranges from 3 months to 96 years. In Table 2.1, we present the age distribution of infected cases separately for females and males.

2.2.2 Descriptive Analysis

Among the 3397 patients, we found that older people have a higher fatality rate compared to younger people. The mean and median of age for deceased cases were found to be 71.5 and 73.5, respectively. The left graph in Figure 2.1 displays the side-by-side barplots for the counts of deceased cases for males and females divided into six age groups, and the right graph in Figure 2.1 records the fatality rate for men and women in the six different age groups, where the fatality rate is calculated as the ratio of the number of deaths in an age group with a given gender to the number of infected cases in that group. It is clear that the fatality rate increases with age, and the fatality rate for men in each age group appears higher than that for women. These results are consistent with those reported by Jin et al. (2020).

We further perform the Chi-square test of independence (Pearson, 1900) to determine whether there is a statistically significant association between age/gender and fatality. For the null hypothesis that the fatality rate is identical for all the age groups, we obtain the p-value of the Chi-square test to be 0.0005. For the null hypothesis that the fatality rate is identical for males and females, we obtain the p-value of the Chi-square test to be 0.0748.

The left plot in Figure 2.2 shows that around 28% of the infected people had a recent travel history. The right plot in Figure 2.2 reports that 13% of the cases had a close contact with other infected people, and the source for the rest large portion (87%) of infections remains unknown, which is very likely due to undetected community transmissions. Among those people with unknown infection sources, about 30% of them had a recent travel history.

To understand what symptoms are most related to infected cases with COVID-19, we perform a text analysis using a *word cloud* (Viégas and Wattenberg, 2008), which typically visualizes word frequencies by using different sizes of the words. The more common a term appears in a text dataset, the larger and bolder it appears in the word cloud. Word clouds are an intuitive tool for visualizing and highlighting words with greater prominence. To generate a word cloud for symptoms of COVID-19, we first collapse the *summary* into a single text document and extract the terms and words describing the symptoms of infected patients, and then store them in a new text document. Thereafter, different medical words and terms that represent a specific symptom are summarized into a single unique word or term. For exam-

ple, in the *summary*, besides *difficulty breathing*, three other terms were alternatively used to describe the same symptom related to breathing: *shortness of breath*, *dyspnea*, and *respiratory distress*. In our text analysis here, we classify them as the same description for the symptom of breathing and then unify them with the term "difficulty breathing". Next, using the obtained text document and the word cloud generator in the package ***wordcloud*** in *R*, we summarize the symptoms for 652 COVID-19 infected patients in Figure 2.3. It is clearly seen that fever, cough, and pneumonia are the most frequent symptoms reported by those patients.

2.3 Examination of Recovery Time

To help the government and health authorities prepare for major spikes in the number of new COVID-19 infected cases, it is important to understand the times for infected patients to recover. In this section, we use survival analysis techniques to study the recovery times of infected patients. Here the recovery time of an infected patient, denoted as T, is taken as the time-to-event, or survival time, using the terminology in survival analysis (e.g., Lawless, 2003). In other words, the *event* is defined to be *recovered*, and hence, patients who die from COVID-19 are treated as censored.

First, we use the distribution-free Kaplan-Meier approach to examine the survivor function $S(t) = P(T > t)$ for the recovery times, where $t \in [0, 45]$ with $[0, 45]$ representing the study period of 45 days, and 0 is defined as the time of symptom onset for an infected patient.

We examine the recovery times from three angles. First, we do not distinguish infected cases; secondly, we classify the infected cases into two groups by gender; thirdly, we divide the infected cases into three age groups: (0, 40], (40, 60], and (60, 96]. The corresponding Kaplan-Meier estimates are reported in Figure 2.4. The top panel of Figure 2.4 illustrates the Kaplan-Meier time-to-recovery survival curve for all the infected cases, where the red curve represents the estimated probabilities, the red shaded areas stand for the 95% confidence region, and patients who are censored are marked with + signs. The dashed dark lines indicate the survivor probability at the median recovery time, saying that with 50% of the probability, an infected patient takes more than 20 days to recover (if they would recover). A 95% confidence interval for the median recovery time is (19, 21).

The middle panel of Figure 2.4 shows the Kaplan-Meier survival curves of recovery times for men and women, which are not considerably different. Furthermore, applying the log-rank test (Harrington, 2005) to assess whether or not the difference between the two curves is statistically significant, we obtain that the p-value is 0.5, clearly showing no evidence that the recovery time differs for men and women. Table 2.2 gives the median recovery times and their corresponding 95% confidence intervals for men and women.

The bottom panel of Figure 2.4 displays the Kaplan-Meier survival curves for the three different age groups. It can be visually concluded that people of older age are more likely to have longer recovery times. The corresponding log-rank test yields the p-value to be 10^{-4}, supporting that the differences in recovery times for different age groups are statistically significant. Median recovery times and their corresponding 95% confidence intervals for the three age groups are summarized in Table 2.3.

Next, we quantify how the recovery time is associated with age and gender. We employ the semiparametric AFT model:

$$\log T = \beta_0 + \beta_1 \times \text{gender} + \beta_2 \times \text{age} + \epsilon,$$

where β_0 is the intercept, β_1 and β_2 are regression parameters, and ϵ is the error term with mean zero and an unspecified probability distribution. For ease of interpretation, we use ten years as the unit of age, as suggested by the editor. Estimation of the parameters can be obtained using the generalized least squares approach (e.g., Chiou et al., 2014); the results are reported in Table 2.4.

The analysis results show no evidence that recovery times differ in women and men. Age is found to be significantly related to the recovery time. Older infected patients need a longer time to recover from COVID-19. Exponentiating the estimate of β_2, we quantify the age effect on the recovery time. With the gender effect adjusted, ten years older in age would extend the recovery time by 9.9%.

2.4 Gap Time Between Exposure and Symptom Onset

One of the major concerns that healthcare workers and the government have been trying to address is on *stealthy transmissions* of COVID-19. Researchers at Columbia University's Mailman School of Public Health used a computer model to show how undetected cases may boost the spread of the COVID-19 outbreak in China. They showed that the virus spread was rapid and its containment was challenging (Li et al., 2020). Understanding the average gap time between the time of exposure to the virus and symptom onset for infected patients is useful for healthcare workers and the government to make effective measures to curb the spread of the virus.

Among the 3397 infected people, 207 reported both the time for exposure and the symptom onset time. The time of exposure is taken as an approximate time a patient contracted the virus by having a close contact with someone who was already infected or traveling to infected areas. The symptom onset date is based on the time when an infected patient experienced flu-like symptoms such as fever, sore throat, and in more severe cases, difficulty breathing. Eighty-five patients reported a time interval for exposure spanning from 1 to 27 days. We treat those exposure intervals with a length of less than one day as a single time point. To understand the underlying incubation times for infected cases who reported different types of information on infection, we estimate the median and average incubation times for the cohort of 3397 infected patients using the following three methods:

- Method 1: the time period between the start time of exposure and symptom onset.
- Method 2: the time period between the end time of exposure and symptom onset.
- Method 3: we use the middle point of the time interval to approximate the exposure time, and take the time period between the approximated exposure time and symptom onset.

For 140 patients who reported only a single time point for exposure, these three methods will yield the same values for them. For the cohort of 3397 infected cases, Method 1 yields that the mean and median incubation times to be 8.4 and 6 days, respectively; Method 2 outputs a lot smaller mean and median incubation times which are respectively 3.3 and 2 days; and Method 3 gives that the mean and median of the incubation period are 5.8 and 5 days, respectively. The

estimates of Method 3 are similar to those reported by Lauer et al. (2020) and Men et al. (2020). Lauer et al. (2020), by conducting a pooled analysis of 181 infections reported between January 4, 2020 and February 24, 2020, found that median incubation period to be 5.1 days. Men et al. (2020) used a chain-of-infection data collected from 10 regions in China to estimate the median incubation period. They employed different statistical approaches, such as Monte-Carlo simulations as well as non-parametric methods, and estimated that the mean and median of incubation times are 5.8 and 5 days, respectively.

To show how incubation times may differ between females and males, in the left panel of Figure 2.5 we report the boxplots of the incubation times obtained from Method 3 for 31 females and 49 males. To see possible age effects, in the right panel of Figure 2.5, we graph the incubation times for three age groups, where 21, 30, and 25 patients are included in the age groups of 0-34, 35-54, and 55-96, respectively. The median incubation period for patients aged within 35-54 is the largest, and the median incubation period for patients over 55 years of age is slightly longer than that of the age 0-34 group. However, incubation times for older patients have more variability than those for younger infected cases.

2.5 Discussion

In this chapter, we explore the epidemiological characteristics of COVID-19 by studying a Kaggle novel coronavirus dataset, dated from January 22, 2020 to March 29, 2020, which includes 3397 infected cases and 83 deaths from COVID-19. We find that the median incubation time of COVID-19 is about 5 days, and older people are more likely to have a longer incubation period. Our text analysis shows that the most dominant symptoms of COVID-19 are fever, cough, and pneumonia. The non-parametric Kaplan-Meier method yields a median recovery time of 20 days for infected patients who are not stratified by their characteristics. Our findings further suggest that the recovery time increases as the age increases, and there is no significant gender-difference in recovery times.

As discussed by He et al. (2020), while many studies examined epidemiological characteristics of COVID-19, those studies do not necessarily reveal the same findings or similar estimates of the same measure. For instance, regarding the estimate of the average incuba-

tion times, He et al. (2020) reviewed five studies conducted between December 31, 2019 and February 24, 2020, and those studies reported varying average incubation time, ranging from 4.9 days to 6.4 days. In addition, we note that our estimate of the median incubation time differs from the estimate 8.1 days provided by Qin et al. (2020). The discrepancies in estimating the same quantity are primarily attributed to the heterogeneity in different studies, including the differences in the time window, the study subjects, the study design, the model assumptions, and the measures of controlling the virus spread by different regions.

We point out that the validity of the analysis results here relies on the quality of the Kaggle data we use. In our analysis, we ignore missing observations, which is basically driven by the perception that missingness arises completely at random. However, when such an assumption is not feasible, proper adjustments of missingness effects are generally expected. On the other hand, as commented by a referee, reporting bias and recall bias should be aware of when analyzing the COVID-19 data. If the degree of such biases is not mild, then proper debiasing adjustments should be introduced in inferential procedures to yield valid or nearly valid analysis results. Methods of addressing effects of error-in-variables can be employed for this purpose. For details, see Carroll et al. (2006), Yi (2017), and Yi et al. (2021).

Finally, we note that our analysis results are obtained from using the reported information for those patients who were assessed by medical personnel. The information for infected patients with mild symptoms or asymptomatic infections was often not available for being included in the dataset because those patients did not go to hospital for assessment. As a result, when interpreting the results, care is needed for the target population.

Table 2.1: Age distribution of infected cases by gender: The entries display the number and the percentage (in parentheses) for each cohort

Age range (in year)	0-19	20-39	40-59	60-79	80-96	Total
Male	28 (3%)	193 (24%)	313 (38%)	242 (30%)	41 (5%)	817
Female	25 (4%)	168 (27%)	212 (34%)	186 (29%)	41 (6%)	632

Table 2.2: Median recovery time for male and female

Gender	The number of infected patients	The number (percentage) of recovery	Median	95% Confidence interval
Female	52	43 (83%)	20	(17, 21)
Male	89	58 (65%)	20	(19, 23)

Table 2.3: Median recovery time (in day) for different age groups

Age group	The number of infected patients	The number (percentage) of recovery	Median	95% Confidence interval
0-40	47	45 (96%)	18	(16, 20)
41-60	50	45 (90%)	20	(17, 22)
61-96	43	10 (23%)	26	(21, 30)

Table 2.4: Analysis results of recovery times under the semiparametric AFT model

Parameters	estimate	standard error	95% Confidence interval
Intercept (β_0)	2.498	0.119	(2.265, 2.731)
gender (β_1)	0.066	0.069	(-0.069, 0.201)
age (β_2)	0.094	0.022	(0.051, 0.137)

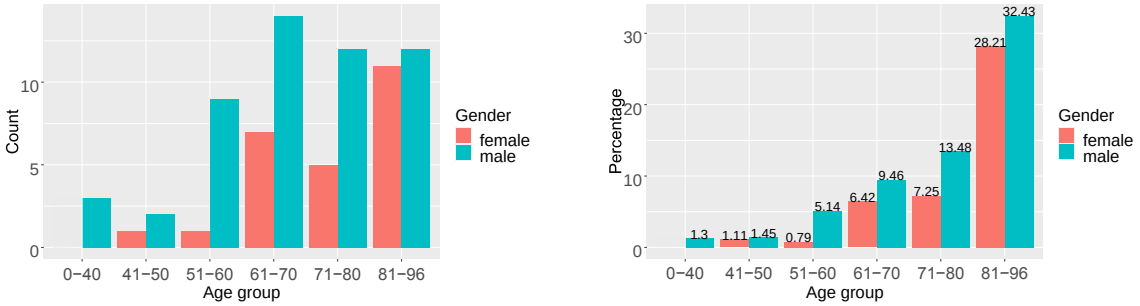

Figure 2.1: Barplots for the number of deceased cases and fatality rate

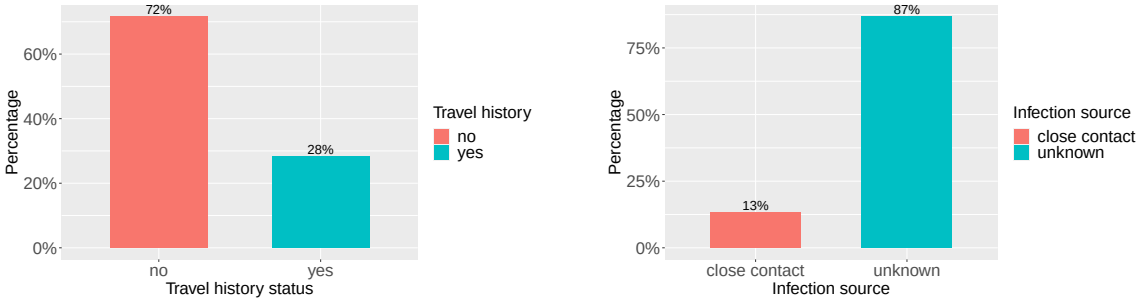

Figure 2.2: Barplots for the recent travel history and infection source

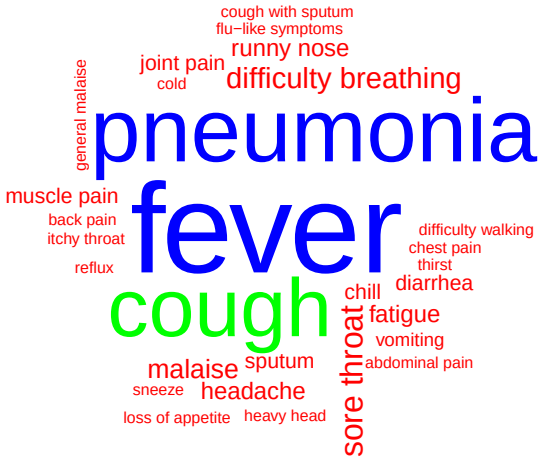

Figure 2.3: Word cloud of the symptoms

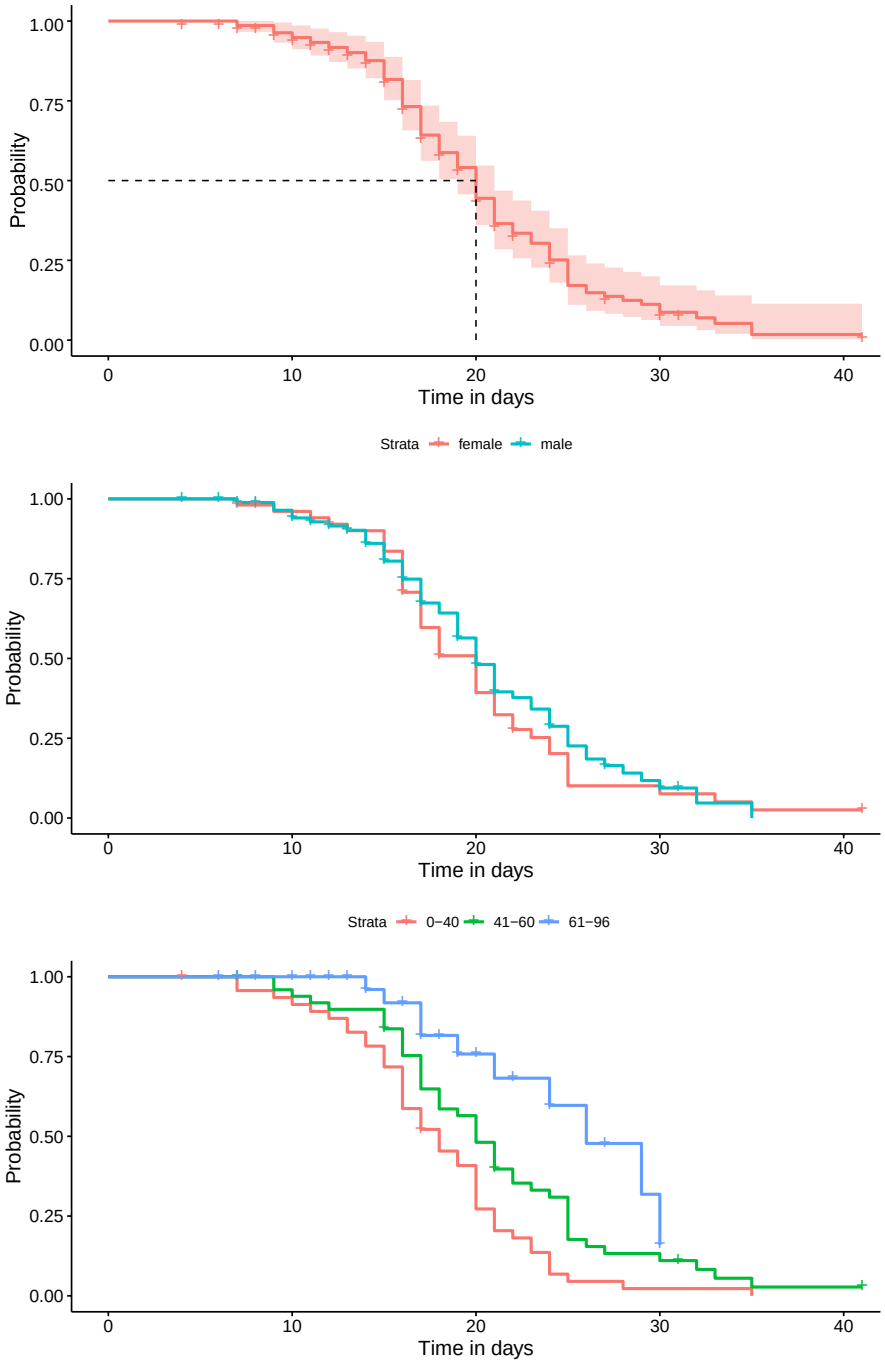

Figure 2.4: Kaplan-Meier time-to-recovery survival curves

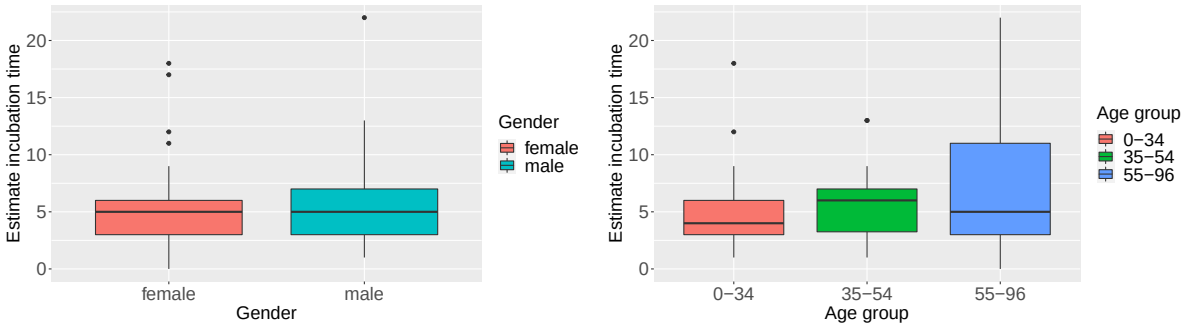

Figure 2.5: Boxplots of estimated incubation times by gender and three age groups

Chapter 3

Understanding Effective Virus Control Policies for COVID-19 with the Q-Learning Method

3.1 Introduction

The *CFR* of COVID-19 is a useful measure to describe the disease severity, which, however, changes considerably from country to country and from time to time. To reduce the detrimental impact of COVID-19, it is imperative to understand how different mitigation policies adopted by different countries may help lower the COVID-19 *CFR*. Using data from 175 countries from January 13 of 2020 to March 9 of 2021, we investigate possible factors associated with the *CFR* and use the Q-learning algorithm to assess optimal preventive policies adopted by individual countries to reduce their COVID-19 *CFR*.

The chapter is structured as follows. In Section 3.2, we describe the Q-learning implementation procedure. Section 3.3 describes the data sources and the variables of interest. In Section 3.4, we apply the procedure in Section 3.2 to analyze the data described in Section 3.3. Sensitivity analyses are further conducted in Section 3.5. We conclude the chapter with discussions in Section 3.6.

3.2 Q-Learning Implementation with Observational Data

3.2.1 Notation and Framework

In this section, we describe the Q-learning algorithm for K pre-determined stages which may involve both stage-invariant and stage-specific covariates. For example, considering the PROmotion of Breastfeeding Intervention Trial (PROBIT) (Kramer et al., 2001), a randomized trial with an interest in assessing the impact of a breastfeeding promotion intervention on children's cognitive development, Moodie et al. (2012) employed the Q-learning method by framing the study period as two stages, with stage 1 covering the period from the birth of study subjects to 3 months and stage 2 spanning from 3 to 6 months of age. The treatment of interest is breastfeeding, measured in each of the stages, and the outcome is taken as the verbal cognitive ability score. Here, covariates include stage-specific variables such as the birthweight of the infant and the infant's 3-month weight, as well as stage-invariant variables or baseline covariates such as geographical location, mother's education, mother's smoking status, mother's age, and children's gender.

We use boldface to denote vectors or spaces, and let capital letters denote random variables and lower case letters denote their realizations. For $k = 1, \cdots, K$, let Y_k denote the outcome at the end of stage k, and assume a smaller Y_k is more desirable. Let $\mathbf{A}_k = \{A_{k,1}, \cdots, A_{k,r}\} \in \boldsymbol{\mathcal{A}}$ denote the vector of r discrete actions taken at stage k, where $\boldsymbol{\mathcal{A}}$ is the action space. Let $\mathbf{C}_j = \{C_{j,1}, \cdots, C_{j,q_j}\} \in \boldsymbol{\mathcal{C}}_j$ denote the stage-invariant features for action $A_{k,j}$ for $j = 1, \cdots, r$, and let $\mathbf{O}_k = \{O_{k,1}, \cdots, O_{k,p_k}\} \in \boldsymbol{\mathcal{O}}_k$ denote the vector of stage k specific features, where $\boldsymbol{\mathcal{C}}_j$ and $\boldsymbol{\mathcal{O}}_k$ are their corresponding spaces. Let $H_k := \left\{\mathbf{C}_1, \cdots, \mathbf{C}_r, \mathbf{O}_1, \cdots, \mathbf{O}_k, \mathbf{A}_1, \cdots, \mathbf{A}_{k-1}\right\}$ denote the collection of covariates at stage k for $k = 1, \cdots, K$. The type of covariates in $\mathbf{O}_k$ and $\mathbf{C}_j$ can be differentiated by whether or not the covariates interact with actions; those variables interacting with actions are called tailoring or prescriptive variables (Chakraborty and Moodie, 2013, Section 3.4.1).

A K stage policy consists of K decision rules, denoted d_k for $k = 1, \cdots, K$, which map the domain of H_k to an action choice in $\boldsymbol{\mathcal{A}}$. To construct a sequence of optimal decision rules, we assume a smaller cumulative outcome is more desirable, and then define the Q-functions for

stage K and stage k with $k = K-1, \cdots, 1$:

$$\begin{aligned} Q_K(H_K, \mathbf{A}_K) &= E[Y_K \mid H_K, \mathbf{A}_K]; \\ Q_k(H_k, \mathbf{A}_k) &= E[Y_k + \min_{\mathbf{a}_{k+1}} Q_{k+1}(H_{k+1}, \mathbf{a}_{k+1}) \mid H_k, \mathbf{A}_k]. \end{aligned} \tag{3.1}$$

In the case where the true Q-functions are known, optimal decision rules are determined by:

$$d_k^{opt}(h_k) \triangleq \arg\min_{\mathbf{a}_k} Q_k(h_k, \mathbf{a}_k) \qquad \text{for } k = K, \cdots, 1. \tag{3.2}$$

In applications, the Q-functions are typically unknown and need to be modeled. As the Q-functions basically represent conditional expectations, it is natural to delineate them using regression modeling techniques such as linear regression. For $k = 1, \cdots, K$, consider the linear regression model

$$Q_k(H_k, \mathbf{A}_k) = \beta_k^{\mathrm{T}} H_{k0} + \sum_{j=1}^{r} (\psi_{k,j}^{\mathrm{T}} H_{k1,j}) A_{k,j}, \tag{3.3}$$

where H_k is divided into H_{k0} and $\{H_{k1,j} : j = 1, \cdots, r\}$, with β_k and $\{\psi_{k,j} : j = 1, \cdots, r\}$ representing the associated regression coefficients. Here H_{k0} contains the variables having a predictive effect on the outcome, and $H_{k1,j}$ denotes the prescriptive components of H_k that interact with the action, $A_{k,j}$. Let $\psi_k = (\psi_{k,1}^{\mathrm{T}}, \cdots, \psi_{k,r}^{\mathrm{T}})^{\mathrm{T}}$ for $k = 1, \cdots, K$.

3.2.2 Estimation Procedure

The determination of (3.2) hinges on the evaluation of the Q-functions that are modeled by (3.3). To estimate the model parameters in the Q-functions, we use the data in a random sample of i.i.d. observations for multiple study units, say n units. We now add i to the symbols Y_k, H_k and $\mathbf{A}_k$ as a subscript to show the corresponding variables for study unit i, where $i = 1, \cdots, n$.

Define $Q_{K+1} \triangleq 0$, and by moving backward through stages, the estimation of the regression

coefficients for each stage can be formed as a minimization problem:

$$(\hat{\beta}_k, \hat{\psi}_k) = \arg\min_{\beta_k, \psi_k} \frac{1}{n} \sum_{i=1}^{n} \Big[\{ Y_{k,i} + \min_{\mathbf{a}_{k+1}} Q_{k+1}(H_{k+1,i}, \mathbf{a}_{k+1,i}; \hat{\beta}_{k+1}, \hat{\psi}_{k+1}) \} - Q_k(H_{k,i}, \mathbf{A}_{k,i}; \beta_k, \psi_k) \Big]^2 \tag{3.4}$$

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

(3.4) can be implemented using the ***lm*** function in *R*. Variance estimates of the least squares estimators can be obtained from the ***lm*** summary in *R*, and W-type CIs for regression parameters can be obtained using ***confint*** function in *R*.

Substituting the resultant estimates from (3.4) into (3.3), we estimate the optimal action

$$\hat{d}_k^{opt}(h_k) = \arg\min_{\mathbf{a}_k} Q_k(h_k, \mathbf{a}_k; \hat{\beta}_k, \hat{\psi}_k) \qquad \text{for } k = K, \cdots, 1. \tag{3.5}$$

With the linear model (3.3) for the Q-functions, the minimums in (3.5) occur at the minimum or maximum of $A_{k,j}$, depending on the sign of the coefficient of $A_{k,j}$, $\psi_{k,j}^{\mathrm{T}} H_{k1,j}$ in (3.3).

3.3 Data Sources and Extraction

3.3.1 Data Descriptions

We extract publicly available COVID-19 data across 175 countries for a period of about ten months with the beginning marked by the date of the first confirmed COVID-19 case in each country; calendar times may vary from country to country because of different times for identifying the first confirmed cases in different countries. The data include information about containment and closure policies, facial covering policy, diagnostic testing policy, contact tracing policy, protection of elderly people policy, the total number of COVID-19 cases per million people, the total number of COVID-19 deaths, care system quality score, obesity prevalence, smoking prevalence, socioeconomic factors, and substance use prevalence. Here, substance use prevalence indicates the age-standardized prevalence of adults with a substance use disorder, including alcohol, opioid, cocaine, amphetamine, cannabis, and other drug use.

Information about various policies, including those of containment and closure, facial cov-

ering, diagnostic testing, contact tracing, and protection of elderly people, is collected from the Oxford COVID-19 Government Response Tracker (OxCGRT) (Hale et al., 2021). The policies are recorded as scores which are aggregated into a suite of policy indices. Since January 1, 2020, policy responses covering more than 180 countries are tracked and updated on a daily basis; they are coded as 23 variables with 0 indicating no enforcement.

From the 23 variables, we opt to focus on eight policies: school closures, workplace closures, restrictions on gatherings, international travel controls, facial covering, diagnostic testing, contact tracing, and protection of elderly people policies. Each of these eight policies is described by two time-varying variables: an ordinal variable of the policy type and a strictness variable showing the implementation level of the policy. For example, for the workplace closures policy, its ordinal variable takes on a value from 0 to 3, with 0 for "no measures", 1 for "recommend closing (or work at home) or all businesses open with alterations resulting in significant differences compared to non-COVID-19 operation", 2 for "require closing (or work at home) for some sectors or categories of workers", and 3 for "require closing (or work at home) for all but essential workplaces (e.g., grocery stores, doctors)"; and its strictness variable, summarized as a single score, takes a value ranging from 0 to 100 with a higher value representing a stricter implementation of the policy.

To be specific, for policy $j = 1, \cdots, M$ with $M = 8$, let $v_{j,t}$ denote the ordinal variable for policy j on day t, with N_j representing its maximum ordinal value; and let $I_{j,t}$ denote the strictness score of implementing the jth policy on day t. Further, let F_j denote a binary value from $\{0, 1\}$ to indicate whether the jth policy has a time-dependent flag variable, denoted $f_{j,t}$, for a geographic scope, with 1 representing "yes" and 0 otherwise; and $f_{j,t}$ is the binary flag variable for the jth policy on day t, taking value 0 if the policy is "geographically targeted" (i.e., being applied only to a sub-region of a jurisdiction) or 1 if the policy is "general" (i.e., being applied throughout that jurisdiction). Among those eight chosen policies, five policies, including school closures, workplace closures, restrictions on gatherings, facial coverings, and protection of elderly people, have a binary flag for geographic scope, whereas three other policies, including international travel controls, diagnostic testing, and contact tracing, do not have a binary flag for geographic scope. For policies without a binary flag for the geographic scope, $F_j = 0$, and consequently $f_{j,t} = 0$. To calculate the strictness score, Hale et al. (2021)

suggested the following formula:

$$I_{j,t} = 100\left\{\frac{v_{j,t} - 0.5(F_j - f_{j,t})}{N_j}\right\}.$$

To provide an overall measure for the preventive policies of the similar nature, we calculate the average of strictness scores for the policies grouped by nature. Considering policies for containment and closure, we put school closures, workplace closures, restrictions on gatherings, and international travel controls in the same group, and let $\mathcal{E}_1$ denote the set of labels for those policies. Considering policies related to health system policies of testing and tracing, we use two ways to combine the policies with protection of elderly people included or not included; specifically, let $\mathcal{E}_2$ denote the index set for diagnostic testing, contact tracing, and protection of elderly people, and let $\mathcal{E}_3$ denote the set for diagnostic testing and contact tracing. Then, for $l = 1, 2, 3$, define Index_l on day t to be

$$\text{Index}_{lt} = \frac{1}{|\mathcal{E}_l|}\sum_{j\in\mathcal{E}_l} I_{j,t}.$$

Basically, a higher value of Index_{lt} indicates a stricter combined policy at time t for $l = 1, 2, 3$.

3.3.2 Analysis Objective

Data on the total number of COVID-19 cases per million people and the total number of COVID-19 deaths are extracted from the website Ourworldindata (Ritchie et al., 2020). Care system quality score, obesity prevalence, smoking prevalence, and substance use prevalence for 2019 are obtained from the Legatum Institute (The Legatum Institute, 2019), and we use "care-score", "obesity-prev", "smoking-prev", and "substance-prev", respectively to represent them for short. In our analyses, we include the following socioeconomic factors: the most recent population weighted geometric mean density ("popu-density" for short) (Edwards et al., 2021), the population proportion of people aged 65 and above for 2019 ("senior-prop" for short) (The World Bank, 2019a), gross domestic product per capita based on purchasing power parity for 2019 ("GDP" for short) (The Global Economy, 2019), government effectiveness score for 2019 ("government-eff" for short) (The World Bank, 2019b), and civic and social

participation score for 2019 ("civic-score" for short) (The Legatum Institute, 2019).

The inclusion of these factors is driven by the following considerations. The "care-score" assesses the ability of a health system to treat and cure diseases and illnesses in the population. It is measured based on a number of indicators, including healthcare coverage, health facilities, health practitioners and staff, satisfaction with healthcare, etc., and takes values from 0 (worst) to 100 (best). As most people and economic agents live much more concentrated in space, the "popu-density" is regarded as a more meaningful measure than the simple population density. The "GDP" is the most commonly used measure of economic activity and represents the total monetary value of the produced goods and services in a country during a specific period. A country with a larger "GDP" tends to have a higher standard of living. The "government-eff" measures perceptions of the quality of public and civil services, the quality of policy formulation and implementation, and the credibility of the government's commitment to such policies. Its value ranges from -2.5 to 2.5, with a higher value indicating better governance. The "civic-score" ranges from 0 to 100, with a higher value indicating a better community involvement and participation; it serves as a proxy for the degree of people's collaborative effort in controlling the spread of the virus.

To protect individuals from contracting the SARS-CoV-2 virus, facial covering, diagnostic testing, contact tracing, protection of elderly people, school closures, workplace closures, restrictions on gatherings and international travel controls have been recommended as effective preventive measures, and they become mandates in many countries or regions. For example, research suggested that universal masking reduced the risk of infection (e.g., Chu et al., 2020). Population testing is another strategy proved to have a significant impact on the mortality rate (e.g., Terriau et al., 2021). While these preventive policies have been widely adopted, the strictness of executing them varies from place to place and from time to time. It is interesting to investigate how the strictness of such policies can be prioritized to effectively reduce COVID-19 deaths, which is the objective of this study here. Using a value in the range [0, 100] to reflect the strictness of the policies with higher values suggesting more stringent policies, we take the outcome variable as the number of COVID-19 deaths per hundred COVID-19 cases, denoted *CFR*, and examine how different combinations of the policies may be more effective to lower *CFR* at different stages. The details are provided in the following sections.

3.4 Statistical Analysis

3.4.1 Data Preparation

The ten-month-period is divided into two parts with the first part containing the first three months and the second part regarded as the study period, where the data extracted from the first part supply the baseline information, which helps accommodate the initial severity of COVID-19 in each country into the analysis. The study period is divided into two stages (i.e., $K = 2$), with stage 1 starting from the first day of the fourth month to the last day of the sixth month, and stage 2 starting from the first day of the seventh month to the last day of the ninth month. We consider two stage-specific features: (1) *CFR*, assessed at the start of each stage, and (2) the total number of infectious COVID-19 cases per million entering each stage, determined by the CDC guideline that people remain infectious for about 5 days after symptom onset (CDC, 2022).

While it is challenging to obtain the accurate values for *CFR*, stringency score of preventive policies, and the total number of infectious COVID-19 cases for each stage, we aim to accommodate the incubation time and the duration from symptom onset to death in the calculations. Figure 3.1 shows the timeline used to determine the relevant quantities. The total number of confirmed COVID-19 cases is obtained for the period marked as TW2, and the total number of deaths from COVID-19 is counted for the period marked as TW3. The stringency score is calculated for the period marked as TW1. To obtain the total number of infectious COVID-19 cases entering a specific stage, we include the total number of confirmed COVID-19 cases a few days prior to the start of the stage as well as those individuals who contracted the coronavirus before the start of the stage but showed symptoms afterward. Consequently, the total number of infectious COVID-19 cases is evaluated over the period marked as TW4.

To calculate the COVID-19 *CFR*, we set the incubation time to be 6 days, an estimated average incubation time (Weng and Yi, 2022); to obtain the total number of infectious COVID-19 cases entering a stage, we take 5 days as the infectious period after the symptom onset (CDC, 2022), and thus TW4 timeline has a length of 11 days. In addition, we take 18 days as the duration from symptom onset to death (Verity et al., 2020). For $k = 1, 2$ and $i = 1, \cdots, n_k$, let T_{ki} and F_{ki} denote the cumulative number of COVID-19 cases and cumulative number of

COVID-19 deaths for country i at stage k, respectively. Then, the *CFR* per hundred COVID-19 cases is given by $\frac{F_{ki}\times 100}{T_{ki}}$. Here, study unit i represents the ith country.

To remove the nonnegativity constraint of *CFR*, we apply the log-transformation to the *CFR* at each stage, which also helps improve the feasibility of using linear regression models (3.3). To be specific, for $k = 1, 2$ and $i = 1, \cdots, n_k$, we let $Y_{ki} = \log\left(\frac{F_{ki}\times 100}{T_{ki}}\right)$ denote the outcome for country i at stage k. In both stages 1 and 2, countries with zero *CFR* (i.e., $Y_{ki} = -\infty$) or having missing values in covariates are removed from the analysis. First, examining the stage 2 data, we remove those 12 countries having zero *CFR* and 19 countries having missing covariates values, and then we examine the data at stage 1, which further yields the removal of 4 countries with zero *CFR*. This gives us that $n_2 = 144$ and $n_1 = 140$, where n_k represents the number of the countries retained for stage k in the analysis with $k = 1, 2$.

To investigate undue effects due to extreme values, we conduct two types of analyses with or without removing those extreme values which may appear in the outcome, predictors, or both of them. The Cook's distance measure with the thresholds $4/n_2$ and $4/(n_1 - I_2)$ is used to decide a data point to be "extreme" for the stages 2 and 1 data, respectively, where I_2 is the total number of extreme values in stage 2. With extreme values removed, the sample sizes in stages 2 and 1 are 131 and 118, respectively.

As the effectiveness of implementing a policy depends on the effort from both the governments and individuals, here we are interested in assessing (1) how the effectiveness of the policy related to individual behavior may differ from that associated with government-level policies, and (2) how different combinations of government-level policies may differ in the effectiveness at different stages.

For each stage, we consider two actions, named Action_1 and Action_2, which are reflected by two binary action variables, $A_{k,j}$ for stage $k = 1, 2$ and $j = 1, 2$. To be specific, let facial-cover_k denote the average strictness score for the facial covering policy over the period of stage k; this is a measure reflecting individual behavior. For $l = 1, 2, 3$ and $k = 1, 2$, let $\text{average-Index}_{l(k)}$ represent the average of the Index_{lt} with t indexing the days in the period of stage k. Action_1 measures the preference of imposing more stringent facial covering policy to diagnostic testing and contact tracing together, and Action_2 compares the effectiveness of imposing strict restriction policies, such as school closure, workplace closure, restrictions on gatherings, and

international travel control lumped together, to implementing other steps, such as diagnostic testing, contact tracing, and protection of elders lumped together. Let $A_{k,1} = 1$ if facial-cover$_k$ is greater than average-Index$_{3(k)}$ and $A_{k,1} = 0$ otherwise; let $A_{k,2} = 1$ if average-Index$_{1(k)}$ is greater than average-Index$_{2(k)}$ and $A_{k,2} = 0$ otherwise. Let "cases-enter$_k$" represent the total number of infectious COVID-19 cases per million people entering the stage k, and let "*CFR*-enter$_k$" represent the recorded *CFR* at the start of the stage k.

Table 3.1 reports descriptive statistics for those variables over stages 1 and 2. The average of the Y_{ki} over all the countries indexed by i decreases from stage 1 to stage 2, probably indicating the potential effectiveness of preventive measures overall. The comparison of the average *CFR* over different countries at the start of stages 2 and 1 shows a more severe overall *CFR* in the beginning than the end of stage 1. The average number of infectious COVID-19 cases per million people across countries increases from stage 1 to stage 2, indicating the surge of COVID-19 infections over the study period.

In the following analyses, the covariates senior-prop, GDP, government-eff, obesity-prev, smoking-prev, substance-prev, popu-density, care-score are taken as confounders, and government-eff, popu-density and civic-score, together with cases-enter and *CFR*-enter, are regarded as prescriptive variables for $A_{k,1}$, while government-eff and popu-density, together with cases-enter and *CFR*-enter, are considered as prescriptive variables for $A_{k,2}$. This is driven by the fact that most preventive policies are made in light of the observed numbers of cases or deaths in the past. Consequently, $H_{k1,j}$ in (3.3) is set as $\{$cases-enter$_k$, *CFR*-enter$_k$, government-eff, popu-density, civic-score$\}$ and $\{$cases-enter$_k$, *CFR*-enter$_k$, government-eff, popu-density$\}$, respectively for $j = 1, 2$. For sensible comparisons, we rescale all the non-binary covariates by dividing them by their standard deviations.

3.4.2 Data Analysis

We now apply the methodology in Section 3.2 to analyze the data described in Section 3.4.1, where K is 2, the log of *CFR* is taken as the outcome, and the Q-functions (3.3) are specified as

$$
\begin{aligned}
Q_2(H_2, A_2) = {} & \beta_{20} + \beta_{21} \times \text{senior-prop} + \beta_{22} \times \text{GDP} + \beta_{23} \times \text{government-eff} \\
& + \beta_{24} \times \text{obesity-prev} + \beta_{25} \times \text{smoking-prev} + \beta_{26} \times \text{substance-prev} \\
& + \beta_{27} \times \text{popu-density} + \beta_{28} \times \text{care-score} \\
& + (\psi_{20,1} + \psi_{21,1} \times \text{cases-enter}_2 + \psi_{22,1} \times \textit{CFR}\text{-enter}_2 \\
& + \psi_{23,1} \times \text{government-eff} + \psi_{24,1} \times \text{popu-density} \\
& + \psi_{25,1} \times \text{civic-score}) A_{2,1} + (\psi_{20,2} + \psi_{21,2} \times \text{cases-enter}_2 \\
& + \psi_{22,2} \times \textit{CFR}\text{-enter}_2 + \psi_{23,2} \times \text{government-eff} \\
& + \psi_{24,2} \times \text{popu-density}) A_{2,2}; \qquad (3.6)
\end{aligned}
$$

$$
\begin{aligned}
Q_1(H_1, A_1) = {} & \beta_{10} + \beta_{11} \times \text{senior-prop} + \beta_{12} \times \text{GDP} + \beta_{13} \times \text{government-eff} \\
& + \beta_{14} \times \text{obesity-prev} + \beta_{15} \times \text{smoking-prev} + \beta_{16} \times \text{substance-prev} \\
& + \beta_{17} \times \text{popu-density} + \beta_{18} \times \text{care-score} \\
& + (\psi_{10,1} + \psi_{11,1} \times \text{cases-enter}_1 + \psi_{12,1} \times \textit{CFR}\text{-enter}_1 \\
& + \psi_{13,1} \times \text{government-eff} + \psi_{14,1} \times \text{popu-density} \\
& + \psi_{15,1} \times \text{civic-score}) A_{1,1} + (\psi_{10,2} + \psi_{11,2} \times \text{cases-enter}_1 \\
& + \psi_{12,2} \times \textit{CFR}\text{-enter}_1 + \psi_{13,2} \times \text{government-eff} \\
& + \psi_{14,2} \times \text{popu-density}) A_{1,2}, \qquad (3.7)
\end{aligned}
$$

where β_{jr}, $\psi_{js,1}$, and $\psi_{jt,2}$ are parameters for $j = 1, 2$; $r = 0, 1, \cdots, 8$; $s = 0, 1, \cdots, 5$; and $t = 0, 1, \cdots, 4$.

As discussed for (3.5), the minimizers of the Q-functions or the optimal value of $A_{k,j}$ are determined by the coefficients of $A_{k,j}$ for $k = 1, 2$ and $j = 1, 2$. That is, for $k = 1, 2$, the optimal value of $A_{k,1}$, $A_{k,1}^{\text{opt}}$, is set as $A_{k,1}^{\text{opt}} = 0$ if

$$(\psi_{k0,1} + \psi_{k1,1} \times \text{cases-enter}_k + \psi_{k2,1} \times \textit{CFR}\text{-enter}_k + \psi_{k3,1} \times \text{government-eff} + \psi_{k4,1} \times \text{popu-density} + \psi_{k5,1} \times \text{civic-score}) > 0, \quad (3.8)$$

and $A_{k,1}^{\text{opt}} = 1$ otherwise; similarly, the optimal value of $A_{k,2}$, $A_{k,2}^{\text{opt}}$, is set as $A_{k,2}^{\text{opt}} = 0$ if

$$(\psi_{k0,2} + \psi_{k1,2} \times \text{cases-enter}_k + \psi_{k2,2} \times \textit{CFR}\text{-enter}_k + \psi_{k3,2} \times \text{government-eff} + \psi_{k4,2} \times \text{popu-density}) > 0, \quad (3.9)$$

and $A_{k,2}^{\text{opt}} = 1$ otherwise.

We now conduct two analyses with and without removing extreme values, and call them Analysis 1 and Analysis 2, respectively. The DB CIs are based on 1000 first-stage and 1000 second-stage bootstrap iterations.

Analysis of Data with Extreme Values Removed

The left panel of Table 3.2 reports on the estimates and 95% CIs of the model parameters obtained from Analysis 1, i.e., analyzing the data with extreme values removed, where both W-type and DB CIs are reported for stage 1 model parameters to address potential non-regularity issues, whereas only W-type CIs are reported for stage 2 model parameters as non-regularity issues do not occur.

At the significance level 5%, the significant covariates shared by both stages include "senior-prop" (β_{21} and β_{11}), "obesity-prev" (β_{24} and β_{14}), "substance-prev" (β_{26} and β_{16}), "care-score" (β_{28} and β_{18}), and the interaction term $\text{Action}_2 \times$ "*CFR*-enter" ($\psi_{22,2}$ and $\psi_{12,2}$). The covariates that are significant for stage 2 but not for stage 1 include three interaction terms: $\text{Action}_1 \times$ "government-eff" ($\psi_{23,1}$), $\text{Action}_1 \times$ "popu-density" ($\psi_{24,1}$), and $\text{Action}_1 \times$ "civic-score" ($\psi_{25,1}$). The covariates that are significant for stage 1 but not for stage 2 include "GDP" (β_{12}) and Action_1 ($\psi_{10,1}$). The significance of those interaction terms suggests the necessity of combining the strictness of implementing preventive policies with the characteristics of individual countries. Interestingly, while W-type CIs and DB CIs reveal the same significance or insignifi-

cance for almost all covariates for stage 1 estimation, the interaction of $\text{Action}_1 \times$ "civic-score" ($\psi_{15,1}$) is found to be statistically significant based on the W-type CI, but insignificant based on the DB CI.

Analysis of Data with Extreme Values

In contrast to Analysis 1, we conduct Analysis 2 by retaining those extreme values, and report in the right panel of Table 3.2 the results in the same manner as for Analysis 1. Clearly, at the significance level 5%, the two analyses do not reveal identical findings.

Regarding stage 2, both analyses indicate that the "senior-prop" and "care-score" are significant predictors of the COVID-19 *CFR*, and that the interaction term $\text{Action}_1 \times$ "government-eff" is statistically significant, suggesting the importance of the efficacy of governmental measures. In addition, the interaction term $\text{Action}_1 \times$ "popu-density" is statistically significant, highlighting the dependence of the COVID-19 *CFR* on the population density. Conversely, the significance of the following predictors is differently revealed by the two analyses: "obesity-prev", "substance-prev", the interaction term $\text{Action}_2 \times$ "*CFR*-enter", and the interaction term $\text{Action}_1 \times$ "civic-score".

In terms of the analysis results for stage 1, both analyses reveal that "senior-prop", "GDP", and "obesity-prev" are significant predictors, irrespective of the type of CI employed. However, other covariates, including "substance-prev", "care-score", Action_1, $\text{Action}_2 \times$ "*CFR*-enter", and $\text{Action}_1 \times$ "civic-score", are not found to be significant from Analysis 2. Additionally, Analysis 2 suggests opposite evidence for the significance of "popu-density" by the W-type and DB CIs; this predictor is, however, found to be insignificant from Analysis 1.

Estimated Optimal Actions

Analyses 1 and 2 do not yield the same findings, demonstrating the significant influence of including/excluding extreme values on the analysis results. These two analyses, however, do reveal a common message that for countries characterized by high population density and low government effectiveness scores, it is beneficial to prioritize facial covering policies over diagnostic testing and contact tracing policies for stage 2.

Regarding the determination of estimated optimal actions for a country, we employ (3.8) and (3.9) in conjunction with the estimates in Table 3.2. For illustrations, in Table 3.3 we report the estimated optimal actions at stages 1 and 2 derived from both Analyses 1 and 2 for some selected countries which reflect across-world diversity, including geographic location, population size, governmental structure, and healthcare infrastructure.

Including or excluding extreme values may or may not change the estimated optimal actions, as shown in Table 3.3, which divides the reported countries into three groups according to the differences in the results of Analyses 1 and 2. We comment that countries having the same estimated optimal actions in Table 3.3 do not necessarily share the same descriptions for the priority of policies because of their dependence on the estimates of the coefficients and the values of the prescriptive variables in (3.8) and (3.9). As examples, here we particularly examine several countries from the three groups.

In the first group with identical estimated optimal actions from the two analyses, consider France as an example. For both stages 1 and 2, diagnostic testing and contact tracing policies should take priority over facial covering policy, and health system policies should be prioritized over containment and closure policies.

In the second group containing estimated optimal actions that differ in only one value from the two analyses, we consider three countries: the United States, Canada, and the United Kingdom. For the United States, both Analyses 1 and 2 suggest that the facial covering policy should take precedence over diagnostic testing and contact tracing policies during stage 1, and the opposite prioritization is for stage 2. While both analyses reveal that for stage 1, health system policies should be prioritized over containment and closure policies, the uncoverings for stage 2 are different. Analysis 1 suggests prioritization of containment and closure policies over health system policies, while Analysis 2 shows the opposite suggestion.

For Canada, both Analyses 1 and 2 recommend prioritizing diagnostic testing and contact tracing policies over facial covering policies during stage 1. However, for stage 2, Analysis 1 suggests the opposite prioritization, while Analysis 2 recommends the same prioritization. Furthermore, both Analyses 1 and 2 recommend prioritizing health system policies over containment and closure policies for both stages.

Concerning the United Kingdom, both Analyses 1 and 2 suggest prioritizing health system

policies over containment and closure policies for both stages, and both analyses concur in recommending the prioritization of facial covering policies over diagnostic testing and contact tracing policies for stage 2. Regarding facial covering for stage 1, the two analyses make different recommendations. Analysis 1 recommends to give priority to facial covering policies over diagnostic testing and contact tracing policies, whereas Analysis 2 suggests the opposite prioritization.

Finally, consider Israel in the third group which has two different values for the estimated optimal actions from Analyses 1 and 2. Both analyses suggest that facial covering policy should be prioritized over diagnostic testing and contact tracing policies for stage 2, and that containment and closure policies should be prioritized over health system policies during stage 1. In terms of differences, Analysis 1 suggests that during stage 1, diagnostic testing and contact tracing policies should be given priority over facial covering policy, and that during stage 2, containment and closure policies should take precedence over health system policies. Analysis 2 recommends the opposite prioritization for the two stages.

3.5 Sensitivity Analyses

Section 3.4.2 presents analyses of the data described in Section 3.4.1 by utilizing linear models for the Q-functions with K set as 2, where the outcome variable is taken as the continuous variable, $\log CFR$. To help understand the associated uncertainty, we further conduct sensitivity analyses for two scenarios. First, we assess the impact of the stage determination by extending the two-stage setting in Section 3.4.2 to a three-stage setting (i.e., set $K = 3$), and report the results in Section 3.5.1. Secondly, we evaluate the impact of different modeling of the Q-functions and report the results in Section 3.5.2. Same as in Analyses 1 and 2, the DB CIs are based on 1000 first-stage and 1000 second-stage bootstrap iterations.

3.5.1 Continuous Outcome with Three Stages

We now repeat Analyses 1 and 2 in Section 3.4.2 by differently dividing the study period into three stages, where each stage spans two months instead of three months as in Section 3.4.2, and we call them Analysis 3 (without extreme values and $K = 3$) and Analysis 4 (with extreme

values and $K = 3$), respectively.

Table 3.4 presents the results in the same manner as for Table 3.2. Similar to those in Section 3.4.2, including or excluding extreme values in the analysis does have a noticeable impact on the results. These two analyses, however, do uncover some common findings, as reported below.

For stage 3, both Analyses 3 and 4 indicate that "senior-prop", "care-score", and the interaction term $\text{Action}_1 \times$ "popu-density" are significant predictors. In terms of differences between the two analyses, we find that the interaction term $\text{Action}_1 \times$ "government-eff" is suggested to be significant by Analysis 3 but not by Analysis 4.

Regarding the results for stage 2 with DB CI considered, both Analyses 3 and 4 find that "senior-prop", "GDP", "obesity-prev", and "care-score" are significant predictors. The two analyses unveil different findings as well. "Substance-prev" and $\text{Action}_1 \times$ "civic-score" are identified as significant predictors by Analysis 3 but not by Analysis 4, whereas $\text{Action}_1 \times$ "*CFR*-enter" is identified as significant by Analysis 4 but not by Analysis 3.

Finally, using the DB CIs, for stage 1, we see that both Analyses 3 and 4 suggest the statistical significance for "senior-prop", "GDP", "obesity-prev", and "care-score", together with the interaction term $\text{Action}_2 \times$ "*CFR*-enter". In terms of differences, Analysis 3 reveals that "substance-prev" and the interaction term $\text{Action}_1 \times$ "government-eff" are statistically significant, yet Analysis 4 suggests the interaction term $\text{Action}_2 \times$ "government-eff" to be statistically significant.

3.5.2 Discrete Outcome with Nonlinear Q-Functions

To assess the effect of different modeling for the Q-functions, we employ a Negative Binomial model with the logarithm link to analyze the data described in Section 3.4.1, for which the outcome variable $\log CFR$ is now replaced by the cumulative number of COVID-19 deaths measured at the end of stage 1 or 2. Let T_k denote the total number of COVID-19 cases at stage k for a country, and take $\log T_k$ as the offset when modeling the Q-functions for $k = 1, 2$. Specifically, we modify (3.6) and (3.7) by replacing their left-hand-side with $\log Q_2(H_2, A_2) - \log T_2$ and $\log Q_1(H_1, A_1) - \log T_1$, respectively, where $Q_k(H_k, A_k)$ is defined as in (3.1) for

$k = 1, 2$, with Y_k representing the cumulative death number at stage k. We call the analysis for the data with the extreme values excluded or included as Analysis 5 or 6.

Table 3.5 presents the results in the same manner as for Table 3.2. The statistical significance of each predictor in stages 2 and 1 is assessed using the W-type CIs and DB CIs at the significance level 5%, respectively. In what follows, we compare the results with those in Table 3.2.

First, we consider stage 2. It is evident that Analysis 1 and Analysis 5 both identify the following covariates to be statistically significant: "senior-prop", "substance-prev", and "care-score", as well as the interaction terms $\text{Action}_2 \times$ "*CFR*-enter", $\text{Action}_1 \times$ "government-eff", $\text{Action}_1 \times$ "popu-density", and $\text{Action}_1 \times$ "civic-score". These findings suggest that countries characterized by a low government effectiveness score, a high population-weighted geometric mean density, and a high civic and social participation score may wish to prioritize facial covering policy over diagnostic testing and contact tracing policies. Countries with a high *CFR* value at the start of stage 2 may be advised to accord higher precedence to health system policies as opposed to containment and closure policies.

The comparison of Analysis 6 to Analysis 2 shows the commonly identified significant predictors, including "senior-prop" and the interaction terms $\text{Action}_1 \times$ "government-eff" and $\text{Action}_1 \times$ "popu-density". These findings highlight that countries characterized by a low government effectiveness score and a high population-weighted geometric mean density may benefit from prioritization of facial covering policy over diagnostic testing and contact tracing policies.

Next, we compare the results for stage 1 using the DB CIs. Both Analyses 5 and 1 find "senior-prop", "GDP", "obesity-prev", "care-score", Action_1, and $\text{Action}_2 \times$ "*CFR*-enter" to be statistically significant. Therefore, countries characterized by a high *CFR* value at the start of the stage may consider prioritizing health system policies over containment and closure policies. Both Analyses 6 and 2 find "GDP" to be statistically significant.

3.6 Discussion

In this chapter, we use the Q-learning method to explore how different COVID-19 preventive policies may be prioritized to lower the *CFR*. Our data analyses suggest that the strictness of preventive policies be tailored to social- and government-related factors to lower *CFR*. Country-level characteristics such as government effectiveness score, population weighted geometric mean density, and civic and social participation score interact with the preventive policies and hence are useful in determining the optimal actions. The determination of the optimal actions also depends on the observed number of deaths at the baseline.

Although the current study provides insight into understanding the effectiveness of different preventive policies, this research has limitations and some issues that warrant further study. The choice of prescriptive variables for implementing the Q-learning method here is dictated by the availability of data. In the analyses here, we consider two actions, Action_1 and Action_2, defined in Section 3.4.1. It is noted that other action variables may be introduced to reflect the availability of other mitigation measures and their combined effects. With different actions introduced, care needs to be taken to interpret the recommended actions, even though the implementation of the Q-learning algorithm can be carried out in the same manner.

As the data size is small with only 175 countries taken as independent study units, we are restricted to explore analyses with more flexible models. If a richer source of data becomes available, it would be interesting to conduct additional in-depth analyses, such as, by including separate actions for each type of policies as well as their interactions.

In our analysis, we take the data collected in the first three months as the baseline measurements for running the Q-learning algorithm. Introducing baseline measurements intends to differentiate different levels of severity of the outbreak for different countries, in addition to assessing their effects on influencing countries to take active steps to curb the virus spread. It is interesting to further evaluate how sensitive the analysis results would be if a shorter or longer time than three months is used to reflect the baseline data.

Although the Q-learning method provides a convenient framework to study the effects of country-specific policies, the validity of the method hinges on the feasibility of the associated assumptions. For example, the *no unmeasured confounders* (NUC) assumption is asso-

ciated with Q-learning, which says that for $k = 1, \cdots, K$ and $j = 1, \cdots, r$, conditional on the observed history H_k, the treatment $A_{k,j}$ is independent of any future covariates or outcomes $\{\mathbf{O}_{k+1}, \cdots, \mathbf{O}_K, Y_k(H_k, \mathbf{A}_k, \mathbf{O}_{k+1})\}$ (Chakraborty and Moodie, 2013). Despite that we hope to let H_{k0} in (3.3) contain all confounding variables, it is cautioned that the violation of the NUC assumption very likely yields erroneous parameter estimates, and thus, invalidating the determination of the optimal actions. The *stable unit treatment value assumption* (SUTVA) is another important assumption, which states that the potential outcome of a subject does not depend on the treatment assigned to other subjects (Chakraborty and Moodie, 2013). Ensuring the SUTVA to hold basically requires preventive measures in a country to not affect the outcome in the neighboring countries. While the implementation of stringent border-closure, travel restriction policies, and mandatory testing at international borders makes the assumption fairly reasonable, departure from SUTVA can yield the identified optimal actions to be suboptimal.

As Q-functions in (3.1) facilitate the stage-dependent conditional expected outcomes, given the history of actions and associated covariates, it is natural to employ regression models to characterize the Q-functions. While linear regression models are commonly used due to their simplicity, such models yield biased results if the linearity assumption is not appropriate. In the analysis in Section 3.4 we impose the log transformation to the outcome variable to help make the linearity assumption approximately true. Alternatively, one may explore more flexible yet more sophisticated models to describe the Q-functions, such as nonlinear models or deep neural networks.

The analysis here tacitly treats each country as an independent study unit without considering possibly mutual influence in taking action to curb the virus spread, which is basically driven by the lack of repeated measurements that can be collected independently. It is important to recognize that the analysis results offer only an approximate picture about the truth by incorporating country-specific characteristics but not inter-countries influential factors.

Table 3.1: Descriptive statistics of the original unscaled variables over stages 1 and 2

Variable name	Stage 1 ($k = 1$)		Stage 2 ($k = 2$)	
	mean	SD	mean	SD
Outcome related variable:				
CFR (log) (Y_k)	0.50	1.00	0.44	0.93
Action of interest:				
Action$_1$ ($A_{k,1}$)[1]	0.42	0.50	0.47	0.50
Action$_2$ ($A_{k,2}$)[2]	0.62	0.49	0.43	0.50
Stage-varying covariate:				
Total number of infectious cases per million people at the start of stage k	287.50	772.81	370.59	571.38
CFR at the start of stage k	4.96	4.44	2.50	2.76
Stage-invariant covariates:				
Population proportion of people aged 65 and above	9.53	6.84		
GDP	22261.94	22102.94		
Government effectiveness score	0.03	0.97		
Obesity prevalence	18.65	8.93		
Smoking prevalence	21.22	9.19		
Substance use prevalence	2137.05	1077.68		
Population weighted geometric mean density	1601.34	4421.29		
Civic and social participation score	39.04	13.92		
Care system quality score	53.63	16.95		

[1] Facial coverings policy vs. diagnostic testing and contact tracing policies together
[2] Containment and closure policies vs. health system policies

Table 3.2: Estimates of regression coefficients and their 95% CIs: Analyses 1 and 2 in Section 3.4.2 with $K = 2$. Results from the two analyses suggesting different types of evidence are highlighted in bold

		Stage 2		Stage 1		
Parameter	Variable	Estimate	W-type CI	Estimate	W-type CI	DB CI
	Analysis 1: without extreme values					
β_{k0}	Intercept	0.871	**(0.137, 1.605)**	0.492	(-0.324, 1.308)	(-1.000, 2.033)
β_{k1}	Population proportion of people aged 65 and above	0.486	(0.116, 0.856)	1.330	(0.889, 1.770)	(0.593, 1.987)
β_{k2}	GDP	-0.138	(-0.473, 0.197)	-0.710	(-1.142, -0.277)	(-1.538, -0.029)
β_{k3}	Government effectiveness score	-0.133	(-0.463, 0.197)	-0.004	(-0.404, 0.396)	(-0.719, 0.759)
β_{k4}	Obesity prevalence	0.364	**(0.012, 0.716)**	1.144	(0.739, 1.550)	(0.508, 1.837)
β_{k5}	Smoking prevalence	0.080	(-0.293, 0.453)	-0.269	(-0.668, 0.130)	(-0.955, 0.386)
β_{k6}	Substance use prevalence	0.457	**(0.134, 0.779)**	0.683	**(0.304, 1.062)**	**(0.174, 1.240)**
β_{k7}	Population weighted geometric mean density	0.311	(-0.198, 0.820)	0.112	**(-0.592, 0.816)**	(-1.025, 1.562)
β_{k8}	Care system quality score	-1.599	(-2.511, -0.688)	-2.252	**(-3.273, -1.232)**	**(-4.227, -0.349)**
$\psi_{k0,1}$	Action_1	0.857	(-0.039, 1.752)	1.284	**(0.294, 2.274)**	**(0.175, 2.338)**
$\psi_{k1,1}$	$\text{Action}_1 \times$ Total number of infectious cases per million people	0.235	(-0.156, 0.626)	-0.480	(-1.487, 0.527)	(-1.208, 0.375)
$\psi_{k2,1}$	$\text{Action}_1 \times$ *CFR*	-0.309	(-1.017, 0.399)	-0.186	(-0.786, 0.414)	(-0.787, 0.320)
$\psi_{k3,1}$	$\text{Action}_1 \times$ Government effectiveness score	0.488	(0.127, 0.850)	0.460	(-0.001, 0.921)	(-0.124, 0.961)
$\psi_{k4,1}$	$\text{Action}_1 \times$ Population weighted geometric mean density	-1.299	(-1.930, -0.668)	-0.182	(-1.694, 1.330)	(-1.620, 1.482)
$\psi_{k5,1}$	$\text{Action}_1 \times$ Civic and social participation score	-0.930	**(-1.610, -0.251)**	-0.852	**(-1.605, -0.099)**	(-1.840, 0.248)
$\psi_{k0,2}$	Action_2	-0.212	(-0.731, 0.306)	-0.304	(-0.797, 0.190)	(-0.903, 0.120)
$\psi_{k1,2}$	$\text{Action}_2 \times$ Total number of infectious cases per million people	-0.122	(-0.524, 0.280)	0.548	(-0.384, 1.480)	(-0.262, 1.612)
$\psi_{k2,2}$	$\text{Action}_2 \times$ *CFR*	0.787	**(0.093, 1.480)**	0.738	**(0.293, 1.183)**	**(0.398, 1.143)**
$\psi_{k3,2}$	$\text{Action}_2 \times$ Government effectiveness score	-0.200	(-0.550, 0.151)	-0.223	(-0.560, 0.114)	(-0.487, 0.113)
$\psi_{k4,2}$	$\text{Action}_2 \times$ Population weighted geometric mean density	0.480	(-0.343, 1.303)	-0.726	(-2.173, 0.720)	(-2.475, 0.589)
	Analysis 2: with extreme values					
β_{k0}	Intercept	0.839	**(-0.092, 1.770)**	0.381	(-0.777, 1.540)	(-1.125, 2.038)
β_{k1}	Population proportion of people aged 65 and above	0.610	(0.177, 1.043)	1.099	(0.535, 1.664)	(0.318, 1.803)
β_{k2}	GDP	-0.229	(-0.630, 0.173)	-1.054	(-1.610, -0.499)	(-1.949, -0.426)
β_{k3}	Government effectiveness score	-0.136	(-0.530, 0.258)	0.194	(-0.356, 0.744)	(-0.552, 1.040)
β_{k4}	Obesity prevalence	0.391	**(-0.034, 0.815)**	1.054	(0.541, 1.568)	(0.389, 1.713)
β_{k5}	Smoking prevalence	0.032	(-0.403, 0.466)	0.006	(-0.520, 0.533)	(-0.694, 0.700)
β_{k6}	Substance use prevalence	0.285	**(-0.110, 0.681)**	0.399	**(-0.103, 0.901)**	**(-0.159, 0.938)**
β_{k7}	Population weighted geometric mean density	0.052	(-0.102, 0.207)	-0.684	**(-0.876, -0.493)**	(-1.612, 0.335)
β_{k8}	Care system quality score	-1.418	(-2.557, -0.279)	-1.240	**(-2.670, 0.190)**	**(-3.109, 0.814)**
$\psi_{k0,1}$	Action_1	0.488	(-0.404, 1.381)	0.177	**(-1.009, 1.363)**	**(-1.677, 1.461)**
$\psi_{k1,1}$	$\text{Action}_1 \times$ Total number of infectious cases per million people	0.027	(-0.459, 0.513)	-0.035	(-0.610, 0.539)	(-0.880, 0.730)
$\psi_{k2,1}$	$\text{Action}_1 \times$ *CFR*	0.289	(-0.234, 0.812)	0.370	(-0.273, 1.013)	(-0.483, 1.448)
$\psi_{k3,1}$	$\text{Action}_1 \times$ Government effectiveness score	0.482	(0.036, 0.928)	-0.122	(-0.698, 0.455)	(-1.146, 0.584)
$\psi_{k4,1}$	$\text{Action}_1 \times$ Population weighted geometric mean density	-0.932	(-1.401, -0.463)	-0.428	(-1.079, 0.224)	(-2.214, 1.943)
$\psi_{k5,1}$	$\text{Action}_1 \times$ Civic and social participation score	-0.678	**(-1.404, 0.047)**	-0.257	**(-1.179, 0.665)**	(-1.431, 0.880)
$\psi_{k0,2}$	Action_2	-0.103	(-0.636, 0.430)	-0.224	(-0.849, 0.401)	(-0.868, 0.620)
$\psi_{k1,2}$	$\text{Action}_2 \times$ Total number of infectious cases per million people	0.057	(-0.445, 0.559)	0.013	(-0.253, 0.279)	(-0.671, 0.828)
$\psi_{k2,2}$	$\text{Action}_2 \times$ *CFR*	0.304	**(-0.156, 0.764)**	0.444	**(-0.135, 1.022)**	**(-0.545, 0.999)**
$\psi_{k3,2}$	$\text{Action}_2 \times$ Government effectiveness score	-0.108	(-0.531, 0.316)	-0.162	(-0.619, 0.296)	(-0.513, 0.229)
$\psi_{k4,2}$	$\text{Action}_2 \times$ Population weighted geometric mean density	0.521	(-0.465, 1.506)	0.046	(-1.060, 1.152)	(-1.745, 1.199)

Table 3.3: Estimated optimal actions for selected countries from Analyses 1 and 2 in Section 3.4.2. Bold entries for Analysis 2 indicate deviations from the corresponding outcomes from Analysis 1

Country	Analysis 1: without extreme values				Analysis 2: with extreme values			
	Action_1		Action_2		Action_1		Action_2	
	Stage 1	Stage 2	Stage 1	Stage 2	Stage 1	Stage 2	Stage 1	Stage 2
	$\hat{A}^{\text{opt}}_{1,1}$	$\hat{A}^{\text{opt}}_{2,1}$	$\hat{A}^{\text{opt}}_{1,2}$	$\hat{A}^{\text{opt}}_{2,2}$	$\hat{A}^{\text{opt}}_{1,1}$	$\hat{A}^{\text{opt}}_{2,1}$	$\hat{A}^{\text{opt}}_{1,2}$	$\hat{A}^{\text{opt}}_{2,2}$
Ghana	1	1	1	0	1	1	1	0
India	0	1	0	0	0	1	0	0
Belgium	0	0	0	0	0	0	0	0
France	0	0	0	0	0	0	0	0
Argentina	0	0	0	0	0	0	0	0
Cameroon	1	1	0	0	**0**	1	0	0
Senegal	0	1	1	0	0	**0**	1	0
Bangladesh	1	1	1	0	1	1	**0**	0
Canada	0	1	0	0	0	**0**	0	0
Denmark	0	0	1	1	**1**	0	1	1
South Korea	0	1	1	0	**1**	1	1	0
United Kingdom	1	1	0	0	**0**	1	0	0
United Arab Emirates	0	0	1	1	**1**	0	1	1
United States	1	0	0	1	1	0	0	**0**
Philippines	1	1	0	0	**0**	1	0	0
Brazil	1	1	0	0	**0**	1	0	0
Norway	0	0	1	1	**1**	0	1	1
Germany	0	1	1	0	**1**	**0**	1	0
Israel	0	1	1	1	**1**	1	1	**0**
Oman	1	0	0	1	1	**1**	**1**	1

Table 3.4: Estimates of regression coefficients and their 95% CIs: Analyses 3 and 4 in Section 3.5.1 with $K = 3$

		Stage 3		Stage 2			Stage 1		
Parameter	Variable	Estimate	W-type CI	Estimate	W-type CI	DB CI	Estimate	W-type CI	DB CI
	Analysis 3: without extreme values								
β_{k0}	Intercept	1.292	(0.465, 2.118)	1.316	(0.540, 2.093)	(-0.126, 2.511)	1.010	(-0.164, 2.184)	(-1.375, 3.128)
β_{k1}	Population proportion of people aged 65 and above	0.596	(0.208, 0.984)	1.337	(0.973, 1.700)	(0.700, 1.838)	2.562	(1.969, 3.155)	(1.626, 3.235)
β_{k2}	GDP	-0.126	(-0.495, 0.243)	-0.831	(-1.255, -0.408)	(-1.798, -0.103)	-0.998	(-1.873, -0.122)	(-2.593, -0.053)
β_{k3}	Government effectiveness score	-0.045	(-0.396, 0.306)	0.019	(-0.337, 0.375)	(-0.480, 0.664)	-0.209	(-0.836, 0.418)	(-1.163, 0.823)
β_{k4}	Obesity prevalence	0.363	(-0.027, 0.754)	1.261	(0.901, 1.622)	(0.682, 2.026)	2.093	(1.499, 2.688)	(1.119, 3.413)
β_{k5}	Smoking prevalence	0.089	(-0.300, 0.478)	0.242	(-0.104, 0.589)	(-0.348, 0.791)	0.048	(-0.489, 0.585)	(-0.711, 0.770)
β_{k6}	Substance use prevalence	0.317	(-0.044, 0.678)	0.735	(0.401, 1.069)	(0.157, 1.260)	1.105	(0.532, 1.679)	(0.283, 1.901)
β_{k7}	Population weighted geometric mean density	0.232	(-0.336, 0.800)	0.918	(0.314, 1.522)	(-0.237, 1.900)	1.330	(0.004, 2.656)	(-0.932, 4.233)
β_{k8}	Care system quality score	-1.976	(-2.989, -0.962)	-3.778	(-4.722, -2.835)	(-5.713, -1.279)	-5.477	(-6.997, -3.957)	(-7.397, -2.001)
$\psi_{k0,1}$	Action_1	0.146	(-0.872, 1.164)	1.044	(0.131, 1.957)	(-0.280, 2.312)	1.318	(-0.191, 2.827)	(-0.052, 2.185)
$\psi_{k1,1}$	$\text{Action}_1 \times$ Total number of infectious cases per million people	-0.004	(-0.258, 0.251)	-0.176	(-0.608, 0.257)	(-0.827, 0.287)	-0.032	(-1.454, 1.389)	(-0.920, 0.871)
$\psi_{k2,1}$	$\text{Action}_1 \times CFR$	0.315	(-0.512, 1.143)	0.658	(0.045, 1.271)	(-0.086, 1.597)	0.198	(-0.842, 1.237)	(-0.631, 0.623)
$\psi_{k3,1}$	$\text{Action}_1 \times$ Government effectiveness score	0.507	(0.066, 0.948)	0.176	(-0.321, 0.674)	(-0.380, 0.597)	0.898	(0.124, 1.671)	(0.088, 1.372)
$\psi_{k4,1}$	$\text{Action}_1 \times$ Population weighted geometric mean density	-0.833	(-1.548, -0.119)	-1.065	(-2.287, 0.157)	(-2.508, 0.531)	-0.970	(-3.815, 1.875)	(-2.189, 1.623)
$\psi_{k5,1}$	$\text{Action}_1 \times$ Civic and social participation score	-0.386	(-1.189, 0.416)	-1.637	(-2.344, -0.930)	(-2.570, -0.585)	-0.935	(-2.103, 0.233)	(-1.699, 0.462)
$\psi_{k0,2}$	Action_2	-0.084	(-0.677, 0.509)	0.042	(-0.495, 0.579)	(-0.804, 0.704)	0.038	(-0.785, 0.861)	(-0.690, 0.426)
$\psi_{k1,2}$	$\text{Action}_2 \times$ Total number of infectious cases per million people	0.107	(-0.309, 0.523)	0.379	(-0.099, 0.856)	(-0.091, 1.102)	-0.471	(-1.913, 0.970)	(-1.511, 0.408)
$\psi_{k2,2}$	$\text{Action}_2 \times CFR$	0.152	(-0.682, 0.986)	0.619	(-0.024, 1.262)	(-0.129, 1.278)	0.530	(-0.342, 1.402)	(0.143, 1.237)
$\psi_{k3,2}$	$\text{Action}_2 \times$ Government effectiveness score	-0.296	(-0.742, 0.150)	0.306	(-0.129, 0.742)	(-0.220, 0.802)	-0.385	(-0.892, 0.123)	(-0.845, 0.126)
$\psi_{k4,2}$	$\text{Action}_2 \times$ Population weighted geometric mean density	0.038	(-0.907, 0.982)	-0.713	(-1.927, 0.501)	(-2.200, 0.807)	-0.429	(-2.857, 1.999)	(-2.481, 1.159)
	Analysis 4: with extreme values								
β_{k0}	Intercept	1.034	(0.020, 2.048)	0.994	(-0.015, 2.003)	(-0.452, 2.429)	0.823	(-0.418, 2.063)	(-1.161, 2.841)
β_{k1}	Population proportion of people aged 65 and above	0.642	(0.181, 1.102)	1.220	(0.751, 1.689)	(0.568, 1.862)	1.976	(1.423, 2.528)	(0.995, 2.991)
β_{k2}	GDP	-0.207	(-0.633, 0.220)	-0.879	(-1.308, -0.449)	(-1.756, -0.184)	-1.499	(-2.053, -0.945)	(-2.688, -0.594)
β_{k3}	Government effectiveness score	-0.126	(-0.550, 0.299)	0.132	(-0.317, 0.582)	(-0.434, 0.830)	0.335	(-0.246, 0.917)	(-0.604, 1.383)
β_{k4}	Obesity prevalence	0.304	(-0.159, 0.767)	1.379	(0.916, 1.841)	(0.621, 2.279)	1.932	(1.398, 2.467)	(0.951, 2.971)
β_{k5}	Smoking prevalence	0.019	(-0.441, 0.480)	0.367	(-0.085, 0.818)	(-0.221, 0.992)	0.121	(-0.400, 0.643)	(-0.859, 1.056)
β_{k6}	Substance use prevalence	0.331	(-0.096, 0.759)	0.458	(0.030, 0.886)	(-0.066, 1.013)	0.766	(0.274, 1.259)	(-0.062, 1.568)
β_{k7}	Population weighted geometric mean density	0.024	(-0.141, 0.189)	-0.305	(-0.463, -0.147)	(-1.548, 0.952)	-0.693	(-0.882, -0.505)	(-2.685, 1.412)
β_{k8}	Care system quality score	-1.582	(-2.815, -0.348)	-3.014	(-4.279, -1.750)	(-5.340, -0.496)	-3.399	(-4.850, -1.948)	(-6.129, -0.330)
$\psi_{k0,1}$	Action_1	-0.018	(-1.042, 1.006)	-0.224	(-1.208, 0.761)	(-1.736, 1.085)	-0.106	(-1.312, 1.101)	(-1.146, 0.953)
$\psi_{k1,1}$	$\text{Action}_1 \times$ Total number of infectious cases per million people	-0.005	(-0.321, 0.311)	0.078	(-0.449, 0.605)	(-0.545, 0.637)	-0.184	(-0.742, 0.374)	(-0.699, 0.437)
$\psi_{k2,1}$	$\text{Action}_1 \times CFR$	0.458	(-0.445, 1.361)	0.877	(0.165, 1.589)	(0.214, 1.788)	-0.196	(-0.896, 0.504)	(-0.789, 0.453)
$\psi_{k3,1}$	$\text{Action}_1 \times$ Government effectiveness score	0.464	(-0.069, 0.996)	-0.067	(-0.537, 0.403)	(-0.632, 0.480)	0.237	(-0.335, 0.809)	(-0.300, 0.814)
$\psi_{k4,1}$	$\text{Action}_1 \times$ Population weighted geometric mean density	-0.552	(-1.075, -0.029)	-0.280	(-1.728, 1.168)	(-1.873, 1.207)	0.633	(-0.873, 2.139)	(-0.587, 2.261)
$\psi_{k5,1}$	$\text{Action}_1 \times$ Civic and social participation score	-0.297	(-1.088, 0.495)	-0.380	(-1.114, 0.353)	(-1.709, 0.960)	0.444	(-0.494, 1.382)	(-0.547, 1.339)
$\psi_{k0,2}$	Action_2	-0.045	(-0.704, 0.615)	0.342	(-0.327, 1.011)	(-0.302, 1.162)	-0.468	(-1.110, 0.174)	(-1.131, 0.147)
$\psi_{k1,2}$	$\text{Action}_2 \times$ Total number of infectious cases per million people	0.027	(-0.475, 0.528)	0.085	(-0.488, 0.657)	(-0.483, 0.588)	0.050	(-0.213, 0.314)	(-0.448, 0.734)
$\psi_{k2,2}$	$\text{Action}_2 \times CFR$	0.110	(-0.810, 1.030)	-0.052	(-0.771, 0.668)	(-1.012, 0.723)	1.048	(0.430, 1.665)	(0.543, 1.571)
$\psi_{k3,2}$	$\text{Action}_2 \times$ Government effectiveness score	-0.204	(-0.745, 0.336)	0.270	(-0.172, 0.713)	(-0.209, 0.782)	-0.582	(-1.065, -0.099)	(-1.163, -0.083)
$\psi_{k4,2}$	$\text{Action}_2 \times$ Population weighted geometric mean density	0.319	(-0.762, 1.399)	-0.292	(-1.762, 1.177)	(-1.926, 1.039)	-0.298	(-1.552, 0.956)	(-1.941, 1.097)

Table 3.5: Estimates of regression coefficients and their 95% CIs: Analyses 5 and 6 in Section 3.5.2 with $K = 2$

		Stage 2		Stage 1		
Parameter	Variable	Estimate	W-type CI	Estimate	W-type CI	DB CI
	Analysis 5: without extreme values					
β_{k0}	Intercept	-3.589	(-4.180, -2.990)	-4.082	(-4.463, -3.701)	(-4.823, -3.430)
β_{k1}	Population proportion of people aged 65 and above	0.599	(0.311, 0.885)	0.706	(0.516, 0.896)	(0.380, 1.023)
β_{k2}	GDP	-0.068	(-0.308, 0.187)	-0.695	(-0.972, -0.418)	(-1.163, -0.219)
β_{k3}	Government effectiveness score	-0.204	(-0.462, 0.055)	0.155	(-0.040, 0.350)	(-0.158, 0.529)
β_{k4}	Obesity prevalence	0.044	(-0.242, 0.329)	0.415	(0.237, 0.593)	(0.103, 0.776)
β_{k5}	Smoking prevalence	0.113	(-0.168, 0.399)	-0.017	(-0.180, 0.149)	(-0.294, 0.259)
β_{k6}	Substance use prevalence	0.394	(0.145, 0.646)	0.250	(0.081, 0.418)	(-0.003, 0.503)
β_{k7}	Population weighted geometric mean density	0.047	(-0.427, 0.570)	-0.341	(-0.761, 0.096)	(-1.159, 0.070)
β_{k8}	Care system quality score	-1.531	(-2.286, -0.772)	-0.981	(-1.484, -0.473)	(-1.854, -0.346)
$\psi_{k0,1}$	Action_1	1.432	(0.683, 2.188)	0.756	(0.304, 1.212)	(0.196, 1.373)
$\psi_{k1,1}$	$\text{Action}_1 \times$ Total number of infectious cases per million people	0.295	(-0.003, 0.605)	0.068	(-0.400, 0.522)	(-0.252, 0.472)
$\psi_{k2,1}$	$\text{Action}_1 \times CFR$	-0.609	(-1.191, -0.046)	-0.200	(-0.493, 0.091)	(-0.558, 0.260)
$\psi_{k3,1}$	$\text{Action}_1 \times$ Government effectiveness score	0.494	(0.208, 0.792)	-0.051	(-0.269, 0.170)	(-0.348, 0.275)
$\psi_{k4,1}$	$\text{Action}_1 \times$ Population weighted geometric mean density	-1.200	(-1.970, -0.544)	-0.342	(-1.070, 0.375)	(-1.060, 0.335)
$\psi_{k5,1}$	$\text{Action}_1 \times$ Civic and social participation score	-1.319	(-1.891, -0.746)	-0.591	(-0.921, -0.259)	(-1.115, -0.120)
$\psi_{k0,2}$	Action_2	-0.393	(-0.814, 0.026)	-0.277	(-0.523, -0.029)	(-0.525, -0.038)
$\psi_{k1,2}$	$\text{Action}_2 \times$ Total number of infectious cases per million people	0.002	(-0.290, 0.316)	0.223	(-0.225, 0.687)	(-0.097, 0.541)
$\psi_{k2,2}$	$\text{Action}_2 \times CFR$	0.964	(0.424, 1.535)	0.526	(0.266, 0.791)	(0.165, 0.823)
$\psi_{k3,2}$	$\text{Action}_2 \times$ Government effectiveness score	-0.083	(-0.359, 0.198)	-0.074	(-0.231, 0.086)	(-0.232, 0.074)
$\psi_{k4,2}$	$\text{Action}_2 \times$ Population weighted geometric mean density	0.557	(-0.117, 1.332)	-0.164	(-0.887, 0.557)	(-0.752, 0.534)
	Analysis 6: with extreme values					
β_{k0}	Intercept	-3.532	(-4.354, -2.694)	-3.939	(-4.521, -3.357)	(-4.551, -3.248)
β_{k1}	Population proportion of people aged 65 and above	0.395	(0.020, 0.767)	0.322	(0.049, 0.598)	(-0.136, 0.686)
β_{k2}	GDP	-0.168	(-0.484, 0.177)	-0.603	(-0.835, -0.356)	(-1.285, -0.163)
β_{k3}	Government effectiveness score	-0.314	(-0.649, 0.020)	0.108	(-0.159, 0.375)	(-0.259, 0.602)
β_{k4}	Obesity prevalence	0.298	(-0.033, 0.628)	0.217	(-0.013, 0.447)	(-0.612, 0.764)
β_{k5}	Smoking prevalence	-0.104	(-0.470, 0.272)	0.001	(-0.244, 0.252)	(-0.286, 0.303)
β_{k6}	Substance use prevalence	0.270	(-0.072, 0.614)	0.047	(-0.184, 0.280)	(-0.388, 0.306)
β_{k7}	Population weighted geometric mean density	0.054	(-0.061, 0.220)	0.014	(-0.069, 0.119)	(-0.683, 0.347)
β_{k8}	Care system quality score	-0.959	(-1.992, 0.071)	-0.120	(-0.803, 0.569)	(-1.486, 1.967)
$\psi_{k0,1}$	Action_1	0.054	(-0.682, 0.809)	0.061	(-0.474, 0.604)	(-0.421, 0.569)
$\psi_{k1,1}$	$\text{Action}_1 \times$ Total number of infectious cases per million people	-0.011	(-0.376, 0.371)	0.112	(-0.151, 0.393)	(-0.339, 0.459)
$\psi_{k2,1}$	$\text{Action}_1 \times CFR$	0.457	(-0.004, 0.918)	0.269	(-0.017, 0.562)	(-0.044, 0.662)
$\psi_{k3,1}$	$\text{Action}_1 \times$ Government effectiveness score	0.370	(0.010, 0.740)	-0.118	(-0.390, 0.160)	(-0.560, 0.191)
$\psi_{k4,1}$	$\text{Action}_1 \times$ Population weighted geometric mean density	-0.846	(-1.438, -0.340)	-0.642	(-0.983, -0.295)	(-1.642, 0.255)
$\psi_{k5,1}$	$\text{Action}_1 \times$ Civic and social participation score	-0.463	(-1.050, 0.131)	-0.200	(-0.618, 0.225)	(-0.785, 0.233)
$\psi_{k0,2}$	Action_2	0.055	(-0.415, 0.533)	-0.277	(-0.568, 0.017)	(-0.803, 0.069)
$\psi_{k1,2}$	$\text{Action}_2 \times$ Total number of infectious cases per million people	-0.158	(-0.520, 0.243)	-0.010	(-0.129, 0.128)	(-0.414, 0.299)
$\psi_{k2,2}$	$\text{Action}_2 \times CFR$	0.176	(-0.245, 0.641)	0.308	(0.041, 0.576)	(-0.007, 0.679)
$\psi_{k3,2}$	$\text{Action}_2 \times$ Government effectiveness score	0.004	(-0.339, 0.362)	-0.053	(-0.278, 0.176)	(-0.297, 0.185)
$\psi_{k4,2}$	$\text{Action}_2 \times$ Population weighted geometric mean density	0.411	(-0.402, 1.309)	-0.114	(-0.637, 0.428)	(-0.771, 0.454)

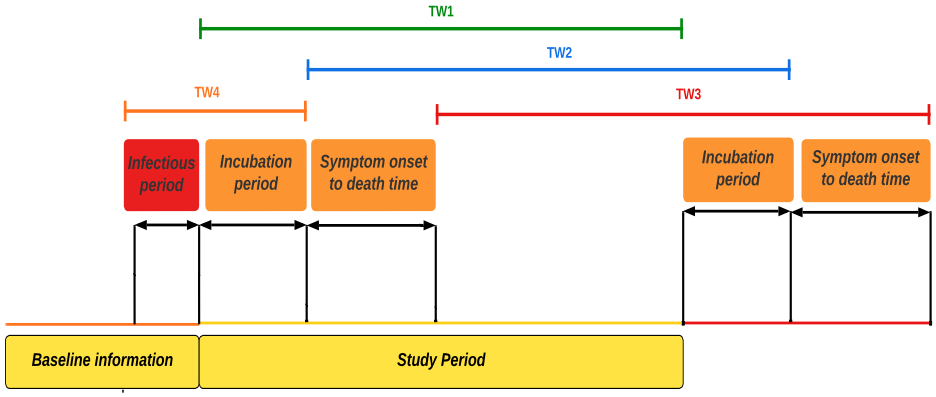

Figure 3.1: The timeline for the determination of the relevant quantities

Chapter 4

Accommodating Misclassification Effects on Optimizing Dynamic Treatment Regimes with Q-Learning

4.1 Introduction

While both direct and indirect modeling strategies have been broadly employed in the area of personalized medicine, those methods are vulnerable to the violation of the critical condition that variables must be precisely measured. This assumption is, however, often not true in the applications, as discussed extensively in the literature, including monographs Carroll et al. (2006), Yi (2017), Yi et al. (2021), and the references therein. Mismeasurement or measurement error is commonly encountered in applications.

Although extensive research has been conducted for regression models with mismeasured variables, there has been limited work on DTRs with error-contaminated data, except for Spicker and Wallace (2020) who investigated the consequences of ignoring covariate measurement error in the context of DTRs. In this chapter, we consider the problem of misclassification in binary covariates with Q-learning.

To highlight the idea, we consider settings where a binary variable is subject to misclassification. This research is partially motivated by the nature of data arising from a multi-level

randomized controlled study of human major depressive disorder (MDD) (Chakraborty et al., 2013), which was designed to evaluate the effectiveness of different treatment regimes on MDD that had 4 levels. In each level, patients were treated by one or a combination of different treatment options for depression. Severity of depression was evaluated using the clinician-rated and self-report versions of the Quick Inventory of Depressive Symptomatology (QIDS) scores. Receipt of a specific treatment option at levels 2, 3, and 4 was driven by the doctor's recommendation as well as the patient's opinion, reflected by the variable termed *patient's preference* to switch or augment his/her previous treatment option. However, it is difficult to precisely record the true value of this variable because of its dependence on the doctor's experience, the level of the patient's trust in the doctor's recommendation, the effectiveness of the communication between the doctor and the patient, and the patient's own knowledge, etc. Challenges in measuring patient's preference in medical contexts have been previously discussed (e.g., Mühlbacher et al., 2016; Soekhai et al., 2019; Janssens et al., 2019). While it is of common interest to develop optimal DTR for patients with MDD (Chakraborty and Moodie, 2013; Chakraborty et al., 2013), ignoring misclassification feature in patient's preference may yield seriously biased results in determining the optimal treatment.

Research about misclassification effects has been studied for various settings, including regression analysis models (e.g., Akazawa et al., 1998; Küchenhoff et al., 2006), survival analysis (e.g., Bang et al., 2013; Yi et al., 2018; Zucker and Spiegelman, 2004), and causal inference (e.g., Greenland, 1988; Kleinbaum et al., 1991; Rothman and Greenland, 1998). However, there has been little work on estimating optimal treatment regimes in the presence of covariate misclassification. Correcting misclassification effects can be more complicated than available work due to the stage interconnectivity in Q-learning.

In this chapter, we investigate how Q-learning may be impacted by misclassified variables in randomized treatment settings, and we present two procedures to ameliorate the bias induced by covariate misclassification. Our research differs from Spicker and Wallace (2020) who considered continuous covariates subject to measurement error in DTRs. While the goal of studying the mismeasurement effects is the same for both error-prone continuous covariates and misclassification-contaminated binary covariates, the inherent differences between continuous and discrete variables make the technical development distinct. Capitalizing on the unique

feature of binary variables, we develop a new method based on using estimating function theory to account for the misclassification effects, in addition to exploring the use of regression calibration explored by Spicker and Wallace (2020) within the dWOLS framework. The former method has the appeal of yielding consistent estimators for parameters involved with the last stage model, yet the latter approach does not possess this property. Unlike that Spicker and Wallace (2020) considered the setting with instrumental data, we examine instances in the presence of validation data to characterize the misclassification process.

The remainder of the chapter is organized as follows. In Section 4.2, we introduce the setting with covariate misclassification and conduct simulation studies to demonstrate misclassification effects on Q-learning. In Sections 4.3 and 4.4, we present two correction methods for debiasing misclassification effects. In Section 4.5, we apply the procedures described in Sections 4.3 and 4.4 to simulated data as well as real data. In Section 4.6, we provide some concluding remarks.

4.2 Misclassification and Naive Analysis

4.2.1 Misclassification and Assumptions

The validity of the Q-learning procedure hinges on the requirement of precisely measured variables as well as the associated conditions such as the SUTVA and the NUC assumption. The SUTVA posits that an individual's outcome is unaffected by the treatment allocation for other individuals, and the NUC assumption says that for $k = 1, \cdots, K$, conditional on the observed history $\{\overline{A}_{k-1}, \overline{X}_k, \overline{Z}_k\}$, the treatment A_k is independent of any future covariates or outcomes $\{X_{k+1}, \cdots, X_K; Z_{k+1}, \cdots, Z_K; Y_k\}$ (Chakraborty and Moodie, 2013). In the presence of error-contaminated data, directly applying the Q-learning algorithm to the observed data, called the naive method, may yield seriously biased results.

To demonstrate this, we consider a simple but illustrative case with only the binary covariate X_k subject to misclassification and other variables being precisely measured. For $k = 1, \cdots, K$, let X_k^* denote the observed version of the true covariate X_k, and let $\overline{X}_k^* = \left\{X_1^*, \cdots, X_k^*\right\}$. While the SUTVA may not be affected by the presence of error-prone covariates, the NUC assump-

tion does not necessarily hold for the observed surrogate measurements. That is, conditional on $\{\overline{A}_{k-1}, \overline{X}^*_k, \overline{Z}_k\}$, A_k is not necessarily independent of future variables $\{X^*_{k+1}, \cdots, X^*_K; Z_{k+1}, \cdots, Z_K; Y_k\}$. Consequently, the naive method may yield biased results.

In the presence of X^*_k, inference about the relationship between $\{Y_k : k = 1, \cdots, K\}$ and $\left\{\{\overline{A}_k, \overline{X}_{k+1}, \overline{Z}_{k+1}\} : k = 1, \cdots, K\right\}$, described by (1.5), becomes more complicated, as it roots in the joint distribution for $\{Y_k : k = 1, \cdots, K\}$ and $\left\{\{A_k, X_k, Z_k, X^*_k\} : k = 1, \cdots, K\right\}$, which is proportional to

$$\begin{aligned} &\Big\{\prod_{k=1}^{K} h(Y_k \mid \overline{Y}_{k-1}, \overline{A}_K, \overline{X}_K, \overline{Z}_K, \overline{X}^*_K)\Big\} h(\overline{X}^*_K, \overline{X}_K, A_K, Z_K) \\ &\propto \Big\{\prod_{k=1}^{K} h(Y_k \mid \overline{Y}_{k-1}, \overline{A}_K, \overline{X}_K, \overline{Z}_K, \overline{X}^*_K)\Big\} \cdot \Big\{\prod_{k=1}^{K} h(X^*_k \mid \overline{X}^*_{k-1}, \overline{X}_K, \overline{A}_K, \overline{Z}_K)\Big\} \end{aligned} \tag{4.1}$$

with $\overline{X}^*_0$ being null, where $h(Y_1 \mid \overline{A}_K, \overline{X}_K, \overline{Z}_K, \overline{X}^*_K)$ and $h(X^*_1 \mid \overline{X}_K, \overline{A}_K, \overline{Z}_K)$ are omitted. In (4.1), $\overline{Y}_{k-1} = \{Y_1, \cdots, Y_{k-1}\}$ for $k = 1, \cdots, K$ with $\overline{Y}_0$ being null; $h(\cdot \mid \cdot)$ represents the conditional probability density or mass function for the random variables indicated by the corresponding arguments. Here for ease of exposition, we use upper case letters to represent both random variables and their realized values.

To link (4.1) to the framework (1.4) together with (1.5), we assume that for $k = K, \cdots, 1$,

$$h(Y_k \mid \overline{Y}_{k-1}, \overline{A}_K, \overline{X}_K, \overline{Z}_K, \overline{X}^*_K) = h(Y_k \mid \overline{A}_k, \overline{X}_{k+1}, \overline{Z}_{k+1}), \tag{4.2}$$

which says that at each stage k, the conditional distribution of Y_k, given the history of outcomes $\overline{Y}_{k-1}$ and the information $\overline{A}_K \cup \overline{X}_K \cup \overline{Z}_K \cup \overline{X}^*_K$ for the entire course, depends only on the history $\overline{A}_k \cup \overline{X}_k \cup \overline{Z}_k$ at stage k as well as the covariates X_{k+1} and Z_{k+1} at the next stage. This assumption is in line with that for the Q-learning method in the misclassification-free context, and it allows us to use the framework (1.5) to study stage-dependent outcome characterized by (1.3). Assumption (4.2) implies that the surrogates $\overline{X}^*_K$ carry no additional information for conducting inference about the response Y_k if the true covariates $\overline{X}_k$ are given. This independence assumption is similar to the nondifferential misclassification mechanism (Yi, 2017, p.50) that is commonly made in the literature of measurement error models with a univariate outcome. It

allows us to conduct inference about the true variables using their surrogate measurements.

4.2.2 Naive Q-Learning Procedure

In the circumstances where the true value of X_k is not observed but only its surrogate value X_k^* is available, it is tempting to still use the naive method by simply repeating the implementation of the Q-learning procedure in Section (1.2.2) with the feature of misclassification in X_k ignored, i.e., replacing X_k with X_k^*. In doing so, we may respectively define the naive Q-functions for stages K and k with $k = K-1, \cdots, 1$ as

$$Q_K^*(\overline{A}_K, \overline{X}_K^*, \overline{Z}_K) = E(Y_K \mid \overline{A}_K, \overline{X}_K^*, \overline{Z}_K);$$
$$Q_k^*(\overline{A}_k, \overline{X}_k^*, \overline{Z}_k) = E\{\hat{Y}_k^* \mid \overline{A}_k, \overline{X}_k^*, \overline{Z}_k\};$$

where $\hat{Y}_k^* \triangleq Y_k + \max_{a_{k+1}} Q_{k+1}^*(\overline{A}_k, \overline{X}_{k+1}^*, \overline{Z}_{k+1}, a_{k+1})$ is taken as a naive pseudo-outcome at stage k. Then we may naively use a counterpart regression model of (1.9) to characterize the $Q_k^*(\overline{A}_k, \overline{X}_k^*, \overline{Z}_k)$ as:

$$Q_k^*(\overline{A}_k, \overline{X}_k^*, \overline{Z}_k) = \beta_k^{*\mathrm{T}} H_{k0}^* + (\psi_k^{*\mathrm{T}} H_{k1}^*) A_k \qquad \text{for} \qquad k = K, \cdots, 1, \tag{4.3}$$

where H_{k0}^* and H_{k1}^* are, respectively, the counterparts of H_{k0} and H_{k1} with X_k replaced by X_k^*, and $\theta_k^* \triangleq (\beta_k^{*\mathrm{T}}, \psi_k^{*\mathrm{T}})^{\mathrm{T}}$ is the vector of the associated regression coefficients which may differ from θ_k in model (1.9).

Consequently, the naive Q-learning algorithm may now be carried out to the observed data, $\mathcal{D}^* \triangleq \left\{\{A_{ki}, X_{ki}^*, Z_{ki}, Y_{ki}\} : k = 1, \cdots, K;\ i = 1, \cdots, n\right\}$, which differs from $\mathcal{D}$ in Section (1.2.2) in the availability of X_{ki}, where X_{ki}^* is the observed version of X_k for subject i. First, regression coefficients for each stage are estimated backward by

$$\hat{\theta}_k^* = \arg\min_{\theta_k^*} \frac{1}{n} \sum_{i=1}^{n} \left[\hat{Y}_{ki}^* - Q_k^*(\overline{A}_{ki}, \overline{X}_{ki}^*, \overline{Z}_{ki}; \theta_k^*)\right]^2$$

for $k = K, \cdots, 1$, where $\hat{Y}_{Ki}^* = Y_{Ki}$; and for $k = K-1, \cdots, 1$, $\hat{Y}_{ki}^* = Y_{ki} + \max_{a_{k+1}} Q_{k+1}^*(\overline{A}_{ki}, \overline{X}_{(k+1)i}^*, \overline{Z}_{(k+1)i}, a_{k+1}; \hat{\theta}_{k+1}^*)$. Then using the naive estimates, we estimate the optimal treatment naively

by

$$\hat{d}_k^* = \arg\max_{a_k} Q_k^*(\overline{A}_{k-1}, \overline{X}_k^*, \overline{Z}_k, a_k; \hat{\theta}_k^*) \qquad \text{for} \qquad k = K, \cdots, 1,$$

where $Q_k^*(\overline{A}_{k-1}, \overline{X}_k^*, \overline{Z}_k, a_k; \hat{\theta}_k^*)$ is determined by (4.3) with θ_k^* replaced by $\hat{\theta}_k^*$.

4.2.3 Simulation Studies

Here we conduct numerical studies to illustrate the misclassification effects on determining optimal DTR under a randomized treatment setting, where we set $K = 2$.

First, the binary treatments A_1 and A_2 are generated independently from the Bernoulli distribution, Bernoulli(0.5). Error-free covariates Z_1 and Z_2 are independently generated by $Z_1 \sim$ Bernoulli(0.5) and $Z_2 \sim$ Bernoulli(0.5). Error-prone binary covariate X_1 is independently generated by $X_1 \sim$ Bernoulli(0.5), and error-prone binary covariate X_2 is generated from the conditional distribution $X_2 \mid A_1, Z_1 \sim$ Bernoulli$(\frac{\exp(\delta_1 A_1 + \delta_2 Z_1)}{1+\exp(\delta_1 A_1 + \delta_2 Z_1)})$, where $\delta_1 = \delta_2 = 0.01$.

The responses Y_2 and Y_1 are respectively generated from

$$Y_2 = \mu_2 + \epsilon_2 \text{ and } Y_1 = \mu_1 + \epsilon_1, \tag{4.4}$$

where

$$\mu_2 = \eta_0 + \eta_1 Z_2 + \eta_2 A_1 + \eta_3 A_2 + \eta_4 X_2 A_2 + \eta_5 A_1 A_2,$$

$$\mu_1 = \gamma_0 + \gamma_1 Z_1 + \gamma_2 X_1 + \gamma_3 A_1 + \gamma_4 X_1 A_1,$$

and ϵ_2 and ϵ_1 are independently generated from $\mathcal{N}(0, 1)$.

We consider three parameter settings to reflect possible occurrence of non-regular and weak non-regular issues, in addition to a regular case. If the issues noted at the end of Section 1.2.2 are present, the occurrence of non-regularity or weak non-regularity is pertinent to the determination of the optimal treatment for the second stage, given the setup considered here. With reference to the present generative model μ_2, the optimal treatment determination in the second stage hinges on the coefficient of A_2, i.e., the linear combination $\eta_3 + \eta_4 X_2 + \eta_5 A_1$. Depending on the values of X_2 and A_1, the linear combination assumes a value of $\eta_3 + \eta_4 + \eta_5$, $\eta_3 + \eta_4$,

$\eta_3 + \eta_5$, and η_3, each with a positive probability. Therefore, setting different values of the coefficients $\{\eta_3, \eta_4, \eta_5\}$ in the model for μ_2, we enable the linear combination $\eta_3 + \eta_4 X_2 + \eta_5 A_1$ to take on value 0 or a value near zero with positive probabilities, thus leading to non-regularity or weak non-regularity issues. Specifically, we consider the following settings. Setting 1 sets $(\eta_0, \eta_1, \eta_2, \eta_3, \eta_4, \eta_5)^{\mathrm{T}} = (0, 0, 0.85, 4, -5, 0.05)^{\mathrm{T}}$, with $(\gamma_0, \gamma_1, \gamma_2, \gamma_3, \gamma_4)^{\mathrm{T}} = (0, 0, 0, -1, 0.8)^{\mathrm{T}}$; Setting 2 takes $(\eta_0, \eta_1, \eta_2, \eta_3, \eta_4, \eta_5)^{\mathrm{T}} = (1.5, 0.8, 0.3, -2, 2.3, 0.5)^{\mathrm{T}}$, with $(\gamma_0, \gamma_1, \gamma_2, \gamma_3, \gamma_4)^{\mathrm{T}} = (0.5, 0.5, -0.75, -1.8, 1.4)^{\mathrm{T}}$; and Setting 3 specifies $(\eta_0, \eta_1, \eta_2, \eta_3, \eta_4, \eta_5)^{\mathrm{T}} = (1.5, 0.8, 0.3, -2, 2, 1)^{\mathrm{T}}$, with $(\gamma_0, \gamma_1, \gamma_2, \gamma_3, \gamma_4)^{\mathrm{T}} = (0.5, 0.5, -0.75, -1.8, 1.4)^{\mathrm{T}}$. Setting 1 ensures $\eta_3 + \eta_4 X_2 + \eta_5 A_1 \neq 0$, with a positive probability, yielding a regular setting. In contrast, in Setting 3, $\eta_3 + \eta_4 X_2 + \eta_5 A_1 = 0$ with a positive probability, rendering it a non-regular scenario; and in Setting 2, $\eta_3 + \eta_4 X_2 + \eta_5 A_1$ assumes values relatively close to zero with a non-negligible probability, showing a weak non-regular scenario. We refer to Settings 1-3 as "regular", "weak non-regular", and "non-regular", respectively.

When employing the Q-learning procedure in Section 1.2.2, model (1.9) is now specified as

$$Q_2(\overline{X}_2, \overline{Z}_2, \overline{A}_2) = \beta_{02} + \beta_{12} Z_2 + \beta_{22} A_1 + (\psi_{02} + \psi_{12} X_2 + \psi_{22} A_1) A_2, \tag{4.5}$$

and

$$Q_1(X_1, Z_1, A_1) = \beta_{01} + \beta_{11} Z_1 + \beta_{21} X_1 + (\psi_{01} + \psi_{11} X_1) A_1, \tag{4.6}$$

where β_{k2}, β_{k1}, and ψ_{k2} are regression parameters for $k = 0, 1, 2$, together with ψ_{k1} for $k = 0, 1$. The optimal DTR is given by the decision rules:

$$d_2 = \mathrm{sign}(\psi_{02} + \psi_{12} X_2 + \psi_{22} A_1);\ d_1 = \mathrm{sign}(\psi_{01} + \psi_{11} X_1), \tag{4.7}$$

where $\mathrm{sign}(t) = 1$ if $t > 0$, and 0 otherwise.

Next, we consider misclassification probabilities. To simplify the exposition of the devel-

opment, assume that

$$h(X_k^* \mid \overline{X}_{k-1}^*, \overline{A}_K, \overline{X}_K, \overline{Z}_K) = h(X_k^* \mid X_k) \quad \text{for } k = 1, \cdots, K,$$

which enables us to directly use a common misclassification matrix, say,

$$\Pi = \begin{pmatrix} 1 - \pi_{10} & \pi_{01} \\ \pi_{10} & 1 - \pi_{01} \end{pmatrix} \tag{4.8}$$

to facilitate the misclassification degree for different stages with $k = 1, \cdots, K$, where $\pi_{jl} = P(X_k^* = j \mid X_k = l)$ for $j = 0, 1$, and $l = 1 - j$. This misclassification is commonly considered in applications (e.g., Yi, 2017; Yi et al., 2018; Akazawa et al., 1998). We use the ***misclass*** function provided in the ***simex*** package in *R* to generate surrogate values X_k^* of X_k with $k = 1, 2$. We consider eight settings for misclassification probabilities, with $(\pi_{10}, \pi_{01})^{\mathrm{T}} = (0, 0.15)^{\mathrm{T}}, (0, 0.3)^{\mathrm{T}}, (0.15, 0)^{\mathrm{T}}, (0.15, 0.15)^{\mathrm{T}}, (0.15, 0.3)^{\mathrm{T}}, (0.3, 0)^{\mathrm{T}}, (0.3, 0.15)^{\mathrm{T}}$, or $(0.3, 0.3)^{\mathrm{T}}$.

To run simulations, we use the proceeding models to generate data of size $n = 1000$, and we repeat 1000 simulations for each parameter configuration. We implement the Q-learning algorithm in Section 4.2.2 to the observed data $\{Z_1, X_1^*, A_1, Y_1, Z_2, X_2^*, A_2, Y_2\}$, called the "naive method", as opposed to the Q-learning procedure in Section 1.2.2 to the true measurements $\{Z_1, X_1, A_1, Y_1, Z_2, X_2, A_2, Y_2\}$, called the "error-free least squares" (EFLS) method.

To evaluate the impacts of misclassification on stage 1 parameter estimation, we define the true values of the population-level parameters in $\theta_1 = (\beta_{01}, \beta_{11}, \beta_{21}, \psi_{01}, \psi_{11})^{\mathrm{T}}$ for the first stage Q-function (4.6) as

$$\arg\min_{\theta_1} E\Big[\{Y_1 + \max_{a_2} Q_2(A_1, X_2, Z_2, a_2; \theta_2) - Q_1(X_1, Z_1, A_1; \theta_1)\}^2\Big], \tag{4.9}$$

with $\theta_2 = (\beta_{02}, \beta_{12}, \beta_{22}, \psi_{02}, \psi_{12}, \psi_{22})^{\mathrm{T}}$ in (4.5) kept as the true values for data generation, where the expectation is taken with respect to the distribution of $\{Z_1, X_1, A_1, Y_1, Z_2, X_2\}$ (Laber et al.,

2014). In particular, by (1.5) and (4.4), at the end stage (i.e., stage 2),

$$\begin{aligned} Q_2(\overline{X}_2, \overline{Z}_2, \overline{A}_2) &= E(Y_2 \mid \overline{X}_2, \overline{Z}_2, \overline{A}_2) \\ &= E(\mu_2 + \epsilon \mid \overline{X}_2, \overline{Z}_2, \overline{A}_2) \\ &= \mu_2, \end{aligned}$$

showing that the true value of θ_2 equals that of $(\eta_0, \eta_1, \eta_2, \eta_3, \eta_4, \eta_5)^{\mathrm{T}}$, specified for each of Settings 1-3 described in Section 4.2.3. Solutions to (4.9) can then be obtained using numerical methods, which allow us to calculate the bias, MSE, and CR estimators for the first stage Q-function parameters described below.

In Tables 4.1-4.2, we report the estimation results for the parameters in the stage 1 model (4.6) and stage 2 model (4.5), respectively. In the tables, "Bias" stands for the difference between the true parameter values and the average of their estimates over 1000 simulations obtained from the EFLS or naive methods, with the true parameter values for the stage 1 model (4.6) determined by (4.9), and with the true parameter values for the stage 2 model (4.5) set as those used in generating data in Section 4.2.3; "SE" represents the average of model-based standard errors (i.e., the standard errors of the least squares parameter estimators over 1000 simulations); "ESE" shows the empirical standard error of the estimates, calculated by $\sqrt{\frac{1}{999}\sum_{k=1}^{1000}(\text{Est.}_k - \overline{\text{Est.}})^2}$ with Est._k representing the estimate for the kth simulation and $\overline{\text{Est.}} = 1000^{-1}\sum_{k=1}^{1000}\text{Est.}_k$; "MSE" displays the mean squared error of the estimates given by $\text{Bias}^2 + \text{SE}^2$; "WTCR" represents the CR of 95% W-type CIs with the model-based standard errors used; "PBCR" represents the CR of 95% PB CIs; and "DBCR" represents the CR of 95% DB CIs. The PB CIs are created based on 1000 bootstrap iterations, and the DB CIs are derived from using 1000 first-stage and 100 second-stage bootstrap iterations, with a reduced number of iterations for the second stage to ease the computational burden.

The results in Table 4.1 suggest that the EFLS method performs well under the regular setting, with small finite sample biases and MSEs as well as reasonable CRs of 95% CIs for both stages 1 and 2; however, under the weak non-regular and non-regular settings, the stage 1 parameter estimates obtained from the EFLS method may incur large biases and poor CRs, except for those derived from double bootstrapping. The W-type CIs (not reported here) incur

severe under-coverage for weak non-regular and non-regular settings.

The results demonstrate the poor performance of the naive method for both stages, reflected by those considerable biases, high MSEs, and unacceptably low CRs. Notably, the induced bias from the covariate misclassification intensifies as the misclassification probabilities increase, and the estimates associated with misclassified covariates suffer a greater impact than those for error-free covariates. These findings emphasize the importance of addressing misclassification effects and developing appropriate correction methods to enhance the accuracy of the estimations.

4.3 Addressing Misclassification Effects: Regression Calibration

Section 4.2.3 demonstrates numerically that in the presence of misclassification, biased results can be produced if the Q-learning procedure is naively implemented with misclassification effects ignored. In this and the next sections, we explore two methods to mitigate the bias due to misclassification in covariates. The goal is basically to optimize the treatment allocation by addressing the misclassification effects on estimating regression coefficients in Q-function at each stage.

We first describe the calibrated Q-learning algorithm by applying the regression calibration method which was initiated by Prentice (1982) for survival analysis with mismeasurement in covariates, where the nondifferential measurement error mechanism is assumed. The basic idea is to employ the usual Q-learning algorithm in Section 1.2.2 with the unobserved true measurement X_k replaced by its conditional expectation, given its surrogate and other variables.

With the observed data $\mathcal{D}^*$, let $X_{ki}^{**} = E(X_{ki} \mid \overline{A}_{(k-1)i}, \overline{X}_{ki}^*, \overline{Z}_{ki})$ for $k = 1, \cdots, K$. The calibrated Q-learning algorithm consists of the following two steps:

Step 1: Repeat the Q-learning algorithm described in Section 1.2.2, with X_{ki} replaced by X_{ki}^{**} or its estimate to obtain point estimates of the regression parameters for each stage.

Step 2: Implement the bootstrap method to obtain standard errors associated with the estimates of the parameters for each stage.

The implementation of the calibrated Q-learning algorithm requires the determination of X_{ki}^{**} for $i = 1, \cdots, n$, which essentially roots in the delineation of the conditional probabilities $P(X_{ki} = 1 \mid \overline{A}_{(k-1)i}, \overline{X}_{ki}^{*}, \overline{Z}_{ki})$ for $k = 1, \cdots, K$. A parametric model, such as the logistic regression model, may be employed for this purpose. To be specific, let $\pi_i^{(k)} = P(X_{ki} = 1 \mid \overline{A}_{(k-1)i}, \overline{X}_{ki}^{*}, \overline{Z}_{ki})$, and we consider the regression model

$$\operatorname{logit} \pi_i^{(k)} = m_k(\overline{A}_{(k-1)i}, \overline{X}_{ki}^{*}, \overline{Z}_{ki}; \boldsymbol{\zeta}_k) \qquad \text{for} \qquad k = 1, \cdots, K, \tag{4.10}$$

where $m_k(\cdot)$ is a specified function, and $\boldsymbol{\zeta}_k$ is the associated parameter vector.

With the availability of internal validation data, denoted $\mathcal{D}_V = \left\{\{A_{ki}, X_{ki}, X_{ki}^{*}, Z_{ki}, Y_{ki}\} : k = 1, \cdots, K;\ i \in \mathcal{V}\right\}$, in addition to the main study data, $\mathcal{D}^{*}$, where $\mathcal{V}$ is a subset of $\mathcal{M} \triangleq \{1, \cdots, n\}$, we may apply the validation data $\mathcal{D}_V$ to model (4.10) to obtain estimates, denoted $\hat{\boldsymbol{\zeta}}_k$, of the model parameter $\boldsymbol{\zeta}_k$, and hence, obtaining the estimate, denoted $\hat{X}_{ki}^{**}$, of X_{ki}^{**} for $i \in \mathcal{M} \setminus \mathcal{V}$. In this case, the implementation of the calibrated Q-learning algorithm can be modified using the measurements of X_{ki} for $i \in \mathcal{V}$ as well as $\hat{X}_{ki}^{**}$ for $i \in \mathcal{M} \setminus \mathcal{V}$.

Let $\hat{\theta}_{kRC} = (\hat{\beta}_{kRC}^{\mathrm{T}}, \hat{\psi}_{kRC}^{\mathrm{T}})^{\mathrm{T}}$ denote the resulting estimate of θ_k for $k = 1, \cdots, K$. Consequently, for $k = K, \cdots, 1$, the optimal treatment is determined by modifying (1.12) with $\hat{\theta}_k$ replaced by $\hat{\theta}_{kRC}$ for subject $i \in \mathcal{V}$, and by,

$$\hat{d}_{kRC} = \arg\max_{a_k} Q_{kRC}(a_k; \hat{\theta}_{kRC}) \tag{4.11}$$

for subject $i \in \mathcal{M} \setminus \mathcal{V}$, where $Q_{kRC}(a_k; \theta)$ is defined to be $Q_k(\overline{A}_{k-1}, \overline{X}_k, \overline{Z}_k, a_k; \theta)$ in (1.12) with X_k replaced by $\hat{X}_{ki}^{**}$ for $i \in \mathcal{M} \setminus \mathcal{V}$.

In the case where no additional data are available to quantify misclassification degrees or to estimate the parameter $\boldsymbol{\zeta}_k$ in model (4.10), we often invoke sensitivity analyses to evaluate the impact of misclassification on the Q-learning outcome. Basically, we consider a set of possibly representative values specified for $\boldsymbol{\zeta}_k$; then use model (4.10) to estimate X_{ki}^{**} for $i = 1, \cdots, n$; finally, we repeat the calibrated Q-learning algorithm to assess how the results may change relative to differently specified misclassification scenarios. Such a study helps us understand the sensitivity of the Q-learning results to the different misclassification degrees.

4.4 Addressing Misclassification Effects: Estimating Equation Approach

Examining the least squares method for (1.10) from the estimation equation perspective, here we develop an estimation equation approach to address misclassification effects. In line with the backward procedure of Q-learning, we proceed with the development stage by stage backward from stage K to stage 1, where the formulation for stage K is separated from that for stage k with $k = K-1, \cdots, 1$, reflecting the differences between the observed outcome at stage K and unobserved pseudo-outcomes for other stages. The following two sections differ in the treatment of the misclassification probabilities in (4.8).

4.4.1 Corrected Estimation Functions with Known Misclassification Probabilities

To highlight the idea, we start with the case where the misclassification probabilities are known. In what follows, we first describe estimation for stage K where the outcome measurements $\{Y_{Ki} : i = 1, \cdots, n\}$ are used, and then describe estimation for stage k with $k = K-1, \cdots, 1$ where pseudo-outcomes in (1.11) are used with modifications to accommodate misclassification effects.

Estimation Related to Stage K

For stage K, set $\ell_{Ki} = \{Y_{Ki} - Q_K(\overline{A}_{Ki}, \overline{X}_{Ki}, \overline{Z}_{Ki}; \theta_K)\}^2$, where $Q_K(\overline{A}_{Ki}, \overline{X}_{Ki}, \overline{Z}_{Ki}; \theta_K)$ is determined by (1.9) with the dependence on parameter θ_K spelled out. Define

$$S_K(\theta_K; Y_{Ki}, \overline{A}_{Ki}, \overline{X}_{Ki}, \overline{Z}_{Ki}) = \left(\frac{\partial \ell_{Ki}}{\partial \beta_K^{\mathrm{T}}}, \frac{\partial \ell_{Ki}}{\partial \psi_K^{\mathrm{T}}}\right)^{\mathrm{T}}$$
$$\triangleq (S_{K\beta}^{\mathrm{T}}(\theta_K; Y_{Ki}, \overline{A}_{Ki}, \overline{X}_{Ki}, \overline{Z}_{Ki}), S_{K\psi}^{\mathrm{T}}(\theta_K; Y_{Ki}, \overline{A}_{Ki}, \overline{X}_{Ki}, \overline{Z}_{Ki}))^{\mathrm{T}},$$

where

$$S_{K\beta}(\theta_K; Y_{Ki}, \overline{A}_{Ki}, \overline{X}_{Ki}, \overline{Z}_{Ki}) = \{Y_{Ki} - Q_K(\overline{A}_{Ki}, \overline{X}_{Ki}, \overline{Z}_{Ki}; \theta_K)\}\frac{\partial Q_K(\overline{A}_{Ki}, \overline{X}_{Ki}, \overline{Z}_{Ki}; \theta_K)}{\partial \beta_K} \qquad (4.12)$$

and

$$S_{K\psi}(\theta_K; Y_{Ki}, \overline{A}_{Ki}, \overline{X}_{Ki}, \overline{Z}_{Ki}) = \{Y_{Ki} - Q_K(\overline{A}_{Ki}, \overline{X}_{Ki}, \overline{Z}_{Ki}; \theta_K)\}\frac{\partial Q_K(\overline{A}_{Ki}, \overline{X}_{Ki}, \overline{Z}_{Ki}; \theta_K)}{\partial \psi_K}. \quad (4.13)$$

By (1.5) and (1.9), it is readily seen that $S_{K\beta}(\theta_K; Y_{Ki}, \overline{A}_{Ki}, \overline{X}_{Ki}, \overline{Z}_{Ki})$ and $S_{K\psi}(\theta_K; Y_{Ki}, \overline{A}_{Ki}, \overline{X}_{Ki}, \overline{Z}_{Ki})$ are unbiased estimating functions, that is, their expectations are zero. Therefore, solving the estimating equations

$$\sum_{i=1}^{n} S_K(\theta_K; Y_{Ki}, \overline{A}_{Ki}, \overline{X}_{Ki}, \overline{Z}_{Ki}) = 0 \quad (4.14)$$

for θ_K yields consistent estimator of θ_K, provided regularity conditions. Here and elsewhere, 0 is used to represent a vector, a matrix, or a real-valued zero without differentiation.

While a consistent estimator of θ_K can be obtained using (4.14), the applicability of (4.14) relies on the availability of precise measurements for X_{ki} with $k = 1, \cdots, K$ and $i = 1, \cdots, n$. When measurements of X_{ki} are unavailable but surrogates X^*_{ki} are collected, i.e., in the absence of $\mathcal{D}$, we cannot apply (4.14) to the observed data $\mathcal{D}^*$ with X_{ki} replaced by X^*_{ki}; otherwise, incorrect results may be produced.

To address the misclassification effects induced by X^*_{ki}, we construct an unbiased estimating function, say $S_{Kc}(\theta_K; Y_{Ki}, \overline{A}_{Ki}, \overline{X}^*_{Ki}, \overline{Z}_{Ki})$, using the observed surrogates X^*_{ki} together with other observed variables. The main idea is to find $S_{Kc}(\theta_K; Y_{Ki}, \overline{A}_{Ki}, \overline{X}^*_{Ki}, \overline{Z}_{Ki})$ such that its conditional expectation recovers the original unbiased estimating function $S_K(\theta_K; y_{Ki}, \overline{A}_{Ki}, \overline{X}_{Ki}, \overline{Z}_{Ki})$ in (4.14). That is, as long as

$$E\{S_{Kc}(\theta_K; Y_{Ki}, \overline{A}_{Ki}, \overline{X}^*_{Ki}, \overline{Z}_{Ki}) \mid Y_{Ki}, \overline{A}_{Ki}, \overline{X}_{Ki}, \overline{Z}_{Ki}\} = S_K(\theta_K; Y_{Ki}, \overline{A}_{Ki}, \overline{X}_{Ki}, \overline{Z}_{Ki}), \quad (4.15)$$

solving

$$\sum_{i=1}^{n} S_{Kc}(\theta_K; Y_{Ki}, \overline{A}_{Ki}, \overline{X}^*_{Ki}, \overline{Z}_{Ki}) = 0$$

for θ_K produces a consistent estimator for θ_K, under certain regularity conditions (Yi, 2017, Section 2.5).

By the linear regression model (1.9):

$$Q_K(\overline{A}_{Ki}, \overline{X}_{Ki}, \overline{Z}_{Ki}; \theta_K) = \beta_K^{\mathrm{T}} H_{K0} + (\psi_K^{\mathrm{T}} H_{K1}) A_K,$$

we simplify (4.12) and (4.13) to be

$$S_{K\beta}(\theta_K; Y_{Ki}, \overline{A}_{Ki}, \overline{X}_{Ki}, \overline{Z}_{Ki}) = \{Y_{Ki} - Q_K(\overline{A}_{Ki}, \overline{X}_{Ki}, \overline{Z}_{Ki}; \theta_K)\} H_{K0} \tag{4.16}$$

and

$$S_{K\psi}(\theta_K; Y_{Ki}, \overline{A}_{Ki}, \overline{X}_{Ki}, \overline{Z}_{Ki}) = \{Y_{Ki} - Q_K(\overline{A}_{Ki}, \overline{X}_{Ki}, \overline{Z}_{Ki}; \theta_K)\} H_{K1} A_K. \tag{4.17}$$

Because X_{Ki} is a binary variable taking value 0 or 1, any polynomials of X_{Ki} equals X_{Ki} itself, thus, the dependence of (4.16) and (4.17) on X_{ki} is merely reflected by X_{ki}.

To find $S_{Kc}(\theta_K; Y_{Ki}, \overline{A}_{Ki}, \overline{X}^*_{Ki}, \overline{Z}_{Ki})$ to meet (4.15), it suffices to find functions of X^*_{ki}, say $U(X^*_{ki})$, which may also involve model parameters, such that

$$E\{U(X^*_{ki}) \mid Y_{Ki}, \overline{A}_{Ki}, \overline{X}_{Ki}, \overline{Z}_{Ki}\} = X_{ki} \qquad \text{for} \qquad k = K, \cdots, 1. \tag{4.18}$$

It is easily shown that setting

$$U(X^*_{ki}) = \frac{X^*_{ki} - \pi_{k10}}{1 - \pi_{k10} - \pi_{k01}} \tag{4.19}$$

makes (4.18) hold (Yi, 2017, Problem 2.10). Consequently, $S_{Kc}(\theta_K; Y_{Ki}, \overline{A}_{Ki}, \overline{X}^*_{Ki}, \overline{Z}_{Ki})$ in (4.15) can be defined as $S_K(\theta_K; Y_{Ki}, \overline{A}_{Ki}, \overline{X}_{Ki}, \overline{Z}_{Ki})$ with X_{ki} replaced by $U(X^*_{ki})$.

Consequently, with the observed data $\mathcal{D}^*$, we solve

$$\sum_{i=1}^{n} S_{Kc}(\theta_K; Y_{Ki}, \overline{A}_{Ki}, \overline{X}^*_{Ki}, \overline{Z}_{Ki}) = 0 \tag{4.20}$$

for θ_K, and let $\hat{\theta}_{Kc} = (\hat{\beta}_{Kc}^{\mathrm{T}}, \hat{\psi}_{Kc}^{\mathrm{T}})^{\mathrm{T}}$ denote the resultant estimator of θ_K. Under regularity conditions (Yi, 2017, Section 1.3),

$$\sqrt{n}(\hat{\theta}_{Kc} - \theta_K) \xrightarrow{d} \mathcal{N}(0, \Sigma(\theta_K)) \quad \text{as} \quad n \to \infty, \tag{4.21}$$

where $\Sigma(\theta_K) = \left\{I(\theta_K)\right\}^{-1} J(\theta_K)\left\{I(\theta_K)\right\}^{-1\mathrm{T}}$, with

$$I(\theta_K) = E\left\{\frac{\partial S_{Kc}(\theta_K; Y_K, \overline{A}_K, \overline{X}^*_K, \overline{Z}_K)}{\partial \theta_K}\right\}$$

and

$$J(\theta_K) = E\Big[S_{Kc}(\theta_K; Y_K, \overline{A}_K, \overline{X}^*_K, \overline{Z}_K)\{S_{Kc}(\theta_K; Y_K, \overline{A}_K, \overline{X}^*_K, \overline{Z}_K)\}^{\mathrm{T}}\Big].$$

Estimation Related to Stage $k < K$

First, we consider (1.10) for $k = K-1, \cdots, 1$ for the error-free setting. Similar to the definitions of ℓ_{Ki} and $S_K(\theta_K; Y_{Ki}, \overline{A}_{Ki}, \overline{X}_{Ki}, \overline{Z}_{Ki})$ introduced earlier, let

$$\hat{\ell}_{ki} = \left\{\hat{Y}_{ki} - Q_k(\overline{A}_{ki}, \overline{X}_{ki}, \overline{Z}_{ki}; \theta_k)\right\}^2$$

and

$$\hat{S}_k(\theta_k; \hat{Y}_{ki}, \overline{A}_{ki}, \overline{X}_{ki}, \overline{Z}_{ki}) = \left(\frac{\partial \hat{\ell}_{ki}}{\partial \beta_k^{\mathrm{T}}}, \frac{\partial \hat{\ell}_{ki}}{\partial \psi_k^{\mathrm{T}}}\right)^{\mathrm{T}},$$

where $\hat{Y}_{ki}$ is the stage k pseudo-outcome for subject i, given by (1.11). For $k = K-1, \cdots, 1$, finding the minimizer in (1.10) may be alternatively viewed as solving the equation

$$\sum_{i=1}^{n} \hat{S}_k(\theta_k; \hat{Y}_{ki}, \overline{A}_{ki}, \overline{X}_{ki}, \overline{Z}_{ki}) = 0 \tag{4.22}$$

for θ_k.

The implementation of (4.22) requires the availability of precise measurements for X_{ki} for $k = 1, \cdots, K-1$ and $i = 1, \cdots, n$. However, precise measurements in $\mathcal{D}$ are unavailable, and we have only the surrogate dataset $\mathcal{D}^*$. Then similar to the consideration for stage K, we modify the function $\hat{S}_k(\theta_k; \hat{Y}_{ki}, \overline{A}_{ki}, \overline{X}_{ki}, \overline{Z}_{ki})$ in (4.22) by replacing X_{ki} with $U(X^*_{ki})$ in (4.18), and let $\hat{S}_{kc}(\theta_k; \hat{Y}^*_{ki}, \overline{A}_{ki}, \overline{X}^*_{ki}, \overline{Z}_{ki})$ denote the resulting function, where the inclusion of symbols $\hat{Y}^*_{ki}$ and X^*_{ki} in the function $\hat{S}_{kc}(\theta_k; \hat{Y}^*_{ki}, \overline{A}_{ki}, \overline{X}^*_{ki}, \overline{Z}_{ki})$ stresses its involvement of the observed surrogate measurements X^*_{ki}, with $\hat{Y}^*_{ki}$ defined by (1.11) with X_{ki} replaced by X^*_{ki}. Then for $k = K -$

$1, \cdots, 1$, we solve

$$\sum_{i=1}^{n} \hat{S}_{kc}(\theta_k; \hat{Y}^*_{ki}, \overline{A}_{ki}, \overline{X}^*_{ki}, \overline{Z}_{ki}) = 0 \tag{4.23}$$

for θ_k, and let $\hat{\theta}_{kc} = (\hat{\beta}^{\mathrm{T}}_{kc}, \hat{\psi}^{\mathrm{T}}_{kc})^{\mathrm{T}}$ denote the resulting estimator of θ_k.

Consequently, with estimators $\hat{\theta}_{Kc}$ and $\hat{\theta}_{kc}$ for $k = K-1, \cdots, 1$, the optimal treatment is estimated by modifying (1.12) with $\hat{\theta}_k$ replaced by $\hat{\theta}_{kc}$ and X_k replaced by $U(X^*_k)$, i.e.,

$$\hat{d}_{kc} = \arg\max_{a_k} Q_{kc}(\overline{A}_{k-1}, \overline{X}^*_k, \overline{Z}_k, a_k; \hat{\theta}_{kc}) \qquad \text{for} \qquad k = K, \cdots, 1, \tag{4.24}$$

where $Q_{kc}(\overline{A}_{k-1}, \overline{X}^*_k, \overline{Z}_k, a_k; \theta)$ is identical to $Q_k(\overline{A}_{k-1}, \overline{X}_k, \overline{Z}_k, a_k; \theta)$ defined in (1.12) except that X_k is replaced by $U(X^*_k)$.

We conclude this subsection with a comment. In contrast to the consistency of $\hat{\theta}_{Kc}$ indicated by (4.21), $\hat{\theta}_{kc}$ derived from (4.23) is not necessarily consistent for θ_{kc} for $k = K-1, \cdots, 1$, because the modified estimation function $\hat{S}_{kc}(\theta_k; \hat{Y}^*_{ki}, \overline{A}_{ki}, \overline{X}^*_{ki}, \overline{Z}_{ki})$ in (4.23) is not necessarily unbiased. In fact, it is difficult to find an unbiased estimating function for θ_k if $k < K$ due to the lack of an analytic form of the associated pseudo-outcome $\hat{Y}_{ki}$ or $\hat{Y}^*_{ki}$. By the similarity in forming (4.23) to that of (4.20), we anticipate implementing (4.23) would yield better estimation results than directly using (4.22) with X_{ki} replaced by X^*_{ki}.

4.4.2 Corrected Estimation Functions with Unknown Misclassification Probabilities

In some applications, the misclassification probabilities are unknown and need to be estimated from an additional data source, such as a validation sample. In this subsection, we modify the method in Section 4.4.1 to incorporate this feature. Consider the main study sample, $\mathcal{D}^*$, and the internal validation subsample $\mathcal{D}_V$ described in Section 4.3, where $\mathcal{V}$ contain m subjects with $m \leq n$.

Estimation of Misclassification Probabilities

For $i \in \mathcal{M}$ and $k = 1, \cdots, K$, let

$$\pi_{ki01} = P(X^*_{ki} = 0 \mid X_{ki} = 1, \overline{A}_{(k-1)i} \cup \overline{X}_{(k-1)i} \cup \overline{Z}_{ki})$$

and

$$\pi_{ki10} = P(X^*_{ki} = 1 \mid X_{ki} = 0, \overline{A}_{(k-1)i} \cup \overline{X}_{(k-1)i} \cup \overline{Z}_{ki})$$

be the misclassification probabilities for the error-prone binary covariate X_{ki}, which may depend on the error-free covariates in $\overline{A}_{(k-1)i} \cup \overline{X}_{(k-1)i} \cup \overline{Z}_{ki}$.

To describe the dependence of misclassification probabilities on the covariates, we consider the logistic regression models

$$\begin{aligned} \operatorname{logit} \pi_{ki10} &= \alpha_{k0}^{\mathrm{T}} \mathcal{W}_{ki0}; \\ \operatorname{logit} \pi_{ki01} &= \alpha_{k1}^{\mathrm{T}} \mathcal{W}_{ki1}, \end{aligned} \tag{4.25}$$

where α_{kl} denotes the vector of regression coefficients and $\mathcal{W}_{kil}$ may include 1 and a subset of covariates $\{X_{ki} = l, \overline{A}_{(k-1)i} \cup \overline{X}_{(k-1)i} \cup \overline{Z}_{ki}\}$ that reflects different misclassification mechanisms for $l = 0, 1$. Having 1 in $\mathcal{W}_{kil}$ allows the inclusion of the intercept in (4.25), and $\mathcal{W}_{kil}$ may contain the entire covariate vector $\{X_{ki} = l, \overline{A}_{(k-1)i} \cup \overline{X}_{(k-1)i} \cup \overline{Z}_{ki}\}$ or just constant 1 alone, where the latter case corresponds to homogeneous misclassification across all subjects. Let $\alpha_k = (\alpha_{k0}^{\mathrm{T}}, \alpha_{k1}^{\mathrm{T}})^{\mathrm{T}}$ denote the parameter vector for $k = 1, \cdots, K$.

For $i = 1, \cdots, n$ and $k = 1, \cdots, K$, let

$$L_{ki}(\alpha_k) = P(X^*_{ki} = x^*_{ki} \mid X_{ki} = x_{ki}, \overline{A}_{(k-1)i} \cup \overline{X}_{(k-1)i} \cup \overline{Z}_{ki}),$$

which equals $\{\pi_{ki10}^{x^*_{ki}}(1 - \pi_{ki10})^{1-x^*_{ki}}\}^{1-x_{ki}} \cdot \{\pi_{ki01}^{1-x^*_{ki}}(1 - \pi_{ki01})^{x^*_{ki}}\}^{x_{ki}}$ for $x_{ki}, x^*_{ki} = 0, 1$. Let $\alpha = (\alpha_1^{\mathrm{T}}, \cdots, \alpha_K^{\mathrm{T}})^{\mathrm{T}}$ and $\vartheta = (\beta_K^{\mathrm{T}}, \psi_K^{\mathrm{T}}, \alpha^{\mathrm{T}})^{\mathrm{T}}$. Let $S_{ki}(\alpha_k) = \partial \log L_{ki}(\alpha_k)/\partial \alpha_k$ and let

$$S_i(\alpha) = \left(S_{1i}^{\mathrm{T}}(\alpha_1), \cdots, S_{Ki}^{\mathrm{T}}(\alpha_K)\right)^{\mathrm{T}}.$$

With internal validation data, solving

$$\sum_{i\in\mathcal{V}} S_i(\alpha) = 0$$

for α yields the maximum likelihood estimate, denoted $\hat{\alpha} = (\hat{\alpha}_1^{\mathrm{T}}, \cdots, \hat{\alpha}_K^{\mathrm{T}})^{\mathrm{T}}$, of α.

Estimation for the Parameters of Q-Functions

For stage K, we estimate θ_K by solving

$$\sum_{i\in\mathcal{M}\setminus\mathcal{V}} S_{Kci}(\theta_K, \hat{\alpha}) + \sum_{i\in\mathcal{V}} S_{Ki}(\theta_K) = 0$$

for θ_K, and let $\hat{\theta}_K$ denote the resultant estimator of θ_K, where with the dependence on the outcome and covariates suppressed in the notation for simplicity, $S_{Kci}(\theta_K, \alpha)$ is $S_{Kc}(\theta_K; Y_{Ki}, \overline{A}_{Ki}, \overline{X}^*_{Ki}, \overline{Z}_{Ki})$ in (4.20) with the dependence on α spelled out, and $S_{Ki}(\theta_K)$ is $S_K(\theta_K; Y_{Ki}, \overline{A}_{Ki}, \overline{X}_{Ki}, \overline{Z}_{Ki})$ in (4.14).

Let $\hat{\vartheta} = (\hat{\theta}_K^{\mathrm{T}}, \hat{\alpha}^{\mathrm{T}})^{\mathrm{T}}$. Adapting the proof of Yi et al. (2018), we show the following theorem.

Theorem 4.4.1 *Under regularity conditions and that the ratio m/n approaches a positive constant, say ρ, as $n \to \infty$, the following results hold:*

(a). $\hat{\vartheta}$ is a consistent estimator of ϑ;

(b).

$$\sqrt{n}(\hat{\vartheta} - \vartheta) \xrightarrow{d} \mathcal{N}(0, \Sigma_V) \quad as \quad n \to \infty,$$

where $\Sigma_V = A_V^{-1} B_V A_V^{-1\mathrm{T}}$, with

$$A_V = -(1-\rho)\begin{pmatrix} E(\frac{\partial S_{Kci}(\theta_K,\alpha)}{\partial \theta_K}) & E(\frac{\partial S_{Kci}(\theta_K,\alpha)}{\partial \alpha}) \\ 0 & 0 \end{pmatrix} - \rho\begin{pmatrix} E(\frac{\partial S_{Ki}(\theta_K)}{\partial \theta_K}) & 0 \\ 0 & E(\frac{S_i(\alpha)}{\partial \alpha}) \end{pmatrix} \text{ and}$$

$$B_V = (1-\rho)\begin{pmatrix} E\{S_{Kci}(\theta_K,\alpha)S_{Kci}^{\mathrm{T}}(\theta_K,\alpha)\} & 0 \\ 0 & 0 \end{pmatrix} + \rho\begin{pmatrix} E\{S_{Ki}(\theta_K)S_{Ki}^{\mathrm{T}}(\theta_K)\} & 0 \\ 0 & E\{S_i(\alpha)S_i^{\mathrm{T}}(\alpha)\} \end{pmatrix}.$$

Finally, for $k = K-1, \cdots, 1$, an estimate of θ_k can be obtained by solving (4.23), where π_{k10} and π_{k01} in (4.19) are determined by (4.25) with α replaced by $\hat{\alpha}$.

4.5 Numerical Studies

4.5.1 Simulation Study

In this subsection, we conduct simulation studies to assess the finite sample performance of the methods described in Sections 4.3 and 4.4. We use the same settings as in Section 4.2.3 to generate the main study data $\mathcal{D}^* = \left\{\{Z_{ki}, X^*_{ki}, A_{ki}, Y_{ki}\} : k = 1, \cdots, K; i \in \mathcal{M}\right\}$ with $|\mathcal{M}| = n = 1000$. Further, we generate an internal validation subsample by randomly selecting 30% of study subjects from $\mathcal{M}$, and record their accurate measurements of $\{X_{ki} : k = 1, \cdots, K; i \in \mathcal{V}\}$ to form the validation subsample $\mathcal{D}_V = \left\{\{A_{ki}, X_{ki}, X^*_{ki}, Z_{ki}, Y_{ki}\} : k = 1, \cdots, K; i \in \mathcal{V}\right\}$.

We consider three methods to analyze the data. The first method, called "RC", applies the method described in Section 4.3 to both $\mathcal{D}^*$ and $\mathcal{D}_V$. The second method, called "EE-known", uses the method in Section 4.4.1 to data $\mathcal{D}^*$, where the misclassification probabilities are taken as known. The third method, called "EE-estimated", applies the method in Section 4.4.2 to both $\mathcal{D}^*$ and $\mathcal{D}_V$, where the misclassification probabilities are estimated using $\mathcal{D}_V$ following the procedure described in Section 4.4.2.

The results for stages 1 and 2 are reported differently due to the difference in knowing the information about these two stages. Table 4.3 reports the results for stage 1 over 1000 simulations for the regular, weak non-regular, and non-regular settings, where "Bias", "ESE", "PBCR", and "DBCR" are the same as in Section 4.2.3, but "SE" and "MSE" may be slightly different, depending on the method used. When the RC method is considered, "SE" represents the average of standard errors obtained from the bootstrap method with 1000 bootstrap samples, and "MSE" represents the mean squared error of the estimates obtained using the bootstrap SE; when the EE-known or EE-estimated method is used, "SE" and "MSE" are the same as in Section 4.2.3. The RC and EE methods exhibit good performance by outputting estimators with fairly small bias, indicating their efficacy in rectifying the effects of misclassification; notably, the RC method's capacity to reduce bias for weak non-regular and non-regular settings deteri-

orates more rapidly than that of the two EE methods as misclassification probabilities increase. As expected, there is a consistent trend of increased variability in parameter estimation for all the three methods as misclassification probabilities increase across all the three settings; such an increase is more pronounced in the two EE methods compared to the RC method; and the EE-estimated method tends to yield more varied estimates than the EE-known method does. PB 95% CIs of the RC and EE methods tend to be over-covered for the regular and weak non-regular settings, yet they may be considerably under-covered for the non-regular setting. On the contrary, DB 95% CIs of the RC and EE methods exhibit under-coverage, regardless of the regularity condition of parameters, and moreover, the severity of under-coverage tends to increase as the misclassification probabilities increase.

The results for stage 2 over 1000 simulations are reported in Tables 4.4-4.6, respectively, for regular, weak non-regular, and non-regular settings, where "Bias", "ESE", "SE", and "MSE" are the same as for Table 4.3, yet "WTCR" represents the CR of 95% CIs obtained using the bootstrap SE for the RC method or model-based SE obtained from the asymptotic result in Section 4.4.1 or 4.4.2 for the EE methods. Similar to the findings for stage 1, the RC and EE-known approaches demonstrate a fairly commendable performance in stage 2. The RC and EE-known methods tend to outperform the EE-estimated method, though the results produced by the EE-estimated method still appear to be fairly satisfactory. The variability of parameter estimation tends to increase as misclassification probabilities increase, and such variability is more noticeable for the EE methods than the RC method. Our limited numerical explorations (not reported here) indicate that the performance of the EE-known and EE-estimated methods can be further improved by increasing the sample size.

4.5.2 Future Treatment Predictions

Here we assess the influence of ignoring misclassification on determining optimal treatments for patients as well as the performance of the proposed methods in terms of prediction. To this end, we first use the procedure in Section 4.5.1 to generate training data for 1000 individuals to estimate model parameters that will be used to do prediction. Next, we simulate a population of 5000 patients for whom we wish to predict optimal treatment regimes; we report the

prediction performance for the proposed methods, applied to the observed surrogate measurements, as opposed to the naive method of ignoring misclassification. Specifically, we employ the following steps:

Step 1: Data generation for prediction:

- First, using the procedure in Section 4.2.3, we generate the true measurements, $\mathcal{D}_{TP} \triangleq \left\{\{A_{ki}, X_{ki}, Z_{ki}\} : i \in \mathcal{M}_P\right\}$ with $|\mathcal{M}_P| = 5000$, for the covariates $\{Z_1, X_1, A_1, Z_2, X_2, A_2\}$ for 5000 patients independently.
- Secondly, we generate corresponding surrogate values, $\left\{X^*_{ki} : k = 1, 2; i \in \mathcal{M}_P\right\}$, of X_{1i} and X_{2i} using the same procedure and misclassification probabilities considered in Section 4.2.3 for Settings 1-3, and obtain observed measurements of covariates and treatments for the main study, $\mathcal{D}^*_P = \left\{\{A_{ki}, X^*_{ki}, Z_{ki}\} : i \in \mathcal{M}_P\right\}$.
- Finally, using the procedure described in Section 4.5.1, we generate measurements of covariates and treatments for an internal validation subsample, $\mathcal{D}_{VP} = \left\{\{A_{ki}, X_{ki}, X^*_{ki}, Z_{ki}\} : k = 1, 2; i \in \mathcal{V}_P\right\}$, with $|V_P| = 30\%|\mathcal{M}_P|$.

Step 2: Parameter estimation from different methods:

- For the naive and EE-known methods, we apply the measurements in $\mathcal{D}^*$ in Section 4.5.1 to estimate the parameters for the stage 1 model (4.6) and the stage 2 model (4.5), where the parameters for the misclassification model (4.8) are taken as the true values in generating $\mathcal{D}^*$.
- For the RC and EE-estimated methods, we use the main study data $\mathcal{D}^*$ and the validation data $\mathcal{D}_V$ in Section 4.5.1 to estimate the parameters for the stage 1 model (4.6) and the stage 2 model (4.5).

Step 3: Prediction:

As a comparison, we introduce the "true" method, which uses the true parameter values θ_1 and θ_2 as well as the true measurements in $\mathcal{D}_{TP}$ to obtain true optimal treatment regimes. That is, using the true values for θ_1 and θ_2 specified in Settings 1-3 in Section 4.2.3 and the true measurements in $\mathcal{D}_{TP}$ for (4.7), we obtain the true

optimal treatments in stages 1 and 2 for 5000 patients, and take these results as reference values. We now examine the following two scenarios, which differ in the use of the true or surrogate covariate measurements.

- Scenario 1: With parameters replaced with their estimates obtained from each method described in Step 2, we predict the corresponding optimal treatments in stages 1 and 2 for 5000 patients using (4.7) in Section 4.2.3, where $\mathcal{D}_{TP}$ is used.
- Scenario 2: With the estimated parameters for each method described in Step 2, we predict the corresponding optimal treatments in stages 1 and 2 for 5000 patients using (4.11) and (4.24) for the RC and EE methods, respectively, where $\mathcal{D}_{VP}$ and $\mathcal{D}_P^*$ are respectively used for the RC and EE methods.

Step 4: Summarizing results:

- Comparing the results from Step 3, we count the number, denoted M, of patients for whom the predicted optimal treatments of each method in Step 2 match the reference values. Then calculate the proportion of correctly specified optimal treatments, denoted PCOT, by dividing M by the number of patients, 5000.

To alleviate Monte Carlo variations, we repeat Steps 1 to 4 for 1000 times, and calculate the average of the proportions obtained in Step 4 across the 1000 repetitions. This average proportion, denoted APCOT, represents an estimated proportion of optimally treated future patients for each method: a larger APCOT indicates better performance of the respective method. In Table 4.7, we report APCOT values under Scenarios 1 and 2 for the four methods described in Step 2. It is evident that the performance of the methods exhibits variation with the change of π_{01} and π_{10}, as well as the regularity condition of the parameters.

Examining the left panel of Table 4.7 (i.e., Scenario 1), we see that for stage 1 optimal treatment prediction, the three proposed methods generally outperform the naive method, except for some settings with $\pi_{10} = 0$ or $\pi_{10} = 0$ for the EE-known method, such as the weak non-regular setting with $\pi_{10} = 0$ and $\pi_{01} = 0.3$, and the regular settings with $\pi_{10} = 0.15$ or 0.3 when $\pi_{01} = 0$. The results produced by the three proposed methods are fairly comparable, and the RC method tends to perform the best, though in some cases, the EE-known or EE-

estimated method outperforms the RC method. Unsurprisingly, as the misclassification degree increases, the performance of the three proposed methods tends to decay, yet this trend does not exhibit for the naive method. For stage 2 optimal treatment prediction, we observe that the three proposed methods outperform the naive method, except for the weak non-regular setting with $\pi_{10} = 0$ and $\pi_{01} = 0.3$. Interestingly, under the non-regular setting with most misclassification probabilities, APCOT values produced by the naive method are larger than those obtained from the three proposed methods and are close to 1. This finding does not necessarily support the good performance of the naive method, but instead, it shows unreliable optimal treatment prediction under the non-regular setting: ignoring the feature of misclassification yields results similar to those produced from using precisely measured data.

Examining the right panel of Table 4.7 (i.e., Scenario 2), we observe that regarding stage 1 optimal treatment prediction, the three proposed methods exhibit superior performance to the naive method. The two EE methods tend to perform better than the RC method, and the EE-estimated method performs the best. Similarly, for stage 2 optimal treatment prediction, the three proposed methods outperform the naive method in both the regular and weak non-regular settings. In the non-regular setting, the naive method outputs the largest APCOT values than the three proposed methods which account for the misclassification effects. However, this does not necessarily indicate the best performance of the naive method, as explained earlier for Scenario 1.

4.5.3 Data Analysis

Sequenced Treatment Alternatives to Relieve Depression (STAR*D) was a multi-site, multi-level randomized clinical trial enrolling 4041 patients with nonpsychotic MDD (Rush et al., 2004; Fava et al., 2003). The objective of this study was to assess the comparative effectiveness of different treatment options for patients. The trial involved four examination levels, each level consisting of a 12-week period of treatment, with scheduled clinic visits at weeks 0, 2, 4, 6, 9, and 12. The primary outcome of the trial is the severity of depression at any clinic visit assessed using the clinician-rated and self-report versions of the *Quick Inventory of Depressive Symptomatology* (QIDS) scores (Rush et al., 2004). Larger values of the QIDS score corre-

spond to higher severity of depression and thus represent a worse outcome, where 5 is taken as a benchmark for a total 12-week clinician-rated QIDS score of a patient. Specifically, at the end of each level, patients with a QIDS score ≤ 5 did not move on to the next examination (i.e., patients achieved clinical remission), whereas a QIDS score > 5 for a patient indicated that he or she did not have an adequate response.

At level 1, all patients were treated with citalopram (CIT). Those without an adequate response at level 1 were eligible to receive one of seven treatment options available at level 2, depending on their preference to switch or augment their level 1 treatment. Level 2 consists of four switch options (venlafaxine (VEN), sertraline (SER), bupropion (BUP), and cognitive therapy (CT)) and three augment options (CIT+CT, CIT+BUP or buspirone (BUS)+CIT). Patients assigned to cognitive therapy (alone or augmented with citalopram) were eligible, in the case of inadequate response, to move to a supplementary level 2A and be randomized to switch to BUP or VEN. Patients showing unsatisfactory responses would continue to level 3 to receive one of the two available switch options (mirtazapine (MIRT) and nortriptyline (NTP)) or to augment their previous treatment with lithium (Li) or thyroid hormone (THY). Patients without a satisfactory response at level 3 continued to level 4 treatments, which included two options: tranylcypromine (TCP) and MIRT +VEN. For a schematic of the STAR*D study design see Chakraborty and Moodie (2013), and for a complete description see Rush et al. (2004).

Similar to Chakraborty and Moodie (2013), here we consider levels 2, 2A, and 3 and cast the problem into the framework of Section 4.2.3. Specifically, we take level 2 (including 2A, if applicable) as stage 1 and let level 3 be stage 2 for the Q-learning framework described in Section 4.2.3 with $K = 2$. We classify treatments at stage 1 into two categories: (i) treatment with either SER or CIT plus one of BUP, BUS and CT, and (ii) treatment with at least one of the options from BUP, CT, and VEN. Treatment options at stage 2 are also classified into two categories: (i) treatment with an augmentation of SER or CIT-containing level 2 treatment with either Li or THY, and (ii) treatment with MIRT or NTP, or augmentation of at least one of the options VEN, BUP, or CT with either Li or THY. For any patient, let A_k denote the treatment status at stage k, which is coded as 1 for category (i) and 0 for category (ii).

Two covariates are included in this analysis. The first covariate, denoted Z_k, is continuous and represents QIDS score measured at the beginning of the level, while the second covariate,

denoted X_k, is binary and indicates the patient's preference for switch ($X_k = 1$) or augmentation ($X_k = 0$).

Some patients without a satisfactory response in stage 1 dropped out of the study without continuing to stage 2. Removing those patients, we analyze a data subset that includes 1396 patients for stage 1 among whom 922 were eligible to move on to stage 2; however, only 369 patients among those 922 patients were present at stage 2 and the rest dropped out. Because of the trial design, at stage 2 we have the outcome data only for the non-remitters from stage 1, therefore, Chakraborty et al. (2013) considered the following overall primary outcome, which is also employed here:

$$Y = I_1 R_1 + (1 - I_1)\left(\frac{R_1 + R_2}{2}\right),$$

where R_1 and R_2 denote the total QIDS scores at the end of stages 1 and 2, respectively; and I_1 is the remission indicator at the end of stage 1, taking 1 for remitters and 0 for non-remitters. That is, using the notation in Section 1.2.2, for each study subject, we take $Y_1 \equiv 0$ and $Y_2 = Y$, where $K = 2$. We employ the following models for Q-functions in (1.9):

$$Q_2(H_2, A_2) = \beta_{02} + \beta_{12} Z_2 + \beta_{22} X_2 + \beta_{32} A_1 + (\psi_{02} + \psi_{12} Z_2 + \psi_{22} X_2) A_2;$$
$$Q_1(H_1, A_1) = \beta_{01} + \beta_{11} Z_1 + \beta_{21} X_1 + (\psi_{01} + \psi_{11} Z_1 + \psi_{21} X_1) A_1,$$

where for $k = 1, 2$, $\beta_{0k}, \beta_{1k}, \beta_{2k}, \psi_{0k}, \psi_{1k}, \psi_{2k}$, and β_{32} are regression coefficients. As mentioned previously, some patients who were eligible to continue to stage 2 dropped out of the study, making the computation of the pseudo-outcomes for them impossible as covariates Z_2 and X_2 are missing for them. To get around this problem, following Chakraborty and Moodie (2013), the value of Z_2 was imputed by the last observed QIDS score in stage 1, and the missing values of the binary variable X_2 were imputed using k nearest neighbor (k-NN) classification with k taken as 5.

We conduct three analyses here. In Analysis 1, we treat both Z_k and X_k as if they were error-free and implement the procedure in Section 1.2.2, and report the estimation results for the model parameters in the first row of Table 4.8 for stages 1 and 2. In the next two analyses, we evaluate the effects of possibly misclassified covariate X_k on parameter estimation with Z_k treated as error-free. As there is no additional data such as a validation subsam-

ple to characterize the degree of misclassification in the STAR*D data, we carry out sensitivity analyses using the two correction methods described in Sections 4.4.1 and 4.3, and respectively call them Analysis 2 and Analysis 3. To reflect possibly different scenarios of misclassification in X_k, in Analysis 2, we consider four sets of misclassification probabilities $(\pi_{10}, \pi_{01})^{\mathrm{T}} = (0.01, 0.01)^{\mathrm{T}}, (0.03, 0.03)^{\mathrm{T}}, (0.07, 0.07)^{\mathrm{T}}$, or $(0.10, 0.10)^{\mathrm{T}}$; and in Analysis 3, we consider model (4.10) with $m_1(H_{1i}^*; \boldsymbol{\zeta}_1) = \zeta_{01} + \zeta_{11}X_1 + \zeta_{21}Z_1$ and $m_2(H_{2i}^*; \boldsymbol{\zeta}_2) = \zeta_{02} + \zeta_{12}X_2 + \zeta_{22}Z_2 + \zeta_{32}A_1 + \zeta_{42}Z_1 + \zeta_{52}X_1$, where four sets of values, listed in Table 4.9, for the model parameters $\boldsymbol{\zeta}_1 = (\zeta_{01}, \zeta_{11}, \zeta_{21})^{\mathrm{T}}$ and $\boldsymbol{\zeta}_2 = (\zeta_{02}, \zeta_{12}, \zeta_{22}, \zeta_{32}, \zeta_{42}, \zeta_{52})^{\mathrm{T}}$ are examined. Numerical results of Analyses 2 and 3 are respectively reported in Tables 4.8 and 4.10.

Unsurprisingly, numerical results reveal varying estimates of the model parameters for both stages among the three analyses. All the three analyses suggest that stage 1 parameter ψ_{11} and the three stage 2 parameters ψ_{01}, ψ_{12} and ψ_{22} are all statistically insignificant. However, the significance for stage 1 parameters ψ_{01} and ψ_{21} are differently revealed with different degrees of misclassification considered. Without considering possible presence of misclassification, Analysis 1 finds evidence of suggesting statistical significance for ψ_{01} but not ψ_{21}. Analyses 2 and 3, addressing the misclassification effects for varying scenarios uncovers different nature of ψ_{01} and ψ_{21}, whose significance is unveiled differently, driven by different amounts of misclassification.

While the specification of misclassification probabilities does not need to be restricted to what is considered here, the sensitivity analyses demonstrate that ignoring misclassification can output results deviating significantly from the ground truth if the issue of misclassification is serious. Although the underlying truth is never known, examining data from different angles helps us reveal a more comprehensive picture to enhance the understanding than conducting a single analysis with data treated as if being error-free.

Finally, the models we consider here differ from those examined by Spicker and Wallace (2020). Unlike Spicker and Wallace (2020), the slope of the QIDS score is not included in the Q-functions here. On the other hand, the interaction term X_2A_2 is included in our second stage Q-function but was not considered by Spicker and Wallace (2020). The study here focuses on investigating the effects stemming from misclassification in discrete covariates, whereas Spicker and Wallace (2020) examined the effects of measurement error in continuous covari-

ates.

4.6 Summary

While careful designs are helpful in collecting good quality data, measurement error and misclassification are still inevitable and they arise ubiquitously in applications. In this chapter, we examine DTRs with misclassification in covariates. Focusing on the Q-learning procedure, we demonstrate biased estimation results through simulation studies. It is necessary to introduce de-biasing adjustments to account for mismeasurement effects in inferential procedures. Here we present two correction methods for Q-learning based on regression calibration and unbiased estimation equation approaches. Results from extensive simulation studies confirm the satisfactory performance of the correction methods in reducing or eliminating bias in parameter estimates.

Table 4.1: Simulation studies for demonstrating biased estimation of the naive method in contrast to the EFLS method: stage 1. Entries in bold are obtained from the setting without misclassification

				regular		weak non-regular		non-regular	
	π_{01}	Method		ψ_{01}	ψ_{11}	ψ_{01}	ψ_{11}	ψ_{01}	ψ_{11}
	0	EFLS	Bias	**0.001**	**0.002**	**0.010**	**0.012**	**0.016**	**0.002**
			SE	**0.202**	**0.285**	**0.100**	**0.142**	**0.102**	**0.144**
			ESE	**0.212**	**0.283**	**0.117**	**0.144**	**0.128**	**0.142**
			MSE	**0.041**	**0.081**	**0.010**	**0.020**	**0.011**	**0.021**
			PBCR	**0.947**	**0.949**	**0.948**	**0.938**	**0.934**	**0.944**
			DBCR	**0.955**	**0.956**	**0.952**	**0.950**	**0.943**	**0.953**
$\pi_{10} = 0$	0.15	Naive	Bias	0.110	0.095	0.157	0.185	0.092	0.179
			SE	0.163	0.250	0.095	0.145	0.096	0.147
			ESE	0.199	0.256	0.117	0.146	0.125	0.149
			MSE	0.039	0.072	0.034	0.055	0.018	0.054
			PBCR	0.918	0.978	0.744	0.792	0.895	0.816
			DBCR	0.852	0.842	0.713	0.678	0.858	0.687
	0.3	Naive	Bias	0.198	0.194	0.242	0.314	0.162	0.330
			SE	0.134	0.227	0.090	0.152	0.091	0.153
			ESE	0.179	0.228	0.110	0.154	0.124	0.151
			MSE	0.057	0.089	0.067	0.122	0.034	0.132
			PBCR	0.846	0.963	0.439	0.430	0.741	0.392
			DBCR	0.747	0.625	0.437	0.296	0.673	0.250
$\pi_{10} = 0.15$	0	Naive	Bias	0.003	0.102	0.011	0.184	0.075	0.184
			SE	0.217	0.286	0.108	0.142	0.109	0.144
			ESE	0.236	0.286	0.131	0.143	0.140	0.141
			MSE	0.047	0.092	0.012	0.054	0.018	0.055
			PBCR	0.976	0.984	0.961	0.772	0.926	0.794
			DBCR	0.907	0.850	0.935	0.668	0.896	0.659
	0.15	Naive	Bias	0.121	0.241	0.179	0.422	0.057	0.424
			SE	0.173	0.244	0.100	0.142	0.102	0.144
			ESE	0.204	0.239	0.133	0.144	0.130	0.142
			MSE	0.044	0.118	0.042	0.198	0.014	0.200
			PBCR	0.949	0.968	0.761	0.075	0.953	0.065
			DBCR	0.858	0.543	0.630	0.095	0.902	0.095
	0.3	Naive	Bias	0.216	0.350	0.285	0.612	0.142	0.614
			SE	0.140	0.215	0.094	0.145	0.095	0.146
			ESE	0.209	0.215	0.122	0.148	0.133	0.141
			MSE	0.066	0.169	0.090	0.395	0.029	0.398
			PBCR	0.846	0.788	0.381	0.000	0.837	0.000
			DBCR	0.769	0.232	0.322	0.006	0.737	0.006
$\pi_{10} = 0.3$	0	Naive	Bias	0.024	0.178	0.071	0.327	0.145	0.323
			SE	0.225	0.279	0.119	0.147	0.120	0.149
			ESE	0.264	0.271	0.139	0.140	0.145	0.147
			MSE	0.051	0.109	0.019	0.129	0.035	0.126
			PBCR	0.990	0.986	0.958	0.362	0.864	0.350
			DBCR	0.895	0.663	0.889	0.237	0.789	0.270
	0.15	Naive	Bias	0.165	0.360	0.120	0.613	0.023	0.612
			SE	0.168	0.221	0.109	0.144	0.110	0.145
			ESE	0.268	0.220	0.143	0.145	0.147	0.140
			MSE	0.055	0.179	0.026	0.396	0.013	0.396
			PBCR	0.936	0.805	0.911	0.000	0.969	0.000
			DBCR	0.839	0.237	0.790	0.002	0.889	0.004
	0.3	Naive	Bias	0.268	0.482	0.225	0.836	0.123	0.839
			SE	0.129	0.183	0.101	0.143	0.102	0.144
			ESE	0.257	0.184	0.132	0.141	0.136	0.146
			MSE	0.089	0.266	0.061	0.719	0.025	0.725
			PBCR	0.834	0.009	0.579	0.000	0.887	0.000
			DBCR	0.787	0.046	0.528	0.000	0.771	0.000

Table 4.2: Simulation studies for demonstrating biased estimation of the naive method in contrast to the EFLS method: stage 2. Entries in bold are obtained from the setting without misclassification

				regular			weak non-regular			non-regular		
	π_{01}	Method		ψ_{02}	ψ_{12}	ψ_{22}	ψ_{02}	ψ_{12}	ψ_{22}	ψ_{02}	ψ_{12}	ψ_{22}
	0	EFLS	Bias	**0.003**	**0.001**	**0.005**	**0.004**	**0.002**	**0.006**	**0.004**	**0.002**	**0.005**
			SE	**0.100**	**0.090**	**0.127**	**0.100**	**0.090**	**0.127**	**0.100**	**0.090**	**0.127**
			ESE	**0.101**	**0.087**	**0.123**	**0.100**	**0.092**	**0.127**	**0.105**	**0.089**	**0.127**
			MSE	**0.010**	**0.008**	**0.016**	**0.010**	**0.008**	**0.016**	**0.010**	**0.008**	**0.016**
			WTCR	**0.952**	**0.954**	**0.946**	**0.951**	**0.947**	**0.960**	**0.939**	**0.952**	**0.947**
$\pi_{10} = 0$	0.15	Naive	Bias	0.655	0.649	0.004	0.292	0.297	0.003	0.250	0.261	0.012
			SE	0.132	0.122	0.171	0.106	0.098	0.137	0.104	0.096	0.135
			ESE	0.153	0.137	0.173	0.112	0.100	0.139	0.110	0.096	0.139
			MSE	0.446	0.436	0.029	0.096	0.098	0.019	0.073	0.077	0.018
			WTCR	0.004	0.000	0.944	0.230	0.142	0.944	0.327	0.236	0.943
	0.3	Naive	Bias	1.157	1.163	0.009	0.536	0.537	0.002	0.460	0.467	0.007
			SE	0.149	0.147	0.198	0.109	0.107	0.145	0.106	0.104	0.141
			ESE	0.173	0.152	0.198	0.114	0.109	0.146	0.117	0.102	0.142
			MSE	1.361	1.374	0.039	0.299	0.300	0.021	0.223	0.229	0.020
			WTCR	0.000	0.000	0.950	0.007	0.002	0.954	0.016	0.009	0.944
$\pi_{10} = 0.15$	0	Naive	Bias	0.000	0.641	0.001	0.002	0.292	0.001	0.001	0.262	0.003
			SE	0.139	0.122	0.170	0.112	0.098	0.137	0.110	0.097	0.135
			ESE	0.119	0.136	0.166	0.110	0.097	0.138	0.105	0.100	0.133
			MSE	0.019	0.426	0.029	0.013	0.095	0.019	0.012	0.078	0.018
			WTCR	0.977	0.000	0.954	0.948	0.147	0.943	0.961	0.244	0.954
	0.15	Naive	Bias	0.748	1.503	0.010	0.349	0.694	0.000	0.301	0.597	0.007
			SE	0.161	0.144	0.204	0.116	0.104	0.146	0.112	0.101	0.142
			ESE	0.171	0.189	0.207	0.120	0.120	0.146	0.116	0.109	0.140
			MSE	0.586	2.280	0.042	0.135	0.492	0.021	0.103	0.367	0.020
			WTCR	0.006	0.000	0.950	0.150	0.000	0.956	0.237	0.000	0.957
	0.3	Naive	Bias	1.321	2.195	0.002	0.600	0.999	0.001	0.523	0.873	0.004
			SE	0.173	0.161	0.225	0.118	0.109	0.153	0.113	0.105	0.147
			ESE	0.193	0.207	0.227	0.123	0.125	0.155	0.116	0.118	0.147
			MSE	1.775	4.844	0.051	0.374	1.010	0.023	0.286	0.773	0.022
			WTCR	0.000	0.000	0.959	0.000	0.000	0.953	0.002	0.000	0.951
$\pi_{10} = 0.3$	0	Naive	Bias	0.001	1.143	0.002	0.001	0.531	0.008	0.003	0.462	0.007
			SE	0.169	0.147	0.198	0.124	0.107	0.145	0.120	0.104	0.141
			ESE	0.131	0.153	0.198	0.114	0.111	0.143	0.112	0.106	0.138
			MSE	0.029	1.328	0.039	0.015	0.294	0.021	0.015	0.224	0.020
			WTCR	0.988	0.000	0.956	0.963	0.000	0.950	0.965	0.009	0.954
	0.15	Naive	Bias	0.885	2.189	0.009	0.410	0.999	0.002	0.350	0.868	0.007
			SE	0.184	0.161	0.225	0.125	0.110	0.153	0.120	0.105	0.147
			ESE	0.195	0.208	0.227	0.129	0.124	0.157	0.121	0.112	0.145
			MSE	0.817	4.818	0.051	0.184	1.010	0.023	0.137	0.764	0.022
			WTCR	0.004	0.000	0.949	0.101	0.000	0.941	0.168	0.000	0.946
	0.3	Naive	Bias	1.500	2.998	0.009	0.687	1.385	0.010	0.595	1.202	0.012
			SE	0.190	0.171	0.241	0.125	0.112	0.158	0.119	0.107	0.151
			ESE	0.210	0.221	0.244	0.130	0.136	0.154	0.122	0.120	0.147
			MSE	2.286	9.017	0.058	0.488	1.931	0.025	0.368	1.456	0.023
			WTCR	0.000	0.000	0.946	0.000	0.000	0.952	0.002	0.000	0.955

Table 4.3: Simulation studies for assessing the performance of the RC, EE-known, and EE-estimated methods: stage 1

		regular						weak non-regular						non-regular					
$\pi_{10}=0$		RC		EE-known		EE-estimated		RC		EE-known		EE-estimated		RC		EE-known		EE-estimated	
		ψ_{01}	ψ_{11}	ψ_{01}	ψ_{11}	ψ_{01}	ψ_{11}	ψ_{01}	ψ_{11}	ψ_{01}	ψ_{11}	ψ_{01}	ψ_{11}	ψ_{01}	ψ_{11}	ψ_{01}	ψ_{11}	ψ_{01}	ψ_{11}
$\pi_{01}=0.15$	Bias	0.004	0.006	0.003	0.013	0.010	0.006	0.014	0.006	0.026	0.003	0.025	0.008	0.072	0.005	0.072	0.007	0.062	0.008
	SE	0.225	0.291	0.218	0.329	0.314	0.456	0.131	0.160	0.118	0.176	0.162	0.234	0.134	0.162	0.118	0.176	0.163	0.235
	ESE	0.236	0.291	0.240	0.339	0.223	0.318	0.130	0.160	0.136	0.178	0.133	0.172	0.138	0.162	0.139	0.177	0.136	0.174
	MSE	0.051	0.085	0.047	0.109	0.099	0.208	0.017	0.026	0.015	0.031	0.027	0.055	0.023	0.026	0.019	0.031	0.030	0.056
	PBCR	0.961	0.981	0.973	0.992	0.968	0.983	0.975	0.973	0.973	0.984	0.952	0.969	0.934	0.969	0.936	0.973	0.921	0.954
	DBCR	0.919	0.912	0.884	0.858	0.910	0.900	0.930	0.920	0.928	0.893	0.930	0.911	0.894	0.914	0.890	0.912	0.895	0.909
$\pi_{01}=0.3$	Bias	0.011	0.007	0.017	0.016	0.004	0.001	0.061	0.007	0.084	0.012	0.053	0.005	0.118	0.008	0.144	0.006	0.111	0.005
	SE	0.238	0.303	0.236	0.377	0.328	0.490	0.139	0.176	0.147	0.230	0.181	0.272	0.142	0.178	0.144	0.225	0.180	0.269
	ESE	0.236	0.300	0.251	0.381	0.252	0.370	0.137	0.180	0.156	0.233	0.160	0.223	0.146	0.176	0.157	0.221	0.159	0.218
	MSE	0.057	0.092	0.056	0.143	0.108	0.240	0.023	0.031	0.029	0.053	0.036	0.074	0.034	0.032	0.041	0.051	0.045	0.072
	PBCR	0.974	0.990	0.997	0.998	0.992	0.993	0.957	0.984	0.961	0.996	0.954	0.988	0.870	0.985	0.887	0.995	0.915	0.985
	DBCR	0.918	0.856	0.823	0.745	0.849	0.813	0.873	0.882	0.842	0.817	0.856	0.816	0.823	0.884	0.745	0.835	0.807	0.826
$\pi_{10}=0.15$																			
$\pi_{01}=0$	Bias	0.002	0.002	0.005	0.003	0.021	0.033	0.002	0.002	0.034	0.001	0.032	0.005	0.060	0.002	0.049	0.001	0.027	0.004
	SE	0.223	0.312	0.254	0.387	0.334	0.490	0.127	0.158	0.109	0.167	0.156	0.228	0.132	0.160	0.110	0.169	0.158	0.231
	ESE	0.228	0.315	0.262	0.388	0.236	0.361	0.128	0.159	0.130	0.169	0.125	0.164	0.134	0.158	0.136	0.168	0.127	0.167
	MSE	0.050	0.097	0.065	0.150	0.112	0.241	0.016	0.025	0.013	0.028	0.025	0.052	0.021	0.026	0.014	0.029	0.026	0.054
	PBCR	0.976	0.991	0.983	0.995	0.983	0.981	0.957	0.974	0.964	0.979	0.958	0.974	0.924	0.975	0.960	0.977	0.954	0.962
	DBCR	0.914	0.903	0.901	0.868	0.936	0.884	0.940	0.921	0.921	0.901	0.939	0.923	0.914	0.928	0.905	0.911	0.929	0.916
$\pi_{01}=0.15$	Bias	0.002	0.006	0.001	0.000	0.011	0.015	0.023	0.002	0.007	0.002	0.001	0.003	0.118	0.003	0.008	0.000	0.003	0.002
	SE	0.236	0.330	0.305	0.499	0.373	0.575	0.139	0.183	0.137	0.223	0.174	0.267	0.143	0.185	0.137	0.222	0.175	0.268
	ESE	0.231	0.329	0.298	0.490	0.285	0.459	0.139	0.180	0.161	0.229	0.152	0.211	0.142	0.188	0.150	0.221	0.153	0.213
	MSE	0.056	0.109	0.093	0.249	0.139	0.331	0.020	0.034	0.019	0.050	0.030	0.071	0.034	0.034	0.019	0.049	0.031	0.072
	PBCR	0.978	0.995	1.000	1.000	0.987	0.998	0.956	0.994	0.987	1.000	0.975	0.995	0.914	0.990	0.991	0.997	0.966	0.981
	DBCR	0.890	0.830	0.808	0.714	0.838	0.768	0.905	0.879	0.867	0.793	0.879	0.851	0.832	0.858	0.885	0.810	0.881	0.824
$\pi_{01}=0.3$	Bias	0.004	0.013	0.001	0.004	0.018	0.018	0.057	0.002	0.038	0.017	0.036	0.011	0.162	0.010	0.079	0.005	0.073	0.029
	SE	0.256	0.357	0.378	0.657	0.435	0.707	0.149	0.206	0.194	0.332	0.220	0.356	0.152	0.207	0.188	0.320	0.214	0.347
	ESE	0.255	0.354	0.374	0.648	0.364	0.607	0.145	0.201	0.207	0.335	0.208	0.319	0.153	0.201	0.196	0.307	0.197	0.315
	MSE	0.066	0.128	0.143	0.432	0.189	0.501	0.025	0.042	0.039	0.111	0.050	0.127	0.049	0.043	0.041	0.103	0.051	0.121
	PBCR	0.982	1.000	1.000	1.000	1.000	1.000	0.983	0.997	0.999	1.000	0.985	1.000	0.852	0.995	0.991	1.000	0.980	0.999
	DBCR	0.886	0.793	0.642	0.543	0.704	0.590	0.890	0.830	0.761	0.636	0.774	0.680	0.733	0.832	0.765	0.676	0.780	0.673
$\pi_{10}=0.3$																			
$\pi_{01}=0$	Bias	0.017	0.004	0.006	0.012	0.033	0.047	0.048	0.005	0.077	0.002	0.053	0.014	0.107	0.008	0.115	0.006	0.077	0.002
	SE	0.244	0.331	0.333	0.539	0.388	0.602	0.135	0.175	0.120	0.198	0.162	0.248	0.139	0.176	0.121	0.200	0.165	0.251
	ESE	0.236	0.325	0.316	0.526	0.305	0.479	0.127	0.168	0.139	0.190	0.138	0.193	0.139	0.180	0.144	0.199	0.140	0.196
	MSE	0.060	0.109	0.111	0.290	0.151	0.364	0.021	0.031	0.020	0.039	0.029	0.062	0.031	0.031	0.028	0.040	0.033	0.063
	PBCR	0.984	0.994	0.997	1.000	0.985	0.998	0.964	0.986	0.969	0.995	0.967	0.979	0.890	0.990	0.936	0.997	0.952	0.986
	DBCR	0.920	0.863	0.850	0.753	0.844	0.778	0.922	0.900	0.859	0.859	0.870	0.849	0.863	0.878	0.786	0.847	0.849	0.840
$\pi_{01}=0.15$	Bias	0.013	0.000	0.005	0.019	0.032	0.035	0.089	0.002	0.031	0.009	0.026	0.016	0.164	0.008	0.067	0.013	0.037	0.030
	SE	0.266	0.361	0.460	0.805	0.492	0.826	0.148	0.206	0.168	0.292	0.198	0.324	0.151	0.207	0.166	0.289	0.198	0.323
	ESE	0.274	0.366	0.455	0.810	0.397	0.707	0.151	0.204	0.190	0.292	0.186	0.284	0.157	0.198	0.189	0.281	0.179	0.279
	MSE	0.071	0.130	0.211	0.648	0.243	0.684	0.030	0.042	0.029	0.085	0.040	0.105	0.050	0.043	0.032	0.084	0.041	0.105
	PBCR	0.976	0.999	1.000	1.000	1.000	1.000	0.935	0.991	0.993	1.000	0.989	0.997	0.832	0.994	0.991	1.000	0.983	0.997
	DBCR	0.885	0.789	0.609	0.508	0.707	0.574	0.837	0.824	0.781	0.671	0.794	0.710	0.732	0.842	0.768	0.690	0.800	0.713
$\pi_{01}=0.3$	Bias	0.027	0.013	0.000	0.014	0.035	0.054	0.125	0.020	0.023	0.045	0.019	0.043	0.210	0.016	0.010	0.025	0.037	0.048
	SE	0.291	0.407	0.717	1.322	0.764	1.379	0.158	0.230	0.289	0.526	0.322	0.573	0.161	0.230	0.275	0.499	0.309	0.548
	ESE	0.299	0.408	0.718	1.353	0.677	1.269	0.156	0.230	0.298	0.532	0.316	0.540	0.158	0.223	0.296	0.496	0.296	0.513
	MSE	0.085	0.165	0.514	1.748	0.585	1.905	0.041	0.053	0.084	0.278	0.104	0.330	0.070	0.053	0.076	0.250	0.097	0.303
	PBCR	0.984	1.000	1.000	1.000	1.000	1.000	0.944	0.998	1.000	1.000	1.000	1.000	0.781	1.000	1.000	1.000	0.999	1.000
	DBCR	0.874	0.743	0.401	0.322	0.500	0.366	0.756	0.766	0.600	0.453	0.608	0.458	0.604	0.790	0.602	0.480	0.656	0.522

Table 4.4: Simulation studies for assessing the performance of the RC, EE-known, and EE-estimated methods: stage 2 and regular case

$\pi_{10} = 0$		RC			EE-known			EE-estimated		
		ψ_{02}	ψ_{12}	ψ_{22}	ψ_{02}	ψ_{12}	ψ_{22}	ψ_{02}	ψ_{12}	ψ_{22}
	Bias	0.001	0.002	0.005	0.009	0.018	0.006	0.014	0.020	0.005
	SE	0.184	0.162	0.192	0.182	0.170	0.184	0.206	0.190	0.226
$\pi_{01} = 0.15$	ESE	0.180	0.158	0.193	0.181	0.173	0.187	0.206	0.198	0.170
	MSE	0.034	0.026	0.037	0.033	0.029	0.034	0.043	0.037	0.051
	WTCR	0.955	0.954	0.945	0.948	0.951	0.942	0.955	0.940	0.990
	Bias	0.019	0.033	0.005	0.024	0.024	0.008	0.042	0.045	0.004
	SE	0.219	0.187	0.227	0.259	0.246	0.245	0.270	0.254	0.272
$\pi_{01} = 0.3$	ESE	0.223	0.185	0.237	0.270	0.249	0.242	0.306	0.299	0.223
	MSE	0.048	0.036	0.052	0.068	0.061	0.060	0.075	0.066	0.074
	WTCR	0.943	0.934	0.924	0.947	0.956	0.958	0.925	0.916	0.981
$\pi_{10} = 0.15$										
	Bias	0.010	0.006	0.004	0.002	0.021	0.002	0.002	0.010	0.000
	SE	0.127	0.160	0.191	0.123	0.169	0.183	0.164	0.188	0.226
$\pi_{01} = 0$	ESE	0.129	0.155	0.190	0.124	0.171	0.182	0.113	0.192	0.161
	MSE	0.016	0.026	0.037	0.015	0.029	0.034	0.027	0.035	0.051
	WTCR	0.954	0.956	0.949	0.943	0.952	0.951	0.993	0.946	0.992
	Bias	0.009	0.016	0.006	0.018	0.022	0.006	0.034	0.053	0.010
	SE	0.205	0.240	0.237	0.228	0.264	0.263	0.242	0.266	0.285
$\pi_{01} = 0.15$	ESE	0.206	0.245	0.239	0.226	0.272	0.267	0.259	0.315	0.230
	MSE	0.042	0.058	0.056	0.052	0.070	0.069	0.060	0.074	0.082
	WTCR	0.952	0.940	0.948	0.955	0.942	0.948	0.945	0.901	0.986
	Bias	0.023	0.035	0.009	0.004	0.044	0.020	0.057	0.074	0.020
	SE	0.238	0.269	0.264	0.340	0.381	0.365	0.342	0.373	0.370
$\pi_{01} = 0.3$	ESE	0.234	0.273	0.259	0.341	0.376	0.366	0.394	0.455	0.321
	MSE	0.057	0.074	0.070	0.116	0.147	0.133	0.120	0.144	0.137
	WTCR	0.951	0.946	0.955	0.947	0.950	0.962	0.939	0.902	0.983
$\pi_{10} = 0.3$										
	Bias	0.018	0.022	0.004	0.002	0.024	0.010	0.004	0.042	0.004
	SE	0.145	0.186	0.228	0.151	0.243	0.244	0.183	0.250	0.271
$\pi_{01} = 0$	ESE	0.144	0.185	0.232	0.150	0.248	0.242	0.136	0.298	0.218
	MSE	0.021	0.035	0.052	0.023	0.059	0.060	0.033	0.064	0.073
	WTCR	0.949	0.945	0.946	0.951	0.939	0.940	0.994	0.915	0.988
	Bias	0.018	0.045	0.003	0.028	0.047	0.004	0.029	0.057	0.005
	SE	0.222	0.270	0.264	0.297	0.380	0.364	0.300	0.373	0.370
$\pi_{01} = 0.15$	ESE	0.217	0.266	0.261	0.304	0.385	0.367	0.325	0.451	0.325
	MSE	0.050	0.075	0.070	0.089	0.147	0.133	0.091	0.142	0.137
	WTCR	0.954	0.949	0.951	0.940	0.942	0.958	0.931	0.904	0.981
	Bias	0.023	0.058	0.007	0.066	0.141	0.020	0.099	0.215	0.014
	SE	0.246	0.284	0.284	0.489	0.598	0.558	0.502	0.618	0.563
$\pi_{01} = 0.3$	ESE	0.252	0.287	0.293	0.497	0.577	0.553	0.596	0.792	0.515
	MSE	0.061	0.084	0.080	0.244	0.378	0.311	0.261	0.428	0.317
	WTCR	0.940	0.942	0.943	0.952	0.949	0.967	0.944	0.914	0.992

Table 4.5: Simulation studies for assessing the performance of the RC, EE-known, and EE-estimated methods: stage 2 and weak non-regular case

$\pi_{10} = 0$		RC			EE-known			EE-estimated		
		ψ_{02}	ψ_{12}	ψ_{22}	ψ_{02}	ψ_{12}	ψ_{22}	ψ_{02}	ψ_{12}	ψ_{22}
	Bias	0.007	0.001	0.002	0.012	0.010	0.001	0.011	0.013	0.000
	SE	0.124	0.115	0.143	0.124	0.121	0.141	0.165	0.155	0.199
$\pi_{01} = 0.15$	ESE	0.125	0.113	0.143	0.126	0.119	0.143	0.128	0.128	0.140
	MSE	0.015	0.013	0.020	0.016	0.015	0.020	0.027	0.024	0.039
	WTCR	0.949	0.956	0.954	0.948	0.957	0.947	0.985	0.986	0.990
	Bias	0.007	0.014	0.005	0.004	0.008	0.007	0.008	0.016	0.009
	SE	0.138	0.130	0.154	0.153	0.157	0.159	0.186	0.180	0.210
$\pi_{01} = 0.3$	ESE	0.139	0.131	0.153	0.153	0.161	0.160	0.171	0.174	0.151
	MSE	0.019	0.017	0.024	0.024	0.025	0.025	0.035	0.033	0.044
	WTCR	0.946	0.949	0.943	0.947	0.952	0.947	0.971	0.964	0.990
$\pi_{10} = 0.15$										
	Bias	0.003	0.004	0.000	0.003	0.010	0.001	0.001	0.005	0.001
	SE	0.108	0.115	0.143	0.108	0.121	0.141	0.155	0.154	0.199
$\pi_{01} = 0$	ESE	0.109	0.113	0.143	0.111	0.116	0.143	0.108	0.126	0.138
	MSE	0.012	0.013	0.020	0.012	0.015	0.020	0.024	0.024	0.040
	WTCR	0.945	0.954	0.945	0.937	0.959	0.939	0.997	0.988	0.996
	Bias	0.004	0.010	0.006	0.005	0.007	0.000	0.013	0.013	0.004
	SE	0.134	0.149	0.157	0.143	0.167	0.165	0.178	0.187	0.214
$\pi_{01} = 0.15$	ESE	0.131	0.150	0.152	0.148	0.172	0.167	0.141	0.178	0.147
	MSE	0.018	0.022	0.025	0.021	0.028	0.027	0.032	0.035	0.046
	WTCR	0.954	0.948	0.957	0.945	0.942	0.941	0.984	0.963	0.992
	Bias	0.005	0.001	0.007	0.022	0.040	0.001	0.018	0.042	0.003
	SE	0.147	0.166	0.166	0.192	0.227	0.203	0.216	0.239	0.242
$\pi_{01} = 0.3$	ESE	0.150	0.170	0.168	0.195	0.230	0.206	0.209	0.258	0.184
	MSE	0.022	0.028	0.028	0.037	0.053	0.041	0.047	0.059	0.059
	WTCR	0.934	0.948	0.948	0.939	0.941	0.947	0.974	0.942	0.995
$\pi_{10} = 0.3$										
	Bias	0.007	0.014	0.005	0.001	0.007	0.003	0.002	0.014	0.004
	SE	0.115	0.129	0.154	0.118	0.156	0.159	0.162	0.179	0.210
$\pi_{01} = 0$	ESE	0.113	0.129	0.151	0.117	0.158	0.155	0.113	0.172	0.154
	MSE	0.013	0.017	0.024	0.014	0.024	0.025	0.026	0.032	0.044
	WTCR	0.956	0.955	0.956	0.950	0.949	0.958	0.997	0.957	0.993
	Bias	0.004	0.001	0.008	0.009	0.036	0.013	0.008	0.040	0.003
	SE	0.141	0.167	0.166	0.173	0.227	0.203	0.202	0.239	0.242
$\pi_{01} = 0.15$	ESE	0.144	0.169	0.172	0.177	0.226	0.203	0.183	0.257	0.189
	MSE	0.020	0.028	0.028	0.030	0.053	0.041	0.041	0.059	0.059
	WTCR	0.942	0.952	0.942	0.945	0.947	0.951	0.969	0.938	0.989
	Bias	0.008	0.019	0.002	0.028	0.051	0.004	0.052	0.126	0.005
	SE	0.152	0.180	0.173	0.260	0.343	0.280	0.291	0.365	0.317
$\pi_{01} = 0.3$	ESE	0.156	0.182	0.170	0.271	0.354	0.286	0.329	0.427	0.273
	MSE	0.023	0.033	0.030	0.068	0.120	0.079	0.087	0.149	0.101
	WTCR	0.938	0.948	0.952	0.946	0.946	0.957	0.962	0.923	0.990

Table 4.6: Simulation studies for assessing the performance of the RC, EE-known, and EE-estimated methods: stage 2 and non-regular case

$\pi_{10} = 0$		RC			EE-known			EE-estimated		
		ψ_{02}	ψ_{12}	ψ_{22}	ψ_{02}	ψ_{12}	ψ_{22}	ψ_{02}	ψ_{12}	ψ_{22}
	Bias	0.007	0.005	0.008	0.014	0.006	0.010	0.008	0.011	0.001
	SE	0.119	0.111	0.139	0.119	0.117	0.137	0.163	0.152	0.197
$\pi_{01} = 0.15$	ESE	0.122	0.108	0.142	0.121	0.113	0.140	0.126	0.121	0.134
	MSE	0.014	0.012	0.019	0.014	0.014	0.019	0.026	0.023	0.039
	WTCR	0.939	0.959	0.937	0.936	0.958	0.950	0.988	0.985	0.998
	Bias	0.003	0.012	0.010	0.008	0.005	0.006	0.007	0.010	0.001
	SE	0.130	0.125	0.148	0.144	0.149	0.152	0.179	0.175	0.206
$\pi_{01} = 0.3$	ESE	0.137	0.124	0.147	0.146	0.145	0.154	0.160	0.160	0.148
	MSE	0.017	0.016	0.022	0.021	0.022	0.023	0.032	0.031	0.042
	WTCR	0.937	0.955	0.954	0.943	0.962	0.950	0.967	0.973	0.992
$\pi_{10} = 0.15$										
	Bias	0.004	0.012	0.001	0.002	0.001	0.003	0.001	0.006	0.000
	SE	0.107	0.111	0.139	0.107	0.117	0.137	0.155	0.152	0.197
$\pi_{01} = 0$	ESE	0.105	0.111	0.135	0.106	0.118	0.137	0.103	0.120	0.135
	MSE	0.011	0.012	0.019	0.011	0.014	0.019	0.024	0.023	0.039
	WTCR	0.954	0.945	0.952	0.950	0.947	0.949	0.996	0.985	0.996
	Bias	0.006	0.004	0.010	0.005	0.015	0.013	0.014	0.026	0.004
	SE	0.127	0.141	0.150	0.136	0.158	0.157	0.173	0.181	0.209
$\pi_{01} = 0.15$	ESE	0.126	0.140	0.150	0.136	0.156	0.155	0.137	0.167	0.145
	MSE	0.016	0.020	0.023	0.019	0.025	0.025	0.030	0.033	0.044
	WTCR	0.959	0.943	0.957	0.951	0.956	0.950	0.986	0.962	0.997
	Bias	0.006	0.010	0.001	0.019	0.029	0.002	0.024	0.037	0.003
	SE	0.138	0.158	0.157	0.177	0.213	0.187	0.205	0.227	0.231
$\pi_{01} = 0.3$	ESE	0.142	0.161	0.157	0.175	0.216	0.183	0.187	0.225	0.173
	MSE	0.019	0.025	0.025	0.032	0.046	0.035	0.042	0.053	0.053
	WTCR	0.942	0.937	0.949	0.961	0.945	0.958	0.978	0.956	0.992
$\pi_{10} = 0.3$										
	Bias	0.009	0.013	0.005	0.001	0.006	0.005	0.002	0.011	0.002
	SE	0.112	0.124	0.148	0.116	0.148	0.152	0.161	0.174	0.206
$\pi_{01} = 0$	ESE	0.111	0.123	0.145	0.116	0.152	0.150	0.116	0.160	0.146
	MSE	0.013	0.016	0.022	0.014	0.022	0.023	0.026	0.030	0.042
	WTCR	0.947	0.954	0.958	0.949	0.947	0.960	0.993	0.976	0.997
	Bias	0.002	0.006	0.004	0.013	0.032	0.003	0.017	0.041	0.003
	SE	0.134	0.157	0.157	0.162	0.212	0.187	0.193	0.227	0.231
$\pi_{01} = 0.15$	ESE	0.139	0.159	0.158	0.158	0.206	0.184	0.171	0.233	0.175
	MSE	0.018	0.025	0.025	0.026	0.046	0.035	0.038	0.053	0.053
	WTCR	0.935	0.951	0.947	0.947	0.952	0.956	0.977	0.954	0.992
	Bias	0.000	0.021	0.012	0.029	0.050	0.001	0.055	0.110	0.001
	SE	0.143	0.171	0.162	0.237	0.319	0.252	0.265	0.337	0.290
$\pi_{01} = 0.3$	ESE	0.148	0.177	0.160	0.236	0.310	0.246	0.272	0.377	0.231
	MSE	0.021	0.030	0.027	0.057	0.104	0.064	0.073	0.126	0.084
	WTCR	0.941	0.937	0.951	0.963	0.956	0.967	0.971	0.941	0.996

Table 4.7: Prediction performance of the naive, RC, EE-known, and EE-estimated methods: proportions of optimally treated individuals

			Scenario 1								Scenario 2							
			Stage 1				Stage 2				Stage 1				Stage 2			
	π_{01}	Setting	Naive	RC	EE-known	EE-estimated	Naive	RC	EE-known	EE-estimated	Naive	RC	EE-known	EE-estimated	Naive	RC	EE-known	EE-estimated
$\pi_{10} = 0$	0.15	Regular	0.772	0.858	0.862	0.864	1.000	1.000	1.000	1.000	0.730	0.769	0.803	0.824	0.925	0.947	0.925	0.947
		Weak non-regular	0.906	0.924	0.907	0.926	1.000	1.000	1.000	1.000	0.852	0.876	0.920	0.921	0.924	0.946	0.925	0.947
		Non-regular	0.995	0.996	0.997	0.997	0.873	0.875	0.874	0.872	0.921	0.946	0.925	0.947	0.865	0.871	0.749	0.789
	0.3	Regular	0.698	0.862	0.855	0.862	1.000	1.000	1.000	1.000	0.629	0.701	0.738	0.771	0.849	0.894	0.849	0.894
		Weak non-regular	0.848	0.884	0.844	0.876	0.999	0.998	0.998	0.998	0.749	0.795	0.850	0.856	0.848	0.894	0.849	0.894
		Non-regular	0.977	0.990	0.981	0.992	0.876	0.879	0.874	0.872	0.831	0.887	0.849	0.893	0.836	0.852	0.749	0.785
$\pi_{10} = 0.15$	0	Regular	0.857	0.859	0.847	0.857	0.994	1.000	1.000	1.000	0.810	0.821	0.847	0.853	0.917	0.930	0.925	0.948
		Weak non-regular	0.716	0.944	0.965	0.961	0.878	0.997	0.997	0.997	0.701	0.764	0.895	0.912	0.824	0.875	0.924	0.946
		Non-regular	0.934	0.999	1.000	1.000	0.998	0.872	0.874	0.869	0.867	0.912	0.925	0.948	0.960	0.934	0.817	0.829
	0.15	Regular	0.768	0.860	0.839	0.844	0.961	1.000	1.000	1.000	0.678	0.737	0.763	0.788	0.820	0.858	0.850	0.895
		Weak non-regular	0.620	0.910	0.922	0.925	0.840	0.998	0.996	0.997	0.586	0.692	0.848	0.876	0.738	0.817	0.850	0.894
		Non-regular	0.819	0.990	0.998	0.999	0.999	0.883	0.875	0.876	0.723	0.826	0.850	0.894	0.924	0.909	0.675	0.736
	0.3	Regular	0.664	0.844	0.774	0.803	0.858	1.000	1.000	1.000	0.592	0.668	0.698	0.736	0.687	0.773	0.774	0.842
		Weak non-regular	0.547	0.864	0.865	0.881	0.795	0.995	0.992	0.991	0.521	0.627	0.773	0.806	0.664	0.762	0.775	0.840
		Non-regular	0.647	0.966	0.975	0.989	0.999	0.879	0.869	0.872	0.581	0.730	0.775	0.840	0.886	0.885	0.674	0.734
$\pi_{10} = 0.3$	0	Regular	0.825	0.851	0.823	0.847	0.573	1.000	1.000	1.000	0.738	0.774	0.784	0.803	0.556	0.705	0.851	0.896
		Weak non-regular	0.515	0.876	0.964	0.958	0.752	0.996	0.994	0.994	0.510	0.623	0.823	0.859	0.677	0.775	0.847	0.890
		Non-regular	0.651	0.991	1.000	0.999	1.000	0.873	0.862	0.867	0.612	0.748	0.850	0.895	0.926	0.910	0.752	0.788
	0.15	Regular	0.703	0.830	0.767	0.780	0.521	1.000	1.000	1.000	0.614	0.688	0.690	0.727	0.511	0.665	0.775	0.842
		Weak non-regular	0.503	0.819	0.923	0.919	0.738	0.994	0.991	0.987	0.501	0.594	0.771	0.819	0.631	0.739	0.775	0.840
		Non-regular	0.536	0.971	1.000	0.999	1.000	0.878	0.878	0.869	0.518	0.675	0.775	0.842	0.888	0.885	0.601	0.682
	0.3	Regular	0.629	0.826	0.707	0.737	0.501	1.000	0.997	0.996	0.556	0.633	0.589	0.636	0.499	0.651	0.700	0.789
		Weak non-regular	0.500	0.767	0.854	0.860	0.705	0.993	0.979	0.977	0.500	0.583	0.692	0.741	0.580	0.701	0.700	0.783
		Non-regular	0.505	0.940	0.957	0.971	0.999	0.879	0.868	0.868	0.503	0.637	0.697	0.779	0.850	0.859	0.599	0.678

Table 4.8: Analyses 1 and 2 results for STAR*D data: DB 95% CIs are included for stage 1 parameters and W-type 95% CIs are reported for stage 2 parameters

		Stage 1			Stage 2		
		ψ_{01}	ψ_{11}	ψ_{21}	ψ_{02}	ψ_{12}	ψ_{22}
	Estimate	-3.225	0.172	0.665	-0.159	0.019	-0.801
Analysis 1	SE	2.097	0.100	1.955	0.845	0.058	0.497
	95% CI	(-5.516, -0.985)	(-0.042, 0.369)	(-2.307, 2.640)	(-1.815, 1.496)	(-0.095, 0.133)	(-1.774, 0.173)
Analysis 2							
(π_{10}, π_{01})							
	Estimate	-3.227	0.171	0.674	-0.156	0.019	-0.819
(0.01, 0.01)	SE	1.203	0.101	1.197	0.785	0.062	0.500
	95% CI	(-3.533, -2.601)	(-0.008, 0.269)	(0.198, 1.271)	(-1.700, 1.380)	(-0.102, 0.141)	(-1.800, 0.161)
	Estimate	-3.240	0.166	0.761	-0.150	0.021	-0.857
(0.03,0.03)	SE	3.607	0.144	5.341	0.786	0.062	0.523
	95% CI	(-4.164, -2.124)	(-0.025, 0.270)	(-0.081, 1.815)	(-1.690, 1.390)	(-0.101, 0.143)	(-1.880, 0.168)
	Estimate	-3.400	0.130	1.434	-0.137	0.024	-0.947
(0.07,0.07)	SE	1.789	0.090	1.091	0.787	0.063	0.578
	95% CI	(-4.152, -3.004)	(-0.108, 0.223)	(0.867, 1.996)	(-1.680, 1.410)	(-0.100, 0.147)	(-2.080, 0.186)
	Estimate	-2.547	0.185	-0.152	-0.128	0.027	-1.028
(0.1,0.1)	SE	1.547	0.093	0.787	0.790	0.064	0.628
	95% CI	(-5.240, 0.520)	(-0.303, 0.297)	(-1.780, 1.460)	(-1.680, 1.420)	(-0.098, 0.152)	(-2.260, 0.202)

Table 4.9: Regression parameter values that are set for Analysis 3 in Section 4.5.3

	Stage 1 parameters			Stage 2 parameters					
	ζ_{01}	ζ_{11}	ζ_{21}	ζ_{02}	ζ_{12}	ζ_{22}	ζ_{32}	ζ_{42}	ζ_{52}
Set 1	-4.61	4.03	0.09	-3.66	4.24	0.11	0.08	-0.06	0.11
Set 2	-3.81	2.47	0.09	-2.78	2.55	0.11	0.17	-0.06	0.06
Set 3	-3.37	1.52	0.09	-2.32	1.57	0.11	0.16	-0.06	0.05
Set 4	-3.01	0.80	0.09	-1.97	0.84	0.11	0.19	-0.06	0.02

Table 4.10: Analysis 3 results for STAR*D data: DB 95% CIs are included for stage 1 parameters and Normal bootstrap 95% CIs are reported for stage 2 parameters

		Stage 1			Stage 2		
		ψ_{01}	ψ_{11}	ψ_{21}	ψ_{02}	ψ_{12}	ψ_{22}
Set 1	Estimate	-2.973	0.162	0.840	-0.271	0.025	-0.945
	Bootstrap SE	1.671	0.375	1.676	0.804	0.064	0.693
	95% CI	(-5.710, 0.447)	(-0.129, 0.443)	(-2.988, 6.563)	(-1.846, 1.305)	(-0.099, 0.150)	(-2.304, 0.414)
Set 2	Estimate	-2.871	0.157	1.011	-0.333	0.031	-1.143
	Bootstrap SE	1.488	1.378	2.046	0.777	0.063	0.951
	95% CI	(-5.513, -0.041)	(-0.161, 0.515)	(-2.739, 7.254)	(-1.857, 1.190)	(-0.093, 0.154)	(-3.008, 0.721)
Set 3	Estimate	-2.827	0.151	1.382	-0.369	0.032	-1.264
	Bootstrap SE	1.660	0.497	2.357	0.810	0.070	1.459
	95% CI	(-5.024, 0.113)	(-0.471, 0.450)	(-4.591, 10.848)	(-1.957, 1.220)	(-0.105, 0.169)	(-4.123, 1.595)
Set 4	Estimate	-2.791	0.133	2.303	-0.388	0.019	-0.779
	Bootstrap SE	1.423	0.427	2.850	0.786	0.074	2.370
	95% CI	(-5.223 -0.154)	(-0.755 0.729)	(-3.506, 7.012)	(-1.928, 1.152)	(-0.125, 0.164)	(-5.424, 3.866)

Chapter 5

Q-Learning with Compound Outcome and Mixed Misclassification and Measurement Error in Covariates

5.1 Introduction

In Chapter 4, we investigate the detrimental impact of covariate misclassification on Q-learning with a univariate outcome, and propose two correction strategies to reduce the bias. Although most studies on covariate measurement error and misclassification investigate these issues separately, it is not uncommon to see real-world data with both covariate measurement error and misclassification. Correcting mixed measurement error and misclassification in covariates have been discussed under some regression models (Yi et al., 2015; Spiegelman et al., 2000; Zhang and Yi, 2023). However, no research on such issues is available for developing optimal DTRs via Q-learning. Unlike the setup in Chapter 4 which considers Q-learning with a univariate outcome, this chapter deals with Q-learning with bivariate outcomes, and we propose correction strategies to account for the potential bias induced by mixed misclassification and measurement error in covariates.

The remainder of the chapter is organized as follows. In Section 5.2, we establish the basic notation and outline the implementation procedure of Q-learning for bivariate outcomes in an

ideal scenario without covariate mismeasurement. In Section 5.3, we introduce a setting with covariate mismeasurement and conduct simulation studies to illustrate the impact of ignoring mismeasurement on the Q-learning estimation process. In Sections 5.4 and 5.5, we respectively present two correction methods for mitigating the bias caused by mismeasurement. In Section 5.6, we assess the performance of the proposed methods through simulation studies. In Section 5.7, the proposed methods are applied to real data to demonstrate their utility. In Section 5.8, we conclude the chapter with discussions.

5.2 Q-Learning with Bivariate Loss/Reward Functions

Most DTR applications focus on univariate outcomes. However, in some DTR applications, a single outcome is not adequate to fully represent the aspects of the problem. In what follows, we describe the Q-learning approach with bivariate outcomes. Suppose that the study has K stages. For $k = 1, \cdots, K$, let A_k denote the binary action taken at stage k, and let X_k and C_k respectively denote error-prone binary and continuous covariates, which are both scalar, where X_k takes on value 0 or 1. Let Z_k denote the vector of precisely measured covariates. Covariates associated with each stage are measured prior to the receipt of the treatment at that stage. Similar to the notations of univariate Q-learning introduced in Section 1.2.2, for $k = 1, \cdots, K$, let $\overline{X}_k = \left\{X_1, \cdots, X_k\right\}$, $\overline{C}_k = \left\{C_1, \cdots, C_k\right\}$, $\overline{Z}_k = \left\{Z_1, \cdots, Z_k\right\}$, and $\overline{A}_k = \left\{A_1, \cdots, A_k\right\}$.

For $k = 1, \cdots, K$, let the first and second outcomes at the end of stage k be denoted by Y_{k1} and Y_{k2}, respectively, and assume larger cumulative outcomes are more desirable. The first and second outcomes can be represented as a function, say $g_j(\cdot)$ with $j = 1, 2$, of the history of the treatment, $\overline{A}_{k-1}$, together with the current treatment A_k, and the history of the covariates, $\overline{X}_k$, $\overline{C}_k$ and $\overline{Z}_k$, as well as the covariates X_{k+1}, C_{k+1} and Z_{k+1} in the next stage. That is,

$$Y_{kj} = g_j(\overline{A}_k, \overline{X}_{k+1}, \overline{C}_{k+1}, \overline{Z}_{k+1}) \tag{5.1}$$

for $j = 1, 2$ and $k = 1, \cdots, K$, where X_{k+1}, C_{k+1} and Z_{k+1} are null when $k = K$.

5.2.1 Composite Q-Function

To construct a sequence of optimal decision rules, we first define the Q-functions for the first and second outcomes for K stages separately for $j = 1, 2$:

$$\begin{aligned} Q_K^j(\overline{A}_K, \overline{X}_K, \overline{C}_K, \overline{Z}_K) &= E(Y_{Kj} \mid \overline{A}_K, \overline{X}_K, \overline{C}_K, \overline{Z}_K); \\ Q_k^j(\overline{A}_k, \overline{X}_k, \overline{C}_k, \overline{Z}_k) &= E\Big\{Y_{kj} + \max_{a_{k+1}} E(Y_{(k+1)j} \mid \overline{A}_k, \overline{X}_{k+1}, \overline{C}_{k+1}, \overline{Z}_{k+1}, a_{k+1}) \,\Big|\, \overline{A}_k, \overline{X}_k, \overline{C}_k, \overline{Z}_k\Big\} \\ &\qquad \text{for } k = K-1, \cdots, 1. \end{aligned} \tag{5.2}$$

To delineate these conditional expectations, we can employ regression approaches, such as linear regression models. For $j = 1, 2$, consider the regression model

$$Q_k^j(\overline{A}_k, \overline{X}_k, \overline{C}_k, \overline{Z}_k) = \beta_{kj}^{\mathrm{T}} H_{k0} + (\psi_{kj}^{\mathrm{T}} H_{k1}) A_k \quad \text{for} \quad k = K, \cdots, 1, \tag{5.3}$$

where we rewrite $\overline{A}_{k-1} \cup \overline{X}_k \cup \overline{C}_k \cup \overline{Z}_k$ as $\{H_{k0}, H_{k1}\}$, with H_{k0} representing the covariates that have a predictive effect on the outcome, and H_{k1} standing for the covariates that interact with the treatment; H_{k0} and H_{k1} may include a constant, or intercept, term, and they may include the same covariates. For $j = 1, 2$ and $k = 1, \cdots, K$, β_{kj} and ψ_{kj} are the regression coefficients, and we write $\theta_{k1} = (\beta_{k1}^{\mathrm{T}}, \psi_{k1}^{\mathrm{T}})^{\mathrm{T}}$ and $\theta_{k2} = (\beta_{k2}^{\mathrm{T}}, \psi_{k2}^{\mathrm{T}})^{\mathrm{T}}$.

When interest lies in finding an optimal treatment in response to either the first outcome or the second outcome separately, we may invoke optimal decision rules to the Q-functions in (5.3) separately for $j = 1, 2$, with $k = K, \cdots, 1$. That is, for $j = 1, 2$, consider $d_{jk} = \arg\max_{a_k} Q_k^j(\overline{A}_{k-1}, \overline{X}_k, \overline{C}_k, \overline{Z}_k, a_k)$, which we call them marginally optimal decision rules. It is noted that the optimal treatment for the first outcome is not necessarily optimal for the second outcome, and vice versa. That is, considering

$$\hat{d}_{k+1} = \arg\max_{a_{k+1}} \begin{pmatrix} Q_{k+1}^1(\overline{A}_k, \overline{X}_{k+1}, \overline{C}_{k+1}, \overline{Z}_{k+1}, a_{k+1}; \hat{\theta}_{(k+1)1}) \\ Q_{k+1}^2(\overline{A}_k, \overline{X}_{k+1}, \overline{C}_{k+1}, \overline{Z}_{k+1}, a_{k+1}; \hat{\theta}_{(k+1)2}) \end{pmatrix} \quad \text{for} \quad k = K-1, \cdots, 0,$$

may turn out to be fruitless because there may not exist a single decision rule that maximizes both objective functions simultaneously. To get around this problem, we aim to identify an

optimal treatment in maximizing both first and second outcomes in a combined format.

To take into account both outcomes simultaneously, we introduce a pre-specified weight parameter δ, taking a value between 0 and 1, and define the composite Q-function as:

$$Q_k(\overline{A}_k, \overline{X}_k, \overline{C}_k, \overline{Z}_k, \delta) = \delta Q_k^1(\overline{A}_k, \overline{X}_k, \overline{C}_k, \overline{Z}_k) + (1-\delta) Q_k^2(\overline{A}_k, \overline{X}_k, \overline{C}_k, \overline{Z}_k) \quad \text{for } k = K, \cdots, 1, \tag{5.4}$$

where $Q_k^j(\overline{A}_k, \overline{X}_k, \overline{C}_k, \overline{Z}_k)$ is determined by (5.2) for $j = 1, 2$ and $k = K, \cdots, 1$.

The weight parameter δ allows us to adjust the relative importance of each outcome in the composite Q-function. When δ is close to 1, the composite Q-function puts more emphasis on the Q-function for the first outcome, whereas setting δ close to 0 gives more weight to the Q-function for the second outcome. By combining the Q-functions using a weight parameter δ, we find a way to balance the two objectives in order to choose a decision rule that is "good enough" with respect to both $Q_{k+1}^1(\cdot)$ and $Q_{k+1}^2(\cdot)$. By adjusting the value of δ, we can prioritize one objective over the other, or find a trade-off between the two.

Consequently, for any given δ, combined optimal decision rules, denoted d_k, are determined by

$$d_k = \arg\max_{a_k} Q_k(\overline{A}_{k-1}, \overline{X}_k, \overline{C}_k, \overline{Z}_k, a_k, \delta) \qquad \text{for} \quad k = K, \cdots, 1. \tag{5.5}$$

The implementation of (5.5) hinges on the knowledge of the Q-functions for the first and second outcomes, $Q_k^j(\overline{A}_k, \overline{X}_k, \overline{C}_k, \overline{Z}_k)$ with $j = 1, 2$ and $k = K, \cdots, 1$, which can be typically modeled by regression models, such as (5.3). That is, the composite Q-function (5.4) for $k = K, \cdots, 1$, can be modeled as

$$Q_k(\overline{A}_k, \overline{X}_k, \overline{C}_k, \overline{Z}_k, \delta) = \delta \left\{ \beta_{k1}^{\mathrm{T}} H_{k0} + (\psi_{k1}^{\mathrm{T}} H_{k1}) A_k \right\} + (1-\delta) \left\{ \beta_{k2}^{\mathrm{T}} H_{k0} + (\psi_{k2}^{\mathrm{T}} H_{k1}) A_k \right\}. \tag{5.6}$$

Parameter Determination

To estimate an optimal DTR, we first estimate the model parameters in the composite Q-functions in (5.4), which can be achieved by separately estimating parameters associated with (5.3) using a dataset consisting of n i.i.d. trajectories, each of the form $\{A_{ki}, X_{ki}, C_{ki}, Z_{ki}, Y_{k1i}, Y_{k2i}\}$,

where $i = 1, \cdots, n$. Let $D \triangleq \Big\{\{A_{ki}, X_{ki}, C_{ki}, Z_{ki}, Y_{k1i}, Y_{k2i}\} : k = 1, \cdots, K; i = 1, \cdots, n\Big\}$.

To be specific, for $\delta = 1$ and $\delta = 0$, we respectively solve

$$\hat{\theta}_{k1} = \arg\min_{\theta_{k1}} \frac{1}{n} \sum_{i=1}^{n} \Big[\hat{Y}_{k1i} - Q_k^1(\overline{A}_{ki}, \overline{X}_{ki}, \overline{C}_{ki}, \overline{Z}_{ki}; \theta_{k1})\Big]^2 \tag{5.7}$$

$$\hat{\theta}_{k2} = \arg\min_{\theta_{k2}} \frac{1}{n} \sum_{i=1}^{n} \Big[\hat{Y}_{k2i} - Q_k^2(\overline{A}_{ki}, \overline{X}_{ki}, \overline{C}_{ki}, \overline{Z}_{ki}; \theta_{k2})\Big]^2, \tag{5.8}$$

where for stage K, $\hat{Y}_{K1i} = Y_{K1i}$, $\hat{Y}_{K2i} = Y_{K2i}$; and for $k = K-1, \cdots, 1$,

$$\hat{Y}_{k1i} = Y_{k1i} + Q_{k+1}^1(\overline{A}_{ki}, \overline{X}_{(k+1)i}, \overline{C}_{(k+1)i}, \overline{Z}_{(k+1)i}, \hat{d}_{(k+1)i}; \hat{\theta}_{(k+1)1}) \tag{5.9}$$

and

$$\hat{Y}_{k2i} = Y_{k2i} + Q_{k+1}^2(\overline{A}_{ki}, \overline{X}_{(k+1)i}, \overline{C}_{(k+1)i}, \overline{Z}_{(k+1)i}, \hat{d}_{(k+1)i}; \hat{\theta}_{(k+1)2}) \tag{5.10}$$

represent stage k pseudo-outcomes for subject i, and the $\hat{d}_{(k+1)i}$ are obtained by solving

$$\begin{aligned}\hat{d}_{k+1} = \arg\max_{a_{k+1}} \{&\delta Q_{k+1}^1(\overline{A}_k, \overline{X}_{k+1}, \overline{C}_{k+1}, \overline{Z}_{k+1}, a_{k+1}; \hat{\theta}_{(k+1)1}) \\ &+ (1-\delta) Q_{k+1}^2(\overline{A}_k, \overline{X}_{k+1}, \overline{C}_{k+1}, \overline{Z}_{k+1}, a_{k+1}; \hat{\theta}_{(k+1)2})\} \quad \text{for} \quad k = K-1, \cdots, 0,\end{aligned} \tag{5.11}$$

which attains at either 0 or 1.

5.2.2 Estimation Equation Method

Alternatively, model parameters in the Q-functions can be estimated using the estimating equation approach. In what follows, we first discuss the estimation for stage K, utilizing the outcome measurements $\{Y_{Kji} : j = 1, 2; i = 1, \cdots, n\}$, and then using the pseudo-outcomes (5.9) and (5.10), we explain the estimation process for stage k where $k = K-1, \cdots, 1$.

For stage K and $j = 1, 2$, set $\ell_{Kji} = \{Y_{Kji} - Q_K^j(\overline{A}_{Ki}, \overline{X}_{Ki}, \overline{C}_{Ki}, \overline{Z}_{Ki}; \theta_{Kj})\}^2$. Define

$$S_{Kj}(\theta_{Kj}; Y_{Kji}, \overline{A}_{Ki}, \overline{X}_{Ki}, \overline{C}_{Ki}, \overline{Z}_{Ki}) = \left(\frac{\partial \ell_{Kji}}{\partial \beta_{Kj}^{\mathrm{T}}}, \frac{\partial \ell_{Kji}}{\partial \psi_{Kj}^{\mathrm{T}}}\right)^{\mathrm{T}} \triangleq$$
$$(S_{K\beta_{Kj}}^{\mathrm{T}}(\theta_{Kj}; Y_{Kji}, \overline{A}_{Ki}, \overline{X}_{Ki}, \overline{C}_{Ki}, \overline{Z}_{Ki}), S_{K\psi_{Kj}}^{\mathrm{T}}(\theta_{Kj}; Y_{Kji}, \overline{A}_{Ki}, \overline{X}_{Ki}, \overline{C}_{Ki}, \overline{Z}_{Ki}))^{\mathrm{T}},$$

where

$$S_{K\beta_{Kj}}(\theta_{Kj}; Y_{Kji}, \overline{A}_{Ki}, \overline{X}_{Ki}, \overline{C}_{Ki}, \overline{Z}_{Ki}) =$$
$$\{Y_{Kji} - Q_K^j(\overline{A}_{Ki}, \overline{X}_{Ki}, \overline{C}_{Ki}, \overline{Z}_{Ki}; \theta_{Kj})\}\frac{\partial Q_K^j(\overline{A}_{Ki}, \overline{X}_{Ki}, \overline{C}_{Ki}, \overline{Z}_{Ki}; \theta_{Kj})}{\partial \beta_{Kj}} \quad (5.12)$$

and

$$S_{K\psi_{Kj}}(\theta_{Kj}; Y_{Ki}, \overline{A}_{Ki}, \overline{X}_{Ki}, \overline{C}_{Ki}, \overline{Z}_{Ki}) =$$
$$\{Y_{Kji} - Q_K^j(\overline{A}_{Ki}, \overline{X}_{Ki}, \overline{C}_{Ki}, \overline{Z}_{Ki}; \theta_{Kj})\}\frac{\partial Q_K^j(\overline{A}_{Ki}, \overline{X}_{Ki}, \overline{C}_{Ki}, \overline{Z}_{Ki}; \theta_{Kj})}{\partial \psi_{Kj}}. \quad (5.13)$$

With (5.3) employed, (5.12) and (5.13) are simplified as

$$S_{K\beta_{Kj}}(\theta_{Kj}; Y_{Kji}, \overline{A}_{Ki}, \overline{X}_{Ki}, \overline{C}_{Ki}, \overline{Z}_{Ki}) = \left[Y_{Kji} - \left\{\beta_{Kj}^{\mathrm{T}} H_{K0} + (\psi_{Kj}^{\mathrm{T}} H_{K1}) A_K\right\}\right] H_{K0} \quad (5.14)$$

and

$$S_{K\psi_{Kj}}(\theta_{Kj}; Y_{Ki}, \overline{A}_{Ki}, \overline{X}_{Ki}, \overline{C}_{Ki}, \overline{Z}_{Ki}) = \left[Y_{Kji} - \left\{\beta_{Kj}^{\mathrm{T}} H_{K0} + (\psi_{Kj}^{\mathrm{T}} H_{K1}) A_K\right\}\right] H_{K1} A_K. \quad (5.15)$$

Both $S_{K\beta_{Kj}}(\theta_{Kj}; Y_{Kji}, \overline{A}_{Ki}, \overline{X}_{Ki}, \overline{C}_{Ki}, \overline{Z}_{Ki})$ and $S_{K\psi_{Kj}}(\theta_{Kj}; Y_{Kji}, \overline{A}_{Ki}, \overline{X}_{Ki}, \overline{C}_{Ki}, \overline{Z}_{Ki})$ are unbiased estimating functions, and thus, solving the estimating equations

$$\sum_{i=1}^{n} S_{Kj}(\theta_{Kj}; Y_{Ki}, \overline{A}_{Ki}, \overline{X}_{Ki}, \overline{C}_{Ki}, \overline{Z}_{Ki}) = 0 \quad (5.16)$$

for θ_{Kj} yields consistent estimator of θ_{Kj}, provided regularity conditions.

Similarly, for stage k with $k = K-1, \cdots, 1$, and $j = 1, 2$, let

$$\hat{\ell}_{kji} = \left\{ \hat{Y}_{kji} - Q_k^j(\overline{A}_{ki}, \overline{X}_{ki}, \overline{C}_{ki}, \overline{Z}_{ki}; \theta_{kj}) \right\}^2$$

and define

$$\hat{S}_{kj}(\theta_{kj}; \hat{Y}_{kji}, \overline{A}_{ki}, \overline{X}_{ki}, \overline{C}_{ki}, \overline{Z}_{ki}) = \left(\frac{\partial \hat{\ell}_{kji}}{\partial \beta_{kj}^{\mathrm{T}}}, \frac{\partial \hat{\ell}_{kji}}{\partial \psi_{kj}^{\mathrm{T}}} \right)^{\mathrm{T}}.$$

Then, an estimator of θ_{kj}, denoted $\hat{\theta}_{kj}$, is obtained by solving

$$\sum_{i=1}^{n} \hat{S}_{kj}(\theta_{kj}; \hat{Y}_{kji}, \overline{A}_{ki}, \overline{X}_{ki}, \overline{C}_{ki}, \overline{Z}_{ki}) = 0 \tag{5.17}$$

for θ_{kj}.

5.3 Mismeasurement and Naive Analysis

In this section, we consider the case where X_k and C_k are subject to mismeasurement for $k = 1, \cdots, K$. We examine numerically the impact of naively implementing the Q-learning procedure in Section 5.2, with the mismeasurement effects ignored.

5.3.1 Measurement Error and Misclassification Models

For $k = 1, \cdots, K$, let X_k^* and C_k^*, respectively, denote the observed versions of the true covariates X_k and C_k, and let $\overline{X}_k^* = \left\{ X_1^*, \cdots, X_k^* \right\}$ and $\overline{C}_k^* = \left\{ C_1^*, \cdots, C_k^* \right\}$.

For $j = 0, 1$ and $l = 1 - j$, let $\pi_{jl,i} = P(X_{ki}^* = j \mid \overline{A}_{ki}, \overline{X}_{(k-1)i}^*, \overline{X}_{(k-1)i}, X_{ki} = l, \overline{C}_{ki}^*, \overline{C}_{ki}, \overline{Z}_{ki})$ denote the misclassification probabilities that may depend on either the true or mismeasured covariates or both. To model the misclassification probabilities, one may employ regression models for binary data, such as logistic regression models. For the misclassification process, it is convenient to assume that $P(X_{ki}^* = j \mid \overline{A}_{ki}, \overline{X}_{(k-1)i}^*, \overline{X}_{(k-1)i}, X_{ki} = l, \overline{C}_{ki}^*, \overline{C}_{ki}, \overline{Z}_{ki}) = P(X_{ki}^* = j \mid X_{ki} = l)$, enabling us to express misclassification probabilities as

$$\Pi = \begin{pmatrix} 1 - \pi_{10} & \pi_{01} \\ \pi_{10} & 1 - \pi_{01} \end{pmatrix}, \tag{5.18}$$

which corresponds to homogeneous misclassification across all subjects.

Now, we describe the measurement error process for the continuous covariates. Let $h(C^*_{ki} \mid \overline{A}_{ki}, \overline{C}^*_{(k-1)i}, \overline{C}_{ki}, \overline{X}^*_{ki}, \overline{X}_{ki}, \overline{Z}_{ki})$ denote the conditional probability density function of C^*_{ki}, given $\overline{A}_{ki} \cup \overline{C}^*_{(k-1)i} \cup \overline{C}_{ki} \cup \overline{X}^*_{ki} \cup \overline{X}_{ki} \cup \overline{Z}_{ki}$. Similar to the misclassification case, one may assume that $h(C^*_{ki} \mid \overline{A}_{ki}, \overline{C}^*_{(k-1)i}, \overline{C}_{ki}, \overline{X}^*_{ki}, \overline{X}_{ki}, \overline{Z}_{ki}) = h(C^*_{ki} \mid C_{ki})$ for simplicity. We can then use a parametric model, say $f(C^*_{ki} \mid C_{ki}; \alpha)$ with parameter α, to modulate $h(C^*_{ki} \mid C_{ki})$. For example, we consider the regression model

$$C^*_{ki} = \xi_0 + \xi_1 C_{ki} + e_{ki},$$

where the error terms e_{ki} have mean 0 and variance σ^2_k, ξ_0 is the intercept, and ξ_1 is the coefficient. In this instance, α included ξ_0, ξ_1 and σ^2_k for $k = 1, \cdots, K$. Setting ξ_0 to be a zero vector and ξ_1 to be the identity matrix gives a classical additive model of the form

$$C^*_{ki} = C_{ki} + e_{ki}. \tag{5.19}$$

Similar to the development in Section 4.2.1, we assume that for $k = 1, \cdots, K$ and $j = 1, 2$,

$$h(Y_{kj} \mid \overline{Y}_{(k-1)j}, \overline{A}_K, \overline{X}_K, \overline{C}_K, \overline{Z}_K, \overline{X}^*_K, \overline{C}^*_K) = h(Y_{kj} \mid \overline{A}_k, \overline{X}_{k+1}, \overline{C}_{k+1}, \overline{Z}_{k+1}). \tag{5.20}$$

Assumption (5.20) says that at stage k, given the history of outcomes $\overline{Y}_{(k-1)j}$ and the information $\overline{A}_K \cup \overline{X}_K \cup \overline{C}_K \cup \overline{Z}_K \cup \overline{X}^*_K \cup \overline{C}^*_K$ over the entire course, the conditional distribution of Y_{kj} depends only on the history $\overline{A}_k \cup \overline{X}_k \cup \overline{C}_k \cup \overline{Z}_k$ at stage k as well as the covariates X_{k+1}, C_{k+1} and Z_{k+1} at the next stage.

5.3.2 Naive Q-Learning Procedure

Ignoring mismeasurement, one may naively use X^*_k and C^*_k to replace the unobserved true covariates X_k and C_k to repeat the Q-learning procedure in Section 5.2, with naive Q-functions defined as

$$
\begin{aligned}
Q_K^{j*}(\overline{A}_K, \overline{X}_K^*, \overline{C}_K^*, \overline{Z}_K) &= E(Y_{Kj} \mid \overline{A}_K, \overline{X}_K^*, \overline{C}_K^*, \overline{Z}_K); \\
Q_k^{j*}(\overline{A}_k, \overline{X}_k^*, \overline{C}_k^*, \overline{Z}_k) &= E\Big\{Y_{kj} + \max_{a_{k+1}} E(Y_{(k+1)j} \mid \overline{A}_k, \overline{X}_{k+1}^*, \overline{C}_{k+1}^*, \overline{Z}_{k+1}, a_{k+1}) \,\Big|\, \overline{A}_k, \overline{X}_k^*, \overline{C}_k^*, \overline{Z}_k\Big\} \\
&\qquad \text{for } k = K-1, \cdots, 1, \text{ and } j = 1, 2.
\end{aligned}
\tag{5.21}
$$

We then use the naive counterpart regression model of (5.3) to characterize $Q_k^{j*}(\overline{A}_k, \overline{X}_k^*, \overline{C}_k^*, \overline{Z}_k)$ as

$$
Q_k^{j*}(\overline{A}_k, \overline{X}_k^*, \overline{C}_k^*, \overline{Z}_k) = \beta_{kj}^{*\mathrm{T}} H_{k0}^* + (\psi_{kj}^{*\mathrm{T}} H_{k1}^*) A_k, \tag{5.22}
$$

where $\{H_{k0}^*, H_{k1}^*\}$ are counterparts of H_{k0} and H_{k1} with X_k and C_k replaced by X_k^* and C_k^*, respectively. Let $\theta_{k1}^* = (\beta_{k1}^{*\mathrm{T}}, \psi_{k1}^{*\mathrm{T}})^{\mathrm{T}}$ and $\theta_{k2}^* = (\beta_{k2}^{*\mathrm{T}}, \psi_{k2}^{*\mathrm{T}})^{\mathrm{T}}$ denote the regression coefficients for the naive Q-functions associated with the first and second outcomes, respectively. Then the naive composite Q-function is constructed as:

$$
\begin{aligned}
Q_k^*(\overline{A}_k, \overline{X}_k^*, \overline{C}_k^*, \overline{Z}_k, \delta) &= \delta Q_k^{1*}(\overline{A}_k, \overline{X}_k^*, \overline{C}_k^*, \overline{Z}_k) + (1-\delta) Q_k^{2*}(\overline{A}_k, \overline{X}_k^*, \overline{C}_k^*, \overline{Z}_k) \\
&\qquad \text{for } k = K, \cdots, 1
\end{aligned}
\tag{5.23}
$$

and subsequently yielding

$$
Q_k^*(\overline{A}_k, \overline{X}_k^*, \overline{C}_k^*, \overline{Z}_k, \delta) = \delta \left\{ {\beta_{k1}^*}^{\mathrm{T}} H_{k0}^* + ({\psi_{k1}^*}^{\mathrm{T}} H_{k1}^*) A_k \right\} + (1-\delta) \left\{ {\beta_{k2}^*}^{\mathrm{T}} H_{k0}^* + ({\psi_{k2}^*}^{\mathrm{T}} H_{k1}^*) A_k \right\}, \tag{5.24}
$$

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

Suppose we have measurements of a random sample with size n, $\mathcal{D}^* = \Big\{\{A_{ki}, X_{ki}^*, C_{ki}^*, Z_{ki}, Y_{k1i}, Y_{k2i}\} : k = 1, \cdots, K; i = 1, \cdots, n\Big\}$. Then the regression coefficients at each stage in (5.2) are naively estimated by repeating (5.7) and (5.8) with X_{ki} and C_{ki} replaced by X_{ki}^* and C_{ki}^*, i.e.,

$$
\hat{\theta}_{k1}^* = \arg\min_{\theta_{k1}^*} \frac{1}{n} \sum_{i=1}^{n} \left[\{\hat{Y}_{k1i}^* - Q_k^{1*}(\overline{A}_{ki}, \overline{X}_{ki}^*, \overline{C}_{ki}^*, \overline{Z}_{ki}; \theta_{k1}^*) \right]^2 \tag{5.25}
$$

and

$$\hat{\theta}^*_{k2} = \arg\min_{\theta_{k2}} \frac{1}{n} \sum_{i=1}^{n} \left[\{ \hat{Y}^*_{k2i} - Q^{2*}_k(\overline{A}_{ki}, \overline{X}^*_{ki}, \overline{C}^*_{ki}, \overline{Z}_{ki}; \theta^*_{k2}) \right]^2, \tag{5.26}$$

where for $k = K$, $\hat{Y}^*_{K1i} = Y_{K1i}$ and $\hat{Y}^*_{K2i} = Y_{K2i}$, and for $k = K-1, \cdots, 1$,

$$\hat{Y}^*_{k1i} = Y_{k1i} + Q^{1*}_{k+1}(\overline{A}_{ki}, \overline{X}^*_{(k+1)i}, \overline{C}^*_{(k+1)i}, \overline{Z}_{(k+1)i}, \hat{d}^*_{(k+1)i}; \hat{\theta}^*_{(k+1)1})$$

and

$$\hat{Y}^*_{k2i} = Y_{k2i} + Q^{2*}_{k+1}(\overline{A}_{ki}, \overline{X}^*_{(k+1)i}, \overline{C}^*_{(k+1)i}, \overline{Z}_{(k+1)i}, \hat{d}^*_{(k+1)i}; \hat{\theta}^*_{(k+1)2}),$$

Consequently, the naive optimal decision rules, denoted $\hat{d}^*_{(k+1)i}$, are determined by (5.11) with θ_{kj} replaced by $\hat{\theta}^*_{kj}$ for $j = 1, 2$ and $\{X_{ki}, C_{ki}\}$ replaced by $\{X^*_{ki}, C^*_{ki}\}$. That is,

$$\begin{aligned} \hat{d}^*_{k+1} = \arg\max_{a_{k+1}} \{ & \delta Q^{1*}_{k+1}(\overline{A}_k, \overline{X}^*_{k+1}, \overline{C}^*_{k+1}, \overline{Z}_{k+1}, a_{k+1}; \hat{\theta}^*_{(k+1)1}) \\ & + (1-\delta) Q^{2*}_{k+1}(\overline{A}_k, \overline{X}^*_{k+1}, \overline{C}^*_{k+1}, \overline{Z}_{k+1}, a_{k+1}; \hat{\theta}^*_{(k+1)2}) \} \quad \text{for} \quad k = K-1, \cdots, 0. \end{aligned} \tag{5.27}$$

5.3.3 Simulation Studies

In this subsection, we conduct a simulation study to investigate the performance of the naive Q-learning procedure described in Section 5.3.2 in the presence of mixed misclassification and measurement error. We consider a randomized treatment setting and set $K = 2$.

For $j = 1, 2$, let $\mu_{Y_{2j}} = E(Y_{2j} \mid X_2, A_1, A_2, C_2, Z_2)$ and let $\mu_{Y_{1j}} = E(Y_{1j} \mid X_1, A_1, C_1, Z_1)$. Then, for $j = 1, 2$ consider models

$$Y_{2j} = \mu_{Y_{2j}} + \epsilon_{2j} \text{ and } Y_{1j} = \mu_{Y_{1j}} + \epsilon_{1j},$$

where

$$\begin{aligned} \mu_{Y_{2j}} &= \eta_{0j} + \eta_{1j} Z_2 + \eta_{2j} X_2 + \eta_{3j} A_1 + \eta_{4j} A_2 + \eta_{5j} X_2 A_2 + \eta_{6j} C_2 A_2, \\ \mu_{Y_{1j}} &= \gamma_{0j} + \gamma_{1j} Z_1 + \gamma_{2j} X_1 + \gamma_{3j} A_1 + \gamma_{4j} X_1 A_1 + \gamma_{5j} C_1 A_1, \end{aligned}$$

and ϵ_{2j} and ϵ_{1j} are the error terms independently generated from $\mathcal{N}(0,1)$ for $j = 1, 2$.

Two binary treatments A_1 and A_2, taking values 0 and 1, are generated independently from the Bernoulli distribution, Bernoulli(0.5). Error-free covariates Z_1 and Z_2 are independently generated by $Z_1 \sim$ Bernoulli(0.5) and $Z_2 \sim$ Bernoulli(0.5). Error-prone binary covariate X_1 is independently generated by $X_1 \sim$ Bernoulli(0.5), and error-prone binary covariate X_2 is generated from the conditional distribution $X_2 \mid A_1 \sim \text{Bernoulli}\left(\frac{\exp(\nu A_1)}{1+\exp(\nu A_1)}\right)$, with ν set as 0.45. Error-prone continuous covariates are independently generated by $C_1 \sim \mathcal{N}(0,1)$ and $C_2 \sim \mathcal{N}(0,1)$.

Similar to the parameter settings considered in Section 4.2.3, we consider three settings for the model parameters in the Q-functions, including regular, weak non-regular, and non-regular settings. In the regular setting, the parameter values are set as $(\eta_{01}, \eta_{11}, \eta_{21}, \eta_{31}, \eta_{41}, \eta_{51}, \eta_{61})^{\mathrm{T}} = (1.5, 0.25, 0.8, -0.25, -2, 1.5, 1.75)^{\mathrm{T}}$, $(\eta_{02}, \eta_{12}, \eta_{22}, \eta_{32}, \eta_{42}, \eta_{52}, \eta_{62})^{\mathrm{T}} = (0.5, 0.75, 1.5, -0.15, -1.2, 0.95, 1.1)^{\mathrm{T}}$, $(\gamma_{01}, \gamma_{11}, \gamma_{21}, \gamma_{31}, \gamma_{41}, \gamma_{51})^{\mathrm{T}} = (0.5, 0.15, 0.5, -1.5, 1.25, 0.95)^{\mathrm{T}}$, and $(\gamma_{02}, \gamma_{12}, \gamma_{22}, \gamma_{32}, \gamma_{42}, \gamma_{52})^{\mathrm{T}} = (0.75, 0.2, 0.85, -1.85, 0.85, 1.45)^{\mathrm{T}}$. In the weak non-regular setting, we take $(\eta_{01}, \eta_{11}, \eta_{21}, \eta_{31}, \eta_{41}, \eta_{51}, \eta_{61})^{\mathrm{T}} = (1.5, 0.25, 0.8, -0.25, -2, 2.02, 0)^{\mathrm{T}}$, $(\eta_{02}, \eta_{12}, \eta_{22}, \eta_{32}, \eta_{42}, \eta_{52}, \eta_{62})^{\mathrm{T}} = (0.5, 0.75, 1.5, -0.15, -1.2, 1.22, 0)^{\mathrm{T}}$, $(\gamma_{01}, \gamma_{11}, \gamma_{21}, \gamma_{31}, \gamma_{41}, \gamma_{51})^{\mathrm{T}} = (0.5, 0.15, 0.5, -1.5, 1.25, 0.95)^{\mathrm{T}}$, and $(\gamma_{02}, \gamma_{12}, \gamma_{22}, \gamma_{32}, \gamma_{42}, \gamma_{52})^{\mathrm{T}} = (0.75, 0.2, 0.85, -1.85, 0.85, 1.45)^{\mathrm{T}}$. In the non-regular setting, the parameter values are specified as $(\eta_{01}, \eta_{11}, \eta_{21}, \eta_{31}, \eta_{41}, \eta_{51}, \eta_{61})^{\mathrm{T}} = (1.5, 0.25, 0.8, -0.25, -0.2, 0.2, 0)^{\mathrm{T}}$, $(\eta_{02}, \eta_{12}, \eta_{22}, \eta_{32}, \eta_{42}, \eta_{52}, \eta_{62})^{\mathrm{T}} = (0.5, 0.75, 1.5, -0.15, -0.1, 0.1, 0)^{\mathrm{T}}$, $(\gamma_{01}, \gamma_{11}, \gamma_{21}, \gamma_{31}, \gamma_{41}, \gamma_{51})^{\mathrm{T}} = (0.5, 0.15, 0.5, -1.5, 1.25, 0.95)^{\mathrm{T}}$, and $(\gamma_{02}, \gamma_{12}, \gamma_{22}, \gamma_{32}, \gamma_{42}, \gamma_{52})^{\mathrm{T}} = (0.75, 0.2, 0.85, -1.85, 0.85, 1.45)^{\mathrm{T}}$.

The Q-functions for the two stages are specified as

$$Q_2^j(A_1, X_2, C_2, Z_2, A_2) = \beta_{02j} + \beta_{12j} Z_2 + \beta_{22j} X_2 + \beta_{32j} A_1 + (\psi_{02j} + \psi_{12j} X_2 + \psi_{22j} C_2) A_2, \quad (5.28)$$

and

$$Q_1^j(X_1, C_1, Z_1, A_1) = \beta_{01j} + \beta_{11j} Z_1 + \beta_{21j} X_1 + (\psi_{01j} + \psi_{11j} X_1 + \psi_{21j} C_1) A_1, \quad (5.29)$$

where for $j = 1, 2$, β_{k2j} (with $k = 0, 1, 2, 3$), β_{k1j} (with $k = 0, 1, 2$), ψ_{k2j} (with $k = 0, 1, 2$), and ψ_{k1j} (with $k = 0, 1, 2$) are regression parameters. Based on the coefficients of A_2 and A_1, the

optimal DTR is given by the decision rules:

$$d_2 = \text{sign}\Big\{\delta(\psi_{021} + \psi_{121}X_2 + \psi_{221}C_2) + (1-\delta)(\psi_{022} + \psi_{122}X_2 + \psi_{222}C_2)\Big\};$$
$$d_1 = \text{sign}\Big\{\delta(\psi_{011} + \psi_{111}X_1 + \psi_{211}C_1) + (1-\delta)(\psi_{012} + \psi_{112}X_1 + \psi_{212}C_1)\Big\},$$

where $\text{sign}(t) = 1$ if $t > 0$, and 0 otherwise.

With the misclassification matrix (5.18) and classical additive model (5.19), we generate surrogate values X_k^* of X_k and C_k^* of C_k with $k = 1, 2$. We consider three different settings for misclassification probabilities and measurement error degree, with $(\pi_{10}, \pi_{01}, \sigma_k^2)^{\mathrm{T}} = (0.1, 0.1, 1)^{\mathrm{T}}, (0.2, 0.2, 1.5)^{\mathrm{T}},$ or $(0.3, 0.3, 2)^{\mathrm{T}}$.

To run simulations, we use the proceeding models to generate data of size $n = 1000$, and we repeat 1000 simulations for each parameter configuration. We implement the Q-learning algorithm in Section 5.3.2 to the observed data $\{Z_1, X_1^*, C_1^*, A_1, Y_{11}, Y_{12}, Z_2, X_2^*, C_2^*, A_2, Y_{21}, Y_{22}\}$, called the "naive method", as opposed to the Q-learning procedure in Section 5.2.1 to the true data $\{Z_1, X_1, C_1, A_1, Y_{11}, Y_{12}, Z_2, X_2, C_2, A_2, Y_{21}, Y_{22}\}$, called the "error-free least squares" (EFLS) method.

In Table 5.1, we report the numerical results for stages 1 and 2 over 1000 simulations for the regular, weak non-regular, and non-regular settings, where "Bias", "SE", "ESE", "MSE", "WTCR", "PBCR", and "DBCR" are defined in the same way as those defined in Section 4.2.3. Same as in Section 4.2.3, we use 1000 bootstrap iterations to calculate the PB CIs and the DB CIs which are based on 1000 first-stage and 100 second-stage bootstrap iterations. We assign the weight parameter δ as 0.9. The results are presented in Table 5.1, which demonstrate the favorable performance of the EFLS method in the regular setting, characterized by minimal biases and MSEs, as well as satisfactory CRs of 95% CIs for both stages 1 and 2. However, for the weak non-regular setting, the stage 1 parameter estimates derived from the EFLS method may exhibit certain levels of bias, particularly notable in the case of $\psi_{0k\delta}$ and $\psi_{1k\delta}$. Noticeably, the CRs of 95% PB CIs indicate satisfactory performance overall, except for $\psi_{0k\delta}$ in the non-regular setting. We also examine W-type CIs for weak non-regular and non-regular settings in stage 1 (results are not reported here) and observe poor coverage.

The results obtained from the naive method for both stages 1 and 2 reveal suboptimal per-

formance characterized by significant biases, high MSEs, and unacceptably low CRs. The extent of biases due to covariate mismeasurement intensifies as the degrees of mismeasurement increase. These findings underscore the significance of addressing the impact of mismeasurement and introducing suitable correction methodologies to enhance the accuracy of estimations.

5.4 Mismeasurement Correction: Regression Calibration

Here we describe the application of the RC method for the case where an additional set of measurements, known as validation data, is available in addition to the main study data with surrogate measurements together with measurements for other variables. The validation data can be used to estimate the magnitude of the mismeasurements associated with the covariates X_j and C_j. To employ the RC approach, we construct approximate measures for these error-prone covariates, which are subsequently used in the Q-learning method as outlined in Section 5.2. This involves replacing the error-prone covariates X_j and C_j with their respective approximated measures, enabling a more accurate estimation process.

Assume internal validation data $\mathcal{D}_V = \left\{\{A_{ki}, X_{ki}, X^*_{ki}, C_{ki}, C^*_{ki}, Z_{ki}, Y_{k1i}, Y_{k2i}\} : k = 1, \cdots, K; i \in \mathcal{V}\right\}$ are available, where $\mathcal{V}$ is a subset of $\mathcal{M} \triangleq \{1, \cdots, n\}$. With the observed data $\mathcal{D}^*$, let

$$X^{**}_{ki} = E(X_{ki} \mid \overline{A}_{(k-1)i}, \overline{X}^*_{ki}, \overline{C}^*_{ki}, \overline{Z}_{ki}) \text{ and } C^{**}_{ki} = E(C_{ki} \mid \overline{A}_{(k-1)i}, \overline{X}^*_{ki}, \overline{C}^*_{ki}, \overline{Z}_{ki})$$

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

Following the same idea given in Section 4.3, we make use of regression modeling techniques to determine X^{**}_{ki} and C^{**}_{ki} for $i = 1, \cdots, n$. Regarding the determination of X^{**}_{ki}, let $\pi^{(k)}_i = P(X_{ki} = 1 \mid \overline{A}_{(k-1)i}, \overline{X}^*_{ki}, \overline{C}^*_{ki}, \overline{Z}_{ki})$ for $k = 1, \cdots, K$ to express X^{**}_{ki}, because X_{ki} is the binary covariate taking value 0 or 1. Then, we consider a logistic regression model to delineate $\pi^{(k)}_i$:

$$\text{logit}\, \pi^{(k)}_i = m_{X_k}(\overline{A}_{(k-1)i}, \overline{X}^*_{ki}, \overline{C}^*_{ki}, \overline{Z}_{ki}; \boldsymbol{\zeta}_{X_k}), \tag{5.30}$$

where $m_{X_k}(\cdot)$ is a specified function, and $\boldsymbol{\zeta}_{X_k}$ is the associated parameter.

To delineate C_{ki}^{**}, we employ a linear regression model

$$C_{ki}^{**} = m_{C_k}(\overline{A}_{(k-1)i}, \overline{X}_{ki}^{*}, \overline{C}_{ki}^{*}, \overline{Z}_{ki}; \boldsymbol{\zeta}_{C_k}), \tag{5.31}$$

where $m_{C_k}(\cdot)$ is a specified function, and $\boldsymbol{\zeta}_{C_k}$ is the associated parameter.

Next, regression models (5.30) and (5.31) are fitted using the validation data $\mathcal{D}_V$, resulting in the estimation of $\boldsymbol{\zeta}_{X_k}$ and $\boldsymbol{\zeta}_{C_k}$, denoted as $\hat{\boldsymbol{\zeta}}_{X_k}$ and $\hat{\boldsymbol{\zeta}}_{C_k}$, respectively. These estimates are then used to determine estimates of X_{ki}^{**} and C_{ki}^{**}. The implementation of the calibrated Q-learning algorithm can be modified in stages K to 1 with different treatments of X_{ki} and C_{ki}, where measurements for X_{ki} and C_{ki} are used for $i \in \mathcal{V}$, and $\hat{X}_{ki}^{**}$ and $\hat{C}_{ki}^{**}$ are used to replace X_{ki} and C_{ki} for $i \in \mathcal{M} \setminus \mathcal{V}$.

5.5 Mismeasurement Correction: Estimating Equation Approach

While regression calibration provides a simple approach to addressing mismeasurement in covariates, this method does not always ensure the consistency for the model parameter estimation. In this section, we develop an alternative approach by employing estimating function theory, and present criteria for developing unbiased estimating functions. We first describe the unbiased estimating function approach for stage K, and then extend it to other stages to create working estimating functions along the lines in Section 4.4.

5.5.1 Corrected Estimation Functions with Known Misclassification Probabilities and Measurement Error Degree

Assume that the misclassification probabilities in (5.18) and the variance of e_{ki} in (5.19) are known. In what follows, we first describe how to correct for mismeasurement-induced bias in stage K, and then we discuss estimation pertinent to stage k for $k = K-1, \cdots, 1$.

Estimation Related to Stage K

If the true covariates X_{ki} and C_{ki} for $k = 1, \cdots, K$ and $i = 1, \cdots, n$ are not available, but surrogate values X^*_{ki} and C^*_{ki} are instead collected, then directly using (5.16) with X_{ki} and C_{ki} replaced by X^*_{ki} and C^*_{ki} may result in inconsistent estimators.

Similar to the idea discussed in Section 4.4.1, our goal here is to construct an unbiased estimating function, say $S^*_{Kj}(\theta_{Kj}; Y_{Kji}, \overline{A}_{Ki}, \overline{X}^*_{Ki}, \overline{C}^*_{Ki}, \overline{Z}_{Ki})$, such that its conditional expectation recovers the unbiased estimating function, i.e., $S_{Kj}(\theta_{Kj}; Y_{Kji}, \overline{A}_{Ki}, \overline{X}_{Ki}, \overline{C}_{Ki}, \overline{Z}_{Ki})$ constructed from using the true covariates together with $\{Y_{Kji}, \overline{A}_{Ki}, \overline{Z}_{Ki}\}$,

$$E\{S^*_{Kj}(\theta_{Kj}; Y_{Kji}, \overline{A}_{Ki}, \overline{X}^*_{Ki}, \overline{C}^*_{Ki}, \overline{Z}_{Ki}) \mid Y_{Kji}, \overline{A}_{Ki}, \overline{X}_{Ki}, \overline{C}_{Ki}, \overline{Z}_{Ki}\} = S_{Kj}(\theta_{Kj}; Y_{Kji}, \overline{A}_{Ki}, \overline{X}_{Ki}, \overline{C}_{Ki}, \overline{Z}_{Ki}). \tag{5.32}$$

For ease of referral, we call $S_{Kj}(\theta_{Kj}; Y_{Kji}, \overline{A}_{Ki}, \overline{X}_{Ki}, \overline{C}_{Ki}, \overline{Z}_{Ki})$ a "true" estimating function, and $S^*_{Kj}(\theta_{Kj}; Y_{Kji}, \overline{A}_{Ki}, \overline{X}^*_{Ki}, \overline{C}^*_{Ki}, \overline{Z}_{Ki})$ a "corrected" estimating function. With (5.32), it is immediate that $S^*_{Kj}(\theta_{Kj}; Y_{Kji}, \overline{A}_{Ki}, \overline{X}^*_{Ki}, \overline{C}^*_{Ki}, \overline{Z}_{Ki})$ is an unbiased estimating function due to that

$$E\{S_{Kj}(\theta_{Kj}; Y_{Kji}, \overline{A}_{Ki}, \overline{X}_{Ki}, \overline{C}_{Ki}, \overline{Z}_{Ki})\} = 0.$$

Then estimating function theory shows that solving

$$\sum_{i=1}^{n} S^*_{Kj}(\theta_{Kj}; Y_{Kji}, \overline{A}_{Ki}, \overline{X}^*_{Ki}, \overline{C}^*_{Ki}, \overline{Z}_{Ki}) = 0$$

for θ_{Kj} yields consistent estimator for θ_{Kj}, provided regularity conditions (Yi, 2017, Section 2.5).

Since the dependence of (5.14) and (5.15) on X_k and C_k is reflected respectively by $\{X_k, X_k X_k^{\mathrm{T}}\}$ and $\{C_k, C_k^2\}$ for $k = 1, \cdots, K$, to find $S^*_{Kj}(\theta_{Kj}; Y_{Kji}, \overline{A}_{Ki}, \overline{X}^*_{Ki}, \overline{C}^*_{Ki}, \overline{Z}_{Ki})$ to meet (5.32), it suffices to find unbiased surrogates for $\{X_k, X_k X_k^{\mathrm{T}}\}$ and $\{C_k, C_k^2\}$ in the following sense. To be specific, we aim to find functions of X^*_k, say X^{**}_k and X^{***}_k, and functions of C^*_k, say C^{**}_k and C^{***}_k, such that

$$E\left\{X^{**}_k \mid Y_{k1i}, Y_{k2i}, \overline{A}_{ki}, \overline{X}_{ki}, \overline{C}_{ki}, \overline{Z}_{ki}\right\} = X_k; \; E\left\{X^{***}_k \mid Y_{k1i}, Y_{k2i}, \overline{A}_{ki}, \overline{X}_{ki}, \overline{C}_{ki}, \overline{Z}_{ki}\right\} = X_k X_k^{\mathrm{T}}; \tag{5.33}$$

and

$$E\left\{C_k^{**} \mid Y_{k1i}, Y_{k2i}, \overline{A}_{ki}, \overline{X}_{ki}, \overline{C}_{ki}, \overline{Z}_{ki}\right\} = C_k;\ E\left\{C_k^{***} \mid Y_{k1i}, Y_{k2i}, \overline{A}_{ki}, \overline{X}_{ki}, \overline{C}_{ki}, \overline{Z}_{ki}\right\} = C_k^2. \quad (5.34)$$

To construct X_k^{**} and X_k^{***}, we apply the technique of Akazawa et al. (1998). For $t = 1, 2$, let e_t denote a 2×1 vector with 1 in the t position and zero elsewhere. We now express the two values, 0 and 1, for the binary variable X_k (or X_k^*) as two 2×1 vectors. If $X_k = 1$, then a 2×1 vector with the first element set as 0 and the second element set as 1; if $X_k = 0$, then we create a 2×1 vector with the first element set to be 1 and the second element set to be 0. That is, $X_k = 1$ and $X_k = 0$ can be represented by e_2 and e_1, respectively. Similarly, $X_k^* = 1$ and $X_k^* = 0$ are represented by e_2 and e_1, respectively.

Theorem 5.5.1 *Define*

$$X_k^{**} = \Pi^{-1} X_k^* \text{ and } X_k^{***} = \sum_{t=1}^{2} \left\{X_k^{**\mathrm{T}} e_t\right\} e_t e_t^{\mathrm{T}}.$$

Then (5.33) holds.

Proof First, for $l = 0, 1$ we have that

$$E\left\{X_{ki}^* \mid X_{ki} = l\right\} = \begin{pmatrix} \pi_{0l} \\ \pi_{1l} \end{pmatrix} \text{ and } \Pi^{-1} \begin{pmatrix} \pi_{0l} \\ \pi_{1l} \end{pmatrix} = e_{l+1}.$$

Consequently,

$$\begin{aligned} E\left\{X_{ki}^{**} \mid Y_{k1i}, Y_{k2i}, \overline{A}_{ki}, \overline{X}_{ki}, \overline{C}_{ki}, \overline{Z}_{ki}\right\} &= E\left\{\Pi^{-1} X_{ki}^* \mid Y_{k1i}, Y_{k2i}, \overline{A}_{ki}, \overline{X}_{(k-1)i}, X_{ki} = l, \overline{C}_{ki}, \overline{Z}_{ki}\right\} \\ &= \Pi^{-1} E\left\{X_{ki}^* \mid Y_{k1i}, Y_{k2i}, \overline{A}_{ki}, \overline{X}_{(k-1)i}, X_{ki} = l, \overline{C}_{ki}, \overline{Z}_{ki}\right\} \\ &= \Pi^{-1} \begin{pmatrix} \pi_{0l} \\ \pi_{1l} \end{pmatrix} = e_{l+1} = X_{ki}, \end{aligned} \quad (5.35)$$

where the associated assumptions are made.

Similarly, using (5.35), we have that

$$\begin{aligned}
&E\{X_{ki}^{***} \mid Y_{k1i}, Y_{k2i}, \overline{A}_{ki}, \overline{X}_{ki}, \overline{C}_{ki}, \overline{Z}_{ki}\} \\
=&E\left\{ \sum_{t=1}^{2} \left\{X_{ki}^{**\mathrm{T}} e_t\right\} e_t e_t^{\mathrm{T}} \;\middle|\; Y_{k1i}, Y_{k2i}, \overline{A}_{ki}, \overline{X}_{(k-1)i}, X_{ki} = l, \overline{C}_{ki}, \overline{Z}_{ki} \right\} \\
=&\left\{ \sum_{t=1}^{2} \left\{X_{ki}^{\mathrm{T}} e_t\right\} e_t e_t^{\mathrm{T}} \right\} \\
=&X_{ki} X_{ki}^{\mathrm{T}}.
\end{aligned}$$

□

The construction of C_k^{**} and C_k^{***} is straightforward by the classical additive error model (5.19). Setting $C_k^{**} = C_k^{*}$ and $C_k^{***} = C_k^{*2} - \sigma_k^2$ makes (5.34) hold.

Consequently, we define $S^*_{Kj}(\theta_{Kj}; Y_{Kji}, \overline{A}_{Ki}, \overline{X}^*_{Ki}, \overline{C}^*_{Ki}, \overline{Z}_{Ki})$ to be $S_{Kj}(\theta_{Kj}; Y_{Kji}, \overline{A}_{Ki}, \overline{X}_{Ki}, \overline{C}_{Ki}, \overline{Z}_{Ki})$ with $\{X_{ki}, X_{ki}X_{ki}^{\mathrm{T}}\}$ and $\{C_{ki}, C_{ki}^2\}$ replaced by their unbiased surrogates $\{X_{ki}^{**}, X_{ki}^{***}\}$ and $\{C_{ki}^{**}, C_{ki}^{***}\}$, respectively. Let $\hat{\theta}_{Kjc} = (\hat{\beta}_{Kjc}^{\mathrm{T}}, \hat{\psi}_{Kjc}^{\mathrm{T}})^{\mathrm{T}}$ denote the resultant estimator of θ_{Kj} by solving

$$\sum_{i=1}^{n} S^*_{Kj}(\theta_{Kj}; Y_{Kji}, \overline{A}_{Ki}, \overline{X}^*_{Ki}, \overline{C}^*_{Ki}, \overline{Z}_{Ki}) = 0 \tag{5.36}$$

for θ_{Kj}. Under regularity conditions, $\sqrt{n}(\hat{\theta}_{Kjc} - \theta_{Kj}) \xrightarrow{d} \mathcal{N}(0, \Sigma(\theta_{Kj}))$ as $n \to \infty$, where $\Sigma(\theta_{Kj}) = \left\{I(\theta_{Kj})\right\}^{-1} J(\theta_{Kj}) \left\{I(\theta_{Kj})\right\}^{-1\mathrm{T}}$, with

$$I(\theta_{Kj}) = E\left\{ \frac{\partial S^*_{Kj}(\theta_{Kj}; Y_{Kji}, \overline{A}_{Ki}, \overline{X}^*_{Ki}, \overline{C}^*_{Ki}, \overline{Z}_{Ki})}{\partial \theta_{Kj}} \right\}$$

and

$$J(\theta_{Kj}) = E\left[S^*_{Kj}(\theta_{Kj}; Y_{Kji}, \overline{A}_{Ki}, \overline{X}^*_{Ki}, \overline{C}^*_{Ki}, \overline{Z}_{Ki}) \{S^*_{Kj}(\theta_{Kj}; Y_{Kji}, \overline{A}_{Ki}, \overline{X}^*_{Ki}, \overline{C}^*_{Ki}, \overline{Z}_{Ki})\}^{\mathrm{T}} \right]$$

(Yi, 2017, Section 1.3).

Estimation Related to Stage $k < K$

In the error-free setting, estimation in stage k can be carried out by solving (5.17). However, in the case where accurate measurements of X_{ki} and C_{ki} are not available, (5.17) may produce unreliable results if directly replacing X_{ki} and C_{ki} with their observed surrogate measurements X^*_{ki} and C^*_{ki}. Similar to the consideration for stage K, we modify $\hat{S}_{kj}(\theta_{kj}; \hat{Y}_{kji}, \overline{A}_{ki}, \overline{X}_{ki}, \overline{C}_{ki}, \overline{Z}_{ki})$ in (5.17) by replacing $\{X_{ki}, X_{ki}X_{ki}^{\mathrm{T}}\}$ and $\{C_{ki}, C_{ki}^2\}$ with their respective unbiased surrogates, $\{X_k^{**}, X_k^{***}\}$ and $\{C_k^{**}, C_k^{***}\}$. Additionally, the pseudo-outcome $\hat{Y}_{kji}$ depends on X_{ki} and C_{ki}, and therefore, it needs to be modified accordingly by replacing X_{ki} and C_{ki} with their unbiased surrogates. Let $\hat{S}^*_{kj}(\theta_{kj}; \hat{Y}^*_{kji}, \overline{A}_{ki}, \overline{X}^*_{ki}, \overline{C}^*_{ki}, \overline{Z}_{ki})$ denote the modified estimating function. Then, for $k = K-1, \cdots, 1$, we solve

$$\sum_{i=1}^{n} \hat{S}^*_{kj}(\theta_{kj}; \hat{Y}^*_{kji}, \overline{A}_{ki}, \overline{X}^*_{ki}, \overline{C}^*_{ki}, \overline{Z}_{ki}) = 0 \tag{5.37}$$

for θ_{kj}, and let $\hat{\theta}_{kjc}$ denote the resulting estimator of θ_{kj}.

5.5.2 Corrected Estimation Functions with Unknown Misclassification Probabilities and Measurement Error Degree

The implementation of the correction procedure described in Section 5.5.1 is contingent upon the availability of prior knowledge regarding the misclassification and measurement error mechanisms. However, in many applications, such prior information is often unavailable, necessitating the estimation of misclassification probabilities and measurement error degree. To this end, using a validation subsample can be instrumental in addressing this challenge. In this section, we propose an adaptation of the estimation equation method proposed in Section 5.5.1, which accounts for the unknown mismeasurement degrees. Our analytical procedure is based on two sets of data, namely, the main study sample $\mathcal{D}^*$ and the internal validation subsample $\mathcal{D}_V$. The validation sample, which comprises a smaller group of m individuals, is drawn from the main study sample, where $m \leq n$.

Estimation of Misclassification Probabilities and Measurement Error Degree

First, we present the process for estimating misclassification probabilities, which is analogous to the procedure discussed in Section 4.4.2, albeit with minor notational modifications.

We start by defining the misclassification probabilities for the error-prone binary covariate X_{ki} as follows: for $i \in \mathcal{M}$ and $k = 1, \cdots, K$,

$$\pi_{ki01} = P(X_{ki}^* = 0 \mid X_{ki} = 1, \overline{A}_{(k-1)i} \cup \overline{X}_{ki} \cup \overline{C}_{ki} \cup \overline{Z}_{ki} \setminus X_{ki})$$

and

$$\pi_{ki10} = P(X_{ki}^* = 1 \mid X_{ki} = 0, \overline{A}_{(k-1)i} \cup \overline{X}_{ki} \cup \overline{C}_{ki} \cup \overline{Z}_{ki} \setminus X_{ki}).$$

To describe how misclassification probabilities are associated with covariates, we employ logistic regression models

$$\begin{aligned} \operatorname{logit} \pi_{ki10} &= \alpha_{k0}^{\mathrm{T}} \mathcal{W}_{ki0}; \\ \operatorname{logit} \pi_{ki01} &= \alpha_{k1}^{\mathrm{T}} \mathcal{W}_{ki1}, \end{aligned} \tag{5.38}$$

where α_{kl} denotes the vector of regression coefficients and $\mathcal{W}_{kil}$ may include 1 and a subset of covariates $\{X_{ki} = l\} \cup \overline{A}_{(k-1)i} \cup \overline{X}_{ki} \cup \overline{C}_{ki} \cup \overline{Z}_{ki} \setminus X_{ki}$ that reflects different misclassification mechanisms for $l = 0, 1$. Having 1 in $\mathcal{W}_{kil}$ allows the inclusion of the intercept in (5.38), and $\mathcal{W}_{kil}$ may contain the entire covariate vector $\{X_{ki} = l\} \cup \overline{A}_{(k-1)i} \cup \overline{X}_{ki} \cup \overline{C}_{ki} \cup \overline{Z}_{ki} \setminus X_{ki}$ or just constant 1 alone, where the latter case corresponds to homogeneous misclassification across all subjects. Let $\alpha_k = (\alpha_{k0}^{\mathrm{T}}, \alpha_{k1}^{\mathrm{T}})^{\mathrm{T}}$ denote the parameter vector for $k = 1, \cdots, K$.

For $i = 1, \cdots, n$ and $k = 1, \cdots, K$, let

$$L_{ki}(\alpha_k) = P(X_{ki}^* = x_{ki}^* \mid X_{ki} = x_{ki}, \overline{A}_{(k-1)i} \cup \overline{X}_{ki} \cup \overline{C}_{ki} \cup \overline{Z}_{ki} \setminus X_{ki})$$

which equals $\{\pi_{ki10}^{x_{ki}^*}(1 - \pi_{ki10})^{1-x_{ki}^*}\}^{1-x_{ki}} \cdot \{\pi_{ki01}^{1-x_{ki}^*}(1 - \pi_{ki01})^{x_{ki}^*}\}^{x_{ki}}$ for $x_{ki}, x_{ki}^* = 0, 1$. Write $\alpha = (\alpha_1^{\mathrm{T}}, \cdots, \alpha_K^{\mathrm{T}})^{\mathrm{T}}$. Let $S_{ki}(\alpha_k) = \partial \log L_{ki}(\alpha_k)/\partial \alpha_k$ and let

$$S_i(\alpha) = \left(S_{1i}^{\mathrm{T}}(\alpha_1), \cdots, S_{Ki}^{\mathrm{T}}(\alpha_K)\right)^{\mathrm{T}}.$$

With internal validation data, solving

$$\sum_{i\in\mathcal{V}} S_i(\alpha) = 0$$

for α yields the maximum likelihood estimate, denoted $\hat{\alpha} = (\hat{\alpha}_1^{\mathrm{T}}, \cdots, \hat{\alpha}_K^{\mathrm{T}})^{\mathrm{T}}$, of α.

Next, we present the procedure for estimating the measurement error degree for the error-prone covariate C_{ki}. Consider the classical additive model (5.19). With a validation sample available, it is possible to estimate the parameters associated with (5.19), regardless of the distribution form of e_{ki}. Here for illustrations, we assume that e_{ki} follows a normal distribution with mean 0 and unknown variance σ_k^2. Our goal is to estimate the unknown parameter σ_k^2 so as to construct the unbiased surrogate C_k^{***}.

For $i \in \mathcal{V}$ and $k = 1, \cdots, K$, let

$$L_{ki}\left(\sigma_k^2\right) = \left(2\pi\sigma_k^2\right)^{-1/2} \exp\left(-\frac{\left(c_{ki}^* - c_{ki}\right)^2}{2\sigma_k^2}\right) \tag{5.39}$$

denote the probability density function of $C_{ki}^* - C_{ki}$ and let $\ell_{ki}(\sigma_k^2) = \partial \log L_{ki}(\sigma_k^2; e_{ki})/\partial\sigma_k^2$. Let $\sigma^2 = \left(\sigma_1^2, \cdots, \sigma_K^2\right)^{\mathrm{T}}$ and define $\ell_i(\sigma^2) = \left(L_{1i}(\sigma_1^2), \cdots, L_{Ki}(\sigma_K^2)\right)^{\mathrm{T}}$.

We then estimate σ^2 by solving

$$\sum_{i\in\mathcal{V}} \ell_i(\sigma^2) = 0, \tag{5.40}$$

and let $\hat{\sigma}^2 = \left(\hat{\sigma}_1^2, \cdots, \hat{\sigma}_K^2\right)^{\mathrm{T}}$ denote the resultant estimator of σ^2.

Estimation for the Parameters of Q-Functions

For notational simplicity, let $S_{Kji}^*(\theta_{Kj}, \alpha, \sigma^2)$ represent $S_{Kj}^*\left(\theta_{Kj}, \alpha, \sigma^2; Y_{Kji}, \bar{A}_{Ki}, \bar{X}_{Ki}^*, \bar{C}_{Ki}^*, \bar{Z}_{Ki}\right)$ in (5.36) with the dependence on α and σ^2 spelled out, and let $S_{Kji}(\theta_{Kj})$ represent $S_{Kj}(\theta_{Kj}; Y_{Kji}, \bar{A}_{Ki}, \bar{X}_{Ki}, \bar{C}_{Ki}, \bar{Z}_{Ki})$ in (5.16). For stage K, we then estimate θ_{Kj} by solving

$$\sum_{i\in\mathcal{M}\setminus\mathcal{V}} S_{Kji}^*(\theta_{Kj}, \hat{\alpha}, \hat{\sigma}^2) + \sum_{i\in\mathcal{V}} S_{Kji}(\theta_{Kj}) = 0$$

for θ_{Kj}, and let $\hat{\theta}_{Kj}$ denote the resultant estimator of θ_{Kj}.

For $j = 1, 2$, let $\vartheta_j = (\beta_{Kj}^{\mathrm{T}}, \psi_{Kj}^{\mathrm{T}}, \alpha^{\mathrm{T}}, \sigma^{2\mathrm{T}})^{\mathrm{T}}$ and let $\hat{\vartheta}_j = (\hat{\theta}_{Kj}^{\mathrm{T}}, \hat{\alpha}^{\mathrm{T}}, \hat{\sigma}^{2\mathrm{T}})^{\mathrm{T}}$. Under regularity conditions and that the ratio m/n approaches a positive constant, say ρ, as $n \to \infty$, $\hat{\vartheta}_j$ is a consistent estimator of ϑ_j, and

$$\sqrt{n}(\hat{\vartheta}_j - \vartheta_j) \xrightarrow{d} \mathcal{N}(0, \Sigma_{Vj}) \qquad \text{as} \qquad n \to \infty$$

where $\Sigma_{Vj} = A_{Vj}^{-1} B_{Vj} A_{Vj}^{-1\mathrm{T}}$, with

$$A_{Vj} = -(1-\rho)\begin{pmatrix} E(\frac{\partial S^*_{Kji}(\theta_{Kj},\alpha,\sigma^2)}{\partial \theta_{Kj}}) & E(\frac{\partial S^*_{Kji}(\theta_{Kj},\alpha,\sigma^2)}{\partial \alpha} & E(\frac{\partial S^*_{Kji}(\theta_{Kj},\alpha,\sigma^2)}{\partial \sigma^2}) \\ 0 & 0 & 0 \end{pmatrix} - \rho \begin{pmatrix} E(\frac{\partial S_{Kji}(\theta_{Kj})}{\partial \theta_{Kj}}) & 0 & 0 \\ 0 & E(\frac{S_i(\alpha)}{\partial \alpha}) & 0 \\ 0 & 0 & E(\frac{\ell_i(\sigma^2)}{\partial \sigma^2}) \end{pmatrix}, \quad (5.41)$$

and

$$B_V = (1-\rho)\begin{pmatrix} E\{S^*_{Kji}(\theta_{Kj},\alpha,\sigma^2) S^{*\mathrm{T}}_{Kji}(\theta_{Kj},\alpha,\sigma^2)\} & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} + \rho \begin{pmatrix} E\{S_{Kji}(\theta_{Kj}) S^{\mathrm{T}}_{Kji}(\theta_{Kj})\} & 0 & 0 \\ 0 & E\{S_i(\alpha) S_i^{\mathrm{T}}(\alpha)\} & 0 \\ 0 & 0 & E\{\ell_i(\sigma^2)\ell_i^{\mathrm{T}}(\sigma^2)\} \end{pmatrix}. \quad (5.42)$$

Finally, for $k = K-1, \cdots, 1$, estimator of θ_{kj} can be obtained by solving (5.37), where π_{k10} and π_{k01} are determined by (5.38) with α replaced by $\hat{\alpha}$, and σ_k^2 is replaced with its estimate determined by (5.40).

5.6 Simulation Study

In this section, we conduct simulation studies to evaluate the finite sample performance of the methods described in Sections 5.4 and 5.5. To accomplish this, we employ the same configura-

tion and parameter settings as in Section 5.3.3 to generate the main study data $\Big\{\{A_{ki}, X^*_{ki}, C^*_{ki}, Z_{ki}, Y_{k1i}, Y_{k2i}\} : k = 1, \cdots, K; i \in \mathcal{M}\Big\}$. Additionally, we create an internal validation subsample by randomly selecting 30% of the study subjects from $\mathcal{M}$ and record their precise measurements of $\{X_{ki}, C_{ki} : k = 1, \cdots, K; i \in \mathcal{V}\}$ to form the validation subsample $\mathcal{D}_V = \Big\{\{A_{ki}, X_{ki}, X^*_{ki}, C_{ki}, C^*_{ki}, Z_{ki}, Y_{k1i}, Y_{k2i}\} : k = 1, \cdots, K;\ i \in \mathcal{V}\Big\}$.

We examine the data using three methods. The first method, referred to as "RC", applies the approach outlined in Section 5.4 to both $\mathcal{D}^*$ and $\mathcal{D}_V$. The second method, termed as "EE-known", applies the method described in Section 5.5.1 to the data in $\mathcal{D}^*$, assuming that the misclassification probabilities and measurement error degrees are known. Lastly, the third method, named "EE-estimated", employs the procedure described in Section 5.5.2 to both $\mathcal{D}^*$ and $\mathcal{D}_V$, where the misclassification probabilities and measurement error degrees are estimated using $\mathcal{D}_V$.

Tables 5.2-5.4 present the numerical results pertaining to stages 1 and 2 for the proposed methods across regular, weak non-regular, and non-regular settings. The definitions of "Bias", "SE", "ESE", "MSE", "WTCR", "PBCR", and "DBCR" correspond to those elucidated in Section 4.5. It is evident that both the RC and EE methods yield satisfactory results in terms of bias, with the RC method exhibiting relatively lower levels of bias compared to the EE methods. Moreover, it is observed that the EE methods yield high SEs, consequently leading to higher MSEs, which tend to escalate as the degrees of mismeasurement increase. Unsurprisingly, the performance of the correction methods deteriorates as mismeasurement degrees increase. Regardless of the regularity condition of the parameters, the PB 95% CIs associated with the RC and EE methods exhibit over-coverage. Conversely, the DB 95% CIs derived from the RC and EE methods display under-coverage, regardless of the regularity condition of the parameters. The W-type 95% CIs for the RC method have reasonable coverage, while for the EE methods, they may have over-coverage.

Following the methodology described in Section 4.5.2, we present the proportion of optimally treated future patients in Table 5.5, where in Scenario 1, the true covariate measurements are treated as available, and in Scenario 2, the true error-prone binary and continuous covariate measurements are unavailable but their surrogate measurements are available. The results obtained from Scenario 1 reveal that, in stage 1, both the RC and EE methods surpass the naive

method. In stage 2, it is evident that both the RC and EE methods outperform the naive method in regular and weak non-regular settings. However, in the non-regular setting, the naive method yields larger APCOT values than the RC and EE methods.

Turning our attention to Scenario 2, the findings demonstrate that the RC method yields superior results for stage 1 estimation compared to the naive method. Furthermore, when the mismeasurement degree is set to be $(0.1, 0.1, 1)$, the EE-estimated method outperforms the naive method. However, for other mismeasurement degrees, the naive method consistently yields larger values of APCOT compared to the EE-estimated method. The comparison of the EE-known method and the naive method shows that the naive method consistently yields larger APCOT values, regardless of the mismeasurement degree. For stage 2, in both regular and weak non-regular settings, both the RC and EE-estimated methods outperform the naive method. However, in the non-regular setting, the naive method results in larger APCOT values compared to both the RC and EE-estimated methods. The EE-known method exhibits superior performance to the naive method only for the weak non-regular setting.

5.7 Data Analysis

We collect publicly available COVID-19 data on 164 countries for a period of about ten months with time windows beginning from the date of the first confirmed COVID-19 case in each country; and the study periods of all those 164 countries span from April 1, 2020 to September 30, 2020. The data contain the information about containment and closure policies including workplace closure and international travel control, as well as health system policies including testing and contact tracing policies. In addition, the data include the information at the country level over the study period concerning the number of COVID-19 cases per million people, the total number of COVID-19 deaths, the economic growth percent change in quarterly real gross domestic product, the care system quality score, obesity prevalence, smoking prevalence, substance use prevalence, and socioeconomic factors.

Information regarding containment and closure policies as well as health system policies is collected from the OxCGRT (Hale et al., 2021). As in Section 3.3.1, the strictness score of

implementing each of the preventive policies on day t, denoted $I_{j,t}$, is calculated by:

$$I_{j,t} = 100\left\{\frac{v_{j,t} - 0.5(F_j - f_{j,t})}{N_j}\right\} \qquad j = 1, \cdots, 7,$$

where $v_{j,t}$, F_j, $f_{j,t}$, and N_j are defined in the same manner as in Section 3.3.1. Let $\mathcal{E}_1$ denote the set of labels for workplace closure and international travel control policies, and let $\mathcal{E}_2$ denote the set of labels for testing and contact tracing policies. Then the overall strictness score for policies of the same nature on a given day, denoted Index_{lt}, is calculated by:

$$\text{Index}_{lt} = \frac{1}{|\mathcal{E}_l|}\sum_{j\in\mathcal{E}_l} I_{j,t},$$

where $l = 1, 2$.

Data on the total number of COVID-19 cases per million people and the total number of COVID-19 deaths are extracted from the website Ourworldindata (Ritchie et al., 2020). Data on economic growth percent change in the quarterly real gross domestic product, denoted *eco-growth*, are extracted from the website "The Global Economy" (The Global Economy, 2020). Care system quality score (care-score), obesity prevalence (obesity-prev), smoking prevalence (smoking-prev), and substance use prevalence (substance-prev) for 2019 are obtained from the Legatum Institute (The Legatum Institute, 2019).

We consider the following socioeconomic factors: the most recent population weighted geometric mean density (popu-density) (Edwards et al., 2021), the population proportion of people aged 65 and above (senior-prop) for 2019 (The World Bank, 2019a), gross domestic product per capita based on purchasing power parity (GDP) for 2019 (The Global Economy, 2019), government effectiveness score (government-eff) for 2019 (The World Bank, 2019b), and infrastructure and market access score (infra-market) for 2019 (The Legatum Institute, 2019). The inclusion of the infra-market stems from acknowledging its role in measuring the quality of the infrastructure that enables trade and distortions in the market for goods and services. Its value ranges from 0 (worst) to 100 (best).

Although strict mitigation policies, such as international travel control policy, have been successful in slowing the spread of the virus (Wells et al., 2020), their inevitable impacts on

economic performance have been immense (Ozili and Arun, 2023). To examine the trade-off between the economic fallout and health costs of the pandemic, we can use the Q-learning approach with bivariate outcomes. This approach allows us to model the interdependence between economic and health outcomes and estimate the optimal policy decisions that minimize the overall costs. Since mismeasurement is ubiquitous in applications, it is interesting to investigate how mismeasurement in covariates can affect estimation results. Now we take the first outcome as the number of COVID-19 deaths per hundred COVID-19 cases, denoted *CFR*, and the second outcome as the negative of *eco-growth*, denoted $-$*eco-growth*, which we want to minimize. We examine how different degrees of covariate mismeasurement can affect the parameter estimates associated with policy decisions.

Since the information on the *eco-growth* is only available on a quarterly basis, the study period is six months long, starting from April 1, 2020 to September 30, 2020, we equally divide it into two stages with $K = 2$, with stage 1 starting from April 1, 2020 to June 30, 2020, and stage 2 starting from July 1, 2020 to September 30, 2020. The information about the total number of COVID-19 cases per million people, COVID-19 *CFR*, and the first quarter *eco-growth* of 2020 gathered at the end of March, 2020 is taken as the baseline features.

Following the same procedure in Section 3.4.1, we obtain *CFR* and stringency score of preventive policies for each stage. Furthermore, we log-transform the *CFR* at each stage to remove the nonnegativity constraint of *CFR*. The *eco-growth* for stages 1 and 2 represents the percent change in the real gross domestic product in the second and third quarters of 2020 compared to the second and third quarters of 2019, respectively.

For $k = 1, 2$, let A_k denote a binary action at stage k, which is defined as follows. For $l = 1, 2$ and $k = 1, 2$, let average-Index$_{l(k)}$ represent the average of the Index$_{lt}$ with t indexing the days in the period of stage k. Let $A_k = 1$ if average-Index$_{1(k)}$ is greater than average-Index$_{2(k)}$ and $A_k = 0$ otherwise. Let "cases-enter$_k$" represent the total number of COVID-19 cases per million people at the start of stage k, let "*CFR*-enter$_k$" represent the recorded *CFR* at the start of the stage k, and let "*eco-growth*-enter$_k$" represent the recorded *eco-growth* at the start of the stage k. As the government-eff value ranges from -2.5 to 2.5, we convert it into a binary variable, taking value 1 or 0, corresponding to "high government effectiveness" if its original value is no smaller than the threshold value 0, and "low government effectiveness" otherwise.

In the following analyses, for sensible comparisons, we normalize all the non-binary covariates by subtracting their means from the values and dividing them by their standard deviations. For the first outcome, *CFR*, we take the covariates senior-prop, GDP, government-eff, obesity-prev, smoking-prev, substance-prev, popu-density, and care-score as confounders, and treat GDP, government-eff, and popu-density, together with *CFR*-enter as prescriptive variables; for the second outcome, $-$*eco-growth*, we take GDP, government-eff, infra-market, and cases-enter as confounders, and consider GDP, government-eff, infra-market, together with *eco-growth*-enter$_k$ as prescriptive variables.

We employ linear regression models to describe the Q-functions for $k = 1, 2$:

$$\begin{aligned}
Q_2(H_2, A_2, \delta) = \delta\Big\{&\beta_{02,1} + \beta_{12,1} \times \text{senior-prop} + \beta_{22,1} \times \text{GDP} + \beta_{32,1} \times \text{government-eff} \\
&+ \beta_{42,1} \times \text{obesity-prev} + \beta_{52,1} \times \text{smoking-prev} + \beta_{62,1} \times \text{substance-prev} \\
&+ \beta_{72,1} \times \text{popu-density} + \beta_{82,1} \times \text{care-score} + (\psi_{02,1} + \psi_{12,1} \times \text{GDP} \\
&+ \psi_{22,1} \times \text{government-eff} + \psi_{32,1} \times \text{popu-density} + \psi_{42,1} \textit{CFR}\text{-enter}_2) A_2\Big\} \\
+ (1-\delta)\Big\{&\beta_{02,2} + \beta_{12,2} \times \text{GDP} + \beta_{22,2} \times \text{government-eff} + \beta_{32,2} \times \text{infra-market} \\
&+ \beta_{42,2} \times \text{cases-enter}_2 + (\psi_{02,2} + \psi_{12,2} \times \text{GDP} + \psi_{22,2} \times \text{government-eff} \\
&+ \psi_{32,2} \times \text{infra-market} + \psi_{42,2} \times \textit{eco-growth}\text{-enter}_2) A_2\Big\},
\end{aligned}$$

and

$$\begin{aligned}
Q_1(H_1, A_1, \delta) = \delta\Big\{&\beta_{01,1} + \beta_{11,1} \times \text{senior-prop} + \beta_{21,1} \times \text{GDP} + \beta_{31,1} \times \text{government-eff} \\
&+ \beta_{41,1} \times \text{obesity-prev} + \beta_{51,1} \times \text{smoking-prev} + \beta_{61,1} \times \text{substance-prev} \\
&+ \beta_{71,1} \times \text{popu-density} + \beta_{81,1} \times \text{care-score} + (\psi_{01,1} + \psi_{11,1} \times \text{GDP} \\
&+ \psi_{21,1} \times \text{government-eff} + \psi_{31,1} \times \text{popu-density} + \psi_{41,1} \textit{CFR}\text{-enter}_1) A_1\Big\} \\
+ (1-\delta)\Big\{&\beta_{01,2} + \beta_{11,2} \times \text{GDP} + \beta_{21,2} \times \text{government-eff} + \beta_{31,2} \times \text{infra-market} \\
&+ \beta_{41,2} \times \text{cases-enter}_1 + (\psi_{01,2} + \psi_{11,2} \times \text{GDP} + \psi_{21,2} \times \text{government-eff} \\
&+ \psi_{31,2} \times \text{infra-market} + \psi_{41,2} \times \textit{eco-growth}\text{-enter}_1) A_1\Big\}.
\end{aligned}$$

We conduct three analyses here by setting $\delta = 0.9$ or $\delta = 0.1$. In Analysis 1, we treat all the

variables to be error-free and implement the procedure in Section 5.2.1. The estimation results for the model parameters are reported in Tables 5.6-5.7 under the heading $\left(\pi_{10}, \pi_{01}, \sigma^2\right) = (0, 0, 0)$ for stages 1 and 2, respectively.

The next two analyses are to assess the effects of possibly mismeasured covariates on each stage parameter estimates. In particular, for the first and second outcomes, we take the government-eff as an error-prone binary covariate. We further take the popu-density as an error-prone continuous covariate for the first outcome, and the infra-market as the error-prone continuous covariate for the second outcome. We carry out sensitivity analyses using the two correction methods described in Sections 5.4 and 5.5, and respectively call them Analysis 2 and Analysis 3. For Analysis 2, we consider three sets of misclassification probabilities as well as measurement error degrees $\left(\pi_{10}, \pi_{01}, \sigma^2\right)^{\mathrm{T}} = (0.02, 0.02, 0.02)^{\mathrm{T}}, (0.03, 0.03, 0.05)^{\mathrm{T}}, (0.04, 0.04, 0.07)^{\mathrm{T}}$. For Analysis 3, we consider models (5.30) and (5.31) with

$$m\,(\text{government-eff}; \zeta) = \zeta_0 + \zeta_1 \times \text{government-eff},$$

$$m\,(\text{popu-density}; \xi_1) = \xi_{10} + \xi_{11} \times \text{popu-density},$$

and

$$m\,(\text{infra-market}; \xi_2) = \xi_{20} + \xi_{21} \times \text{infra-market},$$

where we consider three sets of values for the model parameters $\zeta = (\zeta_0, \zeta_1)^{\mathrm{T}}$, $\xi_1 = (\xi_{10}, \xi_{11})^{\mathrm{T}}$, and $\xi_2 = (\xi_{20}, \xi_{21})^{\mathrm{T}}$ that are listed in Table 5.8. Numerical results of Analysis 2 and Analysis 3 are respectively reported in Tables 5.6-5.7 and 5.9-5.10.

The numerical results of Analyses 1 and 2 reveal different evidence for the significance of some parameters. Consider the case when $\delta = 0.9$. For stage 1 and when the mismeasurement degree is set to be $(0, 0, 0)$, $(0.02, 0.02, 0.02)$, or $(0.04, 0.04, 0.07)$, there is no evidence to support the significance of $\psi_{21,1}$. In contrast, if the mismeasurement degree is $(0.03, 0.03, 0.05)$, $\psi_{21,1}$ is statistically significant. When $\delta = 0.1$ and the mismeasurement degree is set to be $(0, 0, 0)$, there exists no evidence suggesting the significance of $\psi_{21,1}$, $\psi_{01,2}$, and $\psi_{21,2}$, while for all the other mismeasurement degrees, $\psi_{21,1}$, $\psi_{01,2}$, and $\psi_{21,2}$ are found to be statistically significant. These findings suggest that if the government-eff is subject to misclassification

with sensitivity and specificity of 0.98, 0.97, or 0.96, naively estimating parameters may lower statistical power.

On the other hand, the three analyses do find some common evidence. Analyses 1, 2, and 3 collectively underscore the substantive importance of covariates associated with $\psi_{41,1}$ and $\psi_{41,2}$, regardless of δ being 0.9 or 0.1, and the extent of mismeasurement considered here. All the three analyses find the evidence to support that $\psi_{41,1}$ and $\psi_{41,2}$ are statistically significant.

Like the case for stage 1, Analyses 1 and 2 reveal different findings for stage 2. In particular, when the mismeasurement degree is set to be $(0, 0, 0)$, $\psi_{42,1}$ is statistically insignificant, while for all the other mismeasurement degrees, $\psi_{42,1}$ is significantly different from zero. Furthermore, $\psi_{42,2}$ is statistically significant when the mismeasurement degree is set to be $(0, 0, 0)$, while for all the other mismeasurement degrees, $\psi_{42,2}$ is statistically insignificant. Regarding Analysis 3, it is evident that $\psi_{12,1}$, $\psi_{42,1}$, and $\psi_{42,2}$ are statistically significant for all three sets of mismeasurement degrees.

In Tables 5.11-5.12, we report the estimated optimal actions at stages 1 and 2 derived from Analyses 1-3 for some selected countries with δ set to be 0.9 or 0.1, respectively. The countries included in Tables 5.11-5.12 are divided into two groups based on the disparities observed in Analyses 1-3. We consider Italy and the UK from the first group and the UAE, the USA, and Canada from the second group.

Examining Italy and the UK within the first group, it is evident that all conducted analyses support the precedence of health system policies over containment and closure policies in both stage 1 and stage 2. This holds true irrespective of the weight parameter $\delta = 0.9$ or 0.1. These findings underscore the significance of affording greater prominence to health system policies compared to containment and closure measures when prioritizing health outcomes. Moreover, even when accentuating economic considerations, the preeminence of health system policies over containment and closure strategies remains the same.

Now we consider the UAE within the second group. If reducing the first outcome *CFR* is more important than reducing the second outcome $-$*eco-growth* (i.e., when $\delta = 0.9$), Analysis 1 indicates the advantage of emphasizing containment and closure policies over health system policies in both stages 1 and 2; Analysis 2 mirrors these findings, except for the scenario with the most pronounced degree of mismeasurement, where the recommendation deviates by

suggesting a prioritization of health system policies in stage 1 and the opposite priority in stage 2. Conversely, the implications of Analysis 3 point toward the precedence of containment and closure policies in stage 1, which flips in favor of health system policies in stage 2. Shifting the perspective to $\delta = 0.1$, Analyses 1 and 2 both advocate for the prioritization of containment and closure policies in stage 1, followed by a shift in favor of health system policies in stage 2. This consensus remains true for Analysis 3 except with the highest degree of mismeasurement, which recommends a focus on health system policies across both stages 1 and 2.

Regarding the USA, Analyses 1 and 2 suggest giving precedence to health system policies in stage 1, and favoring containment and closure policies in stage 2 when $\delta = 0.9$. However, Analysis 3 suggests the opposite. When we change our perspective to $\delta = 0.1$, both Analyses 1 and 2 reach a common recommendation: prioritizing containment and closure policies in both stages 1 and 2. This consensus is robust, with the exception of Analysis 2 with the highest mismeasurement degree. In this specific case, the counsel pivots towards favoring health system policies in stage 1, and containment and closure policies in stage 2. Analysis 3 maintains its stance irrespective of the shift in δ, advocating the same prioritization strategy as in the scenario where $\delta = 0.9$.

Regarding Canada with δ set as 0.9, all three analyses imply a preference for prioritizing health system policies over containment and closure measures in both stages 1 and 2. This unanimity holds true except for Analysis 3 with the highest degree of mismeasurement, which recommends to prioritize health system policies over containment and closure measures solely in stage 1. As the parameter δ is adjusted to $\delta = 0.1$, Analyses 1 and 2 retain their alignment by suggesting precedence of health system policies for both stages 1 and 2. However, Analysis 3 introduces a nuanced perspective. When the mismeasurement degree is specified by Set 1, it maintains the stance of prioritizing health system policies over containment and closure measures. Yet, when the mismeasurement degree is specified by Set 2 or 3, a distinctive strategy emerges: favoring containment and closure policies over health system policies in stage 1 and vice versa in stage 2.

The variations in the results obtained from different analyses underscore the substantial adverse consequences stemming from covariate mismeasurement. These discrepancies highlight the potential implications of erroneous measurement of relevant factors in determining

optimal strategies. With the uncertainty of quantifying the potential influence of covariate mismeasurement, it is crucial to recognize the associated uncertainties and limitations in policy recommendations derived from such analyses.

5.8 Discussion

In this chapter, we demonstrate the substantially adverse effects of mixed misclassification and measurement error in covariates on parameter estimates using the Q-learning procedure with a composite outcome. We describe two methods for correcting the mismeasurement effects, namely regression calibration and unbiased estimation equation approaches. These proposed methods exhibit favorable performance by effectively reducing bias compared to the naive method.

Here, we consider the case with the availability of an internal validation subsample to characterize the mismeasurement degrees. In situations without additional data to quantify mismeasurement degrees or estimate the parameters $\boldsymbol{\zeta}_{X_k}$ and $\boldsymbol{\zeta}_{C_k}$ in models (5.30) and (5.31), sensitivity analyses are often employed to assess the impact of mismeasurement on the outcomes of the Q-learning algorithm. This involves selecting a set of representative values for $\boldsymbol{\zeta}_{X_k}$ and $\boldsymbol{\zeta}_{C_k}$, and then using models (5.30) and (5.31) to estimate X_{ki}^{**} and C_{ki}^{**} for $i = 1, \cdots, n$. Finally, the calibrated Q-learning algorithm is repeated to evaluate how the results may vary with different mismeasurement scenarios.

Table 5.1: Simulation studies for demonstrating biased estimation of the naive method in contrast to the EFLS method: stages 1-2. Entries in bold are obtained from the setting without mismeasurements

$(\pi_{10}, \pi_{01}, \sigma_k^2)$	Method		regular			weak non-regular			non-regular		
			$\psi_{0k\delta}$	$\psi_{1k\delta}$	$\psi_{2k\delta}$	$\psi_{0k\delta}$	$\psi_{1k\delta}$	$\psi_{2k\delta}$	$\psi_{0k\delta}$	$\psi_{1k\delta}$	$\psi_{2k\delta}$
		Stage 1 ($k = 1$)									
(0,0,0)	EFLS	Bias	**0.009**	**0.008**	**0.002**	**0.012**	**0.014**	**0.002**	**0.001**	**0.003**	**0.000**
		SE	**0.110**	**0.155**	**0.055**	**0.089**	**0.126**	**0.045**	**0.089**	**0.126**	**0.045**
		ESE	**0.125**	**0.158**	**0.058**	**0.105**	**0.126**	**0.045**	**0.117**	**0.132**	**0.044**
		MSE	**0.012**	**0.024**	**0.003**	**0.008**	**0.016**	**0.002**	**0.008**	**0.016**	**0.002**
		PBCR	**0.952**	**0.932**	**0.938**	**0.946**	**0.954**	**0.958**	**0.928**	**0.940**	**0.958**
		DBCR	**0.962**	**0.948**	**0.952**	**0.962**	**0.960**	**0.968**	**0.940**	**0.952**	**0.966**
(0.1,0.1,1)	Naive	Bias	0.129	0.235	0.504	0.152	0.242	0.505	0.134	0.249	0.501
		SE	0.107	0.151	0.038	0.100	0.141	0.035	0.100	0.141	0.035
		ESE	0.148	0.155	0.046	0.123	0.142	0.040	0.123	0.146	0.041
		MSE	0.028	0.078	0.255	0.033	0.078	0.256	0.028	0.082	0.252
		PBCR	0.872	0.748	0.000	0.788	0.664	0.000	0.840	0.690	0.000
		DBCR	0.758	0.488	0.000	0.690	0.432	0.000	0.726	0.412	0.000
(0.2,0.2,1.5)	Naive	Bias	0.268	0.487	0.691	0.303	0.489	0.695	0.255	0.492	0.691
		SE	0.108	0.152	0.030	0.105	0.148	0.029	0.105	0.148	0.029
		ESE	0.148	0.158	0.038	0.132	0.155	0.035	0.132	0.156	0.035
		MSE	0.083	0.260	0.478	0.103	0.261	0.484	0.076	0.264	0.478
		PBCR	0.608	0.020	0.000	0.388	0.034	0.000	0.496	0.014	0.000
		DBCR	0.448	0.046	0.000	0.286	0.052	0.000	0.398	0.050	0.000
(0.3,0.3,2)	Naive	Bias	0.386	0.724	0.799	0.446	0.718	0.799	0.365	0.726	0.800
		SE	0.077	0.108	0.017	0.076	0.108	0.017	0.076	0.108	0.017
		ESE	0.103	0.114	0.021	0.099	0.115	0.021	0.098	0.117	0.022
		MSE	0.155	0.536	0.639	0.205	0.527	0.639	0.139	0.539	0.640
		PBCR	0.018	0.000	0.000	0.002	0.000	0.000	0.004	0.000	0.000
		DBCR	0.052	0.000	0.000	0.008	0.000	0.000	0.034	0.000	0.000
		Stage 2 ($k = 2$)									
(0,0,0)	EFLS	Bias	**0.002**	**0.005**	**0.001**	**0.001**	**0.001**	**0.003**	**0.009**	**0.003**	**0.003**
		SE	**0.086**	**0.116**	**0.041**	**0.086**	**0.116**	**0.041**	**0.086**	**0.116**	**0.041**
		ESE	**0.090**	**0.118**	**0.041**	**0.086**	**0.118**	**0.041**	**0.088**	**0.117**	**0.043**
		MSE	**0.007**	**0.013**	**0.002**	**0.007**	**0.013**	**0.002**	**0.007**	**0.013**	**0.002**
		WTCR	**0.948**	**0.948**	**0.948**	**0.954**	**0.942**	**0.944**	**0.936**	**0.964**	**0.930**
(0.1,0.1,1)	Naive	Bias	0.170	0.284	0.843	0.238	0.394	0.001	0.013	0.035	0.001
		SE	0.121	0.164	0.041	0.100	0.136	0.034	0.088	0.120	0.030
		ESE	0.129	0.174	0.053	0.102	0.135	0.041	0.089	0.119	0.032
		MSE	0.044	0.108	0.712	0.067	0.174	0.001	0.008	0.016	0.001
		WTCR	0.720	0.568	0.000	0.314	0.160	0.902	0.948	0.948	0.926
(0.2,0.2,1.5)	Naive	Bias	0.346	0.595	1.167	0.449	0.778	0.003	0.046	0.073	0.002
		SE	0.134	0.183	0.036	0.109	0.149	0.029	0.089	0.122	0.024
		ESE	0.141	0.190	0.048	0.121	0.156	0.037	0.097	0.131	0.025
		MSE	0.138	0.388	1.363	0.213	0.628	0.001	0.010	0.020	0.001
		WTCR	0.286	0.100	0.000	0.020	0.004	0.874	0.898	0.894	0.946
(0.3,0.3,2)	Naive	Bias	0.499	0.870	1.347	0.676	1.169	0.001	0.066	0.110	0.001
		SE	0.099	0.137	0.022	0.081	0.112	0.018	0.063	0.088	0.014
		ESE	0.102	0.139	0.029	0.090	0.119	0.022	0.065	0.094	0.015
		MSE	0.259	0.776	1.815	0.463	1.379	0.000	0.008	0.020	0.000
		WTCR	0.004	0.000	0.000	0.000	0.000	0.888	0.828	0.728	0.940

Table 5.2: Simulation studies for assessing the performance of the RC, EE-known, and EE-estimated methods: stages 1-2 and regular case

$(\pi_{10}, \pi_{01}, \sigma_k^2)$		RC			EE-known			EE-estimated		
		$\psi_{0k\delta}$	$\psi_{1k\delta}$	$\psi_{2k\delta}$	$\psi_{0k\delta}$	$\psi_{1k\delta}$	$\psi_{2k\delta}$	$\psi_{0k\delta}$	$\psi_{1k\delta}$	$\psi_{2k\delta}$
		Stage 1 ($k = 1$)								
(0.1,0.1,1)	Bias	0.010	0.012	0.001	0.010	0.010	0.008	0.008	0.024	0.011
	SE	0.150	0.190	0.083	0.175	0.271	0.130	0.208	0.265	0.224
	ESE	0.152	0.187	0.083	0.199	0.287	0.142	0.181	0.247	0.140
	MSE	0.023	0.036	0.007	0.031	0.074	0.017	0.043	0.071	0.050
	PBCR	0.984	0.996	1.000	0.994	1.000	1.000	0.994	1.000	0.994
	DBCR	0.896	0.862	0.780	0.808	0.738	0.618	0.844	0.808	0.632
(0.2,0.2,1.5)	Bias	0.011	0.003	0.004	0.035	0.018	0.059	0.013	0.009	0.064
	SE	0.168	0.226	0.095	0.296	0.504	0.269	0.319	0.492	0.434
	ESE	0.164	0.225	0.090	0.316	0.542	0.29	0.291	0.476	0.312
	MSE	0.028	0.051	0.009	0.089	0.254	0.076	0.102	0.242	0.192
	PBCR	0.994	1.000	1.000	1.000	1.000	1.000	1.000	1.000	0.994
	DBCR	0.880	0.784	0.754	0.612	0.490	0.402	0.646	0.518	0.402
(0.3,0.3,2)	Bias	0.019	0.005	0.005	0.013	0.059	0.097	0.010	0.060	0.111
	SE	0.129	0.184	0.072	0.415	0.766	0.382	0.417	0.740	0.488
	ESE	0.124	0.180	0.071	0.436	0.808	0.390	0.398	0.711	0.413
	MSE	0.017	0.034	0.005	0.172	0.590	0.155	0.174	0.551	0.250
	PBCR	0.994	0.998	1.000	1.000	1.000	1.000	1.000	1.000	1.000
	DBCR	0.844	0.740	0.714	0.392	0.290	0.270	0.394	0.346	0.248
		Stage 2 ($k = 2$)								
(0.1,0.1,1)	Bias	0.004	0.001	0.000	0.004	0.020	0.022	0.008	0.004	0.036
	SE	0.138	0.179	0.093	0.170	0.246	0.150	0.188	0.236	0.200
	ESE	0.136	0.175	0.095	0.176	0.249	0.162	0.163	0.227	0.170
	MSE	0.019	0.032	0.009	0.029	0.061	0.023	0.035	0.056	0.041
	WTCR	0.954	0.954	0.934	0.940	0.936	0.934	0.976	0.958	0.990
(0.2,0.2,1.5)	Bias	0.005	0.012	0.002	0.028	0.041	0.108	0.030	0.039	0.118
	SE	0.163	0.217	0.108	0.312	0.483	0.357	0.394	0.562	0.573
	ESE	0.158	0.212	0.106	0.328	0.508	0.359	0.293	0.444	0.399
	MSE	0.027	0.047	0.012	0.098	0.235	0.139	0.156	0.317	0.342
	WTCR	0.960	0.942	0.958	0.954	0.966	0.966	0.984	0.978	0.966
(0.3,0.3,2)	Bias	0.001	0.002	0.004	0.074	0.122	0.151	0.043	0.085	0.104
	SE	0.123	0.165	0.079	0.433	0.715	0.540	0.428	0.710	0.554
	ESE	0.113	0.154	0.078	0.410	0.671	0.514	0.354	0.590	0.464
	MSE	0.015	0.027	0.006	0.193	0.526	0.314	0.185	0.511	0.318
	WTCR	0.974	0.972	0.948	0.976	0.992	0.958	0.992	0.994	0.920

Table 5.3: Simulation studies for assessing the performance of the RC, EE-known, and EE-estimated methods: stages 1-2 and weak non-regular case

$(\pi_{10}, \pi_{01}, \sigma_k^2)$		RC			EE-known			EE-estimated		
		$\psi_{0k\delta}$	$\psi_{1k\delta}$	$\psi_{2k\delta}$	$\psi_{0k\delta}$	$\psi_{1k\delta}$	$\psi_{2k\delta}$	$\psi_{0k\delta}$	$\psi_{1k\delta}$	$\psi_{2k\delta}$
		Stage 1 ($k = 1$)								
(0.1,0.1,1)	Bias	0.007	0.000	0.005	0.014	0.005	0.007	0.026	0.010	0.004
	SE	0.133	0.168	0.074	0.134	0.208	0.107	0.162	0.204	0.181
	ESE	0.130	0.159	0.074	0.156	0.222	0.112	0.146	0.199	0.113
	MSE	0.018	0.028	0.006	0.018	0.043	0.011	0.027	0.042	0.033
	PBCR	0.996	0.988	0.998	0.998	1.000	1.000	0.986	1.000	0.990
	DBCR	0.896	0.892	0.792	0.834	0.766	0.638	0.860	0.810	0.620
(0.2,0.2,1.5)	Bias	0.008	0.004	0.002	0.040	0.005	0.040	0.036	0.013	0.054
	SE	0.150	0.200	0.086	0.220	0.375	0.228	0.242	0.373	0.301
	ESE	0.147	0.198	0.087	0.243	0.405	0.231	0.231	0.372	0.261
	MSE	0.023	0.040	0.007	0.050	0.141	0.054	0.060	0.139	0.094
	PBCR	0.984	0.998	1.000	1.000	1.000	1.000	1.000	1.000	0.996
	DBCR	0.868	0.822	0.708	0.674	0.550	0.422	0.676	0.580	0.376
(0.3,0.3,2)	Bias	0.003	0.003	0.003	0.011	0.138	0.132	0.017	0.036	0.098
	SE	0.114	0.158	0.064	0.340	0.624	0.373	0.342	0.598	0.418
	ESE	0.115	0.156	0.061	0.359	0.662	0.394	0.307	0.555	0.346
	MSE	0.013	0.025	0.004	0.116	0.408	0.157	0.117	0.359	0.184
	PBCR	0.992	1.000	1.000	1.000	1.000	1.000	1.000	1.000	0.992
	DBCR	0.847	0.746	0.732	0.419	0.352	0.237	0.450	0.358	0.268
		Stage 2 ($k = 2$)								
(0.1,0.1,1)	Bias	0.004	0.003	0.001	0.005	0.013	0.000	0.007	0.009	0.002
	SE	0.117	0.160	0.061	0.125	0.173	0.085	0.154	0.172	0.160
	ESE	0.113	0.157	0.065	0.123	0.168	0.096	0.133	0.174	0.077
	MSE	0.014	0.026	0.004	0.016	0.030	0.007	0.024	0.030	0.026
	WTCR	0.956	0.954	0.944	0.960	0.968	0.934	0.986	0.958	1.000
(0.2,0.2,1.5)	Bias	0.009	0.000	0.001	0.048	0.065	0.002	0.031	0.052	0.005
	SE	0.138	0.195	0.073	0.191	0.271	0.178	0.222	0.287	0.293
	ESE	0.140	0.202	0.072	0.198	0.276	0.193	0.200	0.270	0.173
	MSE	0.019	0.038	0.005	0.039	0.078	0.032	0.050	0.085	0.086
	WTCR	0.958	0.950	0.946	0.938	0.942	0.956	0.974	0.944	0.994
(0.3,0.3,2)	Bias	0.005	0.006	0.000	0.061	0.114	0.016	0.060	0.100	0.009
	SE	0.105	0.152	0.056	0.257	0.414	0.398	0.268	0.381	0.348
	ESE	0.109	0.154	0.057	0.238	0.341	0.308	0.244	0.332	0.225
	MSE	0.011	0.023	0.003	0.070	0.184	0.159	0.075	0.155	0.121
	WTCR	0.941	0.951	0.939	0.963	0.963	0.996	0.959	0.959	0.996

Table 5.4: Simulation studies for assessing the performance of the RC, EE-known, and EE-estimated methods: stages 1-2 and non-regular case

$(\pi_{10}, \pi_{01}, \sigma_k^2)$		RC			EE-known			EE-estimated		
		$\psi_{0k\delta}$	$\psi_{1k\delta}$	$\psi_{2k\delta}$	$\psi_{0k\delta}$	$\psi_{1k\delta}$	$\psi_{2k\delta}$	$\psi_{0k\delta}$	$\psi_{1k\delta}$	$\psi_{2k\delta}$
		Stage 1 ($k = 1$)								
(0.1,0.1,1)	Bias	0.005	0.003	0.001	0.007	0.004	0.019	0.005	0.009	0.018
	SE	0.129	0.167	0.074	0.130	0.201	0.106	0.160	0.198	0.181
	ESE	0.131	0.167	0.072	0.150	0.211	0.114	0.149	0.193	0.120
	MSE	0.017	0.028	0.005	0.017	0.040	0.012	0.026	0.039	0.033
	PBCR	0.980	0.986	1.000	0.996	0.994	1.000	0.978	0.994	0.996
	DBCR	0.894	0.872	0.806	0.844	0.788	0.652	0.854	0.820	0.590
(0.2,0.2,1.5)	Bias	0.001	0.000	0.003	0.007	0.020	0.060	0.002	0.018	0.034
	SE	0.144	0.198	0.085	0.207	0.349	0.224	0.222	0.340	0.294
	ESE	0.150	0.201	0.083	0.222	0.363	0.247	0.209	0.343	0.245
	MSE	0.021	0.039	0.007	0.043	0.122	0.054	0.049	0.116	0.088
	PBCR	0.984	0.994	0.998	1.000	1.000	1.000	1.000	1.000	0.988
	DBCR	0.862	0.784	0.712	0.686	0.576	0.332	0.714	0.622	0.362
(0.3,0.3,2)	Bias	0.005	0.004	0.002	0.037	0.064	0.091	0.017	0.026	0.082
	SE	0.109	0.157	0.064	0.285	0.515	0.352	0.293	0.521	0.404
	ESE	0.108	0.152	0.061	0.306	0.545	0.359	0.271	0.480	0.323
	MSE	0.012	0.025	0.004	0.083	0.269	0.132	0.086	0.272	0.170
	PBCR	0.990	1.000	1.000	1.000	1.000	1.000	1.000	1.000	0.990
	DBCR	0.837	0.757	0.712	0.424	0.310	0.245	0.502	0.357	0.255
		Stage 2 ($k = 2$)								
(0.1,0.1,1)	Bias	0.010	0.001	0.001	0.013	0.008	0.000	0.002	0.008	0.003
	SE	0.097	0.138	0.053	0.104	0.151	0.064	0.141	0.154	0.150
	ESE	0.098	0.136	0.054	0.102	0.146	0.068	0.101	0.142	0.058
	MSE	0.010	0.019	0.003	0.011	0.023	0.004	0.020	0.024	0.023
	WTCR	0.932	0.948	0.936	0.946	0.960	0.934	0.992	0.962	1.000
(0.2,0.2,1.5)	Bias	0.008	0.016	0.001	0.009	0.021	0.004	0.001	0.005	0.007
	SE	0.109	0.161	0.059	0.139	0.212	0.098	0.173	0.220	0.260
	ESE	0.113	0.164	0.060	0.149	0.222	0.099	0.131	0.195	0.093
	MSE	0.012	0.026	0.003	0.019	0.045	0.010	0.030	0.048	0.068
	WTCR	0.944	0.956	0.958	0.930	0.942	0.956	0.994	0.974	1.000
(0.3,0.3,2)	Bias	0.002	0.004	0.005	0.024	0.045	0.007	0.010	0.020	0.002
	SE	0.085	0.130	0.045	0.152	0.241	0.121	0.183	0.271	0.252
	ESE	0.078	0.124	0.046	0.151	0.243	0.115	0.132	0.204	0.116
	MSE	0.007	0.017	0.002	0.024	0.060	0.015	0.034	0.074	0.064
	WTCR	0.965	0.971	0.951	0.955	0.937	0.982	0.982	0.980	1.000

Table 5.5: Proportions of optimally treated individuals

		Scenario 1							
		Stage 1				Stage 2			
$(\pi_{10}, \pi_{01}, \sigma_k^2)$	Setting	Naive	RC	EE-known	EE-estimated	Naive	RC	EE-known	EE-estimated
(0.1,0.1,1)	Regular	0.880	0.979	0.970	0.973	0.870	0.986	0.981	0.983
	Weak non-regular	0.886	0.981	0.976	0.978	0.460	0.768	0.770	0.774
	Non-regular	0.878	0.982	0.978	0.979	0.784	0.721	0.717	0.716
(0.2,0.2,1.5)	Regular	0.820	0.976	0.949	0.954	0.775	0.983	0.968	0.969
	Weak non-regular	0.829	0.978	0.960	0.963	0.445	0.771	0.767	0.772
	Non-regular	0.824	0.980	0.962	0.965	0.813	0.711	0.690	0.697
(0.3,0.3,2)	Regular	0.808	0.973	0.908	0.915	0.732	0.983	0.944	0.947
	Weak non-regular	0.817	0.977	0.924	0.928	0.445	0.768	0.768	0.766
	Non-regular	0.816	0.979	0.934	0.939	0.865	0.712	0.662	0.677
		Scenario 2							
(0.1,0.1,1)	Regular	0.850	0.888	0.812	0.861	0.806	0.860	0.783	0.843
	Weak non-regular	0.855	0.893	0.814	0.864	0.461	0.556	0.891	0.860
	Non-regular	0.855	0.893	0.816	0.865	0.780	0.769	0.663	0.676
(0.2,0.2,1.5)	Regular	0.820	0.867	0.738	0.804	0.763	0.829	0.714	0.790
	Weak non-regular	0.827	0.873	0.741	0.808	0.445	0.543	0.796	0.790
	Non-regular	0.828	0.873	0.745	0.811	0.832	0.810	0.606	0.636
(0.3,0.3,2)	Regular	0.810	0.859	0.671	0.744	0.742	0.814	0.662	0.747
	Weak non-regular	0.818	0.866	0.674	0.751	0.444	0.534	0.696	0.714
	Non-regular	0.818	0.866	0.680	0.757	0.851	0.839	0.558	0.591

Table 5.6: Analyses 1 and 2 results for stage 1 parameters with their W-type 95% CIs

$(\pi_{10}, \pi_{01}, \sigma^2)$		$\psi_{01,1}$	$\psi_{11,1}$	$\psi_{21,1}$	$\psi_{31,1}$	$\psi_{41,1}$	$\psi_{01,2}$	$\psi_{11,2}$	$\psi_{21,2}$	$\psi_{31,2}$	$\psi_{41,2}$
						$\delta = \mathbf{0.9}$					
	Estimate	0.14	0.24	0.02	0.35	0.48	5.47	-6.13	-0.48	1.72	-1.21
(0,0,0)	SE	0.34	0.21	0.48	0.43	0.11	3.41	2.43	5.47	3.92	0.33
	95% CI	(-0.54, 0.82)	(-0.18, 0.66)	(-0.93, 0.97)	(-0.49, 1.20)	(0.26, 0.69)	(-1.30, 12.24)	(-10.95, -1.30)	(-11.34, 10.38)	(-6.05, 9.48)	(-1.88, -0.55)
	Estimate	0.15	0.24	0.01	0.38	0.47	4.98	-6.62	0.21	2.03	-1.17
(0.02,0.02, 0.02)	SE	0.17	0.14	0.02	0.26	0.08	2.51	5.47	1.01	8.91	0.32
	95% CI	(-0.18, 0.48)	(-0.02, 0.51)	(-0.03, 0.06)	(-0.14, 0.90)	(0.32, 0.62)	(0.05, 9.91)	(-17.34, 4.11)	(-1.78, 2.19)	(-15.43, 19.49)	(-1.79, -0.55)
	Estimate	0.09	0.24	0.08	0.35	0.47	7.01	-4.47	-1.21	-0.38	-1.26
(0.03, 0.03, 0.05)	SE	0.17	0.14	0.04	0.28	0.08	2.70	5.98	1.56	9.34	0.34
	95% CI	(-0.25, 0.42)	(-0.03, 0.51)	(0.00, 0.16)	(-0.19, 0.90)	(0.32, 0.63)	(1.72, 12.29)	(-16.19, 7.26)	(-4.26, 1.83)	(-18.68, 17.92)	(-1.92, -0.60)
	Estimate	0.14	0.25	-0.04	0.48	0.45	8.66	-4.45	-1.46	-1.03	-1.46
(0.04,0.04, 0.07)	SE	0.17	0.14	0.03	0.30	0.08	3.13	7.13	2.68	10.28	0.44
	95% CI	(-0.20, 0.48)	(-0.03, 0.53)	(-0.09, 0.01)	(-0.12, 1.07)	(0.29, 0.61)	(2.53, 14.78)	(-18.42, 9.53)	(-6.72, 3.80)	(-21.18, 19.13)	(-2.32, -0.60)
						$\delta = \mathbf{0.1}$					
	Estimate	-0.28	-0.01	0.68	0.33	0.47	5.69	-4.38	-0.47	0.03	-1.28
(0,0,0)	SE	0.41	0.25	0.57	0.50	0.13	3.49	2.49	5.60	4.01	0.34
	95% CI	(-1.08, 0.53)	(-0.50, 0.49)	(-0.44, 1.81)	(-0.67, 1.32)	(0.22, 0.71)	(-1.23, 12.62)	(-9.32, 0.56)	(-11.58, 10.64)	(-7.92, 7.98)	(-1.96, -0.60)
	Estimate	-0.31	-0.15	0.82	0.50	0.48	7.32	-4.68	-2.74	1.27	-1.26
(0.02,0.02, 0.02)	SE	0.20	0.20	0.11	0.32	0.09	2.61	5.44	1.32	8.63	0.32
	95% CI	(-0.71, 0.09)	(-0.54, 0.24)	(0.59, 1.04)	(-0.13, 1.13)	(0.31, 0.66)	(2.21, 12.43)	(-15.34, 5.98)	(-5.32, -0.16)	(-15.65, 18.18)	(-1.88, -0.64)
	Estimate	-0.47	-0.15	0.96	0.13	0.50	7.58	-6.24	-4.71	4.46	-1.24
(0.03, 0.03, 0.05)	SE	0.21	0.20	0.14	0.36	0.09	2.68	5.52	1.44	8.59	0.31
	95% CI	(-0.88, -0.05)	(-0.53, 0.24)	(0.67, 1.24)	(-0.58, 0.83)	(0.33, 0.67)	(2.34, 12.83)	(-17.05, 4.58)	(-7.54, -1.87)	(-12.38, 21.30)	(-1.85, -0.63)
	Estimate	-0.69	-0.39	1.38	0.52	0.48	16.66	-1.09	-20.44	5.25	-1.83
(0.04,0.04, 0.07)	SE	0.22	0.23	0.16	0.40	0.10	4.62	10.37	5.80	12.60	0.55
	95% CI	(-1.13, -0.25)	(-0.84, 0.06)	(1.06, 1.70)	(-0.26, 1.30)	(0.29, 0.68)	(7.61, 25.71)	(-21.41, 19.24)	(-31.82, -9.07)	(-19.46, 29.95)	(-2.92, -0.75)

Table 5.7: Analyses 1 and 2 results for stage 2 parameters with their W-type 95% CIs

$(\pi_{10}, \pi_{01}, \sigma^2)$		$\psi_{02,1}$	$\psi_{12,1}$	$\psi_{22,1}$	$\psi_{32,1}$	$\psi_{42,1}$	$\psi_{02,2}$	$\psi_{12,2}$	$\psi_{22,2}$	$\psi_{32,2}$	$\psi_{42,2}$
	Estimate	0.09	-0.80	0.56	0.45	0.28	-2.27	3.36	-2.32	-4.13	-0.48
(0,0,0)	SE	0.32	0.38	0.43	0.84	0.15	2.02	2.18	2.55	2.09	0.08
	95% CI	(-0.54, 0.72)	(-1.56, -0.04)	(-0.30, 1.42)	(-1.21, 2.10)	(-0.02, 0.57)	(-6.27, 1.73)	(-0.97, 7.70)	(-7.38, 2.74)	(-8.27, 0.01)	(-0.64, -0.31)
	Estimate	0.07	-0.85	0.63	0.46	0.28	-2.09	3.35	-2.63	-3.98	-0.47
(0.02,0.02, 0.02)	SE	0.27	0.39	0.48	0.58	0.12	8.44	9.12	10.93	9.50	0.53
	95% CI	(-0.46, 0.59)	(-1.62, -0.08)	(-0.30, 1.56)	(-0.67, 1.59)	(0.05, 0.51)	(-18.63, 14.45)	(-14.52, 21.22)	(-24.05, 18.79)	(-22.59, 14.64)	(-1.52, 0.57)
	Estimate	0.05	-0.82	0.64	0.40	0.30	-2.29	3.39	-2.41	-4.11	-0.48
(0.03, 0.03, 0.05)	SE	0.29	0.44	0.53	0.65	0.12	8.88	9.22	12.12	10.01	0.54
	95% CI	(-0.52, 0.63)	(-1.67, 0.03)	(-0.40, 1.67)	(-0.88, 1.67)	(0.06, 0.53)	(-19.69, 15.11)	(-14.68, 21.46)	(-26.16, 21.35)	(-23.74, 15.51)	(-1.53, 0.57)
	Estimate	-0.04	-1.07	0.83	0.75	0.27	-2.23	3.22	-2.45	-3.87	-0.47
(0.04,0.04, 0.07)	SE	0.34	0.49	0.61	0.75	0.13	9.41	9.26	13.62	10.66	0.53
	95% CI	(-0.71, 0.63)	(-2.04, -0.11)	(-0.37, 2.03)	(-0.73, 2.22)	(0.01, 0.53)	(-20.68, 16.21)	(-14.92, 21.37)	(-29.15, 24.25)	(-24.76, 17.03)	(-1.52, 0.57)

Table 5.8: Values of regression parameters of calibration functions for Analysis 3

	ζ_0	ζ_1	ξ_{10}	ξ_{11}	ξ_{20}	ξ_{21}
Set 1	0.05	0.95	0.05	-0.85	0.05	-0.85
Set 2	0.05	0.85	0.05	-0.75	0.05	-0.75
Set 3	0.05	0.75	0.05	-0.65	0.05	-0.65

Table 5.9: Analysis 3 results for stage 1 parameters with their Normal bootstrap 95% CIs

		$\psi_{01,1}$	$\psi_{11,1}$	$\psi_{21,1}$	$\psi_{31,1}$	$\psi_{41,1}$	$\psi_{01,2}$	$\psi_{11,2}$	$\psi_{21,2}$	$\psi_{31,2}$	$\psi_{41,2}$
						$\delta = \mathbf{0.9}$					
	Estimate	0.12	0.24	0.09	-0.41	0.48	6.70	-6.13	-2.20	-2.02	-1.21
Set 1	Bootstrap SE	2.07	0.29	3.09	0.69	0.11	15.44	3.53	24.80	5.32	0.31
	95% CI	(-3.94, 4.17)	(-0.34, 0.82)	(-5.97, 6.16)	(-1.77, 0.94)	(0.25, 0.70)	(-23.55, 36.96)	(-13.04, 0.79)	(-50.81, 46.41)	(-12.44, 8.41)	(-1.81, -0.62)
	Estimate	0.11	0.24	0.10	-0.47	0.48	6.83	-6.13	-2.43	-2.29	-1.21
Set 2	Bootstrap SE	2.25	0.31	3.44	0.88	0.11	17.05	3.54	27.84	6.25	0.31
	95% CI	(-4.30, 4.53)	(-0.37, 0.85)	(-6.65, 6.85)	(-2.19, 1.25)	(0.26, 0.69)	(-26.58, 40.24)	(-13.07, 0.82)	(-57.00, 52.14)	(-14.54, 9.96)	(-1.83, -0.60)
	Estimate	0.11	0.24	0.12	-0.54	0.48	7.00	-6.13	-2.71	-2.64	-1.21
Set 3	Bootstrap SE	2.30	0.28	3.59	0.97	0.11	18.13	3.65	30.19	7.37	0.31
	95% CI	(-4.39, 4.61)	(-0.32, 0.80)	(-6.93, 7.16)	(-2.44, 1.36)	(0.26, 0.69)	(-28.53, 42.53)	(-13.28, 1.02)	(-61.89, 56.46)	(-17.08, 11.80)	(-1.83, -0.60)
						$\delta = \mathbf{0.1}$					
	Estimate	-1.86	-0.01	3.13	-0.38	0.47	6.80	-4.38	-2.16	-0.03	-1.28
Set 1	Bootstrap SE	2.27	0.37	3.36	0.87	0.14	15.36	3.42	24.85	5.32	0.31
	95% CI	(-6.30, 2.58)	(-0.72, 0.71)	(-3.46, 9.71)	(-2.09, 1.32)	(0.20, 0.73)	(-23.31, 36.91)	(-11.09, 2.32)	(-50.86, 46.54)	(-10.46, 10.39)	(-1.89, -0.68)
	Estimate	-2.02	-0.01	3.44	-0.43	0.47	6.91	-4.38	-2.38	-0.04	-1.28
Set 2	Bootstrap SE	2.52	0.37	3.80	0.97	0.13	16.37	3.41	26.90	5.98	0.30
	95% CI	(-6.96, 2.92)	(-0.74, 0.73)	(-4.00, 10.89)	(-2.34, 1.47)	(0.20, 0.73)	(-25.18, 39.00)	(-11.07, 2.31)	(-55.11, 50.36)	(-11.75, 11.67)	(-1.88, -0.69)
	Estimate	-2.22	-0.01	3.85	-0.50	0.47	7.06	-4.38	-2.66	-0.05	-1.28
Set 3	Bootstrap SE	2.82	0.38	4.38	1.16	0.13	17.73	3.48	29.49	6.90	0.30
	95% CI	(-7.75, 3.30)	(-0.74, 0.73)	(-4.74, 12.45)	(-2.78, 1.78)	(0.21, 0.72)	(-27.69, 41.80)	(-11.19, 2.43)	(-60.46, 55.15)	(-13.58, 13.49)	(-1.88, -0.69)

Table 5.10: Analysis 3 results for stage 2 parameters with their Normal bootstrap 95% CIs

		$\psi_{02,1}$	$\psi_{12,1}$	$\psi_{22,1}$	$\psi_{32,1}$	$\psi_{42,1}$	$\psi_{02,2}$	$\psi_{12,2}$	$\psi_{22,2}$	$\psi_{32,2}$	$\psi_{42,2}$
	Estimate	-1.20	-0.80	2.57	-0.53	0.28	2.92	3.36	-10.60	4.86	-0.48
Set 1	Bootstrap SE	1.21	0.39	1.82	1.38	0.14	7.49	2.19	11.67	2.63	0.09
	95% CI	(-3.57, 1.17)	(-1.56, -0.04)	(-1.00, 6.14)	(-3.24, 2.19)	(0.00, 0.56)	(-11.76, 17.59)	(-0.92, 7.65)	(-33.47, 12.26)	(-0.30, 10.02)	(-0.65, -0.30)
	Estimate	-1.33	-0.80	2.83	-0.60	0.28	3.44	3.37	-11.68	5.51	-0.48
Set 2	Bootstrap SE	1.33	0.40	2.06	1.51	0.14	7.64	2.18	12.15	3.16	0.09
	95% CI	(-3.94, 1.27)	(-1.58, -0.02)	(-1.20, 6.86)	(-3.55, 2.36)	(0.01, 0.55)	(-11.54, 18.41)	(-0.91, 7.64)	(-35.50, 12.14)	(-0.69, 11.71)	(-0.65, -0.31)
	Estimate	-1.50	-0.80	3.17	-0.69	0.28	4.10	3.37	-13.06	6.36	-0.48
Set 3	Bootstrap SE	1.41	0.39	2.21	1.71	0.13	8.92	2.30	14.33	3.62	0.09
	95% CI	(-4.26, 1.26)	(-1.56, -0.04)	(-1.16, 7.49)	(-4.04, 2.67)	(0.02, 0.54)	(-13.38, 21.59)	(-1.15, 7.88)	(-41.14, 15.03)	(-0.73, 13.44)	(-0.65, -0.31)

Table 5.11: Estimated optimal actions for selected countries from Analyses 1, 2 and 3: $\delta = 0.9$

Country	Analysis 1		Analysis 2						Analysis 3					
	(0,0,0)		(0.02,0.02,0.02)		(0.03,0.03,0.05)		(0.04,0.04,0.07)		Set 1		Set 2		Set 3	
	Stage 1	Stage 2	Stage 1	Stage 2	Stage 1	Stage 2	Stage 1	Stage 2	Stage 1	Stage 2	Stage 1	Stage 2	Stage 1	Stage 2
	$\hat{A}_1^{opt}$	$\hat{A}_2^{opt}$	$\hat{A}_1^{opt}$	$\hat{A}_2^{opt}$	$\hat{A}_1^{opt}$	$\hat{A}_2^{opt}$	$\hat{A}_1^{opt}$	$\hat{A}_2^{opt}$	$\hat{A}_1^{opt}$	$\hat{A}_2^{opt}$	$\hat{A}_1^{opt}$	$\hat{A}_2^{opt}$	$\hat{A}_1^{opt}$	$\hat{A}_2^{opt}$
Greece	0	0	0	0	0	0	0	0	0	0	0	0	0	0
India	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Italy	0	0	0	0	0	0	0	0	0	0	0	0	0	0
United Kingdom	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Sweden	0	1	0	1	0	1	0	1	0	0	0	0	0	0
China	0	1	0	1	0	1	0	1	0	0	0	0	0	0
France	0	0	0	0	0	0	0	1	0	0	0	0	0	0
Egypt	0	0	0	0	0	0	0	0	0	1	0	1	0	1
South Korea	0	1	0	1	0	1	0	1	1	0	1	0	1	0
Hong Kong	0	0	0	0	0	0	0	0	1	1	1	1	1	1
USA	0	1	0	1	0	1	0	1	1	0	1	0	1	0
United Arab Emirates	1	1	1	1	1	1	0	1	1	0	1	0	1	0
Canada	0	0	0	0	0	0	0	1	0	0	0	0	0	0
Israel	0	1	0	1	0	1	0	1	1	0	1	0	1	0
Finland	0	1	0	1	0	1	0	1	1	0	1	0	1	0
Germany	0	1	0	1	0	1	0	1	1	0	1	0	1	0

Table 5.12: Estimated optimal actions for selected countries from Analyses 1, 2 and 3: $\delta = 0.1$

Country	Analysis 1		Analysis 2						Analysis 3					
	(0,0,0)		(0.02,0.02,0.02)		(0.03,0.03,0.05)		(0.04,0.04,0.07)		Set 1		Set 2		Set 3	
	Stage 1	Stage 2	Stage 1	Stage 2	Stage 1	Stage 2	Stage 1	Stage 2	Stage 1	Stage 2	Stage 1	Stage 2	Stage 1	Stage 2
	$\hat{A}_1^{opt}$	$\hat{A}_2^{opt}$	$\hat{A}_1^{opt}$	$\hat{A}_2^{opt}$	$\hat{A}_1^{opt}$	$\hat{A}_2^{opt}$	$\hat{A}_1^{opt}$	$\hat{A}_2^{opt}$	$\hat{A}_1^{opt}$	$\hat{A}_2^{opt}$	$\hat{A}_1^{opt}$	$\hat{A}_2^{opt}$	$\hat{A}_1^{opt}$	$\hat{A}_2^{opt}$
Greece	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Italy	0	0	0	0	0	0	0	0	0	0	0	0	0	0
France	0	0	0	0	0	0	0	0	0	0	0	0	0	0
United Kingdom	0	0	0	0	0	0	0	0	0	0	0	0	0	0
China	0	1	0	1	0	1	0	1	0	1	0	1	0	1
Egypt	0	1	0	1	0	1	0	1	0	0	0	0	0	0
India	0	0	0	0	0	0	1	0	0	0	0	0	0	0
South Korea	1	1	1	1	0	1	1	1	1	0	1	0	1	0
Sweden	1	1	1	1	0	1	0	1	1	0	1	0	1	0
Hong Kong	0	1	0	1	0	1	0	1	0	0	0	0	0	0
USA	1	1	1	1	1	1	0	1	1	0	1	0	1	0
United Arab Emirates	1	0	1	0	1	0	0	0	1	0	1	0	1	0
Canada	0	0	0	0	0	0	0	0	0	0	1	0	1	0
Israel	0	1	0	1	0	1	1	1	0	0	0	0	0	0
Finland	0	1	0	1	0	1	0	1	1	0	1	0	1	0
Germany	0	1	0	1	0	1	0	1	0	0	0	0	1	0

Chapter 6

Summary and Future Work

This thesis explores several statistical methods to make sense of COVID-19 data. We, in particular, use a reinforcement learning algorithm, Q-learning, to develop sensible mitigation and suppression strategies when either health outcome or both health and economic outcomes are of primary importance. We further investigate data-related complications that arise in the implementation of Q-learning. In particular, we explore the performance of Q-learning when covariates are subject to misclassification or measurement error. Our analytical and simulation results demonstrate that ignoring this feature can lead to substantial degrees of bias, and that correction strategies are needed for valid inference.

Chapter 2 analyzed the Kaggle novel coronavirus dataset, dated from January 22, 2020 to March 29, 2020, which includes 3397 infected cases and 83 deaths from 39 countries including those in Europe, Asia, and Africa. This chapter summarizes our timely explorations of epidemiological characteristics of COVID-19 in the early stage of the pandemic. We find that prior to March 29, 2020, the median incubation time of COVID-19 is about 5 days, and older people are more likely to have a longer incubation period. Our text analysis shows that the most dominant symptoms of COVID-19 are fever, cough, and pneumonia. The non-parametric Kaplan-Meier method yields a median recovery time of 20 days for infected patients who are not stratified by any of their characteristics. Our findings further suggest that the recovery time increases as the age increases, and there is no significant gender difference in recovery times.

In Chapter 3, we use the Q-learning method to explore how different COVID-19 preventive policies may be prioritized to lower the *CFR*. Our data analysis suggests that in addition to

addressing traditional risk factors to alleviate the risk of death from COVID-19, policymakers should tailor the strictness of preventive policies to country-specific characteristics (e.g., government effectiveness score, population weighted geometric mean density, and civic and social participation score) and evolving situation (e.g., the observed number of deaths) to leverage the salutary effects of prevention strategies. As a future work, it is interesting to explore the application of other regression-based methods such as G-estimation and dynamic weighted ordinary least squares, and then compare the results to help uncover the underlying truth. Furthermore, the analytical approach employed in our analyses assumes that the SUTVA holds. If preventive measures in a country affect the outcome in the neighboring countries, one may employ network-based methods to estimate optimal policies while accounting for the interdependencies among countries. With the availability of inpatient medical records, Q-learning can be used to investigate optimal dosage or order of COVID-19 treatments.

Chapter 4 is partially motivated by the STAR*D study, which is a multi-level randomized controlled study of human MDD. This study was designed to evaluate the effectiveness of different treatment regimes on MDD, and it had 4 levels. In each level, patients were treated by one or a combination of different treatment options for depression. Receipt of a specific treatment option at levels 2, 3 and 4 is driven by the doctor's recommendation as well as the patient's opinion, and it is reflected by the variable termed *patient's preference* to switch or augment his/her previous treatment option. However, it is difficult to precisely record the true value of this variable because of its dependence on the doctor's experience, the level of the patient's trust in the doctor's recommendation, the effectiveness of the communication between the doctor and the patient, and the patient's own knowledge, etc. This suggests that, while careful designs are helpful in collecting good quality data, measurement error and misclassification are still inevitable and they arise ubiquitously in applications. In this chapter, we examine DTRs with misclassification in covariates. Focusing on the Q-learning procedure, we demonstrate how ignoring misclassification in covariates can impact the determination of optimal decision rules, and demonstrate its deleterious effects on Q-learning through empirical studies. Although strategies for handling mismeasurement vary from problem to problem, we present two correction methods for Q-learning based on regression calibration and unbiased estimation equation approaches. Numerical studies reveal that misclassification in covariates

induces non-negligible estimation bias and that the correction methods successfully ameliorate bias in parameter estimation.

Our development in Chapter 4 is directed to the Q-learning approach, which is mainly driven by its widespread applications under DTRs. Other methods such as G-estimation and A-learning are also useful for DTRs. It is interesting to develop de-biasing schemes for those methods to handle error-prone data. While the proposed correction strategies are directed to correct misclassification effects for binary covariates, the estimating equation approach developed in Section 4.4 can be generalized to account for the measurement error effects induced from continuous covariates, where moment identities, as considered by Yi and Lawless (2007), can be used to construct $U(X_{ki}^*)$ to satisfy (4.20). In contrast to covariate mismeasurement considered here, we may face data with measurement error in responses, and it is useful to study problems of mismeasured responses under DTRs.

Another interesting extension to the current development is to incorporate possible drop-out of the study subjects, which may potentially introduce bias and affect the validity of the estimated DTRs. When the drop-out happens *completely at random*, applying the developed methods to the observed data can still output reasonable results, because those data can still be regarded as forming a random sample. However, if the drop-out is *not missing at random*, then the missing measurements cannot be generally ignored due to their inherent association with the outcome process. It is useful to develop valid Q-learning to account for the drop-out effects in such scenarios.

Chapter 5 delves into the examination of mixed misclassification and measurement error in covariates within the framework of DTRs that involve multiple outcome variables. Our analyses reveal that the presence of mixed misclassification and measurement error in covariates can introduce substantial bias, emphasizing the necessity for developing effective correction methods. To mitigate the bias caused by covariate mismeasurement, we propose two correction strategies. It is interesting to apply the development in Chapter 5 to handle more real-world problems which concern two competing outcomes, where it is useful to provide a guideline for properly choosing a value of δ in the formulation of (5.4).

Finally, it is useful to develop software packages, such as *R* functions, to implement the methods developed in Chapters 4 and 5.

Bibliography

Akazawa, K., Kinukawa, N., and Nakamura, T. (1998). A note on the corrected score function adjusting for misclassification. *Journal of the Japan Statistical Society*, 28(1):115–123.

Avriel, M. (2003). *Nonlinear Programming: Analysis and Methods*. Courier Corporation.

Balakrishnan, N., Chimitova, E., Galanova, N., and Vedernikova, M. (2013). Testing goodness of fit of parametric aft and ph models with residuals. *Communications in Statistics-Simulation and Computation*, 42(6):1352–1367.

Bang, H., Chiu, Y.-L., Kaufman, J. S., Patel, M. D., Heiss, G., and Rose, K. M. (2013). Bias correction methods for misclassified covariates in the cox model: Comparison of five correction methods by simulation and data analysis. *Journal of Statistical Theory and Practice*, 7(2):381–400.

Bellman, R. (2010). *Dynamic Programming*. Princeton: Princeton University Press.

Boyd, S. P. and Vandenberghe, L. (2004). *Convex Optimization*. Cambridge university press.

Carroll, R. J., Ruppert, D., Stefanski, L. A., and Crainiceanu, C. M. (2006). *Measurement Error in Nonlinear Models: A Modern Perspective*. CRC press, 2nd edition.

CDC (2022). Ending Isolation and Precautions for People with COVID-19: Interim Guidance. https://www.cdc.gov/coronavirus/2019-ncov/hcp/duration-isolation.html. Accessed: 2023-02-04.

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. and Moodie, E. E. (2013). *Statistical Methods for Dynamic Treatment Regimes*. Springer.

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

Chaudhry, R., Dranitsaris, G., Mubashir, T., Bartoszko, J., and Riazi, S. (2020). A country level analysis measuring the impact of government actions, country preparedness and socioeconomic factors on COVID-19 mortality and related health outcomes. *EClinicalMedicine*, 25:100464. https://doi.org/10.1016/j.eclinm.2020.100464.

Chen, T., Wu, D., Chen, H., Yan, W., Yang, D., Chen, G., Ma, K., Xu, D., Yu, H., Wang, H., Wang, T., Guo, W., Chen, J., Ding, C., Zhang, X., Huang, J., Han, M., Li, S., Luo, X., Zhao, J., and Ning, Q. (2020). Clinical characteristics of 113 deceased patients with coronavirus disease 2019: Retrospective study. *BMJ*, 368:m1091. https://doi.org/10.1136/bmj.m1091.

Cheng, Y., Luo, R., Wang, K., Zhang, M., Wang, Z., Dong, L., Li, J., Yao, Y., Ge, S., and Xu, G. (2020). Kidney disease is associated with in-hospital death of patients with COVID-19. *Kidney International*, 97(5):829–838.

Chiou, S. H., Kang, S., Kim, J., and Yan, J. (2014). Marginal semiparametric multivariate accelerated failure time model with generalized estimating equations. *Lifetime Data Analysis*, 20(4):599–618.

Chu, D. K., Akl, E. A., Duda, S., Solo, K., Yaacoub, S., Schünemann, H. J., El-harakeh, A., Bognanni, A., Lotfi, T., Loeb, M., and SURGE, S. A. (2020). Physical distancing, face masks, and eye protection to prevent person-to-person transmission of SARS-CoV-2 and COVID-19: A systematic review and meta-analysis. *The Lancet*, 395(10242):1973–1987.

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

De Jong, K. (1988). Learning with genetic algorithms: An overview. *Machine Learning*, 3(2):121–138.

Edwards, R., Bondarenko, M., Tatem, A., and Sorichetta, A. (2021). Unconstrained national population weighted density in 2000, 2005, 2010, 2015 and 2020 (100m resolution). *University of Southampton*. https://doi.org/10.5258/SOTON/WP00701.

Fava, M., Rush, A. J., Trivedi, M. H., Nierenberg, A. A., Thase, M. E., Sackeim, H. A., Quitkin, F. M., Wisniewski, S., Lavori, P. W., Rosenbaum, J. F., and Kupfer, D. J. (2003). Background and rationale for the sequenced treatment alternatives to relieve depression (STAR*D) study. *Psychiatric Clinics*, 26(2):457–494.

Fuller, W. A. (1987). *Measurement Error Models*. New York: John Wiley & Sons.

Godambe, V. P. (1960). An optimum property of regular maximum likelihood estimation. *The Annals of Mathematical Statistics*, 31(4):1208–1211.

Godambe, V. P. (1991). *Estimating Functions*. Oxford University Press.

Greenland, S. (1988). Variance estimation for epidemiologic effect estimates under misclassification. *Statistics in Medicine*, 7(7):745–757.

Guillaume, J. H., Jakeman, J. D., Marsili-Libelli, S., Asher, M., Brunner, P., Croke, B., Hill, M. C., Jakeman, A. J., Keesman, K. J., Razavi, S., and Stigter, J. D. (2019). Introductory overview of identifiability analysis: A guide to evaluating whether you have the right type of data for your modeling purpose. *Environmental Modelling & Software*, 119:418–432.

Guo, W., Li, M., Dong, Y., Zhou, H., Zhang, Z., Tian, C., Qin, R., Wang, H., Shen, Y., Du, K., Zhao, L., Fan, H., Luo, S., and Hu, D. (2020). Diabetes is a risk factor for the progression and prognosis of COVID-19. *Diabetes/Metabolism Research and Reviews*, 36(7):e3319. https://doi.org/10.1002/dmrr.3319.

Hale, T., Angrist, N., Goldszmidt, R., Kira, B., Petherick, A., Phillips, T., Webster, S., Cameron-Blake, E., Hallas, L., Majumdar, S., and Tatlow, H. (2021). A global panel database of pandemic policies (Oxford COVID-19 Government Response Tracker). *Nature Human Behaviour*, 5(4):529–538.

Harrington, D. (2005). Linear rank tests in survival analysis. *Encyclopedia of Biostatistics*, 4:2802–2812.

He, W. and Yi, G. Y. (2020). Parametric and semiparametric estimation methods for survival data under a flexible class of models. *Lifetime Data Analysis*, 26(2):369–388.

He, W., Yi, G. Y., and Zhu, Y. (2020). Estimation of the basic reproduction number, average incubation time, asymptomatic infection rate, and case fatality rate for COVID-19: Meta-analysis and sensitivity analysis. *Journal of Medical Virology*, 92(11):2543–2550.

Higazy, M. (2020). Novel fractional order SIDARTHE mathematical model of COVID-19 pandemic. *Chaos, Solitons & Fractals*, 138:110007. https://doi.org/10.1016/j.chaos.2020.110007.

Horst, R. and Tuy, H. (2013). *Global Optimization: Deterministic Approaches*. Springer Science & Business Media.

Janssens, R., Huys, I., van Overbeeke, E., Whichello, C., Harding, S., Kübler, J., Juhaeri, J., Ciaglia, A., Simoens, S., Stevens, H., Smith, M., Levitan, B., Cleemput, I., Bekker-Grob, E. d., and Veldwijk, J. (2019). Opportunities and challenges for the inclusion of patient preferences in the medical product life cycle: A systematic review. *BMC Medical Informatics and Decision Making*, 19:1–16.

Ji, Y., Ma, Z., Peppelenbosch, M. P., and Pan, Q. (2020). Potential association between COVID-19 mortality and health-care resource availability. *The Lancet Global Health*, 8(4):e480. https://doi.org/10.1016/S2214-109X(20)30068-1.

Jin, J.-M., Bai, P., He, W., Wu, F., Liu, X.-F., Han, D.-M., Liu, S., and Yang, J.-K. (2020). Gender differences in patients with COVID-19: Focus on severity and mortality. *Frontiers in Public Health*, 8:152. https://doi.org/10.3389/fpubh.2020.00152.

Jin, Z., Lin, D., and Ying, Z. (2006). On least-squares regression with censored data. *Biometrika*, 93(1):147–161.

Kano, T., Yasui, K., Mikami, T., Asally, M., and Ishiguro, A. (2021). An agent-based model of the interrelation between the COVID-19 outbreak and economic activities. *Proceedings of the Royal Society A*, 477(2245):20200604. https://doi.org/10.1098/rspa.2020.0604.

Kayano, T. and Nishiura, H. (2020). A comparison of case fatality risk of COVID-19 between Singapore and Japan. *Journal of Clinical Medicine*, 9(10):3326. https://doi.org/10.3390/jcm9103326.

Khadem Charvadeh, Y. and Yi, G. Y. (2020). Data visualization and descriptive analysis for understanding epidemiological characteristics of COVID-19: A case study of a dataset from january 22, 2020 to march 29, 2020. *Journal of Data Science*, 18(3):526–535.

Khadem Charvadeh, Y. and Yi, G. Y. (2023a). Accommodating misclassification effects on optimizing dynamic treatment regimes with Q-learning. Submitted for Publication.

Khadem Charvadeh, Y. and Yi, G. Y. (2023b). Understanding effective virus control policies for COVID-19 with the Q-learning method. *Statistics in Biosciences*. https://doi.org/10.1007/s12561-023-09382-w.

Khadem Charvadeh, Y., Yi, G. Y., Bian, Y., and He, W. (2022). Is 14-days a sensible quarantine length for COVID-19? examinations of some associated issues with a case study of COVID-19 incubation times. *Statistics in Biosciences*, 14(1):175–190.

Klang, E., Kassim, G., Soffer, S., Freeman, R., Levin, M. A., and Reich, D. L. (2020). Severe obesity as an independent risk factor for COVID-19 mortality in hospitalized patients younger than 50. *Obesity*, 28(9):1595–1599.

Kleinbaum, D. G., Kupper, L. L., and Morgenstern, H. (1991). *Epidemiologic Research: Principles and Quantitative Methods*. John Wiley & Sons.

Kramer, M. S., Chalmers, B., Hodnett, E. D., Sevkovskaya, Z., Dzikovich, I., Shapiro, S., Collet, J.-P., Vanilovich, I., Mezen, I., Ducruet, T., Shishko, G., Zubovich, V., Mknuik, D., Gluchanina, E., Dombrovskiy, V., Ustinovitch, A., Kot, T., Bogdanovich, N., Ovchinikova, L., and Helsing, E. (2001). Promotion of Breastfeeding Intervention Trial (PROBIT): A

randomized trial in the Republic of Belarus. *Journal of the American Medical Association*, 285(4):413–420.

Küchenhoff, H., Mwalili, S. M., and Lesaffre, E. (2006). A general method for dealing with misclassification in regression: The misclassification SIMEX. *Biometrics*, 62(1):85–96.

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.

Lai, C. C., Shih, T. P., Ko, W. C., Tang, H. J., and Hsueh, P. R. (2020). Severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) and corona virus disease-2019 (COVID-19): The epidemic and the challenges. *International Journal of Antimicrobial Agents*, 55(3). https://doi.org/10.1016/j.ijantimicag.2020.105924.

Lasaulce, S., Zhang, C., Varma, V., and Morărescu, I. C. (2021). Analysis of the tradeoff between health and economic impacts of the COVID-19 epidemic. *Frontiers in Public Health*, 9:620770. https://doi.org/10.3389/fpubh.2021.620770.

Lauer, S. A., Grantz, K. H., Bi, Q., Jones, F. K., Zheng, Q., Meredith, H. R., Azman, A. S., Reich, N. G., and Lessler, J. (2020). The incubation period of coronavirus disease 2019 (COVID-19) from publicly reported confirmed cases: Estimation and application. *Annals of Internal Medicine*, 172(9):577–582.

Lawless, J. F. (2003). *Statistical Models and Methods for Lifetime Data*. John Wiley & Sons, Inc., Hoboken, New Jersey, 2nd edition.

Lehmann, E. L. and Casella, G. (2006). *Theory of Point Estimation*. Springer Science & Business Media.

Li, R., Pei, S., Chen, B., Song, Y., Zhang, T., Yang, W., and Shaman, J. (2020). Substantial undocumented infection facilitates the rapid dissemination of novel coronavirus (SARS-CoV-2). *Science*, 368(6490):489–493.

Liang, L.-L., Tseng, C.-H., Ho, H. J., and Wu, C.-Y. (2020). COVID-19 mortality is negatively associated with test number and government effectiveness. *Scientific Reports*, 10(1):1–7.

Lin, M.-H., Tsai, J.-F., and Yu, C.-S. (2012). A review of deterministic optimization methods in engineering and management. *Mathematical Problems in Engineering*, 2012. https://doi.org/10.1155/2012/756023.

Liu, D., Du, Y., Khadem Charvadeh, Y., Cui, J., Chen, L.-P., Deng, G., Zhang, Q., Cai, K., He, J., He, W., and 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.

Men, K., Wang, X., Yihao, L., Zhang, G., Hu, J., Gao, Y., and Han, H. (2020). Estimate the incubation period of coronavirus 2019 (COVID-19). *Computers in Biology and Medicine*, 158:106794. https://doi.org/10.1016/j.compbiomed.2023.106794.

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.

Mühlbacher, A. C., Juhnke, C., Beyer, A. R., and Garner, S. (2016). Patient-focused benefit-risk analysis to inform regulatory decisions: The european union perspective. *Value in Health*, 19(6):734–740.

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

Murphy, S. A. (2005a). An experimental design for the development of adaptive treatment strategies. *Statistics in Medicine*, 24(10):1455–1481.

Murphy, S. A. (2005b). A generalization error for Q-learning. *Journal of Machine Learning Research*, 6:1073–1097.

Murphy, S. A., van der Laan, M. J., Robins, J. M., and Group, C. P. P. R. (2001). Marginal mean models for dynamic regimes. *Journal of the American Statistical Association*, 96(456):1410–1423.

Nankervis, J. C. (2005). Computational algorithms for double bootstrap confidence intervals. *Computational Statistics & Data Analysis*, 49(2):461–475.

Nelder, J. A. and Mead, R. (1965). A simplex method for function minimization. *The Computer Journal*, 7(4):308–313.

Okhuese, A. V. (2020). Estimation of the probability of reinfection with COVID-19 by the susceptible-exposed-infectious-removed-undetectable-susceptible model. *JMIR Public Health and Surveillance*, 6(2):e19097. https://doi.org/10.2196/19097.

Our World in Data (2021). Case fatality rate vs. total confirmed COVID-19 deaths. https://ourworldindata.org/grapher/deaths-covid-19-vs-case-fatality-rate? Accessed: 2021-10-13.

Ozili, P. K. and Arun, T. (2023). Spillover of COVID-19: Impact on the global economy. In Akkucuk, U., editor, *Managing Inflation and Supply Chain Disruptions in the Global Economy*, pages 41–61. IGI Global.

Pearson, K. (1900). On the criterion that a given system of deviations from the probable in the case of a correlated system of variables is such that it can be reasonably supposed to have arisen from random sampling. *The London, Edinburgh, and Dublin Philosophical Magazine and Journal of Science*, 50(302):157–175.

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

Pribylová, L. and Hajnova, V. (2020). SEIAR model with asymptomatic cohort and consequences to efficiency of quarantine government measures in COVID-19 epidemic. https://doi.org/10.48550/arXiv.2004.02601.

Qian, M., Nahum-Shani, I., and Murphy, S. A. (2012). Dynamic treatment regimes. In Tang, W. and Tu, X., editors, *Modern Clinical Trial Analysis*, pages 127–148. Springer.

Qin, J., You, C., Lin, Q., Hu, T., Yu, S., and Zhou, X.-H. (2020). Estimation of incubation period distribution of COVID-19 using disease onset forward time: A

novel cross-sectional and forward follow-up study. *Science Advances*, 6(33):eabc1202. https://doi.org/10.1101/2020.03.06.20032417.

Ritchie, H., Mathieu, E., Rodés-Guirao, L., Appel, C., Giattino, C., Ortiz-Ospina, E., Hasell, J., Macdonald, B., Beltekian, D., and Roser, M. (2020). Coronavirus pandemic (COVID-19). *Our World in Data*. https://ourworldindata.org/coronavirus.

Robins, J. M. (1997). Causal inference from complex longitudinal data. In Berkane, M., editor, *Latent Variable Modeling and Applications to Causality*, pages 69–117. Springer.

Robins, J. M. (2004). Optimal structural nested models for optimal sequential decisions. In Lin, D. and Heagerty, P. J., editors, *Proceedings of the Second Seattle Symposium in Biostatistics*, pages 189–326. Springer.

Rothman, K. J. and Greenland, S. (1998). *Modern Epidemiology*. Philadelphia, PA: Lippincott-Raven.

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., Kupfer, D. J., Rosenbaum, J. F., Alpert, J., Stewart, J. W., McGrath, P. J., Biggs, M. M., Shores-Wilson, K., Lebowitz, B. D., Ritz, L., and Niederehe, G. (2004). Sequenced treatment alternatives to relieve depression (STAR*D): Rationale and design. *Controlled Clinical Trials*, 25(1):119–142.

Shanno, D. F. (1970). Conditioning of quasi-newton methods for function minimization. *Mathematics of Computation*, 24(111):647–656.

Shieh, M.-S. (2009). *Correction Methods, Approximate Biases, and Inference for Misclassified Data*. PhD thesis, University of Massachusetts Amherst.

Soekhai, V., Whichello, C., Levitan, B., Veldwijk, J., Pinto, C. A., Donkers, B., Huys, I., van Overbeeke, E., Juhaeri, J., and de Bekker-Grob, E. W. (2019). Methods for exploring and eliciting patient preferences in the medical product lifecycle: A literature review. *Drug discovery today*, 24(7):1324–1331.

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.

Spiegelman, D., Rosner, B., and Logan, R. (2000). Estimation and inference for logistic regression with covariate misclassification and measurement error in main study/validation study designs. *Journal of the American Statistical Association*, 95(449):51–61.

Terriau, A., Albertini, J., Montassier, E., Poirier, A., and Le Bastard, Q. (2021). Estimating the impact of virus testing strategies on the COVID-19 case fatality rate using fixed-effects models. *Scientific Reports*, 11(1):1–8.

The Global Economy (2019). GDP per capita based on purchasing power parity in constant 2011 international dollars. https://www.theglobaleconomy.com/download-data.php. Accessed: 2022-02-15.

The Global Economy (2020). Economic growth, quarterly by country: The latest data. https://www.theglobaleconomy.com/rankings/gdp_growth/. Accessed: 2022-02-26.

The Legatum Institute (2019). The Legatum prosperity index. https://www.prosperity.com/about/resources. Accessed: 2022-02-15.

The World Bank (2019a). Population ages 65 and above (% of total population). https://data.worldbank.org/indicator/SP.POP.65UP.TO.ZS. Accessed: 2021-10-09.

The World Bank (2019b). Worldwide governance indicators. https://info.worldbank.org/governance/wgi/Home/Reports. Accessed: 2021-11-24.

Verity, R., Okell, L. C., Dorigatti, I., Winskill, P., Whittaker, C., Imai, N., Cuomo-Dannenburg, G., Thompson, H., Walker, P. G., Fu, H., Dighe MRes, A., Griffin, J. T., Baguelin, M., Bhatia, S., Boonyasiri, A., Cori, A., Cucunubá, Z., FitzJohn, R., Gaythorpe, K., Green, W., Hamlet, A., Hinsley, W., Laydon, D., Nedjati-Gilani, G., Riley, S., Elsland, S. V., Volz, E., Wang, H., Wang, Y., Xi, X., Donnelly, C. A., Ghani, A. C., and Ferguson, N. M. (2020). Estimates of the severity of coronavirus disease 2019: A model-based analysis. *The Lancet infectious diseases*, 20(6):669–677.

Viégas, F. B. and Wattenberg, M. (2008). Timelines tag clouds and the case for vernacular visualization. *Interactions*, 15(4):49–52.

Wagner, E. H., Austin, B. T., Davis, C., Hindmarsh, M., Schaefer, J., and Bonomi, A. (2001). Improving chronic illness care: Translating evidence into action. *Health Affairs*, 20(6):64–78.

Wagner, M. M., Gresham, L. S., and Dato, V. (2006). Case detection, outbreak detection, and outbreak characterization. *Handbook of Biosurveillance*, pages 27–50. https://doi.org/10.1016/B978-012369378-5/50005-3.

Watkins, C. J. C. H. (1989). *Learning From Delayed Rewards*. PhD thesis, King's College, Cambridge United Kingdom.

Wells, C. R., Sah, P., Moghadas, S. M., Pandey, A., Shoukat, A., Wang, Y., Wang, Z., Meyers, L. A., Singer, B. H., and Galvani, A. P. (2020). Impact of international travel and border control measures on the global spread of the novel 2019 coronavirus outbreak. *Proceedings of the National Academy of Sciences*, 117(13):7504–7509.

Weng, Y. and Yi, G. Y. (2022). Estimation of the covid-19 mean incubation time: Systematic review, meta-analysis, and sensitivity analysis. *Journal of Medical Virology*, 94(9):4156–4169.

WHO (2022). Who coronavirus (COVID-19) dashboard. https://covid19.who.int. Accessed: 2022-11-17.

Worldometers (2022). Reported cases and deaths by country, territory, or conveyance. https://www.worldometers.info/coronavirus/countries-where-coronavirus-has-spread/. Accessed: 2022-11-17.

Yi, G. Y. (2017). *Statistical Analysis with Measurement Error or Misclassification: Strategy, Method and Application*. Springer Science+Business Media LLC, New York.

Yi, G. Y., Delaigle, A., and Gustafson, P. (2021). *Handbook of Measurement Error Models*. CRC Press.

Yi, G. Y. and Lawless, J. F. (2007). A corrected likelihood method for the proportional hazards model with covariates subject to measurement error. *Journal of Statistical Planning and Inference*, 137(6):1816–1828.

Yi, G. Y., Ma, Y., Spiegelman, D., and Carroll, R. J. (2015). 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., Yan, Y., Liao, X., and Spiegelman, D. (2018). Parametric regression analysis with covariate misclassification in main study/validation study designs. *The International Journal of Biostatistics*, 15(1). https://doi.org/10.1515/ijb-2017-0002.

Zhang, Q. and Yi, G. Y. (2023). Generalized network structured models with mixed responses subject to measurement error and misclassification. *Biometrics*, 79(2):1073–1088.

Zhang, T., Wu, Q., and Zhang, Z. (2020). Probable pangolin origin of SARS-CoV-2 associated with the COVID-19 outbreak. *Current Biology*, 30(7):1346–1351.

Zhao, Y., Kosorok, M. R., and Zeng, D. (2009). Reinforcement learning design for cancer clinical trials. *Statistics in Medicine*, 28(26):3294–3315.

Zhou, F., Yu, T., Du, R., Fan, G., Liu, Y., Liu, Z., Xiang, J., Wang, Y., Song, B., Gu, X., Guan, L., Wei, Y., Li, H., Wu, X., Xu, J., Tu, S., Zhang, Y., Chen, H., and Cao, B. (2020). Clinical course and risk factors for mortality of adult inpatients with COVID-19 in Wuhan, China: A retrospective cohort study. *The Lancet*, 395(10229):1054–1062.

Zucker, D. M. and Spiegelman, D. (2004). Inference for the proportional hazards model with misclassified discrete-valued covariates. *Biometrics*, 60(2):324–334.

Curriculum Vitae

Name:	Yasin Khadem Charvadeh
Post-Secondary Education	University of Western Ontario, *London, ON, Canada* Ph.D. Statistics, 2019 - 2023
	Memorial University of Newfoundland, *St. John's, Newfoundland, Canada* M.Sc. Statistics, 2017 - 2019
	Allameh Tabataba'i University, *Tehran, Iran* M.Sc. Socioeconomic Statistics, 2014 - 2017
	University of Mazandaran, *Mazandaran, Iran* B.Sc. Statistics, 2010 - 2014
Honours and Awards:	Fallona Family Interdisciplinary Science Award & Lecture, Apr. 18, 2022 Member of SARGC, 2021 - 2023 Statistical Society of Canada Student Travel Award, 2019
Related Work Experience:	Western University, 2019 - Present Memorial University of Newfoundland, 2017 - 2019 Allameh Tabataba'i University, 2015 Teaching Assistantship and Research Assistantship

Publications:

1. **Khadem Charvadeh, Y.** and Yi, G. Y. (2023). Accommodating Misclassification Effects on Optimizing Dynamic Treatment Regimes with Q-Learning. Submitted for Publication.

2. **Khadem Charvadeh, Y.** and Yi, G. Y. (2023). Understanding Effective Virus Control

Policies for COVID-19 with the Q-Learning Method. Published online by *Statistics in Biosciences*. https://doi.org/10.1007/s12561-023-09382-w

3. **Khadem Charvadeh, Y.**, Yi, G. Y., Bian, Y. and He, W. (2022). Is 14-Days a Sensible Quarantine Length for COVID-19? Examinations of Some Associated Issues with a Case Study of COVID-19 Incubation Times. *Statistics in Biosciences*, 14(1), 175-190.

4. **Khadem Charvadeh, Y.** and Yi, G. Y. (2020). Data Visualization and Descriptive Analysis for Understanding Epidemiological Characteristics of COVID-19: A Case Study of a Dataset from January 22, 2020 to March 29, 2020. *Journal of Data Science*, 18(3), 526–535.

5. Liu, D., Du, Y., **Khadem Charvadeh, Y.**, Cui, J., Chen, L. P., Deng, G., Zhang, Q., Cai, K., He, J., He, W. and 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). DOI: 10.5539/ijsp.v9n5p23

6. Pasha, M. A., Moghadam, M. B., Fani, S. and **Khadem Charvadeh, Y.** (2018). Effects of quality characteristic distributions on the integrated model of Taguchi's loss function and economic statistical design of X-bar control charts by modifying the Banerjee and Rahim economic model. *Communications in Statistics-Theory and Methods*, 47(8), 1842-1855.

7. **Khadem Charvadeh, Y.** and Moghadam, M. B. (2017). Economic statistical design of-control charts: Modified version of Rahim and Banerjee (1993) model. *Communications in Statistics-Simulation and Computation*, 48(3), 684-703.

8. Pasha, M. A., **Khadem Charvadeh, Y.** and Moghadam, M. B. (2017). Application of decreasing integrated hazard in Rahim and Banerjee model on economic design of X-bar control charts for systems with increasing failure rate and early replacement. *Communications in Statistics-Theory and Methods*, 46(21), 10481-10494.

9. Moghadam, M. B., **Khadem Charvadeh, Y.**, Fani, S. and Pasha, M. A. (2017). Effects of non-normality on economic and economic statistical designs of X-bar control

charts with multiple assignable causes and Weibull in-control times. *Communications in Statistics-Simulation and Computation*, 47(7), 2055-2069.

10. Pasha, M. A., Moghadam, M. B., **Khadem Charvadeh, Y.** and Fani, S. (2017). An integration of Taguchi's loss function in Banerjee–Rahim model for the economic and economic statistical design of X-bar control charts under multiple assignable causes and Weibull shock model. *Communications in Statistics-Theory and Methods*, 46(24), 12113-12129.

Scientific Meetings, Graduate Departmental Seminars and Papers in Conference Proceedings:

1. Dissertation Proposal Presentation, 2022.

    - **Title:** Optimizing Dynamic Treatment Regimes with Q-Learning: Complications due to Error-Prone Data and Applications to COVID-19 Data.

2. Lighting Talks During the Graduate Research Symposium in Honour of Dr. Raymond Carroll, University of Western Ontario, June 17, 2022.

    - **Presentation Title:** Balancing Health Outcomes against Economic Costs Associated with COVID-19 Preventive Measures.

3. NSERC Emerging Infectious Diseases Modelling (EIDM) Initiative Meeting, Dec. 16, 2021.

    - **Presentation Title:** Understanding Effective Virus Control Policies for COVID-19 with the Q-Learning Method.

4. Graduate Departmental Presentations, University of Western Ontario, 2020-2022.

    - Grace-Wenqing Data Science Research Group (GW-DSRG) Meetings.
        - **Presentation Title:** "Geometric Problems" in *Convex Optimization* by Boyd, S. and Vandenberghe, L. (2004).

- **Presentation Title:** "Q-learning for Estimating Optimal Dynamic Treatment Rules from Observational Data" by Erica E. M. Moodie, Bibhas Chakraborty and Michael S. Kramer (2012).
- **Presentation Title:** Chapters 1-2 of *Statistical Learning with Sparsity: The LASSO and Generalizations* by Travor Hastie, Rob Tibshirani and Martin Wainwright (2015).
- **Presentation Title:** "Matched Learning for Optimizing Individualized Treatment Strategies Using Electronic Health Records" by Peng Wu, Donglin Zeng and Yuanjia Wang (2020).
- **Presentation Title:** "Doubly Robust Estimation of Optimal Dosing Strategies" by Juliana Schulz and Erica E. M. Moodie (2021).
- **Presentation Title:** "Inference for Optimal Treatment Regimes: Model-Assisted Uniformly Honest Inference for Optimal Treatment Regimes in High Dimension" by Yunan Wu, Wang Lan, and Fu Haoda (2021).
- **Presentation Title:** "Reinforcement Learning" in *Foundations of Machine Learning* by Mohri, M., Rostamizadeh, A., and Talwalkar, A. (2018).

5. Graduate Departmental Seminar, Memorial University of Newfoundland, 2019.

   - **Seminar Title:** Propensity Score Matching Methods for Recurrent Events.

6. Comparing the Effectiveness of Virus Control Policies for COVID-19 with the Q-Learning Method. Statistical Society of Canada 2022 Annual Meeting. https://ssc.ca/sites/default/files/imce/ssc2022-program-with-abstracts.pdf

7. Case Study 1: What geographical factors are associated with pipeline incidents that involve spills? Statistical Society of Canada 2021 Annual Meeting. https://ssc.ca/sites/default/files/imce/ssc2021-program-with-abstracts-final.pdf

8. Investigating Misclassification Effects on Estimating Optimal Dynamic Treatment Regimes Using Q-Learning. Statistical Society of Canada 2021 Annual Meeting.

https://ssc.ca/sites/default/files/imce/ssc2021-program-with-abstracts-final.pdf

9. The Use of Propensity Score Matching Methods for Estimating Treatment Effects in Recurrent Events. Statistical Society of Canada 2019 Annual Meeting. https://ssc.ca/sites/default/files/meetings/ssc_2019_program_with_abstracts_0.pdf

10. **Khadem Charvadeh, Y.**, Pasha, M. A., Moghadam, M. B. and Fani, S. Economic and Economic Statistical Design of X-bar Control Charts Under a Bathtub-shaped Shock Model. *Second Seminar on Reliability Theory and its Applications, Department of Statistics, School of Mathematics, Statistics and Computer Science, University of Tehran, Tehran, Iran*, 18-19 May, 2016. http://www.sid.ir/en/seminar/ViewPaper.aspx?FID=543e20160202

11. Pasha, M. A., Yousefi, A. and **Khadem Charvadeh, Y.** Nonlinear Fractional Programming in Optimal Design of X-bar Control Charts Under a Heavy-tailed Shock Model. *9th International Iranian Operation Research conference, Shiraz University of Technology, Iran*, 2016.

Personal Projects:

- Investigating the Association Between COVID-19 Cases/Deaths and Fine Particulate Matter Exposure During the 2020 Wildfires in the United States
  - In this R project, I use a publicly available data set to explore the potential adverse effects of wildfires on COVID-19 cases and fatalities across 92 counties in the western region of the United States. Given the excessive zero values in daily COVID-19 cases/deaths, a Hurdle mixed-effects model is employed to establish a cause-and-effect relationship. The numerical results suggest a significant statistical association between short-term exposure to Particulate Matter < 2.5 ($PM_{2.5}$) and COVID-19 cases/deaths.

- Deep Learning for Time Series: Predicting Hourly Electricity Demand in Ontario

  - The investigation of historical events is imperative for making informed strategic decisions in the future. Time series analysis is a valuable tool for organizations to effectively prepare for upcoming challenges. The Canada Energy Regulator (CER) agency is accountable for ensuring that energy production aligns with energy requirements. To achieve this aim, a robust model of hourly demand is indispensable. In this R project, publicly accessible hourly electricity demand and annual electricity demand information for Ontario are used to forecast hourly electricity demand in the residential sector. Gated Recurrent Unit Networks are employed to address the challenge of hourly electricity demand prediction.

- Predicting Continuous Outcome with Image and Text Using the Keras Functional API

  - In this R project, I leverage the Keras functional API to construct a model that predicts the average ratings of horror movies based on both their descriptions and posters. The data set used is sourced from the TidyTuesday Github repository and comprises numerous variables that depict diverse aspects of horror movies.

- Screening Multiple Models for Predicting Volcanic Explosivity Index (VEI)

  - In this R project, I use a data set on volcano eruptions to screen a number of models with the aim of finding the model with the best predictive capabilities.

- Sentiment and User Network Analysis Using YouTube Comments

  - In this R project, I use data collected from two YouTube videos to perform sentiment analysis and network analysis.

- Image Classification Using Keras in R: Classification of COVID-19 positive/negative cases

  - In this R project, I use publicly available chest X-ray images from a Kaggle repository to build binary/multi-class image classification models that help identify COVID-19 positive cases.

- Predicting Popularity of Spotify Songs

  - In this R project, I use a TidyTuesday data set on Spotify songs to build a machine learning classification model that predicts the popularity of a given Spotify song. Additionally, I showcase the application of partial dependence plots, individual conditional expectation (ICE), and the *LIME* package to facilitate both global and local interpretations of the machine learning model.

- Predicting Coffee Ratings with Beta Regression and Neural Networks

  - This R project centers around constructing a Beta regression model and Neural Network model to predict coffee ratings, utilizing the Coffee Quality Database data set available on the TidyTuesday Github page. As the data set comprises variables with missing values, the project also showcases the implementation of the tidymodels approach to perform missing-data imputation.

- Text Analysis with tidymodels: Predicting Number of Seasons for Netflix shows

  - In this R project, I use a TidyTuesday data set containing information about TV shows available on Netflix to develop a penalized logistic regression model that predicts whether a given TV show will have only one season or more. To build the model, I utilize a range of predictors such as the summary description of the TV show, title of the TV show, genre, and several others. The main objective of this project is to showcase how we can conduct rudimentary text analysis.

- Predicting Bee Colony Losses

  - In this R project, I analyze a TidyTuesday data set on honey bee colonies to examine the factors that contribute to bee colony losses in the United States. The primary objective of this project is to optimize the hyperparameters of an XGBoost model, and subsequently leverage the most optimal XGBoost model to predict the rate of colony losses.

- Interactive COVID-19 dashboard

  - This dynamic web app is created using Flexdashboard + Shiny in R, and automatically updates every 12 hours. Dashboard address: https://yasinkhc.shinyapps

.io/Shiny-COVID-19/ (Please note that it may take up to 80 seconds for the program to generate the map)

Statistical Practice:

1. I am a core member for developing the Canada COVID-19 website. Website address: https://covid-19-canada.uwo.ca/en/about.html

2. I led the SSC Biostatistics COVID-19 case study. The motive for organizing this case study was to allow students to demonstrate initiatives and creativity in contributing to the ongoing needs for data-driven decision making. The work by my team was selected to be featured on the SSC website.

3. Team leader of SSC 2021 case study in data analysis. Case study title: What geographical factors are associated with pipeline incidents that involve spills?

4. Team leader of SSC 2020 case study in data analysis (the meeting was canceled due to the COVID-19 pandemic).

Training and Workshops Attended

- Attended the workshop on "Scientific Analysis of Networks Using Graph Neural Networks" by Prof. María Óskarsdóttir, University of Western Ontario, May 9th to May 11th, 2023.