Continuous Treatment Effects with Surrogate Outcomes

Zhenghao Zeng [1] **David Arbour** [2] **Avi Feller** [3] **Raghavendra Addanki** [2] **Ryan Rossi** [2] **Ritwik Sinha** [2]
Edward H. Kennedy [1]

Abstract

In many real-world causal inference applications, the primary outcomes (labels) are often partially missing, especially if they are expensive or difficult to collect. If the missingness depends on covariates (i.e., missingness is not completely at random), analyses based on fully observed samples alone may be biased. Incorporating surrogates, which are fully observed post-treatment variables related to the primary outcome, can improve estimation in this case. In this paper, we study the role of surrogates in estimating continuous treatment effects and propose a doubly robust method to efficiently incorporate surrogates in the analysis, which uses both labeled and unlabeled data and does not suffer from the above selection bias problem. Importantly, we establish the asymptotic normality of the proposed estimator and show possible improvements on the variance compared with methods that solely use labeled data. Extensive simulations show our methods enjoy appealing empirical performance.

1. Introduction

In many causal inference applications, the primary outcomes are missing for a non-trivial number of observations. For instance, in studies on long-term health effects of medical interventions, some measurements require expensive testing and a loss to follow-up is common (Hogan et al., 2004). In evaluating commercial online ad effectiveness, some individuals may drop out from the panel because they use multiple devices (Shankar et al., 2023), leading to missing revenue measures. In many of these studies, however, there often exist short-term outcomes that are easier and faster to measure, e.g., short-term health measures or an online ad's click-through rate, that are observed for a greater share of the sample. These outcomes, which are typically informative about the primary outcomes themselves, are referred to as *surrogate outcomes* or *surrogates*.

There is a rich causal inference literature addressing missing outcome data. Simply restricting to data with observed primary outcomes may induce strong bias (Hernán & Robins, 2010). Ignoring unlabeled data also reduces the effective sample size for estimating the treatment effects and inflates the variance. Chakrabortty et al. (2022) considered the missing completely at random (MCAR) setting and showed that incorporating unlabeled data reduces variance. Zhang et al. (2023) generalized the results to missing-at-random (MAR) settings where the unlabeled data has a much larger size than the labeled data. Kallus & Mao (2020) further examined the role of surrogates in datasets with limited primary outcomes and showed efficiency gains after including surrogates and unlabeled data in the analysis. Singh (2022) proposed a generalized kernel ridge regression framework, which incorporates information on surrogate outcomes in treatment effect estimation. See also Zhang & Bradic (2022); Hou et al. (2021); Zeng et al. (2023) for relevant discussions.

Continuous treatments appear in many applications; e.g., waiting time before follow-up, percent of discount, and drug dosage. Existing estimation procedures include outcome modeling (Newey, 1994) and treatment process modeling (Galvao & Wang, 2015). Kennedy et al. (2017) proposed doubly robust methods that model both the outcome and the treatment process, and enjoy appealing robustness properties. Bonvini & Kennedy (2022) further examined high-order estimators to achieve faster rates.

In this paper, we consider the estimation of dose-response functions with limited primary outcome data. We propose novel doubly robust methods using both labeled and unlabeled data, with the help of surrogates. Our approach avoids the potential selection bias caused by restricting only to labeled data and provably reduces the variance. Importantly, we study the theoretical properties of the estimator proposed, and establish its asymptotic normality to facilitate statistical inference. Our work serves as a counterpart for Kallus & Mao (2020) in continuous treatment effects

[1]Department of Statistics and Data Science, Carnegie Mellon University, Pittsburgh, PA, USA [2]Adobe Research, San Jose, CA, USA [3]Goldman School of Public Policy and Department of Statistics, University of California, Berkeley, Berkeley, CA, USA. Correspondence to: Zhenghao Zeng <zhenghaz@andrew.cmu.edu>.

Proceedings of the 41st International Conference on Machine Learning, Vienna, Austria. PMLR 235, 2024. Copyright 2024 by the author(s).

settings and enriches the existing dose-response estimation literature (Kennedy et al., 2017; Bonvini & Kennedy, 2022).

The rest of this paper is organized as follows: In Section 2 we introduce the problem setup and notation. Section 3 provides assumptions to identify the continuous treatment effect. Novel methodology and theoretical guarantees are discussed in Section 4. Our simulation study in Section 6 demonstrates the performance of our method. Finally we conclude with a discussion in Section 7. All proofs, additional simulation results and a real data example are included in supplementary materials.

2. Setup and Notation

In this section we introduce the problem of estimating continuous treatment effects with surrogates. We formalize the problem using potential outcomes framework (Splawa-Neyman et al., 1990; Rubin, 1974) and introduce notation to present our results concisely.

2.1. Data Structure

Suppose we have access to two datasets $\mathcal{L}$ and $\mathcal{U}$ from a randomized experiment/observational study. The labeled dataset is $\mathcal{L} = \{\mathbf{Z}_i = (\mathbf{V}_i, \mathbf{S}_i, A_i, Y_i, R_i = 1), 1 \leq i \leq n_1\}$, where $\mathbf{V}$ is the set of covariates, $\mathbf{S}$ is the vector of surrogates, A is the continuous treatment, and Y is the primary outcome. Here R is an indicator with $R = 1$ if a sample is from the labeled data $\mathcal{L}$ and $R = 0$ otherwise. The unlabeled dataset consists of samples without primary outcomes $\mathcal{U} = \{\mathbf{Z}_i = (\mathbf{V}_i, \mathbf{S}_i, A_i, R_i = 0), n_1 + 1 \leq i \leq n_1 + n_2\}$. Hence the total sample size is $n = n_1 + n_2$ and both $\mathcal{L}$ and $\mathcal{U}$ contain information on treatment A, surrogates $\mathbf{S}$ and covariates $\mathbf{V}$, but the primary outcome is missing in the unlabeled dataset $\mathcal{U}$ due to, for example, loss to follow-up. Let $\mathbf{X} = (\mathbf{V}, \mathbf{S})$ be the union of covariates and surrogates. Note that we will present the results for $\mathbf{S} \neq \emptyset$, but $\mathbf{S} = \emptyset$ can be viewed as a special setting where our methodology still applies.

2.2. Estimand and Nuisance Functions

Now we define the continuous treatment effects of interest. We use the random variable Y^a to denote the potential (counterfactual) outcome we would have observed had a subject received treatment $A = a$, which may be contrary to the observed Y. The continuous treatment effect (or dose-response function) is defined as

$$\theta(a) = \mathbb{E}[Y^a]. \tag{1}$$

Without additional causal assumptions introduced in Section 3, $\theta(a)$ involves counterfactual outcome Y^a and cannot be identified by observed data. Note that since $\mathbf{S}$ contains post-treatment surrogate outcomes and may be affected by the treatment, we also use potential outcomes $\mathbf{S}^a$ to denote the surrogate under treatment $A = a$.

We further introduce some nuisance functions that are not of primary interest, but which our estimation method depends on. Let $\mu(A, \mathbf{X}, 1) = \mathbb{E}[Y|A, \mathbf{X}, R = 1]$ be the regression function of the primary outcome in the labeled population $R = 1$. Denote the function obtained by further regressing $\mu(A, \mathbf{X}, 1)$ on $(A, \mathbf{V})$ as $\tau(A, \mathbf{V}) = \mathbb{E}[\mu(A, \mathbf{X}, 1)|A, \mathbf{V}]$, where the expectation is over the conditional distribution of $\mathbf{S}$ given $A, \mathbf{V}$. Denote the conditional density of A given $\mathbf{V}$ as $\pi(a|\mathbf{V})$ and the marginal density of A as $f(a) = \int \pi(a|\mathbf{v}) d\mathbb{P}(\mathbf{v})$. The ratio of the marginal density and conditional density is denoted as $w(a, \mathbf{v}) = f(a)/\pi(a|\mathbf{v})$. $w(a, \mathbf{v})$ is known as a stabilized weight in the literature (Robins et al., 2000; Ai et al., 2021). The propensity score for R (i.e. the conditional probability that the primary outcome is observed) is denoted as $\rho(A, \mathbf{X}) = \mathbb{P}(R = 1|A, \mathbf{X})$.

Let $(\mathcal{V}, \mathcal{S}, \mathcal{A})$ denote the support of $(\mathbf{V}, \mathbf{S}, A)$. For a (possibly random) function f on variables $\mathbf{Z}$ we use $\mathbb{P}_n[f(\mathbf{Z})]$ or $\mathbb{P}_n[f]$ to denote the sample average $\frac{1}{n}\sum_{i=1}^n f(\mathbf{Z}_i)$ on a sample of size n. The sample over which averages are taken should be clear from context. We use $\mathbb{P}[f] = \int f(\mathbf{z}) d\mathbb{P}(\mathbf{z})$ to denote the expectation of $f(\mathbf{Z})$ where only randomness of $\mathbf{Z}$ is considered and f is conditioned on. Finally we use $\|f\|_\infty = \sup_{\mathbf{z} \in \mathcal{Z}} |f(\mathbf{z})|$ to denote the uniform norm, $\|f\|_a = \left\{\int f^2(\mathbf{z}) d\mathbb{P}(\mathbf{z}|A = a)\right\}^{1/2}$ to denote the L_2-norm with respect to the conditional distribution $\mathbf{Z}|A = a$ and $\|f\|_2 = \left\{\int f^2(\mathbf{z}) d\mathbb{P}(\mathbf{z})\right\}^{1/2}$ to denote the usual L_2-norm.

3. Identification

In this section we discuss sufficient conditions to identify the dose-response function (1), summarized as follows:

Assumption 3.1. (Consistency) $Y = Y^a, \mathbf{S} = \mathbf{S}^a$ if $A = a, a \in \mathcal{A}$.

Assumption 3.2. (Exchangeability) $(Y^a, \mathbf{S}^a) \perp\!\!\!\perp A \mid \mathbf{V}$ for $a \in \mathcal{A}$.

Assumption 3.3. (Missing at random) $R \perp\!\!\!\perp Y^a \mid \mathbf{V}, \mathbf{S}^a, A = a$ for $a \in \mathcal{A}$.

Assumption 3.4. (Positivity) $\pi(a|\mathbf{V}) > 0, \rho(a, \mathbf{X}) > 0$ almost surely for $a \in \mathcal{A}$.

Assumption 3.1 is also known as the stable unit treatment value assumption (SUTVA), and requires an absence of interference between different individuals in the study. Assumption 3.2 is commonly used to identify treatment effects and holds in a randomized experiment or observational study with all confounders measured. Assumption 3.3 ensures whether the primary outcome is observed only depends on covariates $\mathbf{V}$, surrogates $\mathbf{S}$, and treatment A so that the distributions of labeled and unlabeled data are comparable after

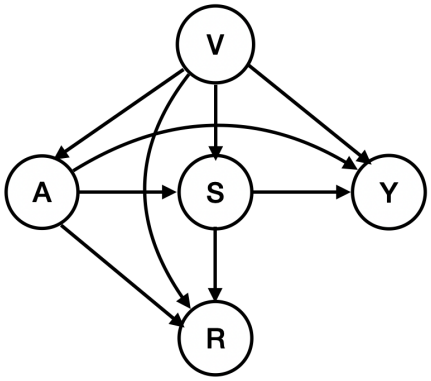


Figure 1. Example of a causal graph with surrogate outcome $\mathbf{S}$.

conditioning on $\mathbf{X}, A$. Assumption 3.4 says every subject has some chance to receive treatment $A = a$ and has the primary outcome observed. An illustrative causal graph is shown in Figure 1. The readers are referred to Gill & Robins (2001) for detailed discussions on identifying continuous treatment effects and Kallus & Mao (2020) for the role of surrogates in identifying treatment effects. With all these assumptions, the treatment effects of interest can be identified using the observable distribution as summarized in Theorem 3.5.

Theorem 3.5. *Under Assumption 3.1–3.4 we have*

$$\begin{aligned}\theta(a) &= \mathbb{E}\{\mathbb{E}[\mathbb{E}(Y|A=a,\mathbf{X},R=1)|A=a,\mathbf{V}]\} \\ &= \mathbb{E}\left\{\mathbb{E}[\mu(a,\mathbf{X},1)|A=a,\mathbf{V}]\right\} \\ &= \mathbb{E}[\tau(a,\mathbf{V})]\end{aligned} \tag{2}$$

for fixed $a \in \mathcal{A}$, where the expectations are over $Y, \mathbf{S}, \mathbf{V}$ in (2).

The identification formula (2) suggests the following plug-in style estimator of $\theta(a)$: first regress Y on $A, \mathbf{X}$ in the labeled dataset $\mathcal{L}$ and obtain an estimator of μ as $\widehat{\mu}$, which is further regressed on $A, \mathbf{V}$ to get an estimator $\widehat{\tau}$. Finally we take an average over all the samples to get an estimator of $\theta(a)$ as

$$\widehat{\theta}(a) = \mathbb{P}_n[\widehat{\tau}(a,\mathbf{V})]. \tag{3}$$

The performance of such a plug-in style estimator highly depends on the estimation error of $\widehat{\tau}$. To see this, assume the nuisance estimator $\widehat{\tau}$ is independent of the samples that we average over, then the conditional bias of $\widehat{\theta}(a)$ given $\widehat{\tau}$ is

$$\mathbb{E}\left[\widehat{\theta}(a)\right] - \theta(a) = \mathbb{E}\left[\widehat{\tau}(a,\mathbf{V}) - \tau(a,\mathbf{V})\right],$$

which solely depends on the estimation error of τ. When τ is hard to estimate (e.g., non-smooth/sparse in high-dimensional problems), the plug-in style estimator may inherit the slow convergence rate of $\widehat{\tau}$ and have sub-optimal performance.

4. Doubly Robust Estimation

In this section we present the novel method of this paper. We begin with an alternative characterization of the dose-response function (1) in Section 4.1, based on which we propose a doubly robust estimator in Section 4.2.

4.1. Doubly Robust Characterization

Since treatment A is continuous, the function $\theta(a)$ in (2) is not pathwise differentiable (Bickel et al., 1993; Díaz & van der Laan, 2013) and we need a novel way to apply semiparametric efficiency theory. The idea is to find a pseudo-outcome $\varphi(\mathbf{Z}) := \varphi(\mathbf{Z};\mu,\pi,\rho)$ depending on nuisance functions such that $\mathbb{E}[\varphi(\mathbf{Z};\mu,\pi,\rho)|A=a] = \theta(a)$, i.e., regressing $\varphi(\mathbf{Z};\mu,\pi,\rho)$ on A yields the target dose-response function ideally with second-order dependence on nuisance estimation error. Following Rubin & van der Laan (2005); Kennedy et al. (2017), we consider the functional $\psi = \mathbb{E}[\theta(A)]$, which is pathwise differentiable and admits an efficient influence function. Then the pseudo-outcome $\varphi(\mathbf{Z};\mu,\pi,\rho)$ is a component in the influence function of ψ. We omit the derivation of the influence function and only present the form of pseudo-outcome. Let $(\bar{\mu},\bar{\pi},\bar{\rho})$ be nuisance functions that may not necessarily equal the true (μ,π,ρ), and define

$$\begin{aligned}&\varphi(\mathbf{Z};\bar{\mu},\bar{\pi},\bar{\rho}) \\ &= \left[\frac{R(Y-\bar{\mu}(A,\mathbf{X},1))}{\bar{\rho}(A,\mathbf{X})} + \bar{\mu}(A,\mathbf{X},1) - \bar{\tau}(A,\mathbf{V})\right] \\ &\quad \times \frac{\int_{\mathcal{V}} \bar{\pi}(A|\mathbf{v})d\mathbb{P}(\mathbf{v})}{\bar{\pi}(A|\mathbf{V})} + \int_{\mathcal{V}} \bar{\tau}(A,\mathbf{v})d\mathbb{P}(\mathbf{v}),\end{aligned}$$

where $\bar{\tau}(A,\mathbf{V}) = \mathbb{E}[\bar{\mu}(A,\mathbf{X},1)|A,\mathbf{V}]$. The following theorem shows an alternative characterization of the dose-response function $\theta(a)$ through $\varphi(\mathbf{Z};\bar{\mu},\bar{\pi},\bar{\rho})$.

Proposition 4.1. *Let $(\bar{\mu},\bar{\pi},\bar{\rho})$ be nuisance functions that may not necessarily equal the true (μ,π,ρ). Then*

$$\mathbb{E}[\varphi(\mathbf{Z};\bar{\mu},\bar{\pi},\bar{\rho})|A=a] = \theta(a)$$

if either $\bar{\mu} = \mu$ or $(\bar{\pi},\bar{\rho}) = (\pi,\rho)$.

Proposition 4.1 gives the first interpretation of the double robustness of our methods. There are two chances to obtain the dose-response function $\theta(a)$ by regressing the pseudo-outcome $\varphi(\mathbf{Z};\bar{\mu},\bar{\pi},\bar{\rho})$ on A: we correctly specify either the outcome regression model μ or both the propensity score of R and conditional density of A given $\mathbf{V}$ (although see Proposition 5.2 for a slightly different parameterization).

4.2. Estimation Procedure

The doubly robust characterization in Proposition 4.1 motivates a two-stage procedure to estimate $\theta(a)$: In the first step

Algorithm 1 Doubly Robust Estimation

Let (D_1^n, D_2^n, T^n) denote three independent samples of n i.i.d observations of $\mathbf{Z}$ and $D^n = (D_1^n, D_2^n)$ denote the training set to train the nuisance functions.

1. Nuisance functions training: Construct estimates of μ, τ, ρ, w using D_1^n. Then use D_2^n to get an initial estimator of $\theta(a)$ as
$$\widehat{\theta}_0(a) = \frac{1}{n} \sum_{i \in D_2^n} \widehat{\tau}(a, \mathbf{V}_i).$$
2. Pseudo-outcome regression: Construct estimated pseudo-outcome
$$\begin{aligned} &\widehat{\varphi}(\mathbf{Z}) \\ &= \left[\frac{R(Y - \widehat{\mu}(A, \mathbf{X}, 1))}{\widehat{\rho}(A, \mathbf{X})} + \widehat{\mu}(A, \mathbf{X}, 1) - \widehat{\tau}(A, \mathbf{V}) \right] \\ &\quad \times \widehat{w}(A, \mathbf{V}) + \widehat{\theta}_0(A) \end{aligned}$$
for each sample in T^n and regress the pseudo-outcomes on the treatment A in T^n using a linear smoother to obtain
$$\widehat{\theta}(a) = \widehat{\mathbb{E}}_n \left[\widehat{\varphi}(\mathbf{Z}) | A = a \right] = \sum_{i \in T^n} W_i(a; \mathbf{A}^n) \widehat{\varphi}(\mathbf{Z}_i),$$
where $\mathbf{A}^n = (A_1, \ldots, A_n)$ are the treatments in T^n and $W_i(a; \mathbf{A}^n)$ is the coefficient of i-th sample.
3. (Optional) Cross-fitting: Swap the role of D_1^n, D_2^n, T^n and repeat steps 1 and 2. Use the average of different estimates as the final estimator of $\theta(a)$.

we model the nuisance functions that appear in the pseudo-outcome $\varphi(\mathbf{Z})$ with flexible nonparametric or machine learning methods. We then construct estimated pseudo-outcomes $\widehat{\varphi}(\mathbf{Z})$ and regress these on A to obtain an estimator of $\theta(a)$. We will use the stability framework developed in Kennedy (2023) to analyze such an estimator, which regresses imputed outcomes $\widehat{\varphi}(\mathbf{Z})$ on treatment. The formal procedure is summarized in Algorithm 1.

Sample splitting is used in Algorithm 1 to avoid complicated empirical process assumptions that are difficult to justify in practice and simplify our theoretical analysis (Robins et al., 2008; Chernozhukov et al., 2018; Kennedy et al., 2020; Levis et al., 2023; Bonvini et al., 2023). In Step 1, one can use any appropriate regression/classification algorithms to estimate μ, τ, ρ. However, there are fewer results on estimating stabilized weight $w(a, \mathbf{v}) = f(a)/\pi(a|\mathbf{v})$. Ai et al. (2021) proposed a method that directly estimates w with entropy maximization. Alternatively, one can estimate the conditional density π (see, e.g., Colangelo & Lee, 2020, for discussions on related methods), and then the marginal density can be estimated by
$$\widehat{f}(a) = \frac{1}{n} \sum_{i \in D_2^n} \widehat{\pi}(a|\mathbf{V}_i)$$
and use their ratio to estimate w.

Our analysis in Section 5 applies to a wide class of nuisance estimators if the product of convergence rates is fast enough. In step 2 we focus on linear smoother-based estimators since they are relatively straightforward to analyze. We believe similar results hold for a wider class of regression estimators under the stability framework in Kennedy (2023) and leave the theoretical analysis of applying general regression algorithms for future investigation. In applications, researchers can choose suitable parametric methods based on their domain knowledge or flexible nonparametric machine learning methods to avoid model misspecification (or ensembles thereof).

5. Theoretical Results

In this section, we provide theoretical guarantees of the proposed method. We first present estimation error guarantees of Algorithm 1, for general linear smoothers. Then we focus on a specific type of linear smoother, namely local linear regression, and study the asymptotic distribution of the estimator in Section 5.2.

5.1. Oracle Estimation Theory

In Algorithm 1 we obtain the estimator $\widehat{\theta}(a)$ by regressing the imputed pseudo-outcome $\widehat{\varphi}(\mathbf{Z})$ on A. It is natural to compare $\widehat{\theta}(a)$ with the "oracle" estimator $\widetilde{\theta}(a)$ defined as
$$\widetilde{\theta}(a) = \sum_{i \in T^n} W_i(a; \mathbf{A}^n) \varphi(\mathbf{Z}_i),$$
which regresses the ground-truth pseudo-outcome $\varphi(\mathbf{Z})$ on A using the same linear smoother. Intuitively it is hard for $\widehat{\theta}$ to have a faster convergence rate than $\widetilde{\theta}$. In the following theorem, we summarize the conditions under which $\widehat{\theta}$ enjoys the same rate as $\widetilde{\theta}$ and hence is "oracle efficient".

Theorem 5.1. *Let $\widehat{\theta}(a)$ denote the doubly robust estimator obtained from Algorithm 1 and $\widetilde{\theta}(a) = \sum_{i \in T^n} W_i(a; \mathbf{A}^n) \varphi(\mathbf{Z}_i)$ denote the oracle estimator with oracle risk $R_n^2(a) = \mathbb{E}[(\widetilde{\theta}(a) - \theta(a))^2]$ at point a. Suppose*

- $\mathrm{Var}(\varphi(\mathbf{Z})|A = a) \geq c, \ \forall\, a \in \mathcal{A}$ *for some constants* $c > 0$.
- *The estimator for nuisance functions in Step 2 is uniformly consistent in the sense that* $\sup_{\mathbf{z}} |\widehat{\varphi}(\mathbf{z}) - \varphi(\mathbf{z})| = o_{\mathbb{P}}(1)$.

Then we have

$$\widehat{\theta}(a)-\theta(a) = \widetilde{\theta}(a)-\theta(a)+\widehat{\mathbb{E}}_n\left[\widehat{b}(A)|A=a\right]+o_{\mathbb{P}}\left(R_n(a)\right)$$

where $\widehat{b}(a) = \mathbb{E}[\widehat{\varphi}(\mathbf{Z}) - \varphi(\mathbf{Z})|D^n, A = a]$ *is the conditional bias of* $\widehat{\varphi}(\mathbf{Z})$ *and* $\widehat{\mathbb{E}}_n\left[\widehat{b}(A)|A=a\right] = \sum_{i\in T^n} W_i(a;\mathbf{A}^n)\widehat{b}(A_i)$. *Further assume*

$$\widehat{\mathbb{E}}_n\left[\widehat{b}(A)|A=a\right] = o_{\mathbb{P}}(R_n(a)),$$

then $\widehat{\theta}$ *is oracle efficient in the sense that*

$$\frac{\widehat{\theta}(a)-\widetilde{\theta}(a)}{R_n(a)} \xrightarrow{P} 0.$$

The conditions in Theorem 5.1 are mild: $\varphi(\mathbf{Z})$ is not a function of A and hence its conditional variance given A should be positive; We do not impose conditions on the convergence rate of $\widehat{\varphi}$ but only require its consistency. The key condition for $\widehat{\theta}$ to be oracle efficient is $\widehat{\mathbb{E}}_n\left[\widehat{b}(A)|A=a\right] = o_{\mathbb{P}}(R_n(a))$ and hence we need a bound on the conditional bias $\widehat{b}(a)$, as summarized in the following proposition.

Proposition 5.2. *Under the conditions in Theorem 5.1, further assume the estimated conditional probability* $\widehat{\rho}(a,\mathbf{x}) \geq c$ *and the estimated stabilized weight* $\widehat{w}(a,\mathbf{v}) \leq C$ *hold for all* $a \in \mathcal{A}, \mathbf{x} \in \mathcal{X}, \mathbf{v} \in \mathcal{V}$ *for some constant* $c, C > 0$. *Then we have*

$$|\widehat{b}(a)| \lesssim \|\widehat{\rho}-\rho\|_a\|\widehat{\mu}-\mu\|_a + \|\widehat{\tau}-\tau\|_a\|\widehat{w}-w\|_a + \frac{1}{n}\sum_{i\in D_2^n}\widehat{\tau}(a,\mathbf{V}_i) - \mathbb{P}[\widehat{\tau}(a,\mathbf{V})],$$

where recall $\|f\|_a^2 = \int f^2(\mathbf{z})d\mathbb{P}(\mathbf{z}|A=a)$. *If we further assume the weights of the linear smoother satisfy* $\sum_{i\in T^n} W_i(a;\mathbf{A}^n) \leq C$ *and there exists a neighborhood* $N(a)$ *around* a *such that* $W_i(a;\mathbf{A}^n) = 0$ *if* $A_i \notin N(a)$. *Then we have*

$$\widehat{\mathbb{E}}_n\left[\widehat{b}(A)|A=a\right] \lesssim \sup_{t\in N(a)}|\widehat{b}(t)|.$$

We note that the condition on linear smoothers will be satisfied by Nadaraya–Watson estimators (Nadaraya, 1964; Watson, 1964) and local polynomial estimators (Fan et al., 1994; Fan & Gijbels, 1996) when the kernel function used has compact support, e.g., the uniform and Epanechnikov kernel. In the bound for $\widehat{b}(a)$, the first two terms depend on the product of the convergence rates of nuisance functions. This phenomenon is commonly observed in influence functions-based doubly robust approaches (Kennedy et al., 2016; Chernozhukov et al., 2018; Meza & Singh, 2021) and gives the second interpretation of double robustness: Introducing an extra term to the plug-in style estimator in φ can correct for the first-order bias, making the remainder second-order and "doubly small". In common examples of ATE estimation, conditional bias only involves estimation error of outcome model and propensity scores. The additional dependence on the convergence rate of $\hat{\tau}$ (compared with Proposition 4.1) appears in the bias since the formula (2) is more complicated compared with ATE-style functional and we use an agnostic estimator of τ in Algorithm 1. In most settings we expect the nuisance error rate $\|\widehat{\mu}-\mu\|_a$ (with respect to measure $d\mathbb{P}(\mathbf{z}|A=a)$) has the same order as the more common conventional rate $\|\widehat{\mu}-\mu\|$ (with respect to measure $d\mathbb{P}(\mathbf{z})$), which is $n^{-\alpha/(2\alpha+d+1)}$ if $\mu(a,\mathbf{x})$ belongs to a Hölder class of order α and $\mathbf{X}$ is d-dimensional. Alternatively one can always upper bound $\|\widehat{\mu}-\mu\|_a$ with $\|\widehat{\mu}-\mu\|_\infty$ at the cost of log factors (Tsybakov, 2009, Section 1.6.2). The empirical process term $\frac{1}{n}\sum_{i\in D_2^n}\widehat{\tau}(a,\mathbf{V}_i) - \mathbb{P}[\widehat{\tau}(a,\mathbf{V})]$ would be $O_{\mathbb{P}}(1/\sqrt{n})$ provided that $\mathbb{E}[\widehat{\tau}^2(a,\mathbf{V})|D^n]$ is bounded. The bound on $\widehat{\mathbb{E}}_n\left[\widehat{b}(A)|A=a\right]$ in Proposition 5.2 involves

$$\sup_{t\in N(a)}\left|\frac{1}{n}\sum_{i\in D_2^n}\widehat{\tau}(t,\mathbf{V}_i) - \mathbb{P}[\widehat{\tau}(t,\mathbf{V})]\right|$$

as a coarse bound and for specific estimators, a tighter bound can be derived. For instance, as we will see in local linear estimation this empirical process term is $o_{\mathbb{P}}\left(1/\sqrt{nh}\right)$ and asymptotically negligible. Hence we can focus on the first two second-order terms in the conditional bias. Theorem 5.1 together with Proposition 5.2 gives conditions under which the estimator $\widehat{\theta}(a)$ has the same rate as the oracle estimator $\widetilde{\theta}(a)$. For instance, assume $\mathbf{V} = \mathbf{X}$ (no surrogates) and $|\mathbf{X}| = d$, suppose the dose-response function $\theta(a)$ is α-smooth (i.e. belongs to a Hölder class of order α) and w and μ are s-smooth, then the rate condition for $\widehat{\theta}(a)$ to be oracle efficient is

$$n^{-\frac{2s}{2s+d+1}} \leq n^{-\frac{\alpha}{2\alpha+1}}$$

or equivalently

$$s \geq \frac{\alpha(d+1)}{2(\alpha+1)}.$$

5.2. Asymptotic Normality

In the following discussions, we will analyze the estimator $\widehat{\theta}(a)$ based on a particular linear smoother (i.e. the local linear regression estimator). For a scalar bandwidth parapmeter h, let $\mathbf{g}_{ha}(t) = [1,(t-a)/h]^\top$ be the local linear basis, $K_{ha}(t) = h^{-1}K((t-a)/h)$ with K being a probability density. The local linear regression solves the following weighted least square problem:

$$\min_{\boldsymbol{\beta}\in\mathbb{R}^2}\sum_{i\in T^n}K_{ha}(A_i)\left[\widehat{\varphi}(\mathbf{Z}_i) - \mathbf{g}_{ha}(A_i)^\top\boldsymbol{\beta}\right]^2$$

which gives the closed-form solution:

$$\widehat{\boldsymbol{\beta}}_h(a) = \widehat{\mathbf{D}}_{ha}^{-1}\mathbb{P}_n[\mathbf{g}_{ha}(A)K_{ha}(A)\widehat{\varphi}(\mathbf{Z})],$$

where $\widehat{\mathbf{D}}_{ha} = \mathbb{P}_n[\mathbf{g}_{ha}(A)K_{ha}(A)\mathbf{g}_{ha}^\top(A)]$ and $\mathbb{P}_n$ is the sample average over T^n. Then the local linear estimator of $\theta(a)$ is $\mathbf{e}_1^\top\widehat{\boldsymbol{\beta}}_h(a)$, i.e. the first component of $\widehat{\boldsymbol{\beta}}_h(a)$. We summarize the asymptotic properties of this local linear estimator in Theorem 5.3, which is the key contribution of our paper.

Theorem 5.3. *(Asymptotic normality of Local Linear Estimator) Let $a \in \mathcal{A}$ be an inner point of the compact support $\mathcal{A}$ of A. Assume*

1. *The bandwidth h satisfies $h = h_n \to 0$ and $nh \to \infty$ as $n \to \infty$.*
2. *The marginal density of the treatment $f(a)$ is continuously differentiable, the conditional variance $\mathrm{Var}(\varphi(\mathbf{Z})|A = a)$ is continuous and the dose response $\theta(a)$ is twice continuously differentiable.*
3. *For any $a \in \mathcal{A}, \mathbf{x} \in \mathcal{X}, \mathbf{v} \in \mathcal{V}$, the conditional variance $\mathrm{Var}(\varphi(\mathbf{Z})|A = a) > 0$, the marginal density of treatment $f(a) \geq c$, the estimated conditional probability $\widehat{\rho}(a, \mathbf{x}) \geq c$ and the estimated stablized weight $\widehat{w}(a, \mathbf{v}) \leq C$ for some constant $C, c > 0$.*
4. *K is a continuous symmetric probability density with support $[-1, 1]$.*
5. *All nuisance functions are estimated consistently in ℓ_∞ norm and the estimated pseudo-outcome also satisfies*

$$\|\widehat{\varphi} - \varphi\|_\infty = o_{\mathbb{P}}(1).$$

Furthermore, the convergence rates of nuisance estimation satisfy

$$\sup_{|t-a|\leq h} \|\widehat{\rho} - \rho\|_t \|\widehat{\mu} - \mu\|_t = o_{\mathbb{P}}\left(\frac{1}{\sqrt{nh}}\right)$$
$$\sup_{|t-a|\leq h} \|\widehat{\tau} - \tau\|_t \|\widehat{w} - w\|_t = o_{\mathbb{P}}\left(\frac{1}{\sqrt{nh}}\right).$$

Then we have

$$\sqrt{nh}\left(\widehat{\theta}(a) - \theta(a) - \frac{h^2\theta''(a)\int u^2K(u)du}{2}\right) \xrightarrow{d} N\left(0, \frac{\sigma^2(a)\int K^2(u)du}{f(a)}\right)$$

where

$$\sigma^2(a) = \mathbb{E}\left\{\left[\frac{\mathrm{Var}(Y|A = a, \mathbf{X}, R = 1)}{\rho(a, \mathbf{X})} + \mathrm{Var}(\mu(a, \mathbf{X}, 1)|A = a, \mathbf{V})\right] w^2(a, \mathbf{V})\middle| A = a\right\}$$

Assumptions 1-4 in Theorem 5.3 are standard in the nonparametric kernel regression literature. Assumption 5 guarantees that the contribution of nuisance estimation error is asymptotically negligible compared with the smoothing error. One can also use a symmetric kernel K supported on $\mathbb{R}$ that is square-integrable and has a finite second-order moment. Then the rate condition would be $\sup_{t\in\mathcal{A}} \|\widehat{\rho} - \rho\|_t\|\widehat{\mu} - \mu\|_t = o_{\mathbb{P}}\left(\frac{1}{\sqrt{nh}}\right)$. Similar to our discussions in Section 5.1, the rate conditions in assumption 5 are imposed on the product of nuisance estimation error since we use a doubly robust estimator and the conditional bias is "second-order small". The theoretically optimal bandwidth to estimate a twice continuously differentiable function is $h \asymp n^{-1/5}$ and yields a root mean square error of order $n^{-2/5}$. With such a choice of h the requirement on the convergence rate becomes $\sup_{t\in\mathcal{A}} \|\widehat{\rho} - \rho\|_t\|\widehat{\mu} - \mu\|_t = o_{\mathbb{P}}\left(n^{-2/5}\right)$, which, for example, can be satisfied when ρ is consistent and μ is estimated with rate $O_{\mathbb{P}}(n^{-2/5})$. In applications, one can select the bandwidth using leave-one-out cross-validation (Härdle et al., 1988) due to its computational ease. Specifically, after we obtain the estimated pseudo-outcome, we treat them as known and select h by

$$\widehat{h}_{opt} = \underset{h\in\mathcal{H}}{\mathrm{argmin}} \sum_i \left\{\frac{\widehat{\varphi}(\mathbf{Z}_i) - \widehat{\theta}_h(A_i)}{1 - \widehat{W}_h(A_i)}\right\}^2. \tag{4}$$

where $\widehat{W}_h(A_i) = \mathbf{e}_1^\top\widehat{\mathbf{D}}_{hA_i}^{-1}\mathbf{e}_1 h^{-1}K(0)$. In the setting of Algorithm 1 we can select the bandwidth on D_2^n to avoid overfitting on T^n.

Similar to most nonparametric inference methods, Theorem 5.3 shows the estimator $\widehat{\theta}$ is centered around $\bar{\theta}(a) = \theta(a) - h^2\theta''(a)\int u^2K(u)du/2$ instead of $\theta(a)$ under optimal smoothing, which is known as the "bias problem" in the literature (Wasserman, 2006, Section 5.7). There are several methods to overcome the bias problem and each of them has its own consideration and trade-offs. For instance, one can estimate the second-order derivative and debias the estimator (Calonico et al., 2018; Takatsu & Westling, 2022) but this requires extra smoothness conditions. Another method is to undersmooth (Fan et al., 2022) and make the bias decrease asymptotically relative to the variance. Unfortunately, there does not seem to be a simple, practical rule for choosing just the right amount of undersmoothing. In this paper we choose to live with the bias and report uncertainty quantification for $\bar{\theta}$. Theoretically, the bias shrinks to 0 as $n \to \infty$ and the proposed estimator $\hat{\theta}(a)$ is still consistent for $\theta(a)$. To construct confidence intervals for $\bar{\theta}(a)$ one needs to estimate the variance. Define a localized functional $\theta_h(a) = \mathbf{e}_1^\top\mathbf{D}_{ha}^{-1}\mathbb{E}[\mathbf{g}_{ha}(A)K_{ha}(A)\theta(A)]$ (which can be viewed as population version of local linear estimator

$\mathbf{e}_1^\top \widehat{\boldsymbol{\beta}}_h(a))$ with efficient influence function

$$\begin{aligned}&\phi_{ha}(\mathbf{Z})\\ =&\mathbf{e}_1^\top \mathbf{D}_{ha}^{-1}\mathbf{g}_{ha}(A)K_{ha}(A)\\ &\times\left(\varphi(\mathbf{Z})-\mathbf{g}_{ha}^\top(A)\mathbf{D}_{ha}^{-1}\mathbb{E}[\mathbf{g}_{ha}(A)K_{ha}(A)\theta(A)]\right)\\ &+\mathbf{e}_1^\top \mathbf{D}_{ha}^{-1}\int \mathbf{g}_{ha}(t)K_{ha}(t)\tau(t,\mathbf{V})f(t)dt-\theta_h(a).\end{aligned}$$

Following Kennedy et al. (2017); Takatsu & Westling (2022), one can show the variance of $\widehat{\theta}(a)$ can be approximated by $\frac{1}{n}\mathbb{P}_n\left[\left(\widehat{\phi}_{ha}(\mathbf{Z})\right)^2\right]$ for

$$\begin{aligned}&\widehat{\phi}_{ha}(\mathbf{Z})\\ =&\mathbf{e}_1^\top \widehat{\mathbf{D}}_{ha}^{-1}\mathbf{g}_{ha}(A)K_{ha}(A)\left(\widehat{\varphi}(\mathbf{Z})-\mathbf{g}_{ha}^\top(A)\widehat{\beta}_h(a)\right)\\ &+\mathbf{e}_1^\top \widehat{\mathbf{D}}_{ha}^{-1}\int \mathbf{g}_{ha}(t)K_{ha}(t)\widehat{\tau}(t,\mathbf{V})d\mathbb{P}_n(t)-\widehat{\theta}(a).\end{aligned}$$

Finally, we compare the asymptotic variance in Theorem 5.3 with the asymptotic variance in Kennedy et al. (2017), where the unlabeled dataset $\mathcal{U}$ and surrogates $\mathbf{S}$ are unavailable. Consider the MCAR setting where R is independent of all other variables so that $\rho(a,\mathbf{x})=\rho\in(0,1)$, and for simplicity assume $n_1=n\rho$ to show how the surrogates and unlabeled data help to reduce the variance in this special setting. Note that since R is independent of all other variables, we have $\mathrm{Var}(Y|A=a,\mathbf{X},R=1)=\mathrm{Var}(Y|A=a,\mathbf{X})$ and $\mu(a,\mathbf{X},1)=\mathbb{E}[Y|A=a,\mathbf{X},R=1]=\mathbb{E}[Y|A=a,\mathbf{X}]=\mu(a,\mathbf{X})$. The asymptotic variance of $\widehat{\theta}(a)$ in our setting (i.e. in Theorem 5.3) is reduced to

$$\begin{aligned}&\frac{1}{nh}\mathbb{E}\left\{\left[\frac{\mathrm{Var}(Y|A=a,\mathbf{X})}{\rho}\right.\right.\\ &\left.\left.+\ \mathrm{Var}(\mu(a,\mathbf{X})|A=a,\mathbf{V})\right]w^2(a,\mathbf{V})\middle|A=a\right\}\end{aligned}\tag{5}$$

under the MCAR assumption (note the additional factor $\int K^2(u)du/f(a)$ is omitted since it appears in both settings). In the setting where the unlabeled data is unavailable, the asymptotic variance is shown to be

$$\frac{1}{n_1h}\mathbb{E}\left\{\mathrm{Var}(Y|A=a,\mathbf{V})w^2(a,\mathbf{V})|A=a\right\}\tag{6}$$

in Kennedy et al. (2017). By the property of conditional variance

$$\begin{aligned}\mathrm{Var}(Y|A=a,\mathbf{V})=&\mathbb{E}[\mathrm{Var}(Y|A=a,\mathbf{X})|A=a,\mathbf{V}]\\ &+\mathrm{Var}(\mathbb{E}[Y|A=a,\mathbf{X}]|A=a,\mathbf{V}),\end{aligned}$$

(6) can be re-written as

$$\begin{aligned}&\frac{1}{n_1h}\mathbb{E}\left\{(\mathrm{Var}(Y|A=a,\mathbf{X})\right.\\ &\left.+\mathrm{Var}(\mu(a,\mathbf{X})|A=a,\mathbf{V}))w^2(a,\mathbf{V})|A=a\right\}.\end{aligned}\tag{7}$$

Comparing (5) with (7) we see the first term is the same since $n\rho=n_1$. However the second term in (7) is improved by a factor of ρ in (5). This shows how the variance of the estimator is smaller after introducing unlabeled data and surrogate outcomes. The amount of improvement depends on the missing rate $1-\rho$ and $\mathrm{Var}(\mu(a,\mathbf{X})|A=a,\mathbf{V})$, which measures the variation of $\mu(A,\mathbf{V},\mathbf{S})$ that cannot be explained by $(A,\mathbf{V})$.

6. Simulation Study

In this section we use simulations to evaluate the performance of the proposed methods. We will illustrate the advantage of doubly robust estimation over naive plug-in style estimators. Consider the following data-generating process: The covariates $\mathbf{V}$ have a multi-variate Gaussian distribution

$$\mathbf{V}=(V_1,V_2,V_3,V_4)^\top\sim N(\mathbf{0},\mathbf{I}_4).$$

Conditioning on $\mathbf{V}$, the continuous treatment A has normal distribution $N(\lambda(\mathbf{V}),1)$ with

$$\lambda(\mathbf{V})=1+0.2V_1+0.2V_2-0.2V_3+0.3V_4.$$

The surrogates $\mathbf{S}$ has a normal distribution

$$\mathbf{S}=(S_1,S_2)^\top\sim N(\mathbf{0},\mathbf{I}_2).$$

The indicator of whether the outcome is observed or not R is Bernoulli(0.5) (so we assume a missing completely at random mechanism and $\rho=0.5$). Finally the outcome Y conditioning on $A,\mathbf{X},R=1$ has a normal distribution $N(\mu(A,\mathbf{X},1),1)$ with

$$\begin{aligned}\mu(A,\mathbf{X},1)=&1+(0.1,-0.1)^\top\mathbf{S}+(0.2,0.2,0.3,-0.1)^\top\mathbf{V}\\ &+A(1-0.1V_1+0.1V_3)-A^2.\end{aligned}$$

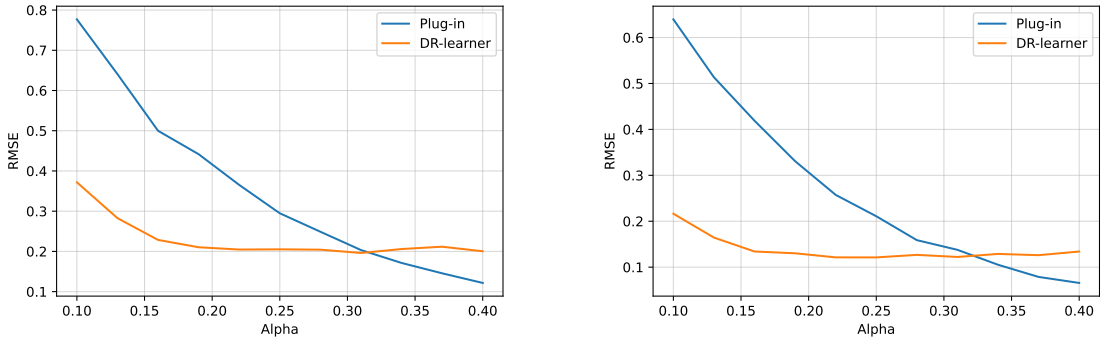
By direct calculations we have

$$\begin{aligned}\tau(A,\mathbf{V})=&1+(0.2,0.2,0.3,-0.1)^\top\mathbf{V}\\ &+A(1-0.1V_1+0.3V_3)-A^2,\end{aligned}$$

The dose-response function of interest is

$$\theta(A)=1+A-A^2.$$

To illustrate the performance of two estimators with different nuisance estimation errors we will manually set the estimation error, which is applicable for simulation purposes. For a fixed α we let $\epsilon_1,\ldots,\epsilon_4\sim N(n^{-\alpha},n^{-2\alpha})$ and set $\widehat{\lambda}(\mathbf{V})=\lambda(\mathbf{V})+\epsilon_1$, the estimated conditional density of A is $N\left(\widehat{\lambda}(\mathbf{V}),1\right)$, $\widehat{\mu}(A,\mathbf{X},1)=\mu(A,\mathbf{X},1)+\epsilon_2$, $\widehat{\tau}(A,\mathbf{V})=\tau(A,\mathbf{V})+\epsilon_3$, $\mathrm{logit}\left(\widehat{\rho}\right)=\mathrm{logit}(\rho)+\epsilon_4$. Such estimates guarantee the nuisance estimation error is of order $n^{-\alpha}$. After



(a) $n = 500$ (b) $n = 2000$

Figure 2. Root mean square error Versus α, where $n^{-\alpha}$ is the estimation error of the nuisance functions.

we generate a sample of size n, the plug-in style estimator is defined as $\widehat{\theta}(a) = \frac{1}{n}\sum_{i=1}^{n}\widehat{\tau}(a, \mathbf{V}_i)$. To implement the doubly robust estimator we split the sample into two parts D, T (since the nuisance estimators are given there is no need to split the sample into three parts as in Algorithm 1). We use D to construct estimator of the initial estimator $\widehat{\theta}_0(a) = \frac{1}{|D|}\sum_{i\in D}\widehat{\tau}(a, \mathbf{V}_i)$, marginal density $\widehat{f}(a) = \frac{1}{|D|}\sum_{i\in D}\widehat{\pi}(a|\mathbf{V}_i)$ and select the optimal bandwidth h^* (i.e. construct pseudo-outcomes and run LOOCV as in equation (4) on D). We then construct pseudo-outcomes on T and perform local linear estimation using the optimal bandwidth h^*. Finally, the roles of D and T are exchanged to obtain another estimator and we average the two estimates as the final doubly robust estimator. For sample size $n \in \{500, 2000\}$ and convergence rate $\alpha \in \{0.1, 0.13, \ldots, 0.4\}$, we repeat the data generation and estimation process $M = 500$ times. We will aim at estimating $\theta(1)$ and compare the RMSE$= \left\{\frac{1}{M}\sum_{m=1}^{M}\left[\widehat{\theta}_m(1) - \theta(1)\right]^2\right\}^{1/2}$ of plug-in estimator and doubly robust estimator, where $\widehat{\theta}_m(1)$ is the estimate from m-th repetition. The results are summarized in Figure 2.

As we see in Figure 2, if the nuisance estimation error is large (α is small), the doubly robust estimator outperforms the naive plug-in estimator. This can be explained by the second-order bias term of the doubly robust estimator in Proposition 5.2, i.e., the conditional bias is the product of nuisance estimation errors and is "doubly small". On the other hand, the plug-in style estimator inherits the slow convergence rate of $\widehat{\tau}$. As α increases, the estimators of nuisance functions are more accurate, and plug-in style estimators finally outperform the doubly robust estimator because the doubly robust estimator may suffer from accumulating error in constructing pseudo-outcomes, bandwidth selection, and local linear regression, which dominates the conditional bias when nuisance estimation is accurate enough. In real applications, there are typically many covariates and parametric models for nuisance functions may not be correct. The convergence rate of nonparametric nuisance estimation can be slow when the dimension of covariates is large so the doubly robust estimator with a smaller bias is recommended for use. Additional simulation results and a real data example are provided in the supplementary materials.

7. Discussion

In this work, we study the estimation of continuous treatment effects when there is limited access to the primary outcome of interest but auxiliary information on surrogate outcomes is available. We propose a doubly robust estimator that is less sensitive to nuisance estimation error and hence incorporates flexible nonparametric machine learning methods. Although nonparametric machine learning methods usually suffer from slow convergence rates, they are widely used in nuisance function estimation, especially when practitioners do not have sufficient domain knowledge to justify parametric models. Our doubly robust estimator facilitates the application of nonparametric methods and enjoys optimal estimation rates under mild conditions. Asymptotic normality is further established, which enables researchers to construct confidence intervals and perform statistical inference. We also show how incorporating information on surrogate outcomes improves the variance, compared with methods solely based on labeled data, as in Kennedy et al. (2017). In summary, our methodology provides a robust and efficient approach to leverage surrogate outcomes in continuous treatment effect estimation. However, our method could not deal with the case where only covariate is available in the unlabeled dataset and continuous treatment is also missing (corresponds to the generalizability problem as in Dahabreh et al. (2019)). It is also interesting to extend our

results to studies with multiple outcomes (Kennedy et al., 2019; Du et al., 2024). We leave these problems for future investigation.

Impact Statement

Our research introduces an efficient method for incorporating surrogate outcomes into the estimation of continuous treatment effects, especially useful in contexts where measuring the primary outcome for all subjects is costly or impractical. This approach is particularly helpful in fields like medicine, where it can facilitate the assessment of new treatments' effects on mortality rates, and in business, for evaluating long-term strategies. By appropriately taking advantage of surrogate outcomes and modeling the missing mechanism of the primary outcome, our method enables more feasible and data-driven decision-making, which provides important implications in difference areas and enhance both the effectiveness and efficiency of research and strategic planning.

References

Ai, C., Linton, O., Motegi, K., and Zhang, Z. A unified framework for efficient estimation of general treatment models. *Quantitative Economics*, 12(3):779–816, 2021.

Bickel, P. J., Klaassen, C. A., Bickel, P. J., Ritov, Y., Klaassen, J., Wellner, J. A., and Ritov, Y. *Efficient and adaptive estimation for semiparametric models*, volume 4. Springer, 1993.

Bonvini, M. and Kennedy, E. H. Fast convergence rates for dose-response estimation. *arXiv preprint arXiv:2207.11825*, 2022.

Bonvini, M., Zeng, Z., Yu, M., Kennedy, E. H., and Keele, L. Flexibly estimating and interpreting heterogeneous treatment effects of laparoscopic surgery for cholecystitis patients. *arXiv preprint arXiv:2311.04359*, 2023.

Calonico, S., Cattaneo, M. D., and Farrell, M. H. On the effect of bias estimation on coverage accuracy in nonparametric inference. *Journal of the American Statistical Association*, 113(522):767–779, 2018.

Chakrabortty, A., Dai, G., and Tchetgen, E. T. A general framework for treatment effect estimation in semi-supervised and high dimensional settings. *arXiv preprint arXiv:2201.00468*, 2022.

Chernozhukov, V., Chetverikov, D., Demirer, M., Duflo, E., Hansen, C., Newey, W., and Robins, J. Double/debiased machine learning for treatment and structural parameters. *The Econometrics Journal*, 21(1):C1–C68, 2018.

Colangelo, K. and Lee, Y.-Y. Double debiased machine learning nonparametric inference with continuous treatments. *arXiv preprint arXiv:2004.03036*, 2020.

Dahabreh, I. J., Robertson, S. E., Tchetgen, E. J., Stuart, E. A., and Hernán, M. A. Generalizing causal inferences from individuals in randomized trials to all trial-eligible individuals. *Biometrics*, 75(2):685–694, 2019.

Díaz, I. and van der Laan, M. J. Targeted data adaptive estimation of the causal dose–response curve. *Journal of Causal Inference*, 1(2):171–192, 2013.

Du, J.-H., Zeng, Z., Kennedy, E. H., Wasserman, L., and Roeder, K. Causal inference for genomic data with multiple heterogeneous outcomes. *arXiv preprint arXiv:2404.09119*, 2024.

Fan, J. and Gijbels, I. *Local polynomial modelling and its applications: monographs on statistics and applied probability 66*, volume 66. CRC Press, 1996.

Fan, J., Hu, T.-C., and Truong, Y. K. Robust non-parametric function estimation. *Scandinavian journal of statistics*, pp. 433–446, 1994.

Fan, Q., Hsu, Y.-C., Lieli, R. P., and Zhang, Y. Estimation of conditional average treatment effects with high-dimensional data. *Journal of Business & Economic Statistics*, 40(1):313–327, 2022.

Flores, C. A. and Flores-Lagunes, A. Identification and estimation of causal mechanisms and net effects of a treatment under unconfoundedness, 2009. URL http://hdl.handle.net/10419/35701. urn:nbn:de:101:1-20090622213.

Frölich, M. and Huber, M. Direct and indirect treatment effects–causal chains and mediation analysis with instrumental variables. *Journal of the Royal Statistical Society Series B: Statistical Methodology*, 79(5):1645–1666, 2017.

Galvao, A. F. and Wang, L. Uniformly semiparametric efficient estimation of treatment effects with a continuous treatment. *Journal of the American Statistical Association*, 110(512):1528–1542, 2015.

Gill, R. D. and Robins, J. M. Causal inference for complex longitudinal data: the continuous case. *Annals of Statistics*, pp. 1785–1811, 2001.

Härdle, W., Hall, P., and Marron, J. S. How far are automatically chosen regression smoothing parameters from their optimum? *Journal of the American Statistical Association*, 83(401):86–95, 1988.

Hernán, M. A. and Robins, J. M. Causal inference, 2010.

Hogan, J. W., Roy, J., and Korkontzelou, C. Handling drop-out in longitudinal studies. *Statistics in medicine*, 23(9): 1455–1497, 2004.

Hou, J., Mukherjee, R., and Cai, T. Efficient and robust semi-supervised estimation of ate with partially annotated treatment and response. *arXiv preprint arXiv:2110.12336*, 2021.

Huber, M. Identifying causal mechanisms (primarily) based on inverse probability weighting. *Journal of Applied Econometrics*, 29(6):920–943, 2014.

Huber, M. Replication data for "Direct and indirect effects of continuous treatments based on generalized propensity score weighting", 2020. URL https://doi.org/10.7910/DVN/AJ7Q9B.

Huber, M., Hsu, Y.-C., Lee, Y.-Y., and Lettry, L. Direct and indirect effects of continuous treatments based on generalized propensity score weighting. *Journal of Applied Econometrics*, 35(7):814–840, 2020.

Kallus, N. and Mao, X. On the role of surrogates in the efficient estimation of treatment effects with limited outcome data. *arXiv preprint arXiv:2003.12408*, 2020.

Kennedy, E. H. Semiparametric doubly robust targeted double machine learning: a review. *arXiv preprint arXiv:2203.06469*, 2022.

Kennedy, E. H. Towards optimal doubly robust estimation of heterogeneous causal effects. *Electronic Journal of Statistics*, 17(2):3008–3049, 2023.

Kennedy, E. H., Ma, Z., McHugh, M. D., and Small, D. S. Non-parametric methods for doubly robust estimation of continuous treatment effects. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, 79(4):1229–1245, 2016. doi: 10.1111/rssb.12212. URL https://rss.onlinelibrary.wiley.com/doi/abs/10.1111/rssb.12212.

Kennedy, E. H., Ma, Z., McHugh, M. D., and Small, D. S. Non-parametric methods for doubly robust estimation of continuous treatment effects. *Journal of the Royal Statistical Society. Series B (Statistical Methodology)*, 79 (4):1229–1245, 2017.

Kennedy, E. H., Kangovi, S., and Mitra, N. Estimating scaled treatment effects with multiple outcomes. *Statistical methods in medical research*, 28(4):1094–1104, 2019.

Kennedy, E. H., Balakrishnan, S., and G'Sell, M. Sharp instruments for classifying compliers and generalizing causal effects. *The Annals of Statistics*, 48(4):2008–2030, 2020.

Laan, M. J. and Robins, J. M. *Unified methods for censored longitudinal data and causality*. Springer, 2003.

Levis, A. W., Bonvini, M., Zeng, Z., Keele, L., and Kennedy, E. H. Covariate-assisted bounds on causal effects with instrumental variables. *arXiv preprint arXiv:2301.12106*, 2023.

Meza, I. and Singh, R. Nested nonparametric instrumental variable regression: Long term, mediated, and time varying treatment effects. *arXiv e-prints*, pp. arXiv–2112, 2021.

Nadaraya, E. A. On estimating regression. *Theory of Probability & Its Applications*, 9(1):141–142, 1964.

Newey, W. K. Kernel estimation of partial means and a general variance estimator. *Econometric Theory*, 10(2): 1–21, 1994.

Robins, J., Li, L., Tchetgen, E., van der Vaart, A., et al. Higher order influence functions and minimax estimation of nonlinear functionals. In *Probability and statistics: essays in honor of David A. Freedman*, volume 2, pp. 335–422. Institute of Mathematical Statistics, 2008.

Robins, J. M., Hernan, M. A., and Brumback, B. Marginal structural models and causal inference in epidemiology. *Epidemiology*, pp. 550–560, 2000.

Rubin, D. and van der Laan, M. J. A general imputation methodology for nonparametric regression with censored data. *UC Berkeley Division of Biostatistics Working Paper Series*, 2005.

Rubin, D. B. Estimating causal effects of treatments in randomized and nonrandomized studies. *Journal of Educational Psychology*, 66(5):688–701, 1974.

Schochet, P. Z., Burghardt, J., and Glazerman, S. National job corps study: The impacts of job corps on participants' employment and related outcomes [and] methodological appendixes on the impact analysis. 2001.

Schochet, P. Z., Burghardt, J., and McConnell, S. Does job corps work? impact findings from the national job corps study. *American economic review*, 98(5):1864–1886, 2008.

Shankar, S., Sinha, R., Mitra, S., Sinha, M., and Fiterau, M. Direct inference of effect of treatment (diet) for a cookieless world. In *International Conference on Artificial Intelligence and Statistics*, pp. 1869–1887. PMLR, 2023.

Singh, R. Generalized kernel ridge regression for long term causal inference: Treatment effects, dose responses, and counterfactual distributions. *arXiv preprint arXiv:2201.05139*, 2022.

Splawa-Neyman, J., Dabrowska, D. M., and Speed, T. P. On the application of probability theory to agricultural experiments. essay on principles. section 9. *Statistical Science*, pp. 465–472, 1990.

Takatsu, K. and Westling, T. Debiased inference for a covariate-adjusted regression function. *arXiv preprint arXiv:2210.06448*, 2022.

Tsiatis, A. A. *Semiparametric theory and missing data*, volume 4. Springer, 2006.

Tsybakov, A. B. *Introduction to Nonparametric Estimation*. Springer, 1st edition, 2009.

Van der Laan, M. J., Polley, E. C., and Hubbard, A. E. Super learner. *Statistical applications in genetics and molecular biology*, 6(1), 2007.

Van der Vaart, A. W. *Asymptotic statistics*, volume 3. Cambridge university press, 2000.

Wasserman, L. *All of nonparametric statistics*. Springer Science & Business Media, 2006.

Watson, G. S. Smooth regression analysis. *Sankhyā: The Indian Journal of Statistics, Series A*, pp. 359–372, 1964.

Zeng, Z., Kennedy, E. H., Bodnar, L. M., and Naimi, A. I. Efficient generalization and transportation. *arXiv preprint arXiv:2302.00092*, 2023.

Zhang, Y. and Bradic, J. High-dimensional semi-supervised learning: in search of optimal inference of the mean. *Biometrika*, 109(2):387–403, 2022.

Zhang, Y., Chakrabortty, A., and Bradic, J. Semi-supervised causal inference: Generalizable and double robust inference for average treatment effects under selection bias with decaying overlap. *arXiv preprint arXiv:2305.12789*, 2023.

A. Background on Efficiency Theory

As discussed in Section 3, the plug-in estimator suffers from first-order error and is sensitive to the estimation accuracy of nuisance functions. To address this problem, one can derive the efficient influence function of the target functional ($\mathbb{E}[\theta(A)]$ in our problem), based on which a one-step estimator can be obtained to reduce the bias. The efficient influence function is critical in non-parametric efficiency theory (Bickel et al., 1993; Tsiatis, 2006; Van der Vaart, 2000; Laan & Robins, 2003; Kennedy, 2022). Mathematically, the influence function is the derivative of the target statistical functional in a Von Mises expansion (i.e., distributional Taylor expansion). In the discrete case, it coincides with the Gateaux derivative of the functional when the contamination distribution is a point mass. Influence functions are important in different respects. First, the variance of the influence function is equal to the efficiency bound of the target statistical functional, which characterizes the inherent estimation difficulty of the target functional and provides a benchmark to compare against when we construct estimators. Moreover, it allows us to correct for first-order bias in the plug-in estimator and obtain doubly robust-style estimators, which enjoy appealing statistical properties even if non-parametric methods with relatively slow rates are used in nuisance estimation.

Suppose the statistical functional of interest $\psi = \psi(\mathbb{P})$ admits the first-order Von Mises expansion. Mathematically, we have

$$\psi(\hat{\mathbb{P}}) - \psi(\mathbb{P}) = -\int \phi_1(\mathbf{z}, \hat{\mathbb{P}}) d\mathbb{P}(\mathbf{z}) + R_2(\hat{\mathbb{P}}, \mathbb{P}), \tag{8}$$

where $\phi_1(\mathbf{z}, \mathbb{P})$ is the influence function of $\psi(\mathbb{P})$, $\psi(\hat{\mathbb{P}})$ is the plug-in estimator and $R_2(\hat{\mathbb{P}}, \mathbb{P})$ is the second-order reminder. Under regularity conditions, Von Mises expansion implies the pathwise differentiability

$$\frac{\partial}{\partial \epsilon} \psi\left(\mathbb{P}_\epsilon\right)\Big|_{\epsilon=0} = \int \phi_1(\mathbf{z}; \mathbb{P}) s_\epsilon(\mathbf{z}) d\mathbb{P}(z)$$

where $s_\epsilon(\mathbf{z}) = \frac{\partial}{\partial \epsilon} \log p_\epsilon(\mathbf{z})\big|_{\epsilon=0}$ is the submodel score function. (8) suggests that the plug-in estimator has first-order bias $-\int \phi_1(\mathbf{z}, \hat{\mathbb{P}}) d\mathbb{P}(\mathbf{z})$. Equivalently, we can write

$$\psi(\mathbb{P}) = \psi(\hat{\mathbb{P}}) + \int \phi_1(\mathbf{z}, \hat{\mathbb{P}}) d\mathbb{P}(\mathbf{z}) + R_2(\hat{\mathbb{P}}, \mathbb{P}),$$

which motivates us to correct for the first-order bias and arrive at the doubly robust estimator

$$\hat{\psi}^{dr} = \psi(\hat{\mathbb{P}}) + \mathbb{P}_n\{\phi_1(\mathbf{Z}, \hat{\mathbb{P}})\}.$$

Under further empirical process assumptions or sample splitting assumptions, we can show the dominating term in the conditional bias of $\hat{\psi}^{dr}$ is $R_2(\hat{\mathbb{P}}, \mathbb{P})$, which is usually a second-order error term and depends on the product of convergence rates of nuisance functions. We refer the readers to Kennedy et al. (2016); Kennedy (2022) for a more complete review. In our problem, the estimand $\theta(a)$ in (2) is not pathwise differentiable and the ideas above do not apply directly. However, the pseudo-outcome $\varphi(\mathbf{Z})$ is obtained by deriving the influence function of $\mathbb{E}[\theta(A)]$. The readers are referred to Section 3.1 of Kennedy et al. (2017) for more discussion on the connection between the pseudo-outcome of $\theta(a)$ and the influence function of $\mathbb{E}[\theta(A)]$.

B. Proofs

B.1. Proof of Theorem 3.5

Proof.

$$\begin{aligned}
&\mathbb{E}[Y^a] \\
=&\mathbb{E}[\mathbb{E}(Y^a|\mathbf{V}, \mathbf{S}^a] \\
=&\mathbb{E}[\mathbb{E}(Y^a|A=a, \mathbf{V}, \mathbf{S}^a)] \\
=&\mathbb{E}[\mathbb{E}(Y^a|A=a, \mathbf{V}, \mathbf{S}^a, R=1)] \\
=&\mathbb{E}\{\mathbb{E}[\mathbb{E}(Y^a|A=a, \mathbf{V}, \mathbf{S}^a, R=1)|\mathbf{V}]\} \\
=&\mathbb{E}\{\mathbb{E}[\mathbb{E}(Y^a|A=a, \mathbf{V}, \mathbf{S}^a, R=1)|A=a, \mathbf{V}]\} \\
=&\mathbb{E}\{\mathbb{E}[\mathbb{E}(Y|A=a, \mathbf{V}, \mathbf{S}, R=1)|A=a, \mathbf{V}]\},
\end{aligned}$$

where the first and fourth equations follow from the property of conditional expectation. The second equation holds since Assumption 3.2 implies $Y^a \perp\!\!\!\perp A|\mathbf{V}, \mathbf{S}^a$. The third equation follows from Assumption 3.3. The fifth equation follows from Assumption 3.2 (specifically $\mathbf{S}^a \perp\!\!\!\perp A|\mathbf{V}$). The last equation follows from Assumption 3.1. Note that positivity is implicitly assumed to guarantee the conditional expectations are well-defined. □

B.2. Proof of Proposition 4.1

Proof. If $\bar{\mu} = \mu$ we have

$$\varphi(\mathbf{Z}; \mu, \bar{\pi}, \bar{\rho}) = \left[\frac{R(Y - \mu(A, \mathbf{X}, 1))}{\bar{\rho}(A, \mathbf{X})} + \mu(A, \mathbf{X}, 1) - \mathbb{E}[\mu(A, \mathbf{X}, 1)|A, \mathbf{V}]\right] \frac{\int_{\mathcal{V}} \bar{\pi}(A|\mathbf{v}) d\mathbb{P}(\mathbf{v})}{\bar{\pi}(A|\mathbf{V})} + \theta(A).$$

Note that

$$\mathbb{E}\left[\frac{R(Y - \mu(A, \mathbf{X}, 1))}{\bar{\rho}(A, \mathbf{X})}\middle| A = a, \mathbf{X}, R\right] = \frac{R}{\bar{\rho}(a, \mathbf{X})}\mathbb{E}[Y - \mu(a, \mathbf{X}, 1)|A = a, \mathbf{X}, R = 1] = 0,$$
$$\mathbb{E}[\mu(A, \mathbf{X}, 1) - \mathbb{E}[\mu(A, \mathbf{X}, 1)|A, \mathbf{V}]|A = a, \mathbf{V}] = 0.$$

These two equations imply

$$\mathbb{E}\left\{\left[\frac{R(Y - \mu(A, \mathbf{X}, 1))}{\bar{\rho}(A, \mathbf{X})} + \mu(A, \mathbf{X}, 1) - \mathbb{E}[\mu(A, \mathbf{X}, 1)|A, \mathbf{V}]\right] \frac{\int_{\mathcal{V}} \bar{\pi}(A|\mathbf{v}) d\mathbb{P}(\mathbf{v})}{\bar{\pi}(A|\mathbf{V})}\middle| A = a\right\} = 0$$

and hence

$$\mathbb{E}[\varphi(\mathbf{Z}; \mu, \bar{\pi}, \bar{\rho})|A = a] = \theta(a).$$

If $(\bar{\pi}, \bar{\rho}) = (\pi, \rho)$,

$$\varphi(\mathbf{Z}; \bar{\mu}, \pi, \rho) = \left[\frac{R(Y - \bar{\mu}(A, \mathbf{X}, 1))}{\rho(A, \mathbf{X})} + \bar{\mu}(A, \mathbf{X}, 1) - \mathbb{E}[\bar{\mu}(A, \mathbf{X}, 1)|A, \mathbf{V}]\right] \frac{\int_{\mathcal{V}} \pi(A|\mathbf{v}) d\mathbb{P}(\mathbf{v})}{\pi(A|\mathbf{V})} + \int_{\mathcal{V}} \mathbb{E}[\bar{\mu}(A, \mathbf{X}, 1)|A, \mathbf{V} = \mathbf{v}] d\mathbb{P}(\mathbf{v}).$$

Note that

$$\mathbb{E}\left[\frac{R(Y - \bar{\mu}(A, \mathbf{X}, 1))}{\rho(A, \mathbf{X})}\middle| A = a, \mathbf{X}, R\right] = \frac{R(\mu(a, \mathbf{X}, 1) - \bar{\mu}(a, \mathbf{X}, 1))}{\rho(a, \mathbf{X})},$$
$$\mathbb{E}\left[\frac{R(Y - \bar{\mu}(A, \mathbf{X}, 1))}{\rho(A, \mathbf{X})}\middle| A = a, \mathbf{X}\right] = \mathbb{E}\left[\frac{R(\mu(a, \mathbf{X}, 1) - \bar{\mu}(a, \mathbf{X}, 1))}{\rho(a, \mathbf{X})}\middle| A = a, \mathbf{X}\right] = \mu(a, \mathbf{X}, 1) - \bar{\mu}(a, \mathbf{X}, 1).$$

By the property of conditional expectation, we have

$$\begin{aligned}
&\mathbb{E}\left\{\left[\frac{R(Y - \bar{\mu}(A, \mathbf{X}, 1))}{\rho(A, \mathbf{X})} + \bar{\mu}(A, \mathbf{X}, 1) - \mathbb{E}[\bar{\mu}(A, \mathbf{X}, 1)|A, \mathbf{V}]\right] \frac{\int_{\mathcal{V}} \pi(A|\mathbf{v}) d\mathbb{P}(\mathbf{v})}{\pi(A|\mathbf{V})}\middle| A = a, \mathbf{V}\right\} \\
=&\mathbb{E}\left\{\mu(a, \mathbf{X}, 1) - \mathbb{E}[\bar{\mu}(a, \mathbf{X}, 1)|A = a, \mathbf{V}]\middle| A = a, \mathbf{V}\right\} \frac{\int_{\mathcal{V}} \pi(a|\mathbf{v}) d\mathbb{P}(\mathbf{v})}{\pi(a|\mathbf{V})} \\
=&\mathbb{E}[\mu(a, \mathbf{X}, 1) - \bar{\mu}(a, \mathbf{X}, 1)\big| A = a, \mathbf{V}] \frac{\int_{\mathcal{V}} \pi(a|\mathbf{v}) d\mathbb{P}(\mathbf{v})}{\pi(a|\mathbf{V})}.
\end{aligned}$$

Due to the following equation on the measure:

$$d\mathbb{P}(\mathbf{v}|a) = \frac{\pi(a|\mathbf{v}) d\mathbb{P}(\mathbf{v})}{\int_{\mathcal{V}} \pi(a|\mathbf{v}) d\mathbb{P}(\mathbf{v})}, \tag{9}$$

we can write

$$\begin{aligned}
&\mathbb{E}\left\{\left[\frac{R(Y - \bar{\mu}(A, \mathbf{X}, 1))}{\rho(A, \mathbf{X})} + \bar{\mu}(A, \mathbf{X}, 1) - \mathbb{E}[\bar{\mu}(A, \mathbf{X}, 1)|A, \mathbf{V}]\right] \frac{\int_{\mathcal{V}} \pi(A|\mathbf{v}) d\mathbb{P}(\mathbf{v})}{\pi(A|\mathbf{V})}\middle| A = a\right\} \\
=& \int_{\mathcal{V}} \mathbb{E}[\mu(a, \mathbf{X}, 1) - \bar{\mu}(a, \mathbf{X}, 1)\big| A = a, \mathbf{V} = \mathbf{v}] \frac{\int_{\mathcal{V}} \pi(a|\mathbf{v}) d\mathbb{P}(\mathbf{v})}{\pi(a|\mathbf{v})} d\mathbb{P}(\mathbf{v}|a) \\
=& \int_{\mathcal{V}} \mathbb{E}[\mu(a, \mathbf{X}, 1) - \bar{\mu}(a, \mathbf{X}, 1)\big| A = a, \mathbf{V} = \mathbf{v}] d\mathbb{P}(\mathbf{v}).
\end{aligned}$$

Finally we get

$$\begin{aligned}
&\mathbb{E}[\varphi(\mathbf{Z};\bar{\mu},\pi,\rho)|A=a]\\
=&\int_{\mathcal{V}}\mathbb{E}[\mu(a,\mathbf{X},1)-\bar{\mu}(a,\mathbf{X},1)\big|A=a,\mathbf{V}=\mathbf{v}]d\mathbb{P}(\mathbf{v})+\int_{\mathcal{V}}\mathbb{E}[\bar{\mu}(a,\mathbf{X},1)|A=a,\mathbf{V}=\mathbf{v}]d\mathbb{P}(\mathbf{v})\\
=&\int_{\mathcal{V}}\mathbb{E}[\mu(a,\mathbf{X},1)|A=a,\mathbf{V}=\mathbf{v}]d\mathbb{P}(\mathbf{v})\\
=&\theta(a).
\end{aligned}$$

□

B.3. Proof of Theorem 5.1

Proof. By Theorem 1 in Kennedy (2023), the linear smoother is stable if the variance $\mathrm{Var}(\varphi(\mathbf{Z})|A=a)$ is bounded away from 0, i.e.

$$\widehat{\theta}(a)-\widetilde{\theta}(a)-\widehat{\mathbb{E}}_n[\widehat{b}(A)|A=a]=o_{\mathbb{P}}(R_n(a))$$

if $\sup_{\mathbf{z}}|\widehat{\varphi}(\mathbf{z})-\varphi(\mathbf{z})|=o_{\mathbb{P}}(1)$. □

B.4. Proof of Proposition 5.2

Proof. For abbreviations, we will omit conditioning on D^n in our notation but all the expectations in this part are conditioning on D^n (recall such expectation is denoted using $\mathbb{P}$). Note that

$$\mathbb{E}[\varphi(\mathbf{Z})|A=a]=\theta(a)$$

by Proposition 4.1. By the property of conditional expectation

$$\begin{aligned}
&\mathbb{P}\left\{\left[\frac{R(Y-\widehat{\mu}(a,\mathbf{X},1))}{\widehat{\rho}(a,\mathbf{X})}+\widehat{\mu}(a,\mathbf{X},1)-\widehat{\tau}(a,\mathbf{V})\right]\widehat{w}(a,\mathbf{V})\Big|A=a\right\}\\
=&\mathbb{P}\left\{\left[\frac{R(\mu(a,\mathbf{X},1)-\widehat{\mu}(a,\mathbf{X},1))}{\widehat{\rho}(a,\mathbf{X})}+\widehat{\mu}(a,\mathbf{X},1)-\widehat{\tau}(a,\mathbf{V})\right]\widehat{w}(a,\mathbf{V})\Big|A=a\right\}\\
=&\mathbb{P}\left\{\left[\frac{\rho(a,\mathbf{X})(\mu(a,\mathbf{X},1)-\widehat{\mu}(a,\mathbf{X},1))}{\widehat{\rho}(a,\mathbf{X})}+\widehat{\mu}(a,\mathbf{X},1)-\widehat{\tau}(a,\mathbf{V})\right]\widehat{w}(a,\mathbf{V})\Big|A=a\right\}\\
=&\mathbb{P}\left\{\left[\left(1-\frac{\rho(a,\mathbf{X})}{\widehat{\rho}(a,\mathbf{X})}\right)(\widehat{\mu}(a,\mathbf{X},1)-\mu(a,\mathbf{X},1))+\mu(a,\mathbf{X},1)-\widehat{\tau}(a,\mathbf{V})\right]\widehat{w}(a,\mathbf{V})\Big|A=a\right\}\\
=&\mathbb{P}\left\{\left[\left(1-\frac{\rho(a,\mathbf{X})}{\widehat{\rho}(a,\mathbf{X})}\right)(\widehat{\mu}(a,\mathbf{X},1)-\mu(a,\mathbf{X},1))\right]\widehat{w}(a,\mathbf{V})\Big|A=a\right\}+\mathbb{P}[(\tau(a,\mathbf{V})-\widehat{\tau}(a,\mathbf{V}))\widehat{w}(a,\mathbf{V})|A=a]
\end{aligned}\tag{10}$$

where the first equation follows by conditioning on $\mathbf{X},R,A=a$, the second equation follows from conditioning on $\mathbf{X},A=a$ and the last equation follows from conditioning on $\mathbf{V},A=a$. Further note that $\theta(a)=\mathbb{E}[\tau(a,\mathbf{V})]$ and

$$\mathbb{P}\left[\widehat{\theta}_0(a)\right]-\theta(a)=\frac{1}{n}\sum_{i\in D_2^n}\widehat{\tau}(a,\mathbf{V}_i)-\mathbb{P}[\widehat{\tau}(a,\mathbf{V})]+\mathbb{P}[\widehat{\tau}(a,\mathbf{V})-\tau(a,\mathbf{V})].\tag{11}$$

By equation 9

$$\mathbb{P}[\widehat{\tau}(a,\mathbf{V})-\tau(a,\mathbf{V})]=\int_{\mathcal{V}}(\widehat{\tau}(a,\mathbf{v})-\tau(a,\mathbf{v}))d\mathbb{P}(\mathbf{v})=\int_{\mathcal{V}}(\widehat{\tau}(a,\mathbf{v})-\tau(a,\mathbf{v}))w(a,\mathbf{v})d\mathbb{P}(\mathbf{v}|A=a).$$

Add equation 10 and equation 11 together we have

$$\begin{aligned}
\widehat{b}(a)=&\mathbb{P}\left\{\left[\left(1-\frac{\rho(a,\mathbf{X})}{\widehat{\rho}(a,\mathbf{X})}\right)(\widehat{\mu}(a,\mathbf{X},1)-\mu(a,\mathbf{X},1))\right]\widehat{w}(a,\mathbf{V})\Big|A=a\right\}\\
&+\mathbb{P}[(\tau(a,\mathbf{V})-\widehat{\tau}(a,\mathbf{V}))(\widehat{w}(a,\mathbf{V})-w(a,\mathbf{V}))|A=a]+\frac{1}{n}\sum_{i\in D_2^n}\widehat{\tau}(a,\mathbf{V}_i)-\mathbb{P}[\widehat{\tau}(a,\mathbf{V})].
\end{aligned}$$

The bound on $\widehat{b}$ then follows from Cauchy-Schwarz's inequality.

For linear smoother $\widehat{\mathbb{E}}_n$, we have

$$\left|\widehat{\mathbb{E}}_n\left[\widehat{b}(A)|A=a\right]\right| = \left|\sum_{i\in T^n} W_i(a;\mathbf{A}^n)\widehat{b}(A_i)\right| \leq \sup_{t\in N(a)} |\widehat{b}(t)| \sum_{i\in T^n} |W_i(a;\mathbf{A}^n)| \leq C \sup_{t\in N(a)} |\widehat{b}(t)|$$

□

B.5. Proof of Theorem 5.3

Proof. We will prove the results with the following decomposition. Let $\widetilde{\theta}(a) = \mathbf{e}_1^\top \widehat{\mathbf{D}}_{ha}^{-1}\mathbb{P}_n[\mathbf{g}_{ha}(A)K_{ha}(A)\varphi(\mathbf{Z})]$ be the oracle estimator. We write

$$\begin{aligned}\widehat{\theta}(a) - \theta(a) =& \widetilde{\theta}(a) - \theta(a) + \mathbf{e}_1^\top \widehat{\mathbf{D}}_{ha}^{-1}\mathbb{P}_n[\mathbf{g}_{ha}(A)K_{ha}(A)(\widehat{\varphi}(\mathbf{Z}) - \varphi(\mathbf{Z}))] \\ =:& \widetilde{\theta}(a) - \theta(a) + R_1 + R_2\end{aligned}$$

where $R_1 = \mathbf{e}_1^\top \widehat{\mathbf{D}}_{ha}^{-1}(\mathbb{P}_n - \mathbb{P})[\mathbf{g}_{ha}(A)K_{ha}(A)(\widehat{\varphi}(\mathbf{Z}) - \varphi(\mathbf{Z}))]$ and $R_2 = \mathbf{e}_1^\top \widehat{\mathbf{D}}_{ha}^{-1}\mathbb{P}[\mathbf{g}_{ha}(A)K_{ha}(A)(\widehat{\varphi}(\mathbf{Z}) - \varphi(\mathbf{Z}))]$.

Step 1: The CLT term. The existing results on local linear estimator (Fan et al., 1994; Fan & Gijbels, 1996) imply

$$\sqrt{nh}\left(\widetilde{\theta}(a) - \theta(a) - \frac{h^2\theta''(a)\int u^2K(u)du}{2}\right) \xrightarrow{d} N\left(0, \frac{\sigma^2(a)\int K^2(u)du}{f(a)}\right)$$

under the conditions stated in the theorem. The conditional variance can be computed as follows:

$$\begin{aligned}\sigma^2(a) =& \operatorname{Var}(\varphi(\mathbf{Z})|A=a) \\ =& \mathbb{E}[(\varphi(\mathbf{Z}) - \theta(a))^2|A=a] \\ =& \mathbb{E}\left\{\left[\frac{R(Y-\mu(a,\mathbf{X},1))}{\rho(a,\mathbf{X})} + \mu(a,\mathbf{X},1) - \tau(a,\mathbf{V})\right]^2 w^2(a,\mathbf{V})\middle|A=a\right\} \\ =& \mathbb{E}\left\{\left[\frac{R(Y-\mu(a,\mathbf{X},1))^2}{\rho^2(a,\mathbf{X})} + (\mu(a,\mathbf{X},1) - \tau(a,\mathbf{V}))^2\right] w^2(a,\mathbf{V})\middle|A=a\right\},\end{aligned}$$

where the last equation follows since by conditioning on $A=a, \mathbf{X}, R$ one can show

$$\mathbb{E}\left\{\left[\frac{R(Y-\mu(a,\mathbf{X},1))(\mu(a,\mathbf{X},1)-\tau(a,\mathbf{V}))}{\rho(a,\mathbf{X})}\right] w^2(a,\mathbf{V})\middle|A=a\right\} = 0.$$

For the first term in $\sigma^2(a)$

$$\begin{aligned}&\mathbb{E}\left\{\frac{R(Y-\mu(a,\mathbf{X},1))^2 w^2(a,\mathbf{V})}{\rho^2(a,\mathbf{X})}\middle|A=a\right\} \\ =& \mathbb{E}\left\{\frac{R\operatorname{Var}(Y|A=a,\mathbf{X},R=1)w^2(a,\mathbf{V})}{\rho^2(a,\mathbf{X})}\middle|A=a\right\} \\ =& \mathbb{E}\left\{\frac{\operatorname{Var}(Y|A=a,\mathbf{X},R=1)w^2(a,\mathbf{V})}{\rho(a,\mathbf{X})}\middle|A=a\right\},\end{aligned}$$

where the first equation follows from conditioning on $A=a, \mathbf{X}, R$ and the second equation follows from conditioning on $A=a, \mathbf{X}$. For the second term in $\sigma^2(a)$ we have

$$\mathbb{E}\left[(\mu(a,\mathbf{X},1) - \tau(a,\mathbf{V}))^2 w^2(a,\mathbf{V})|A=a\right] = \mathbb{E}\left[\operatorname{Var}(\mu(a,\mathbf{X},1)|A=a,\mathbf{V})w^2(a,\mathbf{V})|A=a\right].$$

Hence

$$\sigma^2(a) = \mathbb{E}\left\{\left[\frac{\operatorname{Var}(Y|A=a,\mathbf{X},R=1)}{\rho(a,\mathbf{X})} + \operatorname{Var}(\mu(a,\mathbf{X},1)|A=a,\mathbf{V})\right] w^2(a,\mathbf{V})\middle|A=a\right\}$$

Step 2: Bounding R_1. Then we proceed to analyze $R_1 = \mathbf{e}_1^\top \widehat{\mathbf{D}}_{ha}^{-1}(\mathbb{P}_n - \mathbb{P})[\mathbf{g}_{ha}(A)K_{ha}(A)(\widehat{\varphi}(\mathbf{Z}) - \varphi(\mathbf{Z}))]$. We first show $\mathbf{e}_1^\top \widehat{\mathbf{D}}_{ha}^{-1} = O_{\mathbb{P}}(1)$. Recall $\widehat{\mathbf{D}}_{ha} = \mathbb{P}_n[\mathbf{g}_{ha}(A)K_{ha}(A)\mathbf{g}_{ha}^\top(A)] \in \mathbb{R}^{2\times 2}$. Note that $\mathbb{P}_n[K_{ha}(A)]$ is the kernel density estimator of $f(a)$ and standard results in the literature show

$$\mathbb{E}[(\mathbb{P}_n[K_{ha}(A)] - f(a))^2] = O\left(h^2 + \frac{1}{nh}\right) = o(1),$$

which implies $\widehat{D}_{ha,11} = \mathbb{P}_n[K_{ha}(A)] \xrightarrow{P} f(a)$. For the element of $\widehat{\mathbf{D}}_{ha}$ not on the diagonal

$$\mathbb{E}\left[\left(\frac{A-a}{h}\right)K_{ha}(A)\right] = \int \frac{t-a}{h}\frac{1}{h}K\left(\frac{t-a}{h}\right)f(t)dt = \int uK(u)f(a+hu)du.$$

So we have

$$\left|\mathbb{E}\left[\left(\frac{A-a}{h}\right)K_{ha}(A)\right] - \int uK(u)f(a)du\right| \leq \int |u|K(u)|f(a+hu) - f(a)| \lesssim h\int u^2K(u)du \to 0.$$

Since K is symmetric around 0 we have $\int uK(u)du = 0$ and hence

$$\mathbb{E}\left[\left(\frac{A-a}{h}\right)K_{ha}(A)\right] = O(h).$$

Further notice

$$\begin{aligned}
&\operatorname{Var}\left(\mathbb{P}_n\left[\left(\frac{A-a}{h}\right)K_{ha}(A)\right]\right)\\
&= \frac{1}{n}\operatorname{Var}\left(\left(\frac{A-a}{h}\right)K_{ha}(A)\right)\\
&\leq \frac{1}{n}\mathbb{E}\left[\left(\frac{A-a}{h}\right)^2 K_{ha}^2(A)\right]\\
&= \frac{1}{n}\int\left(\frac{t-a}{h}\right)^2\frac{1}{h^2}K^2\left(\frac{t-a}{h}\right)f(t)dt\\
&= \frac{1}{nh}\int u^2K^2(u)f(a+hu)du\\
&\leq \frac{\|f\|_\infty}{nh}\int u^2K^2(u)du = O\left(\frac{1}{nh}\right).
\end{aligned}$$

Hence we have

$$\mathbb{E}\left\{\left\{\mathbb{P}_n\left[\left(\frac{A-a}{h}\right)K_{ha}(A)\right]\right\}^2\right\} = O\left(h^2 + \frac{1}{nh}\right) = o(1),$$

which implies $\widehat{D}_{ha,12} = \mathbb{P}_n\left[\left(\frac{A-a}{h}\right)K_{ha}(A)\right] \xrightarrow{P} 0$. Finally

$$\mathbb{E}\left[\left(\frac{A-a}{h}\right)^2 K_{ha}(A)\right] = \int\left(\frac{t-a}{h}\right)^2\frac{1}{h}K\left(\frac{t-a}{h}\right)f(t)dt = \int u^2K(u)f(a+hu)du.$$

So we have

$$\left|\mathbb{E}\left[\left(\frac{A-a}{h}\right)^2 K_{ha}(A)\right] - \int u^2K(u)f(a)du\right| \leq \int u^2K(u)|f(a+hu) - f(a)| \lesssim h\int |u|^3K(u)du \to 0.$$

One can similarly show

$$\operatorname{Var}\left(\left(\frac{A-a}{h}\right)^2 K_{ha}(A)\right) = O(1/h)$$

and hence

$$\mathbb{E}\left\{\left\{\mathbb{P}_n\left[\left(\frac{A-a}{h}\right)^2 K_{ha}(A)\right]-\int u^2K(u)f(a)du\right\}^2\right\}=O\left(h^2+\frac{1}{nh}\right)=o(1),$$

which implies $\widehat{D}_{ha,22}=\mathbb{P}_n\left[\left(\frac{A-a}{h}\right)^2 K_{ha}(A)\right]\xrightarrow{P}f(a)\int u^2K(u)du$. Hence we have

$$\widehat{\mathbf{D}}_{ha}^{-1}\xrightarrow{P}\text{diag}\left\{f(a)^{-1},f(a)^{-1}\left(\int u^2K(u)du\right)^{-1}\right\}.$$

This implies $\mathbf{e}_1^\top\widehat{\mathbf{D}}_{ha}^{-1}=O_{\mathbb{P}}(1)$. Then we consider $(\mathbb{P}_n-\mathbb{P})[\mathbf{g}_{ha}(A)K_{ha}(A)(\widehat{\varphi}(\mathbf{Z})-\varphi(\mathbf{Z}))]$. By Lemma 2 in Kennedy et al. (2020) we have for $j=1,2$

$$(\mathbb{P}_n-\mathbb{P})[g_{ha,j}(A)K_{ha}(A)(\widehat{\varphi}(\mathbf{Z})-\varphi(\mathbf{Z}))]=O_{\mathbb{P}}\left(\frac{\|g_{ha,j}(A)K_{ha}(A)(\widehat{\varphi}(\mathbf{Z})-\varphi(\mathbf{Z}))\|_2}{\sqrt{n}}\right)$$

Note that

$$\begin{aligned}
&\|g_{ha,j}(A)K_{ha}(A)(\widehat{\varphi}(\mathbf{Z})-\varphi(\mathbf{Z}))\|_2^2\\
=&\,\mathbb{P}\left[g_{ha,j}^2(A)K_{ha}^2(A)(\widehat{\varphi}(\mathbf{Z})-\varphi(\mathbf{Z}))^2\right]\\
\le&\,\|\widehat{\varphi}-\varphi\|_\infty^2\int\left(\frac{t-a}{h}\right)^{2(j-1)}\frac{1}{h^2}K^2\left(\frac{t-a}{h}\right)f(t)dt\\
\le&\,\frac{\|\widehat{\varphi}-\varphi\|_\infty^2}{h}\int u^{2(j-1)}K^2(u)f(a+hu)du\\
\lesssim&\,\frac{\|\widehat{\varphi}-\varphi\|_\infty^2}{h}.
\end{aligned}$$

This together with $\|\widehat{\varphi}-\varphi\|_\infty=o_{\mathbb{P}}(1)$ implies

$$(\mathbb{P}_n-\mathbb{P})[g_{ha,j}(A)K_{ha}(A)(\widehat{\varphi}(\mathbf{Z})-\varphi(\mathbf{Z}))]=O_{\mathbb{P}}\left(\frac{\|\widehat{\varphi}-\varphi\|_\infty}{\sqrt{nh}}\right)=o_{\mathbb{P}}\left(\frac{1}{\sqrt{nh}}\right).$$

We conclude that $R_1=o_{\mathbb{P}}\left(\frac{1}{\sqrt{nh}}\right)$.

Step 3: Bounding R_2. The last step is to bound $R_2=\mathbf{e}_1^\top\widehat{\mathbf{D}}_{ha}^{-1}\mathbb{P}[\mathbf{g}_{ha}(A)K_{ha}(A)(\widehat{\varphi}(\mathbf{Z})-\varphi(\mathbf{Z}))]$. Since $\mathbf{e}_1^\top\widehat{\mathbf{D}}_{ha}^{-1}=O_{\mathbb{P}}(1)$ we only need to consider

$$\mathbb{P}[\mathbf{g}_{ha}(A)K_{ha}(A)(\widehat{\varphi}(\mathbf{Z})-\varphi(\mathbf{Z}))]=\int\mathbf{g}_{ha}(t)K_{ha}(t)\mathbb{P}[\widehat{\varphi}(\mathbf{Z})-\varphi(\mathbf{Z})|A=t]f(t)dt$$

In Proposition 5.2 we show $\widehat{b}(a)$ is equal to

$$\begin{aligned}
\mathbb{P}[\widehat{\varphi}(\mathbf{Z})-\varphi(\mathbf{Z})|A=t]=&\,\mathbb{P}\left\{\left[\left(1-\frac{\rho(t,\mathbf{X})}{\widehat{\rho}(t,\mathbf{X})}\right)(\widehat{\mu}(t,\mathbf{X},1)-\mu(t,\mathbf{X},1))\right]\widehat{w}(t,\mathbf{V})\middle|A=t\right\}\\
&+\mathbb{P}[(\tau(t,\mathbf{V})-\widehat{\tau}(t,\mathbf{V}))(\widehat{w}(t,\mathbf{V})-w(t,\mathbf{V}))|A=t]+\frac{1}{n}\sum_{i\in D_2^n}\widehat{\tau}(t,\mathbf{V}_i)-\mathbb{P}[\widehat{\tau}(t,\mathbf{V})]
\end{aligned}$$

Plug into the equation above we have

$$\begin{aligned}
&\int g_{ha,j}(t)K_{ha}(t)\mathbb{P}[\widehat{\varphi}(\mathbf{Z})-\varphi(\mathbf{Z})|A=t]f(t)dt\\
=&\int g_{ha,j}(t)K_{ha}(t)\mathbb{P}\left\{\left[\left(1-\frac{\rho(t,\mathbf{X})}{\widehat{\rho}(t,\mathbf{X})}\right)(\widehat{\mu}(t,\mathbf{X},1)-\mu(t,\mathbf{X},1))\right]\widehat{w}(t,\mathbf{V})\middle|A=t\right\}f(t)dt\\
&+\int g_{ha,j}(t)K_{ha}(t)\mathbb{P}[(\tau(t,\mathbf{V})-\widehat{\tau}(t,\mathbf{V}))(\widehat{w}(t,\mathbf{V})-w(t,\mathbf{V}))|A=t]f(t)dt\\
&+(\mathbb{P}_n-\mathbb{P})\int g_{ha,j}(t)K_{ha}(t)\widehat{\tau}(t,\mathbf{V})f(t)dt
\end{aligned}$$

Here we slightly abuse the notation and denote $\mathbb{P}_n$ as the average over D_2^n. By boundedness of nuisance estimates and Cauchy-Schwarz's inequality, we have

$$\begin{aligned}
&\left|\int g_{ha,j}(t)K_{ha}(t)\mathbb{P}\left\{\left[\left(1-\frac{\rho(t,\mathbf{X})}{\widehat{\rho}(t,\mathbf{X})}\right)(\widehat{\mu}(t,\mathbf{X},1)-\mu(t,\mathbf{X},1))\right]\widehat{w}(t,\mathbf{V})\middle|A=t\right\}f(t)dt\right| \\
\lesssim &\int |g_{ha,j}(t)|K_{ha}(t)\|\widehat{\rho}-\rho\|_t\|\widehat{\mu}-\mu\|_t f(t)dt \\
\leq &\sup_{|t-a|\leq h}\|\widehat{\rho}-\rho\|_t\|\widehat{\mu}-\mu\|_t\int\left|\frac{t-a}{h}\right|^{j-1}\frac{1}{h}K\left(\frac{t-a}{h}\right)f(t)dt \\
= &\sup_{|t-a|\leq h}\|\widehat{\rho}-\rho\|_t\|\widehat{\mu}-\mu\|_t\int|u|^{j-1}K(u)f(a+hu)du \\
\lesssim &\sup_{|t-a|\leq h}\|\widehat{\rho}-\rho\|_t\|\widehat{\mu}-\mu\|_t.
\end{aligned}$$

Similarly we have

$$\begin{aligned}
&\left|\int g_{ha,j}(t)K_{ha}(t)\mathbb{P}[(\tau(t,\mathbf{V})-\widehat{\tau}(t,\mathbf{V}))(\widehat{w}(t,\mathbf{V})-w(t,\mathbf{V}))|A=t]f(t)dt\right| \\
\leq &\int |g_{ha,j}(t)|K_{ha}(t)\|\widehat{\tau}-\tau\|_t\|\widehat{w}-w\|_t f(t)dt \\
\lesssim &\sup_{|t-a|\leq h}\|\widehat{\tau}-\tau\|_t\|\widehat{w}-w\|_t.
\end{aligned}$$

For the remaining term $(\mathbb{P}_n-\mathbb{P})\int g_{ha,j}(t)K_{ha}(t)\widehat{\tau}(t,\mathbf{V})f(t)dt$, by Lemma 2 in Kennedy et al. (2020) we have

$$(\mathbb{P}_n-\mathbb{P})\int g_{ha,j}(t)K_{ha}(t)(\widehat{\tau}(t,\mathbf{V})-\tau(t,\mathbf{V}))f(t)dt=O_{\mathbb{P}}\left(\frac{\|\int g_{ha,j}(t)K_{ha}(t)(\widehat{\tau}(t,\mathbf{V})-\tau(t,\mathbf{V}))f(t)dt\|_2}{\sqrt{n}}\right)$$

where

$$\begin{aligned}
&\left\|\int g_{ha,j}(t)K_{ha}(t)(\widehat{\tau}(t,\mathbf{V})-\tau(t,\mathbf{V}))f(t)dt\right\|_2^2 \\
=&\int\left[\int g_{ha,j}(t)K_{ha}(t)(\widehat{\tau}(t,\mathbf{V})-\tau(t,\mathbf{V}))f(t)dt\right]^2 d\mathbb{P}(\mathbf{v}) \\
\leq&\|\widehat{\tau}-\tau\|_\infty^2\left(\int|g_{ha,j}(t)|K_{ha}(t)f(t)dt\right)^2 \\
\lesssim&\|\widehat{\tau}-\tau\|_\infty^2.
\end{aligned}$$

This together with $\|\widehat{\tau}-\tau\|_\infty=o_{\mathbb{P}}(1)$ implies

$$(\mathbb{P}_n-\mathbb{P})\int g_{ha,j}(t)K_{ha}(t)(\widehat{\tau}(t,\mathbf{V})-\tau(t,\mathbf{V}))f(t)dt=o_{\mathbb{P}}\left(\frac{1}{\sqrt{n}}\right). \tag{12}$$

By direct calculations (where we assume the outcome is bounded hence τ is also bounded)

$$\begin{aligned}
&\mathbb{E}\left\{\left[(\mathbb{P}_n-\mathbb{P})\int g_{ha,j}(t)K_{ha}(t)\tau(t,\mathbf{V})f(t)dt\right]^2\right\} \\
\leq&\frac{1}{n}\mathbb{E}\left\{\left(\int g_{ha,j}(t)K_{ha}(t)\tau(t,\mathbf{V})f(t)dt\right)^2\right\} \\
\lesssim&\frac{1}{n}\mathbb{E}\left\{\left(\int|g_{ha,j}(t)|K_{ha}(t)f(t)dt\right)^2\right\} \\
\lesssim&\frac{1}{n},
\end{aligned}$$

which implies

$$(\mathbb{P}_n - \mathbb{P}) \int g_{ha,j}(t) K_{ha}(t) \tau(t, \mathbf{V}) f(t) dt = O_{\mathbb{P}}\left(\frac{1}{\sqrt{n}}\right).$$

Combining this equation with (12) yields

$$(\mathbb{P}_n - \mathbb{P}) \int g_{ha,j}(t) K_{ha}(t) \widehat{\tau}(t, \mathbf{V}) f(t) dt = O_{\mathbb{P}}\left(\frac{1}{\sqrt{n}}\right) = o_{\mathbb{P}}\left(\frac{1}{\sqrt{nh}}\right).$$

Hence under the rate conditions in the theorem, we conclude

$$R_2 = O_{\mathbb{P}}\left(\sup_{|t-a|\le h} \|\widehat{\rho} - \rho\|_t \|\widehat{\mu} - \mu\|_t + \sup_{|t-a|\le h} \|\widehat{\tau} - \tau\|_t \|\widehat{w} - w\|_t\right) + o_{\mathbb{P}}\left(\frac{1}{\sqrt{nh}}\right) = o_{\mathbb{P}}\left(\frac{1}{\sqrt{nh}}\right).$$

The asymptotic normality then follows from Slutsky's theorem. □

C. Additional Simulation Results

C.1. Nuisance Functions Estimated by Parametric Methods

We evaluate the performance of plug-in-style estimator and doubly robust estimator when nuisance functions are estimated using parametric models under the same setting as Section 6. We will fit linear regression models for μ, τ, λ and separately consider correctly specifying the outcome model μ, τ or not, where a misspecified model left out the quadratic term in a but keeps all other main effects and interactions. The conditional density of A given $\mathbf{V}$ is obtained by first estimating the conditional mean of $\mathbb{E}[A|\mathbf{V}]$ and plug-in the normal density (i.e., we assume the model for conditional density is always correct). For the plug-in estimator we randomly separate the sample into two parts D, T. The first part D is used to fit the regression models for all nuisance functions and we take the average on the second part as $\widehat{\theta}(a) = \frac{1}{|T|}\sum_{i\in T} \widehat{\tau}(a, \mathbf{V}_i)$. The roles of D, T are then exchanged to obtain another estimate and the final estimator is the average of two estimates. The doubly robust estimator is implemented according to Algorithm 1, where the bandwidth is selected on D_2. We generate samples with sample size $n \in \{10^{2.6}, 10^{2.8}, \ldots, 10^{4.6}\}$, apply two estimators to estimate $\theta(1)$ and repeat the process 500 times. The results are summarized in Figure 3.

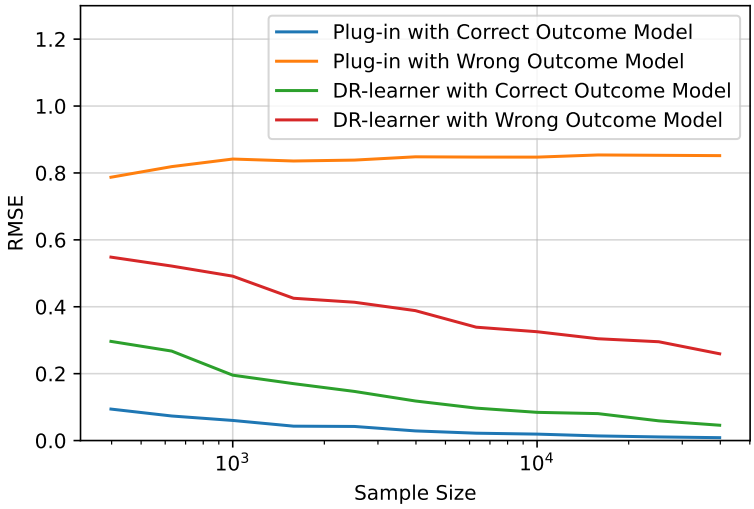


Figure 3. RMSE versus sample size (in log scale) when nuisance functions are estimated by parametric models.

When the outcome model is misspecified, the plug-in estimator (solely based on outcome modeling) is no longer consistent, as shown in Figure 3. The doubly robust estimator, however, models both the outcome and treatment process and has a

smaller estimation error when the outcome model is misspecified. Estimation with correctly specified outcome model corresponds to $\alpha = 0.5$ in Figure 2, where doubly robust estimator with outcome model and propensity score both correctly specified has larger error compared with the plug-in estimator since slow rate of local linear smoothing $O(n^{-2/5})$ dominates the nuisance estimation error $O(n^{-1/2})$.

C.2. Nuisance Functions Estimated by Nonparametric Methods

We further evaluate the performance of the plug-in style estimator and doubly robust estimator when nuisance functions are estimated using nonparametric models under the same setting as Section 6. We will fit nuisance functions μ, τ, ρ by superlearner combining generalized linear models and random forests. The conditional density of A given $\mathbf{V}$ is estimated by kernel density estimation. The plug-in estimator is implemented the same way as in Appendix C.1 and the doubly robust estimator is implemented according to Algorithm 1. We generate samples with sample size $n \in \{10^{2.6}, 10^{2.8}, \ldots, 10^4\}$, apply two estimators to estimate $\theta(1)$ and repeat the process 500 times. The results are summarized in Figure 3.

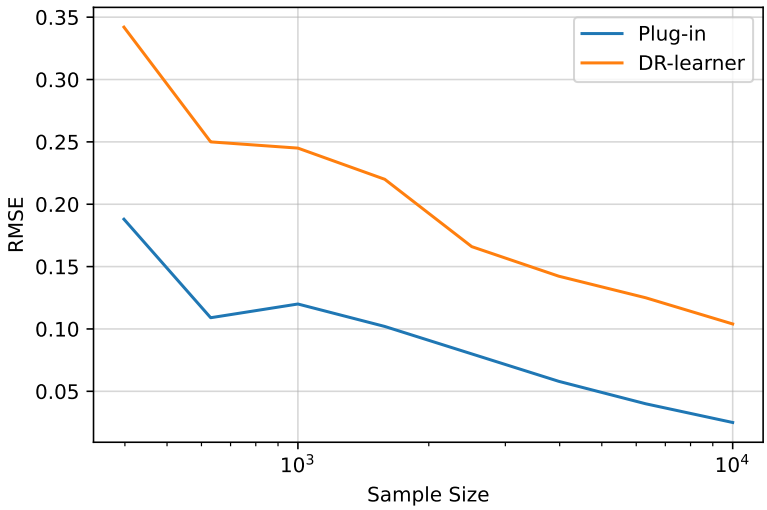


Figure 4. RMSE versus sample size when nuisance functions are estimated by nonparametric models.

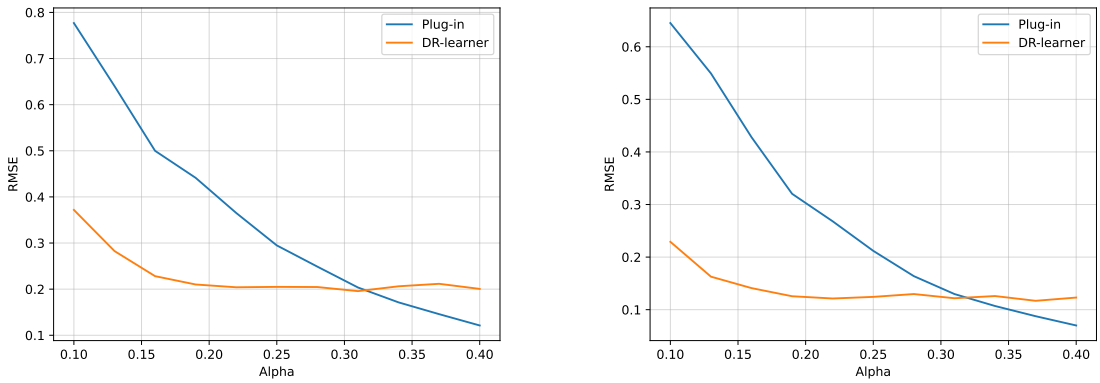
As shown in Figure 4, the plug-in-style estimator has smaller estimation error when all the nuisance functions are fitted by nonparametric methods. This could be explained by the diffculty in nonparametrically estimating conditional density π. Due to the curse of dimensionality, a large sample size is required for the kernel density estimator to estimate π well. In our simulations, we find $\widehat{\pi}$ could be small and hence violate the positivity assumption, yielding variation in the construction of pseudo-outcome $\widehat{\varphi}$ and larger estimation error of dose-response compared with the plug-in estimator. In applications, prior knowledge on the conditional density, including a reliable parametric model on π from domain knowledge or information on the lower bound on π due to study design might help us reduce the variation of $\widehat{\pi}$ and estimate the conditional density better, which could improve the performance of doubly robust estimator proposed.

C.3. Surrogates Dependent on Treatment and Covariates

In this section we provide additional simulation results on the setting where surrogates may depend on the treatment and covariates. In real applications, the surrogates $\mathbf{S}$ are post-treatment short-term outcomes and are likely to depend on pre-treatment covariates and the treatment. We follow the same setting and estimation procedures as in Section 6 (so the nuisance estimation error is set manually) but modify the conditional distribution of $\mathbf{S}$ given $(A, \mathbf{V})$ as

$$\mathbf{S} \sim N\left((V_1 + A, V_2 - A)^\top, \boldsymbol{I}_2\right).$$

The RMSE is estimated from $M = 500$ replications of the data-generating process and estimation procedures. The results of DR-learner and plug-in estimators are summarized in Figure 5.



(a) $n = 500$ (b) $n = 2000$

Figure 5. Root mean square error Versus α, where $n^{-\alpha}$ is the estimation error of the nuisance functions.

Figure 5 shows when surrogate outcomes $\mathbf{S}$ depend on treatment and covariates, the doubly robust estimator enjoys better estimation accuracy compared with the plug-in estimator when α is small (and hence nuisance estimation error is large). As α increases and the nuisance estimation error becomes smaller, the plug-in estimator gradually outperforms the doubly robust estimator. The results are similar to those in Section 6 and readers are referred to Section 6 for more discussion and explanation.

C.4. Comparison Between Supervised and Semi-supervised Methods

In this section we compare the supervised estimator, which implements the method in (Kennedy et al., 2017) on labeled data $\mathcal{L}$, with our semi-supervised estimator incorporating the unlabeled data $\mathcal{U}$ with surrogate outcomes. We follow the same setting in Section 6 and implement method in (Kennedy et al., 2017) via similar cross-fitting techniques as in Algorithm 1 except that there is no need to fit $\rho(a, \mathbf{x}) = \mathbb{P}(R = 1 \mid A = a, \mathbf{X} = \mathbf{x})$ since their method only uses labeled data. All the nuisance functions are estimated by correctly specified parametric models. We generate samples with sample size $n \in \{10^{2.6}, 10^{2.8}, \ldots, 10^{4.6}\}$, apply supervised and semi-supervised methods to estimate $\theta(1)$ and repeat the process $M = 500$ times. The results are summarized in Figure 6.

Figure 6 shows the semi-supervised method incorporating unlabeled data with surrogate outcomes has a smaller estimation error compared with the supervised method only using labeled data. The improvement comes from many aspects: As discussed in Section 5.2, the asymptotic variance of our semi-supervised method is smaller than that of the supervised method only using labeled data; Also in the semi-supervised setting, both labeled and unlabeled data are used to estimate the nuisance functions, making nuisance estimation more accurate than supervised method solely based on labeled data since the effective sample size of semi-supervised method is larger.

D. Real Data Analysis

In this section we apply the proposed method to the Job Corps study, conducted in the 1990s to evaluate the effects of the publicly funded U.S. Job Corps program. The Job Corps program targets a population between 16 and 24 years old living in the U.S. and coming from low-income households, where participants received an average of 1,200 hours of vocational training over approximately eight months. Schochet et al. (2001) and Schochet et al. (2008) discuss the study design in detail and analyze the effects of the Jobs Corps program on various outcomes. They found the Job Corps program effectively increased educational attainment, prevented arrests, and increased employment and earnings. The effects of the Job Corps program have been extensively studied under different causal inference frameworks (Flores & Flores-Lagunes, 2009; Huber, 2014; Frölich & Huber, 2017).

However, these previous studies on the Job Corps program mainly considered binary treatment definitions. In this work, we are interested in estimating the effects of different doses of participation in the program on future involvement in the criminal

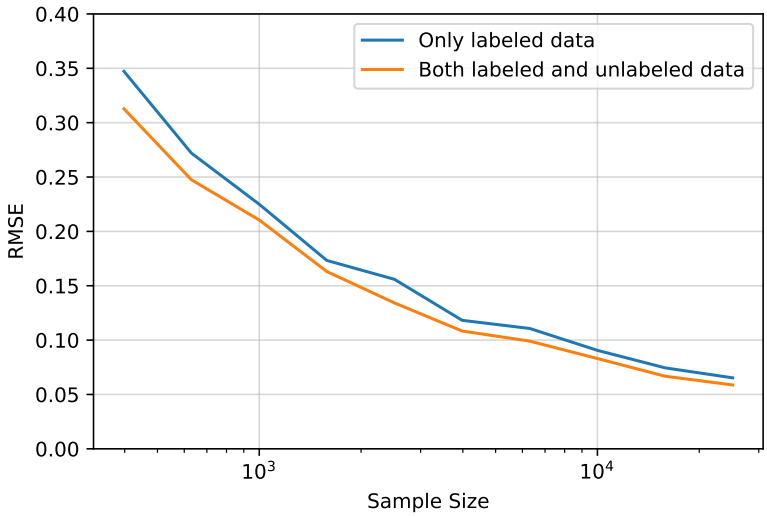


Figure 6. RMSE versus sample size when nuisance functions are estimated by correctly specified parametric models.

justice system, namely the number of arrests in the fourth year after the program (outcome Y). Specifically, our treatment variable A is defined as the total hours spent either in academic or vocational classes of the program. The short-term surrogate outcome S is the proportion of weeks employed in the second year after the program. Huber et al. (2020) used generalized propensity score weighting to estimate the continuous treatment effects of time spent in the Job Corps program under a mediation analysis framework. We re-analyze their dataset publically available on Harvard dataverse (Huber, 2020), using the doubly robust estimator proposed as an illustration of our method.

We follow Huber et al. (2020) and focus on $n = 4000$ samples with a positive treatment (i.e., $A_i > 0, 1 \leq i \leq n$). To identify the causal estimand, we invoke the conditional exchangeability in Section 3, where a set of covariates is conditioned on to adjust for confounding bias. The covariate set $\mathbf{V}$ we adjust for confounding consists of age, gender, ethnicity, education, marital status, previous employment status and income, welfare receipt during childhood, and family background (e.g., parents' education). Missing dummies are created for covariates in $\mathbf{V}$ containing missing values. The readers are referred to Table 4 in Huber et al. (2020) for descriptive statistics of the pretreatment covariates as well as the treatment, surrogate, and outcome variables in the data. Conditioning on a rich set of covariates is important since it enables us to identify the causal estimand by making the conditional exchangeability assumption plausible. In our analysis, the outcome Y (number of arrests in year 4) is omitted for 25% of the samples randomly to mimic the setting where the primary outcome is missing and a surrogate outcome is used as auxiliary information. We estimate the treatment effects $\theta(a)$ for each of $a \in \{100, 150, 200, \ldots, 2000\}$ using both plug-in-style estimator and doubly robust estimator in Algorithm 1, where the nuisance functions ρ, μ, τ are estimated by superlearner (Van der Laan et al., 2007) combining generalized linear model and random forests, conditional density π is estimated by kernel density estimator. The estimated dose response curve is plotted in Figure 7.

Figure 7 shows that, as participants spend more hours in the program, the expected number of arrests in year 4 has a decreasing trend, which confirms the conclusion that such training programs effectively reduce involvement in the criminal justice system. Importantly, the dose-response fitted by the doubly robust estimator is very similar to the results in Huber et al. (2020): the shape of the function is similar to their Figure 2. (Note that the specific values on the y-axis are different since they plot a contrast effect and we instead plot the expectation of potential outcome Y^a.) As pointed out in Huber et al. (2020), the treatment effect is highly nonlinear, which is further verified by the curve estimated from the doubly robust estimator. By contrast, the curve estimated by a simple plug-in estimator in (3) fails to capture this non-linearity, possibly because it suffers from a large first-order bias.

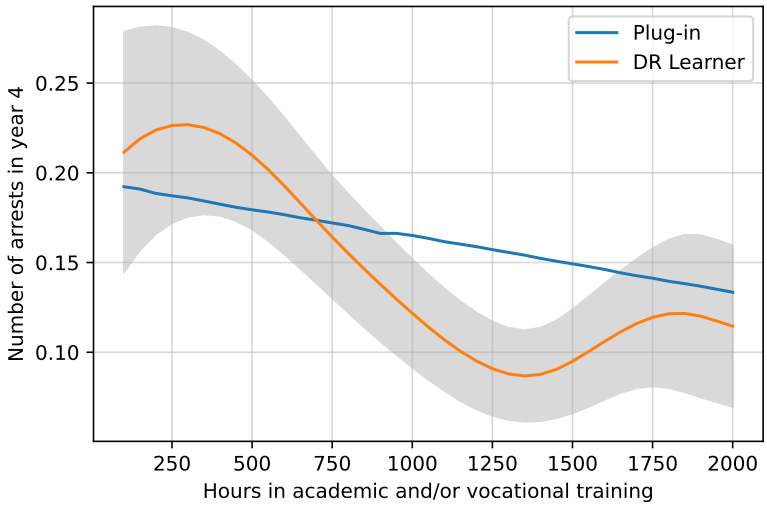


Figure 7. Dose response curve of number of arrests in year 4 (outcome) versus hours in academic and/or vocational training (treatment)