Proceedings of Machine Learning Research vol 140:1–21, 2022 1st Conference on Causal Learning and Reasoning

Integrative R-learner of heterogeneous treatment effects combining experimental and observational studies

Author names withheld

Editor: Under Review for CLeaR 2022

Abstract

The gold-standard approach to estimating heterogeneous treatment effects (HTEs) is randomized controlled trials (RCTs)/controlled experimental studies, where treatment randomization mitigates confounding biases. However, experimental data are usually small in sample size and limited in subjects' diversity due to expensive costs. On the other hand, large observational studies (OSs) are becoming increasingly popular and accessible. However, OSs might be subject to hidden confounding whose existence is not testable. We develop an integrative R-learner for the HTE and confounding function by leveraging experimental data for identification and observational data for boosting efficiency. We form a regularized loss function for the HTE and confounding function that bears the Neyman orthogonality property, allowing flexible models for the nuisance function estimation. The key novelty of the proposed integrative R-learner is to impose different regularization terms for the HTE and confounding function so that the possible smoothness or sparsity of the confounding function can be leveraged to improve the HTE estimation. Our integrative R-learner has two benefits: first, it provides a general framework that can accommodate various HTE models for loss minimization; second, without any prior knowledge of hidden confounding in the OS, the proposed integrative R-learner is consistent and asymptotically at least as efficient as the estimator using only the RCT. The experiments based on extensive simulation and a real-data application adapted from an educational experiment show that the proposed integrative R-learner outperforms alternative approaches.

Keywords: Asymptotic efficiency; Empirical risk minimization; Hidden confounding; Penalized regression; Series estimator;

1. Introduction

Heterogeneous treatment effects (HTEs) are the causal effects of a treatment or an intervention on an outcome of interest given the subjects' characteristics, which is the key query in various areas, such as precision medicine (Hamburg and Collins, 2010), offline contextual bandit evaluation (Dudík et al., 2011) and personalized policy recommendations (Athey, 2017). Randomized controlled trials (RCTs) are the gold-standard method for evaluating the HTEs because the randomization of treatment allocation in an RCT ensures that treatment and control groups are comparable and biases are minimized to the extent possible. However, RCTs might be limited in sample size, due to the possible expensive costs and substantive efforts during the data collection procedure. Besides, RCTs might also be limited in subjects' diversity, due to the restrictive inclusion/exclusion criteria for enrollment. The limitations in both scale and scope of RCTs can compromise the causal effect estimation which leads to under-powered HTE estimators.

On the other hand, due to the larger sample size, broader diversity, and more accessibility, observational studies (OSs) have gained popularity. There is considerable interest in the literature to use OSs to improve the generalizability of the RCT findings (Dong et al., 2020; Wu and Yang, 2021). Besides, OSs have also been widely used for the HTE estimation in various domains such as healthcare and education. For example, clinicians can use electronic health records and disease

© 2022 .

registries to help recommend the individual-level treatment; teachers use school records to assign kids to different classes. However, due to lack of randomization, the OS has a universal concern of unmeasured confounders, which cannot be verified in practice with itself alone (Pearl, 2009). The no unmeasured confounding assumption requires all relevant predictors of treatment and outcome to be measured, which is barely held in the OS. For example, physicians prescribe a medicine based on the patient's conditions of which some are not recorded in medical charts. With the OS alone, the hidden confounding makes the HTE unidentifiable and can lead to a biased HTE estimation. To mitigate confounding bias with the OS alone, classic approaches include instrumental variables (Angrist et al., 1996) and sensitivity analysis (Robins et al., 2000). However, finding the instrumental variables which must be related to the outcome only through the treatment usually depends on domain knowledge and they might not exist. The sensitivity analysis is used to evaluate the robustness of the estimators to no unmeasured confounding, but it could not identify the confounding function. The RCT and OS have their own strengths and weaknesses, therefore, we are motivated to develop an integrative learner by leveraging the RCT for identifying the HTE and the OS for boosting its efficiency.

Under the no unmeasured confounding assumption in the OS, there have been many methods to estimate the HTE, ranging from the classic methods such as regression-based methods and (augmented) inverse probability weighting methods (Robins et al., 1994; Scharfstein et al., 1999) to some recent more flexible methods based on machine learning models such as neural network (Shalit et al., 2017), random forest (Wager and Athey, 2018), boosting (Powers et al., 2018). More recently, motivated by Robinson's transformation (Robinson, 1988), Nie and Wager (2021) formalized a loss function to estimate the HTE which can be optimized by arbitrary loss-minimization procedures. The resulting solution is called R-learner, which bears the Neyman orthogonality property(Neyman, 1959) that the R-learner has a root-N rate of convergence with weaker conditions on nuisance function approximation. Therefore, the R-learner can incorporate a much broader class of flexible machine learning models to approximate nuisance functions. Carrying the benefits of the R-learner, we develop an integrative R-leaner of the HTE combining RCT data and OS data where the OS might be subject to hidden confounding. The key step is to formulate a regularized loss function for the HTE and confounding function with the Neyman orthogonality. More importantly, we impose different regularization terms for the HTE and confounding function, so that the resulting integrative R-learner can capture the possible smoothness or sparsity of the confounding function to improve the HTE estimation. The main benefit of our proposed integrative R-learner is two-folds: first, it can incorporate flexible models for nuisance functions as well as the HTE and confounding function; second, without any prior knowledge of hidden confounding in the OS, the integrative R-learner is consistent to the true HTE and asymptotically at least as efficient as the estimator using the RCT alone.

The rest of the paper is organized as follows. We introduce the basic setup in Section 2, the proposed R-learner in Section 3, and its theoretical analysis in Section 4. We describe two alternative methods for data integration in Section 5 and conduct extensive experiments for comparison based on simulated and real data in Section 6. Finally, we conclude the paper with a discussion in Section 7.

2. Setup

We focus on a setting with a binary treatment variable $A \in \{0, 1\}$, let $X \in \mathcal{X} \subset \mathbb{R}^p$ be a vector of pre-treatment covariates, and let $Y \in \mathbb{R}$ be the outcome of interest. We use the potential outcomes

framework (Neyman, 1923; Rubin, 1974) to define causal effects and let $Y(a)$ be the potential outcome had the subject taken the treatment a, for $a = 0, 1$. Then the HTE can be characterized by $\tau(X) = \mathbb{E}\{Y(1) - Y(0) \mid X\}$.

Two data sources are accessible: the RCT with independent and identically distributed (i.i.d.) n subjects $\{(X_i, A_i, Y_i, S_i)\}_{i\in\mathcal{I}_n}$ and the OS with i.i.d. m subjects $\{(X_i, A_i, Y_i, S_i)\}_{i\in\mathcal{I}_m}$, where $\mathcal{I}_n$ and $\mathcal{I}_m$ are index sets of the RCT and the OS, respectively. Let S_i be the binary indicator of the ith subject in the the RCT: $S_i = 1$ if $i \in \mathcal{I}_n$ and 0 if $i \in \mathcal{I}_m$. We denote two nuisance functions: the propensity score $e(X, S) = P(A = 1 \mid X, S)$, and the conditional mean outcome $\mu(X, S) = \mathbb{E}(Y \mid X, S)$.

One of the fundamental challenges in causal inference is the identification of the HTE because $Y(1)$ and $Y(0)$ cannot be jointly observed on the same subject. Therefore, to overcome this challenge, we invoke the two common assumptions in the causal inference (e.g., Rubin, 1980; Rosenbaum and Rubin, 1983; Tsiatis, 2006) and data integration literature (e.g., Colnet et al., 2020; Degtiar and Rose, 2021).

Assumption 1 *(Ignorability of treatment assignment in the RCT and causal consistency): (a)* $A \perp\!\!\!\perp \{Y(0), Y(1)\} \mid (X, S = 1)$. *(b) There exist* $c_1, c_2 \in (0, 1)$ *such that* $\pi_A(X) \in [c_1, c_2]$ *for all* $X \in \mathcal{X}$. *(c)* $Y = Y(A)$.

Assumption 2 *(Transportability of the HTE):* $\mathbb{E}\{Y(1) - Y(0) \mid X, S = s\} = \tau(X), s = 0, 1$.

Assumption 1*(a)* is also called no unmeasured confounding, which holds in the RCT by default due to the randomized treatment assignment. Under Assumption 1, the HTE can be identified based on the RCT. On the contrary, we do not impose the ignorability of treatment assignment for the OS, which may be stringent in practice. Assumption 2 is needed to extend the definition and identification of causal effects from the RCT to the OS, which holds when X captures all treatment effect modifiers. Its plausibility relies on experts' domain knowledge. It is a common assumption in the data integration literature and the weakest one among the mean exchangeability assumption (Dahabreh et al., 2019), i.e., $\mathbb{E}\{Y(a) \mid X, S = s\}$ and ignorability of study participation, i.e., $Y(a) \perp\!\!\!\perp S \mid X, (a = 0, 1)$ (Stuart et al., 2011; Buchanan et al., 2018).

3. The proposed integrative R-learner

3.1. Identification of the HTE and confounding function and formulation of the loss functions

When lack of the ignorability of treatment assignment, We define the bias between the difference in conditional mean outcomes in the OS and the HTE as the confounding function

$$c(X) = \mathbb{E}(Y \mid X, A = 1, S = 0) - \mathbb{E}(Y \mid X, A = 0, S = 0) - \tau(X). \tag{1}$$

By Assumption 1*(c)*, $\mathbb{E}(Y \mid X, A = a, S = 0) = \mathbb{E}\{Y(a) \mid X, A = a, S = 0\}$. If all the confounding variables are captured in the OS, i.e., $A \perp\!\!\!\perp \{Y(0), Y(1)\} \mid (X, S = 0)$, then $\mathbb{E}\{Y(a) \mid X, A = a, S = 0\} = \mathbb{E}\{Y(a) \mid X, S = 0\}$, which implies $c(X) = 0$ under Assumption 2. The confounding function is not identified based on the OS alone, however, we show that combining the RCT and OS can identify the HTE and confounding function. The key step toward identifying and estimating the HTE and confounding function is introducing the residual

$$\epsilon = Y - \mathbb{E}(Y \mid X, A = 0, S) - A\{\tau(X) + (1 - S)c(X)\}. \tag{2}$$

Under Assumptions 1-2, we can show the fact $\mathbb{E}(\epsilon \mid X, A, S) = 0$ as follows:

a) Conditioning on $(X, A = 0, S)$: $\mathbb{E}(\epsilon \mid X, A = 0, S) = 0$;

b) Conditioning on $(X, A = 1, S = 0)$: $\mathbb{E}(\epsilon \mid X, A = 1, S = 0) = \mathbb{E}(Y \mid X, A = 1, S = 0) - \mathbb{E}(Y \mid X, A = 0, S = 0) - \{\tau(X) + c(X)\} \overset{(1)}{=} 0$;

c) Conditioning on $(X, A = 1, S = 1)$: $\mathbb{E}(\epsilon \mid X, A = 1, S = 1) = \mathbb{E}(Y \mid X, A = 1, S = 1) - \mathbb{E}(Y \mid X, A = 0, S = 1) - \tau(X) = \mathbb{E}\{Y(1) \mid X, S = 1\} - \mathbb{E}\{Y(0) \mid X, S = 1\} - \tau(X) = 0$. The last second equality is implied by Assumptions 1 and the last equality is implied by Assumption 2.

Besides, from the definition of $\mu(X, S)$, we have

$$\begin{aligned} \mu(X, S) &= \mathbb{E}(Y \mid X, A = 1, S)e(X, S) + \mathbb{E}(Y \mid X, A = 0, S)\{1 - e(X, S)\} \\ &= \mathbb{E}(Y \mid X, A = 0, S) + \{\tau(X) + (1 - S)c(X)\}e(X, S). \end{aligned} \tag{3}$$

Combining (2) and (3), we have

$$Y - \mu(X, S) = \{\tau(X) + (1 - S)c(X)\}\{A - e(X, S)\} + \epsilon, \tag{4}$$

which motivates the mean squared error loss function minimization problem for identifying $\{\tau(\cdot), c(\cdot)\}$

$$\{\tau(\cdot), c(\cdot)\} = \underset{\tilde{\tau}, \tilde{c}}{\operatorname{argmin}} \left(\mathbb{E}\left[Y - \mu(X, S) - \{\tilde{\tau}(X) + (1 - S)\tilde{c}(X)\}\{A - e(X, S)\}\right]^2 \right), \tag{5}$$

Naturally, (5) leads to the "empirical" loss function to estimate the HTE and confounding function:

$$\underset{\tilde{\tau}, \tilde{c}}{\operatorname{argmin}} \left(\mathbb{P}_N \left[Y_i - \mu(X_i, S_i) - \{\tilde{\tau}(X_i) + (1 - S_i)\tilde{c}(X_i)\}\{A_i - e(X_i, S_i)\}\right]^2 + \Lambda_\tau + \Lambda_c \right), \tag{6}$$

where $N = n + m, \mathbb{P}_N(X_i) = N^{-1} \sum_{i \in \mathcal{I}_n \bigcup \mathcal{I}_m} X_i$, and Λ_τ and Λ_c are regularizers on the complexity of the $\tau(\cdot)$ and $c(\cdot)$ functions to avoid overfitting. The penalties could be added directly on the coefficients of the features such as penalized regression, or indirectly on the black-box models such as boosting (Mason et al., 1999) or random forest (Breiman, 2001) by restricting the depth of the decision trees.

It is worth discussing the relationship of the proposed estimator with existing approaches. If replacing the functions $\tau(X) + (1 - S)c(X)$ in (4) with finite-dimensional parameters, it coincides with the decomposition proposed in Robinson (1988) to estimate parametric components in partially linear models. This decomposition bears Neyman orthogonality (Neyman, 1959) to obtain $\sqrt{N}$-rate estimation of target parameters while the convergence rates of the nuisance functions are slower than $\sqrt{N}$, which has recently also been called "rate double robustness" property (Rotnitzky et al., 2021). This approach has gained considerable attention in the causal inference community for treatment effect estimation such as G-estimation (Robins, 2004), double machine learning (Chernozhukov et al., 2018), causal forest (Athey et al., 2019) and R-learner (Nie and Wager, 2021). However, the aforementioned work only considers causal effect estimation in the data satisfying the no unmeasured confounding assumption. The proposed loss function can be viewed as an extension (although not straightforward) of the ones of Robinson (1988) and Nie and Wager (2021) to the data integration context combining the RCT and the OS. Thus, we call the proposed method "integrative R-learner".

In order to solve (6), however, two difficulties arise: (i) Except that the propensity score in the RCT $e(X, S = 1)$ might be known, other nuisance functions $\mu(X, S)$ and $e(X, S = 0)$ are unknown; (ii) $\tau(\cdot)$ and $c(\cdot)$ may have different complexities, but rare off-the-shelf software can directly solve (6) with different Λ_τ and Λ_c. We will show that the possible smoothness or sparsity of $c(\cdot)$ can be leveraged to improve the estimation of $\tau(\cdot)$ over using the RCT alone (Theorem 1). Thus, using different Λ_τ and Λ_c is critical to actualize the potential gain from the OS for the HTE estimation. In the next two subsections, we will describe how we address the two concerns.

3.2. Estimation of the nuisance functions

Cross-fitting is an increasingly popular prediction method to incorporate flexible machine learning approaches in classical semiparametrics (Schick, 1986; Robins et al., 2008, 2017; Chernozhukov et al., 2018; Newey and Robins, 2018; Kennedy, 2020; Nie and Wager, 2021). Not only does it help to make theories more elegant, but also to overcome the overfitting/high-complexity phenomena that commonly arise in highly adaptive machine learning methods (Chernozhukov et al., 2018). It is a simple procedure by splitting data into fitting and prediction parts.

We adopt cross-fitting to estimate the nuisance functions. We randomly split the data into K (commonly set to be 5) equal-size folds. Each subject i falls into one and only one data fold. Let $k(i)$ be the index set of the data fold where the subject i belongs. We use the samples that do not belong to $k(i)$ to fit the user-specified models (such as lasso regression, Xgboost, etc), which results in the estimated conditional mean outcome and propensity score models, denoted as $\hat{\mu}^{-k(i)}(\cdot,\cdot)$ and $\hat{e}^{-k(i)}(\cdot,\cdot)$ respectively. Next, using the estimated nuisance functions, we predict the function values on the subjects belonging to $k(i)$. Repeating the procedure on each fold, we finally obtain the nuisance function estimates of all the subjects, i.e., $\hat{\mu}^{-k(i)}(X_i, S_i)$ and $\hat{e}^{-k(i)}(X_i, S_i), i \in \mathcal{I}_n \cup \mathcal{I}_m$.

This step leads to the final empirical loss function for estimating the HTE and confounding function:

$$\{\hat{\tau}(\cdot), \hat{c}(\cdot)\} = \underset{\tilde{\tau},\tilde{c}}{\operatorname{argmin}} \Bigg(\mathbb{P}_N \Big[Y_i - \hat{\mu}^{-k(i)}(X_i, S_i) - \{\tilde{\tau}(X_i) + (1 - S_i)\tilde{c}(X_i)\}\{A_i - \hat{e}^{-k(i)}(X_i, S_i)\} \Big]^2 + \Lambda_\tau + \Lambda_c \Bigg). \tag{7}$$

3.3. Series approximations of the HTE and confounding function

The general framework can accommodate flexible models for nuisance functions as well as the HTE and the confounding function. For exposition, we now describe the integrative R-learner adopting series estimation of the HTE and confounding function. Series estimation is a widely applicable nonparametric method to approximate the unknown conditional expectation that can have a flexible functional form. The approximation consists of D basis functions which have a broad class such as polynomial basis, wavelet basis, and spline, etc, where D grows with the sample size. Besides the versatility of the series estimators, their asymptotic properties have also been studied thoroughly in a large body of literature (see, e.g., Stone, 1985; Eastwood and Gallant, 1991; Newey, 1997; Chen, 2007; Cattaneo and Farrell, 2013; Belloni et al., 2015, and the references therein).

We use the series method to approximate $\tau(X)$ and $c(X)$ as follows,

$$\tau(X) = p_\tau^\intercal(X)\beta_\tau + r_\tau,\ c(X) = p_c(X)^\intercal \beta_c + r_c, \tag{8}$$

where $p_\tau(X) = \{p_{\tau,1}(X), \ldots, p_{\tau,D}(X)\}^\intercal$ and $p_c(X) = \{p_{c,1}(X), \ldots, p_{c,D}(X)\}^\intercal$ are two vectors of basis functions, where D can increase with sample size N, r_τ and r_c are approximation errors.

Combining (4) and (8) leads to the square loss function,

$$L(\beta) = \mathbb{P}[Y - \mu(X,S) - g(X,S;\beta)\{A - e(X,S)\}]^2,$$

where $g(X,S;\beta) = p_\tau^\intercal(X)\beta_\tau + (1-S)p_c(X)^\intercal\beta_c$ and $\beta = (\beta_\tau^\intercal, \beta_c^\intercal)^\intercal$. We denote the empirical loss function as

$$\widehat{L}_N(\beta) = \mathbb{P}_N\big[Y_i - \hat{\mu}^{-k(i)}(X_i,S) - g(X_i,S_i;\beta)\{A - \hat{e}^{-k(i)}(X_i,S_i)\}\big]^2. \tag{9}$$

Since the number of basis functions controls the smoothness of series estimators, $\tau(X)$ and $c(X)$ may require different number of basis functions due to their possibly different underlying complex nature. Thus, we require different regularization parameters. We will show that the possible smooth or sparsity of $c(X)$ leads to the efficiency gain of the integrative R-learner compared to using the RCT alone. Thus, it is critical to impose different regularization parameters for $\tau(X)$ and $c(X)$. Specifically, let $\widehat{\beta} = \text{argmin}_{b\in\mathbb{R}^{2D}}\left\{\widehat{L}_N + \sum_{d=1}^D s_{\lambda_\tau}(|b_{\tau,j}|) + \sum_{d=1}^D s_{\lambda_c}(|b_{c,j}|)\right\}$, where $s_\lambda(\cdot)$ is the smoothly clipped absolute deviation (SCAD) penalty function (Fan and Li, 2001) and the tuning parameters λ_τ and λ_c can be chosen through cross-validation by grid searching two pre-specified ranges. One way to simplify the tuning process is to add a scale tuning parameter between λ_τ and λ_c. For example, let $sc = \lambda_\tau/\lambda_c$, and set the search range of the scale tuning parameter as $sc \in \{0, 0.5, 1, 1.5\}$, which represents that no penalty on $\tau(\cdot)$, the penalty on $\tau(\cdot)$ less than, equal to, and larger than that on $c(\cdot)$ respectively. For each sc, we can use cross-validation to choose λ_c with the off-the-shelf software `cv.ncvreg` function in the R package `ncvreg` (Breheny and Huang, 2011). The final fitted model corresponds to the scale tuning parameter that has the minimum mean cross-validated error (see more details in Algorithm 1).

When the shrinkage parameters are complex and high-dimensional such as those in deep neural network, Xgboost, etc, we propose a variant of integrative R-learner by iteratively learning $\tau(\cdot)$ and $c(\cdot)$ for the real data experiment (see Appendix C).

Algorithm 1 Penalty scale searching for the integrative R-learner

Data: The RCT and the OS data $\{(X_i, A_i, Y_i, S_i)\}_{i\in\mathcal{I}_n\cup\mathcal{I}_m}$; the estimated nuisance functions $\hat{\mu}^{-k(i)}(X_i,S_i)$ and $\hat{e}^{-k(i)}(X_i,S_i), i \in \mathcal{I}_n \cup \mathcal{I}_m$

Initialize $cmv = \infty$

Result: β_τ and the HTE $\tau(\cdot) = p_\tau(\cdot)^\intercal\beta_\tau$

for $sc = 0, 0.5, 1, 1.5$ **do**

  Fit $Y_i - \hat{\mu}^{-k(i)}(X_i,S_i)$ on $\{p_\tau(X_i), (1-S_i)p_c(X_i)\}\{A_i - \hat{e}^{-k(i)}(X_i,S_i)\}$ with penalized regression where the corresponding shrinkage parameters are $(sc\lambda_c, \lambda_c)$ and λ_c is chosen by cross-validation;

  Record the above mean cross-validated error cvm_{sc} and the fitted coefficients $\beta_{sc,\tau}$;

  if $cvm_{sc} < cvm$ **then**

    $\beta_\tau \leftarrow \beta_{sc,\tau}$

    $cvm \leftarrow cvm_{sc}$

  end

end

4. Theoretical analysis

The main goal of our theoretical analysis is to show that our integrative R-learner is consistent and at least as efficient as the R-learner based only on the RCT. For concreteness, we focus on the R-

learner based on series estimation. First of all, under the regularity conditions given in Fan and Li (2001), $\widehat{\beta}$ satisfies the selection consistency and oracle properties, i.e., $\|\widehat{\beta} - \beta\| = O_p(N^{-1/2})$. Thus, it suffices to focus on $\widehat{L}_N(\beta)$ in (9) for theoretical analysis. Replacing its estimated nuisance functions with the true ones, we denote it as $L_N(\beta) = \mathbb{P}_N\big[Y_i - \mu(X_i, S_i) - g(X_i, S_i; \beta)\{A_i - e(X_i, S_i)\}\big]^2$. The following two assumptions are used to assure the difference between $\widehat{L}_N(\beta)$ and $L_N(\beta)$ diminishes with a relatively fast rate with N under which the Neyman orthogonality of the loss function renders the impact of the estimated nuisance functions negligible.

Assumption 3 *$\|\tau(X)\|_\infty$, $\|c(X)\|_\infty$ and $\mathbb{E}[\{Y - \mu(X, S)\}^2 \mid X, S]$ are bounded, for any $X \in \mathcal{X}, S \in \{0, 1\}$.*

Assumption 4 *$\mathbb{E}\left[\{\mu(X, S) - \hat{\mu}(X, S)\}^2\right] = O(a_N^2)$ and $\mathbb{E}\left[\{e(X, S) - \hat{e}(X, S)\}^2\right] = O(a_N^2)$ for some sequence a_N such that $a_N = O(N^{-r})$ with $r > 1/4$.*

Assumption 3 is plausible in many practical problems. Assumption 4 relaxes the usual $\sqrt{N}$-consistency to a converge rate required to be only faster than $N^{-1/4}$, which can occur in a broad class of models such as single-index models, generalized additive models, partially linear models, and lasso regression, etc (Horowitz, 2009; Belloni et al., 2014; Kennedy, 2016). Besides, we impose the standard assumptions for series estimation (Belloni et al., 2015). For simplicity, let $p_i = p(X_i, A_i, S_i) = \{A_i - e(X_i, A_i)\}\{p_\tau^\intercal(X_i), (1 - S_i)p_c(X_i)^\intercal\}^\intercal$, $\hat{p}_i = \{A_i - \hat{e}(X_i, A_i)\}\{p_\tau^\intercal(X_i), (1 - S_i)p_c(X_i)^\intercal\}^\intercal$, and $\xi_q = \sup_{x,a,s} \|p(x, a, s)\|$.

Assumption 5 *Uniformly over all N, eigenvalues of $\Gamma := \mathbb{E}(p_i p_i^\intercal)$ are bounded above and away from zero.*

Assumption 6 *(a) For each N and q, there are finite constants c_q and l_q such that for each $f \in \mathcal{F}$, $\|r_f\|_{F,2} := \sqrt{\int_{x\in\mathcal{X}} r_f^2(x)dF(x)} \le c_q$ and $\|r_f\|_{F,\infty} := \sup_{x\in\mathcal{X}} |r_f(x)| \le l_q r_q$. (b) $\sqrt{\xi_q^2 \log(q/N)}(1 + \sqrt{q} l_q c_q) \to 0$. (c) $l_q c_q \to 0$. (d) $\sqrt{N/q} l_q c_q \to 0$.*

Assumption 7 *$\mathbb{E}(\epsilon^2 \mid X, S, A) = \sigma^2$ for some constant σ.*

By the proposed method, the estimated treatment effect is $\hat{\tau}(X) = \hat{p}_\tau^\intercal \hat{\beta}_\tau$, where $\hat{\beta} := (\hat{\beta}_\tau^\intercal, \hat{\beta}_c^\intercal) = \operatorname{argmin}_b \mathbb{P}_N \left\{Y_i - \hat{\mu}^{-k(i)}(X_i, S_i) - \hat{p}_i^\intercal b\right\}^2$. If only using the RCT data, then the estimated treatment effect is $\hat{\tau}_{\text{rct}} = \hat{p}_\tau^\intercal \hat{\beta}_{\text{rct}}$, where $\hat{\beta}_{\text{rct}} = \operatorname{argmin}_b \mathbb{P}_N S_i \left\{Y_i - \hat{\mu}^{-k(i)}(X_i, S_i) - \hat{p}_i^\intercal b\right\}^2$. Denote $\mathbb{V}(Z)$ as the asymptotic variance of the random variable Z or the asymptotic covariance matrix of the random vector Z.

Theorem 1 *Under Assumptions 1-7, $\hat{\tau}(x) - \tau(x) = O_p(N^{-1/2})$ for any $x \in \mathcal{X}$, and $\mathbb{V}(\hat{\tau}) \le \mathbb{V}(\hat{\tau}_{rct})$. Moreover, the equality holds if and only if $p_\tau(X) = M p_c(X)$, for some constant matrix M.*

Theorem 1 shows that the integrative R-learner is consistent and at least as efficient as the HTE estimator with only the RCT. More interestingly, when the underlying HTE $\tau(\cdot)$ is less restrictive than $c(\cdot)$, the integrative R-learner has strictly larger efficiency than the estimator with the RCT alone. We present a proof sketch here (the details are presented in Appendix A):

a) Under Assumptions 3-4, we can show: $\widehat{L}_N(\beta) = L_N(\beta) + O_P(a_N^2)$, where $a_N = O(N^{-r})$ with $r > 1/4$, which implies $\hat{\beta} = \text{argmin}_b \, \mathbb{P}_N \left(Y_i - \mu - p_i^\intercal b\right)^2 + O_P(a_N^2)$.

b) By Pointwise Normality of series (see Theorem 4.2 in Belloni et al., 2015)

$$\sqrt{N}(\hat{\beta} - \beta) = \mathbb{E}(p_i p_i^\intercal)^{-1} \sqrt{N}(\mathbb{P}_N - \mathbb{P}) \begin{pmatrix} p_\tau \epsilon \\ p_c \epsilon \end{pmatrix} + o_P(1) + O_P\left(\sqrt{N} a_N^2\right).$$

c) By algebra, we can show $\mathbb{V}^{-1}(\hat{\beta}_\tau) - \mathbb{V}^{-1}(\hat{\beta}_{\text{rct}})$ is non-negative definitive.

5. Other methods

We describe two alternative integrative HTE estimators. They both start with constructing an adjusted outcome $\widetilde{Y}$ such that $\mathbb{E}(\widetilde{Y} \mid X, S = 1) = \tau(X)$ under Assumptions 1-2. A common choice of the adjusted outcome is

$$\widetilde{Y} = \frac{A\{Y - Q_S(X,1)\}}{e(X,S)} + Q_S(X,1) - \frac{(1-A)\{Y - Q_S(X,0)\}}{1 - e(X,S)} - Q_S(X,0), \tag{10}$$

where $Q_S(X,a) = \mathbb{E}(Y \mid X, A = a, S)$. Then, one can fit $\widetilde{Y}$ on X to obtain the HTE estimator. This approach is also a variant of the U-learner (Nie and Wager, 2021). In the following, the first integrative estimator constructs the similar outcome-adjusted equality for $c(X)$ in the OS, while the second one proposed by Kallus et al. (2018) estimates $c(X)$ based on imposing the linear model assumption and using the $\widetilde{Y}_i$ in RCT as the unbiased estimates of $\tau(X_i)$.

5.1. Outcome-adjusted method

We show that the adjusted outcome for the OS can be used to identify $c(X)$ once $\tau(X)$ is identified from the RCT as stated in Proposition 2 (see the proof in Appendix B).

Proposition 2 *Under Assumptions 1-2, we have*

$$\mathbb{E}(\widetilde{Y} \mid X, S = 0) = \tau(X) + c(X), \tag{11}$$

where $\widetilde{Y}$ is (10).

Here is one way to get the integrative estimator for the HTE by combining the facts $\mathbb{E}(\widetilde{Y} \mid X, S = 1) = \tau(X)$ and (11). We first fit $\widetilde{Y}$ on X using the RCT to find important features, and then fit $\widetilde{Y}$ on X again but using all the RCT and the OS samples with incorporating the selected features in $\tau(X)$ and all the features in $c(X)$. The procedure is summarized below.

a) Estimate the adjusted outcomes, denoted as $\widetilde{Y}_i$, $i \in \mathcal{I}_n \cup \mathcal{I}_m$;

b) Fit the adjusted outcomes on basis functions $b_{Kn}(X)$ with lasso regression using the RCT samples: $\widetilde{Y}_i \sim \alpha^\intercal b_{Kn}(X_i)$, $i \in \mathcal{I}_n$; obtain the selected basis functions denoted as $b^*_{Kn}(X_i)$;

c) Lasso regression with the RCT and the OS samples on the selected basis functions $b^*_{Kn}(X_i)$ and basis functions $b_{Lm}(X_i)$ of the OS: $Y_i \sim \alpha^\intercal b^*_{Kn}(X_i) + \beta^\intercal b_{Lm}(X_i)$. Note that $b_{Lm}(X_i) = 0, i \in \mathcal{I}_n$, and the penalty is only added on β;

d) Obtain the estimated HTE $\hat{\tau}(X) = \hat{\alpha}^\intercal b^*_{Kn}(X)$, where $\hat{\alpha}$ is got from step c).

However, this approach relies on the estimation of the adjusted outcomes and can be highly unstable since the inverse of the estimated propensity scores might bring in extreme values.

5.2. 2-step procedure proposed by Kallus et al. (2018)

Assuming a linear model for the confounding function $c(X) = \theta^\intercal X$, Kallus et al. (2018) proposed an integrative 2-step HTE estimation procedure summarized below.

a) Use any regression methods, e.g., random forest, to estimate $Q_0(X, 1)$ and $Q_0(X, 0)$ with the OS samples, then take its difference, denoted as $\hat{\omega}(\cdot) = \widehat{Q}_0(\cdot, 1) - \widehat{Q}_0(\cdot, 1)$;

b) Apply the definition in (1) to learn $c(X)$: $\hat{\theta} = \text{argmin}_\theta \sum_{i \in \mathcal{I}_n} \left\{ \widetilde{Y}_i - \hat{\omega}(X_i) + \theta^\intercal X_i \right\}^2$, which gives the estimated HTE immediately $\hat{\tau}(x) = \hat{\omega}(x) - \hat{\theta}^\intercal x$.

This approach has a different loss function from ours and it shows the consistency of $\hat{\tau}(\cdot)$ at a rate that is governed by the rate of $\hat{w}(\cdot)$, but it has restrictions on the form of $c(X)$.

6. Experiments

We evaluate the finite-sample performances of our proposed method through simulated data and a real data application of the Tennessee STAR study which aims to measure the effect of class size on test scores (Word et al., 1990; Krueger, 1999). To illustrate the generality of our proposed framework, we use lasso regression with polynomial basis functions with degree 2 for the simulated data and machine learning approaches including Xgboost (Chen and Guestrin, 2016) and random forest (Breiman, 2001) for the application.

We evaluate four categories of methods:

- Outcome adjusted methods as introduced in Section 5.1 with either the RCT, the OS or the integrated data, denoted as "*rct_oa*", "*os_oa*" or "*int_oa*" respectively.
- R-learner methods with either the RCT, the OS or the integrated data, denoted as "*rct_rlearner*", "*os_rlearner*" or "*int_rlearner*" (our method).
- Learning two regression functions with random forest for the treated and control and taking their difference using either the RCT or the OS, denoted as "*rct_rf*", "*os_rf*" respectively.
- The method proposed in Kallus et al. (2018) where the confounding function is estimated by linear regression or lasso regression with polynomial basis functions, denoted as "*NK_linear*" or "*NK_lasso*", respectively.

The performances are measured by the square root of the mean square errors (RMSE) between the estimated condition treatment effect and the ground truth over the test data, i.e. $\text{RMSE}(\hat{\tau}) = \sqrt{T^{-1} \sum_{t=1}^{\intercal} \{\hat{\tau}(X_{\text{test}}^{(t)}) - \tau(X_{\text{test}}^{(t)})\}^2}$, where T is the sample size of the test data.

6.1. Simulation study

We modify two simulation settings considered in the literature. For the first set of simulation studies, we consider the simulation setting with scalar covariates as in Kallus et al. (2018), where the RCT and the OS have only partial overlap and the true confounding function is linear. For the second set of simulation studies, we follow the simulation setting with multiple covariates in Yang et al. (2020) where the true confounding function is non-linear and the strength of confounding bias can be adjusted.

6.1.1. SIMULATION STUDY I

We generate the RCT data with sample size $n = 200$ as follows:

$$X_{\text{rct}} \sim \text{Uniform}[-1, -1], U_{\text{rct}} \sim \mathcal{N}(0, 1), A_{\text{rct}} \sim \text{Bernoulli}(0.5).$$

Next we generate the OS data with sample size $m = 500$ as follows: $A_{\text{os}} \sim \text{Bernoulli}(0.5)$, then the observed covariate X_{os} and the unobserved covariate U_{os} are sampled from a bivariate Normal distribution with sample size m

$$\begin{pmatrix} X_{\text{os}} \\ U_{\text{os}} \end{pmatrix} \mid A_{\text{os}} \sim \mathcal{N}\left\{ \begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 & A_{\text{os}} - 0.5 \\ A_{\text{os}} - 0.5 & 1 \end{pmatrix} \right\}.$$

For both datasets, we generate the outcomes $Y = 1 + X + 0.5X^2 + U + A\tau(X) + 0.5\epsilon$, where $\epsilon \sim \mathcal{N}(0, 1)$. Then by the definition of the confounding function, we can get $c(X) = \mathbb{E}(U \mid X, A = 1, S = 0) - \mathbb{E}(U \mid X, A = 0, S = 0) = X$. Considering both linear and non-linear conditional treatment effect, we let $\tau(X) = 1 + 2X + lX^2$, $l \in \{0, 0.75\}$. The test data X_{test} is sampled from $\mathcal{N}(0, 1)$ with size $T = 10^5$. The experiment results of RMSE($\hat{\tau}$) for different methods are presented in Table 1, our proposed method has lowest RMSE among all, which has dramatic decrease compared with the methods only using the RCT or the OS and the method in Kallus et al. (2018), while the outcome-adjusted integrative estimators also gain decent improvement. While l is changed from 0 to 0.75, the treatment effect becomes non-linear and all the RMSEs have increased, among which the methods based on random forest have the largest increase. *"NK_linear"* gains the improvement compared with *"rct_rf"* and *"os_rf"* due to the correctly specified confounding function, but since its consistency at a relatively slow rate based on random forest, it is hard to beat the outcome-adjusted method and the integrative R-learner. Besides, *"int_oa"* has pretty decent performance in such a low dimensional setting.

Table 1: The averaged RMSE($\hat{\tau}$) over 100 experiment replicates (standard error in parentheses).

l		rct_oa	os_oa	int_oa	rct_rlearner	os_rlearner	int_rlearner	rct_rf	os_rf	NK_lasso	NK_linear
0	avg	0.67	0.97	0.39	0.73	1.00	**0.36**	1.21	1.29	1.40	0.95
	se	(0.05)	(0.01)	(0.04)	(0.04)	(0.01)	(0.04)	(0.02)	(0.01)	(0.07)	(0.01)
0.75	avg	0.80	1.01	0.62	0.74	1.00	**0.51**	1.56	1.31	1.43	0.98
	se	(0.05)	(0.01)	(0.04)	(0.04)	(0.01)	(0.05)	(0.02)	(0.01)	(0.07)	(0.01)

6.1.2. SIMULATION STUDY II

We first generate population of size $T = 10^5$, independent p-dimensional covariates $X_j \sim \mathcal{N}(0, 1)$, $j = 1, \ldots, p$. Let the treatment modifier $W = (X_1, X_2, \ldots, X_{p-1})$, where $p = 10$. Then we generate the potential outcomes $Y(a)$ by $Y(a) \mid X = X_1 + X_p + a\tau(W) + \epsilon(a)$, where $\epsilon(a) \sim \mathcal{N}(0, 1)$, for $a = 0, 1$. Now we generate the RCT: the RCT slection indicator is generated by $S \mid X \sim \text{Bernoulli}\{\pi_S(X)\}$, where $\text{logit}\{\pi_S(X)\} = -7 - X_1 + X_2$, which results in around 200 RCT samples. Then an OS sample is simply randomly selected from the population with size $m = 500$. The treatments are assigned as follows: $A \mid X, S = 1 \sim \text{Bernoulli}\{e(X, S = 1)\}$, where $\text{logit}\{e(X, S = 1)\} = -1 - X_1 + X_2$, and $A \mid X, S = 0 \sim \text{Bernoulli}\{e(X, S = 0)\}$, where $\text{logit}\{e(X, S = 0)\} = -1 - X_1 + X_2 + bX_p$. We let X_p is unobserved in both the RCT and the OS, thus b indicates the strength of unmeasured confounding: the larger absolute value of b induces the larger confounding bias. We vary $b \in \{0, 2.5\}$, and the confounding function equals to zero when

$b = 0$. The observed outcomes $Y = AY(1) + (1 - A)Y(0)$. We also consider different dimensions of the treatment modifiers by letting $\tau(W) = 1 + \sum_{j=1}^{J} X_j, J \in \{1, 9\}$. The test data X_{test} is all the first $p - 1$ dimensional covariates in the initially generated population of size $T = 10^5$. As shown in Table 2, when b increasing, the RMSEs of methods with the OS alone will increase. The proposed integrative R-learner performs the best among the all. *"NK_linear"* suffers high loss due to its incorrectly specified model for the non-linear $c(X)$ and *"int_oa"* also perform badly due to the instability induced from the inverse of the propensity scores estimated by the high dimensional covariates.

Table 2: The averaged RMSE($\hat{\tau}$) over 100 experiment replicates (standard error in parentheses).

J	b		rct_oa	os_oa	int_oa	rct_rlearner	os_rlearner	int_rlearner	rct_rf	os_rf	NK_lasso	NK_linear
9	0	avg	1.86	1.07	1.46	1.63	1.00	**0.58**	1.99	1.87	1.92	1.90
		se	(0.13)	(0.04)	(0.09)	(0.03)	(0.02)	(0.02)	(0.02)	(0.01)	(0.02)	(0.12)
	2.5	avg	1.86	1.55	2.00	1.63	1.53	**1.10**	1.99	2.12	1.81	1.86
		se	(0.13)	(0.01)	(0.08)	(0.03)	(0.01)	(0.05)	(0.02)	(0.01)	(0.02)	(0.12)
1	0	avg	1.11	0.55	0.88	0.88	0.50	**0.27**	1.31	0.95	1.02	1.61
		se	(0.14)	(0.03)	(0.07)	(0.04)	(0.02)	(0.03)	(0.03)	(0.01)	(0.02)	(0.11)
	2.5	avg	1.11	1.41	1.37	0.88	1.33	**0.49**	1.31	1.28	0.87	1.55
		se	(0.14)	(0.01)	(0.06)	(0.04)	(0.01)	(0.05)	(0.03)	(0.01)	(0.02)	(0.11)

6.2. Real data application

The Tennessee Student/Teacher Achievement Ratio (STAR) experiment is a randomized controlled experiment aiming to study the effect of class size on students' standardized test scores. We take two kinds of class sizes (small classes containing 13-17 pupils and regular classes containing 22-25 pupils) as a binary treatment A. The outcome Y is the sum of the math, reading, and listening standardized test scores. And covariates X = (gender, race, birth month, birthday, birth year, free lunch given or not, teacher id). Among all the 4218 students, 2413 were randomly assigned to regular-size classes $(A = 0)$ and 1805 to small classes $(A = 1)$.

We extract the RCT, the OS, and test data from the original data as Kallus et al. (2018) in the following. We split all samples over a binary variable U: rural or inner-city ($U = 1$; 2811 students) vs. urban or suburban ($U = 0$; 1407 students) which is known to strongly affect outcome (Krueger, 1999). The RCT is randomly sampled half of the students with $U = 1$. Then the OS consist of two parts: (a) From samples with $U = 1$, taking samples with $T = 0$ but not in the RCT and samples with $T = 1$ whose outcomes are in the lower half of outcomes among samples with $T = 1$ & $U = 1$; (b) From samples with $U = 0$, taking all samples with $T = 0$ and samples with $T = 1$ whose outcomes are in the lower half of outcomes among samples with $T = 1$ & $U = 0$. This procedure results in the RCT and the OS not fully overlapping since the RCT only has rural or inner-city students, and it also biases the treatment effect estimates in the OS downward since only the lower half of scores among samples with $T = 1$ are selected into the OS. Let U be the unmeasured confounder, the outcome and treatment in the OS are confounded significantly. Since the ground truth, τ_i is inaccessible in practice, it is replaced with an unbiased estimate from all samples, which is defined as $A_i\{Y_i - \mathbb{E}(Y)\}/P(A = 1) - (1 - A_i)\{Y_i - \mathbb{E}(Y)\}/\{1 - P(A = 1)\}$, $i \in$ all samples, where $\mathbb{E}(Y)$ is estimated by the average of the outcomes and $P(A = 1)$ is estimated by the proportion of treated samples among all samples. The test data is the held-out sample of all samples excluding the RCT.

Besides the regression with the SCAD penalty implemented in the simulation study of Section 6.1, we can also incorporate a broad class of off-the-shelf machine learning approaches such as

Xgboost, random forest, etc. To realize different penalties on $\tau(\cdot)$ and $c(\cdot)$, we propose an iterative learning procedure in Appendix C.

We initialize the confounding function values as the ones estimated from the method in Kallus et al. (2018), do the iteration 20 times, and obtain the results of RMSE in Figure 1. With the number of the iteration increasing, the RMSE of our integrative R-learner increases to a steady value. The estimators with only the RCT or the OS give the RMSE around 80, and the method in Kallus et al. (2018) has significant improvement (= 71.01) while still worse than ours.

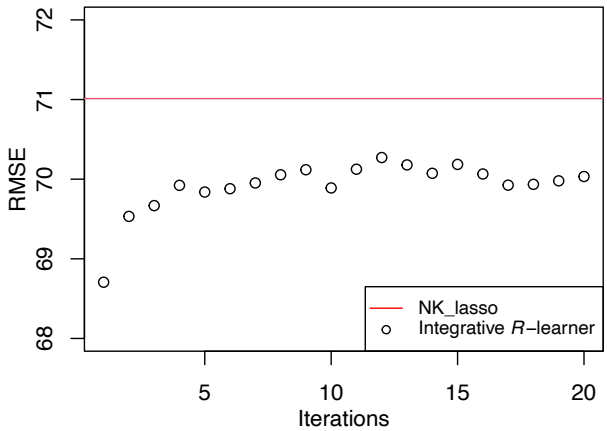

Figure 1: The plot of RMSE for the integrative R-learner at each iteration vs. RMSE of the estimator proposed in Kallus et al. (2018)

7. Discussion

We propose the integrative R-learner which can identify both the HTE and confounding function, and improve the efficiency of the HTE estimator. It is a general framework that has a broad class of choices for employing flexible models in either nuisance function approximation or loss-minimization procedures. An interesting direction of future work is to extend the integrative R-learner to multiple treatments which commonly arise in reality. It can be constructed similarly based on the multivariate version of Robinson's transformation and the corresponding R-learner (Nie and Wager, 2021). Besides, it would also be interesting to extend the integrative R-learner to off-policy evaluation under covariate shifts (Uehara et al., 2020) or generalizable individualized decision rules learning (Zhao et al., 2019; Wu and Yang, 2021) to improve the efficiency of the estimated policy values or the estimated optimal decision rules.

References

Joshua D Angrist, Guido W Imbens, and Donald B Rubin. Identification of causal effects using instrumental variables. *Journal of the American Statistical Association*, 91(434):444–455, 1996.

Susan Athey. Beyond prediction: Using big data for policy problems. *Science*, 355(6324):483–485, 2017.

Susan Athey, Julie Tibshirani, and Stefan Wager. Generalized random forests. *The Annals of Statistics*, 47(2):1148–1178, 2019.

Alexandre Belloni, Victor Chernozhukov, and Christian Hansen. Inference on treatment effects after selection among high-dimensional controls. *The Review of Economic Studies*, 81(2):608–650, 2014.

Alexandre Belloni, Victor Chernozhukov, Denis Chetverikov, and Kengo Kato. Some new asymptotic theory for least squares series: Pointwise and uniform results. *Journal of Econometrics*, 186 (2):345–366, 2015.

Patrick Breheny and Jian Huang. Coordinate descent algorithms for nonconvex penalized regression, with applications to biological feature selection. *Annals of Applied Statistics*, 5(1):232–253, 2011.

Leo Breiman. Random forests. *Machine Learning*, 45(1):5–32, 2001.

Ashley L Buchanan, Michael G Hudgens, Stephen R Cole, Katie R Mollan, Paul E Sax, Eric S Daar, Adaora A Adimora, Joseph J Eron, and Michael J Mugavero. Generalizing evidence from randomized trials using inverse probability of sampling weights. *Journal of the Royal Statistical Society. Series A (Statistics in Society)*, 181(4):1193, 2018.

Matias D Cattaneo and Max H Farrell. Optimal convergence rates, bahadur representation, and asymptotic normality of partitioning estimators. *Journal of Econometrics*, 174(2):127–143, 2013.

Tianqi Chen and Carlos Guestrin. Xgboost: A scalable tree boosting system. In *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, pages 785–794, 2016.

Xiaohong Chen. Large sample sieve estimation of semi-nonparametric models. *Handbook of Econometrics*, 6:5549–5632, 2007.

Victor Chernozhukov, Denis Chetverikov, Mert Demirer, Esther Duflo, Christian Hansen, Whitney Newey, and James Robins. Double/debiased machine learning for treatment and structural parameters. *The Econometrics Journal*, 21(1):C1–C68, 01 2018.

Bénédicte Colnet, Imke Mayer, Guanhua Chen, Awa Dieng, Ruohong Li, Gaël Varoquaux, Jean-Philippe Vert, Julie Josse, and Shu Yang. Causal inference methods for combining randomized trials and observational studies: a review. *arXiv preprint arXiv:2011.08047*, 2020.

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

Irina Degtiar and Sherri Rose. A review of generalizability and transportability. *arXiv preprint arXiv:2102.11904*, 2021.

Lin Dong, Shu Yang, Xiaofei Wang, Donglin Zeng, and Jianwen Cai. Integrative analysis of randomized clinical trials with real world evidence studies. *arXiv preprint arXiv:2003.01242*, 2020.

Miroslav Dudík, John Langford, and Lihong Li. Doubly robust policy evaluation and learning. *Proceedings of the 28th International Conference on International Conference on Machine Learning*, page 1097–1104, 2011.

Brian J Eastwood and A Ronald Gallant. Adaptive rules for seminonparametric estimators that achieve asymptotic normality. *Econometric Theory*, 7(3):307–340, 1991.

Jianqing Fan and Runze Li. Variable selection via nonconcave penalized likelihood and its oracle properties. *Journal of the American Statistical Association*, 96(456):1348–1360, 2001.

Margaret A Hamburg and Francis S Collins. The path to personalized medicine. *New England Journal of Medicine*, 363(4):301–304, 2010.

Joel L Horowitz. *Semiparametric and nonparametric methods in econometrics*, volume 12. Springer, 2009.

Nathan Kallus, Aahlad Manas Puli, and Uri Shalit. Removing hidden confounding by experimental grounding. In *Proceedings of the 32nd International Conference on Neural Information Processing Systems*, pages 10911–10920, 2018.

Edward H Kennedy. Semiparametric theory and empirical processes in causal inference. In *Statistical Causal Inferences and their Applications in Public Health Research*, pages 141–167. Springer, 2016.

Edward H Kennedy. Optimal doubly robust estimation of heterogeneous causal effects. *arXiv preprint arXiv:2004.14497*, 2020.

Alan B Krueger. Experimental estimates of education production functions. *The Quarterly Journal of Economics*, 114(2):497–532, 1999.

Llew Mason, Jonathan Baxter, Peter Bartlett, and Marcus Frean. Boosting algorithms as gradient descent in function space. volume 12, pages 512–518, 1999.

Whitney K Newey. Convergence rates and asymptotic normality for series estimators. *Journal of Econometrics*, 79(1):147–168, 1997.

Whitney K Newey and James R Robins. Cross-fitting and fast remainder rates for semiparametric estimation. *arXiv preprint arXiv:1801.09138*, 2018.

Jerzy Neyman. On the application of probability theory to agricultural experiments. essay on principles. section 9. (translated and edited by dm dabrowska and tp speed, statistical science (1990), 5, 465-480). *Annals of Agricultural Sciences*, 10:1–51, 1923.

Jerzy Neyman. Optimal asymptotic tests of composite hypotheses. *Probability and Statsitics*, pages 213–234, 1959.

Xinkun Nie and Stefan Wager. Quasi-oracle estimation of heterogeneous treatment effects. *Biometrika*, 108(2):299–319, 2021.

Judea Pearl. *Causality*. Cambridge university press, 2009.

Scott Powers, Junyang Qian, Kenneth Jung, Alejandro Schuler, Nigam H Shah, Trevor Hastie, and Robert Tibshirani. Some methods for heterogeneous treatment effect estimation in high dimensions. *Statistics in Medicine*, 37(11):1767–1787, 2018.

James Robins, Lingling Li, Eric Tchetgen, and Aad van der Vaart. Higher order influence functions and minimax estimation of nonlinear functionals. In *Probability and Statistics: Essays in Honor of David A. Freedman*, pages 335–421. Institute of Mathematical Statistics, 2008.

James M Robins. Optimal structural nested models for optimal sequential decisions. In *Proceedings of the second Seattle Symposium in Biostatistics*, pages 189–326. Springer, 2004.

James M Robins, Andrea Rotnitzky, and Lue Ping Zhao. Estimation of regression coefficients when some regressors are not always observed. *Journal of the American Statistical Association*, 89 (427):846–866, 1994.

James M Robins, Andrea Rotnitzky, and Daniel O Scharfstein. Sensitivity analysis for selection bias and unmeasured confounding in missing data and causal inference models. In *Statistical Models in Epidemiology, the Environment, and Clinical Trials*, pages 1–94. Springer, 2000.

James M Robins, Lingling Li, Rajarshi Mukherjee, Eric Tchetgen Tchetgen, and Aad van der Vaart. Minimax estimation of a functional on a structured high-dimensional model. *The Annals of Statistics*, 45(5):1951–1987, 2017.

Peter Robinson. Root- n-consistent semiparametric regression. *Econometrica*, 56(4):931–54, 1988.

Paul R Rosenbaum and Donald B Rubin. The central role of the propensity score in observational studies for causal effects. *Biometrika*, 70(1):41–55, 1983.

Andrea Rotnitzky, Ezequiel Smucler, and James M Robins. Characterization of parameters with a mixed bias property. *Biometrika*, 108(1):231–238, 2021.

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

Donald B Rubin. Randomization analysis of experimental data: The fisher randomization test comment. *Journal of the American Statistical Association*, 75(371):591–593, 1980.

Daniel O Scharfstein, Andrea Rotnitzky, and James M Robins. Adjusting for nonignorable drop-out using semiparametric nonresponse models. *Journal of the American Statistical Association*, 94 (448):1096–1120, 1999.

Anton Schick. On asymptotically efficient estimation in semiparametric models. *The Annals of Statistics*, 14(3):1139–1151, 1986.

Uri Shalit, Fredrik D. Johansson, and David Sontag. Estimating individual treatment effect: generalization bounds and algorithms. In *Proceedings of the 34th International Conference on Machine Learning*, volume 70, pages 3076–3085, 2017.

Charles J Stone. Additive regression and other nonparametric models. *The Annals of Statistics*, 13 (2):689–705, 1985.

Elizabeth A Stuart, Stephen R Cole, Catherine P Bradshaw, and Philip J Leaf. The use of propensity scores to assess the generalizability of results from randomized trials. *Journal of the Royal Statistical Society: Series A (Statistics in Society)*, 174(2):369–386, 2011.

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

Masatoshi Uehara, Masahiro Kato, and Shota Yasui. Off-policy evaluation and learning for external validity under a covariate shift. *Advances in Neural Information Processing Systems*, 33:49–61, 2020.

Stefan Wager and Susan Athey. Estimation and inference of heterogeneous treatment effects using random forests. *Journal of the American Statistical Association*, 113(523):1228–1242, 2018.

Elizabeth Word, John Johnston, Helen P Bain, B DeWayne Fulton, Jayne B Zaharias, Charles M Achilles, Martha N Lintz, John Folger, and Carolyn Breda. The state of tennessee's student/teacher achievement ratio (star) project. *Tennessee Board of Education*, 1990.

Lili Wu and Shu Yang. Transfer learning of individualized treatment rules from experimental to real-world data. *arXiv preprint arXiv:2108.08415*, 2021.

Shu Yang, Donglin Zeng, and Xiaofei Wang. Elastic integrative analysis of randomized trial and real-world data for treatment heterogeneity estimation. *arXiv preprint arXiv:2005.10579*, 2020.

Ying-Qi Zhao, Donglin Zeng, Catherine M Tangen, and Michael L LeBlanc. Robustifying trial-derived optimal treatment rules for a target population. *Electronic Journal of Statistics*, 13(1): 1717–1743, 2019.