Estimating individual treatment effects under unobserved confounding using binary instruments

Anonymous Author(s)
Affiliation
Address
email

Abstract

Estimating individual treatment effects (ITEs) from observational data is relevant in many fields such as personalized medicine. However, in practice, the treatment assignment is usually confounded by unobserved variables and thus introduces bias. A remedy to remove the bias is the use of instrumental variables (IVs). Such settings are widespread in medicine (e.g., trials where compliance is used as binary IV). In this paper, we propose a novel, multiple robust machine learning framework, called MRIV, for estimating ITEs using binary IVs and thus yield an unbiased ITE estimator. Different from previous work for binary IVs, our framework estimates the ITE directly via a pseudo outcome regression. (1) We provide a theoretical analysis where we show that our framework yields multiple robust convergence rates: our ITE estimator achieves fast convergence even if several nuisance estimators converge slowly. (2) We further show that our framework asymptotically outperforms state-of-the-art plug-in IV methods for ITE estimation. (3) We build upon our theoretical results and propose a tailored deep neural network architecture called MRIV-Net for ITE estimation using binary IVs. Across various computational experiments, we demonstrate empirically that our MRIV-Net achieves state-of-the-art performance. To the best of our knowledge, our MRIV is the first multiple robust machine learning framework tailored to estimating ITEs in the binary IV setting.

1 Introduction

Individual treatment effects (ITEs) are relevant across many disciplines such as marketing [30] and personalized medicine [38]. Knowledge about ITEs provides insights into the heterogeneity of treatment effects, and thus help in potentially better treatment decisions.

Many recent works that use machine learning to estimate ITEs assume unconfoundedness [1, 12, 20, 26, 31]. That is, all variables that affect both treatment assignment and outcome are observed in the data. In practice, however, it is common that some confounders are not reported in the data. Typical, examples are race, income, gender, or the socioeconomic status of patients, which are not stored in medical files. If the confounding is sufficiently strong, standard methods for estimating ITEs suffer from confounding bias [22], which may lead to inferior treatment decisions.

To handle unobserved confounders, instrumental variables (IVs) can be leveraged to relax the assumption of unconfoundedness and still compute reliable ITE estimates. IV methods were originally developed in economics [36], but, only recently, there is a growing interest in combining IV methods with machine learning (see Sec. 3). Importantly, IV methods outperform classical ITE estimators if a sufficient amount of confounding is not observed [14]. We thus aim at estimating ITEs from observational data under unobserved confounding using IVs.

In this paper, we consider the setting where a single binary instrument is available. This setting is widespread in personalized medicine (and other applications such as marketing or public policy) [8]. In fact, the setting is encountered in essentially all observational or randomized studies with observed non-compliance [15]. As an example, consider a randomized controlled trial, where treatments are randomly assigned to patients and their outcomes are observed. Due to some potentially unobserved confounders (e.g., income, education), some patients refuse to take the treatment initially assigned to them. Here, the treatment assignment serves as a binary IV. Moreover, such RCTs have been widely used by public decision-makers, e.g., to analyze the effect of health insurance on health outcome (see the so-called *Oregon health insurance experiment*) [13] or the effect of military service on lifetime earnings [2].

We propose a novel machine learning framework (called MRIV) for estimating ITEs using binary IVs. Importantly, our framework is multiple robust, i.e., it is consistent even if several inputs are misspecified. Existing methods aimed at binary IVs are plug-in estimators [5, 15], yet which are not multiple robust and, on top of that, it is well known that such estimators suffer from plug-in bias [11]. To address this, we develop MRIV: Different from existing works, we estimate the ITE directly by performing a pseudo outcome regression on a multiple robust parametrization of the efficient influence function. As such, we refrain from estimating the different components of the ITE separately but, instead, leverage the pseudo outcome.

We provide a theoretical analysis where we show that our framework achieves a multiple robust convergence rate, i.e., our MRIV converges with a fast rate even if several nuisance parameters converge slowly. We further show that, compared to existing plug-in IV methods, the performance of our framework is asymptotically superior. Finally, we leverage our framework and, on top of it, build a tailored deep neural network called MRIV-Net.

Contributions:[1] (1) We propose a novel, multiple robust machine learning framework (called MRIV) to learn the ITE using the binary IV setting. To the best of our knowledge, ours is the first that is multiple robust, i.e., consistent even if several inputs are misspecified. (2) We prove that MRIV achieves a multiple robust convergence rate. We further show that our MRIV is asymptotically superior to existing plug-in estimators. (3) We propose a tailored deep neural network, called MRIV-Net, which builds upon our framework to estimate ITEs. We demonstrate that MRIV-Net achieves state-of-the-art performance.

2 Problem setup

Data generating process: We observe data $\mathcal{D} = (x_i, z_i, a_i, y_i)_{i=1}^n$ consisting of $n \in \mathbb{N}$ observations of the tuple (X, Z, A, Y). Here, $X \in \mathcal{X}$ are observed confounders, $Z \in \{0, 1\}$ is a binary instrument, $A \in \{0, 1\}$ is a binary treatment, and $Y \in \mathbb{R}$ is an outcome of interest. Furthermore, we assume the existence of unobserved confounders $U \in \mathcal{U}$, which affect both the treatment A and the outcome Y. The causal graph is shown in Fig. 1.

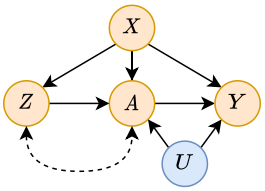

Figure 1: Underlying causal graph. The instrument Z has a direct influence on the treatment A, but does not have a direct effect on the outcome Y. Note that we allow for unobserved confounders for both Z–A (dashed line) and A–Y (given by U).

The above setting is common in practice, e.g., in personalized medicine. In particular, modeling treatments as binary variables is consistent with previous literature on causal effect estimation and standard in medical practice [24]. For example, our setting is always encountered in (all) settings where individuals do not comply with their treatment assignment. Our setting thus includes essentially all randomized controlled trials (RCT) with non-compliance. Such RCTs have been extensively used by public decision-makers, e.g., to analyze the effect of health insurance on health outcome [13] or the effect of military service on lifetime earnings [2]. In these cases, the instrument Z is the treatment assignment and indicates the decision whether a patient should receive treatment.

We build upon the potential outcomes framework [25] for modeling causal effects. Let $Y(a, z)$ denote the potential outcome that would have been observed under $A = a$ and $Z = z$. Following

[1] Codes are in the supplementary materials. Codes are also available at https://anonymous.4open.science/r/MRIV-Net-0AC4 (Upon acceptance, we replace the link and point to a public GitHub repository).

previous literature on IV estimation [34], we impose the following standard IV assumptions on the data generating process.

Assumption 1 (Standard IV assumptions [34, 35])**.** We assume: (1) *Exclusion:* $Y(a, z) = Y(a)$ for all $a, z \in \{0, 1\}$, i.e., the instrument has no direct effect on the patient outcome; (2) *Independence:* $Z \perp\!\!\!\perp U \mid X$; (3) *Relevance:* $Z \not\perp\!\!\!\perp A \mid X$, (iv) *The model includes all A–Y confounder:* $Y(a) \perp\!\!\!\perp (A, Z) \mid (X, U)$ for all $a \in \{0, 1\}$.

Assumption 1 is fairly unrestrictive and thus widely fulfilled in practice [2, 4, 15]. Note that Assumption 1 does not prohibit the existence of unobserved Z–A confounders. On the contrary, it merely prohibits the existence of unobserved counfounders that affect all Z, A, and Y simultaneously, as it is standard in IV settings [35]. A practical and widespread example where Assumption 1 is satisfied are randomized controlled trials (RCTs) with non-compliance [15]. Here, the treatment assignment Z is randomized, but the actual relationship between treatment A and outcome Y may still be confounded. For instance, in the *Oregon health insurance experiment* [13], people were given access to health insurance (Z) by a lottery with aim to study the effect of health insurance (A) on health outcome (Y) [13]. Here, non-compliance information is observed because the lottery winners needed to sign up for health insurance.

Objective: In this paper, we are interested in estimating the *individual treatment effect* (ITE)

$$\tau(x) = \mathbb{E}[Y(1) - Y(0) \mid X = x]. \tag{1}$$

If there is no unobserved confounding ($U = \emptyset$), the ITE is identifiable from observational data under mild positivity assumptions [26]. However, in practice, it is often unlikely that all confounders are observable. To account for this, we leverage the instrument Z to identify the ITE. We state the following assumption for identifiability.

Assumption 2 (Identifiability of the ITE)**.** At least one of the following two statements holds true: (1) There is no additive U–Z interaction in $\mathbb{E}[A \mid Z, X, U]$, i.e., $\mathbb{E}[A \mid Z = 1, X, U] - \mathbb{E}[A \mid Z = 0, X, U] = \mathbb{E}[A \mid Z = 1, X] - \mathbb{E}[A \mid Z = 0, X]$; or (2) there is no additive U–a interaction in $\mathbb{E}[Y(a) \mid X, U]$, i.e., $\mathbb{E}[Y(1) - Y(0) \mid X, U] = \mathbb{E}[Y(1) - Y(0) \mid X]$.

Under Assumptions 1 and 2, the ITE is identifiable [34]. It can be written as

$$\tau(x) = \frac{\mu_1^Y(x) - \mu_0^Y(x)}{\mu_1^A(x) - \mu_0^A(x)} = \frac{\delta_Y(x)}{\delta_A(x)}, \tag{2}$$

where $\mu_i^Y(x) = \mathbb{E}[Y \mid Z = i, X = x]$ and $\mu_i^A(x) = \mathbb{E}[A \mid Z = i, X = x]$. Even if Assumption 2 does not hold, the quantity on the right-hand side of Eq. (2) still allows for interpretation. If no unobserved Z–A confounders exist, it can be interpreted as conditional version of the *local average treatment effect* (LATE) [15, 5] under a monotonicity assumption. Furthermore, under a no-current-treatment-value-interaction assumption, it can be interpreted as conditional *treatment effect on the treated* (ETT) [34]. [2] This has an important implication for our results: If Assumption 2 does not hold in practice, our estimates still provide conditional LATE or ETT estimates under the respective assumptions because they are based on Eq. (2). If Assumption 2 does hold, all three – i.e., ITE, conditional LATE, and ETT – coincide [34].

3 Related work

ITE methods without unconfoundedness: Various machine learning methods for estimating ITEs *without* unobserved confounding have been proposed in recent literature [1, 12, 18, 20, 26, 31, 39, 40]. To remove plug-in bias, the DR learner performs a second stage regression on the uncentered influence function of the average treatment effect [16, 11]. However, under unobserved confounding, all of these methods are biased (see Appendix). As a result, this hampers their performance in our setting.

ITE methods for unobserved confounding: IV methods address the problem of unobserved confounding by exploiting the variance in treatment and outcome induced by the instruments. Traditionally, two-stage least squares (2SLS) has been used for estimating causal effects [36, 4]. 2SLS was

[2] The conditional LATE measures the ITE for individuals which are part of the complier subpopulation, i.e., the subpopulation for which it holds that $Z = A$. The conditional ETT measures the ITE for treated individuals.

originally developed in economics, and follows a two-stage procedure: it performs a first stage regression of treatment A on the instrument Z, and then uses the fitted values for a second stage regression to predict the outcome Y. Several nonparametric methods have been developed in econometric to generalize 2SLS in order to account for non-linearities within the data [21, 33], yet these are limited to low-dimensional settings.

Only recently, machine learning has been integrated into IV methods. These are: [27] and [37] generalize 2SLS by learning complex feature maps using kernel methods and deep learning, respectively. [14] adopts a two-stage neural network architecture that performs the first stage via conditional density estimation. [6] and [29] leverage moment conditions for IV estimation. Finally, [29] is a doubly robust machine learning framework to estimate ITEs via a pseudo outcome regression. However, the aforementioned methods are not specifically designed for the binary IV setting but, rather, for multiple IVs or treatment scenarios. Hence, they require additive confounding in order to identify the ITE. We later show for our setting that, even under additive confounding, our framework outperforms the above IV methods.

Binary IVs: In the binary IV setting, current methods proceed by estimating $\mu_i^Y(x)$ and $\mu_i^A(x)$ separately, before plugging them in Eq. 2 [15, 3, 5]. As result, these suffer from plug-in bias and do *not* offer robustness properties. In contrast, multiple robust estimators based on semiparametric efficiency theory have been proposed only for average treatment effects (ATEs) and <u>not</u> for ITEs. For example, for ATEs, Wang and Tchetgen Tchetgen [34] derive a multiple robust parametrization of the efficient influence curve for the uncentered ATE in the binary setting. However, there exists no similar approach for ITE estimation.

Research gap: To the best of our knowledge, there exists no method for ITE estimation under unobserved confounding that is multiple robust. To fill this gap, we propose MRIV: a multiple robust machine learning framework tailored to the binary IV setting.

4 MRIV for estimating ITEs using binary instruments

In the following, we present our MRIV framework for estimating ITEs under unobserved confounding (Sec. 4.1). We then derive we derive an asymptotic convergence rates for MRIV (Sec. 4.2) and finally use our framework to develop a tailored deep neural network called MRIV-Net (Sec. 4.4).

4.1 Framework

Motivation: A naïve approach to estimate the ITE is to leverage the identification result in Eq. (2). Assuming that we have estimated the nuisance components $\hat{\mu}_i^Y$ and $\hat{\mu}_i^A$ for $i \in \{0, 1\}$, we can simply plug them into Eq. (2) to obtain the so-called (plug-in) Wald estimator $\hat{\tau}_{\mathrm{W}}(x)$ [32].

However, in practice, the true ITE curve $\tau(x)$ is often simpler (e.g., smoother, more sparse) than its complements $\mu_i^Y(x)$ or $\mu_i^A(x)$ [18]. In this case, $\hat{\tau}_{\mathrm{W}}(x)$ is inefficient because it models all components separately, and, to address this, our proposed framework estimates τ directly using a pseudo outcome regression.

Overview: We now propose MRIV. MRIV is a two-stage meta learner that takes any base method for ITE estimation as input. For instance, the base method could be the Wald estimator from Eq. (2), any other IV method such as 2SLS, or a deep neural network (as we propose in our MRIV-Net later in Sec. 4.4). In Stage 1, MRIV produces nuisance estimators $\hat{\mu}_0^Y(x)$, $\hat{\mu}_0^A(x)$, $\hat{\delta}_A(x)$, and $\hat{\pi}(x)$, where $\hat{\pi}(x)$ is an estimator of the propensity score $\pi(x) = \mathbb{P}(Z = 1 \mid X = x)$. In Stage 2, MRIV estimates $\tau(x)$ directly using a pseudo outcome $\hat{Y}_0$ as a regression target.

Given an arbitrary initial ITE estimator $\hat{\tau}_{\mathrm{init}}(x)$ and nuisance estimates $\hat{\mu}_0^Y(x)$, $\hat{\mu}_0^A(x)$, $\hat{\delta}_A(x)$, and $\hat{\pi}(x)$, we define the pseudo outcome

$$\hat{Y}_0 = \left(\frac{Z - (1 - Z)}{\hat{\delta}_A(X)} \right) \left(\frac{Y - A\,\hat{\tau}_{\mathrm{init}}(X) - \hat{\mu}_0^Y(X) + \hat{\mu}_0^A(X)\,\hat{\tau}_{\mathrm{init}}(X)}{Z\,\hat{\pi}(X) + (1 - Z)(1 - \hat{\pi}(X))} \right) + \hat{\tau}_{\mathrm{init}}(X). \tag{3}$$

Once we have obtained the pseudo outcome $\hat{Y}_0$, we regress it on X to obtain the Stage 2 MRIV estimator $\hat{\tau}_{\mathrm{MRIV}}(x)$ for $\tau(x)$. The pseudocode for MRIV is given in Algorithm 1.

Algorithm 1: MRIV

Input: data (X, Z, A, Y), initial ITE estimator $\hat{\tau}_{\text{init}}(x)$

`// Stage 1: Estimate nuisance components`

$\hat{\pi}(x) \leftarrow \hat{\mathbb{E}}[Z \mid X = x], \quad \hat{\mu}_0^Y(x) \leftarrow \hat{\mathbb{E}}[Y \mid X = x, Z = 0], \quad \hat{\mu}_0^A(x) \leftarrow \hat{\mathbb{E}}[A \mid X = x, Z = 0]$

$\hat{\delta}_A(x) \leftarrow \hat{\mathbb{E}}[A \mid X = x, Z = 1] - \hat{\mathbb{E}}[A \mid X = x, Z = 0]$

`// Stage 2: Pseudo outcome regression`

$\hat{Y}_0 \leftarrow \left(\frac{Z-(1-Z)}{\hat{\delta}_A(X)}\right)\left(\frac{Y - A\,\hat{\tau}_{\text{init}}(X) - \hat{\mu}_0^Y(X) + \hat{\mu}_0^A(X)\,\hat{\tau}_{\text{init}}(X)}{Z\,\hat{\pi}(X) + (1-Z)(1-\hat{\pi}(X))}\right) + \hat{\tau}_{\text{init}}(X)$

$\hat{\tau}_{\text{MRIV}}(x) \leftarrow \hat{\mathbb{E}}[\hat{Y}_0 \mid X = x]$

The pseudo outcome $\hat{Y}_0$ in Eq. (3) is a multiple robust parameterization of the (uncentered) efficient influence curve for the average treatment effect $\mathbb{E}_X[\tau(X)]$ (see the derivation in [34]). Hence, MRIV can be interpreted as a way to remove plug-in bias from $\hat{\tau}_{\text{W}}(x)$ via influence functions [11]. Using this observation, we derive the multiple robustness property of $\hat{\tau}_{\text{MRIV}}(x)$ in Theorem 1.

Theorem 1 (Multiple robustness property)**.** *Let $\hat{\mu}_0^Y(x)$, $\hat{\mu}_0^A(x)$, $\hat{\delta}_A(x)$, $\hat{\pi}(x)$, and $\hat{\tau}_{\text{init}}(x)$ denote estimators of $\mu_0^Y(x)$, $\mu_0^A(x)$, $\delta_A(x)$, $\pi(x)$, and $\tau(x)$, respectively. Then, for all $x \in \mathcal{X}$, it holds that $\mathbb{E}[\hat{Y}_0 \mid X = x] = \tau(x)$,if least one of the following conditions is satisfied: (1) $\hat{\mu}_0^Y = \mu_0^Y$, $\hat{\mu}_0^A = \mu_0^A$, $\hat{\delta}_A = \delta_A$; or (2) $\hat{\pi} = \pi$ and $\hat{\delta}_A = \delta_A$; or (3) $\hat{\pi} = \pi$ and $\hat{\tau}_{\text{init}} = \tau$.*

Theorem 1 implies that $\hat{\tau}_{\text{MRIV}}(x)$ is consistent for $\tau(x)$ even if several nuisance components are misspecified. It is consistent even when the initial ITE estimator $\hat{\tau}_{\text{init}}$ is misspecified as long as the propensity estimator $\hat{\pi}(x)$ and $\hat{\delta}_A(x)$ are correctly specified.

Our MRIV is directly applicable to RCTs with non-compliance: Then, the treatment assignment is randomized and the propensity score $\pi(x)$ is known. Our MRIV framework can be thus adopted by plugging in the known $\pi(x)$ into the pseudo outcome in Eq. (3). Moreover, $\hat{\tau}_{\text{MRIV}}(x)$ is already consistent if either $\hat{\tau}_{\text{init}}$(x) or $\hat{\delta}_A(x)$ are.

4.2 Theoretical analysis

In the following, we derive the asymptotic convergence rate of MRIV under smoothness assumptions. For this, we define s-smooth functions as functions contained in the Hölder class $\mathcal{H}(s)$, associated with Stone's minimax rate [28] of $n^{-2s/(2s+p)}$, where p is the dimension of $\mathcal{X}$.

Assumption 3 (Smoothness)**.** We assume that (1) the nuisance components $\mu_i^Y(\cdot)$ are α-smooth, $\mu_i^A(\cdot)$ and $\delta_A(\cdot)$ are β-smooth, and $\pi(\cdot)$ is δ-smooth; (2) all nuisance components are estimated with their respective minimax rate of $n^{\frac{-2k}{2k+p}}$, where $k \in \{\alpha, \beta, \delta\}$; and (3) the oracle ITE $\tau(\cdot)$ is γ-smooth and the initial ITE estimator $\hat{\tau}_{\text{init}}$ converges with rate $r_\tau(n)$.

Assumption 3 for smoothness provides us with a way to quantify the difficulty of the underlying nonparametric regression problems. Similar assumptions have been imposed for asymptotic analysis of previous ITE estimators in [16, 12]. They can be replaced with other assumptions such as assumptions on the level of sparsity of the ITE components. We also provide an asymptotic analysis under sparsity assumptions (see Appendix B).

We additionally impose the following boundedness assumptions on the the underlying data generating process and estimators.

Assumption 4 (Boundedness)**.** We assume that there exist constants $C, \rho, \widetilde{\rho}, \epsilon, K > 0$ such that for all $x \in \mathcal{X}$ it holds that: (1) $|\mu_i^Y(x)| \leq C$; (2) $|\delta_A(x)| = |\mu_1^A(x) - \mu_0^A(x)| \geq \rho$ and $|\hat{\delta}_A(x)| \geq \widetilde{\rho}$; (3) $\epsilon \leq \hat{\pi}(x) \leq 1 - \epsilon$; and (4) $|\hat{\tau}_{\text{init}}(x)| \leq K$.

Assumptions 4.1, 4.3, and 4.4 are standard and in line with previous works on theoretical analyses of ITE estimators [12, 16]. Assumption 4.2 ensures that both the oracle ITE and the estimator are bounded. Violations of Assumption 4.2 may occur when working with so-called "weak" instruments, which are IVs that are only weakly correlated with the treatment. Using IV methods with weak instruments should generally be avoided [19]. However, in many applications such as RCTs with non-compliance, weak instruments are unlikely to occur as patients' decisions to follow the treatment are generally correlated with the initial treatment assignments.

We state now our main theoretical result: an upper bound on the oracle risk of the MRIV estimator. To derive our bound, we leverage sample splitting [16]. Sample splitting has been used initially in [16]

to analyze the DR-learner for ITE estimation under unconfoundedness, and has later been adapted to several other meta learners [12], yet not for IV methods. Note that we use sample splitting only for our theoretical analysis. In practice, all models may be fitted on the same dataset or cross-fitting may be used in small sample regimes [9].

Theorem 2 (Oracle upper bound under sample splitting)**.** *Let $\mathcal{D}_\ell$ for $\ell \in \{1,2,3\}$ be independent samples of size n. Let $\hat{\tau}_{init}(x)$, $\hat{\mu}_0^Y(x)$, and $\hat{\mu}_0^A(x)$ be trained on $\mathcal{D}_1$, and let $\hat{\delta}_A(x)$ and $\hat{\pi}(x)$ be trained on $\mathcal{D}_2$. We denote $\hat{Y}_0$ as the pseudo outcome from Eq.* (3) *and Y_0 as the corresponding oracle. Let $\hat{\tau}_{\mathrm{MRIV}}(x) = \hat{\mathbb{E}}_n[\hat{Y}_0 \mid X = x]$ and $\widetilde{\tau}_{\mathrm{MRIV}}(x) = \hat{\mathbb{E}}_n[Y_0 \mid X = x]$ denote the (oracle) pseudo outcome regression on $\mathcal{D}_3$ for some generic estimator $\hat{\mathbb{E}}_n[\cdot \mid X = x]$ of $\mathbb{E}[\cdot \mid X = x]$.*

If the second-stage estimator $\hat{\mathbb{E}}_n$ yields the minimax rate $n^{-\frac{2\gamma}{2\gamma+p}}$ and satisfies the mild assumptions of Theorem 1 in [16] (see Appendix A), the oracle risk is upper bounded by

$$\mathbb{E}\left[\left(\hat{\tau}_{\mathrm{MRIV}}(x) - \tau(x)\right)^2\right] \lesssim n^{\frac{-2\gamma}{2\gamma+p}} + r_\tau(n)\left(n^{\frac{-2\beta}{2\beta+p}} + n^{\frac{-2\delta}{2\delta+p}}\right) + n^{-2\left(\frac{\alpha}{2\alpha+p}+\frac{\delta}{2\delta+p}\right)} + n^{-2\left(\frac{\beta}{2\beta+p}+\frac{\delta}{2\delta+p}\right)}.$$

Proof. See Appendix A. □

Recall that the first summand of the lower bound in Eq. (2) is the minimax rate for the oracle ITE $\tau(x)$ which cannot be improved upon. Hence, for a fast convergence rate of $\hat{\tau}_{\mathrm{MRIV}}(x)$, it is sufficient if either: (1) $r_\tau(n)$ decreases fast and δ is large; (2) $r_\tau(n)$ decreases fast and α and β are large; or (3) all α, β, and δ are large. This is in line with the multiple robustness property of MRIV and means that MRIV achieves a fast rate of convergence even if the initial estimator or several nuisance estimators converge slowly.

From the bound in Eq. (2), it follows that $\hat{\tau}_{\mathrm{MRIV}}(x)$ improves on the convergence rate of the initial ITE estimator $\hat{\tau}_{\mathrm{init}}(x)$ if its rate $r_\tau(n)$ is lower bounded by

$$r_\tau(n) \gtrsim n^{\frac{-2\gamma}{2\gamma+p}} + n^{-2\left(\frac{\alpha}{2\alpha+p}+\frac{\delta}{2\delta+p}\right)} + n^{-2\left(\frac{\beta}{2\beta+p}+\frac{\delta}{2\delta+p}\right)}. \tag{4}$$

Hence, our MRIV estimator is more likely to improve on the initial estimator for large α, β, and δ, i.e., if the nuisance components are smooth. Note that it is sufficient if either (1) *only* the propensity score $\pi(x)$ is relatively smooth (large δ) *or* (2) that *all* other nuisance components are (large α *and* β). In fact, this is widely fulfilled in practice. For example, the former is fulfilled for RCTs with non-compliance, where $\pi(x)$ is often some known, fixed number $p \in (0,1)$. Hence, for RCTs with non-compliance, MRIV should (at least asymptotically) improve the performance of most estimators.

4.3 MRIV vs. Wald estimator

In the following, we compare $\hat{\tau}_{\mathrm{MRIV}}(x)$ to the Wald estimator $\hat{\tau}_{\mathrm{W}}(x)$. First, we derive corresponding upper bound under smoothness.

Theorem 3 (Wald oracle upper bound)**.** *Given estimators $\hat{\mu}_i^Y(x)$ and $\hat{\mu}_i^A(x)$. Let $\hat{\delta}_A(x) = \hat{\mu}_1^A(x) - \hat{\mu}_0^A(x)$ satisfy Assumption 4. Then, the oracle risk of the Wald estimator $\hat{\tau}_W(x)$ is bounded by*

$$\mathbb{E}\left[\left(\hat{\tau}_{\mathrm{W}}(x) - \tau(x)\right)^2\right] \lesssim n^{-\frac{2\alpha}{2\alpha+p}} + n^{-\frac{2\beta}{2\beta+p}}. \tag{5}$$

Proof. See Appendix A. □

We now consider the MRIV estimator $\hat{\tau}_{\mathrm{MRIV}}(x)$ with $\hat{\tau}_{\mathrm{init}} = \hat{\tau}_{\mathrm{W}}(x)$, i.e., initialized with the Wald estimator (under sample splitting). Plugging the Wald rate from Eq. (5) into the Eq. (2) yields

$$\mathbb{E}\left[\left(\hat{\tau}_{\mathrm{MRIV}}(x) - \tau(x)\right)^2\right] \lesssim n^{\frac{-2\gamma}{2\gamma+p}} + n^{\frac{-4\beta}{2\beta+p}} + n^{-2\left(\frac{\alpha}{2\alpha+p}+\frac{\beta}{2\beta+p}\right)} + n^{-2\left(\frac{\delta}{2\delta+p}+\frac{\alpha}{2\alpha+p}\right)} + n^{-2\left(\frac{\delta}{2\delta+p}+\frac{\beta}{2\beta+p}\right)}. \tag{6}$$

For $\alpha = \beta = \delta$, the rates of $\hat{\tau}_{\mathrm{MRIV}}(x)$ and $\hat{\tau}_{\mathrm{W}}(x)$ reduce to

$$\mathbb{E}\left[\left(\hat{\tau}_{\mathrm{MRIV}}(x) - \tau(x)\right)^2\right] \lesssim n^{\frac{-2\gamma}{2\gamma+p}} + n^{\frac{-4\alpha}{2\alpha+p}} \quad \text{and} \quad \mathbb{E}\left[\left(\hat{\tau}_{\mathrm{W}}(x) - \tau(x)\right)^2\right] \lesssim n^{\frac{-2\alpha}{2\alpha+p}}. \tag{7}$$

Hence, $\hat{\tau}_{\mathrm{MRIV}}(x)$ outperforms $\hat{\tau}_{\mathrm{W}}(x)$ asymptotically for $\gamma > \alpha$, i.e., when the ITE $\tau(x)$ is smoother than its components, which is usually the case in practice [18]. For $\gamma = \alpha$, the rates of both estimators coincide. Hence, we should expect MRIV to improve on the Wald estimator in real-world settings with sufficiently large sample size.

4.4 MRIV-Net

Based on our MRIV framwork, we develop a tailored deep neural network called MRIV-Net for ITE estimation using IVs. Our MRIV-Net produces both an initial ITE estimator $\hat{\tau}_{\text{init}}(x)$ and nuisance estimators $\hat{\mu}_0^Y(x)$, $\hat{\mu}_0^A(x)$, $\hat{\delta}_A(x)$, and $\hat{\pi}(x)$.

For MRIV-Net, we choose deep neural networks for the nuisance components due to their predictive power and their ability to learn complex shared representations for several nuisance components. Sharing representations between nuisance components has been exploited previously for ITE estimation, yet only under unconfoundedness [26, 12]. Building shared representations is more efficient in finite sample regimes than estimating all nuisance components separately as they usually share some common structure.

In MRIV-Net, not all nuisance components should share a representation. Recall that, in Theorem 2, we assumed that (1) $\hat{\tau}_{\text{init}}(x)$, $\hat{\mu}_0^Y(x)$, and $\hat{\mu}_0^A(x)$; and (2) $\hat{\delta}_A(x)$ and $\hat{\pi}(x)$ are trained on two independent samples in order to derive the upper bound on the oracle risk. Hence, we propose to build two separate representations Φ_1 and Φ_2, so that (i) Φ_1 is used to learn $\hat{\tau}_{\text{init}}(x)$, $\hat{\mu}_0^Y(x)$, and $\hat{\mu}_0^A(x)$, and (ii) Φ_2 is used to learn $\hat{\delta}_A(x)$ and $\hat{\pi}(x)$. This ensures that the nuisance estimators (1) share minimal information with nuisance estimators (2) even though they are estimated on the same data. Intuitively, this should lead to a faster decay of the oracle upper bound (cf. [12]).

The architecture of MRIV-Net is shown in Fig. 2. MRIV-Net takes the observed covariates X as input to build the two representations Φ_1 and Φ_2. The first representation Φ_1 is used to output estimates $\hat{\mu}_1^Y(x)$, $\hat{\mu}_0^Y(x)$, $\hat{\mu}_1^A(x)$, and $\hat{\mu}_0^A(x)$ of the ITE components. The second representation Φ_2 is used to output estimates $\widetilde{\mu}_1^A(x)$, $\widetilde{\mu}_0^A(x)$, and $\hat{\pi}(x)$. MRIV-Net is trained by minimizing an overall loss

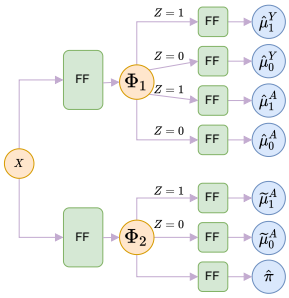

Figure 2: Architecture of MRIV-Net.

$$\mathcal{L}(\theta) = \sum_{i=1}^{n} \left[\left(\hat{\mu}_{z_i}^Y(x_i) - y_i\right)^2 + \text{BCE}\left(\hat{\mu}_{z_i}^A(x_i), a_i\right) + \text{BCE}\left(\widetilde{\mu}_{z_i}^A(x_i), a_i\right) + \text{BCE}\left(\hat{\pi}(x_i), z_i\right) \right], \tag{8}$$

where θ denotes the neural network parameters and BCE is the binary cross entropy loss. After training MRIV-Net, we obtain the $\hat{\tau}_{\text{init}}(x) = \frac{\hat{\mu}_1^Y(x) - \hat{\mu}_0^Y(x)}{\hat{\mu}_1^A(x) - \hat{\mu}_0^A(x)}$ and obtain the nuisance estimators $\hat{\mu}_0^Y(x)$, $\hat{\mu}_0^A(x)$, $\hat{\delta}_A(x) = \widetilde{\mu}_1^A(x) - \widetilde{\mu}_0^A(x)$ and $\hat{\pi}(x)$. Then, we perform, we perform the pseudo regression (Stage 2) of MRIV to obtain $\hat{\tau}_{\text{MRIV}}(x)$.

Implementation: We use PyTorch Lightning for our implementation and train MRIV-Net with the Adam optimizer [17]. Details on the network architecture and hyperparameter tuning are in Appendix F. We perform both the training of MRIV-Net and the pseudo outcome regression on the full training data. Needless to say, MRIV-Net can be easily adopted for sample splitting or cross-fitting procedures as in [9], namely, by learning separate networks for each representation Φ_1 and Φ_2. However, in our experiments, we do not use sample splitting or cross-fitting, as this can affect the performance in finite sample regimes. Of note, our choice is consistent with previous work [12].

5 Computational experiments

5.1 Simulated data

In causal inference literature, it is common practice to use simulated data for performance evaluations [7, 12, 14]. Simulated data offers the crucial benefit that it provides ground-truth information on the counterfactual outcomes and thus allows for direct benchmarking against the oracle ITE.

Data generation: We generate simulated data by sampling the oracle ITE $\tau(x)$ and the nuisance components $\mu_i^Y(x)$, $\mu_i^A(x)$, and $\pi(x)$ from Gaussian process priors. Using Gaussian processes has the following advantages: (1) It allows for a fair method comparison, as there is no need to explicitly specify the nuisance components, which could lead to unwanted inductive biases favoring a specific

method; (2) the sampled nuisance components are non-linear and thus resemble real-world scenarios where machine learning methods would be applied; and, (3) by sampling from the prior induced by the Matérn kernel [23], we can control the smoothness of the nuisance components, which allows us to confirm our theoretical results from Sec. 4.2. For a detailed description of our data generating process, we refer to Appendix C.

Baselines: We compare our MRIV-Net with the following state-of-the-art baselines: (1) ITE methods for unconfoundedness: **TARNet** [26] and TARNet combined with the **DR learner** [16]; (2) general IV methods: **2SLS** [36], kernel IV (**KIV**) [27], **DFIV** [37], **DeepIV** [14], **DeepGMM** [6], **DMLIV** [29], and DMLIV combined with **DRIV** (as described in [29]); (3) the (plug-in) Wald estimator using **linear models** and Bayesian additive regression trees (**BART**) [10]. Of note, the DR learner assumes unconfoundedness, which is why we only combine it TARNet in our experiments. Implementation details regarding baselines and nuisance parameter estimation are in Appendix E.

Performance evaluation: For all experiments, we use a 80/20 split as training/test set. We calcalute the root mean squared errors (RMSE) between the ITE estimates and the oracle ITE on the test set. We report the mean RMSE and the standard deviation over five data sets generated from random seeds.

Table 1: Performance comparison: our MRIV-Net vs. existing baselines.

Method	$n = 3000$	$n = 5000$	$n = 8000$
(1) STANDARD ITE			
TARNet [26]	0.76 ± 0.14	0.70 ± 0.12	0.69 ± 0.17
TARNet + DR [26, 16]	0.78 ± 0.10	0.66 ± 0.09	0.70 ± 0.10
(2) GENERAL IV			
2SLS [35]	1.22 ± 0.23	0.79 ± 0.37	1.12 ± 0.29
KIV [27]	1.54 ± 0.53	1.18 ± 1.14	3.80 ± 4.71
DFIV [37]	0.43 ± 0.11	0.40 ± 0.21	0.46 ± 0.54
DeepIV [14]	0.96 ± 0.30	0.28 ± 0.09	0.23 ± 0.04
DeepGMM [6]	0.95 ± 0.38	0.37 ± 0.09	0.42 ± 0.14
DMLIV [29]	1.92 ± 0.71	0.92 ± 0.41	1.14 ± 0.24
DMLIV + DRIV [29]	0.41 ± 0.12	0.22 ± 0.04	0.21 ± 0.06
(3) WALD ESTIMATOR [32]			
Linear	1.06 ± 0.63	0.62 ± 0.22	0.81 ± 0.34
BART	0.95 ± 0.30	0.63 ± 0.33	0.88 ± 0.28
MRIV-Net (ours)	$\mathbf{0.26 \pm 0.11}$	$\mathbf{0.15 \pm 0.03}$	$\mathbf{0.13 \pm 0.03}$

Reported: RMSE for base methods (mean $\pm$ standard deviation). Lower = better (best in bold)

Results: Table 1 shows the results for all baselines. Here, the DR learner does not improve the performance of TARNet, which is reasonable as both the DR learner and TARNet assume unconfoundedness and are thus biased in our setting. Our MRIV-Net outperforms all baselines. Our MRIV-Net also achieves a smaller standard deviation. For additional results, we refer to Appendix G.

We further compare the performance of two different meta-learner frameworks – DRIV [29] and our MRIV– across different base methods. The nuisance parameters are estimated using feed forward neural networks (DRIV) or TARNets with either binary or continuous outputs (MRIV). The results are in Table 2. Our MRIV improves over the variant without any meta-learner framework across all base methods (both in terms of RMSE and standard deviation). Furthermore, MRIV is clearly superior over DRIV. This demonstrates the effectiveness of our MRIV across different base methods (note: MRIV with an arbitrary base model is typically superior to DRIV with our custom network from above). MRIV-Net is overall best.

Table 2: Base model with different meta-learners (i.e., none, DRIV, and our MRIV).

Base methods \ Meta-learners	$n = 3000$			$n = 5000$			$n = 8000$		
	None	DRIV	MRIV (ours)	None	DRIV	MRIV (ours)	None	DRIV	MRIV (ours)
(1) STANDARD ITE									
TARNet [26]	0.76 ± 0.14	$\mathbf{0.31 \pm 0.05}$	0.34 ± 0.13	0.70 ± 0.12	$\mathbf{0.17 \pm 0.06}$	$\mathbf{0.17 \pm 0.05}$	0.69 ± 0.17	0.21 ± 0.04	$\mathbf{0.16 \pm 0.04}$
(2) GENERAL IV									
2SLS [35]	1.22 ± 0.23	0.40 ± 0.11	$\mathbf{0.31 \pm 0.08}$	0.79 ± 0.37	$\mathbf{0.17 \pm 0.09}$	0.19 ± 0.05	1.12 ± 0.29	0.21 ± 0.05	$\mathbf{0.16 \pm 0.02}$
KIV [27]	1.54 ± 0.53	0.40 ± 0.10	$\mathbf{0.39 \pm 0.11}$	1.18 ± 1.14	0.20 ± 0.08	$\mathbf{0.17 \pm 0.06}$	3.80 ± 4.71	0.31 ± 0.18	$\mathbf{0.28 \pm 0.19}$
DFIV [37]	0.43 ± 0.11	$\mathbf{0.26 \pm 0.05}$	0.27 ± 0.07	0.40 ± 0.21	0.18 ± 0.09	$\mathbf{0.16 \pm 0.04}$	0.46 ± 0.54	0.21 ± 0.06	$\mathbf{0.18 \pm 0.05}$
DeepIV [14]	0.96 ± 0.30	0.27 ± 0.03	$\mathbf{0.26 \pm 0.05}$	0.28 ± 0.09	$\mathbf{0.18 \pm 0.08}$	$\mathbf{0.18 \pm 0.05}$	0.23 ± 0.04	0.21 ± 0.03	$\mathbf{0.16 \pm 0.03}$
DeepGMM [6]	0.95 ± 0.38	0.40 ± 0.15	$\mathbf{0.36 \pm 0.13}$	0.37 ± 0.09	0.24 ± 0.12	$\mathbf{0.16 \pm 0.05}$	0.42 ± 0.14	0.21 ± 0.03	$\mathbf{0.17 \pm 0.03}$
DMLIV [29]	1.92 ± 0.71	0.41 ± 0.12	$\mathbf{0.37 \pm 0.11}$	0.92 ± 0.41	0.22 ± 0.05	$\mathbf{0.16 \pm 0.05}$	1.14 ± 0.24	0.21 ± 0.06	$\mathbf{0.18 \pm 0.05}$
(3) WALD ESTIMATOR [32]									
Linear	1.06 ± 0.63	0.42 ± 0.15	$\mathbf{0.38 \pm 0.14}$	0.62 ± 0.22	$\mathbf{0.19 \pm 0.09}$	0.25 ± 0.09	0.81 ± 0.34	0.19 ± 0.09	$\mathbf{0.18 \pm 0.04}$
BART	0.95 ± 0.30	0.48 ± 0.14	$\mathbf{0.46 \pm 0.12}$	0.63 ± 0.33	0.26 ± 0.13	$\mathbf{0.20 \pm 0.07}$	0.88 ± 0.28	0.31 ± 0.08	$\mathbf{0.29 \pm 0.04}$
MRIV-Net\w network only (ours)	0.39 ± 0.13	0.35 ± 0.12	$\mathbf{0.26 \pm 0.11}$	0.31 ± 0.04	0.19 ± 0.13	$\mathbf{0.15 \pm 0.03}$	0.26 ± 0.06	0.18 ± 0.08	$\mathbf{0.13 \pm 0.03}$

Reported: RMSE (mean $\pm$ standard deviation). Lower = better (best improvement over none meta-learner in bold)

Ablation study: Table 3 compares different variants of our MRIV-Net. These are: (1) MRIV but network only; (2) MRIV-Net with a single representation for all nuisance estimators; and (3) our MRIV-Net from above. We observe that MRIV-Net is best. This justifies our proposed network architecture for MRIV-Net. Hence, combing the re-

Table 3: Ablation study.

Method	$n = 3000$	$n = 5000$	$n = 8000$
MRIV-Net\w network only	0.39 ± 0.13	0.31 ± 0.04	0.26 ± 0.06
MRIV-Net\w single repr.	0.28 ± 0.12	0.21 ± 0.04	0.32 ± 0.10
MRIV-Net (ours)	$\mathbf{0.26 \pm 0.11}$	$\mathbf{0.15 \pm 0.03}$	$\mathbf{0.13 \pm 0.03}$

Reported: RMSE (mean $\pm$ standard deviation). Lower = better (best in bold)

sult from above, our performance gain must be attributed to both our framework and the architecture of our deep neural network.

Robustness checks for unobserved confounding and smoothness: Here, we demonstrate the importance of handling unobserved confounding (as we do in our MRIV framework). For this, Fig. 3 plots the results for our MRIV-Net vs. standard ITE without customization for confounding (i.e., TARNet with and without the DR learner) over over different levels of unobserved confounding. The RMSE of both TARNet variants increase almost linearly with increasing confounding. In contrast, the RMSE of our MRIV-Net only marginally. Even for low confounding regimes, our MRIV-Net performs competitively.

Fig. 4 varies the smoothness level. This is given by α of $\mu_i^Y(\cdot)$ (controlled by the Matérn kernel prior). Here, the performance decreases for the baselines, i.e., DeepIV and our network without MRIV framework. In contrast, the peformance of our MRIV-Net remains robust and outperforms the baselines. This confirms our theoretical results from above. It thus indicates that our MRIV framework works best when the oracle ITE $\tau(x)$ is smoother than the nuisance parameters $\mu_i^Y(x)$.

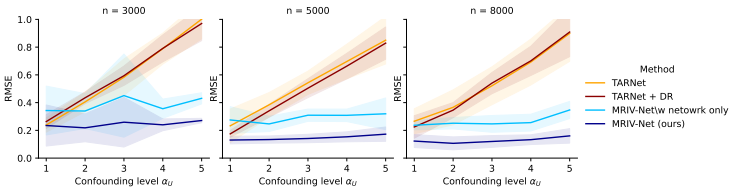

Figure 3: Results over different levels of confounding α_U. Shaded area shows standard deviation.

5.2 Case study with real-world data

Setting: We demonstrate effectiveness of our framework using a case study with real-world, medical data. Here, we use medical data from the so-called *Oregon health insurance experiment* (OHIE) [13]. It provides data for an RCT with non-compliance: In 2008, ∼30,000 low-income, uninsured adults in Oregon were offered participation in a health insurance program by a lottery. Individuals whose names were drawn could decide to sign up for health insurance. After a period of 12 months, in-person interviews took place to evaluate the health condition of the respective participant. In our analysis, the lottery assignment is the instrument Z, the decision to sign up for health insurance is treatment A, and an overall health score is the outcome Y. We also include five covariates X (age, gender, etc.). For details, we refer to Appendix D.

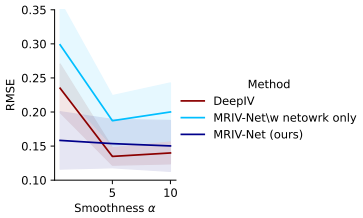

Figure 4: Results over different levels of smoothness α of $\mu_i^Y(\cdot)$, sample size $n = 8000$. Larger α = smoother. Shaded areas show standard deviation.

We estimate the ITE for both genders, fixing the other covariates while varying the age. We repeat the same procedure for our neural network architecture without the MRIV-Net framework and TARNet. The results are in Fig. 5.

Results: Our MRIV-Net estimates larger ITEs for an older age. In contrast, TARNet does not estimate positive ITEs even for an older age. Even though we cannot evaluate the estimation quality on real-world data, our estimates seem reasonable in light of the medical literature: the benefit of health insurance should increase with older age. This showcases that TARNet may suffer from bias induced by unobserved confounders. We also report the results for DRIV with DMLIV as base method, and observe that in contrast to MRIV-Net, the corresponding ITE does not vary much between ages. Interestingly, both our MRIV-Net estimate a somewhat smaller ITE for middle ages (around 30–50 yrs). One explanation might be that individual in this age group are more likely to have stable jobs and, thus, are also more likely to be able to afford medical care, decreasing the direct effect of health insurance on individuals health. In sum, the findings from our case study are of direct relevance for decision-makers in public health [15], and highlight the practical value of our framework.

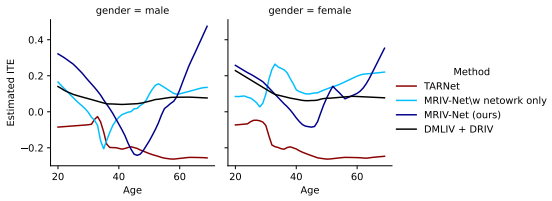

Figure 5: Results on real-world medical data.

References

[1] Ahmed M. Alaa and Mihaela van der Schaar. "Bayesian inference of individualized treatment effects using multi-task Gaussian processes". In: *NeurIPS*. 2017.

[2] Joshua D. Angrist. "Lifetime earnings and the vietnam era draft lotter: Evidence from social security administrative records". In: *The American Economic Review* 80.3 (1990), pp. 313–336.

[3] Joshua D. Angrist, Guido W. Imbens, and Donald B. Rubin. "Identification of causal effects using instrumental variables". In: *Journal of the American Statistical Association* 91.434 (1996), pp. 444–455.

[4] Joshua D. Angrist and Alan B. Krueger. "Does compulsory school attendance affect schooling and earnings?" In: *The Quarterly Journal of Economics* 106.4 (1991), pp. 979–1014.

[5] Falco J. Bargagli-Stoffi, Kristof de Witte, and Giorgio Gnecco. "Heterogeneous causal effects with imperfect compliance: A Bayesian machine learning approach". In: *Annals of Applied Statistics* (2021).

[6] Andrew Bennett, Nathan Kallus, and Tobias Schnabel. "Deep generalized method of moments for instrumental variable analysis". In: *NeurIPS*. 2019.

[7] Ioana Bica et al. "Estimating counterfactual treatment outcomes over time through adversarially balanced representations". In: *ICLR*. 2020.

[8] Howard S. Bloom et al. "The benefits and costs of JTPA title II-A programs: Key Findings from the National Job Training Partnership Act Study". In: *Journal of Human Resources* 32.32 (1997), pp. 549–586.

[9] Victor Chernozhukov et al. "Double/debiased machine learning for treatment and structural parameters". In: *The Econometrics Journal* 21.1 (2018), pp. C1–C68.

[10] Hugh A. Chipman, Edward I. George, and Robert E. McCulloch. "BART: Bayesian additive regression trees". In: *The Annals of Applied Statistics* 4.1 (2010), pp. 266–298.

[11] Alicia Curth, Ahmed M. Alaa, and Mihaela van der Schaar. "Estimating structural target functions using machine learning and influence functions". In: *arXiv preprint* (2020).

[12] Alicia Curth and Mihaela van der Schaar. "Nonparametric estimation of heterogeneous treatment effects: From theory to learning Algorithms". In: *AISTATS*. 2021.

[13] Amy Finkelstein et al. "The oregon health insurance experiment: Evidence from the first year". In: *The Quarterly Journal of Economics* 127.3 (2012), pp. 1057–1106.

[14] Jason Hartford et al. "Deep IV: A flexible approach for counterfactual prediction". In: *ICML*. 2017.

[15] Guido W. Imbens and Joshua D. Angrist. "Identification and estimation of local average treatment effects". In: *Econometrica* 62.2 (1994), pp. 467–475.

[16] Edward H. Kennedy. "Optimal doubly robust estimation of heterogeneous causal effects". In: *arXiv preprint* (2020).

[17] Diederik P. Kingma and Jimmy Ba. "Adam: A method for stochastic optimization". In: *ICLR*. 2015.

[18] Sören R. Künzel et al. "Metalearners for estimating heterogeneous treatment effects using machine learning". In: *Proceedings of the National Academy of Sciences (PNAS)* 116.10 (2019), pp. 4156–4165.

[19] Chunxiao Li, Cynthia Rudin, and Typer H. McCormick. "Rethinking nonlinear instrumental variable models through prediction validity". In: *Journal of Machine Learning Research* 23 (2022), pp. 1–55.

[20] Bryan Lim, Ahmed M. Alaa, and Mihaela van der Schaar. "Forecasting treatment responses over time using recurrent marginal structural networks". In: *NeurIPS*. 2018.

[21] Whitney K. Newey and James L. Powell. "Instrumental variable estimation of nonparametric models". In: *Econometrica* 71.5 (2003), pp. 1565–1578.

[22] Judea Pearl. *Causality*. New York City: Cambridge University Press, 2009.

[23] Carl Edward Rasmussen and Christopher K. I. Williams. *Gaussian processes for machine learning*. 3. print. Adaptive computation and machine learning. Cambridge, Mass.: MIT Press, 2008.

[24] James M. Robins, Miguel A. Hernán, and Babette Brumback. "Marginal structural models and causal inference in epidemiology". In: *Epidemiology* 11.5 (2000), pp. 550–560.

[25] Donald B. Rubin. "Estimating causal effects of treatments in randomized and nonrandomized studies". In: *Journal of Educational Psychology* 66.5 (1974), pp. 688–701.

[26] Uri Shalit, Fredrik D. Johansson, and David Sontag. "Estimating individual treatment effect: Generalization bounds and algorithms". In: *ICML*. 2017.

[27] Rahul Singh, Maneesh Sahani, and Arthur Gretton. "Kernel instrumental variable regression". In: *NeurIPS*. 2019.

[28] Charles J. Stone. "Optimal rates of convergence for nonparametric estimators". In: *Annals of Statistics* 8.6 (1980).

[29] Vasilis Syrgkanis et al. "Machine learning estimation of heterogeneous treatment effects with instruments". In: *NeurIPS*. 2019.

[30] Hal R. Varian. "Causal inference in economics and marketing". In: *Proceedings of the National Academy of Sciences (PNAS)* 113.27 (2016), pp. 7310–7315.

[31] Stefan Wager and Susan Athey. "Estimation and inference of heterogeneous treatment effects using random forests". In: *Journal of the American Statistical Association* 113.523 (2018), pp. 1228–1242.

[32] Abraham Wald. "The fitting of straight lines if both variables are subject to error". In: *Annals of Mathematical Statistics* 11.3 (1940), pp. 284–300.

[33] Guihua Wang, Jun Li, and Wallace J. Hopp. "An instrumental variable forest approach for detecting heterogeneous treatment effects in observational studies". In: *Management Science* (2021).

[34] Linbo Wang and Eric J. Tchetgen Tchetgen. "Bounded, efficient and multiply robust estimation of average treatment effects using instrumental variables". In: *Journal of the Royal Statistical Society: Series B* 80.3 (2018), pp. 531–550.

[35] Jeffrey M. Wooldridge. *Introductory Econometrics: A modern approach*. Routledge, 2013.

[36] Phillip G. Wright. *The tariff on animal and vegitable oils*. New York: Macmillan, 1928.

[37] Liyuan Xu et al. "Learning deep features in instrumental variable regression". In: *ICLR*. 2021.

[38] Azam M. Yazdani and Eric Boerwinkle. "Causal inference in the age of decision medicine". In: *Journal of Data Mining in Genomics & Proteomics* 6.1 (2015).

[39] Jinsung Yoon, James Jordon, and Mihaela van der Schaar. "GANITE: Estimation of individualized treatment effects using generative adversarial nets". In: *ICLR*. 2018.

[40] Yao Zhang, Alexis Bellot, and Mihaela van der Schaar. "Learning overlapping representations for the estimation of individualized treatment effects". In: *AISTATS*. 2020.

Checklist

1. For all authors...
   (a) Do the main claims made in the abstract and introduction accurately reflect the paper's contributions and scope? [Yes] See here
   (b) Did you describe the limitations of your work? [Yes] See Sec. 5.2.
   (c) Did you discuss any potential negative societal impacts of your work? [No]
   (d) Have you read the ethics review guidelines and ensured that your paper conforms to them? [Yes]
2. If you are including theoretical results...
   (a) Did you state the full set of assumptions of all theoretical results? [Yes] See Sec. 4.2 and Appendix A..
   (b) Did you include complete proofs of all theoretical results? [Yes] See Appendix A.
3. If you ran experiments...
   (a) Did you include the code, data, and instructions needed to reproduce the main experimental results (either in the supplemental material or as a URL)? [Yes] (both).
   (b) Did you specify all the training details (e.g., data splits, hyperparameters, how they were chosen)? [Yes] See Appendix E and F
   (c) Did you report error bars (e.g., with respect to the random seed after running experiments multiple times)? [Yes] One standard deviation using 5 random seeds, see Sec. 5.

(d) Did you include the total amount of compute and the type of resources used (e.g., type of GPUs, internal cluster, or cloud provider)? [Yes] See Appendix E.

4. If you are using existing assets (e.g., code, data, models) or curating/releasing new assets...
   (a) If your work uses existing assets, did you cite the creators? [Yes] See Sec. 5.2.
   (b) Did you mention the license of the assets? [No] No license is provided on the OHIE website.
   (c) Did you include any new assets either in the supplemental material or as a URL? [Yes] See Appendix D.
   (d) Did you discuss whether and how consent was obtained from people whose data you're using/curating? [No] We only used simulated and publicly available data.
   (e) Did you discuss whether the data you are using/curating contains personally identifiable information or offensive content? [No] The OHIE data is publicly available, personally identifiable information are censored.
5. If you used crowdsourcing or conducted research with human subjects...
   (a) Did you include the full text of instructions given to participants and screenshots, if applicable? [No] Not applicable.
   (b) Did you describe any potential participant risks, with links to Institutional Review Board (IRB) approvals, if applicable? [No] Not applicable.
   (c) Did you include the estimated hourly wage paid to participants and the total amount spent on participant compensation? [No] Not applicable.