ABSTRACT

DONG, LIN. Semiparametric Methods for Decision Making and Causal Effect Generalization. (Under the direction of Dr. Eric Laber and Dr. Shu Yang).

Semiparametric method is a broad term that encompasses many estimation and modeling approaches. In the simplest terms, all such approaches assume that some features of the data can be constructed parametrically, and the others can be modeled nonparametrically. We study its usage in the context of dynamic treatment regimes, causal effect generalization, and predictive modeling. Chapter 1 gives an overview of the research work. Chapter 2 is dedicated to the use of semiparametric method in dynamic treatment regimes, where we present an estimating equation framework to tackle the missing data issue in estimating optimal treatment regimes. We propose an augmented weighting estimating equation approach that applies to a broad class of estimators including both Q-learning and outcome weighted learning. We establish consistency and demonstrate its advantages in finite samples using simulation experiments and application to a schizophrenia study. In Chapter 3, we develop a new semiparametric framework to evaluate the average treatment effect of a target population integrating the complementary features of the randomized clinical trial and real world evidence studies. We establish related asymptotic results and derive the semiparametric efficiency bound under this setting. We propose an augmented calibration weighting estimator that achieves such bound when the nuisance models are correctly specified. We apply our proposed methods to estimate the effect of adjuvant chemotherapy in early-stage non-small-cell lung cancer where we utilize the data from trial CALGB 9633 and the National Cancer Database. In Chapter 4, we focus on the use of semiparametric methods in predictive modeling. We introduce a new class of neural network architecture to represent some of the widely-used semiparametric regression models, including additive models, single index and additive index models. A hypothesis testing based model ranking is proposed and validated by simulation experiments. These developments help fill the methodology gaps that remain in the emerging fields of precision medicine, data integration and supervised learning.

Semiparametric Methods for Decision Making and Causal Effect Generalization

by
Lin Dong

A dissertation submitted to the Graduate Faculty of
North Carolina State University
in partial fulfillment of the
requirements for the Degree of
Doctor of Philosophy

Statistics

Raleigh, North Carolina

2019

APPROVED BY:

Dr. Eric Laber Co-chair of Advisory Committee	Dr. Shu Yang Co-chair of Advisory Committee
Dr. Rui Song	Dr. Yair Goldberg External Member

Dr. Jamian (Krishna) Pacifici

DEDICATION

To my family and friends.

BIOGRAPHY

Lin Dong (Dǒng Lín in Chinese pinyi) was born and grew up in Beijing, China. She obtained her B.S. in Statistics and Operations Research from Hong Kong Baptist University in May, 2014. She then joined North Carolina State University to pursue her Ph.D. in Statistics. Lin will defend her dissertation in June 2019.

ACKNOWLEDGEMENTS

I would like to express my deepest appreciation to my advisors. I am deeply indebted to Dr. Eric Laber for his continuous support and guidance. He is a great advisor and leader, who has been providing me and Laber-Labs with valuable resources. I thank him for giving me the chances to tackle various tasks and challenges, which greatly broaden my horizon. I would like to extend my deepest gratitude to Dr. Shu Yang for her kindly guidance and patience that cannot be underestimated. She is a great scholar and is always prompt on answering my questions. The completion of my dissertation would not have been possible without the support of them.

I would like to extend my sincere thanks to my committee members, Dr. Rui Song, Dr. Yair Goldberg and Dr. Jamian Pacific for their advises and time devoted to serve in the committee. I must also thank Dr. Vijay Nair and Dr. Joel Vaughan at Wells Fargo for their constructive advises in the structured neural network project. I would like to thank Dr. Shannon Holloway for helping me polish the writing of my dissertation. I also wish to thank Dr. Sujit Ghosh for his guidance in a biosimilar project.

The department of statistics at North Carolina State University is a great place to conduct collaborative research. Many thanks also to the entire faculty and staffs in the department for offering such a friendly environment.

I owe my long overdue thanks to my professors in HKBU. My gratitude goes to Dr. Lixing Zhu, Dr. Man-Lai Tang and Dr. Xiaonan Wu for encouraging me to pursue a doctoral degree and their recommendations.

I want to express my sincere appreciation to my Laber-Labs lab-mates for the valuable discussions we had over the years. Special thanks to the friends I met during my graduate study and internships. We had a lot of good-times together.

Last but not the least, I thank my parents for their everlasting love and support and I thank Zhentao for his constant companion.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

CHAPTER 1

INTRODUCTION

1.1 Semiparametric methods

Semiparametric method is a broad term that encompasses many estimation and modeling approaches. In the simplest terms, all such approaches assume that some features of the data can be constructed parametrically (with nite-dimensional parameters), and the others can be modeled nonparametrically (with in nite-dimensional parameters). Over decades of developments, semiparametric methods and theory have become a classic branch of statistics and been applied to a wide range of scienti c questions, such as missing data and causal inference. The methods popularity is due in part to the inherent balance struck in the bias-variance trade-o . Speci cally, semiparametric methods fall between fully parametric methods, which are subject to misspeci cation (bias), and fully nonparametric methods, which lose e ciency (variance). Though semiparametric theory is well-established, several important methodology developments remain. Particularly, we study the the application of semiparametric theory in the context of estimating optimal treatment regimes with missing data, causal e ect generalization, and predictive modeling. These areas are of importance in the emerging elds of precision medicine,

data integration and machine learning and are the focus of this dissertation.

1.2 Overview of research work

A summary of the chapters that follow is given here as an overview of the research work included in this dissertation.

Chapter 2 is dedicated to semiparametric methods in the context of missing data. We explore this scenario by application to the subject area of dynamic treatment regimes (DTRs), an area of growing importance in the age of precision medicine. Generally speaking, dynamic treatment regimes operationalize precision medicine as a sequence of decision rules, one per stage of clinical intervention, that map up-to-date patient information to a recommended intervention. Of primary importance in this eld is the identi cation of an "optimal" treatment regime, i.e., one that maximizes a mean utility function when applied to the population of interest. Current methods for estimating an optimal treatment regime assume that the data is fully observed, which rarely occurs in practice. A common approach to overcome this shortcoming of available data is to use multiple imputation and pool estimators across imputed (complete) datasets. However, this approach requires estimating the joint distribution of patient trajectories, which can be high-dimensional, especially when there are multiple stages of intervention. To avoid the underlying modeling step required to obtain complete data, we propose a weighted estimating equation based approach to estimate optimal treatment regimes. Our approach applies to a broad class of estimators including Q-learning and a generalization of outcome weighted learning, which are among the most popular estimators of an optimal treatment regime. In addition, we establish consistency under mild regularity conditions and demonstrate the advantages of our proposed methods in nite samples using simulation experiments and application to a schizophrenia study.

Randomized clinical trials (RCTs) are designed to and are regarded as the gold standard to evaluate treatment e ects; however, due to their restricted inclusion and exclusion criteria, the ndings from RCTs often lack generalizability to some real-world population of interest, which

we refer to as the target population. Another data source, called real world evidence (RWE) studies, often include large samples that are representative of a target population, but such studies are often observational and subject to complex confounding. In Chapter 3, we leverage the complementing features RCT and RWE to estimate the average treatment e ect of the target population. First, we propose a calibration weighting estimator that uses only covariate information from the RWE study. Because this estimator enforces the covariate balance between the RCT and RWE study, the generalizability of the trial-based estimator is improved. We further propose an augmented calibration weighting estimator that can be applied in the event that treatment and outcome information is also available from the RWE study. This estimator achieves a semiparametric e ciency bound that we derived under the identi cation and outcome mean function transportability assumptions when the nuisance models are correctly speci ed. To resolve the misspeci cation issue associated with parameteric approaches, a data-adaptive nonparametric sieve method is provided as an alternative. The sieve method guarantees good approximation of the nuisance models. We establish related asymptotic results under mild regularity conditions. Finite sample performances of the proposed estimators are veri ed in simulation studies. Finally, we apply our proposed methods to estimate the e ect of adjuvant chemotherapy in early-stage resected non-small-cell lung cancer, where we utilize the data from trial CALGB 9633 and a sample from the National Cancer Database.

Chapter 4 is inspired by the use of semiparametric regression models in predictive modeling. We switch gears away from DTRs and causal inference to a statistical learning setting with two competing objectives, predictive performance and model explainability. Arti cial neural networks have attracted much attention because of their high predictive power and computational scalability to large datasets. However, a key hindrance of using neural networks in practice is their lack of interpretability. We propose a new class of neural networks called the structured neural networks (sNNs). In the sNN class, we design the architecture of the feedforward neural network to represent the model structures of some commonly used semiparametric regression

models, including but not limited to the generalized additive model, the single index model, and the additive index model. The proposed sNNs leverage semiparametric regression models and neural networks to achieve a better balance between model explainability and performance. A hypothesis testing based model ranking procedure, which can be used to aid the model selection in (but not limited to) this class, is proposed and validated by simulation experiments.

CHAPTER 2

DYNAMIC TREATMENT REGIMES WITH MISSING DATA

2.1 Introduction

Dynamic treatment regimes operationalize clinical decision making as a sequence of decision rules, one per stage of intervention, that map current patient information to a recommended intervention (Murphy, 2003; Robins, 2004).An optimal treatment regime maximizes the mean utility if applied to select interventions in the patient population of interest (for alternative de nitions of optimality, see Kosorok & Moodie, 2015; Linn et al., 2017; Wang et al., 2018). Optimal treatment regimes have been estimated across a wide range of application areas including anticoagulation (Henderson et al., 2010; Barrett et al., 2014; Rich et al., 2014), cancer (Wang et al., 2012b; Xu et al., 2019), mental disorders (Nahum-Shani et al., 2012; Laber et al., 2014a; Zhang et al., 2017), and HIV (Laan & Petersen, 2007; Petersen et al., 2012; Young et al., 2011). In these and nearly all other biomedical application areas, the observed data are subject to missingness, which can include missing measurements, treatments, and outcomes (Shortreed et al., 2014; Kosorok & Moodie, 2015).

There is a large body of literature on estimation of optimal treatment regimes using complete

data. This body of research includes: approximate dynamic programming methods like Q- and A-learning (Murphy, 2003; Robins, 2004; Blatt et al., 2004; Murphy, 2005a; Moodie et al., 2007; Schulte et al., 2014) and its many variants (e.g., Zhao et al., 2009; Goldberg & Kosorok, 2012; Lu et al., 2013; Moodie et al., 2014; Tian et al., 2014; Laber et al., 2014b; Zhou & Kosorok, 2017; Jeng et al., 2018; Shi et al., 2018; Kosorok & Laber, 2019, and references therein) ; direct-search methods including outcome weighted learning (Orellana et al., 2010; Zhang et al., 2012a; Zhao et al., 2012; Zhang et al., 2013; Zhao et al., 2014; Zhao et al., 2015; Zhou et al., 2017; Athey & Wager, 2017; Zhang et al., 2017; Zhang & Zhang, 2018; Liu et al., 2018; Luckett et al., 2018); and model-based planning via g-computation (see Robins, 1997; Yu & Laan, 2002; Xu et al., 2016; Xu et al., 2019; Guan et al., 2018; Laber et al., 2018, and references therein). Because these methods require complete data, it is often necessary to employ methods to address missing data. A common approach is to apply multiple imputation to complete the data, compute a given estimator of an optimal regime on each of the imputed data sets, and then aggregate these estimators. e.g., by averaging or voting (Almirall et al., 2016; Lu et al., 2016; Ertefaie et al., 2016; Nahum-Shani et al., 2017; Kilbourne et al., 2018; Kidwell et al., 2018). Though estimation of an optimal treatment regime is often but one part of a suite of secondary analyses, the requirement to develop a complete dataset is convenient as it can be used for a variety of other analyses.

Despite its appealing features, multiple imputation can be problematic with complex longitudinal data arising in the context of sequential decision problems because the data can be high-dimensional and subjects are often missing large segments of data. Further, multiple imputation requires estimating the joint distribution of patient trajectories and constructing a high-quality imputation model in this context is di cult. If a misspeci ed model is used to impute large amounts of missing data, the estimators may be biased and inferences inaccurate (Seaman et al., 2012). We examine a class of augmented inverse probability weighted estimators of the optimal treatment regime that applies to approximate dynamic programming and to direct-search methods when the data have a monotone missingness pattern. The application of

standard arguments of semiparametric e ciency theory establishes a double robustness property for this class of estimators (e.g., Tsiatis, 2007). We show that augmented weighting performs favorably as compared to multiple imputation and to simple inverse probability weighting in simulation examples. These results suggest that investigators should give serious consideration to using weighting methods as an alternative to multiple imputation in the context of estimating optimal treatment regimes in practice.

The remainder of this chapter is organized as follows. In Section 2.2, we set notation and de ne an optimal treatment regime using potential outcomes. In Section 2.3, we describe a class of estimators of an optimal treatment regimes for use with complete data; this class includes Q-learning and outcome weighted learning as special cases. In Section 2.4, we derive an augmented inverse probability weighted estimator for the proposed class of estimators that applies under a monotone missingness pattern. In Section 2.5, we present simulation examples and an application to data from a sequential multiple assignment randomized trial on schizophrenia. A brief discussion of future work is given in Section 2.6.

2.2 Setup and notation

We consider longitudinal data arising from an observational study or a sequential multiple assignment randomized trial (SMART, Lavori & Dawson, 2004; Murphy, 2005b; Kidwell, 2014). The complete data are assumed to be of the form $\{X_{1i}, A_{1i}, X_{2i}, A_{2i}, \ldots, X_{Ti}, A_{Ti}, Y_i\}_{i=1}^{n}$, which comprise n independent replicates of $(X_1, A_1, X_2, A_2, \ldots, X_T, A_T, Y)$; where: T is the number of treatment stages, $X_1 \in \mathbb{R}^{p_1}$ is baseline patient information and $X_t \in \mathbb{R}^{p_t}$ is information collected during stage $(t-1)$ for $t = 2, \ldots, T$, $A_t \in \mathcal{A}_t = \{-1, 1\}$ is the treatment assigned during stage $t = 1, \ldots, T$, and $Y \in \mathcal{Y} = \mathbb{R}$ is the terminal outcome coded so that higher values are better. The restriction to binary treatments is not necessary for approximate dynamic programming methods; however, most variants of outcome weighted learning require binary treatments (exceptions include Laber & Zhao, 2015; Chen et al., 2016) so we impose this

restriction to allow for a simple and uni ed notation.

De ne $H_1 = X_1$, and recursively de ne $H_t = (H_{t-1}; A_{t-1}; X_t)$ for $t = 2; \ldots; T$. Thus, H_t is the information available to the decision maker at time $t = 1; \ldots; T$. Let $\mathcal{H}_t$ denote the support of H_t, and for each $h_t \in \mathcal{H}_t$, de ne $\mathcal{A}_t(h_t) \subseteq \mathcal{A}_t$ to be the set of allowable treatments for a patient presenting with history $H_t = h_t$ at time t. A treatment regime in this context is a sequence of maps, $\pi = (\pi_1; \ldots; \pi_T)$, with $\pi_t : \mathcal{H}_t \to \mathcal{A}_t$ and $\pi_t(h_t) \in \mathcal{A}_t(h_t)$ for all $h_t \in \mathcal{H}_t$ and $t = 1; \ldots; T$. Under π, a patient with history $H_t = h_t$ at time t would be recommended to receive treatment $\pi_t(h_t)$. Let Π denote the space of all feasible regimes. An optimal treatment regime, say $\pi^{opt} \in \Pi$, maximizes the mean outcome if applied to the population of interest. We formalize this de nition using potential outcomes (Rubin, 1978; Neyman, 1923).

For each $t = 1; \ldots; T$, de ne $\bar{a}_t = (a_1; \ldots; a_t)$. Let $H_t(\bar{a}_{t-1})$ denote the potential history under treatment sequence $\bar{a}_{t-1}$, and let $Y(\bar{a}_T)$ denote the potential outcome under treatment sequence $\bar{a}_T$. Therefore, the set of all potential outcomes is

$$W = \{H_t(\bar{a}_{t-1}); Y(\bar{a}_T) : a_t \in \mathcal{A}_t\{H_t(\bar{a}_{t-1})\}; t = 1; \ldots; T\};$$

where we have de ned $H_1(\bar{a}_0) \equiv H_1$. Let $1_{(\cdot)}$ be the indicator function. The potential outcome under a regime $\pi \in \Pi$ is

$$Y(\pi) = \sum_{\bar{a}_T} Y(\bar{a}_T) \prod_{v=1}^{T} 1_{[\pi_v\{H_v(\bar{a}_{v-1})\} = a_v]}.$$

De ne the value of regime π to be $V(\pi) = EY(\pi)$, i.e., the marginal mean outcome if all subjects were assigned treatment according to π. The optimal regime, $\pi^{opt} \in \Pi$, satis es $V(\pi^{opt}) \geq V(\pi)$ for all $\pi \in \Pi$. To identify π^{opt} in terms of the data-generating model, we make the following assumptions: (i) consistency, $H_t = H_t(\bar{A}_{t-1})$ for $t = 2; \ldots; T$ and $Y = Y(\bar{A}_T)$, (ii) strong ignorability, $A_t \perp W \mid H_t$ for $t = 1; \ldots; T$, and (iii) positivity, $P(A_t = a_t \mid H_t = h_t) > 0$ for all $a_t \in \mathcal{A}_t(h_t)$ and $t = 1; \ldots; T$. These assumptions are standard in the dynamic treatment regimes

literature (see Robins, 2004; Chakraborty & Moodie, 2013; Schulte et al., 2014; Kosorok & Moodie, 2015, for additional discussion). Hereafter, we implicitly assume that these conditions hold.

2.3 Estimation with complete data

In this section, we review estimation of an optimal treatment regime when the data are completely observed. We consider a class of estimators that are representable as solutions to a set of estimating equations. This class is quite broad and includes most of the estimators commonly used in practice. To illustrate this point, we show in the Appendix that Q-learning and a generalization of outcome weighted learning belong to this class.

We consider treatment regimes of the form $= f_1(\ ;\ _1);\ldots;\ _T(\ ;\ _T)g$ in which the decision rules composing the regime are indexed by parameters$= (\ _1^T;\ _2^T;\ldots;\ _T^T)^T \ 2\ B$, where B is a normed linear space with norm$jj\ \ jj_B$. For example, one might consider linear decision rules of the form $_t(h_t;\ _t) = \text{sign}\ \ _t^T h_{t;0}$, where $h_{t;0}$ is a feature vector constructed fromh_t and sign(u) is 1 if u is positive and $\ 1$ otherwise. We do not exclude the case in which $_t$ includes nuisance parameters so that $_t(\ :\ _t)$ depends only on a subvector of $_t$; however, we do not make any special distinction for such nuisance parameters as it is not important for our purposes. We assume that an estimator$b_n = (\ b_{1;n}^T;\ldots;b_{T;n}^T)^T$ of is constructed by solving the estimating equation

$$P_n m_n(H_T;A_T;Y;\ ) = 0 \tag{2.1}$$

over $\ 2\ B$, where P_n denotes the empirical measure, and$m_n : H_T\ \ A_T\ \ Y\ !\ \ R^J$. The dependence of$m_n$ on n is to allow for regularization or other factors that may vary with the sample size (Qian & Murphy, 2011; Zhao et al., 2012; Zhao et al., 2015; Jeng et al., 2018).

Many estimators of an optimal treatment regime are based on backwards recursion. For these estimators, some components ofm_n will depend on the partial history $S_t\ ,\ (H_t^T;A_t)^T$ for

$t = 1, \ldots, T$ rather than the complete data $S_{T+1} \equiv (H_T^T, A_T^T, Y)^T$. For $t = 1, \ldots, T$ define J_t to be the indices of m_n such that $m_{n,j}$ depends on S_{T+1} only through S_t, i.e.,

$$J_t = \{1 \leq j \leq J : m_{n,j}(h_T, a_T, y, \psi) = m_{n,j}(h_t, a_t, \psi) \text{ for some } m_{n,j} : \mathcal{H}_t \times \mathcal{A}_t \to \mathbb{R}\},$$

and J_{T+1} are the indices that rely on the complete data. Under this representation, the estimation equation in (2.1) can be equivalently expressed as

$$\mathbb{P}_n \dot{m}_{n,j}(S_t, \psi) = 0 \quad \text{for all} \quad j \in J_t, \quad t = 1, \ldots, T+1. \tag{2.2}$$

We will exploit this representation to use more of the observed data in constructing weighted complete case estimators in Section 2.4.

Let ψ_n denote the population analog of $\hat{\psi}_n$, i.e., the solution to (2.2) with $\mathbb{P}_n$ replaced by $\mathbb{P}$. We say that $\hat{\psi}_n$ is consistent if $\|\hat{\psi}_n - \psi_n\|_B$ converges to zero in probability. Because our objective is not to propose new estimators of an optimal treatment regime, we will assume that the estimating equation has been suitably constructed to ensure consistency under the data-generating model in the complete data case and avoid stating specific conditions under which such consistency holds. Giving such general conditions would be cumbersome; for example, the conditions under which Q-learning with linear models is consistent are quite different from those under which kernel-based outcome weighted learning is consistent. Before describing how to adjust the estimating equations to accommodate missing data we first briefly recount how to express Q-learning and outcome weighted learning in the form given in (2.1).

2.3.1 Q-learning with complete data

Q-learning is an approximate dynamic programming algorithm that has been applied to estimate optimal treatment regimes in a variety of biomedical and engineering applications (Murphy, 2005a; Busoniu et al., 2010; Geramifard et al., 2013; Schulte et al., 2014; Sutton & Barto,

2018). The basis for Q-learning is the dynamic programming characterization of an optimal regime. De ne $Q_T(h_T; a_T) = E(Y | H_T = h_T; A_T = a_T)$, and recursively for $t = T - 1; T - 2; \ldots; 1$ de ne $Q_t(h_t; a_t) = E\left[\max_{a_{t+1} \in \mathcal{A}_{t+1}(H_{t+1})} Q_{t+1}(H_{t+1}; a_{t+1}) \mid H_t = h_t; A_t = a_t\right]$. It follows from dynamic programming that the optimal regime satis es $\eta_t^{opt}(h_t) = \arg\max_{a_t \in \mathcal{A}_t(h_t)} Q_t(h_t; a_t)$ (Bellman, 1957). Let $Q_t(h_t; a_t; \beta_t)$ denote a posited class of models for $Q_t(h_t; a_t)$ indexed by $\beta_t \in B_t$ for $t = 1; \ldots; T$. The induced class of treatment regimes is thus of the form $\eta_t(h_t; \beta_t) = \arg\max_{a_t \in \mathcal{A}_t(h_t)} Q_t(h_t; a_t; \beta_t)$ (Zhang et al., 2012b). If $B_t = R^{p_t}$ and $Q(h_t; a_t; \beta_t)$ is di erentiable in β_t for all $h_t; a_t \in H_t \times A_t$, then $\hat{\beta}_{T;n}$ solves

$$P_n \{Y - Q_T(H_T; A_T; \beta_T)\} \nabla_{\beta_T} Q_T(H_T; A_T; \beta_T) = 0; \tag{2.3}$$

and for $t = T - 1; T - 2; \ldots; 1$ the estimators $\hat{\beta}_{t;n}$ solve

$$P_n \left[\max_{a_{t+1} \in \mathcal{A}_{t+1}(H_{t+1})} Q_{t+1}(H_{t+1}; a_{t+1}; \hat{\beta}_{t+1;n}) - Q_t(H_t; A_t; \beta_t)\right] \nabla_{\beta_t} Q_t(H_t; A_t; \beta_t) = 0: \tag{2.4}$$

Thus, the estimator $\hat{\beta}_n$ is obtained by solving $P_n m_n(\beta) = 0$, where m_n is constructed by stacking (2.3) and (2.4) for $T - 1; \ldots; 1$ so that $\hat{\beta}_n$ is a root of

$$P_n \begin{bmatrix} \{Y - Q_T(H_T; A_T; \beta_T)\} \nabla_{\beta_T} Q_T(H_T; A_T; \beta_T) \\ \left[\max_{a_T \in \mathcal{A}_T(H_T)} Q_T(H_T; a_T; \beta_T) - Q_{T-1}(H_{T-1}; A_{T-1}; \beta_{T-1})\right] \nabla_{\beta_{T-1}} Q_{T-1}(H_{T-1}; A_{T-1}; \beta_{T-1}) \\ \vdots \\ \left[\max_{a_2 \in \mathcal{A}_2(H_2)} Q_2(H_2; a_2; \beta_2) - Q_1(H_1; A_1; \beta_1)\right] \nabla_{\beta_1} Q_1(H_1; A_1; \beta_1) \end{bmatrix}:$$

The estimated optimal decision at staget is thus $\hat{\eta}_{n;t}(h_t) = \arg\max_{a_t \in \mathcal{A}_t(h_t)} Q_t(h_t; a_t; \hat{\beta}_{t;n})$.

Similar expressions can be obtained for nonparametric variants of Q-learning (Zhao et al., 2009; Moodie et al., 2014; Zhang et al., 2017). For the purpose of illustration, we brie y describe kernel ridge regression for Q-learning. Suppose that $H_t \subset R^{p_t}$ for all $t = 1; \ldots; T$. For each t, let $K_t : R^{p_t} \times R^{p_t} \to R$ be symmetric and positive de nite and write $\mathcal{H}_t$ to denote the corresponding reproducing kernel Hilbert space with norm $\|\cdot\|_{\mathcal{H}_t}$ (for an introduction see Cristianini & Shawe-

Taylor, 2000; Moguerza & Muñoz, 2006; Berlinet & Thomas-Agnan, 2011; Nosedal-Sanchez et al., 2012). To approximate Q_T within H_T, one solves for each $a \in \{-1, 1\}$

$$\hat{Q}_{T;n}(\cdot\,; a) = \arg\min_{f_a \in \mathcal{H}_T} \mathbb{P}_n \mathbb{1}_{A_T = a} \{Y - f_a(H_T)\}^2 + \lambda_{T;a;n} \|f_a\|^2_{\mathcal{H}_T}; \tag{2.5}$$

where $\lambda_{T;a;n} > 0$ is a tuning parameter. For each $a \in \{-1, 1\}$ define $I_{T;a} = \{i : A_{T;i} = a\}$ to be the subset of patients to receive treatment a at time T and define $Z^T_{T;a}(h_T) = \{K_T(H_{T;i}; h_T)\}_{i \in I_{T;a}}$. It follows that $\hat{Q}_{T;n}(h_T; a_T) = Z^T_{T;a_T}(h_T)\hat{\beta}_{T;a;n}$, where $\hat{\beta}_{T;a;n}$ is a solution of

$$\mathbb{P}_n \mathbb{1}_{A_T = a_T} \left\{ \prod_{} (1 + \epsilon_{T;a;n}) Z^T_{T;a_T}(H_T) - \lambda_{T;a} Z_{T;a_T}(H_T) \right\} = 0.$$

Notice that in the previous estimating equation, we have replaced $\lambda_{T;a;n}$ by $\epsilon_{T;a;n}$ to reflect in re-writing the estimator the penalty has been scaled by the number of subjects receiving treatment a_T. Constructing $Z_{t;a_t}(h_t)$ analogously for $t = T-1, T-2, \ldots, 1$, one can construct estimators $\hat{Q}_{t;n}(h_t; a_t) = Z^T_{t;a_t}(h_t)\hat{\beta}_{t;a;n}$ where $\hat{\beta}_{t;a;n}$ is a solution of

$$\mathbb{P}_n \mathbb{1}_{A_t = a_t} \left\{ \max_{a_{t+1} \in \mathcal{A}_{t+1}(H_{T+1})} \hat{Q}_{t+1;n}(H_{t+1}; a_{t+1}) - (1 + \epsilon_{t;a;n}) Z^T_{t;a_t} \lambda_{t;a} Z_{t;a_t}(H_t) \right\} = 0.$$

The preceding estimating equations can be stacked to obtain a single estimating equation; see Zhang et al. (2017) for additional details.

2.3.2 Outcome weighted learning with complete data

Direct-search methods, also known as value-search methods, estimate an optimal treatment regime by maximizing an estimator of the marginal mean outcome over a pre-specified class of regimes (Orellana et al., 2010; Zhang et al., 2012a; Rubin & Laan, 2012; Zhang et al., 2013; Zhang & Zhang, 2018). Outcome weighted learning (OWL) comprises a subclass of these methods, which rely on the use of a convex surrogate to carryout the proposed maximization (see below for details). Outcome weighted learning was introduced for single-stage decisions by Zhao et al.

(2012) and has since been generalized to multi-stage decisions (Zhao et al., 2015) and undergone a number of other modi cations and re nements (Zhao et al., 2014; Chen et al., 2016; Zhou et al., 2017; Liu et al., 2018; Qi & Liu, 2018a; Qi et al., 2018b).

We consider a variant of outcome weighted learning that uses a convex relaxation of the augmented inverse probability weighted estimator of the marginal mean outcome (Zhang et al., 2013; Liu et al., 2018; Zhao et al., 2019a). We use a backwards recursive procedure (see Zhao et al., 2015) to extend the single-stage procedure proposed by Zhao et al. (2019a) to multiple-stages; while this is our not main methodological contribution, it may of be interest in its own right (see Zhang & Zhang, 2018; Davidian et al., 2019, for additional details).

We describe the estimator as a sequence of models $_t$, each indexed by their own parameters, before stacking the models and parameters into a single estimating equation. To ease bookkeeping and clarify development, we begin by using $= (^T_1; \ldots; ^T_T)^T$ solely to index the decision rules and use $= (^T_1; \ldots; ^T_T)^T$ to index nuisance models; we later pool these into a single collection of parameters, $= (^T_1; \ldots; ^T_T)^T$, to match the notation used in our general framework. For simplicity, we assume linear decision rules of the form $_t(h_t; _t) = \text{sign}(^T_t h_{t;0})$, where $h_{t;0}$ is known feature vector constructed from h_t. Furthermore, let $Q_T(h_T; a_T; _T)$ be a posited model for $Q_T(h_T; a_T) = E(Y \mid H_T = h_T; A_T = a_T)$. For each $_T$, de ne $V_T(h_T; _T) = Q_T\{h_T; _T(h_T; _T)\}$ with corresponding posited model $V_T(h_T; _T; _T) = Q_T\{h_t; _T(h_T; _T); _T\}$. De ne

$$Q_{T-1}(h_{T-1}; a_{T-1}; _{T-1}; _T) = E\{V_T(H_T; _T; _T) \mid H_{T-1} = h_{T-1}; A_{T-1} = a_{T-1}\};$$

thus $Q_{T-1}(h_{T-1}; a_{T-1}; _{T-1}; _T)$ is the mean outcome for a patient presenting with $H_{T-1} = h_{T-1}$, treated with $A_{T-1} = a_{t-1}$, and subsequently treated according to $_T(; _T)$ assuming that the model $Q_T(h_T; a_T; _T)$ is correct. Let $Q_{T-1}(h_{T-1}; a_{T-1}; _{T-1}; _T; _{T-1})$ be a posited model for $Q_{T-1}(h_{T-1}; a_{T-1}; _{T-1}; _T)$. Using an underbar to denote future, e.g., $_t = (^T_t; \ldots; ^T_T)^T$, de ne

$$V_{T-1}(h_{T-1}; _{T-1}; _{T-1}) = Q_{T-1}\{h_{T-1}; _{T-1}(h_{T-1}; _{T-1}); _{T-1}; _T; _{T-1}\}$$

to be the expected outcome for a patient presenting with $H_{T-1} = h_{T-1}$ and treated according to $\beta_{T-1}(\cdot\,; \psi_{T-1})$ and $\beta_T(\cdot\,; \psi_T)$ at times $T-1$ and T under the models indexed by ψ_{T-1}. Recursively, for $t = T-2, \ldots, 1$ define

$$Q_t(h_t, a_t; \psi_{t+1}, \beta_{t+1}) = E\left[V_{t+1}(H_{t+1}; \psi_{t+1}, \beta_{t+1}) \mid H_t = h_t, A_t = a_t\right],$$

and let $Q_t(h_t, a_t; \psi_{t+1}, \beta_{t+1}, \psi_t)$ denote a posited model. Subsequently, define

$$V_t(h_t; \psi_t, \beta_t) = Q_t\left(h_t; \psi_t, \beta_t(h_t; \psi_t); \psi_{t+1}, \beta_{t+1}, \psi_t\right).$$

Thus, for any regime β indexed by $\psi = \psi_1$ the marginal mean outcome under the models indexed by $\psi = \psi_1$ is $V(\psi) = \mathbb{E}V_1(H_1; \psi, \beta)$. This construction is the basis for Q-learning with policy-search (Taylor et al., 2015; Zhang et al., 2017); however, we will not be using the Q-functions in this way. Our purpose in deriving them is to include them as augmentation terms in a doubly robust estimator of the incremental (stagewise) regret.

We assume that $Q_t(h_t, a_t; \psi_{t+1}, \beta_{t+1}, \psi_t)$ is continuously differentiable in ψ_t for all h_t, a_t, ψ_{t+1} and β_{t+1}. For each n, let $\hat{\beta}_n(\cdot) = \left\{\hat{b}_{1;n}(\cdot\,; \psi_2), \hat{b}_{2;n}(\cdot\,; \psi_3), \ldots, \hat{b}_{T-1;n}(\cdot\,; \psi_T), \hat{b}_{T;n}\right\}$ be a root of the estimating equation

$$\mathbb{P}_n \begin{bmatrix} \{Y - Q_T(H_T, A_T; \psi_T)\}\, \dot{r}_{\psi_T} Q_T(H_T, A_T; \psi_T) \\ \{V_T(H_T; \psi_T, \beta_T) - Q_{T-1}(H_{T-1}, A_{T-1}; \psi_T, \beta_T, \psi_{T-1})\}\, \dot{r}_{\psi_{T-1}} Q_{T-1}(H_{T-1}, A_{T-1}; \psi_T, \beta_T, \psi_{T-1}) \\ \vdots \\ V_{t+1}(H_{t+1}; \psi_{t+1}, \beta_{t+1}) - Q_t(H_t, A_t; \psi_{t+1}, \beta_{t+1}, \psi_t)\, \dot{r}_{\psi_t} Q_t(H_t, A_t; \psi_{t+1}, \beta_{t+1}, \psi_t) \\ \vdots \\ V_2(H_2; \psi_2, \beta_2) - Q_1(H_1, A_1; \psi_2, \beta_2, \psi_1)\, \dot{r}_{\psi_1} Q_1(H_1, A_1; \psi_2, \beta_2, \psi_1) \end{bmatrix} = 0, \quad (2.6)$$

where we note that ψ_1 does not affect $\hat{b}_n(\cdot)$; however, to estimate the performance of β_o with the estimated Q-functions one would use $\mathbb{P}_n Q_1\left(H_1; \psi_1(H_1; \psi_1), \psi_2, \hat{b}_{2;n}(\cdot\,; \psi_3), \hat{b}_{1;n}(\cdot\,; \psi_2)\right)$, which can be seen to depend on all of ψ.

The second set of estimating equations are based on a backwards recursive representation of an augmented inverse probability weighted estimator of the incremental regret. As noted previously, the estimated Q-functions derived above serve as augmentation terms. De ne $_T(H_T; A_T; _T) = Q_T(H_T; A_T; _T)$ $Q_T(H_T; A_T; _T)$ and $\hat{}_{T;n}(H_T; A_T) = _T(H_T; A_T; \hat{}_{T;n})$. De ne the estimated incremental regret at stage T as

$$\begin{aligned} J_{T;n}(_T; _T) &= P_n 1_{\text{signf } W_T(H_T; A_T; _T) g A_T \in _T(H_T; _T)} W_T(H_T; A_T; Y; _T) \\ &= P_n 1_{\text{signf } W_T(H_T; A_T; Y; _T) g A_T \ _T^T H_{T;0} < 0} W_T(H_T; A_T; Y; _T) ; \end{aligned} \tag{2.7}$$

where

$$W_T(H_T; A_T; Y; _T) = \frac{Y \quad Q_T(H_T; \ A_T; _T) \quad 1 \quad P(A_T \ H_T) \quad _T(H_T; A_T; _T)}{P(A_T j H_T)}:$$

De ne $\hat{J}_{T;n}(_T) = J_{T;n}(_T; \hat{}_{T;n})$. It can be shown that $\hat{J}_{T;n}(_T)$ is, up to an additive constant that does not depend on $_T$, a doubly robust estimator for the di erence between: (i) the marginal mean outcome for a patient receiving treatment as per protocol (i.e., under the data-generating model) for the rst T 1 time points, followed by treatment under an optimal regime and (ii) the marginal mean outcome for a patient receiving treatment per protocol for the rst T 1 time points followed by treatment under $_T(; _T)$ (this is a variant of the so-called estimated regret, see Davidian et al., 2019, for detailed discussion and justi cation of this seemingly unusual estimand). Thus, $\hat{J}_{T;n}(_T)$ is a measure of the loss incurred by treatment of patients at time T, who had theretofore been treated per protocol, with $_T(; _T)$ rather than an optimal treatment. Because of the indicator function, direct minimization of (2.7) over $_T$ is a mixed integer program (Wolsey & Nemhauser, 2014), which is NP-hard in general and thus typically requires the use of either specialized software or the use of heuristics (Zhang et al., 2013; Zhang & Zhang, 2018). A

distinguishing feature of outcome weighted learning is the relaxation of (2.7) by replacing the indicator function with a convex function. Given a convex function $L : R \to R$, called a convex surrogate, backwards recursive outcome weighted learning uses the objective

$$J^L_{T;n}(\beta_T; \gamma_T) = P_n L\left[\text{sign}\{W_T(H_T, A_T, Y; \gamma_T)\}(A_T - \beta_T^T H_{T;0})\right] W_T(H_T, A_T, Y; \gamma_T),$$

which can be seen to be a convex function of β_T for each γ_T. Define $\hat{J}^L_{T;n}(\beta_T) = J^L_{T;n}(\beta_T; \hat{b}_{T;n})$. We define $\hat{b}_{T;n}$ to be the solution to $\min_{\beta_T} \hat{J}^L_{T;n}(\beta_T) = 0$, where the gradient can be replaced with a sub-gradient if L is not differentiable (Boyd & Vandenberghe, 2004). Thus, the estimated optimal rule at stage T is $b_T(\cdot; \hat{b}_{T;n})$.

Estimators of the optimal decision rules at stages $t = T-1, T-2, \ldots, 1$ solve analogous estimating equations, which are defined recursively as follows. For $t = T-1$ define

$$\begin{aligned}&\gamma_{T-1}(H_{T-1}, A_{T-1}, \beta_T, \gamma_T)\\ &\quad = Q_{T-1}\{H_{T-1}, A_{T-1}, \beta_T, \gamma_T, \gamma_{T-1}\} - Q_{T-1}\{H_{T-1}, A_{T-1}, \beta_T, \gamma_T, \gamma_{T-1}\},\end{aligned}$$

and define $\hat{b}_{T-1;n}(H_{T-1}, A_{T-1}) = \left\{\beta_{T-1} H_{T-1}, A_{T-1}, \hat{b}_{T;n}(b_{T;n})\right\}$, where $\hat{b}_{T;n}(b_{T;n}) = \left\{\hat{b}_{T-1;n}(b_{T;n}), \hat{b}_T\right\}$. Subsequently, define the relaxed loss function

$$\begin{aligned}&J^L_{T-1;n}(\beta_{T-1}; \gamma_{T-1}) =\\ &\quad P_n L\left[\text{sign}\{W_{T-1}(H_T, A_T, Y; \gamma_{T-1}, \gamma_T)\}(A_{T-1} - \beta_{T-1}^T H_{T-1;0})\right] W_{T-1}(H_T, A_T, Y; \gamma_{T-1}, \gamma_T),\end{aligned}$$

where the weights are given by

$$W_{T-1}\{H_T, A_T, Y, \gamma_{T-1}, \gamma_T\} = \frac{1_{A_T = \gamma_T(H_T,\gamma_T)} Y}{P(A_T|H_T)P(A_{T-1}|H_{T-1})} \frac{Q_{T-1}\{H_{T-1}, A_{T-1}, \gamma_T, \gamma_{T-1}\}}{P(A_{T-1}|H_{T-1})} \left(\frac{\{1 - P(A_{T-1}|H_{T-1})\} g_{\gamma_{T-1};n}(H_{T-1}, A_{T-1}, \gamma_T, \gamma_{T-1})}{P(A_{T-1}|H_{T-1})} - \frac{1_{A_T = \gamma_T(H_T,\gamma_T)}}{P(A_T|H_T)P(A_{T-1}|H_{T-1})} + \frac{1_{A_T = \gamma_T(H_T,\gamma_T)}}{P(A_T|H_T)} \right) Q_{T;n}\{H_T, \gamma_T(H_T, \gamma_T), \gamma_T\}.$$

Define

$$J^L_{T-1;n}(\gamma_{T-1}) = J_{T-1}\left\{\gamma_{T-1}, b_{T;n}, b_{T-1;n}(b_{T;n})\right\},$$

and let $b_{T-1;n}$ be a solution to $r_{\gamma_{T-1}} J^L_{T-1;n}(\gamma_{T-1}) = 0$. For a generic $t < T - 1$, define

$$_t(H_t, A_t, \gamma_t, \gamma_{t+1}) = Q_t\{H_t, A_t, \gamma_{t+1}, \gamma_{t+1}, \gamma_t\} - Q_t\{H_t, A_t, \gamma_{t+1}, \gamma_{t+1}, \gamma_t\},$$

and define $b_{t;n}(H_t, A_t) = \left\{H_t, A_t, b_{t;n}(b_{t+1;n}), b_{t+1;n}\right\}$, where $b_{t;n}(b_{t+1;n}) = \{b_{t;n}(b_{t+1;n}), b_{t+1;n}(b_{t+2;n}), \ldots, b_{T;n}\}$. Subsequently, define

$$J^L_t(\gamma_t, \gamma_t) = P_n L\left[\text{sign}\{W_t(H_T, A_T, Y, \gamma_{t+1}, \gamma_t) A_t = \gamma_t(H_t, \gamma_t)\} W_t(H_T, A_T, Y, \gamma_{t+1}, \gamma_t)\right],$$

where the weights are given by

$$W_t(H_T, A_T, Y, \gamma_{t+1}, \gamma_t) = \frac{Y \prod_{s=t+1}^{T} 1_{A_s = \gamma_s(H_s,\gamma_s)}}{\prod_{s=t}^{T} P(A_s|H_s)} \frac{Q_t(H_t, A_t, \gamma_{t+1}, \gamma_{t+1}, \gamma_t)}{P(A_t|H_t)} \left(\frac{1 - P(A_t|H_t)}{P(A_t|H_t)} g_{t;n}(H_t, A_t, \gamma_{t+1}, \gamma_t) + \sum_{r=t+1}^{T} \left[\frac{\prod_{s=t+1}^{r-1} 1_{A_s = \gamma(H_s,\gamma_s)}}{\prod_{s=t}^{r-1} P(A_s|H_s)} \left(\frac{1_{A_r = \gamma_r(H_t,\gamma_r)}}{P(A_r|H_r)} - 1 \right) Q_r\{H_r, \gamma_r(H_r, \gamma_r), \gamma_{r+1}, \gamma_{r+1}, \gamma_r\} \right] \right).$$

Define $J^L_{t;n}(\gamma_t) = J^t\left\{\gamma_t, b_{t+1;n}, b_{t;n}(b_{t+1;n})\right\}$, and let $b_{t;n}$ be a solution to $r_{\gamma_t} J^L_{t;n}(\gamma_t) = 0$.

The joint estimating equations for $\psi = (\gamma^T, \gamma^T)^T$ are obtained by stacking (2.6) and the

estimating equations $r_{\psi_t} J_t(\psi_t, \eta_t) = 0$ for $t = 1, \ldots, T$.

2.4 Estimation with incomplete data

2.4.1 Missingness mechanism

We assume that baseline covariate information and initial treatment assignment, (X_1, A_1), are always observed. This assumption generally holds in practice because patients who do not receive an initial treatment assignment are often unenrolled from the study and excluded from subsequent analyses. We further assume that the missingness pattern is nearly monotone; i.e., any item missingness that violates this monotone pattern is sparse, and the missingness pattern has been made monotone through artificial censoring or single imputation. Because patient dropout is the primary cause for missing data in longitudinal studies, e.g. SMARTS, this assumption is common in the literature (e.g., Shortreed et al., 2014, and references therein).

Let $C \in \{1, \ldots, T+1\}$ denote the dropout time so that $C = t$ if the patient dropped out after assignment of A_t for $t = 1, \ldots, T$ and $C = T+1$ if the patient's trajectory is fully observed, i.e.,

$$C = \begin{cases} T+1; & \text{observe } (X_1, A_1, \ldots, X_T, A_T, Y) = (H_T, A_T, Y); \\ T; & \text{observe } (X_1, A_1, \ldots, X_T, A_T) = (H_T, A_T); \\ \vdots & \vdots \\ t; & \text{observe } (X_1, A_1, \ldots, X_t, A_t) = (H_t, A_t); \\ \vdots & \vdots \\ 1; & \text{observe } (X_1, A_1) = (H_1, A_1). \end{cases}$$

We further assume that the data are missing at random (MAR, Rubin, 1976; Little & Rubin, 2014) so that $1_{C=t} \perp (X_{t+1}, A_{t+1}, \ldots, X_T, A_T, Y) \mid H_t, A_t$ for all $t = 1, \ldots, T$.

The simplest strategy for adapting the estimating equations presented in Section 2.3 for

missing data is through inverse probability weighting of complete cases. Note that `complete case' for a term $m_{n;j}(S_t;\ )$, where $j\ 2\ J_t$, is the one in which S_t is observed and not necessarily the one for which the complete trajectory, S_{T+1}, is observed.

2.4.2 Inverse probability weighted estimating equations

Inverse probability weighted complete case (IPWCC) estimators re-weight the terms in the estimating equation for an optimal regime by their respective probabilities of being observed (Tsiatis, 2007; Tsiatis et al., 2014). De ne the discrete hazard of dropout at time $t = 1;:::;T$ to be $_t(s_t) = P(C = tjC\ \ t; S_t = s_t)$. Thus, $_t(s_t)$ is the probability of dropping out at stage t for a patient with covariate and treatment history $S_t = s_t$. The survivor function at time t is thus $K_t(s_t) = P(C > t\,jS_t = s_t) = \prod_{v=1}^{t} f1\ \ _v(s_v)g$. Under the MAR assumption, we can model the hazards using a binary regression model. For concreteness, we use a logistic regression model so that

$$_t(s_t;\ _t) = \text{expit}\ f g_t(s_t;\ _t)g;$$

where $\text{expit}(u)\ ,\ \exp(u)=f1+\exp(u)g$, $_t\ 2\ K_t\ \ R^{q_t}$ is a vector of parameters, and $g_t(s_t;\ _t)$ is continuously di erentiable in $_t$ for all s_t. De ne $\ ,\ (\ _1^T;:::;\ _T^T)^T\ 2\ K\ \ R^q$; where $q = q_1 + \ \ + q_T$. De ne $\&_t(c;s_t;\ _t)\ ,\ 1_{c=t}r\ _t g_t(s_t;\ _t)\ \ 1_{c\ t}r\ _t g_t(s_t;\ _t)\text{expit}\,f g_t(s_t;\ _t)g$ to be the score function of the posited logistic regression model, and let $\hat{}_{t;n}$ be a solution to $P_n\&_t(C;S_t;\ _t) = 0$. Let $\bar{}_t = (\ _1^T;\ _2^T;:::;\ _t^T)^T$ so that $\hat{}_{t;n} = (\ b_{1;n}^T; b_{2;n}^T;:::; b_{t;n}^T)^T$. The estimated survivor function is $K_t(s_t;\hat{}_{t;n}) = \prod_{v=1}^{t}\ 1\ \ _v(s_v; b_{v;n})\ .$

De ne the complete case weights at level $t = 2;:::;T+1$ and $j\ 2\ J_t$ under parameters $\ $ to be $w_j^{cc}(c;s_{t\ 1};\bar{}_{t\ 1}) = 1_{c>(t\ 1)}=K_{t\ 1}(s_{t\ 1};\bar{}_{t\ 1})$. The IPWCC estimator of an optimal treatment regime based on (2.2) solves

$$P_n w_j^{cc}(C;S_{t\ 1};\hat{}_{t\ 1;n})\,m_{n;j}(S_t;\ ) = 0;\quad j\ 2\ J_t;\quad t = 1;:::;T+1;$$

with the understanding that $w_j^{cc}(c; s_0; \bar{\ }_0)$ 1 for j 2 J_1. The preceding equations along with those for could be expressed as a single stacked estimating equation by concatenating the score equation for the logistic regression models onto m_n. Let $P_n m_n(S_{T+1}; ;) = 0$ denote this joint estimating equation. It follows from the derivations given in the Appendix that both the Q-learning and outcome weighted learning estimators can thus be constructed under a monotone missingness pattern using IPWCC by means of the preceding estimating equation. The following result can be used to establish consistency of the IPWCC estimator when combined with standard conditions for Z-estimators, e.g., the estimating equation has a unique isolated minimizer (see Kosorok, 2007).

Theorem 1 Assume that the survivor function is correctly speci ed so that $K_t(s_t) = K_t(s_t; \bar{\ }_t)$ for all t and s_t, for some 2 K and that the $\hat{\ }_n$! in probability. If $_n$ satis es $\|P m_n(S_{T+1}; _n)\| = o(1)$, then $\|P m_n(S_{T+1}; _n; \hat{\ }_n)\| = o_p(1)$.

2.4.3 Augmented inverse probability weighted estimating equations

Using semi-parametric e ciency theory for monotone coarsening, of which monotone missingness is a special case, we derive an augmented inverse probability weighted complete case (AIPWCC) estimator (Robins & Rotnitzky, 1995; Tsiatis, 2007). De ne for each $t = 1; : : : ; T$, the conditional mean $d_{n;t}(s_t)$, $E\ m_n(S_{T+1};)\ S_t = s_t$ for which we posit parametric models $d_{n;t}(s_t; _t)$ indexed by $_t$ 2 R^{v_t}. Let = ($^T_1; : : : ; ^T_T)^T$. An estimator of $d_{n;t}$ can thus be obtained by regressing $m_n(S_{T+1}; \hat{\ }_n)$ on $d_{n;t}(S_t; _t)$ restricted to patients with $C = T + 1$, which gives $\hat{\ }_{t;n}$ and subsequently $\hat{d}_{n;t}(s_t) = d_{n;t}(s_t; \hat{\ }_{t;n})$ for each $t = 1; : : : ; T$. Let $_t(s_t; _t)$ and $K_t(s_t; \bar{\ }_t)$ be as de ned in the previous section. For each $t = 2; : : : ; T + 1$, j 2 J_t, and $r = 1; : : : ; t$ 1, de ne the augmentation weights

$$w_{r;j}^{aug}(c; s_r; \bar{\ }_r) = \frac{1_{c=r} \quad _r(s_r; _r)}{K_r(s_r; \bar{\ }_r)}:$$

The AIPWCC estimating equations are

$$P_n \; w_j^{cc}(C; S_{t\ 1}; \hat{b}_{t\ 1;n})\, \hat{m}_{n;j}(S_t; \ ) + \sum_{r=1}^{t\ 1} w_{r;j}^{aug}(C; S_r; \hat{b}_{r;n})\, d_{n;r;j}(S_r; b_{r;n}) \ = 0; \quad \text{for all} \quad j \ 2 \ J_t; t = 1; : : : ; T + 1:$$

To obtain a single set of estimating equations one could concatenate the estimating equations for and to those given above. Let $P_n m_n(S_{T+1}; \ ; \ ; \ ) = 0$ denote the joint estimating equations. Explicit forms of these estimating equations for both Q-learning and outcome weighted learning are provided in the Appendix.

Theorem 2 Assume that the hazard functions for dropout are correctly speci ed so that $_t(s_t) = \ _t(s_t; \ _t)$ for all s_t for some $_t \ 2 \ K_t$ and $\hat{b}_n \ !$ in probability or that the regression functions are correctly speci ed so that $d_{n;t}(s_t) = d_{n;t}(s_t; \ _t)$ for some $_t \ 2 \ a_t$ and $b_n \ !$ in probability. If $_n$ satis es $jjP m_n(S_{T+1}; \ _n)jj = o(1)$, then $jjP m_n(S_{T+1}; \ _n; \hat{b}_n; b_n)jj = o_p(1)$. Thus, the AIPWCC estimator is doubly robust.

Remark 1 When the dimension of the trajectory space is high, the solutions to the estimating equations may be unstable. In these cases, regularization may be necessary to stabilize the solution and to reduce over- tting. (see Fu, 2003; Johnson et al., 2008, for additional discussion of regularization with estimating equations). In our simulation experiments, we regularize the estimating equation using an adaptive ridge penalty; i.e., given a preliminary estimator b_n^0 of based on the un-penalized estimating equation, we compute $\hat{b}_n$ as the solution to $P_n m_n(S_{T+1}; \ ; \ b_n; b_n) + \ _n = jb_n j = 0$, where the division is taken elementwise and $_n \ 0$ is a tuning parameter.

2.5 Simulation and Data Application

2.5.1 Simulation Studies

We compare the performances of IPWCC, AIPWCC, and multiple imputation (MI) in terms of the value of the estimated optimal regime; these methods are applied with monotone missing data and the estimation is done using either Q-learning or outcome weighted learning. Data are simulated to mimic a two-stage SMART with binary treatments at each stage. The complete data are generated as follows:

$$
\begin{aligned}
&X_1 = (X_{11}, \ldots, X_{1p})^T; X_{1k} \sim \text{Bernoulli}(0.5), k = 1, \ldots, p, \\
&A_1 \sim \text{Uniform}\{-1, 1\}, \\
&X_2 \mid X_1 = x_1, A_1 = a_1 \sim N_p(\beta_0 + \beta_1 a_1) x_1, \sigma^2 I_p), \\
&A_2 \sim \text{Uniform}\{-1, 1\}, \\
&Y \mid X_1 = x_1, A_1 = a_1, X_2 = x_2, A_2 = a_2 \sim N(\mu_Y(x_1, a_1, x_2, a_2), \sigma_Y^2),
\end{aligned}
\tag{2.8}
$$

where $\mu_Y(x_1, a_1, x_2, a_2) = \alpha_{20} + \alpha_{21} a_1 + \alpha_{22}^T x_1 a_1 + \alpha_{23}^T x_2 + (\beta_{20} + \beta_{21} a_1 + \beta_{22}^T x_2) a_2$. Thus, the model is indexed by the matrices $\beta_0, \beta_1 \in \mathbb{R}^{p \times p}$, coefficients $\alpha_{20}, \alpha_{21}, \alpha_{22}, \alpha_{23}, \beta_{20}, \beta_{21}, \beta_{22}$, and variance components $\sigma^2, \sigma_Y^2 > 0$.

For the missingness mechanism, we consider hazard functions of the form

$$
P(C = t \mid C \geq t, H_t, A_t) = \pi_t(H_t, A_t) = \text{expit}\{\psi_{t1} + \psi_{t2} X_{t1} + \psi_{t3} A_t + \psi_{t4} X_{t2}^T \psi_{t5}\}, \quad t = 1, 2.
$$

We vary the parameters ψ_1 and ψ_2 to obtain 35% and 65% missingness. The actual parameter values are provided in the Appendix. All simulation experiments use training sets of size $n = 1000$ and 500 Monte Carlo replications.

In our implementation of Q-learning, we use linear models of the form $\text{set} Q_t(H_t, A_t, \beta_t) =$

$_t^T B_{t;0}$, where $B_{1;0} = (1; X_1^T; A_1; X_1^T A_1)^T$ and $B_{2;0} = (1; X_1^T; A_1; X_1^T A_1; A_2; A_1 A_2; X_2^T A_2)^T$. For outcome weighted learning, we use linear decision rules of the form $f_t(H_t; \ _t) = \ _t^T H_{t;0}$, where $H_{1;0} = (1; X_1^T)^T$ and $H_{2;0} = (1; X_1^T; A_1; X_1^T A_1; X_2^T)$. These rules are estimated using logistic loss, also known as the entropy loss, as the convex surrogate (Zhao et al., 2019a; Jiang et al., 2019).

To illustrate the double-robustness property of the AIPWCC estimator, we consider both correctly and incorrectly speci ed models for the hazard functions. In the correctly speci ed case, we t a logistic regression model at each stage with the correct features, i.e., $(1; X_{t1}; A_t X_{t2})$ for $t = 1; 2$. For the incorrectly speci ed model, we t a logistic regression model with features $(1; X_1; X_{11} X_{12})$ at stage 1 and $(1; X_{21}^2; X_{22}^2)$ at stage 2. The conditional mean model is not correctly speci ed throughout.

We evaluate the performance of an estimated optimal regime $\hat{b}$, in terms of its relative value, which is de ned as

$$RV(\hat{b}) = \frac{V(\hat{b}) \quad V(\ _0)}{V(\ _{opt}) \quad V(\ _0)};$$

where $_0$ is the stochastic policy that assigns treatments at each stage using a fair-sided coin ip. The reason for using the relative value, for instance, instead of the raw value, is to allow for comparison across a range of generative models, e.g., di erent problem dimensions. The requisite values are estimated using Monte Carlo methods with 40,000 simulated patients (see Section A.6. of Schulte et al., 2014, for additional details on simulation-based value estimation).

We approximate the roots of the estimating equations using R package nleqslv with multiple starts. We implement MI using R package MICE with default settings and 10 imputed datasets (Buuren & Groothuis-Oudshoorn, 2011). The conditional expectation model in the AIPWCC estimator is tted using ridge regression and tuned using the 5-fold cross-validation estimator of the value under the optimal regime.

The results for the correctly speci ed hazard models with Q-learning and outcome weighted learning are displayed in Figures 2.1 and 2.2, respectively. It can be seen that both the IPWCC and AIPWCC estimators generally outperform MI when the dimension of the covariates at

each stage, p, is large. The results for the incorrectly specified hazard models estimated using Q-learning and outcome weighted learning are displayed in Figures 2.3 and 2.4, respectively. The results are qualitatively similar though the robustness of the AIPWCC shows improved performance relative to the IPWCC in some scenarios.

2.5.2 CATIE Trial Analysis

We use data from the CATIE schizophrenia study (Stroup et al., 2003) to illustrate the proposed methods. The CATIE study is a SMART, which enrolled 1460 schizophrenia patients. This dataset was chosen in part because it was used as an illustrative case study with MI by others (Shortreed et al., 2011; Shortreed & Moodie, 2012; Laber et al., 2014c; Shortreed et al., 2014).

As done elsewhere in the literature, we compare two treatments of primary clinical interest at each stage: Perphenazine (coded -1) and Olanzapine (coded 1). The dataset consisting of 506 patients receiving these treatments , 46% of whom followed the entire course (i.e., they are complete cases), 34% dropped out after stage 1, and 20% dropped out after stage 2. The missingness pattern is shown in Figure 2.5. Item missingness was sparse (less than 2% of the observed data) and singly imputed using mean imputation.

The Positive and Negative Syndrome Scale (PANSS) score is the standard medical scale for measuring symptom severity in schizophrenia. This score is a time-varying variable and was measured at each stage: baseline (PANSS0); stage 1 (PANSS1); and stage 2 (PANSS2). A higher PANSS score is associated with more severe symptoms, so we use 100 - PANSS2 as the final outcome to match our convention of higher values representing better clinical outcomes. We include four baseline covariates in our analyses: PANNS0, baseline PANSS; EXACER, an indicator that the patient has been recently hospitalized; SEX; and TD, an indicator that the patient has Tardive Dyskinesia, a serious movement disorder associated with some antipscyhotic medications. In addition, we include PANNS1, first stage PANSS, in our second stage models. As in the simulation study, we used linear models for the Q-functions of Q-learning and linear

Figure 2.1 Relative value of Q-learning with IPWCC estimator, AIPWCC estimator and MI when the missingness model is correctly speci ed. The length of the corresponding vertical bar is the Monte Carlo standard deviation of the relative value estimates

Figure 2.2 Relative value of outcome weighted learning with IPWCC estimator, AIPWCC estimator and MI when the missingness model is correctly specified. The length of the corresponding vertical bar is the Monte Carlo standard deviation of the relative value estimates.

Figure 2.3 Relative value of Q-learning with IPWCC estimator, AIPWCC estimator and MI when the missingness model is misspeci ed. The length of the corresponding vertical bar is the Monte Carlo standard deviation of the relative value estimates.

Figure 2.4 Relative value of outcome weighted learning with IPWCC estimator, AIPWCC estimator and MI when the missingness model is misspecified. The length of the corresponding vertical bar is the Monte Carlo standard deviation of the relative value estimates.

Table 2.1 Cross-validated value estimates of the optimal regimes estimated using di erent methods.

	Q-MI	Q-IPW	Q-AIPW	OWL-MI	OWL-IPW	OWL-AIPW
$\hat{V}(b^{opt})$	35.406	40.684	41.147	48.220	49.366	48.164

Figure 2.5 Missiningness pattern of CATIE study after cleaning.

decision rules for outcome weighted learning.

We compared the following six approaches: Q-learning with MI (Q-MI), Q-learning with IPWCC (Q-IPW), Q-learning with AIPWCC (Q-AIPW), outcome weighted learning with MI (OWL-MI), outcome weighted learning with IPWCC (OWL-IPW) and outcome weighted learning with AIPWCC (OWL-AIPW). The cross-validated value for each approach is reported in Table 2.1. With the CATIE data, outcome weighted learning generally performed favorably to Q-learning. In terms of adjustment for missing data, MI appears to be worse than AIWPCC/IPWCC with Q-learning but about the same with outcome weighted learning.

2.6 Discussion

Missing data is essentially unavoidable with SMARTS and other longitudinal study designs commonly used to estimate optimal treatment regimes. Multiple Imputation has been shown to be an e ective tool for accommodating missing data in such studies. However, in some settings, imputation can involve modeling complex processes, which may be prone to misspeci cation and high variance. We examined the use of inverse probability weighted methods and showed such methods are consistent for a broad class of estimators of an optimal treatment regime. Furthermore, in empirical experiments, these methods outperformed imputation with the gap in performance widening with the increasing trajectory dimension as well as with the increasing amount of missingness. Thus, we recommend that such weighting methods be given serious consideration by researchers estimating optimal treatment regimes from randomized or observational studies. Alternatively, it may be bene cial to forgo choosing between imputation and weighting and instead combine them (e.g., Shortreed et al., 2014). We leave such a hybrid approach to future work.

CHAPTER 3

INTEGRATIVE ANALYSIS OF RANDOMIZED CLINICAL TRIALS WITH REAL WORLD EVIDENCE STUDIES

3.1 Introduction

Randomized clinical trials (RCTs) are the gold standard for evaluation of treatment e ects. However, due to restrictive inclusion and exclusion criteria for patient eligibility, the trial sample is narrowly de ned and can be systematically di erent from the real-world patient population to which the new treatment is supposed to be given. Therefore, the ndings from RCTs often lack generalizability to the target population of interest. Real world evidence (RWE) studies often include large samples that are representative of real-world patient populations; however, there are concerns about whether or not confounding has been addressed adequately in the analyses of RWE studies. In cancer research, there is an ongoing discussion on the strengths and limitations of utilizing data from RCT and RWE in comparative e ectiveness analyses (Hahn & Schilsky, 2012; Hershman & Wright, 2012; Korn & Freidlin, 2012). The cancer research community has reached a consensus on the need to incorporate evidence from real word data, but how to e ectively integrate them in pooled analysis of individual patients data from RCTs

remains under-developed.

There is considerable interest in bridging the findings from a RCT to the target population. This problem has been termed as generalizability (Cole & Stuart, 2010; Stuart et al., 2011; Hernan & VanderWeele, 2011; Tipton, 2013; O'Muircheartaigh & Hedges, 2014; Stuart et al., 2015; Keiding & Louis, 2016; Buchanan et al., 2018), external validity (Rothwell, 2005) or transportability (Pearl & Bareinboim, 2011; Rudolph & Laan, 2017) in the statistics literature and has connections to the covariate shift problem in machine learning (Sugiyama & Kawanabe, 2012). Most of the existing methods rely on direct modeling of the sampling score, which is the sampling analog of the propensity score (Rosenbaum & Rubin, 1983a). The subsequent sampling score adjustments include inverse probability of sampling weighting (IPSW, Cole & Stuart, 2010; Buchanan et al., 2018), stratification (Tipton, 2013; O'Muircheartaigh & Hedges, 2014), and augmented IPSW (Dahabreh et al., 2018). There are two major drawbacks in the sampling score adjustment approaches. Firstly, they require correct model specification of the sampling score. The IPSW estimators are unstable or inconsistent if the sampling score is too extreme or misspecified. Secondly, they assume the RWE sample to be a simple random sample from the target population, and implicitly require either the population size or all the baseline information of the population to be available. For example, Dahabreh et al. (2018) set the target population to be all trial-eligible individuals and assumed that all population baseline covariates are known, which is rarely the case in practice.

In this chapter, we consider combining a RCT sample and a RWE sample to estimate the average treatment effect (ATE) of the target population, where the RCT sample is subject to selection bias and the RWE sample is representative of the target population with a known sampling mechanism. It is worth noting that we allow the RWE sample to be a general random sample from the target population. This relaxation is particularly useful when real-world cohort studies are based on stratified random sampling in order to sample sufficient representations of some subgroups.

To address the selection bias of the RCT sample, we estimate the sampling score weights directly by calibrating covariate balance between the RCT sample and the design-weighted RWE sample, in contrast to the dominant approaches that focus on predicting sample selection probabilities. Calibration weighting (CW) is widely used to integrate auxiliary information in survey sampling (Wu & Sitter, 2001; Chen et al., 2002; Kott, 2006; Chang & Kott, 2008; Kim et al., 2016), and causal inference, such as in Constrained Empirical Likelihood (Qin & Zhang, 2007), Entropy Balancing (Hainmueller, 2012), Inverse Probability Tilting (Graham et al., 2012), and Covariate Balance Propensity Score (Imai & Ratkovic, 2014; Fan et al., 2016). Chan et al. (2015) showed that estimating ATE by empirical balancing calibration weighting can achieve global e ciency. To the best of our knowledge, our paper is the rst to use calibration in combining a RCT sample and a RWE sample to construct ATE estimators. We show that calibration weighting has several advantages in the data integration problem. First, its estimation does not require the population baseline covariates or the population size to be available, and it allows the RWE sample to be a general random sample, not necessarily a simple random sample, from the target population. Second, the weights are estimated directly from an optimization problem instead of inverting the estimated sampling probability, which requires careful monitoring to avoid extreme weights that result in highly variable estimates. Third, similar to Zhao & Percival (2017), we show that the calibration weighting estimators are doubly robust in the sense that the estimators are consistent if the parameterization for either the outcome or the sampling score model is correctly speci ed.

The e ciency of the CW estimators can be further improved if we have additional treatment and outcome information from the RWE sample. We derive the rst semiparametric e ciency bound for the ATE under identi cation assumptions and the outcome mean function transportability assumption. We propose an augmented CW estimator that is doubly robust and achieves the semiparametric e ciency bound if both nuisance models are correctly speci ed. In the presence of complex confounding, we adopt the method of sieves (Shen, 1997; Chen, 2007),

which allows flexible data-adaptive estimation of the nuisance functions while retaining usual root-n consistency under mild regularity conditions.

The rest of this chapter is organized as follows. We formalize the causal framework and assumptions in Section 3.2. The CW estimators are introduced in Section 3.3. We provide the semiparametric efficiency bound and propose the augmented CW estimator in Section 3.4. The finite sample performances of the proposed estimators are evaluated and compared in simulations studies in Section 3.5. We apply the proposed methods to estimate the effect of adjuvant chemotherapy by integrating the RCT data for early-stage resected non{small-cell lung cancer with the real-world data from the National Cancer Database in Section 3.6. Section 3.7 concludes with future directions.

3.2 Basic setup

3.2.1 Notation: causal effect and two data sources

We let X be the p-dimensional vector of covariates; let A be the treatment assignment with two levels $\{0, 1\}$, where 0 and 1 are the labels for control and active treatments, respectively; and let Y be the outcome of interest, which can be either continuous or binary. We use the potential outcomes framework (Neyman, 1923; Rubin, 1974) to formulate the causal problem. Under the Stable Unit Treatment Value Assumption (Rubin, 1980), for each level of treatment a, we assume that each subject in the target population has a potential outcome $Y(a)$, representing the outcome had the subject, possibly counterfactual, been given treatment a. The conditional average treatment effect (CATE) is defined as $\tau(X) = E\{Y(1) - Y(0) \mid X\}$. We are interested in estimating the population ATE $\tau_0 = E\{\tau(X)\}$, where the expectation is taken with respect to the target population. If the outcome is binary, the ATE is referred to as the average causal risk difference $\tau_0 = P\{Y(1) = 1\} - P\{Y(0) = 1\}$.

We consider a scenario where we have access to two samples: one sample is from a RCT

that compares the two treatments, and the other sample is from an observational RWE study that can be used to characterize the target population. The target population of size N consists of all patients with certain diseases to whom the new treatment is intended to be given. In practice, it is often difficult to identify the entire population, and the population size, N, is not necessarily known. Let $= 1$ denote RCT participation, and let $^{e} = 1$ denote the RWE study participation. The RWE sample is assumed to be a random sample drawn from the target population with a known sampling mechanism. Let $d = 1=P(^{e} = 1 \mid X)$ be the design weight in the RWE sample. For example, in health research, many cohort studies utilized stratified sampling to over-represent some subgroups. Suppose the population consists of L strata with sizes $f N_1; : : : ; N_L g$, and suppose that the RWE study selects a fixed number of subjects n_l from the N_l subjects in the lth stratum. Then, the design weight for subject j from the lth stratum is $d_j = N_l=n_l$. We denote data from the RCT of size n to be $f (X_i; A_i; Y_i; _i = 1) : i = 1; :::; ng$; the data from the RWE study of size m to be either $\overset{n}{(X_j; ^{e}_j = 1) : j = n + 1; : : : ; n + m}\overset{o}{}$ if we only have covariate information, or $\overset{n}{(X_j; A_j; Y_j; ^{e}_j = 1) : j = n + 1; : : : ; n + m}\overset{o}{}$ if treatment and outcome information are also available in the study. The data structure is demonstrated in Figure 3.1.

3.2.2 Identification assumptions

A fundamental problem in causal inference is that we can observe at most one of the potential outcomes for an individual subject. To identify the ATE from the observed data, we make the following assumptions throughout this chapter.

Assumption 1 (Consistency) The observed outcome is the potential outcome under the actual received treatment: $Y = AY (1) + (1 \quad A)Y (0)$.

Assumption 2 (Randomization and Transportability) (i) $Y (a) ? A \mid (X; \quad = 1)$ for $a = 0; 1$; and (ii) $E f Y (1) \quad Y (0) \mid X; \quad = 1 g = \quad (X)$.

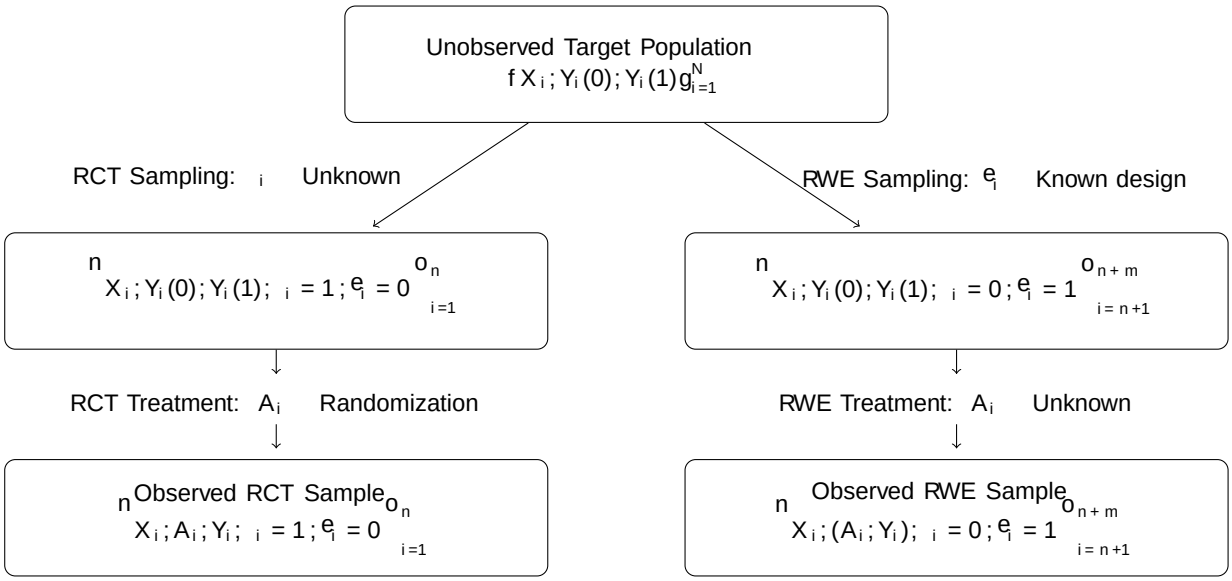

Figure 3.1 Demonstration of the sampling and treatment assignment regimes for the RCT and RWE samples within the target population.

Assumption 2 (i) holds for the RCT by default. Assumption 2 (ii) requires that the CATE function is transportable from the RCT to the target population. This assumption relaxes the ignorability assumption on trial participation (Stuart et al., 2011; Buchanan et al., 2018), i.e., $\{Y(0), Y(1)\} \perp S \mid X$, and the mean exchangeability assumption over treatment assignment and trial participation (Dahabreh et al., 2018), i.e. $E\{Y(a) \mid X, S = 1, A = a\} = E\{Y(a) \mid X, S = 1\}$ as well as $E\{Y(a) \mid X, S = 1\} = E\{Y(a) \mid X\}$ for $a = 0, 1$.

Moreover, we require adequate overlap of the covariate distribution between the trial sample and the target population, and also the treatment groups over the trial sample, formalized by the following assumption. Define the sampling score as $\pi(X) = P(S = 1 \mid X)$.

Assumption 3 (Positivity) There exists a constant c such that with probability 1, $\pi(X) > c > 0$; and $0 < P(A = 1 \mid X = x, S = 1) < 1$ for all x such that $P(X = x \mid S = 1) > 0$.

Under Assumptions 1{3, the ATE satis es

$$\begin{aligned}
{}_0 &= E\{ (X)g = E[E\{Y(1) \quad Y(0) \mid X; \quad = 1g] \\
&= E\left[\frac{}{(X)} \{E(Y \mid X; A = 1; \quad = 1) \quad E(Y \mid X; A = 0; \quad = 1)g \right] \\
&= E\left[\frac{}{(X)} \{E(Y \mid X; A = 1; \quad = 1) \quad E(Y \mid X; A = 0; \quad = 1)g \right];
\end{aligned}$$

and thus is identi able.

3.2.3 Existing estimation methods

Because the RCT assigns treatments randomly to the participants, (X) is identi able and can be estimated by standard estimators solely from the RCT. However, the covariate distribution of the RCT sample $f(X \mid = 1)$ is di erent from that of the target population $f(X)$ in general; therefore, $E\{(X) \mid = 1g$ is di erent from ${}_0$, and the ATE estimator using trial data only is biased of ${}_0$ generally.

A widely-used approach is the IPSW estimator that predicts the sampling score (X) and uses the inverse of the estimated sampling score to account for the shift of the covariate distribution from the RCT sample to the target population. Speci cally, most of the empirical literature assumes that (X) follows a logistic regression model $(X;)$ and can be estimated by $(X; b)$. The IPSW estimator of the ATE is

$$\hat{}^{IPSW} = \frac{\sum_{i=1}^{n} (X_i; b)^{1} A_i Y_i}{\sum_{i=1}^{n} (X_i; b)^{1} A_i} \quad \frac{\sum_{i=1}^{n} (X_i; b)^{1} (1 \quad A_i) Y_i}{\sum_{i=1}^{n} (X_i; b)^{1} (1 \quad A_i)}: \tag{3.1}$$

The IPSW estimator has several drawbacks as discussed in Section 3.1. In the next two sections, we propose approaches to (i) calibrate the covariate distribution of the trial sample to that of the design-weighted RWE sample so that the estimated treatment e ects can be generalized to the target population, and (ii) leverage the predictive power of the RWE sample to improve the precision for the generalized ATE estimator.

3.3 Calibration weighting estimators

We propose to use calibration originated in survey sampling to eliminate the selection bias in the trial-based ATE estimator. The calibration weighting approach is similar to the idea of entropy balancing weights introduced by Hainmueller (2012). We calibrate subjects in the RCT sample so that after calibration, the covariate distribution of the RCT sample empirically matches the target population. Our insight is based on the observation that for any vector-valued function $g(X)$,

$$E\left\{\frac{R}{e(X)}g(X)\right\} = E\{d\,g(X)\} = E\{g(X)\}.$$

Here, $g(X)$ contains the covariate functions to be calibrated, which could be moment functions of the original covariate X or any sensible transformations of X.

To this end, we assign a weight q_i to each subject i in the RCT sample so that

$$\sum_{i=1}^{N} q_i g(X_i) = \hat{\mu}_g, \tag{3.2}$$

where $\hat{\mu}_g = \sum_{i=1}^{N} e_i d_i g(X_i) / \sum_{i=1}^{N} e_i d_i$ is a design-weighted estimate of $E\{g(X)\}$ from the RWE sample. Constraint (3.2) is referred to as the balancing constraint, and weights $Q = \{q_i : i = 1\}$ are the calibration weights. The balancing constraint calibrates the covariate distribution of the RCT sample to the target population in terms of $g(X)$. The choice of $g(X)$ is important for both bias and variance considerations, which we will discuss in Section 3.4.2.

We estimate Q by solving the following optimization problem:

$$\min_{Q} \sum_{i=1}^{n} q_i \log q_i, \tag{3.3}$$

subject to $q_i \geq 0$, for all i; $\sum_{i=1}^{n} q_i = 1$, and the balancing constraint (3.2).

The objective function in (3.3) is the entropy of the calibration weights; thus, minimizing

this criteria ensures that the empirical distribution of calibration weights are not too far away from the uniform, such that it minimizes the variability due to heterogeneous weights. This optimization problem can be solved using convex optimization with Lagrange multiplier. By introducing Lagrange multiplier λ, the objective function becomes

$$L(\lambda; Q) = \sum_{i=1}^{n} q_i \log q_i - \lambda^{\top} \left(\sum_{i=1}^{n} q_i g(X_i) - g \right) : \tag{3.4}$$

Thus by minimizing (3.4), the estimated weights are

$$q_i = q(X_i; \beta) = \frac{\exp\{\beta^{\top} g(X_i)\}}{\sum_{i=1}^{n} \exp\{\beta^{\top} g(X_i)\}};$$

and β solves the equation

$$U(\beta) = \sum_{i=1}^{n} \exp\{\beta^{\top} g(X_i)\} \{g(X_i) - g\} = 0; \tag{3.5}$$

which is the dual problem to the optimization problem (3.3).

There are two general types of weighted estimators for population means, namely the Horvitz-Thompson estimator (HT, Horvitz & Thompson, 1952) and the Hajek estimator (Hajek, 1971). We propose both estimators with the superscript 0 for the HT estimator and the superscript 1 for the Hajek estimator. Let $\pi_{Ai} = P(A_i = 1 | X_i; \gamma_i = 1)$ be the treatment propensity score for subject i. For RCTs, it is common that the propensity score is known with $\pi_{Ai} = 0.5$, for all $i = 1; \ldots; n$.

Based on the calibration weights, we propose the HT CW estimator

$$\hat{\tau}_{CW0} = \sum_{i=1}^{n} q_i \left(\frac{A_i Y_i}{\pi_{Ai}} - \frac{(1 - A_i) Y_i}{1 - \pi_{Ai}} \right); \tag{3.6}$$

and the Hajek CW estimator

$$\hat{\tau}^{CW1} = \sum_{i=1}^{n} \hat{q}_i \left(\frac{A_i Y_i}{\sum_{i=1}^{n} A_i \hat{q}_i} - \frac{(1 - A_i) Y_i}{\sum_{i=1}^{n} (1 - A_i) \hat{q}_i} \right) : \tag{3.7}$$

To investigate the properties of the proposed CW estimators, we impose the following regularity conditions on the sampling designs for both the RWE and the RCT samples.

Assumption 4 Let $\mu_{g0} = E\{g(X)\}$. The design weighted estimator $\hat{\mu}_g = N^{-1} \sum_{i=1}^{N} e_i d_i g(X_i)$ satisfies $V(\hat{\mu}_g) = O(m^{-1})$, and $\{V(\hat{\mu}_g)\}^{-1/2}(\hat{\mu}_g - \mu_{g0}) \to N(0,1)$ in distribution, as $m \to \infty$.

Assumption 5 The sampling score of RCT participation follows a loglinear model, i.e. $\pi(X) = \exp\{\gamma_0^\top g(X)\}$ for some γ_0.

Note that if the sampling score follows a logistic regression model $\pi(X;\gamma) = \exp\{\gamma^\top g(X)\}/[1+\exp\{\gamma^\top g(X)\}]$ and the sampling fraction of the RCT is small, the loglinear model in Assumption 5 is close to the logistic regression model. Our simulation studies demonstrate that the proposed estimators perform well under a logistic regression model.

In addition, in the estimation of calibration weights we only require specifying $g(X)$. Thus, CW evades explicitly modeling either the sampling score model or the outcome mean models. Under Assumption 5, we show that there is a direct correspondence between calibration weight $q(X_i;\beta)$ and the estimated sampling score $\pi(X_i;\beta)$, i.e. $q(X_i;\beta) = \{N\pi(X_i;\beta)\}^{-1} + o_p(N^{-1})$; see the proof of Theorem 3 in the Appendix.

The following assumption is on the linearity of the CATE in $g(X)$.

Assumption 6 $\tau(X) = \gamma_0^\top g(X)$.

Based on the above assumptions, we establish the double robustness property of the CW estimators in the following theorem and relegate all proofs to the Appendix.

Theorem 3 (Double robustness of the CW estimators) Under Assumptions 1–4, if either Assumption 5 or Assumption 6 holds, not necessarily both, $\hat{\tau}^{CW0}$ and $\hat{\tau}^{CW1}$ in (3.6) and (3.7) are consistent for τ_0.

The Hajek-type estimator often improves the HT-type estimator empirically (Särndal et al., 2003). However, this improvement is not insured in our context. We compare their asymptotic variances in the proof of Theorem 3 to show that either one could outperform the other in certain scenarios; see Section B.2.1 for details.

We provide a simple bootstrap procedure to estimate the variance of the CW estimators. The first step is to draw B bootstrap samples from both the RCT sample and the RWE sample respectively, which results in B pairs of bootstrapped samples. Then for each resampled pair, we can obtain a replicate of the CW estimator. The variance of the CW estimator is given by the sample variance of the B bootstrapped estimators.

3.4 Semiparametric efficient estimator when Y and A are available in RWE

3.4.1 Augmented calibration weighting estimator

We have utilized the design weighted covariate distribution of the RWE sample to adjust for selection bias of the RCT sample. Now we consider the setting where we have access to additional treatment and outcome information (Y, A) from the RWE sample.

Define the trial conditional outcome mean function as $\mu_a(X) = E(Y \mid X, A = a, \tilde{} = 1)$ for $a = 0, 1$. To leverage the predictive power of the RWE sample, we assume the transportability of $\mu_a(X)$ from the RCT sample to the RWE sample, as formulated below.

Assumption 7 For $a = 0, 1$, $E(Y \mid X, A = a, \tilde{} = 1) = \mu_a(X)$.

Assumption 7 is plausible if X includes a rich set of covariates that influence the outcome, and is also testable because it is based only on the observed data.

The following theorem gives the semiparametric efficiency bound for β_0 in our data integration setting.

Theorem 4 (Semiparametric efficiency bound) Under Assumptions 1 - 4 and 7, the semiparametric efficiency score for β_0 is

$$\varphi(X; A; Y; \pi; e) = \frac{1}{\tau(X)} \left\{ \frac{A\{Y - \mu_1(X)\}}{\pi} - \frac{(1-A)\{Y - \mu_0(X)\}}{1-\pi} \right\} + \{\tau(X) - \tau_0\}.$$

The semiparametric efficiency bound for β_0 is

$$V_e = E\left[\frac{1}{\tau(X)^2} \left\{ \frac{V\{Y(1)|X\}}{\pi} + \frac{V\{Y(0)|X\}}{1-\pi} \right\} + \{\tau(X) - \tau_0\}^2 \right].$$

The semiparametric efficiency bound is decomposed naturally into two components as shown in its form. The first term accounts for the variance of potential outcomes adjusted for treatment assignment and sample selection; the second term quantifies the treatment heterogeneity in the RWE sample. Rudolph & Laan (2017) also provided a semiparametric efficiency score for transporting the ATE from one study site to another, where one site is regarded as a population without the sampling assumption.

The result in Theorem 4 serves as a foundation to derive efficient estimators combining two data sources. The score $\varphi(X; A; Y; \pi; e)$ has unknown nuisance functions $\tau(X)$ and $\mu_a(X)$, $(a = 0; 1)$. Therefore, to estimate β_0, we posit models for the nuisance functions, denoted by $\tau(X; \gamma)$ and $\mu_a(X; \alpha_a)$. For example, we assume $\tau(X)$ is a loglinear model as in Assumption 5. By the correspondence between the loglinear model and the calibration weighing algorithm, we can estimate β_0 following the optimization algorithm in (3.3). We also posit models $\mu_a(X; \alpha_a)$; $a = 0; 1$. By Assumption 7, we are able to obtain a consistent estimator $\hat{\alpha}_a$ based on the RWE sample.

Based on the semiparametric efficiency score, we propose a new estimator for the ATE. As the

outcome mean models in the semiparametric e ciency score can be viewed as an augmentation to the CW estimator, we refer to the proposed estimator as the augmented calibration weighting (ACW) estimator, and it is given by

$$\hat{\tau}^{ACW} = \sum_{i=1}^{N} \omega_i \left\{ \frac{A_i \left[Y_i - \mu_1(X_i; \hat{\beta}_1) \right]}{A_i} - \frac{(1 - A_i) \left[Y_i - \mu_0(X_i; \hat{\beta}_0) \right]}{1 - A_i} \right\}$$
$$+ \left(\sum_{i=1}^{N} e_i d_i \right)^{-1} \sum_{i=1}^{N} e_i d_i \left[\mu_1(X_i; \hat{\beta}_1) - \mu_0(X_i; \hat{\beta}_0) \right]. \quad (3.8)$$

We show in the following theorem that $\hat{\tau}^{ACW}$ achieves double robustness and local e ciency. For a vector v, we use $\|v\|_2 = (v^\top v)^{1/2}$ to denote its Euclidean norm. For a function $f(V)$, where V is a generic random variable, we de ne its L_2-norm as $\|f(V)\| = \left\{ \int f(v)^2 dP(v) \right\}^{1/2}$.

Theorem 5 (Double robustness and local e ciency of the ACW estimator) Under Assumptions 1–4 and 7, if either Assumption 5 or Assumption 6 holds, not necessarily both, $\hat{\tau}^{ACW}$ is consistent for τ_0. When both Assumption 5 and Assumption 6 hold, $N^{1/2}(\hat{\tau}^{ACW} - \tau_0) \to N(0, V_e)$ in distribution, as $n \to \infty$, where V_e is de ned in Theorem 4, i.e. $\hat{\tau}^{ACW}$ is locally e cient with its asymptotic variance achieves the semiparametric e ciency bound.

By the empirical processes theory, the e ect of nuisance parameter estimation in $\hat{\tau}^{ACW} - \tau_0$ is bounded by $\|\pi(X; \beta) - \pi(X)\| \sum_{a=0}^{1} \|\mu_a(X; \hat{\beta}_a) - \mu_a(X)\|$; see Section B.2.4 in the Appendix for details. If this bound is of rate $o_p(n^{-1/2})$, then it is asymptotically negligible, and thus $\hat{\tau}^{ACW}$ is semiparametric e cient. This rate can occur in various scenarios. For example, if all nuisance models are correctly speci ed parametric models, both $\|\pi(X; \beta) - \pi(X)\|$ and $\sum_{a=0}^{1} \|\mu_a(X; \hat{\beta}_a) - \mu_a(X)\|$ are $o_p(n^{-1/2})$ and their product is $o_p(n^{-1})$, then $\hat{\tau}^{ACW}$ is locally e cient. Another instance is to approximate the nuisance functions by exible machine learning methods. In such case, if both $\|\pi(X; \beta) - \pi(X)\|$ and $\sum_{a=0}^{1} \|\mu_a(X; \hat{\beta}_a) - \mu_a(X)\|$ reach $o_p(n^{-1/4})$, then their product still achieves $o_p(n^{-1/2})$, and therefore is asymptotically negligible. In general, there exist di erent combinations of convergence rates of $\pi(X; \beta)$ and $\mu_a(X; \hat{\beta}_a)$ $(a = 0, 1)$

that result in a negligible error bound accommodating di erent smoothness conditions of the underlying true nuisance functions. The following theorem formalizes the above statement.

Theorem 6 Let $(X; b)$ and $_a(X; b_a)$ $(a = 0; 1)$ be general semiparametric models for (X) and $_a(X)$ $(a = 0; 1)$, respectively. Assume the following regularity conditions hold:

Condition 1 $k \quad (X; b) \quad (X)k = o_p(1)$, and $k \ _a(X; b_a) \quad _a(X)k = o_p(1)$, for $a = 0; 1$;

Condition 2 $k \quad (X; b) \quad (X)k \ \sum_{a=0}^{1} k \ _a(X; b_a) \quad _a(X)k = o_p(n^{-1=2})$.

Then $\hat{}^{ACW}$ is consistent for $_0$ and achieves the semiparametric e ciency bound.

The semiparametric e ciency bound is attained as long as either b or $(b_0; b_1)$ approximate the underlying sampling score model or the outcome models well. Condition 1 states that we require that the posited models to be consistent. Condition 2 states that the combined rate of convergence of the posited models is of $o_p(n^{-1=2})$. In Section 3.4.2, we construct such estimators using the method of sieves, which satis es Condition 1 and 2 under mild regularity conditions.

For locally e cient estimator $\hat{}^{ACW}$, the variance estimator can be calculated empirically as

$$\hat{V}^{ACW} = \sum_{i=1}^{N} \left[\frac{i q_i}{(X_i; b)} \left(\frac{\hat{V}(Y_i|X_i; A_i = 1)}{A} + \frac{\hat{V}(Y_i|X_i; A_i = 0)}{1 \quad A} \right) + \left(\sum_{i=1}^{N} e_i d_i \quad e_i d_i^2 f\hat{}(X_i) \quad \hat{}g^{25} \right)^{1} \right]^{3} : \quad (3.9)$$

The bootstrap variance estimator introduced in Section 3.3 can be applied here as well. The bootstrap variance estimator is more straightforward than (3.9), and it can accommodate situations where either one of the nuisance models is misspeci ed. Thus, the bootstrap variance estimator is recommended in practice and adopted in the simulation study.

3.4.2 Semiparametric models by the method of sieves

To overcome the model misspeci cation issue inherent to parametric models, we consider the method of sieves (Geman & Hwang, 1982), which allows exible models for (X) and $_a(X); (a = 0; 1)$ and also achieves Conditions 1 and 2 under mild regularity conditions. In comparison with other nonparametric methods such as the kernel method, the sieve method is particularly well-suited to calibration weighting. Although general sieve basis functions such as Fourier series, splines, wavelets, and arti cial neural networks (see Chen, 2007, for a comprehensive review) are applicable, the power series is the most common class of functions. For a p-vector of non-negative integers $= (_1; : : : ; _p)$, let $j j = \sum_{l=1}^{p} _l$ and $X = \prod_{l=1}^{p} X_l^{\ l}$: De ne a series $f (k) : k = 1; 2; : : :g$ for all distinct vectors of such that $j (k)j \quad j (k + 1)j$. Based on this series, we consider a K-vector $g(X) = f g_1(X); : : : ; g_K(X)g^> = f X^{(1)}; : : : ; X^{(K)}g^>$.

To accommodate di erent type of variables, we approximate (X) and $_a(X)$ by the generalized sieves functions

$$\log (X;) = ^> g(X); \quad _a(X; _a) = m_a f _a^> g(X)g \ (a = 0; 1);$$

where for a continuous outcome, $m_a( )$ is the identity link function, and for a binary outcome, $m_a( )$ is the expit (inverse logit) link function with

$$^> = \arg\min E[(X) \quad \exp f ^> g(X)g]^2;$$

$$_a^> = \arg\min E[_a(X) \quad m_a f ^> g(X)g]^2; \ (a = 0; 1):$$

We assume that (X) and $_a(X); (a = 0; 1)$ are su ciently smooth, with (x) s-times continuously di erentiable, and $_a(x)$ s_a-times continuously di erentiable for any x in the support of X with a condition of $\min(s ; s_0; s_1) > 2p$. Under standard regularity conditions speci ed in Section B.3 in the Appendix, the deterministic di erences between the true functions

and the sieves approximations are bounded (Newey, 1997)

$$\sup_{x \in X} | \quad (x) \quad \exp\{ \quad ^{>} g(x)\} | = O\left(K^{1 \quad s \ =(2p)} \right); \tag{3.10}$$
$$\sup_{x \in X} | \ _a(x) \quad m_a\{ \ _a^{>} g(x)\} | = O\left(K^{1 \quad s \ _a=(2p)} \right); \ (a = 0; 1):$$

Thus, the approximation errors can be made su ciently small by choosing a large K. However, it is also necessary that K increases slowly with the sample size n to control the variance of the sieves estimators. Formally, for the sieves estimators to be consistent, K should satisfy that $1=K + K=n \ ! \ 0$ as K and n grow (Chen, 2007). Therefore, in the presence of data with complex confounding, choosing K becomes important. In spite of the importance, there has limited study of the selection of K in sieves estimation. Imbens et al. (2005) proposed a selection method through minimizing the mean-squared-error (MSE) of the ATE over a pre-de ned candidate set of K, which requires a rather complicated estimation of the population MSE and is not automatic.

To cope with this issue, we propose a new basis (variable) selection procedure for sieves estimation of (X) and $_a(X)$. The number of basis functions controls the smoothness of sieves estimators. From this viewpoint, we can specify a su ciently large K as an initial number and apply the penalization technique to regularize the variability of the estimators.

For penalized sieves estimation of (X), the key insight is that the dual problem of calibration leads to solving a system of estimating equations given by equation (3.5). Therefore, we adopt the penalized estimating equation approach (Johnson et al., 2008; Wang et al., 2012a) to facilitate basis selection. We consider solving the penalized estimating equations

$$U(\ ) = U(\ ) \quad q(j \ j)\text{sign}(\ );$$

for $= (\ _1; : : : ; \ _K)^T$, where $q(j \ j) = f q(j \ _1 j); : : : ; q(j \ _K j)g^{>}$ is some continuous function, $q(j \ j)\text{sign}(\ )$ is the element-wise product of $q(j \ j)$ and $\text{sign}(\ )$. We let $q(x) = d \, p(x)=dx$,

where $p_\lambda(x)$ is some penalization function. In this paper, we specify $p_\lambda(x)$ to be a folded-concave SCAD penalty function (Fan & Li, 2001), although the same discussion applies to different penalty functions, such as adaptive LASSO (Zou, 2006). Accordingly, for the SCAD penalty, we have

$$q_\lambda(|\beta_k|) = \lambda I(|\beta_k| < \lambda) + \frac{(b\lambda - |\beta_k|)_+}{(b-1)} I(|\beta_k| \geq \lambda), \quad k = 1, \ldots, K,$$

and we specify $b = 3.7$ following the suggestion of the literature.

Implicitly, we turn the problem of choosing the number of basis functions K to the problem of choosing the tuning parameter λ. The tuning parameter λ controls the magnitude of the regularization. To help understand the penalized estimating equation, we discuss two scenarios. If $|\beta_k|$ is large, then $q_\lambda(|\beta_k|)$ is zero, and the kth component of the estimating equation $U_k(\beta)$ is not penalized. Whereas, if $|\beta_k|$ is small but not zero, then $q_\lambda(|\beta_k|)$ is nonzero and $U_k(\beta)$ is penalized. Consequently, the penalty term forces $\hat{\beta}_k$ to be zero and excludes the kth element in $g(X)$ from the final selected set of variables.

In the simulation studies, we use 5-fold cross validation to select tuning parameter λ. The loss function is specified to be $l(\lambda) = \sum_{i=1}^{n} q(X_i; \lambda) \| g(X_i) - \bar{g} \|_2 / K$, which quantifies the average degree of covariate balancing with calibration weights $q(X_i; \lambda), i = 1, \ldots, n$. Under standard regularity conditions, the estimator that solves the penalized estimating equation with the SCAD penalty would possess the oracle property (see Wang et al., 2012a, for details). Therefore, with $\lambda = \hat{\lambda}$, we can obtain $\| \hat{\beta} - \beta \|_2 = O_p(n^{-1/2})$.

For penalized sieves estimation of $\alpha_a(X)$, we can apply the standard penalization technique for regression models with the pre-specified basis functions based on the RWE sample. Specifically, let

$$\hat{\beta}_a = \arg\min_{\beta \in R^K} \left(\sum_{i=1}^{N} e_i d_i I(A_i = a) \left[Y_i - m_a^{-1} \{ g(X)^\top \beta \} \right]^2 + \sum_{j=1}^{K} p_{\lambda_a}(|\beta_j|) \right),$$

where $p_{\lambda_a}(\cdot)$ is the SCAD penalty function, for $a = 0, 1$. Under certain regularity conditions given in Fan & Li (2001), $\hat{\beta}_a$ satisfies the selection consistency and oracle properties under

penalized likelihood for both linear regression or logistic regression, i.e. $\|\hat{\beta}_a - \beta_a\|_2 = O_p(n^{-1/2})$, for $a = 0, 1$.

Assuming $\min(s, s_0, s_1)/(2p) > 3/4$ in (3.10), the penalized sieves estimators of $\pi(X)$ and $\mu_a(X)$ satisfies the two conditions in Theorem 6. Therefore, $\hat{\tau}^{ACW}$ with flexible approximation of the two nuisance functions still achieves root-n consistency and is semiparametric efficient.

3.5 Simulation

In this section, we evaluate the finite sample performances of the CW estimators and the ACW estimator via a set of simulation experiments. We first generate a target population of size $N = 50000$. Covariate $X \in R^4$ is generated from $X_j \sim N(1, 1)$ for each $j = 1, \ldots, 4$. For continuous outcomes, the potential outcome model is

$$Y(a)|X = 100 + 27.4aX_1 + 13.7X_2 + 13.7X_3 + 13.7X_4 + \varepsilon,$$

where $\varepsilon \sim N(0, 1)$ for $a = 0, 1$. Therefore, the true ATE under this setting is $\tau_0 = 27.4E(X_1) = 27.4$. For binary outcomes, we generate potential outcome according to

$$Y(a)|X \sim \text{Bernoulli}\{p_a(X)\},$$

where

$$\text{logit}\{p_a(X)\} = 1 - 2aX_1 - X_2 - X_3 + X_4, \tag{3.11}$$

under which the average causal risk difference is $\tau_0 = -0.24$. We generate the indicator of selection into the RCT sample according to $\delta|X \sim \text{Bernoulli}\{\pi(X)\}$, where $\text{logit}\{\pi(X)\} = -2.5 - 0.5X_1 - 0.3X_2 - 0.5X_3 - 0.4X_4$. By this design, the RCT selection rate is around 2%, which results in a roughly $n = 1000$ subjects in the RCT sample. The treatment assignment in the RCT sample is $A \sim \text{Bernoulli}(0.5)$. For the remaining subjects in the population, we take a random

sample of size $m = 5000$ to form a RWE sample. For patients in the RWE sample, treatment assignment is $A \mid X \sim \text{Bernoulli}\{e_A(X)\}$, where $\text{logit}\{e_A(X)\} = -X_1 + 0.5X_2 - 0.25X_3 - 0.1X_4$. The actual observed outcome Y in both samples is generated by $Y = AY(1) + (1-A)Y(0)$.

To study the impact of model misspecification, following Kang & Schafer (2007), we define a nonlinear transformation of X to be

$$X^* = (\exp(X_1/3), X_2/\{1+\exp(X_1)\} + 10, (X_1X_3/25 + 0.6)^3, (X_1 + X_4 + 20)^2)^\top,$$

and further scale and center X^* such that $E(X_j) = 1$ and $V(X_j) = 1$, for $j = 1, \ldots, 4$. Throughout, we use X for fitting models. We assume X^* to be unobserved, but can be used in the true generative models, in which cases the fitted models are misspecified. We compare the following four model specification scenarios:

- Scenario 1 (O:C/S:C): both outcome and sampling score models are correctly specified;
- Scenario 2 (O:C/S:W): the outcome model is correctly specified; the sampling score model is incorrectly specified by using X^* in the generative model;
- Scenario 3 (O:W/S:C): the outcome model is incorrectly specified by using X^* in the generative model; the sampling score model is correctly specified;
- Scenario 4 (O:W/S:W): both outcome model and sampling score models are incorrectly specified by using X^* in the generative model.

To demonstrate the double robustness of the ACW estimator against parametric model misspecification, we consider the calibration variables $g_1(X) = (X_1, X_2, X_3, X_4)^\top$ in all four scenarios. Moreover, we consider the ACW estimator using sieves estimation, with an initial large number of basis functions. Specifically, we extend the basis functions in $g_1(X)$ to its second order power series that includes all two-way interaction terms and quadratic terms, i.e. $g_2(X) = (X_1, \ldots, X_p, X_1X_2, \ldots, X_{p-1}X_p, X_1^2, \ldots, X_p^2)^\top$. For $p = 4$, $g_2(X)$ contains 14 basis functions.

We compare the following estimators for ATE:

1. Naive: the di erence in sample means of the two treatment groups in the RCT sample to demonstrate the degree of selection bias;
2. IPSW: the inverse probability of sampling weighting estimator de ned by (3.1), where the sampling weights estimated by logistic regression;
3. CW0: the HT calibration weighting estimator de ned by (3.6) with $g(X) = g_1(X)$;
4. CW1: the Hajek calibration weighting estimator de ned by (3.7) with $g(X) = g_1(X)$;
5. ACW: the augmented calibration weighting estimator de ned by (3.8) with $g(X) = g_1(X)$;
6. ACW(S): the penalized augmented calibration weighting estimator using the method of sieves with $g(X) = g_2(X)$.

We use bootstrap variance estimation for all estimators with $B = 50$. All simulations are all based on 1000 Monte Carlo replications.

Table 3.1 and Figure 3.2 summarize the results for continuous outcome; Table 3.2 and Figure 3.3 summarize the results for binary outcome. It can be seen that the CW estimators are doubly robust compared to the IPSW estimator for continuous outcome. The CW estimators for binary outcome demonstrate some bias in Scenario 2. This phenomenon is because of Assumption 6 does not hold under the data generating mechanisms(3.11). With additional information on $(A; Y)$ in the RWE sample, the ACW estimator is shown to be doubly robust and more e cient than the IPSW estimator and the CW estimators. In Scenario 4, where both outcome and sampling score models are misspeci ed, we show that the ACW estimator using the method of sieves is still unbiased and e cient. Moreover, the empirical coverage rates for the unbiased ACW estimator are close to the nominal level. As the sample size increases, which we show in the Appendix, they become closer to the nominal level.

Table 3.1 Simulation results for continuous outcome: bias of point estimates, Monte Carlo variance, relative bias of bootstrap variance estimate, and the coverage rate of 95% Wald confidence interval.

	Naive	IPSW	CW0	CW1	ACW	ACW(S)
	Bias					
1. O:C/S:C	13.03	0.00	0.15	0.20	0.16	0.15
2. O:C/S:W	11.45	0.93	0.10	0.17	0.12	0.12
3. O:W/S:C	10.99	0.08	0.19	0.13	0.18	0.21
4. O:W/S:W	10.30	1.92	1.53	1.61	1.55	0.00
	Monte Carlo variance					
1. O:C/S:C	3.45	9.41	1.81	8.13	0.15	0.15
2. O:C/S:W	2.94	5.89	1.54	5.32	0.15	0.15
3. O:W/S:C	3.34	14.25	6.07	11.98	4.78	0.72
4. O:W/S:W	3.00	5.94	1.69	5.42	0.98	0.31
	Relative bias (%) of bootstrap variance estimate					
1. O:C/S:C	3.7	0.2	2.6	7.7	0.1	1.4
2. O:C/S:W	4.2	3.8	0.3	1.0	2.3	3.2
3. O:W/S:C	6.9	4.1	3.0	4.5	5.6	0.3
4. O:W/S:W	1.0	2.4	8.2	2.9	5.9	10.6
	95% Wald CI coverage rate					
1. O:C/S:C	0.0	93.9	94.1	93.2	94.1	93.6
2. O:C/S:W	0.0	93.4	95.6	95.5	94.3	95.7
3. O:W/S:C	0.0	94.7	94.3	93.6	94.3	92.8
4. O:W/S:W	0.0	85.3	80.1	87.4	59.5	92.1

Figure 3.2 Boxplot of estimators for continuous outcome under four model speci cation scenarios.

Table 3.2 Simulation results for binary outcome: bias of point estimates, Monte Carlo variance, relative bias of bootstrap variance estimate, and the coverage rate of 95% Wald con dence interval.

	Naive	IPSW	CW0	CW1	ACW	ACW(S)
			Bias 1000			
1. O:C/S:C	91:6	0:4	1:4	1:9	1:9	1:7
2. O:C/S:W	83:4	7:1	17:1	16:4	1:5	2:0
3. O:W/S:C	77:0	0:1	1:4	1:9	2:1	2:0
4. O:W/S:W	72:7	-25:1	29:8	29:1	19:8	4:4
			Monte Carlo variance 1000			
1. O:C/S:C	0:92	1:51	1:42	1:44	0:92	0:97
2. O:C/S:W	0:81	1:19	1:23	1:22	0:85	1:02
3. O:W/S:C	0:90	1:56	1:86	1:51	1:25	1:24
4. O:W/S:W	0:90	1:31	1:78	1:32	1:19	1:21
			Relative bias (%) of bootstrap variance estimate			
1. O:C/S:C	2:6	1:7	6:9	0:3	1:1	2:4
2. O:C/S:W	12:0	10:9	5:4	5:9	0:7	2:6
3. O:W/S:C	0:6	0:1	0:2	1:7	1:5	0:2
4. O:W/S:W	2:4	3:8	1:6	0:7	1:5	2:4
			95% Wald CI coverage rate			
1. O:C/S:C	15:4	94:0	95:1	93:2	94:5	94:6
2. O:C/S:W	20:6	94:5	92:6	90:8	94:6	93:1
3. O:W/S:C	26:6	93:8	93:8	93:3	94:8	94:5
4. O:W/S:W	31:7	89:6	87:3	85:9	89:3	95:3

Figure 3.3 Boxplot of estimators for binary outcome under four model speci cation scenarios.

Table 3.3 Number of patients by treatment of the CALGB 9633 trial sample and the NCDB sample.

	Observation ($A = 0$)	Adjuvant chemotherapy ($A = 1$)	Total
RCT: CALGB 9633	163	156	319
RWE: NCDB	10936	4271	15207

3.6 Real data application

We apply the proposed estimators to evaluate the effect of adjuvant chemotherapy for early-stage resected non-small cell lung cancer (NSCLC). Adjuvant chemotherapy for resected NSCLC is shown to be effective in stages II and IIIA disease on the basis of RCTs (Massarelli et al., 2003); however, its utility in the early-stage disease remains unclear. Cancer and Leukemia Group B (CALGB) 9633 is the only trial designed specifically for stage IB NSCLC (Strauss et al., 2008). Nonetheless, it consists of only 319 patients, which is undersized to detect clinically meaningful improvements. National Cancer Database (NCDB), as another data source, is a large clinical oncology database that contains information of more than 70% of the newly diagnosed cancer patients in the US with more than 34 million historical records.

The comparable sample from the NCDB includes 15207 patients diagnosed with NSCLC between years 2004 - 2016 with stage IB disease who first had surgery and then received either adjuvant chemotherapy or on observation (i.e. no chemotherapy) and with age greater than 20; these patients also did not receive any of the neoadjuvant chemotherapy, radiation therapy, induction therapy, immunotherapy, hormone therapy, transplant/endocrine procedures, or systemic treatment before their surgery. Thus, the patients in both samples are stage IB NSCLC patients who either received adjuvant chemotherapy or were on observation after their surgery without any other interventions.

In our analysis, the treatment indicator A is coded as 1 for adjuvant chemotherapy and 0 for on observation. A summary of the sample sizes in both samples by treatment groups is given in Table 3.3. We include four covariates in the analysis: X_1 is gender (1 = male, 0 = female); X_2 is

age; X_3 is the indicator for histology (1 = squamous, 0 = non-squamous); X_4 is the tumor size in cm. The outcome is the indicator of cancer recurrence within 3 years after the surgery, i.e. $Y = 1$ if recurrence occurred and $Y = 0$ otherwise. Table 3.4 reports the covariate and outcome means of the two samples. It can be seen that the patients in the CALGB 9633 trial are healthier compared to the population represented by the NCDB sample, i.e. the trial patients are younger and have smaller tumor sizes.

We compare the same six methods as in the simulation studies with the same set of basis functions. Bootstrap variance estimation is applied to estimate the standard errors. In the estimation we standardize age by its mean 67.7 and standard deviation 10.2, and tumor size by its mean 4.9 and standard deviation 3.0. Table 3.5 reports the results.

The results indicate that in the RCT sample there is a 8.3% decrease in the risk of recurrence for adjuvant chemotherapy over observation. The IPSW, CW, and ACW estimators, which utilized RWE sample information, show a 9%-11% decrease in the risk of recurrence. However, the causal effect is not significant according to the 95% confidence interval. The ACW estimator with the method of sieves gives an estimate of 17% risk decrease; the risk difference is significant at 0.05 level. The difference between the nonparametric ACW estimator and other adjusted estimators is because the former has second order polynomials selected and used for both calibration and outcome regression, which also indicates that the sampling score or/and outcome mean functions used in IPSW, CW, and ACW estimators are likely to be misspecified. The bootstrap variance estimates are not distinctive from one another, which is consistent with our simulation results that there is no clear improvement in efficiency for binary outcomes. All of the RWE-adjusted estimators have deeper decline in recurrence risk compared to the trial-based Naive estimator, which suggests that the causal risk difference in the target population is larger than the one of the RCT sample, i.e. the effect of adjuvant chemotherapy is more profound in the real world population.

Table 3.4 Covariate and outcome means comparison of the CALGB 9633 trial sample and the NCDB sample.

	Gender (X_1)	Age (X_2)	Histology(X_3)	Tumor size(X_4)	Recurrence (Y)
CALGB 9633	0.64	60.83	0.40	4.60	0.25
NCDB	0.55	67.87	0.39	4.94	0.33

Table 3.5 Point estimate, standard error and 95% Wald con dence interval of the causal risk di erence between adjuvant chemotherapy and observation based on the CALGB 9633 trial sample and the NCDB sample.

	Est.	S.E	95% Wald con dence interval
Naive	0.083	0.048	(0.177, 0.011)
IPSW	0.088	0.060	(0.205, 0.029)
CW0	0.107	0.062	(0.227, 0.014)
CW1	0.105	0.061	(0.225, 0.014)
ACW	0.110	0.065	(0.236, 0.017)
ACW(S)	0.172	0.082	(0.333, 0.011)

3.7 Concluding remarks

In this chapter, we have developed a new semiparametric framework to evaluate the average treatment e ects integrating the complementary features of the RCT and RWE study. There are several directions for future work: (i) we will extend this framework to the setting with survival outcomes. Additional challenges arises due to possible right censoring. We will follow the technique in Bai et al. (2013) to establish the corresponding semiparametric e ciency theory and estimators in the context of data integration; (ii) RCTs are often underpowered for the heterogeneity of treatment e ect (HTE); on the contrary, RWE studies provide rich information on individual treatment responses (see Wu et al., 2018; Zhao et al., 2019b, for such developments in precision medicine). We will develop a semiparametric framework for integrative analyses of the HTE utilizing randomization in the RCT and power of treatment response prediction in the RWE study.

CHAPTER 4

MODEL BUILDING WITH STRUCTURED NEURAL NETWORKS

4.1 Introduction

Neural networks (NNs) have attracted much attention in the areas of machine learning and arti cial intelligence because of their high predictive power and computational scalability to large datasets; however, often being referred to as black-box models, a key hindrance of using NNs in practice is their lack of interpretability. Model explainability/interpretability is curcial and is even mandated by law in some settings (e.g. EU law on explainable AI, Koszegi, 2019). The restriction of using machine learning models is especially reinforced in areas with strict regulations, such as the banking industry. The US Federal Reserve has been advocating the development of explainable machine learning tools for the use in nancial services (Brainard, 2018).

To balance the two competing objectives, model performance and interpretability, it is favorable to use semiparametric regression models that contain both parametric and nonparametric components in their model structures. The parametric components help in model interpretation, and the nonparametric components add exibility to the model so that it can capture the

underlying complex features of data. We survey a list of popular semiparametric regression models in econometrics and machine learning: the single index model (SIM, Ichimura, 1993; Hardle et al., 1993; Xia & Li, 1999); the (generalized) additive model (AM, Hastie & Tibshirani, 1987; Hastie, 2017) and its regularized versions (Lin & Zhang, 2006; Ravikumar et al., 2009); and the additive index model (AIM, Xia et al., 2002; Ruan & Yuan, 2010), which is also known as the projection pursuit regression (PPR, Friedman & Stuetzle, 1981). Because of their straightforward model structures, those models have similar level of interpretability as linear models.

In this chapter, we represent a continuum of models spanning linear model, AM, SIM and AIM using a new neural network architecture. We design the architecture of feedforward neural networks (FNN) to represent the model structures of the aforementioned semiparametric regression models. We name the proposed NN class as the structured neural network (sNN). The additive-index model based explainable neural network in Vaughan et al. (2018) is included as a special case in the sNN.

The sNN class is designed to leverage the NNs and classical semiparametric regression models to strike a better balance between model explainability and predictive performance. Though we essentially t the same model structures, there are several advantages of using the sNN representation and thus the NN training procedure over semiparametric regression methods. First of all, the universal approximation theorem of neural networks justi es its usage in functional approximation (Hornik et al., 1989). In addition, NN can be classi ed as nonlinear parametric models from a statistical view, which means that the gradient of its cost function can be easily computed through backpropagation (Rumelhart et al., 1985) and therefore gradient-based optimization algorithms can be directly applied. On the contrary, the existing estimation methods for semiparametric regression models mostly based on back tting algorithm or its modi cations, which are of higher cost computationally. The back tting algorithm iterates between the estimation functional forms and the parametric components of the model. The functional forms are estimated by nonparametric smoothing operators, including smoothing

splines or kernel regression (Wahba, 1990; Gu, 2013), which are subject to curse of dimensionality and thus are computationally burdensome for large and high-dimensional datasets. Lastly, using standard deep learning software, the training process of the sNNs are uni ed and convenient, whereas the tting methods for semiparametric regression models have many variants in the literature.

Within the sNN class, a remaining issue lies in how to determine the optimal model structure from the observed data. The existing selection criteria include information criteria (AIC, BIC and NIC Akaike, 1973; Akaike, 1977; Murata, 1991) and cross-validated errors (see Anders & Korn, 1999, for details). These methods do not have satisfactory performance in neural network models (Curry & Morgan, 2006). More importantly, they are not designed to suit in the context of balancing performance and explainability of the models. To this end, we propose a model ranking procedure that can be applied (but not limited) to compare the predictive performances of NN models. We provide a partial ordering on the structural complexity of models within the sNN class. Then by sample split, we construct hypothesis tests on the predictive errors that leads to valid comparative inference for models along this path.

The rest of this chapter is organized as the following. In Section 4.2 we set up the problem and notations. In Section 4.3 we propose the sNN model class. Section 4.4 presents a model ranking procedure based on hypothesis testing. Section 4.5 presents the simulation studies that compare the sNNs with semiparametric regression models and evaluate of the proposed model ranking procedure. Finally, Section 4.6 concludes and provides future directions.

4.2 Set-up and notations

We consider the predictive modeling or supervised learning setting. Denote $X \in \mathbb{R}^p$ as a p-dimensional covariate/predictors. We focus on a continuous outcome $Y \in \mathbb{R}$, though generalization to other outcome types is feasible. As in a typical supervised learning setting, the data observed are $\{(Y_i, X_i)\}_{i=1}^n$, and $(Y_i, X_i) \overset{i.i.d.}{\sim} P_0$, where P_0 is the underlying joint distribution of (Y_i, X_i).

Particularly, we consider the setting of large sample size with moderate number of predictors (< 30), which is common in credit models or credit scoring models in the banking industry. We assume that the conditional expectation of outcome satisfies $E(Y|X = x) = \mu(x)$.

The goal is to model $\mu(x)$ such that the estimated model can be used to predict future observation with high accuracy. In the next section, we propose a novel class of NN architectures in Section 4.3 to approximate the unknown function $\mu(x)$ that combines the predictive power of NN and the interpretability of semiparametric regression models.

4.3 Structured neural networks

4.3.1 Architecture of structured neural networks

To start with, we briefly review the model structures of some commonly used semiparametric regression models. For the unified of notation, we refer to the linear coefficients as indices and refer to the unknown functions as ridge functions in all the following models. The single index model has the form of $\mu(x) = f(\beta^T x)$. It has one index, β, and one ridge function $f(\cdot)$. The additive model has the form $\mu(x) = \sum_{j=1}^{p} c_j f_j(x_j)$. It assumes that each predictor X_j has its own ridge function $f_j(\cdot)$ and is additive to each other. The additive index model, which generalized both the single index model and the additive model, has the form $\mu(x) = \sum_{k=1}^{q} c_k f_k(\beta_k^T x)$. It has q ridge functions, additive to each other, and within each there is linear combination of the predictors.

To take advantage of the explainability of semiparametric regression models inherited to their structures, we design the architecture of the FNNs to represent these models. We refer to such FNN models as the structured neural networks (sNN) to highlight the special design in their structures. We list the semiparametric models and their sNN counterparts in Table 4.1. Linear model is also listed as the simplest model form and a special case.

The key idea in sNN is to learn both the indices and the ridge functions by NNs. The

Table 4.1 Summary of structured neural networks and their model structures.

(Semi)parametric Regression Model	sNN	Model Structure
Linear Model		$(x) = {}^{T}x$
Single Index Model	SIM- Net	$(x) = f({}^{T}x)$
Additive Model	AM-Net	$(x) = \sum_{j=1}^{p} c_j f_j(x_j)$
Additive Index Model	AIM-Net	$(x) = \sum_{k=1}^{q} c_k f_k({}^{T}_{k}x)$

Figure 4.1 Architecture of a subnetwork with 5 hidden layers.

indices are represented by the weights of a projection layer. The ridge functions are learned by subnetworks. A subnetwork serves as a nonlinear function approximator. The structure of a l-layer subnetwork is a fully-connected FNN that takes a single neural input and a single neural output. In between the first and the last hidden layer, the subnetwork consists of $l-2$ fully-connected layers of arbitrary number of hidden units in each layer. We denote the architecture of a l layer fully-connected FNN by $[n_1; n_2; \ldots; n_l]$. For example, $[30; 10; 5]$ is a 3-layer fully-connected FNN with 30, 10 and 5 hidden units within each layer, respectively. Figure 4.1 shows a 5-layer subnetwork with architecture $[1; 5; 5; 3; 1]$.

In spirit, subnetwork can be related to the modular neural networks first appeared in Jacobs et al. (1991), where the network's architecture is designed to contain multiple split neural networks. Each split neural network serves as an expert of a sub-task (see Chapter 16, Rojas, 2013, for a comprehensive review). Watanabe et al. (2018) takes the other way by first training

the FNN and then decomposing it into multiple small networks to gain insights of FNN. In either case, the idea is to explain the model behavior by capsulize the structures into several divided pieces.

To be more speci c, Figure 4.2 illustrates the architectures of the three sNN classes in Table 4.1, i.e. SIM-Net, AM-Net and AIM-Net. The predictors rst enter a projection layer, which then produces univariate inputs to be fed into the subnetwork(s). The weights of the projection layer can be either trained or assigned, depending on the class of models. For example, AM-Net has the projection layer weights xed to map each predictor to a separate subnetwork. The subnetwork learns a nonlinear functional form. The nal output of the sNN is produced from a combination layer, which could be a direct output from the subnetwork (as in SIM-Net), or linear combination of the outputs of several subnetworks.

We also give a uni ed representation of the sNNs as

$$\underset{\text{input}}{X} \;!\; \underset{\text{projection layer}}{BX} \;!\; \underset{\text{subnetworks}}{g(BX)} \;!\; \underset{\text{combination layer}}{c^T g(BX)} \;!\; \underset{\text{output}}{\hat{Y}}; \tag{4.1}$$

where $B \in R^{q \times p}$ is the weight matrix of the projection layer that maps X to a lower dimension: $R^p \mapsto R^q$ where in general $q < p$. g is a vector of nonlinear transformation(s) learned by the subnetwork(s), and each applied to one element of BX. $BX \in R^q$. $c \in R^q$ is the weight vector of the combination layer that combines the output of the learned subnetworks. Explicit forms of each sNN class under this representation is given in Appendix C.1.

Given the structural design, sNN combines the strengths of NN and semiparametric regression models. We show that the sNNs and the semiparametric regression model have similar predictive performance in Section 4.5.1.

We note that there is a partial ordering in terms of structural complexity among these models, as shown in Figure 4.3. From top to bottom, the models become more exible but also more complex in structure. Speci cally, linear model is the simplest model; AM and SIM are two generalizations of linear model; AIM model generalizes both SIM and AM and Fully-connected

(a)

(b)

(c)

Figure 4.2 Network architectures of (a) SIM-Net, (b) AM-Net, and (c) AIM-Net.

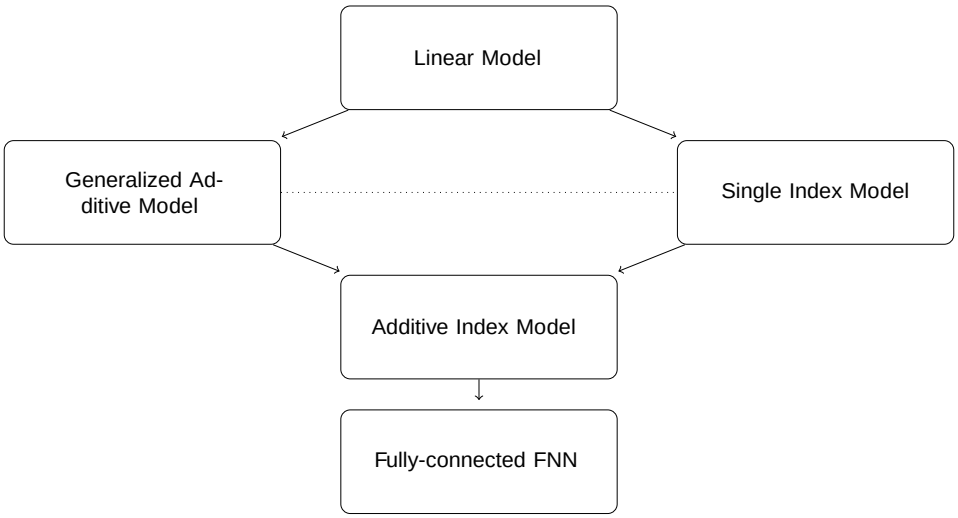

Figure 4.3 From top to bottom, partially ordered simple to complex structures.

FNN is the most flexible model under consideration. The partially ordered hierarchy helps in model selection, which we discuss in Section 4.4.

4.3.2 Regularization

Though we do not focus on high-dimensional X, we point out that penalty can be added to the projection layer and the combination layer to achieve variable/ridge function selection. Based on representation (4.1), the regularized sNN can be written as

$$\underset{\text{input}}{X} \;\to\; \underset{\text{projection layer}}{BX + p_\lambda(B)} \;\to\; \underset{\text{subnetworks}}{B X \;\to\; g(BX)} \;\to\; \underset{\text{combination layer}}{c^T g(BX) + p_\gamma(c)} \;\to\; \underset{\text{output}}{\hat{Y}}:$$

Here, $p_\lambda(B)$ is a q-dimensional vector of penalty functions added to the projection layer, one for each row of B, and $p_\gamma(c)$ is the penalty added to the combination layer. The tuning parameters λ and γ control the magnitude of of the penalties. The choice of penalty functions is flexible, e.g. LASSO, adaptive LASSO (Tibshirani, 1996; Zou, 2006) can be applied.

4.3.3 Computation

We use the python deep learning library Keras with TensorFlow backend to implement the sNN. We use the tanh function (i.e. hyperbolic tangent) as the activation function for the subnetwork. In simulation study, we set batch size to be 1000 and number of epochs to be 500 for both training and tting. The optimizer used is ADAM algorithm (Kingma & Ba, 2014).

A remarkable advantage of the sNN over semiparametric regression algorithm is that the training and validating process of the sNN can be paralleled, same as other NN models. Therefore, with the power of GPU computing, the speed of training and tuning over sNN models for large scale datasets can be escalated. This parallelization is readily applicable using the existing software to all sNN models.

4.4 Prediction error based model ranking

4.4.1 Prediction error

Model selection is a very practical issue in predictive modeling. In particular, for modeling with sNN that contains multiple competing structures, the issue remains in which structure to use in practice. We propose a prediction error based model ranking approach to facilitate the model selection process.

We start by constructing a set of candidate models $M = \{f_j(x) : j = 1, \ldots, J;\ \theta_j \in \Theta_j\}$; parameter θ_j indexes model j and Θ_j is its parameter space, which can be of either nite or in nite dimensional. The number of candidate models, J is assumed to be xed. Among the models in M, we want to nd the best performing model. We de ne \best" with respect to the prediction squared error, also called the prediction risk. For a model indexed by , the prediction risk is

$$= \int (y - f(x))^2 dP_0. \tag{4.2}$$

The bene ts for choosing this criterion are twofold:(i). it measures the expected performance of the model on future (unobserved) data, representing our main interest in predictive modeling; and (ii). it can be easily estimated by the means of sample split, which we will describe soon.

To construct valid inference, we use sample split in training and validating models. We split observed data into two datasets: training set and testing set. Normally, we make an 80%/20% training and testing split as the sample size is assumed to be large. The training set $\mathcal{D}_{\text{train}}$ with size m is used to build a candidate set: $\mathcal{M}_{m,\gamma} = \{\hat{\beta}_{m,j}(x) : j = 1, \ldots, J;\ \beta_{m,j} \in \beta_j\}$. Denote the testing set to be $\mathcal{D}_{\text{test}}$ and its size to be $r = n - m$. An estimate of the prediction risk is the mean squared error (MSE) evaluated on $\mathcal{D}_{\text{test}}$,

$$\hat{R}_{r,\gamma} = \frac{1}{r} \sum_{i \in \mathcal{D}_{\text{test}}} \{y_i - \hat{\beta}_\gamma(x_i)\}^2.$$

Based on MSE, we introduce a simple yet valid hypothesis testing procedure to facilitate model selection.

4.4.2 Hypothesis test

Let $\hat{R}_{r,\gamma} = (\hat{R}_{r,\gamma_1}, \ldots, \hat{R}_{r,\gamma_J})^T$ be the vector of MSEs of all the models in the candidate set evaluated on $\mathcal{D}_{\text{test}}$, and $R_\gamma = (R_{\gamma_1}, \ldots, R_{\gamma_J})^T$ be the vector of the true predictive risks under the data generating distribution P_0, as de ned in (4.2).

For a xed candidate model set indexed by γ, by the central limit theorem, we have an asymptotic multivariate normal distribution

$$\sqrt{r}(\hat{R}_{r,\gamma} - R_\gamma) \xrightarrow{d} N_J\{0, \Omega(\gamma)\}. \tag{4.3}$$

The covariance matrix $\Omega(\gamma)$ can be estimated empirically as

$$\hat{\Omega}(\gamma) = \frac{1}{r} \sum_{i \in \mathcal{D}_{\text{test}}} \{\epsilon^2_{\gamma,i} - \bar{\epsilon}^2_{\gamma,i}\}\{\epsilon^2_{\gamma,i} - \bar{\epsilon}^2_{\gamma,i}\}^T;$$

where $^{2}_{;i} = \overset{h}{f} y_i \quad {}_1(x_i)g^2; : : : ; f y_i \quad {}_J(x_i)g^2 \overset{i}{}_T$:

As we favor the model (x) with the smallest in M, the challenge is to identify such \best" model from limited observed data. We utilize the asymptotic result (4.3) to construct a hypothesis testing procedure to compare the model performance and consequently select the best model in the candidate set. The procedure is able to provide (i) pointwise con dence intervals for each b ; and (ii) pairwise comparisons between models with family-wise error control.

The hypothesis test compares the prediction risk of model ${}_i(x)$ to the prediction risk of model ${}_j(x)$ for $i < j$, i.e. test the null hypothesis

$$H_{0;i;j} : \quad {}_j \quad {}_i \quad 0;$$

against the alternative

$$H_{A;i;j} : \quad {}_j \quad {}_i < 0;$$

for all $j > i;\ i = 1; : : : ; J \quad 1$. For each pairwise test, the test statistic is

$$t_{ij} = q \frac{b_i \quad b_j}{b_{ii}( ) + b_{jj}( ) \quad 2b_{ij}( )}; \tag{4.4}$$

which is a one-sided t-test comparing the di erence of two means. If the null hypothesis is rejected, we conclude that ${}_j(x)$ has a signi cant smaller prediction risk compared to ${}_i(x)$ and thus is a better model for prediction purposes.

When comparing J models, there are $K = (J \quad 1)J=2$ number of pairwise comparisons. Therefore, we need to control family-wise error rate (FWER). We use Holm step-down procedure to control FWER (Lehmann & Romano, 2006). Let be the target FWER level. A brief review of Holm step-down procedure is the following:

Step 1: Order p-values of the K hypothesis tests from small to large

$$p_{(1)} \quad p_{(2)} \quad ; : : : ; \quad p_{(K)}$$

and the corresponding null hypotheses as $H_{(1)}; H_{(2)}; : : : ; H_{(K)}$;

Step 2: For a given target signi cant level , let l be the minimal index such that $p_{(l)} > \frac{}{K+1 \quad l}$;

Step 3: Reject $H_{(1)}; H_{(2)}; : : : ; H_{(l \quad 1)}$, do not reject $H_{(l)}; H_{(l+1)}; : : : ; H_{(K)}$;

Based on the pairwise comparison results, we can provide a partially ordered list of models. It is partially ordered when one or some of the models are not signi cantly di erent from some of the others; and then those models are grouped as a set of similar models. On the cases where one model stands out, it is the selected \best" model. When no such distinction exists, we can still identify a \best" set of models. We can then choose among them incorporating domain knowledge and explainability concerns, i.e. the most parsimonious model is preferred within the set of equally performing models.

Remark 2 We make an important remark on the model ranking procedure that we do not aim to uncover the true data generating model since in practice it is mostly unknown; instead we aim to select the best performing model within a pre-de ned candidate set. As we only specify model con gurations and train the models by the training data, the actual trained models and thus their performance rankings are data dependent, which means that whether each H_0 holds is conditional on the candidate set built with one speci c dataset.

In simulation study, we are able to identify which hypothesis holds given a tted candidate model set by the aid of a large hold-out dataset D_{val} generated under the same mechanism that generates the training and testing set. The true prediction risk is approximated by the MSE on D_{val}. As the sample size of D_{val} grows arbitrarily large, it will be very close to the true prediction risk inherited to model (x). Therefore, in simulation study we are able to detect whether H_0

holds in each run, and thus the power and type I error rate regarding the whole sample-splitting and hypothesis testing procedure are accessible. The simulation study for such examinations is in Section 4.5.2.

4.5 Simulation studies

4.5.1 Comparative study of sNN and semiparametric regression models

In this section, we benchmark the sNN models against their semiparametric counterparts to demonstrate the equivalence. Specically, SIM-net is compared with projection pursuit regression with one ridge function using python packageprojection-pursuit (Komarov, 2018); AM-Net is compared with generalized additive model using python packagepyGAM(Servén & Brummitt, 2018); AIM-Net is compared with projection pursuit regression with multiple ridge functions. The smooth function used in semiparametric models is polynomials of degree 2. The subnetwork architecture is $[1; 30; 10; 5; 1]$ throughout.

We simulate $X \in R^p$ in the following fashion. First we generate Z_k, $k = 1; \ldots; p$ and an auxiliary variable U i.i.d. from $\text{Uniform}(0; 1)$; then the covariate is generated by $X_k = (Z_k + tU)=(1 + t)$, for $k = 1; \ldots; p$. By this formulation, the pairwise correlation between two covariate is $\text{Corr}(X_i; X_j) = t^2=(1 + t^2)$ for $i \neq j$.

The metric used for comparison is MSE on the testing set, as predictive performance is our primary goal. We compare results based on two levels of pairwise correlation, i.e. $t = 0 (\text{Corr} = 0)$ and $t = 1 (\text{Corr} = 0:5)$, and two sample sizes of $n = 10; 000$ and $n = 80; 000$. We make an 80%/20% sampling splitting for training and testing set respectively. Within each class of the sNN model, we generate data according to two outcome models. Results are based on 500 Monte Carlo replications.

SIM-Net : we generate $X \in R^8$ and use the following two SIM models to generate outcomes:

SIM1 $Y = \sin(2 \quad {}_1^T X)=f 2 \quad \sin(2 \quad {}_1^T X)g + e;$

SIM2 $Y = \exp(3\ {}_2^T X/10) + e$.

${}_1 = (1; 2; 3; 4; 0; 0; 0; 0)^T$, ${}_2 = (1; 1; 1; 1; 1; 1; 1; 1)^T$, and $e \sim N(0; 1)$.

AM-Net : we use the following two additive models to generative outcomes:

AM1 $Y = 2X_1 + 3(2X_2 - 1)^2 + 3\sin(2X_3)/(2 - \sin(2X_3)) + 3\log(X_4 + 0.5) + \varepsilon$;

AM2 $Y = X_1 + 2X_2 + X_3 + 2X_4 + 3X_5 + (2X_6 - 1)^2 + \sin(2X_7)/(2 - \sin(2X_7)) + \exp(3X_8/10) + \log(2X_9 + 0.5) + \cos(X_{10}) + e$.

We generate $X \in R^8$ for AM1, $X \in R^{20}$ for AM2 and $e \sim N(0; 1.5^2)$.

AIM-Net : we generate $X \in R^8$. Outcomes are generated from

AIM1 $Y = {}_3^T X + ({}_4^T X)^2 + \sin(2\ {}_5)^T X/(2 - \sin(2\ {}_5^T X)) + e$;

AIM2 $Y = \exp(3\ {}_3^T X/10) - \log(2\ {}_4^T X + 0.5) + \cos({}_5^T X) + e$.

${}_3 = (1; 0; 1; 0; \ldots; 0)^T$, ${}_4 = (0; 1; 0; 1; \ldots; 0)^T$, ${}_5 = (0; 0; 1; 0; 1; 0; \ldots; 0)^T$ and $e \sim N(0; 2^2)$.

Results are summarized in Table 4.2. In addition to MSE comparison, we also check the shape of learned ridge functions, the results are regulated to Appendix C.2. From the results, sNN and semiparametric regression produce similar results though one outperforms the other in different scenarios. However, we made an observation that the performance of sNN exhibits more prominent improvement when the sample size increases. This verifies that sNNs are more suitable for large dataset as compared to the semiparametric regression methods.

4.5.2 Simulation studies on model ranking

In this section, we examine the performance of our proposed model ranking procedure.

We generate covariate $X \in R^8$ and $X_i \overset{i.i.d}{\sim} U[0; 1]$ for $i = 1; \ldots; 8$. Throughout this section, we set the total sample size to be $n = 10{,}000$. Again, we make 80% / 20% training-testing splitting, which results in $m = 8{,}000$ training data D_{train} and $r = 2{,}000$ testing data D_{test}. In each simulation, we use D_{train} to train the J models in the pre-defined candidate set. Each model

takes training data as input, and outputs a tted model $\hat{g}_{m;j}(x)$ for $j = 1;:::;J$. Then we evaluate $\hat{g}_{m;j}(x)$ using D_{test}, which results in an estimated prediction risk $\hat{r}_{m;j}$. An enormous validation dataset D_{val} of size 1,000,000 is created to assess the true prediction risk $r_{m;j}$ for $j = 1:::;J$. Results in this section are all based on 10,000 Monte Carlo simulations. The nominal level for hypothesis test is $= 0:05$.

The metrics for the pointwise estimate of prediction risk are: MSE$\hat{r}_{m;j}$ and its standard error; 95% coverage of the Wald-type con dence interval for$\hat{r}_{m;j}$. The metrics for pairwise comparisons are: MC mean of$p$-values of each pairwise test$p$-value; the number of times H_0 is true # H_0 and the number of rejections givenH_0 is true # Rej$_0$; the number of times H_A is true # H_A and the number of rejections givenH_A is true # Rej$_A$; the empirical type 1 error rate $\hat{} = $ # Rej$_0$=# H_0, and the empirical power 1 $\hat{}$ = # Rej$_A$=# H_A with and without Holm step-down FWER adjustment. We also provide a sanity check with respect to the percentage that the estimated best model set contains the best model and the average length (number of models) of the estimated best model set.

We provide three cases with di erent candidate model sets and data generating mechanisms. The candidate model sets are all formed from the following models:

Model 1: LM - ordinary least square;

Model 2: SIM-Net - subnetwork layer architecture [1; 30; 10; 5; 1];

Model 3: AM-Net - subnetwork layer architecture [1; 30; 10; 5; 1];

Model 4: AIM-Net - subnetwork layer architecture [1; 30; 10; 5; 1];

Model 5: FFNN1 - fully-connect FNN with hidden layer architecture [5; 5];

Model 6: FFNN2 - fully-connect FNN with hidden layer architecture [30; 10; 5].

Case 1: we start with a simple case, where only three models are compared. The candidate

model set is $\{1.\text{AIM-Net}; 2.\text{FFNN1}; 3.\text{FFNN2}\}$. Outcomes are generated from

$$Y = \frac{q}{\exp(3\,\beta_1^T X - 10)\cos(\beta_2^T X) + 2X_2X_8} + e;$$

where $\beta_1 = (0.5, 0.5, 0.5, 0, \ldots, 0)^T$, $\beta_2 = (0.1, 0, 0.1, 0, 0.1, 0, 0)^T$ and $e \sim N(0, 1)$.

Case 2: we expand the candidate set to contain five models: $\{1.\text{SIM-Net}; 2.\text{AM-Net}; 3.\text{AIM-Net}; 4.\text{FFNN1}; 5.\text{FFNN2}\}$. Therefore, there are 10 hypothesis tests to be conducted simultaneously. Outcomes are generated by the same model as in Case 1. I

Case 3: we generate outcomes from a more complex model

$$Y = \frac{q}{\exp(3\,\beta_1^T X - 10)\cos(\beta_2^T X) + 2X_2X_8 + 2X_3\sin(\beta_3^T X) + e;}$$

where $e \sim N(0, 1)$, $\beta_1 = (0.5, 0.5, 0.5, 0, \ldots, 0)^T$, $\beta_2 = (0, 0.1, 0, 0.1, 0, 0.1, 0, 0)$ and $\beta_3 = (0, 0, 0, 0, 0.2, 0.2, 0, 0)^T$. The candidate model set is $\{1.\text{LM}; 2.\text{SIM-Net}; 3.\text{AM-Net}; 4.\text{AIM-Net}; 5.\text{FFNN2}\}$. Again, we conduct 10 hypothesis tests simultaneously. The purpose of this case is to see the performance when the performance among models are more distinct.

The density plot of the prediction risk for models in the candidate set are shown in Figure 4.4 - 4.6 for Case 1 - 3 respectively. It shows the distribution of the prediction risk evaluated using D_{val} for each model configuration.

Simulation results of each case are summarized in Table 4.3 - 4.5. The results show that the coverage rate of the Wald confidence intervals is very close to the nominal level, and the estimated best set can almost always contains the best model. Besides, there are additional insightful observations from the results. The empirical type I error rate is always smaller than the nominal level, which indicates that the test is conservative. As known, the power of the test is contingent on the effect size, which in our case is the difference between two prediction risks. Note that the metric p-value quantifies the effect size. It can be seen from the results that, the power of the test is higher when p-value is either close to 0 or 1, which is consistent with the fact

that the larger the effect size, the higher the power.

Recall that the actual fitted candidate set is data-dependent. Even though the model configurations in the candidate set are fixed, it is possible to observe either event "H_0 holds" or event "H_A holds" in each run of simulation. Notice that the larger the p-value, the more likely the occurrence of "H_0 holds"; the small the p-value, the more likely the occurrence of "H_A holds". To see this better, Figure 4.4 shows that FFNN2 performs the best and FFNN1 performs the worst on average in Case 1. This is consistent with the result that the third test in Case 1 has p-value close to 0. The effect size in the third test is the largest, thus it has the highest power. Also notice that the event "H_A holds" dominates with 9853 occurrences in 10,000 Monte Carlo runs. Other results can be interpreted in the same manner.

Comparing the results in Case 2 and 3, we show the impact of effect size on the model selection. In Case 2, the distributions of prediction risk are similar among models, except for the AM-Net, as shown in Figure 4.5. On contrary, the distributions are more separated in Case 2. It is thus more easily to select the best model from the candidate set in Case 3. Therefore, the average length of the estimated best model set in Case 3 (6.18) is shorter than that of Case 2 (3.31).

4.6 Concluding remarks and future work

We have presented a new model framework that uses neural networks to represent the semiparametric regression models, along with a model ranking procedure. The proposed sNN class is highly practical to predictive modeling objectives and in the meanwhile maintains the model explainability. The sNN class can be generalized to other semiparametric model of interest easily.

For future work, we propose to construct a visualization tool to aid the decision making in model selection for practitioners. An illustration of such tool is shown in Figure 4.7. From the top to bottom, we have models that have four to one ridge function(s). The left panel shows the heat maps that correspond to the estimated indices; the right panel shows the solution paths for each

Figure 4.4 Empirical distribution and tted density of prediction risks based on the validation dataset of Case 1.

of the ridge functions for each model. With the aid of such tool, we can enable an interactive model building process that helps explain the model behaviour to the practitioners.

Figure 4.5 Empirical distribution and tted density of prediction risks based on the validation dataset of Case 2.

Figure 4.6 Empirical distribution and tted density of prediction risks based on the validation dataset of Case 3.

Figure 4.7 Illustration of the proposed visualization tool.

Table 4.2 Comparison on MSE (S.E.) of sNN with semiparametric regressions: (a). SIM-Net verses PPR with one ridge function; (b). AM-Net verses GAM; (c). AIM-Net verses PPR with three ridge functions.

(a) SIM-Net vs PPR(1)

	Corr = 0		Corr = 0.5	
	SIM-Net	PPR(M = 1)	SIM-Net	PPR(M = 1)
	n = 10;000			
SIM1	1:058(0:048)	1:122(0:034)	1:054(0:044)	1:095(0:035)
SIM2	1:026(0:032)	1:002(0:029)	1:020(0:035)	1:000(0:033)
	n = 80;000			
SIM1	1:011(0:036)	1:123(0:012)	1:002(0:014)	1:094(0:012)
SIM2	1:005(0:013)	1:001(0:011)	1:001(0:010)	1:002(0:010)

(b) AM-Net vs GAM

	AM-Net	GAM	AM-Net	GAM
	n = 10;000			
AM1	2:275(0:083)	2:305(0:045)	2:267(0:089)	2:264(0:055)
AM2	2:278(0:332)	2:277(0:055)	2:271(0:114)	2:269(0:054)
	n = 80;000			
AM1	2:256(0:045)	2:298(0:063)	2:258(0:071)	2:254(0:020)
AM2	2:263(0:064)	2:261(0:022)	2:257(0:070)	2:252(0:026)

(c) AIM-Net vs PPR(3)

	AIM-Net	PPR	AIM-Net	PPR
	n = 10;000			
AIM1	4:158(0:078)	4:140(0:007)	4:089(0:023)	4:073(0:008)
AIM2	4:137(0:176)	4:030(0:128)	4:053(0:159)	4:024(0:133)
	n = 80;000			
AIM1	4:056(0:076)	4:099(0:002)	4:023(0:007)	4:034(0:001)
AIM2	4:022(0:071)	4:029(0:044)	4:005(0:045)	4:009(0:046)

Table 4.3 Summary of Case 1: point estimate of prediction risk, 95% coverage of the Wald-type con- dence intervals for each model; mean of p-values, empirical type I error rate and power for each pair-wise comparison w/o Holm adjustment; percentage of the estimated best set contains the true best model w/o Holm adjustment and the average length of the the estimated best set.

Model	1. AIM-Net	2. FFNN1	3. FFNN2
MSE (S.E.)	1:012(0:033)	1:023(0:033)	1:007(0:032)
95% Wald CI Coverage	0:951	0:951	0:950

H_0	p-value	$\#H_0$(# Rej $_0$)	$\hat{b}$	$\hat{b}^{adj}$	$\#H_A$(# Rej $_A$)	$1-\hat{b}$	$1-\hat{b}^{adj}$
$_2 \quad _1 \quad 0$	0:796	8717(31)	0:004	0:001	1283(505)	0:894	0:337
$_3 \quad _1 \quad 0$	0:281	1538(28)	0:018	0:007	8462(3249)	0:384	0:286
$_3 \quad _2 \quad 0$	0:079	147(4)	0:027	0:007	9853(7252)	0:736	0:616

Percentage of the estimated best set contains the true best model			
Unadjusted	(Ave. length)	Holm Adjusted	(Ave. length)
99:64%	(2:08)	99:88%	(2:25)

Table 4.4 Summary of Case 2: point estimate of prediction risk and 95% coverage of the Wald-type con dence interval; mean of p-value, empirical type I error rate and power w/o Holm adjustment; per-centage of the estimated best set contains the true best model w/o Holm adjustment and the average length of the the estimated best set.

Model	1. SIM-Net	2.AM-Net	3.AIM-Net	4.FFNN1	5.FFNN2
MSE (S.E.)	1:03(0:06)	1:03(0:03)	1:01(0:03)	1:01(0:03)	1:01(0:03)
95% Wald CI Coverage	0:949	0:951	0:950	0:951	0:951

H_0	p-value	$\#H_0$(# Rej $_0$)	$\hat{b}$	$\hat{b}^{adj}$	$\#H_A$(# Rej $_A$)	$1-\hat{b}$	$1-\hat{b}^{adj}$
$_2 \quad _1 \quad 0$	0:818	8471(7)	0:001	0	1529(1242)	0:812	0:761
$_3 \quad _1 \quad 0$	0:260	1120(44)	0:039	0:003	8880(3639)	0:410	0:272
$_3 \quad _2 \quad 0$	0:027	0(0)	-	-	10,000(8641)	0:864	0:630
$_4 \quad _1 \quad 0$	0:387	3629(52)	0:014	0:002	6371(2970)	0:466	0:336
$_4 \quad _2 \quad 0$	0:032	92(3)	0:033	0	9908(8870)	0:885	0:673
$_4 \quad _3 \quad 0$	0:650	7511(98)	0:013	0:001	2489(428)	0:172	0:047
$_5 \quad _1 \quad 0$	0:222	321(8)	0:025	0:003	9679(4066)	0:420	0:264
$_5 \quad _2 \quad 0$	0:021	0(0)	-	-	10,000(8884)	0:888	0:662
$_5 \quad _3 \quad 0$	0:435	3587(87)	0:024	0:004	6413(1289)	0:201	0:062
$_5 \quad _4 \quad 0$	0:297	1555(37)	0:024	0:006	8445(2853)	0:338	0:157

Percentage of the estimated best set contains the true best model			
Unadjusted	(Ave. length)	Holm Adjusted	(Ave. length)
98:67%	(3:31)	99:77%	(3:85)

Table 4.5 Summary of Case 3: point estimate of prediction risk and 95% coverage of the Wald-type con dence interval; mean of p-value, empirical type I error rate and power w/o Holm adjustment; percentage of the estimated best set contains the true best model w/o Holm adjustment and the average length of the the estimated best set.

Model			1.LM	2.SIM-Net	3.AM-Net	4.AIM-Net	5.FFNN2
MSE (S.E.)			1:04(0:03)	1:04(0:04)	1:05(0:03)	1:02(0:03)	1:01(0:03)
95% Wald CI Coverage			0.951	0.951	0.950	0.948	0.950

H_0			p-value	$\#H_0$(# Rej $_0$)	$\hat{\beta}$	$\hat{\beta}^{adj}$	$\#H_A$(# Rej $_A$)	1 $\hat{\beta}$	1 $\hat{\beta}^{adj}$
2	1	0	0:199	1080(5)	0:005	0:	8920(4896)	0:549	0:322
3	1	0	0:672	9846(132)	0:013	0:003	154(13)	0:084	0:060
3	2	0	0:810	9144(21)	0:002	0:	856(408)	0:477	0:380
4	1	0	0:033	11(0)	0:	0:	9989(8513)	0:466	0:652
4	2	0	0:098	204(2)	0:001	0:	9796(6669)	0:852	0:468
4	3	0	0:028	6(0)	0:	0:	9994(8745)	0:681	0:688
5	1	0	0:010	0(0)	-	-	10;000(9502)	0:950	0:803
5	2	0	0:027	0(0)	-	-	10;000(8702)	0:870	0:678
5	3	0	0:007	0(0)	-	-	10;000(9671)	0:967	0:843
5	4	0	0:183	535(24)	0:045	0:007	9465(4612)	0:487	0:291

Percentage of the estimated best set contains the true best model			
Unadjusted	(Ave. length)	Holm Adjusted	(Ave. length)
99:76%	(1:68)	99:96%	(2:30)

BIBLIOGRAPHY

Akaike, H. (1973). \Information theory and an extension of the maximum likelihood principle". In: 2nd International Symposium on Information Theory, 1973. Akademiai Kaido, pp. 267{281.

| (1977). \On entropy maximization principle". Application of Statistics, pp. 27{41.

Almirall, D. et al. (2016). \Longitudinal e ects of adaptive interventions with a speech-generating device in minimally verbal children with ASD". Journal of Clinical Child & Adolescent Psychology 45.4, pp. 442{456.

Anders, U. & O. Korn (1999). \Model selection in neural networks". Neural networks 12.2, pp. 309{323.

Athey, S. & S. Wager (2017). \E cient policy learning". arXiv preprint arXiv:1702.02896.

Bai, X. et al. (2013). \Doubly-Robust Estimators of Treatment-Speci c Survival Distributions in Observational Studies with Strati ed Sampling". Biometrics 69, pp. 830{839.

Barrett, J. K. et al. (2014). \Doubly robust estimation of optimal dynamic treatment regimes". Statistics in Biosciences6.2, pp. 244{260.

Bellman, R. (1957). \Dynamic programming". Princeton University Press, p. 151.

Berlinet, A. & C. Thomas-Agnan (2011). Reproducing Kernel Hilbert Spaces in Probability and Statistics. Springer Science & Business Media.

Bickel, P. J. et al. (1993). E cient and Adaptive Inference in Semiparametric Models. Johns Hopkins University Press Baltimore.

Blatt, D et al. (2004). A-learning for approximate planning. Tech. rep. University of Michigan, Ann Arbor, MI, pp. 48109{2122.

Boos, D. D. & L. A. Stefanski (2013). Essential Statistical Inference: Theory and Methods Vol. 120. Springer Science & Business Media.

Boyd, S. & L. Vandenberghe (2004).Convex Optimization. Cambridge university press.

Brainard, L. (2018). What Are We Learning about Arti cial Intelligence in Financial Services? https://www.federalreserve.gov/newsevents/speech/brainard20181113a.htm .

Buchanan, A. L. et al. (2018). \Generalizing evidence from randomized trials using inverse probability of sampling weights". Journal of the Royal Statistical Society: Series A (Statistics in Society) 181, pp. 1193{1209.

Busoniu, L. et al. (2010). Reinforcement Learning and Dynamic Programming using Function Approximators. CRC press.

Buuren, S. & K. Groothuis-Oudshoorn (2011). \mice: Multivariate imputation by chained equations in R". Journal of Statistical Software 45.3.

Chakraborty, B. & E. Moodie (2013). Statistical Methods for Dynamic Treatment Regimes. Springer.

Chan, K. C. G. et al. (2015). \Globally e cient non-parametric inference of average treatment e ects by empirical balancing calibration weighting". Journal of the Royal Statistical Society: Series B (Statistical Methodology) 78, pp. 673{700.

Chang, T. & P. S. Kott (2008). \Using calibration weighting to adjust for nonresponse under a plausible model". Biometrika 95, pp. 555{571.

Chen, G. et al. (2016). \Personalized dose nding using outcome weighted learning" Journal of the American Statistical Association 111.516, pp. 1509{1521.

Chen, J et al. (2002). \Using empirical likelihood methods to obtain range restricted weights in regression estimators for surveys" Biometrika 89, pp. 230{237.

Chen, X. (2007). \Large sample sieve estimation of semi-nonparametric models" Handbook of Econometrics 6, pp. 5549{5632.

Cole, S. R. & E. A. Stuart (2010). \Generalizing evidence from randomized clinical trials to target populations: The ACTG 320 trial". American Journal of Epidemiology 172, pp. 107{115.

Cristianini, N., J. Shawe-Taylor, et al. (2000). An Introduction to Support Vector Machines and other Kernel-based Learning Methods Cambridge university press.

Curry, B. & P. H. Morgan (2006). \Model selection in neural networks: some di culties". European Journal of Operational Research 170.2, pp. 567{577.

Dahabreh, I. J. et al. (2018). \Generalizing causal inferences from individuals in randomized trials to all trial-eligible individuals". Biometrics 0. eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1111/biom.13009 .

Davidian, M. et al. (2019). Introduction to Treatment Regimes. Chapman Hall (forthcoming).

Ertefaie, A. et al. (2016). \Q-learning residual analysis: application to the e ectiveness of sequences of antipsychotic medications for patients with schizophrenia" Statistics in Medicine 35.13, pp. 2221{2234.

Fan, J. & R. Li (2001). \Variable selection via nonconcave penalized likelihood and its oracle properties". Journal of the American Statistical Association 96, pp. 1348{1360.

Fan, J. et al. (2016). Improving covariate balancing propensity score: A doubly robust and e cient approach . Tech. rep. Technical report, Princeton University.

Friedman, J. H. & W. Stuetzle (1981). \Projection pursuit regression". J Am Stat Assoc 76, pp. 817{823.

Fu, W. J. (2003). \Penalized estimating equations". Biometrics 59.1, pp. 126{132.

Geman, S. & C.-R. Hwang (1982). \Nonparametric maximum likelihood estimation by the method of sieves". The Annals of Statistics 10, pp. 401{414.

Geramifard, A. et al. (2013). \A tutorial on linear function approximators for dynamic programming and reinforcement learning". Foundations and Trends in Machine Learning 6.4, pp. 375{451.

Goldberg, Y. & M. R. Kosorok (2012). \Q-learning with censored data". Annals of statistics 40.1, p. 529.

Graham, B. S. et al. (2012). \Inverse probability tilting for moment condition models with missing data". The Review of Economic Studies 79, pp. 1053{1079.

Gu, C. (2013). Smoothing Spline ANOVA Models. Vol. 297. Springer Science & Business Media.

Guan, Q. et al. (2018). \Bayesian Nonparametric Policy Search with Application to Periodontal Recall Intervals". arXiv preprint arXiv:1810.04338.

Hahn, J. (1998). \On the role of the propensity score in e cient semiparametric estimation of average treatment e ects". Econometrica 66, pp. 315{331.

Hahn, O. M. & R. L. Schilsky (2012). \Randomized controlled trials and comparative e ectiveness research". Journal of Clinical Oncology 30, pp. 4194{4201.

Hainmueller, J. (2012). \Entropy balancing for causal e ects: A multivariate reweighting method to produce balanced samples in observational studies". Political Analysis 20, pp. 25{46.

Hajek, J (1971). \Comment on a paper by D. Basu". Foundations of Statistical Inference 236.

Hardle, W. et al. (1993). \Optimal smoothing in single-index models". The Annals of Statistics 21, pp. 157{178.

Hastie, T. & R. Tibshirani (1987). \Generalized additive models: some applications". Journal of the American Statistical Association 82.398, pp. 371{386.

Hastie, T. J. (2017). \Generalized additive models". In: Statistical models in S. Routledge, pp. 249{307.

Henderson, R. et al. (2010). \Regret-regression for optimal dynamic treatment regimes"Biometrics 66.4, pp. 1192{1201.

Hernan, M. A. & T. J. VanderWeele (2011). \Compound treatments and transportability of causal inference".Epidemiology 22, p. 368.

Hershman, D. L. & J. D. Wright (2012). \Comparative e ectiveness research in oncology methodology: observational data". Journal of Clinical Oncology 30, pp. 4215{4222.

Hirano, K. et al. (2003). \E cient estimation of average treatment e ects using the estimated propensity score". Econometrica 71, pp. 1161{1189.

Hornik, K. et al. (1989). \Multilayer feedforward networks are universal approximators". Neural networks 2.5, pp. 359{366.

Horvitz, D. G. & D. J. Thompson (1952). \A generalization of sampling without replacement from a nite universe". Journal of the American Statistical Association 47, pp. 663{685.

Ichimura, H. (1993). \Semiparametric least squares (SLS) and weighted SLS estimation of single-index models".Journal of Econometrics 58, pp. 71{120.

Imai, K. & M. Ratkovic (2014). \Covariate balancing propensity score". Journal of the Royal Statistical Society: Series B (Statistical Methodology) 76, pp. 243{263.

Imbens, G. W. et al. (2005). \Mean-square-error calculations for average treatment e ects". Department of Economics, UC Berkeley, unpublished manuscript

Jacobs, R. A. et al. (1991). \Adaptive mixtures of local experts." Neural computation 3.1, pp. 79{87.

Jeng, X. J. et al. (2018). \High-dimensional inference for personalized treatment decision". Electronic Journal of Statistics 12.1, pp. 2074{2089.

Jiang, B. et al. (2019). \Entropy Learning for Dynamic Treatment Regimes". Statistica Sinica.

Johnson, B. A. et al. (2008). \Penalized estimating functions and variable selection in semiparametric regression models".Journal of the American Statistical Association 103, pp. 672{680.

Kang, J. D. Y. & J. L. Schafer (2007). \Demystifying Double Robustness: A Comparison of Alternative Strategies for Estimating a Population Mean from Incomplete Data". Statistical Science22, pp. 523{539.

Keiding, N. & T. A. Louis (2016). \Perils and potentials of self-selected entry to epidemiological studies and surveys". Journal of the Royal Statistical Society: Series A (Statistics in Society) 179, pp. 319{376.

Kennedy, E. H. (2016). \Semiparametric theory and empirical processes in causal inference". In: Statistical Causal Inferences and Their Applications in Public Health Research. Springer, pp. 141{167.

Kidwell, K. M. (2014). \SMART designs in cancer research: Past, present, and future". Clinical trials 11.4, pp. 445{456.

Kidwell, K. M. et al. (2018). \Design and analysis considerations for comparing dynamic treatment regimens with binary outcomes from sequential multiple assignment randomized trials". Journal of Applied Statistics 45.9, pp. 1628{1651.

Kilbourne, A. M. et al. (2018). \Adaptive School-based Implementation of CBT (ASIC): clustered-SMART for building an optimized adaptive implementation intervention to improve uptake of mental health interventions in schools". Implementation Science 13.1, p. 119.

Kim, J. K. et al. (2016). \Calibrated propensity score method for survey nonresponse in cluster sampling". Biometrika 103, pp. 461{473.

Kingma, D. P. & J. Ba (2014). \Adam: A method for stochastic optimization". arXiv preprint arXiv:1412.6980

Komarov, P. (2018). Projection-pursuit: An implementation of multivariate projection pursuit regression and univariate classi cation https://github.com/pavelkomarov/projection-pursuit .

Korn, E. L. & B. Freidlin (2012). \Methodology for Comparative E ectiveness Research: Potential and Limitations". Journal of Clinical Oncology 30, pp. 4185{4187.

Kosorok, M. R. (2007). Introduction to Empirical Processes and Semiparametric Inference Springer Science & Business Media.

Kosorok, M. R. & E. E. Moodie (2015). Adaptive Treatment Strategies in Practice: Planning Trials and Analyzing Data for Personalized Medicine. Vol. 21. SIAM.

Kosorok, M. & E. Laber (2019). \Precision Medicine". Annual Review of Statistics and Its Application In press .

Koszegi, S. T. (2019). High-Level Expert Group on Arti cial Intelligence .

Kott, P. S. (2006). \Using calibration weighting to adjust for nonresponse and coverage errors". Survey Methodology 32, pp. 133{142.

Laan, M. J. van der & M. L. Petersen (2007). \Causal e ect models for realistic individualized treatment and intention to treat rules". The International Journal of Biostatistics 3.1.

Laber, E. & Y. Zhao (2015). \Tree-based methods for individualized treatment regimes". Biometrika 102.3, pp. 501{514.

Laber, E. B. et al. (2014a). \Dynamic treatment regimes: Technical challenges and applications". Electronic journal of statistics 8.1, p. 1225.

Laber, E. B. et al. (2014b). \Interactive model building for Q-learning". Biometrika 101.4, pp. 831{847.

Laber, E. B. et al. (2014c). \Set-valued dynamic treatment regimes for competing outcomes". Biometrics 70.1, pp. 53{61.

Laber, E. B. et al. (2018). \Optimal treatment allocations in space and time for on-line control of an emerging infectious disease"Journal of the Royal Statistical Society: Series C (Applied Statistics) 67.4, pp. 743{789.

Lavori, P. W. & R. Dawson (2004). \Dynamic treatment regimes: practical design considerations". Clinical trials 1.1, pp. 9{20.

Lehmann, E. L. & J. P. Romano (2006). Testing Statistical Hypotheses Springer Science & Business Media.

Lin, Y., H. H. Zhang, et al. (2006). \Component selection and smoothing in multivariate nonparametric regression".The Annals of Statistics 34.5, pp. 2272{2297.

Linn, K. A. et al. (2017). \Interactive Q-learning for Quantiles". Journal of the American Statistical Association 112.518, pp. 638{649.

Little, R. J. & D. B. Rubin (2014). Statistical Analysis with Missing Data . Vol. 333. John Wiley & Sons.

Liu, Y. et al. (2018). \Augmented outcome-weighted learning for estimating optimal dynamic treatment regimens". Statistics in Medicine.

Lu, W. et al. (2013). \Variable selection for optimal treatment decision". Statistical Methods in Medical Research22.5, pp. 493{504.

Lu, X. et al. (2016). \Comparing dynamic treatment regimes using repeated-measures outcomes: modeling considerations in SMART studies". Statistics in Medicine 35.10, pp. 1595{1615.

Luckett, D. J. et al. (2018). \Estimating Dynamic Treatment Regimes in Mobile Health Using V-learning". Journal of the American Statistical Association just-accepted, pp. 1{39.

Massarelli, E et al. (2003). \A retrospective analysis of the outcome of patients who have received two prior chemotherapy regimens including platinum and docetaxel for recurrent non-small-cell lung cancer". Lung Cancer 39, pp. 55{61.

Moguerza, J. M., A. Mu-noz, et al. (2006). \Support vector machines with applications". Statistical Science 21.3, pp. 322{336.

Moodie, E. E. et al. (2007). \Demystifying optimal dynamic treatment regimes". Biometrics 63.2, pp. 447{455.

Moodie, E. E. et al. (2014). \Q-learning: Flexible learning about useful utilities". Statistics in Biosciences 6.2, pp. 223{243.

Murata, N. (1991). \A criterion for determining the number of parameters in an arti cial neural network model". Arti cial Neural Networks , pp. 9{14.

Murphy, S. A. (2003). \Optimal dynamic treatment regimes". Journal of the Royal Statistical Society: Series B (Statistical Methodology) 65.2, pp. 331{355.

| (2005a). \A generalization error for Q-learning". Journal of Machine Learning Research 6.Jul, pp. 1073{1097.

| (2005b). \An experimental design for the development of adaptive treatment strategies". Statistics in Medicine 24.10, pp. 1455{1481.

Nahum-Shani, I. et al. (2012). \Q-learning: A data analysis method for constructing adaptive interventions." Psychological methods 17.4, p. 478.

Nahum-Shani, I. et al. (2017). \A SMART data analysis method for constructing adaptive treatment strategies for substance use disorders". Addiction 112.5, pp. 901{909.

Newey, W. K. (1997). \Convergence rates and asymptotic normality for series estimators". Journal of Econometrics 79, pp. 147{168.

Neyman, J. (1923). \Sur Les applications de la thar des probabilities aux experiences Agaricales: Essay de principle. English translation of excerpts by Dabrowska, D. and Speed, T." Statistical Science 5, pp. 465{472.

Nosedal-Sanchez, A. et al. (2012). \Reproducing kernel Hilbert spaces for penalized regression: A tutorial". The American Statistician 66.1, pp. 50{60.

O'Muircheartaigh, C. & L. V. Hedges (2014). \Generalizing from unrepresentative experiments: a strati ed propensity score approach". Journal of the Royal Statistical Society: Series C (Applied Statistics) 63, pp. 195{210.

Orellana, L. et al. (2010). \Dynamic regime marginal structural mean models for estimation of optimal dynamic treatment regimes, part I: main content". The International Journal of Biostatistics 6.2.

Pearl, J. & E. Bareinboim (2011). \Transportability of causal and statistical relations: A formal approach". In: Data Mining Workshops (ICDMW), 2011 IEEE 11th International Conference on. IEEE, pp. 540{547.

Petersen, M. L. et al. (2012). \Diagnosing and responding to violations in the positivity assumption". Statistical Methods in Medical Research21.1, pp. 31{54.

Qi, Z., Y. Liu, et al. (2018a). \D-learning to estimate optimal individual treatment rules". Electronic Journal of Statistics 12.2, pp. 3601{3638.

Qi, Z. et al. (2018b). \Multi-armed Angle-based Direct Learning for Estimating Optimal Individualized Treatment Rules with Various Outcomes". Journal of the American Statistical Association just-accepted, pp. 1{35.

Qian, M. & S. A. Murphy (2011). \Performance guarantees for individualized treatment rules". Annals of Statistics 39.2, p. 1180.

Qin, J. & B. Zhang (2007). \Empirical-likelihood-based inference in missing response problems and its application in observational studies". Journal of the Royal Statistical Society: Series B (Statistical Methodology) 69, pp. 101{122.

Ravikumar, P. et al. (2009). \Sparse additive models". Journal of the Royal Statistical Society: Series B (Statistical Methodology) 71.5, pp. 1009{1030.

Rich, B. et al. (2014). \Simulating sequential multiple assignment randomized trials to generate optimal personalized warfarin dosing strategies".Clinical Trials 11.4, pp. 435{444.

Robins, J. M. (1997). \Causal inference from complex longitudinal data". In: Latent variable modeling and applications to causality. Springer, pp. 69{117.

| (2004). \Optimal structural nested models for optimal sequential decisions". In: Proceedings of the second seattle Symposium in Biostatistics Springer, pp. 189{326.

Robins, J. M. & A. Rotnitzky (1995). \Semiparametric e ciency in multivariate regression models with missing data". Journal of the American Statistical Association 90.429, pp. 122{129.

Rojas, R. (2013). Neural Networks: A Systematic Introduction. Springer Science & Business Media.

Rosenbaum, P. R. & D. B. Rubin (1983a). \The Central Role of the Propensity Score in Observational Studies for Causal E ects". Biometrika 70, pp. 41{55.

Rothwell, P. M. (2005). \External validity of randomised controlled trials: \to whom do the results of this trial apply?"" The Lancet 365, pp. 82{93.

Ruan, L. & M. Yuan (2010). \Dimension reduction and parameter estimation for additive index models". Statistics and Its Interface 3.4, pp. 493{499.

Rubin, D. B. & M. J. van der Laan (2012). \Statistical issues and limitations in personalized medicine research with clinical trials". The International Journal of Biostatistics 8.1.

Rubin, D. B. (1974). \Estimating causal e ects of treatments in randomized and nonrandomized studies". Journal of Educational Psychology 66, pp. 688{701.

| (1976). \Inference and missing data". Biometrika 63.3, pp. 581{592.

| (1978). \Bayesian inference for causal e ects: The role of randomization". Annals of statistics, pp. 34{58.

| (1980). \Comment on "Randomization analysis of experimental data: The Fisher randomization test" by D. Basu". Journal of the American Statistical Association 75, pp. 591{593.

Rudolph, K. E. & M. J. van der Laan (2017). \Robust estimation of encouragement design intervention e ects transported across sites". Journal of the Royal Statistical Society: Series B (Statistical Methodology) 79, pp. 1509{1525.

Rumelhart, D. E. et al. (1985). Learning internal representations by error propagation Tech. rep. California Univ San Diego La Jolla Inst for Cognitive Science.

Särndal, C.-E. et al. (2003). Model assisted survey sampling Springer Science & Business Media.

Schulte, P. J. et al. (2014). \Q-and A-learning methods for estimating optimal dynamic treatment regimes". Statistical Science 29.4, p. 640.

Seaman, S. R. et al. (2012). \Combining multiple imputation and inverse-probability weighting". Biometrics 68.1, pp. 129{137.

Servén, D. & C. Brummitt (2018). pyGAM: Generalized Additive Models in Python . https://github.com/dswah/pyGAM .

Shen, X. (1997). \On methods of sieves and penalization". The Annals of Statistics 25, pp. 2555{2591.

Shi, C. et al. (2018). \High-dimensional A-learning for optimal dynamic treatment regimes". The Annals of Statistics 46.3, pp. 925{957.

Shortreed, S. M. & E. E. Moodie (2012). \Estimating the optimal dynamic antipsychotic treatment regime: evidence from the sequential multiple-assignment randomized Clinical Antipsychotic Trials of Intervention and E ectiveness schizophrenia study". Journal of the Royal Statistical Society: Series C (Applied Statistics)61.4, pp. 577{599.

Shortreed, S. M. et al. (2011). \Informing sequential clinical decision-making through reinforcement learning: an empirical study". Machine Learning 84.1-2, pp. 109{136.

Shortreed, S. M. et al. (2014). \A multiple imputation strategy for sequential multiple assignment randomized trials". Statistics in Medicine 33.24, pp. 4202{4214.

Strauss, G. M. et al. (2008). \Adjuvant paclitaxel plus carboplatin compared with observation in stage IB non{small-cell lung cancer: CALGB 9633 with the Cancer and Leukemia Group B, Radiation Therapy Oncology Group, and North Central Cancer Treatment Group Study Groups". Journal of Clinical Oncology 26, p. 5043.

Stroup, T. S. et al. (2003). \The National Institute of Mental Health Clinical Antipsychotic Trials of Intervention E ectiveness (CATIE) project: schizophrenia trial design and protocol development." Schizophrenia Bulletin 29.1, p. 15.

Stuart, E. A. et al. (2011). \The use of propensity scores to assess the generalizability of results from randomized trials". Journal of the Royal Statistical Society: Series A (Statistics in Society) 174, pp. 369{386.

Stuart, E. A. et al. (2015). \Assessing the generalizability of randomized trial results to target populations". Prevention Science16, pp. 475{485.

Sugiyama, M. & M. Kawanabe (2012). Machine Learning in Non-stationary Environments: Introduction to Covariate Shift Adaptation . MIT press.

Sutton, R. S. & A. G. Barto (2018). Reinforcement Learning: An Introduction. MIT press.

Taylor, J. M. et al. (2015). \Reader reaction to A robust method for estimating optimal treatment regimes by Zhang et al.(2012)".Biometrics 71.1, pp. 267{273.

Tian, L. et al. (2014). \A simple method for estimating interactions between a treatment and a large number of covariates". Journal of the American Statistical Association 109.508, pp. 1517{1532.

Tibshirani, R. (1996). \Regression shrinkage and selection via the lasso"Journal of the Royal Statistical Society: Series B (Methodological)58, pp. 267{288.

Tipton, E. (2013). \Improving generalizations from experiments using propensity score subclassi cation: Assumptions, properties, and contexts". Journal of Educational and Behavioral Statistics 38, pp. 239{266.

Tsiatis, A. (2007). Semiparametric Theory and Missing Data Springer Science & Business Media.

Tsiatis, A. A. et al. (2014). Handbook of Missing Data Methodology. Chapman and Hall/CRC.

Vaart, A. W. van der & J. A. Wellner (1996). Weak Convergence and Emprical Processes: With Applications to Statistics. New York: Springer.

Vaughan, J. et al. (2018). \Explainable Neural Networks based on Additive Index Models". arXiv e-prints, arXiv:1806.01933.

Wahba, G. (1990). Spline models for observational data Vol. 59. Siam.

Wang, L. et al. (2012a). \Penalized generalized estimating equations for high-dimensional longitudinal data analysis". Biometrics 68, pp. 353{360.

Wang, L. et al. (2018). \Quantile-optimal treatment regimes". Journal of the American Statistical Association 113.523, pp. 1243{1254.

Wang, L. et al. (2012b). \Evaluation of viable dynamic treatment regimes in a sequentially randomized trial of advanced prostate cancer".Journal of the American Statistical Association 107.498, pp. 493{508.

Watanabe, C. et al. (2018). \Modular representation of layered neural networks". Neural Networks 97, pp. 62{73.

Wolsey, L. A. & G. L. Nemhauser (2014). Integer and Combinatorial Optimization. John Wiley & Sons.

Wu, C. & R. R. Sitter (2001). \A model-calibration approach to using complete auxiliary information from survey data". Journal of the American Statistical Association 96, pp. 185{193.

Wu, P. et al. (2018). \Matched Learning for Optimizing Individualized Treatment Strategies Using Electronic Health Records". Journal of the American Statistical Association just-accepted, pp. 1{35.

Xia, Y. & W. Li (1999). \On single-index coe cient regression models". Journal of the American Statistical Association 94.448, pp. 1275{1285.

Xia, Y. et al. (2002). \An adaptive estimation of dimension reduction space". Journal of the Royal Statistical Society: Series B (Statistical Methodology) 64.3, pp. 363{410.

Xu, Y. et al. (2016). \Bayesian nonparametric estimation for dynamic treatment regimes with sequential transition times". Journal of the American Statistical Association 111.515, pp. 921{950.

Xu, Y. et al. (2019). \Bayesian non-parametric survival regression for optimizing precision dosing of intravenous busulfan in allogeneic stem cell transplantation".Journal of the Royal Statistical Society: Series C (Applied Statistics) 68.3, pp. 809{828.

Young, J. G. et al. (2011). \Comparative e ectiveness of dynamic treatment regimes: an application of the parametric g-formula". Statistics in Biosciences3.1, p. 119.

Yu, Z. & M. J. van der Laan (2002). \Construction of counterfactuals and the G-computation formula".

Zhang, B. & M. Zhang (2018). \C-learning: A new classi cation framework to estimate optimal dynamic treatment regimes". Biometrics 74.3, pp. 891{899.

Zhang, B. et al. (2012a). \A robust method for estimating optimal treatment regimes". Biometrics 68.4, pp. 1010{1018.

Zhang, B. et al. (2012b). \Estimating optimal treatment regimes from a classi cation perspective". Stat 1.1, pp. 103{114.

Zhang, B. et al. (2013). \Robust estimation of optimal dynamic treatment regimes for sequential treatment decisions". Biometrika 100.3, pp. 681{694.

Zhang, Y. et al. (2017). \Estimation of Optimal Treatment Regimes Using Lists". Journal of the American Statistical Association just-accepted.

Zhao, Q. & D. Percival (2017). \Entropy balancing is doubly robust". Journal of Causal Inference 5, DOI: https://doi.org/10.1515/jci{2016{0010.

Zhao, Y.-Q. et al. (2014). \Doubly robust learning for estimating individualized treatment with censored data".Biometrika 102.1, pp. 151{168.

Zhao, Y.-Q. et al. (2015). \New statistical learning methods for estimating optimal dynamic treatment regimes". Journal of the American Statistical Association 110.510, pp. 583{598.

Zhao, Y.-Q. et al. (2019a). \E cient augmentation and relaxation learning for individualized treatment rules using observational data". Journal of Machine Learning ResearchTo appear .

Zhao, Y.-Q. et al. (2019b). \Robustifying trial-derived optimal treatment rules for a target population". Electronic Journal of Statistics 13, pp. 1717{1743.

Zhao, Y. et al. (2012). \Estimating individualized treatment rules using outcome weighted learning". Journal of the American Statistical Association 107.499, pp. 1106{1118.

Zhao, Y. et al. (2009). \Reinforcement learning design for cancer clinical trials".Statistics in Medicine 28.26, pp. 3294{3315.

Zhou, X. & M. R. Kosorok (2017). \Causal nearest neighbor rules for optimal treatment regimes". arXiv preprint arXiv:1711.08451.

Zhou, X. et al. (2017). \Residual weighted learning for estimating individualized treatment rules". Journal of the American Statistical Association 112.517, pp. 169{187.

Zou, H. (2006). \The adaptive lasso and its oracle properties".Journal of the American Statistical Association 101, pp. 1418{1429.

APPENDICES

APPENDIX A

APPENDIX FOR CHAPTER 2

A.1 IPWCC estimating equations

We give the explicit forms of the IPWCC estimating equations for Q-learning and OWL as the following. Denote

$$\hat{\gamma}_t = \max_{a_{t+1} \in \mathcal{A}_{t+1}(H_{t+1})} Q_{t+1}(H_{t+1}, a_{t+1}, \hat{\beta}_{t+1}), \; t = 1, \ldots, T-1.$$

For Q-learning, the stacked IPWCC estimating equation for β is

$$\mathbb{P}_n \begin{bmatrix} \left\{ Y - Q_T(S_T, \beta_T) \right\} \nabla_{\beta_T} Q_T(S_T, \beta_T) 1_{C>T} / K_T(S_T, \bar{C}_T) \\ \left\{ \hat{\gamma}_{T-1} - Q_{T-1}(S_{T-1}, \beta_{T-1}) \right\} \nabla_{\beta_{T-1}} Q_{T-1}(S_{T-1}, \beta_{T-1}) 1_{C>T-1} / K_t(S_{T-1}, \bar{C}_{T-1}) \\ \vdots \\ \left\{ \hat{\gamma}_1 - Q_1(S_1, \beta_1) \right\} \nabla_{\beta_1} Q_1(S_1, \beta_1) 1_{C>1} / K_t(S_1, \bar{C}_1) \end{bmatrix} = 0.$$

For OWL, the stacked IPWCC estimating equation for $= (\ ^T;\ ^T)^T$ is

$$P_n \begin{bmatrix} \{Y - Q_T(H_T; A_T;\ _T)\}\ r\ _T Q_T(H_T; A_T;\ _T) 1_{C>T} = K_T(S_T; \bar{}_T) \\ \vdots \\ V_2(H_2;\ _2;\ _2) - Q_1(H_1; A_1;\ _2;\ _2;\ _1) - r\ _1 Q_1(H_1; A_1;\ _2;\ _2;\ _1) 1_{C>1} = K_1(S_1; \bar{}_1) \\ r\ _T J_T(\ _T;\ _T) 1_{C=T+1} = K_T(S_T; \bar{}_T) \\ \vdots \\ r\ _1 J_1(\ _1;\ _1) 1_{C=T+1} = K_T(S_T; \bar{}_T) \end{bmatrix} = 0:$$

A.2 AIPWCC estimating equations

For Q-learning, at stage $t = T; : : : ; 1$, the AIPWCC estimating equation for $_t$ is

$$P_n \left\{ \frac{1_{C>t}}{K_t(S_t; \bar{}_t)} \left\{ \ _t - Q_t(H_t; A_t;\ _t) - r\ _t Q_t(H_t; A_t;\ _t) \right\} + \right.$$
$$\left. \sum_{r=1}^{t} \frac{1_{C=r}\ _r(S_r;\ _r) 1_{C\ r}}{K_r(S_r; \bar{}_r)} E\left[\left\{ \ _t - Q_t(H_t; A_t;\ _t) - r\ _t Q_t(H_t; A_t;\ _t) \right\} j S_r;\ _{r;t} \right] \right\} = 0: \quad \text{(A.1)}$$

A remark is that for Q-learning, the estimating equation at stage t, i.e. (A.1) belongs to J_{t+1}. The joint estimating equation is formed by stacking (A.1) for $t = 1; : : : ; T$ with the estimating equations for and .

For OWL, at stage $t = T; : : : ; 1$, the AIPWCC estimating equation for $_t$ is

$$P_n \left\{ \frac{1_{C=T+1}}{K_T(S_T; \bar{}_T)} r\ _t J_t^L(\ _t;\ _t) + \right.$$
$$\left. \sum_{r=1}^{T} \frac{1_{C=r}\ _r(S_r;\ _r) 1_{C=r}}{K_r(S_r; \bar{}_r)} E\left[r\ _t J_t^L(\ _t;\ _t) \, j \, S_r;\ _{r;t} \right] \right\} = 0: \quad \text{(A.2)}$$

It can be seen that (A.2) belongs to J_{T+1} because outcome Y is required to get the outcome-based weights $W_t(H_t; A_T; Y;\ _{t+1};\ _t)$ in $J_t(\ _t;\ _t)$ at every stage $T; : : : ; 1$. Again, the joint estimating equation is formed by stacking (A.2) with the ones for , and .

A.3 Simulation settings

Let $\beta = (\beta_{20}, \beta_{21}, \beta_{22}^T, \beta_{23}^T)^T$ and $\gamma = (\gamma_{20}, \gamma_{21}, \gamma_{22}^T)^T$.

For $p = 2$:

$\beta_0 = (1.0, 0.2, 0.2, 0.2, 0.1, 0.1)^T$, $\beta_1 = (0.1, 0.1, 0.1, 0.1)^T$,

$\alpha_0 = \begin{pmatrix} 1 & 0.2 \\ 0.3 & 1 \end{pmatrix}$ and $\alpha_1 = \begin{pmatrix} 0.2 & 0.0 \\ 0.0 & 0.2 \end{pmatrix}$.

For $p = 10$, $p = 30$ and $p = 50$,

$\beta = (1, 0.2, 0.2, 0.2, \underbrace{0, \ldots, 0}_{p-2 \text{ zeros}}, 0.1, 0.1, \underbrace{0, \ldots, 0}_{p-2 \text{ zeros}})^T$;

$\gamma = (0.1, 0.1, 0.1, 0.1, \underbrace{0, \ldots, 0}_{p-2 \text{ zeros}})^T$, $\alpha_0 = \begin{pmatrix} 1 & 0.2 & 0 & \cdots & 0 \\ 0.3 & 1 & 0 & \cdots & 0 \\ \vdots & & & \ddots & \vdots \\ 0 & \cdots & 0 & & 1 \end{pmatrix}_{p \times p}$, and

$\alpha_1 = \text{diag}(0.2, \ldots, 0.2)$

For missing rate 65%: $\gamma_1 = (-0.5, 0.3, 0.3)^T$, $\gamma_2 = (-0.4, 0.2, 0.2)^T$.

For missing rate 35%: $\gamma_1 = (1.2, -0.6, 0.2)^T$, $\gamma_2 = (1.2, -0.6, 0.3)^T$.

A.4 Proof of Theorem 1

By stacking the estimating equations for β and IPWCC estimating equations for γ, we construct a Z-estimator $(\hat{\beta}_n, \hat{\gamma}_n)$ that solves

$$P_n \begin{bmatrix} 1_{C=T} r_{\gamma_T} g_T(S_T, \beta_T) - 1_{C \geq T} r_{\gamma_T} g_T(S_T, \beta_T) \exp\{\gamma^T g_T(S_T, \beta_T)\} \\ \vdots \\ 1_{C=1} r_{\gamma_1} g_1(S_1, \beta_1) - 1_{C \geq 1} r_{\gamma_1} g_1(S_1, \beta_1) \exp\{\gamma^T g_1(S_1, \beta_1)\} \end{bmatrix} w^{cc}(C, S_T, \gamma) \, m_n(H_T, A_T, Y, \beta) = P_n m_n(S_T, Y, \beta, \gamma) = 0,$$

where $w^{cc}(C; S_T; \ )$ and $m_n(H_T; A_T; Y; \ )$ are constructed by aggregating $w_j^{cc}(C; S_{t\ 1}; \bar{}_{t\ 1})$ and $\dot{m}_{n;j}(S_t; \ )$ in the order of $j \ 2 \ J_{T+1} \ [\ J_T \ [\ \ [\ J_1$ respectively; the $\ $ notation is the elementwise product operator for vectors.

Assume that the following standard regularity conditions for Z-estimator hold.

1. $P m_n(S_T; Y; \ ; \ )$ exists for all $\ ; \ \ 2 \ K \ \ B$; there exists $\ ; \ _n \ 2 \ K \ \ B$ such that $jjP m_n(S_T; Y; \ ; \ _n)jj = o(1)$, and $jjP m_n(S_T; Y; \ ; \ )jj \ 6= o(1)$ for $\ 6= \ ; \ 6= \ _n$.

2. Each function in $m_n(\ )$ is continuous in $K \ \ B$ and bounded by an integrable function of the data that does not depend on $(; \ )$.

We show the consistency of $\hat{}_n$ given that $\ _t(s_t) = \ _t(s_t; \ _t)$. Assuming the correctness of the hazard models, it follows that $K_t(s_t) = K_t(s_t; \bar{}_t)$ for all t and s_t, for some $\ 2 \ K$ and that the $\hat{}_n \ ! \ $ in probability. We only need to show

$$E\{w_j^{cc}(C; S_{t\ 1}; \bar{}_{t\ 1})\dot{m}_{n;j}(S_t; \ )g = E\{\dot{m}_{n;j}(S_t; \ )g; j \ 2 \ J_t; t = 1; : : : T + 1:$$

Take the double expectation with the inner expectation condition on S_t and by MAR assumption, we have

$$\begin{aligned} E\left[\frac{1_{C>t\ 1}}{K_{t\ 1}(S_{t\ 1})}\dot{m}_{n;j}(S_t; \ )\right] &= E\left[E\left[\frac{1_{C>t\ 1}}{K_{t\ 1}(S_{t\ 1})}\dot{m}_{n;j}(S_t; \ ) \ S_t\right]\right] \\ &= E\left[\dot{m}_{n;j}(S_t; \ )\frac{E(1_{C>t\ 1}|S_t)}{K_{t\ 1}(S_{t\ 1})}\right] \\ &= E\{\dot{m}_{n;j}(S_t; \ )g \end{aligned}$$

Therefore, if $\ _n$ satisfies $jjP m_n(S_{T+1}; \ _n)jj = o(1)$, we have $jjP m_n(S_{T+1}; \ _n; \hat{}_n)jj = o_p(1)$.

A.5 Proof of Theorem 2

We prove Theorem 2 by showing the consistency in either one of two scenarios: 1. the missingness model is correctly speci ed, i.e. $_t(s_t) = \ _t(s_t; \ _t)$ for all s_t for some $_t \ 2 \ K_t$ and $\hat{}_n \ !$ in probability; 2. the conditional expectation is correctly speci ed, i.e. $d_{n;t}(s_t) = \ d_{n;t}(s_t; \ )$ for some $\ 2 \ a$.

By stacking together the AIPWCC estimation equation for , estimating equation for and , we form a uni ed Z-estimator representation $P_n m_n(S_{T+1}; \ ; \ ; \ ) = 0$. Under same regularity conditions as in proof of Theorem 1 for general Z-estimator, it is su cient to show for $t = 1; : : : ; T + 1$, one has

$$E \ w_j^{cc}(C; S_{t \ 1}; \hat{}_{t \ 1;n}) \hat{}_{n;j}(S_t; \ ) + \sum_{r=1}^{t \ 1} w_{r;j}^{aug}(C; S_r; \hat{}_{r;n}) d_{n;r;j}(S_r; b_{r;n}) \ = E \ _{n;j}(S_t; \ ) \quad \text{for all} \quad j \ 2 \ J_t: \quad (A.3)$$

By Lemma 10.4 of (Tsiatis, 2007), we have

$$1 \quad \frac{1_{C>t}}{K_t(S_t)} = \sum_{r=1}^{t} \frac{1_{C=r} \ _r(S_r; \ _r)1_{C \ r}}{K_r(S_r; \bar{}_r)} \ ; \quad \text{for all} \quad t = 1; : : : ; T: \quad (A.4)$$

Therefore

$$\frac{1_{C>t \ 1}}{K_{t \ 1}(S_{t \ 1}; \bar{}_{t \ 1})} \hat{}_{n;j}(S_t; \ ) = \ _{n;j}(S_t; \ ) + \left(\frac{1_{C>t \ 1}}{K_{t \ 1}(S_{t \ 1}; \bar{}_{t \ 1})} \quad 1 \right) \ _{n;j}(S_t; \ );$$

and the AIPWCC estimation equation on the left of (A.3) can be rearranged as

$$\ _{n;j}(S_t; \ ) \quad \sum_{r=1}^{t \ 1} \frac{1_{C=r} \ _r(S_r; b_{r;n})1_{C \ r}}{K_r(S_r; \hat{}_{r;n})} \{ \ _{n;j}(S_t; \ ) \quad d_{n;r;j}(S_r; b_{r;n}) \}$$

As we assume that the first term $jjPrem_{n;j}(S_t; \ )jj = o(1)$, now it is sufficient to show that

$$E\left[\frac{1_{C=r} \quad {}_r(S_r; b_{r;n})1_{C \ r}}{K_r(S_r; b_{r;n})} f rem_{n;j}(S_t; \ ) \quad d_{n;r;j}(S_r; b_{r;n})g\right] = 0; \qquad (A.5)$$

for $r = 1; : : : ; t \quad 1$ for a specific $t = 1; : : : ; T + 1$

We first prove for scenario 1 where the hazards models are correctly specified so that $b_n \to \ $. Define a sequence of random vectors $G_r = (S_t^T; 1_{C=1}; : : : ; 1_{C=r \ 1})^T$ for $r = 1; : : : t \quad 1$. By the law of iterated expectations, we take conditional expectation on G_r. Noting that the only unknown random variable is $1_{C=r}$ given G_r, we have

$$E\left[\frac{E(1_{C=r} \mid G_r) \quad {}_r(S_r)1_{C \ r}}{K_r(S_r)} f rem_{n;j}(S_t; \ ) \quad d_{n;r;j}(S_r; b_{r;n})g\right]$$

Because of MAR assumption, we have

$$Ef1_{C=r} \mid G_r g = P(C = r \mid C \quad r; S_t)1_{C \ r} = P(C = r \mid C \quad r; S_r)1_{C \ r} = {}_r(S_r)1_{C \ r}:$$

Thus, we prove that (A.5) holds.

Now we consider scenario 2 - the conditional expectation models are correctly specified so that $d_{n;t}(s_t) = d_{n;t}(s_t; \ _t)$. Rewrite the left hand side of (A.5) as

$$E\left[\frac{1_{C=r}}{K_r(S_r; b_{r;n})} f rem_{n;j}(S_t; \ ) \quad d_{n;t}(S_t)g\right] \quad E\left[\frac{{}_r(S_r; b_r)1_{C \ r}}{K_r(S_r; b_r)} f rem_{n;j}(S_t; \ ) \quad d_{n;t}(S_t)g\right]:$$

Applying the law of iterated expectation again, we take conditional expectation of the first term in the above equation on $(1_{C=r}; S_r^T)^T$. This leads to

$$E\left[\frac{1_{C=r}}{K_r(S_r; \ _r)}[Ef rem_{n;j}(S_t; \ ) \mid 1_{C=r}; S_r g \quad Ef rem_{n;j}(S_t; \ ) \mid S_r g]\right]:$$

By MAR, we have $P(S_t \mid 1_{C=r}, S_r) = P(S_t \mid S_r)$, i.e. $E[\hat{m}_{n,j}(S_t, \ ) \mid 1_{C=r}, S_r] = E[\hat{m}_{n,j}(S_t, \ ) \mid S_r]$. This implies the first term in is zero. Similarly, we take conditional expectation of the second term on $(1_{C \ r}, S_r^T)^T$ and obtain

$$E \quad \frac{_r(S_r, \ _r)1_{C \ r}}{K_r(S_r, \overline{\ }_r)} [E[\hat{m}_{n,j}(S_t, \ ) \mid 1_{C \ r}, S_r] \quad E[\hat{m}_{n,j}(S_t, \ ) \mid S_r]] \ :$$

By MAR again, we have $P(S_t \mid 1_{C \ r}, S_r) = P(S_t \mid S_r)$, which implies the second term in is zero.

Therefore we show that (A.5) holds under the two scenarios and the doubly robust conclusion follows.

APPENDIX B

APPENDIX FOR CHAPTER 3

B.1 Additional treatment calibration

We have developed methods that use calibration to enforce the covariate balance between the RCT sample and the target population. We can also apply the calibration technique to balance the covariate distributions between the two treatment groups in the RCT. Toward this end, we further assign weight w_{1i} to each subject i in the treatment group with $A_i = 1$ and assign weight w_{0i} to subject i in the treatment group with $A_i = 0$. The weights w_{1i} and w_{0i} are obtained by solving the following optimization problem:

$$\min_{w_{1i};w_{0i};8i} \sum_{i2A} I(A_i = 1)\, w_{1i} \log w_{1i} + I(A_i = 0)\, w_{0i} \log w_{0i}$$

$$\text{subject to } w_{ai} \quad 0; \sum_{i2A} I(A_i = a) \frac{w_{ai}}{a \;\; A_i + (1 \quad a)(1 \quad A_i)} = 1 \;\; (a = 0; 1);$$

$$\sum_{i2A} I(A_i = 1) \frac{w_{1i}}{A_i} g(X_i) = \sum_{i2A} I(A_i = 0) \frac{w_{0i}}{1 \quad A_i} g(X_i):$$

Similar as in the estimation of Q in (3.3), we can estimate w_{ai} using Lagrangian multiplier as

well and obtain $\hat{w}_{ai}$ for $a = 0; 1$ and $i = 1; : : : ; n$. We refer to the estimators that calibrate both the sampling and the treatment covariate balance as the Double Covariate Balancing (DCB) estimators. The HT-type DCB estimator is

$$\hat{\tau}_{DCB0} = \sum_{i=1}^{n} \left(\frac{\hat{w}_{1i} A_i Y_i}{A_i} - \frac{\hat{w}_{0i}(1 - A_i) Y_i}{1 - A_i} \right) :$$

The Hajek-type DCB estimator is

$$\hat{\tau}_{DCB1} = \sum_{i=1}^{n} \left(\frac{\hat{w}_{1i} A_i Y_i}{\sum_{i=1}^{n} \hat{w}_{1i} A_i} - \frac{\hat{w}_{0i}(1 - A_i) Y_i}{\sum_{i=1}^{n} \hat{w}_{0i}(1 - A_i)} \right) :$$

Similar weights can be added to the ACW estimator as well. It is straightforward, so we omit the formula here.

We examined the performances of the DCB estimators through simulation study, where we found that their improvement over the CW estimators are not evident. This is likely due to that the RCT sample already has a good balance between the two treatment groups. Therefore, we exclude the DCB estimators from the main text. However, we recommend the practitioners to consider using the DCB estimators if there is a belief on the unbalance of covariate distributions between the two treatment groups.

B.2 Proofs

B.2.1 Proof of Theorem 3

Proof of the double robustness of the CW estimators

Let $_{g0} = E\,f\,g(X)g$, $g_0 = g(X)$ $_{g0}$. To use the M-estimator theory (Boos & Stefanski, 2013), we write (3.5) as the following estimating equations

$$\frac{1}{N}\sum_{i=1}^{N} C(X_i; e_i; \ _g) = \frac{1}{N}\sum_{i=1}^{N} e_i d_i f\,g(X_i) \quad _g g = 0; \tag{B.1}$$

$$\frac{1}{N}\sum_{i=1}^{N} (X_i; \ _i; \ ; \ _g) = \frac{1}{N}\sum_{i=1}^{N} \ _i \exp\left\{ ^{>} g(X_i)\right\} f\,g(X_i) \quad _g g = 0: \tag{B.2}$$

First consider the case where Assumption 5 holds, we have $(X) = \exp f\ ^{>}_0 g(X)g$. Notice that $_{g0}$ is the solution to $E\,f\,C(X\,; \ _g)g = 0$. Taking expectation on the left hand side of (B.2) with $_g = \ _{g0}$ leads to

$$E\,f\ (X; \ ; \ ; \ _{g0})g == E\ \ (X)\exp f\ ^{>} g(X)g[g(X) \ \ E\,f\,g(X)g] \ :$$

For the above conditional expectation to be zero, one needs $(X)\exp f\ ^{>} g(X)g$ to be constant. As $(X) = \exp f\ ^{T}_0 g(X)g$, we have $(X)\exp f\ ^{>} g(X)g = \exp f(\ _0 + \ )^{>} g(X)g$. Thus $= \ _0$ makes (B.2) a system of unbiased estimating equations. We point out that denominator in b_q is

an estimator of population size N, i.e.

$$\begin{aligned}
\frac{1}{N}\sum_{i=1}^{n} \exp\{\gamma^{\top} g(X_i)\} &= \frac{1}{N}\sum_{i=1}^{N} \pi_i \exp\{\gamma^{\top} g(X_i)\} \\
&= \frac{1}{N}\sum_{i=1}^{N} \pi_i \exp\left\{\gamma_0^{\top} g(X_i)\right\} + O_p(n^{-1/2}N^{-1}) \\
&= 1 + O_p(N^{-1/2}) + O_p(n^{-1/2}N^{-1}) \\
&= 1 + o_p(1).
\end{aligned}$$

Therefore,

$$\eta_i = q(X_i;\beta) = \frac{\exp\left\{\beta^{\top} g(X_i)\right\}}{\sum_{i=1}^{n} \exp\left\{\beta^{\top} g(X_i)\right\}} = \frac{1}{N}\frac{1}{(X_i;\gamma_0)} + O_p(n^{-1/2}N^{-1}); \qquad \text{(B.3)}$$

i.e. $\eta_i \{N \cdot (X_i;\gamma_0)\} \to 1$ as $n \to \infty$. Based on (B.3), we have

$$\begin{aligned}
\hat{\tau}_{CW0} &= \sum_{i=1}^{N} \eta_i \frac{A_i Y_i}{A_i} - \frac{(1-A_i)Y_i}{1-A_i} \qquad \text{(B.4)} \\
&= \frac{1}{N}\sum_{i=1}^{N} \frac{\pi_i}{(X_i;\gamma_0)} \frac{A_i Y_i}{A_i} - \frac{(1-A_i)Y_i}{1-A_i} = \tau_0 + O_p(N^{-1/2}) + O_p(n^{-1/2}) \\
&= \tau_0 + o_p(1).
\end{aligned}$$

Therefore, $\hat{\tau}_{CW0}$ is consistent for τ_0.

The Hajek-type CW estimator assumes that $\pi_{Ai} = \pi_1$; for all i in the RCT. Under this assumption, based on (B.3), we have

$$\sum_{i=1}^{N} \eta_i A_i = \frac{1}{N}\sum_{i=1}^{N} \frac{\pi_i A_i}{(X_i;\gamma_0)} + O_p(n^{-1/2}) = \pi_1 + O_p(N^{-1/2}) + O_p(n^{-1/2}) = \pi_1 + o_p(1);$$

and thus $\sum_{i=1}^{N} \eta_i (1 - A_i) = 1 - \pi_1 + o_p(1)$. Based on (B.4), it follows that $\hat{\tau}_{CW1}$ is consistent

for $_0$.

Now consider the case where Assumption 6 holds. Then we have

$$
\begin{aligned}
E\left(\sum_{i=1}^{n} \frac{A_i Y_i}{A_i} - \frac{(1-A_i)Y_i}{1-A_i}\right) &= E\left[\sum_{i=1}^{n} E\left(\frac{A_i Y_i}{A_i} - \frac{(1-A_i)Y_i}{1-A_i} \mid X_i; A_i = 1\right)\right] \\
&= E\left[\sum_{i=1}^{n} E\{Y(1) - Y(0) \mid X_i; A_i = 1\}\right] \\
&= {}_0 E\left(\sum_{i=1}^{n} g(X_i)\right) = {}_0 E\left(\frac{1}{N}\sum_{i=1}^{N} d_i g(X_i)\right) \\
&= E\{{}_0 g(X)\} = {}_0;
\end{aligned}
$$

where the equation on the third line is obtained by the balancing constraint (3.2). Under mild regularity conditions for unbiased M-estimators, $\hat{}^{CW0}$ is consistent for $_0$.

Again, for the Hajek CW estimator,

$$
E\left(\sum_{i=1}^{N} {}_i A_i\right) = E\left(\sum_{i=1}^{N} {}_i E(A_i \mid X_i)\right) = E\left(\sum_{i=1}^{N} {}_i E(A_i \mid X_i)\right) = {}_1;
$$

And similarly, $E\left(\sum_{i=1}^{N} {}_i (1-A_i)\right) = 1 - {}_1$. Under regularity conditions for unbiased M-estimators, we have $\sum_{i=1}^{N} {}_i A_i \to {}_1$ as $n \to 1$. Combining this with the consistency results for $\hat{}^{CW0}$, $\hat{}^{CW1}$ is consistent for $_0$. We thus conclude the double robustness of both $\hat{}^{CW0}$ and $\hat{}^{CW1}$.

Proof of the asymptotic variances for the CW estimators

We compare the asymptotic variances of $\hat{}^{CW0}$ and $\hat{}^{CW1}$ under Assumption (5) and (6), because the comparison under the doubly robust condition (i.e. either Assumption (5) or Assumption (6) holds) is cumbersome and unavailing.

Let $= (\ _g^{>}; \ ^{>}; \ _A; \ )^{>}$ to denote all parameters. The estimating function for is

$$(X; A; Y; \ ; \ e; \ ) = \begin{pmatrix} C(X; \ e; \ _g) \\ (X; \ ; \ ; \ _g) \\ h(X; A; \ ; \ ; \ _A) \\ t(X; A; Y; \ ; \ ; \ _A; \ ) \end{pmatrix};$$

where the rst to functions are (B.1), (B.2) and

$$h(X; A; \ ; \ ; \ _A) = \ \exp f \ ^{>} g(X) g(A \ \ _A);$$

$$t(A; X; Y; \ ; \ ; \ _A; \ ) = \ \exp f \ ^{>} g(X) g \ \frac{AY}{_A} \ \ \frac{(1 \ \ A)Y}{1 \ \ _A} \quad :$$

Then $b = (b_g^{>}; b^{>}; \hat{}_A; \hat{}^{CW1})^{>}$ solves the joint estimating equation

$$\frac{1}{N} \sum_{i=1}^{N} \ (X_i; A_i; Y_i; \ _i; \ ) = 0:$$

Under standard regularity conditions in the M-estimator theory, we have

$$b \quad _0 = \frac{1}{N} \sum_{i=1}^{N} \ (X_i; A_i; Y_i; \ _i; \ _0) + o_p(N \ ^{1=2});$$

where $A(\ _0) = E f r \ _{\ _0} \ (\ _0) g$, and $_0 = (\ _{g0}^{>}; \ _0^{>}; \ _A; \ _0)^{>}$. The asymptotic variance of $N \ ^{1=2}(b \ \ _0)$ is $A \ ^{1}(\ _0) B(\ _0) A \ ^{1}(\ _0)^{>}$, where $B(\ _0) = E \ \ (\ _0) \ \ (\ _0)^{>}$.

To further express the asymptotic variance, we denote $q_0 = q_0(X) = \exp f \ \ _0^{>} g(X) g$ and $(Y; A) = f AY = \ _A \ \ (1 \ \ A)Y = (1 \ \ \ _A) \ \ _0 g$. Note that $E f \ (Y; A) j X; \ = 1 g = \ (X) \ \ _0$ and $E \ ^{\theta} d \ = 1$. Under Assumption 5 and 6, $\ (X) q_0(X) = 1$ and $\ (X) \ \ _0 = \ _0^{>} g_0$. In the following derivation we use $)$ to indicate equality when both Assumption 5 and 6 hold.

Using iterated expectation, we have

$$A(\theta_0) = E\{r_\theta(\theta_0)\}$$

$$= E\begin{pmatrix} dI_K & 0_{K\times K} & 0_{K\times 1} & 0_{K\times 1} \\ q_0 I_K & q_0 gg^\top & 0_{K\times 1} & 0_{K\times 1} \\ 0_{1\times K} & q_0(A - A)g^\top & q_0 & 0 \\ 0_{1\times K} & q_0\,(Y;A)g^\top & q_0\left\{\frac{AY}{A^2} + \frac{(1-A)Y}{(1-A)^2}\right\} & q_0 \end{pmatrix}$$

$$= \begin{pmatrix} I_K & 0_{K\times K} & 0_{K\times 1} & 0_{K\times 1} \\ E(q_0) I_K & E\{q_0(g - g_0)g^\top g & 0_{K\times 1} & 0_{K\times 1} \\ 0_{1\times K} & 0_{1\times K} & E(q_0) & 0 \\ 0_{1\times K} & E[(X)q_0\{(X) - {}_0 gg^\top & E\left\{q_0\left\{\frac{AY}{A^2} + \frac{(1-A)Y}{(1-A)^2}\right\}\right\} & E(q_0) \end{pmatrix};$$

By block matrix inversion,

$$A(\theta_0)^{-1} = \begin{pmatrix} I_K & 0_{K\times K} & 0_{K\times 1} & 0_{K\times 1} \\ E\{q_0 g_0 g^\top\}^{-1} E(q_0) & E\{q_0 g_0 g^\top\}^{-1} & 0_{K\times 1} & 0_{K\times 1} \\ 0_{1\times K} & 0_{1\times K} & E(q_0)^{-1} & 0 \\ A_{41} & A_{42} & A_{43} & A_{44} \end{pmatrix};$$

where

$$A_{41} = E\left[q_0(X)\{(X) - {}_0 gg^\top\right] E\{q_0(X)g_0 g^\top\}^{-1}) \; {}_0^\top I_K;$$

$$A_{42} = -E\{q_0(X)g\}^{-1} E\left[(X)q_0(X)\{(X) - {}_0 gg^\top\right] E\{q_0(X)g_0 g^\top\}^{-1}) \; {}_0^\top I_K;$$

$$A_{43} = -E\{q_0(X)g\}^{-2} E\left\{q_0(X)\left\{\frac{AY}{A^2} + \frac{(1-A)Y}{(1-A)^2}\right\}\right\} E\left\{q_0(X)\left\{\frac{Y(1)}{A} + \frac{Y(0)}{1-A}\right\}\right\};$$

$$A_{44} = E\{q_0(X)g\}^{-1}) \; 1;$$

Taking iterated expectation again, we have

$$B(\ _0) = E\left\{ (\ _0)\ (\ _0)^{>} \right\}$$
$$= \begin{pmatrix} B_{11} & 0_{K\ K} & 0_{K\ 1} & 0_{K\ 1} \\ 0_{K\ K} & B_{22} & 0_{K\ 1} & B_{24} \\ 0_{1\ K} & 0_{1\ K} & B_{33} & B_{34} \\ 0_{1\ K} & B_{24}^{>} & B_{34} & B_{44} \end{pmatrix};$$

where

$$B_{11} = E\ \ q_0^2 g_0 g_0^{>}\ ;$$

$$B_{22} = E\{\ q_0(X)^2 g_0 g_0^{>} g\ ) \quad E\{\ q_0(X) g_0 g_0^{>} g;$$

$$B_{24} = E\ \ q_0(X)^2 f\ (X) \quad {}_0 g g_0\ ) \quad E\left\{ q_0(X)\ {}_0^{>} g_0 g_0 \right\};$$

$$B_{33} = E\{\ q_0(X)^2 V(A|X) g\ ) \quad {}_A(1 \quad {}_A) E\{\ q_0(X) g;$$

$$B_{34} = E\{\ q_0^2(X)\ (Y;A)(A \quad {}_A) g = E\ \ q_0^2(X) f(1 \quad {}_A) Y(1) + \ {}_A Y(0) g\ ;$$

$$B_{44} = E\{\ q_0^2(X)\ (Y;A)^2 g = E\ \ q_0^2(X)\ \left\{ \frac{Y(1)^2}{A} + \frac{Y(0)^2}{1 \quad A} \right\} \quad 2\ (X)\ {}_0 + \ {}_0^2 \quad :$$

Notice that $\hat{}^{CW1}$ is the last entry of b, the asymptotic variance of $N^{1=2}(\hat{}^{CW1} \quad {}_0)$ is the bottom rightmost entry of the matrix $A^{1}(\ _0) B(\ _0) A^{1}(\ _0)^{>}$, which is

$$V^{CW1} = A_{41} B_{11} A_{41}^{>} + (A_{42} B_{22} + A_{44} B_{24}^{>}) A_{42}^{>} + (A_{43} B_{33} + A_{44} B_{34}) A_{43}$$
$$+ (A_{42} B_{24} + A_{43} B_{34} + A_{44} B_{44}) A_{44}:$$

Similarly, we can express the asymptotic variance of $N^{1=2}(\hat{}^{CW0} \quad {}_0)$ as $V^{CW0} = A_{41} B_{11} A_{41}^{>} + A_{44}^2 B_{44}$.

Under Assumption 5 and 6, we have

$$A_{41}B_{11}A_{41}^{>} = E\left[\theta_0^2 f_A(X) \quad {}_0\right]^2;$$

$$(A_{42}B_{22} + A_{44}B_{24}^{>})A_{42}^{>} = 0;$$

$$A_{42}B_{24} = E\left[q_0(X) f_A(X) \quad {}_0\right]^2;$$

$$A_{43}B_{34} = \quad E\{Y(1)= {}_A + Y(0)=(1 \quad {}_A)\}E\left[q_0^2(X)\{(1 \quad {}_A)Y(1) + {}_A Y(0)\}\right];$$

$$A_{43}^2 B_{33} = [E\{Y(1)= {}_A + Y(0)=(1 \quad {}_A)\}]^2 \quad {}_A(1 \quad {}_A)E\{q_0(X)\}:$$

Therefore, V^{CW0} can be simplified as

$$V^{CW0} = E\left[\theta_0^2 f_A(X) \quad {}_0\right]^2 + E\left[q_0^2(X)\left(\frac{Y(1)^2}{{}_A} + \frac{Y(0)^2}{1 \quad {}_A}\right) \quad 2 {}_A(X) {}_0 + {}_0^2\right]:$$

The difference in the asymptotic variance of the two CW estimators is given by

$$V^{CW1} \quad V^{CW0} = E\left[\frac{Y(1)}{{}_A} + \frac{Y(0)}{1 \quad {}_A}\right]^2 {}_A(1 \quad {}_A)E\{q_0(X)\} \quad E\left[q_0(X) f_A(X) \quad {}_0\right]^2$$
$$2E\left[\frac{Y(1)}{{}_A} + \frac{Y(0)}{1 \quad {}_A}\right] E\left[q_0^2(X)\{(1 \quad {}_A)Y(1) + {}_A Y(0)\}\right]:$$

We now compare the asymptotic variances under the special case where $_A = 0:5$. Define the sum of two potential outcomes to be $Y^+ = Y(1) + Y(0)$. The asymptotic variances are

$$V^{CW0} = E\left[\theta_0^2 f_A(X) \quad {}_0\right]^2 + 2E\left[q_0^2(X)\{Y(1)^2 + Y(0)^2 \quad 2 {}_A(X) {}_0 + {}_0^2\right];$$

$$V^{CW1} = V^{CW0} + \left[E\{Y^+\}\right]^2 E\{q_0(X)\} \quad E\left[q_0(X) f_A(X) \quad {}_0\right]^2 \quad 2E\{Y^+\} E\left[q_0^2(X) Y^+\right];$$

and their differences is

$$V^{CW1} \quad V^{CW0} = E^2\{Y^+\} E\{q_0(X)\} \quad E\left[q_0(X) f_A(X) \quad {}_0\right]^2 \quad 2E\{Y^+\} E\left[q_0^2(X) Y^+\right]:$$

It is important to note that the difference can be negative, zero, or positive. Therefore, there is no guarantee that which one of the two estimators is better than the other. An observation is that the expected value of Y^+, which is the sum of the two potential outcomes, plays an important role in the asymptotic variance. In the simulation studies provided in the Section B.4, we demonstrate that one outperforms the other in different scenarios, which confirms our findings.

B.2.2 Proof of Theorem 4

Let $Z = (X; A; Y; ; \theta_e)$ be a vector of random variables. Assumptions 1 - 3 and 7 constitute the semiparametric model. The semiparametric likelihood based on a single Z is

$$f(Z) = f(X)\{f(A; Y|X; \quad = 1) \quad (X)\}^{\{f(A; Y|X; \theta_e = 1) \quad _e(X)\}^{e}};$$

where $f(\,)$ is a density function for a continuous random variable and is a probability mass function for a discrete random variable.

Assuming that $\theta_e = 0$, the score function(Hahn, 1998) satisfies

$$S(X; A; Y; ; \theta_e) = S(X; A; Y; \quad) + S(X; A; Y; \theta_e):$$

We first list four identities that are used in the following derivation of the efficiency bound:

1. For any function $h(X; A; \quad)$, we have $E\{h(X; A; \quad)S(Y|X; A; \quad)\} = 0$;
2. For any function $h(X; A; \quad)$, we have $E[h(X; A; \quad)\{Y \quad E(Y|X; A; \quad)\}] = 0$;
3. For any $h(X; A; Y)$, if $E\{\theta h(X; A; Y)\} = 0$, we have $E\{\theta h(X; A; Y)S(X; A; Y; \quad)\} = 0$
4. For any $h(X; A; Y)$, if $E\{h(X; A; Y)\} = 0$, we have $E\{h(X; A; Y)S(X; A; Y; \theta_e)\} = 0$:

To derive the semiparametric e ciency score, we use the method of parametric submodel (Bickel et al., 1993). Let $\{f_t(Z) : t \in R\}$ be a regular parametric submodel which contains the truth at $t = 0$, i.e. $f_t(Z)|_{t=0} = f(Z)$.

Note that $(X) = E(Y|X; A = 1; = 1) \quad E(Y|X; A = 0; = 1)$ and $_0 = E\{ (X)\}$. Let $_t = E_t\{ (X)\}$ denote the parameter evaluated with respect to the regular parametric submodel $f_t(Z)$. Following Bickel et al. (1993), the semiparametric e ciency score $= (Z)$ is the pathwise derivative of the target parameter in the sense that

$$\frac{@}{@t} _t \Big|_{t=0} = E\{ S(Z)\};$$

where $S(Z) = @\log f_t(Z)=@t|_{t=0}$. Toward this end, we express

$$\frac{@}{@t} _t \Big|_{t=0} = E\{ (X)S(X)\} + E\left\{ \frac{@_t(X)}{@t} \Big|_{t=0} \right\} : \tag{B.5}$$

To express (B.5) further,

$$\begin{aligned}
\frac{@_t(X)}{@t} \Big|_{t=0} &= \int y \frac{@}{@t} f_t(y|X; = 1; A = 1)\, dy \quad \int y \frac{@}{@t} f_t(y|X; = 1; A = 0)\, dy \Big|_{t=0} \\
&= \int yS(y|X; A = 1; = 1) f_t(y|X; A = 1; = 1)\, dy \Big|_{t=0} \\
&\quad \int yS(y|X; A = 0; = 1) f_t(y|X; A = 0; = 1)\, dy \Big|_{t=0} \\
&= E\left\{ \frac{AY}{(X)^A} S(Y|A; X;)|X \right\} \quad E\left\{ \frac{(1 \quad A)Y}{(X)(1 \quad ^A)} S(Y|A; X;)|X \right\} \\
&= E\left\{ \frac{}{(X)} \left(\frac{AY}{^A} \quad \frac{(1 \quad A)Y}{(1 \quad ^A)} \right) S(Y|A; X;)|X \right\} :
\end{aligned}$$

Therefore,

$$
\begin{aligned}
E\left[\frac{\partial \tau_t(X)}{\partial t}\right]_{t=0} &= E\left[\frac{}{(X)}\left(\frac{AY}{A}-\frac{(1-A)Y}{(1-A)}\right)S(Y|A,X;\ )\right] \\
&= E\left[\frac{}{(X)}\left(\frac{Af(Y=1)-\mu_1(X)g}{A}-\frac{(1-A)f(Y=0)-\mu_0(x)g}{1-A}\right)S(Y|A,X;\ )\right] \quad \text{(B.6)} \\
&= E\left[\frac{}{(X)}\left(\frac{Af(Y=1)-\mu_1(X)g}{A}-\frac{(1-A)f(Y=0)-\mu_0(x)g}{1-A}\right)S(X,A,Y;\ )\right] \quad \text{(B.7)} \\
&= E\left[\frac{}{(X)}\left(\frac{Af(Y=1)-\mu_1(X)g}{A}-\frac{(1-A)f(Y=0)-\mu_0(x)g}{1-A}\right)S(X,A,Y;\ ;\ e)\right]: \\
& \qquad \text{(B.8)}
\end{aligned}
$$

In the above derivation, (B.6) follows by identity 1, (B.7) follows by identity 2, and (B.8) follows by identity 3 and 4. Therefore,

$$
\begin{aligned}
E\left[\left\{d\ f(X)-\mu_0(X)\right\}S(X)\right] &= E\left[\left\{d\ f(X)-\mu_0(X)\right\}S(X)\right] \\
&= E\left[\left\{d\ f(X)-\mu_0(X)\right\}\left\{S(X)+S(X,A,Y;\ )+S(X,A,Y;\ e)\right\}\right] \quad \text{(B.9)} \\
&= E\left[\left\{d\ f(X)-\mu_0(X)\right\}S(X,A,Y;\ ;\ e)\right];
\end{aligned}
$$

where(B.9) holds because

$$
\begin{aligned}
E\left[\left\{d\ f(X)-\mu_0(X)\right\}S(X,A,Y;\ e)\right] &= E\left[\left\{d\ f(X)-\mu_0(X)\right\}S(X,A,Y;\ e=1)\right] \\
&= E\left[\left\{d\ f(X)-\mu_0(X)\right\}\left\{E\left[S(X,A,Y;\ e=1)\ \middle|\ X,\ e\right]\right\}\right] \\
&= E\left[\left\{d^2 f(X)-\mu_0(X)\right\}\left\{E\left[S(X,A,Y;\ e=1)\ \middle|\ X,\ e=1\right]\right\}\right] \\
&= 0:
\end{aligned}
$$

Substituting back to (B.5), we have

$$\frac{@}{@t^t}_{t=0} = E \quad {}^{e}d\, f\ (X) \quad {}_0 g + \frac{}{(X)} \quad \frac{A f Y \quad {}_1(X) g}{A} \quad \frac{(1 \quad A) f Y \quad {}_0(x) g}{1 \quad A} \quad S(X; A; Y; ; {}^{e})$$

Thus, the semiparametric e ciency score is

$$= {}^{e}d\, f\ (X) \quad {}_0 g + \frac{}{(X)} \quad \frac{A f Y \quad {}_1(X) g}{A} \quad \frac{(1 \quad A) f Y \quad {}_0(x) g}{(1 \quad {}_A)} \ :$$

It follows that the semiparametric e ciency bound is

$$E(\ ^2) = E \quad {}^{e}d^2 f\ (X) \quad {}_0 g^2 + \frac{}{^2(X)} \quad \frac{V f Y(1) j X g}{A} + \frac{V f Y(0) j X g}{1 \quad A} \ :$$

B.2.3 Proof of Theorem 5

Proof of the double robustness of the ACW estimator

Let $= (\ ; {}_0; {}_1)$ denote the vector of nuisance parameters. Note that $\hat{}^{ACW}$ is the solution to the estimating equation $N^{1} \sum_{i=1}^{N} (X_i; A_i; Y_i; {}_i; {}^{e}_i; ; {}^{b}) = 0$, where

$$(X; A; Y; ; {}^{e}; ;) = \frac{}{(X;)} \quad \frac{A f Y \quad {}_1(X; {}_1) g}{A} \quad \frac{(1 \quad A) f Y \quad {}_0(X; {}_0) g}{1 \quad A} + {}^{e}d\, f\ {}_1(X; {}_1) \quad {}_0(X; {}_0) g \ :$$

Let be the probability limits of ^{b}. It su ces to show that $E^{n} (X; A; Y; ; {}^{e}; {}_0;)^{o} = 0$ if either $(X;)$ or ${}_a(X; {}_a)\ (a = 0; 1)$ is correctly speci ed. Under standard regularity conditions

for M-estimators, $\hat{}^{ACW}$ is consistent for . Use iterated expectation, we can write

$$E\left\{ (X;A;Y;\ ;\ e;\ _0;\ )\right\} = E\left\{ \frac{}{(X\ ;\ )}\left(\frac{AY}{A} - \frac{(1\quad A)Y}{1\quad A} \right)\quad _0 \right\} +$$

$$E\left\{ \frac{}{(X\ ;\ )}\ f\ _1(X\ ;\ _1)\quad _0(X\ ;\ _0)g\ :\right. \quad \text{(B.10)}$$

The rst term on the left-hand-side of (B.10) is 0 either one of the $(X\ ;\ )$ or $_a(X\ ;\ _a)\,(a = 0;1)$ is correctly speci ed, as shown in the proof of consistency in the CW estimators. Now consider the second term on the left-hand-side of (B.10).

Firstly, if $(X\ ;\ )$ is correctly speci ed, we have $(X\ ;\ ) = (X)$. Take iterated expectation conditional on X, we have the second term on the left-hand-side of (B.10)

$$E\left\{ \frac{}{(X)}\ f\ _1(X\ ;\ _1)\quad _0(X\ ;\ _0)g\right\} = E\left\{ f\ _1(X\ ;\ _1)\quad _0(X\ ;\ _0)gE\left\{ \frac{}{(X)}jX\right.\right.$$

$$= 0;$$

as $E\left\{ \right\} = (X)jX = 0$. Thus, (B.10) equals to zero.

Secondly, if outcome model $_a(X\ ;\ _a)\,(a = 0;1)$ is correctly speci ed, we have $_1(X\ ;\ _1)$ $_0(X\ ;\ _0) = \ _0^{>}g(X)$. Then the second term on the left-hand-side of (B.10) satis es

$$E\left\{ \frac{}{(X)}\ f\ _1(X\ ;\ _1)\quad _0(X\ ;\ _0)g\right\} = E\left\{ \ _0^{>}\left[= (X\ ;\ )\right] g(X)\right\} = 0$$

by the balancing constraint (3.2). Thus, (B.10) equals to zero under this scenario as well. This completes the proof of the double robustness of $\hat{}^{ACW}$.

B.2.4 Proof of Theorem 5 and Theorem 6

Proof of local e ciency

Following the empirical process literature, let P_N denote the empirical measure. For a random variable V, $Pf\,f(V)g = \int f(v)dP$ is the expectation of $f(V)$ under the true data-generating process. Recall that $Z = (X; A; Y; ; e)$, $= (\ ; \ _0; \ _1)$, is the probability limits of b and $_0$ is the corresponding true parameter value. Let

$$
\begin{aligned}
(Z; \ ) = & \frac{}{(X; \ )} \frac{A f Y \quad _1(X; \ _1)g}{A} \quad \frac{(1 \quad A) f Y \quad _0(X; \ _0)g}{1 \quad A} \\
& + \frac{}{} d f \ _1(X; \ _1) \quad _0(X; \ _0)g \\
= & \frac{}{(X; \ )} \frac{A f Y \quad _1(X; \ _1)g}{A} + \frac{}{} d \ _1(X; \ _1) \\
& \frac{}{(X; \ )} \frac{(1 \quad A) f Y \quad _0(X; \ _0)g}{1 \quad A} \quad \frac{}{} d \ _0(X; \ _0) \\
=: & \ _1(Z; \ ) \quad _0(Z; \ ):
\end{aligned}
$$

Under the conditions speci ed in Theorem 5 or the conditions speci ed in Theorem 6, and assume that $(Z; \ )$ belongs to Donsker classes (Vaart & Wellner, 1996; Kennedy, 2016), $P \ _1(Z; \ ) = \ _1$, $P \ _0(Z; \ ) = \ _0$ and $P \ (Z; \ ) = \ _1 \quad _0 = \ _0$. Thus,

$$
\begin{aligned}
\hat{}^{ACW} \quad _0 &= P_N \ (Z; b) \quad P \ (Z; \ ) \\
&= (P_N \quad P) \ (Z; b) + Pf \ (Z; b) \quad (Z; \ )g \\
&= (P_N \quad P) \ (Z; \ ) + Pf \ (Z; b) \quad (Z; \ )g + o_p(N^{-1=2}): \qquad \text{(B.11)}
\end{aligned}
$$

We now show that

$$
Pf \ (Z; b) \quad (Z; \ )g = Pf \ _1(Z; b) \quad _1(Z; \ )g \quad Pf \ _0(Z; b) \quad _0(Z; \ )g
$$

is a small order term under conditions in Theorem 5 or Theorem 6. We write

$$\begin{aligned}
Pf\ _1(Z;\hat{b}) \quad _1(Z; \quad)g &= P\left\{ \frac{\hat{A}}{(X;\hat{b})} \left(\frac{Y \quad _1(X;\hat{b}_1)}{\hat{A}} + \left(\hat{d} \quad _1(X;\hat{b}_1) \right) \right) \quad _1 \right\} \\
&= P\left\{ \frac{}{(X;\hat{b})} \quad 1 \quad _1(X) \quad _1(X;\hat{b}_1) \right\} + \left(\hat{d} \quad 1 \quad _1(X;\hat{b}_1) \right) \\
&= P\left\{ \frac{(X) \quad (X;\hat{b})}{(X;\hat{b})} \quad _1(X) \quad _1(X;\hat{b}_1) \right\} :
\end{aligned}$$

Similarly, we have

$$Pf\ _0(Z;\hat{b}) \quad _0(Z; \quad)g = P\left\{ \frac{(X) \quad (X;\hat{b})}{(X;\hat{b})} \quad _0(X) \quad _0(X;\hat{b}_0) \right\} :$$

Therefore, by Cauchy-Schwarz inequality and the positivity of $(X;b)$, $jPf\ (Z;\hat{b}) \quad (Z; \quad)gj$ is bounded above by

$$k \quad (X) \quad (X;\hat{b})k \sum_{a2f\ 0;1g} k \ _a(X) \quad _a(X;\hat{b}_a)k: \tag{B.12}$$

Under the conditions in Theorem 5, if $(X; \quad)$ is a correctly speci ed parametric model for (X), then $k \quad (X) \quad (X;\hat{b})k = O_p(n^{\ 1=2})$; and if $_a(X; \ _a)$ is a correctly speci ed parametric model for $_a(X)$, then $k \ _a(X) \quad _a(X;\hat{b}_a)k = O_p(n^{\ 1=2})$. Therefore, the product (B.12) is $O_p(n^{\ 1})$, which makes $Pf\ (Z;\hat{b}) \quad (Z; \quad)g$ in (B.11) asymptotically negligible. Under the conditions in Theorem 6, the product (B.12) is $o_p(n^{\ 1=2})$ and therefore the term $Pf\ (Z;\hat{b}) \quad (Z; \quad)g$ in (B.11) is asymptotically negligible. The result follows.

B.3 Conditions for the sieves estimator

Following Hirano et al. (2003), we assume the following regularity conditions on the data generating process and the nuisance functions.

Condition 3 (Distribution of X) Let $X \subset R^p$ be the support of X. Assume that X is a Cartesian product of compact intervals, i.e. $X = \prod_{j=1}^{p} [l_j, u_j]$, $l_j, u_j \in R$. The density of X, $f(X)$, is bounded above and below away from 0 on X.

Condition 4 (Basis functions) There exist constant l and u such that

$$l \quad \min \left\{ g(X)^{\top} g(X) \right\} \quad \max \left\{ g(X)^{\top} g(X) \right\} \quad u;$$

almost surely where λ_{min} and λ_{max} denote the minimum and maximum eigenvalues of a matrix.

Condition 5 (Potential outcomes) The second moment of the potential outcomes are finite. i.e. $E\left[Y(a)^2\right] < 1$, for $a = 0, 1$.

Condition 6 (Smoothness) The log sampling score function $\log \pi(x)$ is s_π-times continuously differentiable and the outcome mean function $\mu_a(x)$ is s_a-times continuously differentiable, $\forall x \in X$; $a = 0, 1$; The sieves estimators of $\log \pi(X)$ and $\mu_a(X)$ use a power series and the smoothness condition is $1 - s/2p > 1/4$, for $s = s_\pi$ and $s = s_a$ respectively.

B.4 Additional simulation study

We conduct additional simulation studies to evaluate the effect of potential outcome model and sample sizes on the performance of our proposed estimators.

B.4.1 Comparison of the CW estimators

In this simulation study, we compare the performance of $\hat{\tau}^{CW0}$ and $\hat{\tau}^{CW1}$. Under the data generating mechanism in Section 3.5, $\hat{\tau}^{CW0}$ is more efficient than $\hat{\tau}^{CW1}$. In this section, the simulation setting is the asme as that in Section 3.5, except that we generate the potential

outcomes according to

$$Y(a)|X = 100 + 27{:}4X_1 \quad a + 13{:}7X_2 + 13{:}7X_3 + 13{:}7X_4 + \quad ;$$

where $\quad N(0;1)$ for $a = 0;1$. Figure B.1 reports the simulation results. Under this setting, it can be seen that the variances of $\hat{}^{CW0}$ become dramatically larger in all four scenarios than the variances of $\hat{}^{CW1}$. Combining with the simulation results in Section 3.5, we show that $\hat{}^{CW0}$ and $\hat{}^{CW1}$ can outperform each other in di erent scenarios.

B.4.2 Increased sample sizes

In this simulation study, we increase the population size to $N = 500000$, the RWE sample size to $m = 10000$, and the RCT selection rate to 0.45% that leads to $n = 2250$ to study the e ect of sample size. Results are shown in Figure B.2 and Table B.1 for continuous outcome and . It can be seen that when the sample sizes are larger, the advantages of our proposed estimators become more obvious. The coverage rate for both continuous outcome and binary outcome are closer to the nominal rate.

Figure B.1 Boxplot of estimators under four model speci cation scenario:$\hat{}^{CW0}$ is worsen than $\hat{}^{CW1}$.

Table B.1 Simulation results for continuous outcome with population size N = 500000: bias of point estimates, Monte Carlo variance, relative bias of bootstrap variance estimate, and the coverage rate of 95% Wald con dence interval.

	Naive	IPSW	CW0	CW1	ACW	ACW(S)
	Bias					
1. O:C/S:C	13:64	0:09	0:07	0:00	0:05	0:05
2. O:C/S:W	11:80	1:20	0:06	0:08	0:06	0:06
3. O:W/S:C	11:39	0:10	0:03	0:04	0:02	0:00
4. O:W/S:W	11:01	2:68	1:59	1:56	1:58	0:02
	Monte Carlo variance					
1. O:C/S:C	1:53	5:23	1:02	4:42	0:08	0:08
2. O:C/S:W	1:38	2:98	0:74	2:79	0:08	0:08
3. O:W/S:C	1:51	8:87	4:10	7:34	3:23	0:09
4. O:W/S:W	1:42	2:98	0:88	2:84	0:48	0:10
	Relative bias (%) of bootstrap variance estimate					
1. O:C/S:C	2:3	0:6	4:4	4:3	1:2	0:1
2. O:C/S:W	4:3	1:0	8:2	3:1	2:1	2:5
3. O:W/S:C	1:1	0:3	11:0	9:5	13:7	1:2
4. O:W/S:W	1:9	0:8	12:6	4:4	2:8	1:4
	95% Wald CI coverage					
1. O:C/S:C	0:0	94:0	93:6	93:2	94:6	94:6
2. O:C/S:W	0:0	88:0	95:1	94:5	93:9	93:8
3. O:W/S:C	0:0	92:3	94:1	91:8	91:7	94:9
4. O:W/S:W	0:0	62:3	63:2	81:6	33:7	94:9

Figure B.2 Boxplot of estimators for continuous outcomewith population size N = 500000 under four model speci cation scenarios.

Table B.2 Simulation results for continuous outcome with population size N = 500000: bias of point estimates, Monte Carlo variance, relative bias of bootstrap variance estimate, and the coverage rate of 95% Wald con dence interval.

	Naive	IPSW	CW0	CW1	ACW	ACW(S)
	Bias 1000					
1. O:C/S:C	94:75	0:49	1:08	0:22	0:22	0:28
2. O:C/S:W	87:52	8:73	16:20	16:37	1:59	1:58
3. O:W/S:C	80:35	1:93	0:50	1:35	1:23	0:93
4. O:W/S:W	76:53	24:52	26:57	26:72	17:82	2:95
	Monte Carlo variance 1000					
1. O:C/S:C	0:40	0:70	0:69	0:68	0:43	0:44
2. O:C/S:W	0:38	0:55	0:54	0:54	0:36	0:44
3. O:W/S:C	0:38	0:71	0:85	0:70	0:61	0:56
4. O:W/S:W	0:38	0:57	0:75	0:55	0:50	0:52
	Relative bias (%) of bootstrap variance estimate					
1. O:C/S:C	1:8	4:8	2:5	6:0	3:4	4:2
2. O:C/S:W	2:7	5:2	4:6	3:9	2:8	4:3
3. O:W/S:C	1:1	2:8	5:6	4:5	10:0	7:3
4. O:W/S:W	1:9	8:3	3:4	8:3	2:9	0:9
	95% Wald CI coverage					
1. O:C/S:C	0:4	94:3	94:5	93:9	94:4	93:8
2. O:C/S:W	1:0	94:0	90:1	89:1	95:3	94:2
3. O:W/S:C	2:3	94:0	93:6	94:8	93:2	93:4
4. O:W/S:W	2:1	82:9	83:2	79:6	86:7	95:3

Figure B.3 Boxplot of estimators for binary outcome with population size N = 500000 under four model speci cation scenarios.

APPENDIX C

APPENDIX FOR CHAPTER 4

C.1 sNN specifications

To be more explicit, we give the exact form of each class of the sNNs introduced in Section 4.3.1.

For SIM-Net, the weights are $B = (\ _{11}\ _{12} \cdots\ _{1p})^T$ and $c = c_1$. Nonlinear transformation $g(\ ) = g_1(\ )$ corresponds to one subnetwork.

For AM-Net, the project weights are fixed (do not need to be learned) to be $B = (I_p; 0_{p\ q})^T$. The combination weight is $c = (c_1; \cdots; c_p; 0_{q\ p})^T$. Each predictor has its own nonlinear transformation $g_j(X_j)$, for $j = 1; \cdots; p$, corresponding to p subnetworks.

For AIM-Net, the weights are

$$B = \begin{bmatrix} {}^T{}_1 \\ {}^T{}_2 \\ \vdots \\ {}^T{}_q \end{bmatrix} = \begin{bmatrix} {}_{11} & {}_{12} & \cdots & {}_{1p} \\ {}_{21} & {}_{22} & \cdots & {}_{2p} \\ \vdots & & & \\ {}_{q1} & {}_{q2} & \cdots & {}_{qp} \end{bmatrix}; \qquad c = \begin{bmatrix} c_1 \\ \vdots \\ c_q \end{bmatrix}.$$

Figure C.1 Ridge functions learned from AM with polynomial splines.

Nonlinear transformation $g(\cdot)$ contains q subnetworks, one for each row of $\mathbf{B}$.

C.2 Ridge functions

We take the AM2 model in Section 4.5.1 as a showcase. Both AM and AM-Net can recover the true ridge functions, as shown in Figure C.1 and C.2.