Meta-Learners for Partially-Identified Treatment Effects Across Multiple Environments

Jonas Schweisthal [*1,2] **Dennis Frauen** [*1,2] **Mihaela van der Schaar** [3] **Stefan Feuerriegel** [1,2]

Abstract

Estimating the conditional average treatment effect (CATE) from observational data is relevant for many applications such as personalized medicine. Here, we focus on the widespread setting where the observational data come from multiple environments, such as different hospitals, physicians, or countries. Furthermore, we allow for violations of standard causal assumptions, namely, overlap within the environments and unconfoundedness. To this end, we move away from point identification and focus on *partial identification*. Specifically, we show that current assumptions from the literature on multiple environments allow us to interpret the environment as an instrumental variable (IV). This allows us to adapt bounds from the IV literature for partial identification of CATE by leveraging treatment assignment mechanisms across environments. Then, we propose different model-agnostic learners (so-called meta-learners) to estimate the bounds that can be used in combination with arbitrary machine learning models. We further demonstrate the effectiveness of our meta-learners across various experiments using both simulated and real-world data. Finally, we discuss the applicability of our meta-learners to partial identification in instrumental variable settings, such as randomized controlled trials with non-compliance.

1. Introduction

Estimating conditional average treatment effects (CATEs) from observational data is relevant for many applications. Examples include medicine (Glass et al., 2013; Feuerriegel et al., 2024), economics (Angrist, 1990; Kuzmanovic et al., 2024), or marketing (Varian, 2016). For example, medical professionals are interested in leveraging electronic health records to personalize care by understanding the estimated CATE of treatments.

In this paper, we are interested in the CATE in a setting where we have access to observational data, that are collected from *multiple environments*. Furthermore, we allow for *violations of standard causal assumptions*, namely, overlap within the environments and unconfoundedness. Our setting is particularly relevant for medical applications for two reasons.

First, CATE estimation in practice often involves settings where the observational data come from *multiple environments* (Shi et al., 2021). Common examples in medicine are settings where patient data come from multiple hospitals, multiple physicians, or multiple countries (e.g., Huang et al., 2022). As a result, each environment is characterized by unique patient demographics (e.g., a specialized hospital may have more severe cases of disease) and/or unique treatment policies (e.g., the default treatment option may vary across countries).

Second, identification of the CATE from observational data is challenging and typically requires strong *assumptions* such as (i) *overlap (within environments)* and (ii) *unconfoundedness* (Wager & Athey, 2018). (i) Overlap ensures that each individual has a positive probability of receiving any treatment (D'Amour et al., 2021) within each environment. (ii) Unconfoundedness implies that all potential confounders that influence both the treatment assignment and the outcome of interest are contained within the observational data. Under violations of these assumptions, the CATE is generally unidentifiable and can not be estimated consistently from observational data (Pearl, 2009). Notwithstanding, *violations* are likely to occur in settings with multiple environments. (i) When patient characteristics and/or treatment policies vary across environments, then it is likely that certain patient characteristics and/or certain treatments are rarely or even never observed in some environments (e.g., certain treatments are only administered by specialized physicians). As such, overlap is often violated. (ii) Confounders such as socioeconomic status are often

[*]Equal contribution [1]LMU Munich, Germany [2]Munich Center for Machine Learning (MCML), Germany [3]University of Cambridge, UK. Correspondence to: Jonas Schweisthal <jonas.schweisthal@lmu.de>, Dennis Frauen <frauen@lmu.de>.

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

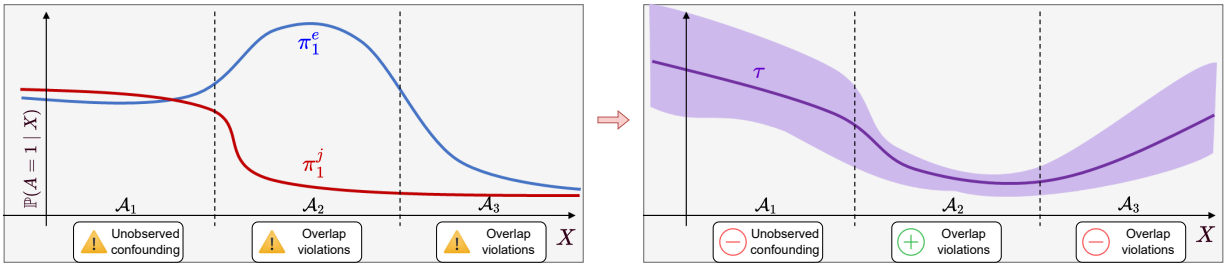

Figure 1. Intuition for our bounds. *Left:* Two propensity scores $\pi_1^e(x) = \mathbb{P}(A^e = 1 \mid X^e = x)$ and $\pi_1^j(x) = \mathbb{P}(A^j = 1 \mid X^j = x)$ corresponding to different environments e and j are plotted over observed confounders X. Large values of $\pi_1^e(x)$ and $\pi_1^j(x)$ correspond to a high probability of receiving treatment and vice versa. *Right:* CATE $\tau(x)$ together with bounds depending on violations of overlap and unconfoundedness. In region $\mathcal{A}_1$, no overlap violations occur, leading to wide bounds for the CATE τ due to potential unobserved confounding. In region $\mathcal{A}_2$, overlap violations occur at *opposite ends* across environments, leading to tight bounds for τ. In region $\mathcal{A}_3$, overlap violations occur on the same end across environments, leading to wide bounds for τ due to a lack of data for treated individuals.

not recorded in medical studies (Adler & Stead, 2015). As such, unconfoundedness is often violated. Consequently, standard CATE estimators may be biased and thus unreliable, motivating the need for methods to relax standard assumptions.

To address the challenges of the setting above, we move away from *point* identification and focus on *partial* identification. Partial identification means that one relaxes assumptions on the underlying data-generating process to estimate bounds for a causal query of interest (Jesson et al., 2021). Knowing that the bounds are above or below zero is often sufficient for consequential decision-making (Kallus et al., 2019). For example, knowledge about a positive treatment effect may be sufficient for a physician to prescribe a treatment. To the best of our knowledge, our work is unique in two ways. First, while there is rich literature on partial identification (e.g., Manski, 1990; Kilbertus et al., 2020; Padh et al., 2023), no previous works allow for observational datasets from *multiple environments*. Second, the *derivation* of bounds for specific settings is a common theme in the partial identification literature (e.g., Duarte et al., 2023), yet the effective estimation of such bounds is often not the focus. Here, a novelty of ours is that we provide flexible meta-learners that can be combined with arbitrary machine learning models for estimating such bounds in our setting.

In this paper, we make three **contributions**:[1] (1) We show that under similar assumptions as in related work (Kallus et al., 2018), the variable representing the environment can be interpreted as an instrument variable (IV). Then, we generalize previous results from the partial identification literature for binary IVs (Balke & Pearl, 1997; Swanson et al., 2018) to our setting to obtain bounds for CATE. (2) We propose novel meta-learners to *estimate* the bounds from observational data. Importantly, our meta-learners are model-agnostic and can be used in combination with any machine learning model. (3) We provide theoretical results for our meta-learners by showing consistency and double robustness properties. Finally, we confirm the effectiveness of our meta-learners by performing various experiments using both simulated and real-world data.

We now give an **intuition** behind our bounds, and, thereby, we explain that access to a discrete variable representing the environment can indeed help with partial identification of the CATE, even under violations of overlap and unobserved confounding. Fig. 1 shows two different environments, namely, hospital e and hospital j. Both have different treatment assignment mechanisms (i.e., π_1^e and π_1^j, respectively), which are plotted against some patient characteristic X (e.g., patient age). Let us highlight three regions: In $\mathcal{A}_1$ and $\mathcal{A}_3$, the treatment assignments are similar, because of which, unfortunately, we have wide bounds for the CATE. However, in $\mathcal{A}_2$, the overlap assumption is violated in both environments, but at *opposite ends*: in hospital e, middle-aged patients are almost always treated, while, in hospital j, they are rarely treated. Hence, we can overcome the limited overlap by *combining* data across both environments. Furthermore, we can deduce claims regarding the "strength" of unobserved confounding in the region $\mathcal{A}_2$, as there can *not* exist an unobserved variable with a strong influence on the treatment assignment because this would imply more observed variation in the prescribed treatments. Our intuition is that violations of overlap and unconfoundedness *exclude each other*, so that we obtain tight bounds for the CATE.

2. Related work

Our work draws from multiple streams of related literature, which we list in the following. Additional literature is in Appendix A.

Treatment effect estimation across different environments: Several works focus on treatment effect estimation

[1] Code is available at https://github.com/JSchweisthal/BoundMetaLearners.

from different environments. (i) One stream focuses on randomized data, typically to combine randomized and observational datasets (e.g., Kallus et al., 2018; Athey et al., 2020; Ghassami et al., 2022; Hatt et al., 2022; Imbens et al., 2022). (ii) Another stream uses purely observational data from different environments. For example, previous works offer a theory for the transportability of causal effects across different environments (Bareinboim & Pearl, 2016) or methods for transfer learning (Bica & van der Schaar, 2022) and for detecting unobserved confounding (Karlsson & Krijthe, 2023). However, none of these works estimates *bounds* for the CATE across multiple environments *under violations of assumptions*.

Meta-learners for CATE estimation: Model-agnostic meta-learners, in particular two-stage learners, are commonly used in the standard setting for CATE estimation without unobserved confounding and achieve state-of-the-art performance both theoretically and empirically (Foster & Syrgkanis, 2019; Curth & van der Schaar, 2021a). The basic idea of two-stage learners is to estimate the CATE directly via an additional second-stage regression using a constructed pseudo-outcome. Several learners have been developed, which primarily differ in terms of the pseudo-outcome used: (i) the regression-adjustment learner (called RA- or X-learner) (Künzel et al., 2019; Curth & van der Schaar, 2021a); the inverse-propensity weighted learner (IPW-learner) (Curth & van der Schaar, 2021a); the doubly robust learner (DR-learner) (Kennedy et al., 2023); and the R-learner (Nie & Wager, 2021). There also exist some efforts to generalize the standard meta-learners to different causal inference settings (e.g., with instrumental variables, Syrgkanis et al., 2019; Frauen & Feuerriegel, 2023). However, these meta-learners are all designed for estimating CATE under *point* identification but <u>not</u> *partial* identification.

We are only aware of one work that proposes a meta-learner for partial identification of CATE, namely, the B-learner (Oprescu et al., 2023). However, the B-learner has been proposed for causal sensitivity analysis, a setting with a fundamentally different set of assumptions: the B-learner requires prior knowledge that limits the maximum strength of the confounding. In contrast, we do *not* make assumptions that limit the strength of the confounding but instead assume that observational data is collected from multiple environments. Because of that, the B-learner is <u>not</u> applicable in our work.

Partial identification of treatment effects: In the context of causal inference from a single observational dataset with unobserved confounding, several works have proposed procedures for partial identification of treatment effects.

Manski (1990) was the first to obtain bounds for (conditional) average treatment effects under the assumption of bounded outcomes. Balke & Pearl (1997) derived tighter bounds for average treatment effects in a similar setting than ours, but with binary variables. Duarte et al. (2023) proposed a general procedure for deriving bounds in discrete structural causal models. In settings with continuous variables, recent works leverage machine learning approaches to learn bounds (Kilbertus et al., 2020; Hu et al., 2021; Balazadeh et al., 2022; Chen et al., 2023; Padh et al., 2023). Note that these methods focus primarily on *identification* results (i.e., deriving bounds), while we develop meta-learners for *estimating* bounds. An exception is the estimator from (Levis et al., 2023), which however targets bounds for *average* treatment effects.

Finally, a related but distinct stream of literature leverages machine learning for causal sensitivity analysis to obtain bounds under so-called sensitivity models, which make assumptions that limit the strength of unobserved confounding (Jesson et al., 2021; Dorn & Guo, 2022; Dorn et al., 2022; Yin et al., 2024; Frauen et al., 2023; Jin et al., 2023; Frauen et al., 2024). Again, we do not make an assumption that limits the strength of unobserved confounding, because of which the tasks/bounds are not comparable. On top of that, to the best of our knowledge, none of these works considers effective *estimation* of bounds for CATE via meta-learners.

Research gap: We focus on CATE estimation from observational data across multiple environments under violations of standard assumptions. To the best of our knowledge, we are the first to study partial identification in settings with *multiple* environments. Further, we are the first to develop effective *meta-learners* in our partial identification setting to estimate bounds.

3. Problem setup

Setting: We consider a setting with multiple environments (e.g., different hospitals, different physicians, or different countries), which, for simplicity, we denote by a discrete environment variable $E \in \{0, \dots, k\}$. We further have access to an observational dataset $\mathcal{D} = \{e_i, x_i, a_i, y_i\}_{i=1}^n$ of size n. The data are sampled i.i.d. from a population $(E, X, A, Y) \sim \mathbb{P}$, with patient covariates $X \in \mathcal{X} \subseteq \mathbb{R}^p$, discrete treatments $A \in \mathcal{A} \subseteq \mathbb{N}$, and bounded outcomes $Y \in \mathcal{Y} \subseteq [s_1, s_2] \subseteq \mathbb{R}$. The causal structure is shown in Fig. 2. In particular, we assume that E is an instrumental variable that has an effect on the treatment A but no direct effect on the outcome Y that is not mediated via A.

Our setting is relevant for a variety of practical applications where observational data are collected from different environments. Consider electronic healthcare records from patients, where X includes patient characteristics such as age and gender, A is an indicator of whether a patient has been prescribed medical treatment, and Y is a health out-

come such as heart rate or blood pressure. Further, let E correspond to an environment in which the electronic health records are collected such as the hospital. Such settings are common in medical research to understand the effectiveness of treatments (e.g., Huang et al., 2022).

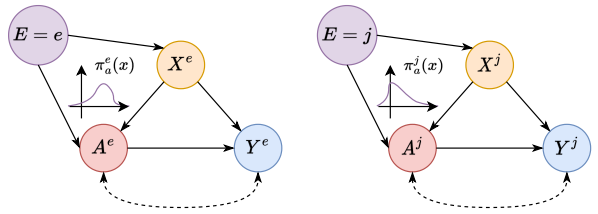

Figure 2. Causal graphs for different environments e and j. We assume that the causal structure between e and j remains unchanged but we allow for different treatment assignment mechanisms (propensity scores) $\pi_a^e(x)$ and $\pi_a^j(x)$. The dotted arrows indicate potential unobserved confounding.

Notation: We define the environment-specific *propensity score* as $\pi_a^e(x) = \mathbb{P}(A = a \mid X = x, E = e)$, which denotes the treatment assignment mechanism that is observed in environment e. Furthermore, we define the environment-specific *response surfaces* via $\mu_a^e(x) = \mathbb{E}[Y \mid X = x, A = a, E = e]$. Importantly, we assume that both the propensity scores and the response surfaces may differ across environments, i.e., $\pi_a^e(x) \neq \pi_a^j(x)$ and $\mu_a^e(x) \neq \mu_a^j(x)$ for environments $e \neq j$. For example, the standard of care might vary across countries, leading to different treatment policies and thus propensity scores. Finally, we define the environment probability as $\delta_e(x) = \mathbb{P}(E = e \mid X = x)$.

Assumptions: We use the potential outcomes framework to formalize our causal inference problem (Rubin, 1974). Let $Y^e(a) \in \mathcal{Y}$ denote the potential outcome in environment $E = e$ for a treatment intervention $A = a$. In this paper, we impose the following assumptions on the data-generating process across environments.

Assumption 3.1 (Consistency across environments)**.** For each environment e, observing $A = a$ implies $Y^e(a) = Y$.

Assumption 3.2 (Environment-agnostic oracle response surfaces)**.** We assume that $\mathbb{E}[Y^e(a) \mid X = x, E = e] = \mathbb{E}[Y^j(a) \mid X = x, E = j]$ holds for all environments $0 \leq e, j \leq k$, treatments $a \in \mathcal{A}$, and covariates $x \in \mathcal{X}$.

Assumption 3.3 (Common support)**.** We assume that $\delta_e(x) > 0$ holds for all environments $0 \leq e \leq k$ and covariates $x \in \mathcal{X}$.

Assumption 3.1 (consistency) excludes spillover effects across individuals and is a standard assumption in the causal inference literature (Shalit et al., 2017; Wager & Athey, 2018). Assumption 3.2 implies that a treatment results in the same expected potential outcome across environments.[2] In particular, the outcome-generating causal mechanisms coincide in expectation for each environment. Similar assumptions are commonly made in the literature to ensure the transferability of treatment effect estimates. Finally, Assumption 3.3 ensures that the populations of all environments are supported on a common domain, i.e., there are no types of patients that only exist in certain environments. Thus, we are consistent with prior literature on causal inference using combined datasets (Kallus et al., 2018; Shi et al., 2021; Hatt et al., 2022).

Target estimand: Using Assumption 3.2, we can define the oracle response surface as $\widetilde{\mu}_a(x) = \mathbb{E}[Y^e(a) \mid X = x, E = e]$, which is independent of the environment e. We are interested in the (environment-agnostic) conditional average treatment effect (CATE)

$$\tau_{a_1,a_2}(x) = \widetilde{\mu}_{a_1}(x) - \widetilde{\mu}_{a_2}(x), \tag{1}$$

which quantifies the difference in expected potential outcomes for two treatments $a_1, a_2 \in \mathcal{A}$. The CATE is of interest in various applications, e.g., personalized medicine (Feuerriegel et al., 2024), because it captures treatment effect heterogeneity by conditioning on the covariates x.

Violations of standard causal inference assumptions: The standard causal inference literature imposes the following two additional assumptions to identify the CATE $\tau_{a_1,a_2}(x)$ from the observational data distribution (Wager & Athey, 2018; Curth & van der Schaar, 2021a):

(i) *Overlap within envionments*: $\pi_a^e(x) > 0$; and
(ii) *Unconfoundedness*: $Y^e(a) \perp\!\!\!\perp A \mid X = x$

for all $x \in \mathcal{X}$ and $a \in \mathcal{A}$.

Overlap implies that each individual with covariates x must have a non-zero probability of receiving every treatment a. Unconfoundedness implies that the observed covariates X capture all confounding between the treatment A and outcome Y. Under overlap and unconfoundedness, the oracle response functions are identified via $\widetilde{\mu}_a(x) = \mu_a^e(x)$ and, hence, coincide with their observed counterparts for all environments e (Shalit et al., 2017). In particular, the CATE is identified via $\tau_{a_1,a_2}(x) = \mu_{a_1}^e(x) - \mu_{a_2}^e(x)$.

However, both overlap and unconfoundedness are violated in many practical settings. (i) The overlap violation is almost always violated in settings with high-dimensional covariates X, which is common in modern settings with available text/image data, or other electronic healthcare records (D'Amour et al., 2021; Schweisthal et al., 2023). (ii) The unconfoundedness assumption is highly unrealistic in medical settings, where confounders such as socio-economic status are rarely recorded in electronic health

[2]Prior literature often defines instrumental variables E via $Y(a) \perp\!\!\!\perp E \mid X$ (Pearl, 2009). Note that this assumption is stronger than our Assumption 3.2. Hence, our results can be applied to any setting where E is a discrete instrument in the classical sense, e.g., RCTs with non-compliance.

records (Adler & Stead, 2015). If either overlap or unconfoundedness (or both) are violated, neither response functions $\widetilde{\mu}_a(x) \neq \mu_a^e(x)$ nor the CATE are identified. Then, unbiased estimation is impossible (Pearl, 2009).

In the following, we relax assumptions related to overlap and unconfoundedness and focus on *partial identification*.

4. Partial identification of CATE

We now leverage the idea that the environment variable E acts as an instrumental variable (IV) and generalize results from the literature on partial identification for IVs (Balke & Pearl, 1997; Swanson et al., 2018) that hold for binary variables and average treatment effects. To do so, we first recall a known result from the literature on bounding treatment effects without multiple environments.

Lemma 4.1. *(Manski, 1990) For any environment e, the oracle response surfaces are bounded under Assumption 3.1 via*

$$\widetilde{\mu}_a(x) \leq \pi_a^e(x)\, \mu_a^e(x) + (1 - \pi_a^e(x))\, s_2 \tag{2}$$

and

$$\widetilde{\mu}_a(x) \geq \pi_a^e(x)\, \mu_a^e(x) + (1 - \pi_a^e(x))\, s_1, \tag{3}$$

where $[s_1, s_2]$ denotes the support of Y.

Proof. See Appendix B. □

The intuition behind Lemma 4.1 is as follows: Whenever we have overlap violations and the propensity score $\pi_a^e(x)$ is large for a treatment a, most of the randomness in the treatment is removed by conditioning on the observed covariates x. Hence, there can *not* be unobserved confounders that have a strong influence on the treatment assignment. This is reflected in the bounds from Lemma 4.1, which become tighter whenever $\pi_a^e(x)$ becomes larger.

Previously, Manski (1990) proposed to obtain an upper bound for the CATE $\tau_{a_1,a_2}(x) = \widetilde{\mu}_{a_1}(x) - \widetilde{\mu}_{a_2}(x)$ by combining the upper and lower bound from Lemma 4.1 for the different treatments a_1 and a_2. This results in

$$\begin{aligned}\tau_{a_1,a_2}(x) \leq\ & \pi_{a_1}^e(x)\, \mu_{a_1}^e(x) + (1 - \pi_{a_1}^e(x))\, s_2 \\ & - \pi_{a_2}^e(x)\, \mu_{a_2}^e(x) - (1 - \pi_{a_2}^e(x))\, s_1\end{aligned} \tag{4}$$

for the upper bound. An analogous result can be obtained for the lower bound by swapping the support points s_1 and s_2.

One drawback of the Manski bound in Eq. (4) is that it is not particularly tight. When subtracting the lower bound from the upper bound, we obtain a constant tightness of $s_2 - s_1$, which does *not* depend on the covariates x and *not* on the propensity score $\pi_a^e(x)$. In particular, the tightness equals the full support of the outcome variable Y.

Extending the Manski bounds: The above motivates us to leverage the different propensity scores across environments to obtain tighter bounds (see Fig. 1). By combining the bounds from Lemma 4.1 for the oracle response functions across environments, we derive the following result.

Theorem 4.2. *Under Assumptions 3.1 and 3.2, the CATE is bounded via*

$$b^-(x) \leq \tau_{a_1,a_2}(x) \leq b^+(x), \tag{5}$$

where

$$b^+(x) = \min_{e,j} b_{e,j}^+(x) \quad \textit{and} \quad b^-(x) = \max_{e,j} b_{e,j}^-(x) \tag{6}$$

with

$$\begin{aligned}b_{e,j}^+(x) =\ & \pi_{a_1}^e(x)\mu_{a_1}^e(x) + (1 - \pi_{a_1}^e(x))s_2 \\ & - \pi_{a_2}^j(x)\mu_{a_2}^j(x) - (1 - \pi_{a_2}^j(x))s_1,\end{aligned} \tag{7}$$

$$\begin{aligned}b_{e,j}^-(x) =\ & \pi_{a_1}^e(x)\mu_{a_1}^e(x) + (1 - \pi_{a_1}^e(x))s_1 \\ & - \pi_{a_2}^j(x)\mu_{a_2}^j(x) - (1 - \pi_{a_2}^j(x))s_2.\end{aligned} \tag{8}$$

Proof. See Appendix B. □

The bounds from Theorem 4.2 are a generalization of the so-called *natural bounds* from the literature on partial identification with IVs and discrete variables (Swanson et al., 2018). Unlike other bounds such as the ones from Balke & Pearl (1997), these can be straightforwardly extended to continuous outcomes and thus our multiple environments setting via Theorem 4.2. We refer to Swanson et al. (2018) for a detailed overview on bounds in instrumental settings (with discrete variables). Finally, another advantage of the natural bounds is that they are in closed-form and thus allow for estimation via meta-learners (see Sec. 5).

Tightness of the bounds: Note that our bounds from Theorem 4.2 satisfy

$$\begin{aligned}& b^+(x) - b^-(x) \\ & \quad \leq \min_{e,j} \left\{(s_2 - s_1)(2 - \pi_{a_1}^e(x) - \pi_{a_2}^j(x))\right\},\end{aligned} \tag{9}$$

which shows that they improve on the Manski bounds from Eq. (4) in terms of tightness whenever $\pi_{a_1}^e(x) + \pi_{a_2}^j(x) > 1$.

Remark: Our bounds become tight whenever there exist two environments e and j so that both $\pi_{a_1}^e(x)$ and $\pi_{a_2}^j$ are large. In other words, we have a rather surprising implication: *violations of overlap can be beneficial to increase the tightness* of our bounds as long as the respective treatment assignments across environments change sufficiently.

Intuitively, the degree to which the propensity scores across environments differ can also be viewed as the strength of dependence between the environment variable E and the treatment A. In the literature on IV regression, it is well-known

that so-called weak instruments E that are only weakly correlated with A may lead to biased estimates (Angrist & Pischke, 2008). Our bounds from Theorem 4.2 automatically account for this issue and become wider whenever E is a weak instrument and the propensity scores $\pi^e_{a_1}(x)$ and $\pi^j_{a_2}$ across environments become similar.

5. Meta-learners for estimating the bounds

We now develop meta-learners for estimating our bounds from Theorem 4.2, which denote quantities in population. In the following, we thus propose model-agnostic meta-learners that can be used in combination with any machine-learning method. In this section, we describe the estimation of our bounds for binary treatments $A \in \{0, 1\}$, so that $a_1 = 1$ and $a_2 = 0$. We also provide a theoretical analysis (Sec. 5.3) and an implementation using neural networks (Sec. 5.4).

For simplicity, we will focus on obtaining estimators $\hat{b}^+(x)$ for the upper bound $b^+(x)$. Estimators for the lower bound $b^-(x)$ can be obtained the same way by interchanging s_1 and s_2 (see Theorem 4.2).

5.1. Naïve plug-in learner

A straightforward way to obtain an estimator for $b^+(x)$ is the so-called naïve *plug-in* approach. In this, we first obtain estimators $\hat{\mu}^e_a(x)$ and $\hat{\pi}^e_a(x)$ of the *nuisance functions* $\mu^e_a(x)$ and $\pi^e_a(x)$ for all $a, e \in \{0, 1\}$. Note that estimating $\mu^e_a(x)$ is a regression task and estimating $\pi^e_a(x)$ is a classification task, which means that off-the-shelf machine-learning algorithms can be applied. Then, we directly plug the estimated nuisance functions into Eq. (6), which yields

$$\hat{b}^+(x) = \min_{e,j} \Big\{ \hat{\pi}^e_1(x)\hat{\mu}^e_1(x) + \hat{\pi}^e_0(x)\, s_2 - \hat{\pi}^j_0(x)\, \hat{\mu}^j_0(x) - \hat{\pi}^j_1(x) s_1 \Big\}. \quad (10)$$

We call the estimator from Eq. (10) the *naïve plugin-learner*.

The general approach behind plug-in learners can suffer from a so-called plug-in bias (Kennedy, 2022) that can limit estimation performance. As a remedy, we develop novel two-stage learners for our task in the following.

5.2. Two-stage learners

We now aim to address the drawbacks of plug-in learners by proposing so-called *two-stage learners* that directly estimate the bounds from Theorem 4.2. While the plug-in learner plugs the estimated nuisance function into Eq. (10) to estimate $b^+_{e,j}(x)$, we now aim to estimate $b^+_{e,j}(x)$ directly via a second-stage learner for all environment combinations $e, j \in \{0, 1\}$. If the bounds $b^+_{e,j}(x)$ are easier to estimate directly than the corresponding nuisance functions (e.g., due to cancellation effects), such an approach should perform better than the plug-in learner. We do this by constructing pseudo-outcomes that are equal to the bound we want to estimate in expectation, which then can be used as a second-stage regression objective. As a result, we obtain novel two-stage learners for estimating $b^+_{e,j}(x)$ and thus $b^+(x)$ (instead of the non-identifiable CATE $\tau_{1,0}(x)$).

Types of learners: We distinguish between two different cases: (1) within-environment bounds $b^+_{e,e}(x)$, which combine the response function bounds from Lemma 4.1 within the same environment e, and (2) cross-environment bounds $b^+_{e,j}(x)$, which combine the response function bounds across different environments $e \neq j$. Consequently, we yield two different approaches for the second-stage learners, which we call (1) **WB-learner** and (2) **CB-learner** (where the latter comes in different variants called CB-PI, CB-RA-, CB-IPW-, and CB-DR-learner). Once all second-stage learners $\hat{b}^+_{e,j}(x)$ are fitted, the within-environment bounds and the cross-environment bounds are combined via $\hat{b}^+(x) = \min_{e,j} \hat{b}^+_{e,j}(x)$ to obtain the final bound estimator for CATE. The full procedure is shown in Algorithm 1.

5.2.1. WB-LEARNER

Whenever we consider a single environment $e = j$, we define the pseudo-outcome

$$\hat{B}^{+\mathrm{WB}}_e = \mathbb{1}\{E = e\}\,(AY + (1 - A)s_2 - (1 - A)Y - As_1)\,. \quad (11)$$

Then, we use the pseudo-outcome $\hat{B}^{+\mathrm{WB}}_e$ to estimate $\hat{b}^+_{e,e}(x) = \hat{\mathbb{E}}[\hat{B}^+_e \mid X = x, E = e]$ via a second-stage learner $\hat{\mathbb{E}}$ conditional on $E = e$. That is, we only use the data from environment e for the estimator $\hat{b}^+_{e,e}(x)$. We call this the *within-environment bound-learner* (WB-learner).

5.2.2. CB-LEARNERS

Whenever $e \neq j$, we need to combine data from different environments e and j to estimate the cross-environment bounds $b^+_{e,j}(x)$. For this purpose, we show that the estimation of $b^+_{e,j}(x)$ can be cast into a standard CATE estimation problem with a transformed outcome and treatment variable $\widetilde{Y}$ and using the environment E as a (multi-valued) treatment. That is, we define

$$\widetilde{Y}^+_{e,j} = \mathbb{1}\{E = e\}\,(AY + (1 - A)s_2) + \mathbb{1}\{E = j\}\,((1 - A)Y + As_1)\,, \quad (12)$$

and we denote the *transformed response functions* as $r^+_\ell(x) = \mathbb{E}[\widetilde{Y}^+_{e,j} \mid X = x, E = \ell]$ for $\ell \in \{e, j\}$.

High-level approach: We proceed in two stages. In **stage 1**, we obtain estimators $\hat{r}^+_\ell$ and $\hat{\delta}_e(x)$ of the nuisance functions r^+_ℓ and $\delta_\ell(x)$. In **stage 2**, we treat the estimation of $b^+_{e,j}(x)$ as a standard CATE estimation task using $\hat{r}^+_\ell$ as estimated

response functions and $\hat{\delta}_\ell(x)$ as the estimated propensity score. Then, we adopt the classical meta-learners (plugin-, RA-, IPW-, and DR-learner (Curth et al., 2020; Kennedy, 2023)) for multi-valued treatments (Acharki et al., 2023) (i.e., our environments). For guarantees on this approach, we refer to Sec. 5.3.

(i) **CB-PI-learner**: The *cross-environment-bound plugin-learner* is defined via

$$\hat{b}^{+,\mathrm{PI}}_{e,j}(x) = \hat{r}^+_e(x) - \hat{r}^+_j(x), \tag{13}$$

which is analogous to the plugin-learner for standard CATE estimation (Curth et al., 2020).

(ii) **CB-RA-learner**: The *cross-environment-bound regression-adjustment learner* is defined via the pseudo outcome

$$\hat{B}^{+\mathrm{RA}}_{e,j} = \mathbb{1}\{E=e\}\left(\widetilde{Y}^+_{e,j} - \hat{r}^+_j(x)\right) + \mathbb{1}\{E=j\}\left(\hat{r}^+_e(x) - \widetilde{Y}^+_{e,j}\right) + \mathbb{1}\{E\neq e\}\mathbb{1}\{E\neq j\}\left(\hat{r}^+_e(x) - \hat{r}^+_j(x)\right), \tag{14}$$

and the second-stage pseudo-outcome regression $\hat{b}^+_{e,j}(x) = \hat{\mathbb{E}}[\hat{B}^{+\mathrm{RA}}_{e,j} \mid X = x]$.

(iii) **CB-IPW-learner**: The *cross-environment-bound inverse propensity weighted learner* is defined via

$$\hat{B}^{+\mathrm{IPW}}_{e,j} = \left(\frac{\mathbb{1}\{E=e\}}{\hat{\delta}_e(x)} - \frac{\mathbb{1}\{E=j\}}{\hat{\delta}_j(x)}\right)\widetilde{Y}^+_{e,j}, \tag{15}$$

and the second-stage pseudo-outcome regression $\hat{b}^+_{e,j}(x) = \hat{\mathbb{E}}[\hat{B}^{+\mathrm{IPW}}_{e,j} \mid X = x]$.

(iv) **CB-DR-learner**: The *cross-environment-bound doubly robust learner* is defined via

$$\hat{B}^{+\mathrm{DR}}_{e,j} = \hat{B}^{+\mathrm{IPW}}_{e,j} + \left(1 - \frac{\mathbb{1}\{E=e\}}{\hat{\delta}_e(x)}\right)\hat{r}^+_e(x) - \left(1 - \frac{\mathbb{1}\{E=j\}}{\hat{\delta}_j(x)}\right)\hat{r}^+_j(x) \tag{16}$$

and the second-stage pseudo-outcome regression $\hat{b}^+_{e,j}(x) = \hat{\mathbb{E}}[\hat{B}^{+\mathrm{DR}}_{e,j} \mid X = x]$.

(v) **Further meta-learners**: We emphasize that, in principle, *any* existing meta-learner for CATE estimation can be applied for estimating the cross-environment bounds. Further learners not specifically discussed here include the U-learner (Künzel et al., 2019) and the R-learner (Nie & Wager, 2021) (in the case of a binary environmental variable E).

5.3. Theoretical guarantees

The following result analyzes the consistency of our proposed second-stage learners depending on whether the nuisance parameters are correctly specified.

Theorem 5.1 (Consistency and double robustness)**.** *The meta-learners are consistent estimators of the within and cross-environment bounds. For the WB-learner, it holds that*

$$b^+_{e,e}(x) = \mathbb{E}[\hat{B}^{+\mathrm{WB}}_{e,e} \mid X = x, E = e] \tag{17}$$

Algorithm 1: Two-stage learners for estimating bounds

Input : observational data (E, X, A, Y), method $m \in \{\mathrm{RA}, \mathrm{IPW}, \mathrm{DR}\}$
Output : estimated upper bound $\hat{b}^+(x)$ (for $\hat{b}^-(x)$, interchange s_1 and s_2)
// Stage 1 (nuisance estimation)
$\hat{r}^+_\ell(x) \leftarrow \hat{\mathbb{E}}[\widetilde{Y}^+_{e,j} \mid X = x, E = \ell]$
$\hat{\delta}_\ell(x) \leftarrow \hat{\mathbb{P}}(E = \ell \mid X = x)$
// Stage 2
for $e, j \in \{0, \dots, k\}$ **do**
  if $e = j$ **then**
    $\hat{b}^+_{e,e}(x) = \hat{\mathbb{E}}[\hat{B}^{+\mathrm{WB}}_e \mid X = x, E = e]$
  else
    $\hat{b}^+_{e,j}(x) = \hat{\mathbb{E}}[\hat{B}^{+m}_{e,j} \mid X = x]$
  end
end
// Final bound
$\hat{b}^+(x) = \min_{e,j} \hat{b}^+_{e,j}(x)$

for $e \in \{0,1\}$. *For the CB-m-learner with* $m \in \{\mathrm{RA}, \mathrm{IPW}, \mathrm{DR}\}$, *we obtain*

$$b^+_{e,j}(x) = \mathbb{E}[\hat{B}^{+m}_{e,j} \mid X = x] \tag{18}$$

for $e, j \in \{0,1\}$ *whenever, for all* $e, a \in \{0,1\}$, *one of the following cases hold: (i)* $\hat{r}^+_\ell(x) = r^+_\ell(x)$ *for* $m = \mathrm{RA}$*; (ii)* $\hat{\delta}_\ell(x) = \delta_\ell(x)$ *for* $m = \mathrm{IPW}$*; or (iii)* <u>*either*</u> $\hat{r}^+_\ell(x) = r^+_\ell(x)$ <u>*or*</u> $\hat{\delta}_\ell(x) = \delta_\ell(x)$ *for* $m = \mathrm{DR}$.

Proof. See Appendix B. □

Note that the CB-DR-learner satisfies a double robustness property analogously to the standard DR-learner for CATE estimation (Kennedy et al., 2023). In particular, it allows for a misspecification of either $\hat{r}^+_\ell(x)$ or $\hat{\delta}^\ell(x)$ as long as the other nuisance function is correctly specified. Furthermore, the WB-learner is always consistent because the corresponding pseudo-outcome from Eq. (11) does not depend on any nuisance estimators but just on the observational data. This is different for cross-environment bounds $b^+_{e,j}(x)$, which would require observing the same data in two different environments. In that sense, the problem of estimating the $b^+_{e,j}(x)$ gives rise to the *fundamental problem of causal inference*: potential outcomes corresponding to different environments are never observed simultaneously.

5.4. Implementation using neural networks

In general, our proposed meta-learners are model-agnostic in both stages and thus work with arbitrary machine learning models. Here, we provide a flexible implementation using neural networks. Specifically, we adapt the implementations used in previous works for evaluating meta-learners for point-identified CATE estimation (Curth & van der Schaar, 2021a;b). Importantly, we use similar network architectures and settings for all meta-learners. Thus, performance differences are decoupled from model choice, so that the source of gain is given by the meta-learners. We report the implementation details in Appendix D. Here, we want to emphasize

method	Synthetic Data 1	Synthetic Data 2
WB naïve	0.073 ± 0.031	0.075 ± 0.045
WB	0.142 ± 0.069	0.130 ± 0.077
CB naïve	0.148 ± 0.098	0.156 ± 0.105
CB-PI	0.125 ± 0.059	0.127 ± 0.063
CB-RA	0.179 ± 0.089	0.119 ± 0.037
CB-IPW	**0.117 ± 0.057**	0.165 ± 0.072
CB-DR	0.132 ± 0.061	**0.111 ± 0.069**

Table 1. Mean and standard deviation of the RMSE over 5 random runs for synthetic datasets 1&2.

that our experiments serve as a "proof-of-demonstration" to show the behavior of the different meta-learners in different settings, while the neural network architectures are not optimized for the respective tasks. Instead, in practice, the exact modeling choices can be flexibly adjusted to fit data properties and desired inductive biases.

6. Experiments

We perform experiments on synthetic and real-world data to demonstrate that our meta-learners can effectively learn bounds in our setting with multiple environments. Synthetic data are commonly used to evaluate causal inference methods as it ensures that the causal ground-truth is available (Curth & van der Schaar, 2021a; Xu et al., 2021) and thus allows us to examine the performance. Since, as in CATE estimation (Curth & van der Schaar, 2021a), the performance of our different individual meta-learners depends strongly on the exact data-generating process, we refrain from heavy benchmarking. Instead, we show that our meta-learners reliably learn valid bounds and give short intuition to their behavior.

Synthetic data: For synthetic data, unlike for real-world data, we have access to the ground-truth data-generating process. Hence, we can compare our meta-learners against the oracle CATE and the oracle bounds from Theorem 4.2 calculated with the ground-truth nuisance estimators. Here, we consider two settings where we vary the complexity of the environment probability $\delta_e(x)$ from rather simple (dataset 1) to more complex (dataset 2). Details on the data-generating mechanisms are in Appendix E. In both cases, the treatment assignment tends to become more different across environments at the border of X, so that the cross-environment bounds should become tighter and more informative in these areas.

We report the results in terms of the root mean squared error (RMSE) to the oracle bounds in Table 1. We further display illustrative insights into the true and predicted bounds of different meta-learners for dataset 2 in Figure 3. We can observe the following: (i) All of our meta-learners learn valid bounds reliably, as shown by low average and variation of the RMSE and comparable performances between the meta-learners. (ii) Depending on the setting, different meta-learners may perform better or worse. As expected, in dataset 1 with simple $\delta_e(x)$, the CB-IPW learner performs best, while, in the more complex dataset 2, the CB-DR-learner shows the best performance. This is in line with previous work around meta-learners for CATE estimation (Curth & van der Schaar, 2021a), while we show that the same holds true for meta-learners aimed at partial identification. (iii) The cross-environment bounds can be especially helpful for learning tight and informative bounds, as demonstrated in the border regions of X in Figure 3. Hence, reliable estimation is particularly important in these areas, and, in the considered settings, our CB-learners help to improve performance over the naïve plug-in learner.

Furthermore, in Figure 3, we can give an intuition into the differences in bounds estimation between the naïve plug-in learner and the two-stage learners. For the latter, we use the WB-learner in combination with the CB-DR-learner as a representative example for illustration. While still performing well, we can observe that the naïve plugin-learner yields slightly less stable estimates with higher stand deviations, which is expected due to estimation errors in the nuisance estimation. The two-stage learners, in contrast, are less prone to minor errors in the nuisance estimators and, as expected, yield slightly smoother and more robust estimates of the bounds, as shown by lower stand deviations. This is also reflected by lower RMSE and lower variation in Table 1. However, note that this can result in oversmoothing when the ground-truth bounds have a rather complex form, e.g., as for the within-environment bounds. Interestingly, as stated above, also here a similar behavior was observed for meta-learners for point identification. In sum, our results show that, depending on the data properties, different learners can be preferable for estimating bounds, and we provide a selection of different options with our novel meta-learners.

Real-world data: We now demonstrate the applicability of our meta-learners to real-world data. Since the ground-truth CATE is unobservable for real-world data, we refrain from benchmarking, but, instead, our primary aim is to demonstrate the practical value of our meta-learners. Here, we perform a case study using a dataset with COVID-19 hospitalizations in Brazil across different regions (Baqui et al., 2020). We are interested in predicting the effect of comorbidity on the mortality of COVID-19 patients. For the environments, we use the regions of the hospitals in Brazil, which are split into North and Central-South. As observed confounders, we include age, sex, and ethnicity.

We report the estimates of the CATE bounds wrt. to age and averaged over the other confounders ($mean \pm std$),

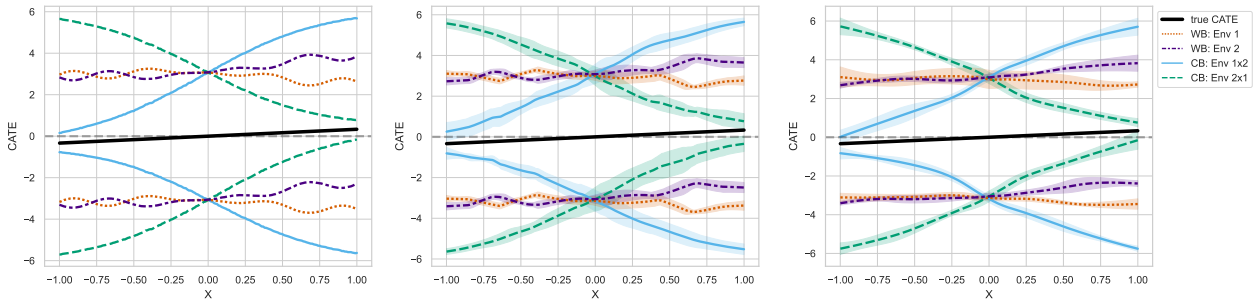

Figure 3. Comparison of estimation methods for bounds based on synthetic dataset 2 (predicted bounds $mean \pm 3std$ over 5 runs). **Left:** Oracle bounds for within- (WB) and cross- (CB) environments. **Center:** Estimated bounds by the naïve plug-in learner. **Right:** Estimated bounds by our two-stage meta-learners (here: WB-learner with CB-DR-learner).

estimated by (i) the naïve plug-in learner and (ii) the CB-PI-learner as another example for the two-stage learners. The results are displayed in Fig. 4. As expected, for both learners, the best lower bounds are closer to zero than the best upper bounds for all ages, indicating that comorbidity has no (large) negative effect on mortality (i.e., no large positive effect on survival probability). However, given our data, we cannot prove that comorbidity has an effect on mortality at all. Further, we observe that the cross-environment bounds help to tighten the bounds. In comparison, our two-stage learners yield similar predictions to the naïve plug-in learner, indicating robust estimation. In sum, the results demonstrate the applicability of our meta-learners to real-world data.

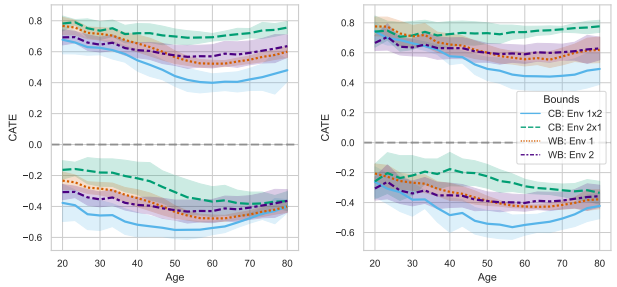

Figure 4. Insights from real-world data: Effect of comorbidity on mortality in COVID-19 patients in Brazil. **Left:** Estimated bounds by naïve plug-in learner. **Right:** Estimated bounds by our two-stage meta-learners (here: WB-learner and CB-PI-learner).

7. Applicability to IV settings

Our meta-learners apply to any setting in which a discrete instrumental variable (IV), E, is available, which does not necessarily need to correspond to an environment indicator. A prominent example are randomized controlled trials (RCTs) with non-compliance. Note that RCTs with non-compliance do not need multiple environments but can have just a single environment. Then, E corresponds to a randomized treatment prescription (e.g., by a physician) and A indicates whether the patient complied and followed the treatment decision. Here, the randomized treatment assignment automatically renders E independent of potential unobserved confounders, thus fulfilling Assumption 3.2.

In settings with discrete IVs, there is again rich literature on CATE estimation using point estimates (Syrgkanis et al., 2019; Frauen & Feuerriegel, 2023). There are also different efforts aimed at the *derivation* of bounds (Swanson et al., 2018), but methods for *estimation*, such as meta-learners, are scarce. To the best of our knowledge, our work is the first to propose meta-learners for estimating such bounds.

8. Discussion

Conclusion: In this paper, we derive tight bounds for partial identification of the CATE with data from multiple environments by showing the relation of our setting to IV settings. Further, we propose novel model-agnostic meta-learners for estimating these bounds from observational data. We find that the cross-environment bounds are especially helpful for learning tight bounds in areas where the treatment assignment policy varies clearly between the environments, and that the performance ranking of the respective meta-learners depends on the data generating process. Hence, our proposed meta-learners provide a valuable tool for tailoring the estimation of bounds for the CATE with data from multiple environments to the respective properties of different settings.

Future directions: We envision that our ideas could be used to develop meta-learners in other partial identification settings, for which effective estimation procedures are often lacking. Examples are meta-learners for settings with continuous instruments, leaky mediation, or sensitivity analysis.

Impact statement

We acknowledge our meta-learners build on formal assumptions that are standard in the causal inference literature. Hence, we recommend a cautious and responsible use of our meta-learners to ensure that the assumptions are met, so that reliable inferences can be made. Even though we aim to relax standard assumptions such as unconfoundedness, we still implicitly assume that the environmental variable E acts as an instrument, i.e., is not correlated with any potential unobserved confounders. This assumption can generally not be tested and must be justified by domain knowledge. However, we note that there are many real-world scenarios where this assumption holds, for example, RCTs with non-compliance (Finkelstein et al., 2012).

Notwithstanding, our work aims to relax the (strict) assumptions of prior research by allowing for partial identification (instead of point identification). This can help to make causal inference more robust. As such, it can even help to improve the reliability of inferences for marginalized groups. For example, marginalized groups are different from the rest of the population in terms of their socio-economic status, which is a known confounder in medical studies, so that causal inference that ignores such confounder may be biased, while our meta-learners can help to make more reliable inferences.

References

Acharki, N., Lugo, R., Bertoncello, A., and Garnier, J. Comparison of meta-learners for estimating multi-valued treatment heterogeneous effects. In *International Conference on Machine Learning*, 2023.

Adler, N. E. and Stead, W. W. Patients in context: EHR capture of social and behavioral determinants of health. *The New England Journal of Medicine*, 372(8):695–698, 2015.

Angrist, J. D. Lifetime earnings and the vietnam era draft lotter: Evidence from social security administrative records. *The American Economic Review*, 80(3):313–336, 1990.

Angrist, J. D. and Pischke, J.-S. *Most harmless econometrics*. Princeton University Press, Princeton, 2008.

Athey, S., Chetty, R., and Imbens, G. Combining experimental and observational data to estimate treatment effects on long term outcomes. *arXiv preprint*, arXiv:2006.09676, 2020.

Balazadeh, V., Syrgkanis, V., and Krishnan, R. G. Partial identification of treatment effects with implicit generative models. In *NeurIPS*, 2022.

Balke, A. and Pearl, J. Bounds on treatment effects from studies with imperfect compliance. *Journal of the American Statistical Association*, 92(439):1171–1176, 1997.

Baqui, P., Bica, I., Marra, V., Ercole, A., and van Der Schaar, M. Ethnic and regional variations in hospital mortality from covid-19 in brazil: a cross-sectional observational study. *The Lancet Global Health*, 8(8):e1018–e1026, 2020.

Bareinboim, E. and Pearl, J. Causal inference and the data-fusion problem. *Proceedings of the National Academy of Sciences of the United States of America*, 113(27): 7345–7352, 2016.

Bica, I. and van der Schaar, M. Transfer learning on heterogeneous feature spaces for treatment effects estimation. In *NeurIPS*, 2022.

Chen, K., Wang, B., and Small, D. S. A differential effect approach to partial identification of treatment effects. *arXiv preprint*, arXiv:2303.06332, 2023.

Curth, A. and van der Schaar, M. Nonparametric estimation of heterogeneous treatment effects: From theory to learning algorithms. In *AISTATS*, 2021a.

Curth, A. and van der Schaar, M. On inductive biases for heterogeneous treatment effect estimation. In *NeurIPS*, 2021b.

Curth, A., Alaa, A. M., and van der Schaar, M. Estimating structural target functions using machine learning and influence functions. *arXiv preprint*, arXiv:2008.06461, 2020.

D'Amour, A., Ding, P., Feller, A., Lei, L., and Sekhon, J. Overlap in observational studies with high-dimensional covariates. *Journal of Econometrics*, 221(2):644–654, 2021.

Dorn, J. and Guo, K. Sharp sensitivity analysis for inverse propensity weighting via quantile balancing. *Journal of the American Statistical Association*, 2022.

Dorn, J., Guo, K., and Kallus, N. Doubly-valid/ doubly-sharp sensitivity analysis for causal inference with unmeasured confounding. *arXiv preprint*, arXiv:2112.11449, 2022.

Duarte, G., Finkelstein, N., Knox, D., Mummolo, J., and Shpitser, I. An automated approach to causal inference in discrete settings. *Journal of the American Statistical Association*, 2023.

Feuerriegel, S., Frauen, D., Melnychuk, V., Schweisthal, J., Hess, K., Curth, A., Bauer, S., Kilbertus, N., Kohane, I. S., and van der Schaar, M. Causal machine learning for

predicting treatment outcomes. *Nature Medicine*, 30(4): 958–968, 2024.

Finkelstein, A., Taubman, S., Wright, B., Bernstein, M., Gruber, J., Newhouse, J. P., Allen, H., and Baicker, K. The oregon health insurance experiment: Evidence from the first year. *The Quarterly Journal of Economics*, 127 (3):1057–1106, 2012.

Foster, D. J. and Syrgkanis, V. Orthogonal statistical learning. *arXiv preprint*, arXiv:1901.09036, 2019.

Frauen, D. and Feuerriegel, S. Estimating individual treatment effects under unobserved confounding using binary instruments. In *ICLR*, 2023.

Frauen, D., Melnychuk, V., and Feuerriegel, S. Sharp bounds for generalized causal sensitivity analysis. In *NeurIPS*, 2023.

Frauen, D., Imrie, F., Curth, A., Melnychuk, V., Feuerriegel, S., and van der Schaar, M. A neural framework for generalized causal sensitivity analysis. In *The Twelfth International Conference on Learning Representations*, 2024.

Ghassami, A., Yang, A., Richardson, D., Shpitser, I., and Tchetgen, E. T. Combining experimental and observational data for identification and estimation of long-term causal effects. *arXiv preprint*, arXiv:2201.10743, 2022.

Glass, T. A., Goodman, S. N., Hernán, M. A., and Samet, J. M. Causal inference in public health. *Annual Review of Public Health*, 34:61–75, 2013.

Hatt, T., Berrevoets, J., Curth, A., Feuerriegel, S., and van der Schaar, M. Combining observational and randomized data for estimating heterogeneous treatment effects. *arXiv preprint*, 2022.

Heinze-Deml, C., Peters, J., and Meinshausen, N. Invariant causal prediction for nonlinear models. *Journal of Causal Inference*, 6(2), 2018. ISSN 2193-3677.

Hu, Y., Wu, Y., and Wu, X. A generative adversarial framework for bounding confounded causal effects. In *AAAI*, 2021.

Huang, Y., Talwar, A., Lin, Y., and Aparasu, R. R. Machine learning methods to predict 30-day hospital readmission outcome among us adults with pneumonia: analysis of the national readmission database. *BMC Medical Informatics and Decision Making*, 22(1):288, 2022.

Ilse, M., Forré, P., Welling, M., and Mooij, J. M. Combining interventional and observational data using causal reductions. *arXiv preprint*, arXiv:2103.04786, 2023.

Imbens, G., Kallus, N., Mao, X., and Wang, Y. Long-term causal inference under persistent confounding via data combination. *arXiv preprint*, 2022.

Jesson, A., Mindermann, S., Gal, Y., and Shalit, U. Quantifying ignorance in individual-level causal-effect estimates under hidden confounding. In *ICML*, 2021.

Jin, Y., Ren, Z., and Candès, E. J. Sensitivity analysis of individual treatment effects: A robust conformal inference approach. *Proceedings of the National Academy of Sciences (PNAS)*, 120(6), 2023.

Johansson, F. D., Shalit, U., and Sonntag, D. Learning representations for counterfactual inference. In *ICML*, 2016.

Jung, Y., Diaz, I., Tian, J., and Bareinboim, E. Estimating causal effects identifiable from a combination of observations and experiments. In *NeurIPS*, 2023.

Kallus, N., Pulim Aahlad Manas, and Shalit, U. Removing hidden confounding by experimental grounding. In *NeurIPS*, 2018.

Kallus, N., Mao, X., and Zhou, A. Interval estimation of individual-level causal effects under unobserved confounding. In *AISTATS*, 2019.

Karlsson, R. K. A. and Krijthe, J. H. Combining observational datasets from multiple environments to detect hidden confounding. In *NeurIPS*, 2023.

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

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

Kennedy, E. H., Balakrishnan, S., and Wasserman, L. Semiparametric counterfactual density estimation. *Biometrika*, 2023. ISSN 0006-3444.

Kilbertus, N., Kusner, M. J., and Silva, R. A class of algorithms for general instrumental variable models. In *NeurIPS*, 2020.

Künzel, S. R., Sekhon, J. S., Bickel, P. J., and Yu, B. Metalearners for estimating heterogeneous treatment effects using machine learning. *Proceedings of the National Academy of Sciences (PNAS)*, 116(10):4156–4165, 2019.

Kuzmanovic, M., Frauen, D., Hatt, T., and Feuerriegel, S. Causal machine learning for cost-effective allocation of development aid. In *KDD*, 2024.

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

Manski, C. F. Nonparametric bounds on treatment effects. *The American Economic Review*, 80(2):319–323, 1990.

Mooij, J. M., Magliacane, S., and Claassen, T. Joint causal inference from multiple contexts. *Journal of Machine Learning Research*, 21:1–108, 2020.

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

Oprescu, M., Dorn, J., Ghoummaid, M., Jesson, A., Kallus, N., and Shalit, U. B-learner: Quasi-oracle bounds on heterogeneous causal effects under hidden confounding. In *ICML*, 2023.

Padh, K., Zeitler, J., Watson, D., Kusner, M., Silva, R., and Kilbertus, N. Stochastic causal programming for bounding treatment effects. In *CLeaR*, 2023.

Pearl, J. *Causality*. Cambridge University Press, New York City, 2009. ISBN 9780521895606.

Peters, J., Bühlmann, P., and Meinshausen, N. Causal inference by using invariant prediction: Identification and confidence intervals. *Journal of the Royal Statistical Society: Series B*, 78(5):947–1012, 2016. ISSN 1467-9868.

Rubin, D. B. Estimating causal effects of treatments in randomized and nonrandomized studies. *Journal of Educational Psychology*, 66(5):688–701, 1974. ISSN 0022-0663.

Schweisthal, J., Frauen, D., Melnychuk, V., and Feuerriegel, S. Reliable off-policy learning for dosage combinations. In *NeurIPS*, 2023.

Shalit, U., Johansson, F. D., and Sontag, D. Estimating individual treatment effect: Generalization bounds and algorithms. In *ICML*, 2017.

Shi, C., Blei, D. M., and Veitch, V. Adapting neural networks for the estimation of treatment effects. In *NeurIPS*, 2019.

Shi, C., Veitch, V., and Blei, D. M. Invariant representation learning for treatment effect estimation. In *UAI*, 2021.

Swanson, S. A., Hernán, M. A., Miller, M., Robins, J. M., and Richardson, T. S. Partial identification of the average treatment effect using instrumental variables: review of methods for binary instruments, treatments, and outcomes. *Journal of the American Statistical Association*, 113(522):933–947, 2018.

Syrgkanis, V., Lei, V., Oprescu, M., Hei, M., Battocchi, K., and Lewis, G. Machine learning estimation of heterogeneous treatment effects with instruments. In *NeurIPS*, 2019.

Varian, H. R. Causal inference in economics and marketing. *Proceedings of the National Academy of Sciences (PNAS)*, 113(27):7310–7315, 2016.

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

Xu, L., Chen, Y., Srinivasan, S., Freitas, N. d., Doucet, A., and Gretton, A. Learning deep features in instrumental variable regression. In *ICLR*, 2021.

Yin, M., Shi, C., Wang, Y., and Blei, D. M. Conformal sensitivity analysis for individual treatment effects. *Journal of the American Statistical Association*, 119(545):1–14, 2024.

Zhang, J., Tian, J., and Bareinboim, E. Partial counterfactual identificatiton from observational and experimental data. In *ICML*, 2022.

A. Extended related work

Causal inference using multiple datasets: A related stream of literature focuses on causal inference from multiple datasets and can be roughly separated into works that leverage (i) randomized or (ii) purely observational data.

Works from (i) often leverage small amounts of randomized data in combination with observational data to obtain unbiased causal effect estimates. For example, Kallus et al. (2018) and Hatt et al. (2022) use the observational data to obtain a biased CATE estimator, which is then subsequently debiased using the randomized data. Similar approaches exist for long-term effects (Athey et al., 2020; Ghassami et al., 2022; Imbens et al., 2022). An estimation theory for more general causal effects has been proposed by Jung et al. (2023). Finally, bounds for causal effects have been derived by Zhang et al. (2022) and Ilse et al. (2023). Note that all of these works require randomized data, while our paper relies on purely observational data.

Works from (ii) include invariant causal prediction methods (Peters et al., 2016; Heinze-Deml et al., 2018; Mooij et al., 2020) that aim to discover the causal predictors of a target variable. This idea has been adapted to CATE estimation to find a valid adjustment set of covariates (Shi et al., 2021). Bareinboim & Pearl (2016) proposed a theory for identifiability and transportability of causal effects across different environments under potential selection bias. Other works that leverage observational datasets from different environments include transfer learning for CATE under unconfoundedness (Bica & van der Schaar, 2022) and methods for detecting unobserved confounding (Karlsson & Krijthe, 2023). However, none of these works learns bounds for the CATE under violations of assumptions.

Model-based CATE estimation: Another stream of literature focuses on adapting specific machine learning models for CATE estimation. Examples include forest-based methods (Wager & Athey, 2018) and neural networks, which leverage shared representations across response surfaces and propensity scores to improve finite-sample performance (Shi et al., 2019; Curth & van der Schaar, 2021a). A complementary approach is to learn balanced representations that are unpredictable of the treatment (Johansson et al., 2016; Shalit et al., 2017). These methods are in contrast to model-agnostic meta-learners, which can be used with any *arbitrary* machine learning model.

B. Proofs

B.1. Proof of Lemma 4.1

Proof. The result follows from decomposing the oracle response function into a factual and counterfactual part, which can be bounded. That is,

$$\begin{aligned}\widetilde{\mu}_a(x) &= \pi_a^e(x)\mathbb{E}[Y(a) \mid X = x, A = a] + (1 - \pi_a^e(x))\mathbb{E}[Y(a) \mid X = x, A \neq a] && (19)\\ &= \pi_a^e(x)\mu_a^e(x) + (1 - \pi_a^e(x))\mathbb{E}[Y(a) \mid X = x, A \neq a] && (20)\\ &\leq \pi_a^e(x)\mu_a^e(x) + (1 - \pi_a^e(x))s_2, && (21)\end{aligned}$$

where we used the definition of s_1 and s_2 as the support boundary of Y. □

B.2. Proof of Theorem 4.2

Proof. Note that

$$\tau_{a_1,a_2}(x) = \widetilde{\mu}_{a_1}(x) - \widetilde{\mu}_{a_2}(x). \tag{22}$$

Hence, Lemma 4.1 and Assumptions 3.1 and 3.2 imply that

$$b_{e,j}^-(x) \leq \tau_{a_1,a_2}(x) \leq b_{e,j}^+(x) \tag{23}$$

for all environment e and j. Hence, taking the minimum and maximum over e and j yields the result. □

B.3. Proof of Theorem 5.1

Proof. Without loss of generality, we show the result for the upper bounds. For lower bounds, the proof works analogously by interchanging the support points s_1 and s_2. We proceed by calculating $\mathbb{E}[\hat{B}_{e,j}^+ \mid X = x]$ for each pseudo-outcome $\hat{B}_{e,j}^+$, which corresponds to an oracle second stage regression. We start with the WB-learner, which uses pseudo-outcomes defined via

$$\hat{B}_e^{+\mathrm{WB}} = \mathbb{1}\{E = e\}\left(AY + (1 - A)s_2 - (1 - A)Y - As_1\right). \tag{24}$$

Hence, we obtain

$$\begin{aligned}\mathbb{E}[\hat{B}_e^{+\mathrm{WB}} \mid X = x, E = e] &= \mathbb{E}[AY \mid X = x, E = e] + s_2\mathbb{E}[1 - A \mid X = x, E = e] && (25)\\ &\quad - \mathbb{E}[(1 - A)Y \mid X = x, E = e] - s_1\mathbb{E}[A \mid X = x, E = e] && (26)\\ &= \pi_1^e(x)\mu_1^e(x) + \pi_0^e(x)s_2 - \pi_0^e(x)\mu_0^e(x) - \pi_1^e(x)s_1 = b_{e,e}^+(x). && (27)\end{aligned}$$

as desired. The CB-RA-learner uses pseudo-outcomes defined via

$$\begin{aligned}\hat{B}_{e,j}^{+\mathrm{RA}} &= \mathbb{1}\{E = e\}\left(\widetilde{Y}_{e,j}^+ - \hat{r}_j^+(x)\right) + \mathbb{1}\{E = j\}\left(\hat{r}_e^+(x) - \widetilde{Y}_{e,j}^+\right) && (28)\\ &\quad + \mathbb{1}\{E \neq e\}\mathbb{1}\{E \neq j\}\left(\hat{r}_e^+(x) - \hat{r}_j^+(x)\right). && (29)\end{aligned}$$

Taking expectation conditional on $X = x$ yields

$$\begin{aligned}\mathbb{E}[\hat{B}_{e,j}^{+\mathrm{RA}} \mid X = x] &= \delta_e(x)\left(\mathbb{E}[AY \mid X = x, E = e] + s_2\mathbb{E}[1 - A \mid X = x, E = e] - \hat{r}_j^+(x)\right) && (30)\\ &\quad + \delta_j(x)\left(\hat{r}_e^+(x) - \mathbb{E}[(1 - A)Y \mid X = x, E = j] - s_1\mathbb{E}[A \mid X = x, E = j]\right) && (31)\\ &\quad + (1 - \delta_e(x) - \delta_j(x))\left(\hat{r}_e^+(x) - \hat{r}_j^+(x)\right) && (32)\\ &= \delta_e(x)\left(\pi_1^e(x)\mu_1^e(x) + \pi_0^e(x)s_2 - \hat{r}_j^+(x)\right) && (33)\\ &\quad + \delta_j(x)\left(\hat{r}_e^+(x) - \pi_0^j(x)\mu_0^j(x) - \pi_1^j(x)s_1\right) && (34)\\ &\quad + (1 - \delta_e(x) - \delta_j(x))\left(\hat{r}_e^+(x) - \hat{r}_j^+(x)\right). && (35)\end{aligned}$$

Hence, whenever $\hat{r}_e^+(x) = r_e^+(x) = \pi_1^e(x)\mu_1^e(x) + \pi_0^e(x)s_2$ and $\hat{r}_j^+(x) = r_j^+(x) = \pi_0^j(x)\mu_0^j(x) - \pi_1^j(x)s_1$, we obtain

$$\mathbb{E}[\hat{B}_{e,j}^{+\mathrm{RA}} \mid X = x] = \delta_e(x)b_{e,j}^+(x) + \delta_j(x)b_{e,j}^+(x) + +(1 - \delta_e(x) - \delta_j(x))b_{e,j}^+(x) = b_{e,j}^+(x). \tag{36}$$

The pseudo-outcomes of the CB-IPW-learner are defined via

$$\hat{B}_{e,j}^{+\mathrm{IPW}} = \frac{\mathbb{1}\{E=e\}}{\hat{\delta}_e(x)}\left(AY + (1-A)s_2\right) - \frac{\mathbb{1}\{E=j\}}{\hat{\delta}_j(x)}\left((1-A)Y + As_1\right), \tag{37}$$

which yields

$$\mathbb{E}[\hat{B}_{e,j}^{+\mathrm{IPW}} \mid X=x] = \frac{\delta_e(x)}{\hat{\delta}_e(x)}\left(\mathbb{E}[AY \mid X=x, E=e] + s_2\mathbb{E}[1-A \mid X=x, E=e]\right) \tag{38}$$

$$- \frac{\delta_j(x)}{\hat{\delta}_j(x)}\left(\mathbb{E}[(1-A)Y \mid X=x, E=j] + s_1\mathbb{E}[A \mid X=x, E=j]\right) \tag{39}$$

$$= \frac{\delta_e(x)}{\hat{\delta}_e(x)}\left(\pi_1^e(x)\mu_1^e(x) + \pi_0^e(x)s_2\right) - \frac{\delta_j(x)}{\hat{\delta}_j(x)}\left(\pi_0^j(x)\mu_0^j(x) + \pi_1^j(x)s_1\right). \tag{40}$$

Hence, $\hat{\delta}_\ell(x) = \delta_\ell(x)$ implies

$$\mathbb{E}[\hat{B}_{e,j}^{+\mathrm{IPW}} \mid X=x] = b_{e,j}^{+}(x). \tag{41}$$

Finally, the pseudo outcomes of the CB-DR-learner are defined via

$$\hat{B}_{e,j}^{+\mathrm{DR}} = \hat{B}_{e,j}^{+\mathrm{IPW}} + \left(1 - \frac{\mathbb{1}\{E=e\}}{\hat{\delta}_e(x)}\right)\hat{r}_e^{+}(x) - \left(1 - \frac{\mathbb{1}\{E=j\}}{\hat{\delta}_j(x)}\right)\hat{r}_j^{+}(x). \tag{42}$$

Again, by taking expectation conditional on $X=x$, we obtain

$$\mathbb{E}[\hat{B}_{e,j}^{+\mathrm{DR}} \mid X=x] = \frac{\delta_e(x)}{\hat{\delta}_e(x)}\left(\pi_1^e(x)\mu_1^e(x) + \pi_0^e(x)s_2\right) - \frac{\delta_j(x)}{\hat{\delta}_j(x)}\left(\pi_0^j(x)\mu_0^j(x) + \pi_1^j(x)s_1\right) \tag{43}$$

$$+ \left(1 - \frac{\delta_e(x)}{\hat{\delta}_e(x)}\right)\hat{r}_e^{+}(x) - \left(1 - \frac{\delta_j(x)}{\hat{\delta}_j(x)}\right)\hat{r}_j^{+}(x). \tag{44}$$

$$\tag{45}$$

Under $\hat{\delta}_\ell(x) = \delta_\ell(x)$, this reduces to

$$\mathbb{E}[\hat{B}_{e,j}^{+\mathrm{DR}} \mid X=x] = \left(\pi_1^e(x)\mu_1^e(x) + \pi_0^e(x)s_2\right) - \left(\pi_0^j(x)\mu_0^j(x) + \pi_1^j(x)s_1\right) + 0\,\hat{r}_e^{+} - 0\,\hat{r}_j^{+} \tag{46}$$

$$= b_{e,j}^{+}(x). \tag{47}$$

Under $\hat{r}_e^{+}(x) = r_e^{+}(x) = \pi_1^e(x)\mu_1^e(x) + \pi_0^e(x)s_2$ and $\hat{r}_j^{+}(x) = r_j^{+}(x) = \pi_0^j(x)\mu_0^j(x) - \pi_1^j(x)s_1$, the expression reduces to

$$\mathbb{E}[\hat{B}_{e,j}^{+\mathrm{DR}} \mid X=x] = \left(1 + \frac{\delta_e(x)}{\hat{\delta}_e(x)} - \frac{\delta_e(x)}{\hat{\delta}_e(x)}\right)\left(\pi_1^e(x)\mu_1^e(x) + \pi_0^e(x)s_2\right) - \left(1 + \frac{\delta_j(x)}{\hat{\delta}_j(x)} - \frac{\delta_j(x)}{\hat{\delta}_j(x)}\right)\left(\pi_0^j(x)\mu_0^j(x) + \pi_1^j(x)s_1\right) \tag{48}$$

$$= b_{e,j}^{+}(x), \tag{49}$$

which proves the result. □

C. Additional Insights

In Figures 3 and 4, we plot the learned bounds for (i) both within-environment bounds and (ii) both cross-environment bounds to demonstrate the general performance of our learners. However, in practice, we are usually only interested in the *tightest* bounds as stated in Theorem 4.2. Hence, for completeness, we plot the learned tightest bounds from Figures 3 and 4 separately in Figures 5 and 6. We observe the same patterns for the tightest bounds as for all possible bounds: For the synthetic dataset, the CB-DR-learner results in valid and slightly more stable estimates with lower variation between runs compared to the naïve plugin learner.

For the real-world dataset, where we cannot compare with ground-truth bounds, both selected learners yield similar estimates. Also, in both considered settings, almost solely our proposed cross-environment bounds result in the tightest bounds, while the within-environment bounds are not helpful for yielding informative bounds. Here, this implies that one should focus on proper estimation of the CB-bounds, which can be done by our different proposed two-stage CB-meta-learners. In sum, this demonstrates that both contributions, (i) the bounds making use of data from different environments, and (ii) the meta-learners, are useful for learning informative bounds robustly.

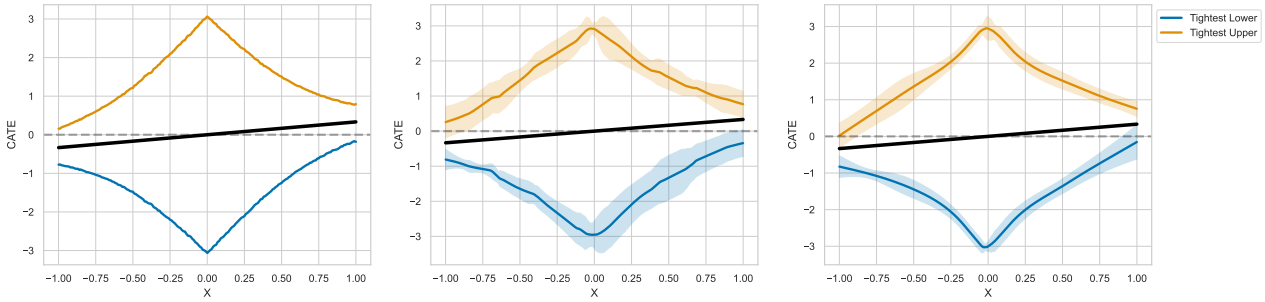

Figure 5. Comparison of estimation methods for *tightest* predicted bounds based on synthetic dataset 2 (predicted bounds $mean \pm 3std$ over 5 runs). **Left:** Oracle bounds for within- (WB) and cross- (CB) environments. **Center:** Estimated bounds by the naïve plug-in learner. **Right:** Estimated bounds by our two-stage meta-learners (here: WB-learner with CB-DR-learner).

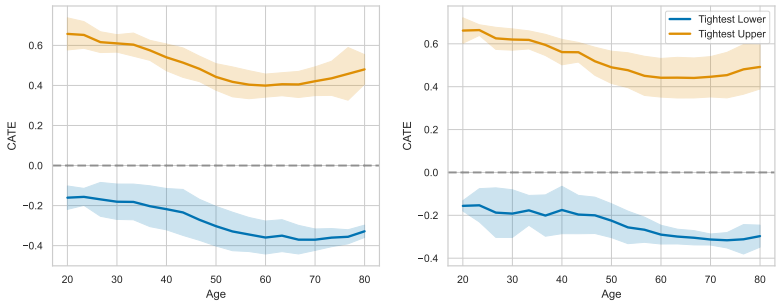

Figure 6. Insights from real-world data: Effect of comorbidity on mortality in COVID-19 patients in Brazil. **Left:** Estimated *tightest* bounds by naïve plug-in learner. **Right:** Estimated *tightest* bounds by our two-stage meta-learners (here: WB-learner and CB-PI-learner).

D. Implementation details

Model architecture and parameters: For our experiments, we adapt the implementations used in previous works for evaluating meta-learners for point-identified CATE estimation (Curth & van der Schaar, 2021a;b). Importantly, we use similar network architectures and parameter settings for all meta-learners. Thus, performance differences are decoupled from model choice and can be explained by the meta-learners as the source of gain. In detail, we use the software package `https://github.com/AliciaCurth/CATENets` and all of the default settings of the PyTorch CATE meta-learners provided in this package. Here, the networks for the first- and second-stage models are simple MLPs with 2 hidden layers and hidden neuron size of 100. For the nuisance function estimation of our naïve plugin estimator, we implement similar architectures.

Implementation: By using our pseudo-outcomes for bound estimation as described in Sec. 5.2, we can learn the bounds as follows: For the CB-learners, we define the pseudo-outcomes of Eq. (12) for each possible environment combination e, j and for upper and lower bounds, respectively. Then, by simply replacing the factual outcomes with the pseudo-outcomes, and the treatment assignment with the environment assignment, we can train the respective two-stage-learners by using the above-mentioned implementation of the CATE meta-learners. We train an own meta-learner for each environment combination and each upper and lower bound. For the WB-learner, we simply directly learn models to predict the pseudo-outcomes of Eq. (11) for each environment e, and upper and lower bound, respectively. We provide a wrapper module for our bound meta-learners in our code.[3]

[3] `https://github.com/JSchweisthal/BoundMetaLearners`

E. Details regarding simulated data

Data-generating process: For our synthetic dataset 1, we simulate an observed confounder $X \sim \text{Uniform}[-1, 1]$ and an unobserved confounder $U \sim \text{Uniform}[0, 1]$. We define $\delta(x) = \mathbb{P}(E = 1 \mid x) = \sigma(x)$, where $\sigma(\cdot)$ denotes the sigmoid function $\sigma(x) = \frac{1}{1+exp(-x)}$. Then, we simulate the environment E via

$$E = 1 \mid X = x \sim \text{Bernoulli}\left(p = \delta(x)\right)). \tag{50}$$

We then sample our treatment assignments from the environment-specific propensity scores as

$$A = 1 \mid X = x, U = u, E = 1 \sim \text{Bernoulli}\left(p = \sigma\left(2.5x + u\right)\right), \tag{51}$$

and

$$A = 1 \mid X = x, U = u, E = 0 \sim \text{Bernoulli}\left(p = \left(1 - \sigma\left(2.5x + u\right)\right)\right). \tag{52}$$

We define the treatment effect as

$$\tau(x) = \frac{1}{3}x. \tag{53}$$

Finally, we simulate a continuous outcome

$$Y = \tau(X)A + \frac{1}{3}(\sin(12X) + X) + \frac{1}{60}\cos(2X) + U + 0.3\varepsilon, \tag{54}$$

where $\varepsilon \sim \text{Laplace}(0, 1)$.

For our synthetic dataset 2, we model a more complex environment probability. Here, we model $\delta(x) = 0.15\sin(5x) + 0.5$, while keeping the remaining process unchanged, such that we can isolate and show the effect of the environment probability on the performance of the different meta-learners.

To create the simulated data used in Sec. 6, for both datasets, we sample $n = 10000$ from the data-generating process above. We then split the data into train (70%), val (10%), and test (20%) sets.

F. Details regarding real-world data

We perform a case study using a dataset with COVID-19 hospitalizations in Brazil across different regions (Baqui et al., 2020). We are interested in predicting the effect of comorbidity on the mortality of COVID-19 patients. For the environments, we use the regions of the hospitals in Brazil, which are split into North and Central-South. As observed confounders, we include age, sex, and ethnicity. Further, we exclude patients younger than 20 or older than 80 years. To define comorbidity as a binary variable, we define comorbidity as 1 if at least one of the following conditions were diagnosed for the patient: cardiovascular diseases, asthma, diabetes, pulmonary disease, immunosuppression, obesity, liver diseases, neurological disorders, renal disease. We then split the data into train (70%), val (10%), and test (20%) sets. For training, we use the same settings as reported in Appendix D.