Unveiling the Potential of Robustness in Selecting Conditional Average Treatment Effect Estimators

Anonymous Author(s)
Affiliation
Address
email

Abstract

The growing demand for personalized decision-making has led to a surge of interest in estimating the Conditional Average Treatment Effect (CATE). Various types of CATE estimators have been developed with advancements in machine learning and causal inference. However, selecting the desirable CATE estimator through a conventional model validation procedure remains impractical due to the absence of counterfactual outcomes in observational data. Existing approaches for CATE estimator selection, such as plug-in and pseudo-outcome metrics, face two challenges. First, they must determine the metric form and the underlying machine learning models for fitting nuisance parameters (e.g., outcome function, propensity function, and plug-in learner). Second, they lack a specific focus on selecting a robust CATE estimator. To address these challenges, this paper introduces a Distributionally Robust Metric (DRM) for CATE estimator selection. The proposed DRM is nuisance-free, eliminating the need to fit models for nuisance parameters, and it effectively prioritizes the selection of a distributionally robust CATE estimator. The experimental results validate the effectiveness of the DRM method in selecting a CATE estimator that is robust to the distribution shift incurred by covariate shift and unmeasured confounders.

1 Introduction

The escalating demand for decision-making has sparked an increasing interest in *Causal Inference* across various research domains, such as economics [19, 11, 36, 1], statistics [62, 42, 21, 34], healthcare [69, 22, 53, 7, 35], and financial application [9, 12, 31, 17, 20]. The primary goal in personalized decision-making is to quantify the individualized causal effect of a specific treatment (or policy/intervention) on the target outcome, and understanding such causal effects is closely connected with identifying the *Conditional Average Treatment Effect (CATE)*. In observational studies, identifying the CATE inevitably faces a significant challenge due to the absence of *counterfactual* knowledge. According to Rubin Causal Model [56], the CATE is determined by comparing *potential outcomes* under different treatment assignments (i.e., treat and control) for a specific individual. Nonetheless, in real-world applications, we can only observe the potential outcome under the actual treatment (i.e., *factual outcome*), while the potential outcome under the alternative treatment (i.e., *counterfactual outcome*) remains unobserved. The unavailability of the counterfactual outcome is widely recognized as the fundamental problem in causal inference [28], making it difficult to accurately determine the true value of the CATE.

The advancement of machine learning (ML) has opened up a promising opportunity to improve the CATE estimation from observational data. Several innovative CATE estimation approaches, such as meta-learners and causal ML models, have been proposed to tackle the fundamental challenge in causal inference and enhance the predictive accuracy of CATE estimates (as discussed in Section 2). Nevertheless, the emergence of various CATE estimation methods has brought forth a new question:

Submitted to 38th Conference on Neural Information Processing Systems (NeurIPS 2024). Do not distribute.

Given multifarious options for CATE estimators, which should be chosen? Conventional model validation procedures, unfortunately, are not suitable for CATE estimator selection due to the absence of ground truth CATE labels. Therefore, exploring proper metrics for CATE estimator selection remains an essential yet challenging research topic in causal inference.

Recent research has emphasized the significance of model selection for CATE estimators, as highlighted in [58, 16, 45]. These works have proposed and summarized two types of criteria for CATE estimator selection: *plug-in* and *pseudo-outcome* metrics. The large-scale empirical studies have shown that these metrics offer some assistance in identifying well-performing CATE estimators. However, one may still face two challenges when using these metrics for CATE estimator selection, as thoroughly discussed in Section 3.1. First, there is a dilemma in determining the form of evaluation metric and its underlying ML algorithm. Second, these metrics do not prioritize the selection of robust CATE estimators. Given these challenges, we propose a Distributionally Robust Metric (DRM) for CATE estimator selection. The contributions of this paper are summarized as follows.

Contributions. (1) The proposed DRM method is nuisance-free, eliminating the need to fit models for nuisance parameters (outcome function, propensity function, and plug-in learner). (2) The DRM method is designed to prioritize selecting a distributionally robust CATE estimator. (3) We provide a finite sample analysis of the proposed distributionally robust value $\hat{\mathcal{V}}^t(\hat{\tau})$ for $t \in \{0, 1\}$, showing it decays to $\mathcal{V}^t(\hat{\tau})$ at a rate of $n^{-1/2}$. (4) Experimental results validate the effectiveness of the DRM method in selecting a CATE estimator that is robust to the distribution shift incurred by covariate shift and unmeasured confounders.

2 Related Work

CATE estimation. Recent advancements in ML have emerged as powerful tools for estimating CATE from observational data, and researchers pay particular attention to *meta-learners* and *causal ML* models. Existing meta-learners mainly include traditional learners such as S-learner, T-learner, PS-learner, and IPW-learner, as well as new learners such as X-learner [40], DR-learner [34, 21], R-learner [47], and RA-learner [14]. The specific details of these meta-learners are stated in Appendix A.1. Additionally, some studies also focus on developing innovative causal ML models for CATE estimation, such as Causal BART [25], Causal Forest [62, 6, 50], generative models like CEVAE [43] and GANITE [68], representation learning nets including SITE [67], TARNet [60], Dragonnet [61], FlexTENet [15], and HTCE [8], disentangled learning nets like D^{2}VD [37, 38], DeR-CFR [65], and DR-CFR [26], and representation balancing nets such as BNN [32], CFRNet [60], DKLITE [70], IGNITE [24], BWCFR [4], and DRRB [30]. Recent surveys [23, 66, 48] have also conducted a systematic review of various causal inference methods.

CATE estimator selection. Compared to the diverse range of CATE estimation methods, selecting CATE estimators has received limited attention in existing causal inference research. Current methods for selecting CATE estimators can be broadly classified into two main categories. **The first category**, which is also considered in this paper, involves using plug-in and pseudo-outcome methods to evaluate CATE estimators. These methods share two common characteristics: 1) Both methods require fitting ML models for nuisances (e.g., outcome function, propensity function, CATE function) on a validation set and then implementing the learned ML models in either the plug-in surrogate or the pseudo-outcome surrogate; 2) Both methods serve as surrogates for the expected error between the CATE estimator and the true CATE, i.e., $\mathcal{R}^{oracle}(\hat{\tau})$ in equation (1). The difference between the two methods is that the plug-in method directly approximates the true CATE function, where only covariate variables are involved, while the pseudo-outcome method typically constructs a specific formula incorporating covariates, treatment, and outcome variables. For example, a pseudo-DR proposed in [57] is constructed by the outcome predictors learned with representation balancing objective [60, 33]. Recent research [58, 16, 45] has conducted thorough empirical investigations into exploring these two methods for selecting CATE estimators. Their findings suggest that no single selection criterion can universally outperform others in all scenarios in the task of selecting CATE estimators. More details of the two selection methods are stated in Appendix A.2. **The second category** considers leveraging the data generating process (DGP) to generate synthetic data with a known true CATE, allowing the validation of CATE estimators' performance on this synthetic data. For example, authors in [2] find that placebo and structured empirical Monte Carlo methods are helpful for estimator selection under some restrictive conditions. In addition, researchers in [59, 5, 51] focus on training generative models to enforce the generated data to approximate the distribution of the observed data. However, the DGP-based method still faces some limitations in CATE estimator selection due to two key factors: i) it only guarantees the resemblance of the generated data to the factual distribution, without considering the counterfactual distribution, and ii) there is a potential risk of the method favoring estimators that closely resemble the generative models [13].

3 Background of CATE Estimator Selection

Suppose the observational data contain n i.i.d. samples $\{(x_i, t_i, y_i)\}_{i=1}^n$, with the associated random variables being $\{(X_i, T_i, Y_i)\}_{i=1}^n$. For each unit i, $X_i \in \mathcal{X} \subset \mathbb{R}^d$ is d-dimensional covariates and $T_i \in \{0, 1\}$ is the binary treatment. Potential outcomes for treat ($T = 1$) and control ($T = 0$) are denoted by $Y^1, Y^0 \in \mathcal{Y} \subset \mathbb{R}$. The observed (factual) outcome is $Y = TY^1 + (1 - T)Y^0$. The propensity score [55] is defined as $\pi(x) := P(T = 1 \mid X = x)$. The conditional mean potential outcome surface is defined as $\mu_t(x) := \mathbb{E}[Y^t \mid X = x]$ for $t \in \{0, 1\}$. The true CATE is defined as

$$\tau_{true}(x) := \mathbb{E}\left[Y^1 - Y^0 \mid X = x\right] = \mu_1(x) - \mu_0(x).$$

Following the standard and necessary assumptions in potential outcome framework [56], we impose Assumption 3.1 that ensure treatment effects are identifiable.

Assumption 3.1 (Consistency, Overlap, and Unconfoundedness)**.** Consistency: If the treatment is t, then the observed outcome Y equals Y^t. Overlap: The propensity score is bounded away from 0 to 1, i.e., $0 < \pi(x) < 1, \forall x \in \mathcal{X}$. Unconfoundedness [1]: $Y^t \perp\!\!\!\perp T \mid X,\ \forall t \in \{0, 1\}$.

The goal of CATE estimator selection is to select the best CATE estimator, denoted by $\hat{\tau}_{best}$, from a set of J candidate estimators $\{\hat{\tau}_1, \ldots, \hat{\tau}_J\}$:

$$\hat{\tau}_{best} = \underset{\hat{\tau} \in \{\hat{\tau}_1, \ldots, \hat{\tau}_J\}}{\arg\min} \mathcal{R}^{oracle}(\hat{\tau}), \quad \mathcal{R}^{oracle}(\hat{\tau}) := \sqrt{\frac{1}{n}\sum_{i=1}^{n}(\hat{\tau}(X_i) - \tau_{true}(X_i))^2}. \tag{1}$$

Here, $\mathcal{R}^{oracle}(\hat{\tau})$ is associated with $\mathbb{E}[(\hat{\tau}(X) - \tau_{true}(X))^2]$, known as the Precision of Estimating Heterogeneous Effects (PEHE) w.r.t. $\hat{\tau}$ [27, 60]. Note that $\mathcal{R}^{oracle}(\hat{\tau})$ cannot be employed to evaluate CATE estimators' performances in real applications as we do not have access to τ_{true}. Previous studies have introduced plug-in and pseudo-outcome metrics to aid in CATE estimator selection:

$$\mathcal{R}^{plug}_{\tilde{\tau}}(\hat{\tau}) = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(\hat{\tau}(X_i) - \tilde{\tau}(X_i))^2}, \quad \mathcal{R}^{pseudo}_{\tilde{Y}}(\hat{\tau}) = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(\hat{\tau}(X_i) - \tilde{Y}_i)^2}. \tag{2}$$

One can establish a plug-in estimator $\tilde{\tau}$ or construct a pseudo-outcome estimator $\tilde{Y}$ using the validation data. Then, an estimator can be selected based on the following two criteria: $\hat{\tau}_{select} = \arg\min_{\hat{\tau} \in \{\hat{\tau}_1, \ldots, \hat{\tau}_J\}} \mathcal{R}^{plug}_{\tilde{\tau}}(\hat{\tau})$ or $\hat{\tau}_{select} = \arg\min_{\hat{\tau} \in \{\hat{\tau}_1, \ldots, \hat{\tau}_J\}} \mathcal{R}^{pseudo}_{\tilde{Y}}(\hat{\tau})$. Notably, both the plug-in and pseudo-outcome metrics necessitate the fitting of nuisance parameters $\tilde{\eta}$ (e.g., $\tilde{\eta} = (\tilde{\mu}_1, \tilde{\mu}_0, \tilde{\pi})$) using off-the-shelf ML models. For the plug-in metric, $\tilde{\tau}$ can be constructed using any CATE estimator discussed in Appendix A.1, yielding metrics such as plug-T, plug-DR, etc. For the pseudo-outcome metric, $\tilde{Y}$ can be constructed using a specific formula discussed in Appendix A.2, yielding metrics such as pseudo-DR, pseudo-R, etc. In line with [16], the metrics based on the influence function [3] and the R-learner objective [47] are categorized into the pseudo-outcome metric.

3.1 Motivation

The previous high-quality study [16] has provided valuable insights into the advantages and disadvantages of different plug-in and pseudo-outcome metrics, highlighting the need to further explore CATE estimator selection methods. Standing on the shoulders of giants, our paper is motivated by the following two potential challenges faced by existing CATE estimator selection metrics.

The first challenge lies in determining the metric form and underlying ML models for nuisance parameters. As previously discussed, plug-in and pseudo-outcome metrics have various forms, and

[1] Note that in the setting C of our experiments, the unconfoundedness assumption is violated, leading to misspecified nuisance parameters in CATE estimators, plug-in selectors, and pseudo-outcome selectors.

both of them rely on estimating nuisance parameters $\tilde{\eta}$ using ML algorithms such as linear models, tree-based models, etc. Plug-in metrics even need to fit an additional ML model for the plug-in learner $\tilde{\tau}$. However, selecting the suitable metric form and ML algorithms can be very difficult without the knowledge of true data generating process. Consequently, we might go round in circles as this challenge leads us back to the original estimator selection problem [16].

The second challenge is that these metrics are not well-targeted for selecting robust a CATE estimator. In potential outcome framework [56], the factual distribution P^F and the counterfactual distribution P^{CF} for $t \in \{0, 1\}$ can be defined as follows:

$$\begin{aligned} P^F &:= P(X, Y^t | T = t) = P(Y^t | X, T = t) P(X | T = t); \\ P^{CF} &:= P(X, Y^t | T = 1 - t) = P(Y^t | X, T = 1 - t) P(X | T = 1 - t). \end{aligned} \tag{3}$$

The above (3) reveals that the covariate shift $P(X|T = t) \neq P(X|T = 1 - t)$ leads to a distribution shift between P^F and P^{CF} - and such distribution shift can be further exacerbated once the unconfoundedness assumption $P(Y^t|X, T = t) = P(Y^t|X, T = 1-t)$ is violated. It is widely recognized that ML models often struggle when the training and test data do not adhere to the same distribution. Therefore, it becomes essential to select a CATE estimator learned on P^F that demonstrates robust performance to the counterfactual distribution P^{CF}. This need for robustness holds even greater significance than the pursuit of an ideal "stellar" estimator because striving for the perfect estimator can be futile in the absence of ground truth counterfactual labels.

Given the two challenges, our aim is to explore a CATE estimator selection metric that satisfies the following two requirements: ***(1) Nuisance-free***: The metric does not require fitting models for nuisance parameters (outcome function, propensity function, and plug-in learner); ***(2) Robustness***: The metric prioritizes the selection of a CATE estimator that demonstrates robustness to the distribution shift incurred by the covariate shift and unconfoundedness violation.

4 The Distributionally Robust Metric

In this section, we introduce the Distributionally Robust Metric (DRM) for CATE estimator selection. First, we capture the uncertainty in PEHE in a distributionally robust manner (Section 4.1). We then establish the DRM based on the distributionally robust value of PEHE (Section 4.2).

4.1 Capturing the Uncertainty in PEHE

Proposition 4.1. *The PEHE w.r.t. the CATE estimator $\hat{\tau}$ can be decomposed as follows:*

$$\mathbb{E}[(\hat{\tau}(X) - \tau_{true}(X))^2] = \mathbb{E}[\hat{\tau}(X)^2] + 2\mathbb{E}[\hat{\tau}(X)Y^0] + 2\mathbb{E}[-\hat{\tau}(X)Y^1] + \zeta, \tag{4}$$

where $\zeta = \mathbb{E}[(\mu_1(X) - \mu_0(X))^2]$. The proof is deferred to Appendix B.1.

Proposition 4.1 indicates that the PEHE is equal to four terms, where $\mathbb{E}[\hat{\tau}(X)^2]$, $\mathbb{E}[\hat{\tau}(X)Y^0]$, and $\mathbb{E}[-\hat{\tau}(X)Y^1]$ depend on $\hat{\tau}$, while ζ is a constant that is independent of $\hat{\tau}$. The term $\mathbb{E}[\hat{\tau}(X)Y^t]$ for $t \in \{0, 1\}$ can be further decomposed as follows:

$$\mathbb{E}[\hat{\tau}(X)Y^t] = \underbrace{\mathbb{E}[\hat{\tau}(X)Y^t|T = t]}_{\text{(a) Empirically computable}} P(T = t) + \underbrace{\mathbb{E}[\hat{\tau}(X)Y^t|T = 1 - t]}_{\text{(b) Empirically uncomputable}} P(T = 1 - t). \tag{5}$$

Equation (5a) can be computed empirically since the potential outcome Y^t is observable in the group of $T = t$. However, equation (5b) is empirically uncomputable due to the unavailability of Y^t in the group of $T = 1 - t$. The unknown term $\mathbb{E}[\hat{\tau}(X)Y^t|T = 1 - t]$ therefore determines the uncertainty in PEHE. To capture such an uncertainty, we therefore establish distributionally robust values for $\mathbb{E}[\hat{\tau}(X)Y^0|T = 1]$ and $\mathbb{E}[-\hat{\tau}(X)Y^1|T = 0]$ based on a Kullback-Leibler (KL) ambiguity set.

Definition 4.2 (KL ambiguity set)**.** Given two distributions Q and P and the ambiguity radius $\epsilon > 0$. The KL ambiguity (uncertainty) set $\mathcal{B}_\epsilon(P)$ is defined as

$$\mathcal{B}_\epsilon(P) := \{Q : D_{KL}(Q||P) \leq \epsilon\}, \quad \text{where } D_{KL}(Q||P) = \int_{\mathcal{X}} q(x) \log \frac{q(x)}{p(x)} dx. \tag{6}$$

Here, $D_{KL}(Q||P)$ denotes the KL divergence of some arbitrary distribution Q from the reference distribution P. Now we define the distribution of (X, Y^0, Y^1) in the treated and controlled groups as

$$P_T := P(X, Y^0, Y^1|T = 1); \; P_C := P(X, Y^0, Y^1|T = 0). \tag{7}$$

By setting an adequately large ambiguity radius in Definition 4.2, the following inequalities hold for $\mathbb{E}[\hat{\tau}(X)Y^0|T=1] = \mathbb{E}^{P_T}[\hat{\tau}(X)Y^0]$ and $\mathbb{E}[-\hat{\tau}(X)Y^1|T=0] = \mathbb{E}^{P_C}[-\hat{\tau}(X)Y^1]$:

$$\begin{aligned}\mathbb{E}[\hat{\tau}(X)Y^0|T=1] = \mathbb{E}^{P_T}[\hat{\tau}(X)Y^0] &\leq \sup_{Q\in B_{\epsilon_0}(P_C)} \mathbb{E}^{Q}[\hat{\tau}(X)Y^0] =: \mathcal{V}^0(\hat{\tau});\\ \mathbb{E}[-\hat{\tau}(X)Y^1|T=0] = \mathbb{E}^{P_C}[-\hat{\tau}(X)Y^1] &\leq \sup_{Q\in B_{\epsilon_1}(P_T)} \mathbb{E}^{Q}[-\hat{\tau}(X)Y^1] =: \mathcal{V}^1(\hat{\tau}).\end{aligned} \tag{8}$$

To provide a clearer understanding, let us consider the example of $\mathbb{E}^{P_T}[\hat{\tau}(X)Y^0]$. Since the term $\mathbb{E}[\hat{\tau}(X)Y^0]$ is computable on its factual distribution P_C but uncomputable on its counterfactual distribution P_T, we can construct an ambiguity set centered around the distribution P_C such that it is large enough to contain the distribution P_T. By doing so, we can capture the uncertainty of $\mathbb{E}^{P_T}[\hat{\tau}(X)Y^0]$ w.r.t. $\hat{\tau}$. In other words, the value of the uncomputable quantity $\mathbb{E}^{P_T}[\hat{\tau}(X)Y^0]$ will be **at most** $\mathcal{V}^0(\hat{\tau})$. Similarly, the value of the uncomputable quantity $\mathbb{E}^{P_C}[-\hat{\tau}(X)Y^1]$ will be **at most** $\mathcal{V}^1(\hat{\tau})$. Obviously, the uncertainty in PEHE will be larger if the distribution shift between factual and counterfactual distribution is severer. Consequently, we can obtain the distributionally robust value of PEHE in Corollary 4.3, which measures the uncertainty in PEHE.

Corollary 4.3. *Let $\mathcal{V}^0(\hat{\tau})$ and $\mathcal{V}^1(\hat{\tau})$ be the quantities defined in equation* (8)*, ζ be the constant given in Proposition 4.1, $u_1 := P(T=1)$, and $u_0 = 1-u_1 = P(T=0)$. The distributionally robust value of PEHE w.r.t. $\hat{\tau}$ is defined as $\mathcal{V}_{PEHE}(\hat{\tau})$ such that*

$$\begin{aligned}&\mathbb{E}[(\hat{\tau}(X)-\tau_{true}(X))^2] \leq \mathcal{V}_{PEHE}(\hat{\tau})\\ &= \mathbb{E}[\hat{\tau}(X)^2] + 2\left(u_0\mathbb{E}^{P_C}[\hat{\tau}(X)Y^0] + u_1\mathbb{E}^{P_T}[-\hat{\tau}(X)Y^1]\right) + 2\left(u_0\mathcal{V}^1(\hat{\tau}) + u_1\mathcal{V}^0(\hat{\tau})\right) + \zeta.\end{aligned} \tag{9}$$

4.2 Establishing Distributionally Robust Metric

As Corollary 4.3 provides the distributionally robust (worst-case) value of PEHE, it can naturally measure the robustness of the CATE estimator $\hat{\tau}$ against distribution shift between counterfactual distribution and factual distribution. In this section, we will provide two steps involved in using Corollary 4.3 to construct the DRM method for CATE estimator selection.

Step 1: Establishing computational tractability of $\mathcal{V}^t(\hat{\tau})$. The distributionally robust values $\mathcal{V}^0(\hat{\tau})$ and $\mathcal{V}^1(\hat{\tau})$ in equation (9) are initially defined as supremum problems over infinite support, presenting a substantial computational challenge. Theorem 4.4 reformulates the infeasible supremum problems into tractable minimum problems.

Theorem 4.4. *The distributionally robust values $\mathcal{V}^0(\hat{\tau})$ and $\mathcal{V}^1(\hat{\tau})$ in equation* (8) *are equivalent to*

$$\begin{aligned}\mathcal{V}^0(\hat{\tau}) &= \min_{\lambda_0>0} \lambda_0\epsilon_0 + \lambda_0 \log \mathbb{E}^{P_C}[\exp(\hat{\tau}(X)Y^0/\lambda_0)];\\ \mathcal{V}^1(\hat{\tau}) &= \min_{\lambda_1>0} \lambda_1\epsilon_1 + \lambda_1 \log \mathbb{E}^{P_T}[\exp(-\hat{\tau}(X)Y^1/\lambda_1)].\end{aligned} \tag{10}$$

The proof is deferred to Appendix B.3.

In the finite-sample scenario, $\mathcal{V}^0(\hat{\tau})$ and $\mathcal{V}^1(\hat{\tau})$ can be empirically approximated as follows:

$$\begin{aligned}\hat{\mathcal{V}}^0(\hat{\tau}) &= \min_{\lambda_0>0} \lambda_0\epsilon_0 + \lambda_0 \log \frac{1}{n_c}\sum_{i=1}^{n}(1-T_i)\exp(\hat{\tau}(X_i)Y_i/\lambda_0);\\ \hat{\mathcal{V}}^1(\hat{\tau}) &= \min_{\lambda_1>0} \lambda_1\epsilon_1 + \lambda_1 \log \frac{1}{n_t}\sum_{i=1}^{n}T_i\exp(-\hat{\tau}(X_i)Y_i/\lambda_1).\end{aligned} \tag{11}$$

Note that in equation (11), the potential outcomes Y^0 and Y^1 are replaced by the observed outcome Y due to the fact that $(1-T)Y^0 = (1-T)Y$ and $TY^1 = TY$, which aligns with the Consistency assumption in Assumption 3.1. We then provide a finite-sample analysis of the gap between $\hat{\mathcal{V}}^t(\hat{\tau})$ and $\mathcal{V}^t(\hat{\tau})$ in the following Theorem 4.5, which suggests the gap decays at a rate of $n^{-1/2}$.

Theorem 4.5. *Let $u_t := P(T=t)$ for $t\in\{0,1\}$. Assume $0 < \underline{\lambda} \leq \lambda_0, \lambda_1 \leq \bar{\lambda}$ and $\hat{\tau}(X)Y$ is bounded within the range of $\underline{M}$ to $\bar{M}$. Define $C_{exp} = \mathbf{1}_{\{\underline{M}\leq\bar{M}\leq 0\}}\exp\left(\bar{M}/\bar{\lambda} - \underline{M}/\underline{\lambda}\right) + \mathbf{1}_{\{\underline{M}\leq 0,\bar{M}\geq 0\}}\exp\left(\bar{M}/\underline{\lambda} - \underline{M}/\underline{\lambda}\right) + \mathbf{1}_{\{0\leq\underline{M}\leq\bar{M}\}}\exp\left(\bar{M}/\underline{\lambda} - \underline{M}/\bar{\lambda}\right)$. For $n \geq 2/u^2\log(2/\delta)$ and*

Algorithm 1 Using DRM for CATE Estimator Selection

Input: The candidate CATE estimators $\{\hat{\tau}_1, \ldots, \hat{\tau}_J\}$. The validation dataset with n i.i.d. observational samples $\{(X_i, T_i, Y_i)\}_{i=1}^n$. The number of iterations K. The initialization $\lambda_0^{(0)}$ and $\lambda_1^{(0)}$. The ambiguity radius ϵ_0 and ϵ_1.

```
for j = 1 to J do
    for k = 0 to K − 1 do
        Compute F̂_t(λ_t^(k), ε_t; τ̂_j) for t ∈ {0, 1} by equation (13a).
        Compute ∂F̂_t(λ_t^(k), ε_t; τ̂_j)/∂λ_t^(k) for t ∈ {0, 1} by equation (13b).
        λ_t^(k+1) ← max{λ_t^(k) − F̂_t(λ_t^(k), ε_t; τ̂_j)/(∂F̂_t(λ_t^(k), ε_t; τ̂_j)/∂λ_t^(k)), 0} for t ∈ {0, 1}.
        Save V̂^t(τ̂_j)[k] = F̂_t(λ_t^(k+1), ε_t; τ̂_j) for t ∈ {0, 1}.
    Return V̂^t(τ̂_j) = argmin_{k∈{0,...,K−1}} V̂^t(τ̂_j)[k] for t ∈ {0, 1}.
    Use V̂^0(τ̂_j) and V̂^1(τ̂_j) to compute R^DRM(τ̂_j) by equation (14).
```

Output: $\hat{\tau}_{select} = \arg\min_{\hat{\tau}\in\{\hat{\tau}_1,\ldots,\hat{\tau}_J\}} \mathcal{R}^{DRM}(\hat{\tau})$.

$t \in \{0, 1\}$, *with probability* $1 - \delta$, *we have*

$$|\hat{\mathcal{V}}^t(\hat{\tau}) - \mathcal{V}^t(\hat{\tau})| \leq \mathcal{O}\left(\sqrt{\frac{8\bar{\lambda}^2 \log\frac{2}{\delta}}{nu_t^2}C_{exp}^2}\right) + \mathcal{O}\left(\sqrt{\frac{2\bar{\lambda}^2 \log(\frac{2}{\delta})}{nu_t^2}}\right). \tag{12}$$

The proof is deferred to Appendix B.4.

Step 2: Finalizing Distributionally Robust Metric for CATE estimator selection. We first define two functions that are useful in obtaining $\mathcal{V}^0(\hat{\tau})$ and $\mathcal{V}^1(\hat{\tau})$:

$$\hat{F}_0(\lambda_0, \epsilon_0; \hat{\tau}) = \lambda_0\epsilon_0 + \lambda_0 \log \frac{1}{n_c}\sum_{i=1}^{n_c} e^{\frac{Z_i}{\lambda_0}},\ \hat{F}_1(\lambda_1, \epsilon_1; \hat{\tau}) = \lambda_1\epsilon_1 + \lambda_1 \log \frac{1}{n_t}\sum_{i=1}^{n_t} e^{\frac{-Z_i}{\lambda_1}}; \tag{13a}$$

$$\frac{\partial \hat{F}_0}{\partial \lambda_0} = \epsilon_0 + \log\sum_{i=1}^{n_c}\frac{e^{\frac{Z_i}{\lambda_0}}}{n_c} - \frac{\sum_{i=1}^{n_c} Z_i e^{\frac{Z_i}{\lambda_0}}}{\lambda_0\sum_{i=1}^{n_c} e^{\frac{Z_i}{\lambda_0}}},\ \frac{\partial \hat{F}_1}{\partial \lambda_1} = \epsilon_1 + \log\sum_{i=1}^{n_t}\frac{e^{\frac{-Z_i}{\lambda_1}}}{n_t} - \frac{\sum_{i=1}^{n_t} -Z_i e^{\frac{-Z_i}{\lambda_1}}}{\lambda_1\sum_{i=1}^{n_t} e^{\frac{-Z_i}{\lambda_1}}}. \tag{13b}$$

Here, Z denotes $\hat{\tau}(X)Y$ for notational simplicity. We then use the Newton-Raphson method to find the empirical solution for $\hat{\mathcal{V}}^t(\hat{\tau})$, exploiting the convexity of $\hat{F}_t(\lambda_t, \epsilon_t; \hat{\tau})$ w.r.t. λ_t. Based on the distributionally robust value of PEHE, i.e., $\mathcal{V}_{PEHE}(\hat{\tau})$ in equation (9), we finally obtain the selected estimator $\hat{\tau}_{select} = \arg\min_{\hat{\tau}\in\{\hat{\tau}_1,\ldots,\hat{\tau}_J\}} \mathcal{R}^{DRM}(\hat{\tau})$ such that

$$\mathcal{R}^{DRM}(\hat{\tau}) = \frac{1}{n}\sum_{i=1}^{n}\hat{\tau}(X_i)^2 + \frac{2}{n}\left(\sum_{i=1}^{n_c}\hat{\tau}(X_i)Y_i + \sum_{i=1}^{n_t} -\hat{\tau}(X_i)Y_i + n_c\hat{\mathcal{V}}^1(\hat{\tau}) + n_t\hat{\mathcal{V}}^0(\hat{\tau})\right). \tag{14}$$

Algorithm 1 provides complete procedure of using the DRM method for CATE estimator selection.

Discussion on the ambiguity radius ϵ. The ambiguity radius ϵ plays a critical role in real-world applications [46, 44, 52]. However, determining an appropriate value for ϵ can be challenging as it requires striking a balance between ensuring the bound in equation (8) holds and maintaining its tightness. Specifically, if ϵ is set too small, it fails to guarantee that the counterfactual distribution is contained within the ambiguity set centered at factual distribution (the bound in Corollary 4.3 can hold). On the other hand, if ϵ is set too large, even though the ambiguity set can encompass more distributions to ensure the counterfactual distribution is contained, the bound in Corollary 4.3 can be less tight. In general, selecting a proper ambiguity radius is an open problem in distributioanlly robust optimization (DRO) literature [29, 46, 39, 41, 63].

In this paper, we provide useful guidance for determining the ambiguity radius for our DRM method. Based on the above discussion, an ideal radius should be $\epsilon^* = D_{KL}(P_C||P_T)$, which ensures that the bound in Corollary 4.3 holds and is tight. However, as defined in equation (7), both P_C and P_T involve counterfactual information, making it unattainable to directly compute $D_{KL}(P_C||P_T)$. To overcome this challenge, we demonstrate that Proposition 4.6 provides an intriguing alternative approach to acquire $D_{KL}(P_C||P_T)$ when unconfoundedness in Assumption 3.1 is satisfied.

Proposition 4.6. *Let $P_X^T := P(X|T=1)$ and $P_X^C := P(X|T=0)$ denote the covariates distribution in the treat and control group, respectively. Assuming that random variables (X, T, Y^1, Y^0) satisfy the unconfoundedness in Assumption 3.1, we have*

$$D_{KL}(P_C||P_T) = D_{KL}(P_X^C||P_X^T). \tag{15}$$

The proof is deferred to Appendix B.2.

Proposition 4.6 provides an important insight that the uncomputable term $D_{KL}(P_C||P_T)$ can be replaced by a computable quantity $D_{KL}(P_X^C||P_X^T)$, where P_X^C and P_X^T are empirically observable. As a result, we suggest setting the ambiguity radius with $\epsilon^* = D_{KL}(P_X^C||P_X^T)$. Note that while the KL divergence can be approximated with empirical algorithm (e.g, Nearest-Neighbor [64, 49]), the DRM remains nuisance-free, as this serves merely as a means to determine the ambiguity radius and does not involve learning the outcome function, propensity function, or any plug-in learner.

5 Experiments

5.1 Experimental Setup.

Estimators & Selectors. We consider a total of **24 CATE estimators**, comprising the combination of 3 base ML models and 8 meta-learners. Specifically, the chosen base ML models are Linear Regression (LR), Support Vector Machine (SVM), and Random Forests (RF). We consider these ML models for CATE estimators because they are representative of both rigid and flexible models, with each encoded distinct inductive biases, as highlighted by [15, 16]. Note that for the LR method, we employ Ridge regression for regression tasks and Logistic regression for classification tasks. As for the remaining methods, we utilize their corresponding regressors and classifiers for regression and classification tasks, respectively. Regarding the meta-learners, we select a set of both traditional basic learners (S-, T-, PS-, and IPW-learners) and recently developed learners (X-, DR-, R-, and RA-learners), as detailed in Appendix A.1. We consider **13 CATE selectors**, consisting of 8 plug-in methods that rely on the above 8 learners, 3 pseudo-outcome methods (pseudo-DR, -R, and -IF), the random selection, the factual selection (from the 6-learner pool with S-, T-), the Nearest-Neighbor matching [54], and our proposed DRM. The specific details of baseline selectors are stated in Appendix A.2. We employ the eXtreme Gradient Boosting (XGB) [10] as the underlying ML model for both plug-in and pseudo-outcome methods. We choose XGB because: i) it demonstrates superior performance in various scenarios, ensuring a good performance of baseline selectors; ii) the need to avoid potential congeniality bias that may arise from using the similar ML models employed in CATE estimators [16]; iii) aligning with [3] where XGB is used for their proposed pseudo-IF metric. Following [16], we adopt grid search for hyperparameter tuning whenever model training is required.

Dataset. Since the ground truth of CATE is unavailable in real-world data, previous studies commonly utilize semi-synthetic datasets to compare model performance. In line with [15, 16], we collect the covariates with $n = 4802$ data points from ACIC2016 dataset [18]. Then, we generate treatment with $T_i|X_i \sim Bern(1/(1+\exp(-\xi(\beta_T' X_i + 3))))$, where $Bern$ indicates the Bernoulli distribution. The potential outcomes are generated by a linear function with interaction terms:

$$Y_i = \sum_{j}^{d} \beta_j' X_{i;j} + \sum_{j=1}^{d}\sum_{k=j}^{d} \beta_{j,k}' X_{i;j}X_{i;k} + \sum_{j=1}^{d}\sum_{k=j}^{d}\sum_{l=k}^{d} \beta_{j,k,l}' X_{i;j}X_{i;k}X_{i;l} + T_i \sum_{j=1}^{d} \gamma_j X_{i;j} + \epsilon_i.$$

The coefficient values are set as follows: $\beta_T, \beta_j, \beta_{j,k}, \beta_{j,k,l} \sim Bern(0.2)$, $\gamma_j \sim Bern(\rho)$, and $\epsilon_i \sim \mathcal{N}(0, 0.1)$. The parameter ξ in treatment assignment represents the level of selection bias, and the parameter ρ in γ_j represents the complexity of the CATE function. We adopt the above data generating process to randomly generate 100 datasets, each with a training/validation/testing ratio of 49%/21%/30%. All the experiments are run on Dell 3640 with Intel Xeon W-1290P 3.60GHz CPU.

Settings. In this section, we mainly investigate whether the estimator selected by DRM can demonstrate robustness to the selection bias and unobserved confounders. In addition, as demonstrated in [15, 16], the complexity of CATE function also affects relative performance of estimators and selectors. Given these considerations, we design the following three settings to compare the CATE selectors. **Setting A:** With the unconfoundedness assumption, let ρ vary in $\{0, 0.1, 0.3\}$ with fixing $\xi = 1$. **Setting B:** With the unconfoundedness assumption, let ξ vary in $\{0, 1, 2\}$ with fixing $\rho = 0.1$.

Table 1: Comparison of Regret for different selectors across Settings A, B, and C (Note that B ($\xi = 1$) matches A ($\rho = 0.1$)). Reported values (mean $\pm$ standard deviation) are computed over 100 experiments. Bold denotes the best three results among all selectors. Smaller value is better.

	A ($\rho = 0$)	A ($\rho = 0.1$)	A ($\rho = 0.3$)	B ($\xi = 0$)	B ($\xi = 2$)	C ($m = 0.1$)	C ($m = 0.5$)	C ($m = 0.9$)
Plug-S	3.40±5.87	2.88±5.45	2.34±5.16	**2.04**±7.24	8.84±13.90	**3.62**±5.66	6.16±8.19	**11.01**±11.89
Plug-PS	3.40±5.87	2.88±5.45	2.34±5.16	**1.83**±7.00	8.84±13.90	**3.62**±5.66	**6.15**±8.20	**11.01**±11.90
Plug-T	38.12±21.95	36.39±19.95	34.38±19.43	15.19±22.79	45.88±22.91	37.87±20.12	35.50±20.50	32.22±14.27
Plug-X	8.49±7.93	7.38±7.03	6.57±6.72	7.59±12.66	15.13±16.42	10.89±13.44	14.34±17.87	16.84±12.56
Plug-IPW	29.47±23.33	26.57±22.82	23.79±21.12	12.34±25.46	33.60±19.05	33.62±27.27	25.43±21.62	27.08±18.18
Plug-DR	35.48±22.12	34.84±21.53	32.92±19.68	13.03±22.93	44.54±24.01	37.05±20.52	33.85±21.55	29.19±15.79
Plug-R	**1.86**±6.11	**1.27**±5.84	**1.19**±5.71	**0.70**±4.21	**4.38**±8.42	**2.17**±5.50	**2.86**±7.66	**4.18**±8.67
Plug-RA	38.84±21.75	36.89±19.94	34.47±19.30	13.86±22.96	46.46±23.10	37.86±19.85	35.57±20.20	32.62±14.59
Pseudo-DR	38.06±21.82	35.76±20.84	33.36±20.35	15.39±22.32	45.92±23.24	37.14±19.99	34.84±20.42	32.05±15.43
Pseudo-R	**1.21**±3.47	**2.60**±9.89	**1.23**±4.04	4.78±16.03	**7.74**±17.47	4.78±13.33	8.88±12.61	15.97±12.97
Pseudo-IF	32.01±10.68	31.18±11.46	30.09±10.75	17.39±18.28	34.02±13.38	32.49±13.24	29.12±9.10	23.62±6.72
Random	37.44±50.22	37.47±30.92	28.71±31.30	14.45±15.96	42.42±48.37	37.06±33.27	33.25±36.36	30.14±28.29
Fact	39.25±31.25	38.64±30.99	36.62±31.03	5.05±10.18	58.54±40.09	38.00±28.40	36.53±28.55	39.27±27.29
Matching	35.79±18.98	33.89±18.47	32.27±16.86	13.85±20.49	40.92±21.19	34.53±17.60	32.83±15.52	29.43±13.18
DRM	**0.23**±0.97	**0.11**±0.28	**0.13**±0.44	2.73±16.30	**1.21**±7.93	**0.80**±6.24	**0.32**±0.76	**1.27**±2.04

Setting C: Without unconfoundedness assumption, fix $\rho = 0.1$ and $\xi = 1$. Then randomly remove $\lfloor m \cdot d \rfloor$ covariates such that the dimension of observed covariates is $d - \lfloor m \cdot d \rfloor$, where m denotes the ratio of missing covariates varying in $\{0.1, 0.5, 0.9\}$.

Comparison criteria. The CATE estimator $\hat{\tau}$ is believed better if it achieves a smaller difference between $\mathcal{R}^{oracle}(\hat{\tau})$ and $\mathcal{R}^{oracle}(\hat{\tau}_{best})$, where $\hat{\tau}_{best}$ is the actual best estimator in equation (1). We therefore use the following Regret criteria to compare estimators chosen by different selectors:

$$\text{Regret} = \mathcal{R}^{oracle}(\hat{\tau}_{select}) - \mathcal{R}^{oracle}(\hat{\tau}_{best}).$$

To further assess the ranking ability of each selector, we calculate the Spearman rank correlation between the rank order determined by the oracle metric $\mathcal{R}^{oracle}(\hat{\tau})$ and the rank order determined by each selector. All the reported values (Mean $\pm$ Standard deviation) are computed over 100 runs.

5.2 Experimental Results

Regret comparison. Based on the results presented in Table 1, we observe consistent good performance from both DRM and Plug-R across various settings. Specifically, in setting A, the DRM selector consistently outperforms other selectors as the CATE complexity (ρ) increases. Additionally, Table 1 suggests that Plug-R performs well in terms of the Regret criterion, aligning with the findings in [58] that R-objective is excellent in many cases. Additionally, we also make a comparison of PEHE performance (i.e., $\mathcal{R}^{oracle}(\hat{\tau}_{select})$) of different selectors in Table 3 of Section C.1. The result indicates that both DRM and Plug-R tend to exhibit better performance in terms of PEHE as the CATE complexity decreases, which aligns with the result in [16] that the R-based objective achieves better PEHE performance as the CATE complexity decreases. Moving on to setting B, we observe that DRM demonstrates significant robustness against selection bias (controlled by ξ). Notably, the advantage of the DRM selector becomes more pronounced when the level of selection bias is strong ($\xi = 2$). In this case, all baseline selectors, except for Plug-R and Pseudo-R, exhibit large Regret. In the scenario $\xi = 0$ where no selection bias is present, there is distribution shift between factual distribution and counterfactual distribution. Consequently, the factual selection criterion performs better in this specific setting compared to others. However, despite its good performance in this case, it is not the best selector among all selectors as it excludes most CATE estimators from the candidate pool. Simultaneously, as we would expect, DRM does not demonstrate significant advantage in this case, since there is no distribution shift caused by selection bias. Considering setting C where the unconfoundedness assumption is violated, we observe that most selectors exhibit inferior performance. In contrast, DRM demonstrates consistent outperformance across all three cases, and its superiority becomes particularly significant as m increases to 0.9. This showcases the robustness of DRM against the distribution shift arising from unobserved confounders.

Ranking ability. In Table 2, the DRM method demonstrates favorable performance in ranking estimators, surpassing certain Plug- (e.g., T, IPW, DR, RA) and Pseudo- (e.g., DR, IF) selectors. In comparison to other nuisance-free baselines (Random, Fact, and Matching), DRM achieves

Table 2: Comparison of rank correlation for different selectors across Settings A, B, and C (Note that B ($\xi = 1$) matches A ($\rho = 0.1$)). Bold denotes the best three results among all selectors. Reported values (mean $\pm$ standard deviation) are computed over 100 experiments. Larger is better.

	A ($\rho = 0$)	A ($\rho = 0.1$)	A ($\rho = 0.3$)	B ($\xi = 0$)	B ($\xi = 2$)	C ($m = 0.1$)	C ($m = 0.5$)	C ($m = 0.9$)
Plug-S	**0.95**$\pm$0.06	**0.95**$\pm$0.06	**0.95**$\pm$0.06	**0.88**$\pm$0.09	**0.92**$\pm$0.09	**0.94**$\pm$0.05	**0.90**$\pm$0.13	**0.84**$\pm$0.18
Plug-PS	**0.95**$\pm$0.06	**0.95**$\pm$0.06	**0.95**$\pm$0.06	**0.88**$\pm$0.09	**0.92**$\pm$0.09	**0.94**$\pm$0.05	**0.90**$\pm$0.13	**0.84**$\pm$0.18
Plug-T	0.30$\pm$0.35	0.30$\pm$0.35	0.29$\pm$0.35	0.66$\pm$0.21	0.26$\pm$0.36	0.27$\pm$0.30	0.27$\pm$0.39	0.22$\pm$0.40
Plug-X	0.89$\pm$0.08	0.89$\pm$0.13	0.88$\pm$0.14	0.78$\pm$0.15	0.89$\pm$0.09	0.86$\pm$0.11	0.80$\pm$0.18	0.70$\pm$0.26
Plug-IPW	0.58$\pm$0.31	0.58$\pm$0.31	0.59$\pm$0.31	0.75$\pm$0.20	0.49$\pm$0.29	0.50$\pm$0.31	0.57$\pm$0.30	0.54$\pm$0.31
Plug-DR	0.38$\pm$0.35	0.37$\pm$0.36	0.36$\pm$0.35	0.74$\pm$0.16	0.30$\pm$0.36	0.31$\pm$0.31	0.36$\pm$0.39	0.36$\pm$0.40
Plug-R	**0.96**$\pm$0.08	**0.96**$\pm$0.08	**0.96**$\pm$0.08	**0.88**$\pm$0.08	**0.95**$\pm$0.10	**0.96**$\pm$0.05	**0.95**$\pm$0.07	**0.92**$\pm$0.13
Plug-RA	0.30$\pm$0.35	0.30$\pm$0.35	0.29$\pm$0.35	0.70$\pm$0.18	0.26$\pm$0.36	0.26$\pm$0.30	0.27$\pm$0.39	0.23$\pm$0.40
Pseudo-DR	0.32$\pm$0.35	0.31$\pm$0.35	0.31$\pm$0.35	0.65$\pm$0.22	0.28$\pm$0.35	0.27$\pm$0.29	0.29$\pm$0.38	0.23$\pm$0.41
Pseudo-R	0.94$\pm$0.05	0.92$\pm$0.17	0.94$\pm$0.05	0.83$\pm$0.12	0.87$\pm$0.20	0.89$\pm$0.18	0.85$\pm$0.13	0.72$\pm$0.25
Pseudo-IF	0.42$\pm$0.31	0.41$\pm$0.31	0.40$\pm$0.31	0.36$\pm$0.35	0.52$\pm$0.32	0.40$\pm$0.29	0.41$\pm$0.30	0.52$\pm$0.33
Random	-0.18$\pm$0.12	-0.19$\pm$0.13	-0.17$\pm$0.11	-0.18$\pm$0.13	-0.21$\pm$0.14	-0.18$\pm$0.13	-0.16$\pm$0.14	-0.21$\pm$0.17
Fact	-0.03$\pm$0.10	-0.03$\pm$0.10	-0.03$\pm$0.10	-0.04$\pm$0.08	-0.07$\pm$0.12	-0.02$\pm$0.10	-0.04$\pm$0.11	-0.11$\pm$0.14
Matching	0.30$\pm$0.30	0.29$\pm$0.29	0.28$\pm$0.29	0.68$\pm$0.19	0.27$\pm$0.30	0.29$\pm$0.26	0.28$\pm$0.33	0.28$\pm$0.36
DRM	0.85$\pm$0.09	0.85$\pm$0.08	0.85$\pm$0.08	0.86$\pm$0.10	0.80$\pm$0.14	0.85$\pm$0.09	0.87$\pm$0.07	**0.84**$\pm$0.10

significantly superior ranking ability. However, compared to Plug-S, -PS, and -R, it does not exhibit remarkable performance in ranking CATE estimators, possibly due to the fact that DRM selects estimators based on their distributionally robust (worst-case) performance. Indeed, the definition of ranking inherently involves the concept of expected (average) performance, which is not determined solely by either the best or worst performance. While distributionally robust performance serves as a suitable criterion for selecting players to participate in the Olympics, it may not be a reasonable standard for ranking players' average performance. Therefore, it would be intriguing to explore some ways in future research that can enhance the ranking ability of our DRM selector.

Additional experiments. We also conduct analysis for examining the best and worst performance of each selector. Specifically, in each of the 100 experiments, we sort all 24 estimators in ascending order based on their $\mathcal{R}^{oracle}(\hat{\tau})$ values, resulting in the sorted list: $[\mathcal{R}^{oracle}(\hat{\tau}_1), \ldots, \mathcal{R}^{oracle}(\hat{\tau}_J)]$. We then determine the actual rank of the selected estimator within this list and visualize the distribution of these 100 ranks using a stacked bar chart (Figure 1 of Appendix C.1). The results reveal that DRM is able to select higher-ranked estimators while mitigate the risk of selecting lower-ranked estimators. Additionally, we conduct similar comparisons when the candidate pool comprises 8 candidate estimators in C.2, C.3, and C.4 of Appendix, with the underlying ML models of each comparison fixed as LR, SVM, and RF, respectively.

6 Conclusion

This paper sheds lights on the potential of robustness in CATE estimator selection. We propose a distributionally robust metric (DRM). The proposed metric is nuisance-free, eliminating the need to fit models for nuisance parameters (outcome function, propensity function, and plug-in learner). Additionally, it is well-targeted for selecting a robust CATE estimator. We provide a finite sample analysis that demonstrates the gap between $\hat{\mathcal{V}}^t(\hat{\tau})$ and $\mathcal{V}^t(\hat{\tau})$ reduces at a rate of $n^{-1/2}$ for $t \in \{0, 1\}$. The experimental results showcase that the CATE estimator selected by DRM demonstrate robustness to the distribution shift incurred by covariate shift and unconfoundedness violation.

Limitations. This paper uncovers the potential of robustness in CATE estimator selection. However, we acknowledge that our DRM method is not a one-size-fits-all solution and still faces several challenges that should be addressed in future research. For instance, compared to baseline selectors, our method does not exhibit a significant advantage in cases where the CATE function is simple and there is no selection bias. If one already knows that there is no (or low) selection bias in observational data, we recommend using the plug-R or factual metric for CATE estimator selection. Further, as mentioned in Section 5.2, enhancing the ranking ability of DRM is also an intriguing avenue for further exploration. Moreover, while our results are based on KL-divergence, considering the ambiguity set constructed with other divergence such as Wasserstein may contain more diverse distributions [29, 46, 39, 41, 63]. We hope our methods and results will stimulate increased attention towards CATE estimator selection and provide valuable insights for future insightful research.

References

[1] Alberto Abadie, Susan Athey, Guido W Imbens, and Jeffrey M Wooldridge. When should you adjust standard errors for clustering? *The Quarterly Journal of Economics*, 138(1):1–35, 2023.

[2] Arun Advani, Toru Kitagawa, and Tymon Słoczyński. Mostly harmless simulations? using monte carlo studies for estimator selection. *Journal of Applied Econometrics*, 34(6):893–910, 2019.

[3] Ahmed Alaa and Mihaela Van Der Schaar. Validating causal inference models via influence functions. In *International Conference on Machine Learning*, pages 191–201. PMLR, 2019.

[4] Serge Assaad, Shuxi Zeng, Chenyang Tao, Shounak Datta, Nikhil Mehta, Ricardo Henao, Fan Li, and Lawrence Carin. Counterfactual representation learning with balancing weights. In *International Conference on Artificial Intelligence and Statistics*, pages 1972–1980. PMLR, 2021.

[5] Susan Athey, Guido W Imbens, Jonas Metzger, and Evan Munro. Using wasserstein generative adversarial networks for the design of monte carlo simulations. *Journal of Econometrics*, 2021.

[6] SUSAN ATHEY, JULIE TIBSHIRANI, and STEFAN WAGER. Generalized random forests. *The Annals of Statistics*, 47(2):1148–1178, 2019.

[7] Ioana Bica, Ahmed M Alaa, Craig Lambert, and Mihaela Van Der Schaar. From real-world patient data to individualized treatment effects using machine learning: current and future methods to address underlying challenges. *Clinical Pharmacology & Therapeutics*, 109(1):87–100, 2021.

[8] Ioana Bica and Mihaela van der Schaar. Transfer learning on heterogeneous feature spaces for treatment effects estimation. *Advances in Neural Information Processing Systems*, 35:37184–37198, 2022.

[9] Léon Bottou, Jonas Peters, Joaquin Quiñonero-Candela, Denis X Charles, D Max Chickering, Elon Portugaly, Dipankar Ray, Patrice Simard, and Ed Snelson. Counterfactual reasoning and learning systems: The example of computational advertising. *Journal of Machine Learning Research*, 14(11), 2013.

[10] Tianqi Chen and Carlos Guestrin. Xgboost: A scalable tree boosting system. In *Proceedings of the 22nd acm sigkdd international conference on knowledge discovery and data mining*, pages 785–794, 2016.

[11] Victor Chernozhukov, Denis Chetverikov, Mert Demirer, Esther Duflo, Christian Hansen, Whitney Newey, and James Robins. Double/debiased machine learning for treatment and structural parameters, 2018.

[12] Zhixuan Chu, Stephen L Rathbun, and Sheng Li. Graph infomax adversarial learning for treatment effect estimation with networked observational data. In *Proceedings of the 27th ACM SIGKDD Conference on Knowledge Discovery & Data Mining*, pages 176–184, 2021.

[13] Alicia Curth, David Svensson, Jim Weatherall, and Mihaela van der Schaar. Really doing great at estimating cate? a critical look at ml benchmarking practices in treatment effect estimation. In *Thirty-fifth conference on neural information processing systems datasets and benchmarks track (round 2)*, 2021.

[14] Alicia Curth and Mihaela van der Schaar. Nonparametric estimation of heterogeneous treatment effects: From theory to learning algorithms. In *International Conference on Artificial Intelligence and Statistics*, pages 1810–1818. PMLR, 2021.

[15] Alicia Curth and Mihaela van der Schaar. On inductive biases for heterogeneous treatment effect estimation. *Advances in Neural Information Processing Systems*, 34:15883–15894, 2021.

[16] Alicia Curth and Mihaela Van Der Schaar. In search of insights, not magic bullets: Towards demystification of the model selection dilemma in heterogeneous treatment effect estimation. In *Proceedings of the 40th International Conference on Machine Learning*, volume 202 of *Proceedings of Machine Learning Research*, pages 6623–6642. PMLR, 23–29 Jul 2023.

[17] Robert Donnelly, David Blei, Susan Athey, et al. Correction to: Counterfactual inference for consumer choice across many product categories. *Quantitative Marketing and Economics*, 19(3-4):409–409, 2021.

[18] Vincent Dorie, Jennifer Hill, Uri Shalit, Marc Scott, and Dan Cervone. Automated versus do-it-yourself methods for causal inference: Lessons learned from a data analysis competition. *Statistical Science*, 34(1):43–68, 2019.

[19] Max H Farrell. Robust inference on average treatment effects with possibly more covariates than observations. *Journal of Econometrics*, 189(1):1–23, 2015.

[20] Carlos Fernández-Loría, Foster Provost, Jesse Anderton, Benjamin Carterette, and Praveen Chandar. A comparison of methods for treatment assignment with an application to playlist generation. *Information Systems Research*, 34(2):786–803, 2023.

[21] Dylan J Foster and Vasilis Syrgkanis. Orthogonal statistical learning. *The Annals of Statistics*, 51(3):879–908, 2023.

[22] Jared C Foster, Jeremy MG Taylor, and Stephen J Ruberg. Subgroup identification from randomized clinical trial data. *Statistics in medicine*, 30(24):2867–2880, 2011.

[23] Ruocheng Guo, Lu Cheng, Jundong Li, P Richard Hahn, and Huan Liu. A survey of learning causality with data: Problems and methods. *ACM Computing Surveys (CSUR)*, 53(4):1–37, 2020.

[24] Ruocheng Guo, Jundong Li, Yichuan Li, K Selçuk Candan, Adrienne Raglin, and Huan Liu. Ignite: A minimax game toward learning individual treatment effects from networked observational data. In *Proceedings of the Twenty-Ninth International Conference on International Joint Conferences on Artificial Intelligence*, pages 4534–4540, 2021.

[25] P Richard Hahn, Jared S Murray, and Carlos M Carvalho. Bayesian regression tree models for causal inference: Regularization, confounding, and heterogeneous effects (with discussion). *Bayesian Analysis*, 15(3):965–1056, 2020.

[26] Negar Hassanpour and Russell Greiner. Learning disentangled representations for counterfactual regression. In *International Conference on Learning Representations*, 2019.

[27] Jennifer L Hill. Bayesian nonparametric modeling for causal inference. *Journal of Computational and Graphical Statistics*, 20(1):217–240, 2011.

[28] Paul W Holland. Statistics and causal inference. *Journal of the American statistical Association*, 81(396):945–960, 1986.

[29] Zhaolin Hu and L Jeff Hong. Kullback-leibler divergence constrained distributionally robust optimization. *Available at Optimization Online*, 1(2):9, 2013.

[30] Yiyan Huang, Cheuk Hang Leung, Shumin Ma, Zhiri Yuan, Qi Wu, Siyi Wang, Dongdong Wang, and Zhixiang Huang. Towards balanced representation learning for credit policy evaluation. In *International Conference on Artificial Intelligence and Statistics*, pages 3677–3692. PMLR, 2023.

[31] Yiyan Huang, Cheuk Hang Leung, Xing Yan, Qi Wu, Nanbo Peng, Dongdong Wang, and Zhixiang Huang. The causal learning of retail delinquency. In *Proceedings of the AAAI Conference on Artificial Intelligence*, volume 35, pages 204–212, 2021.

[32] Fredrik Johansson, Uri Shalit, and David Sontag. Learning representations for counterfactual inference. In *International conference on machine learning*, pages 3020–3029. PMLR, 2016.

[33] Fredrik D Johansson, Uri Shalit, Nathan Kallus, and David Sontag. Generalization bounds and representation learning for estimation of potential outcomes and causal effects. *The Journal of Machine Learning Research*, 23(1):7489–7538, 2022.

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

[35] Newton Mwai Kinyanjui and Fredrik D Johansson. Adcb: An alzheimer's disease simulator for benchmarking observational estimators of causal effects. In *Conference on Health, Inference, and Learning*, pages 103–118. PMLR, 2022.

[36] Toru Kitagawa and Aleksey Tetenov. Who should be treated? empirical welfare maximization methods for treatment choice. *Econometrica*, 86(2):591–616, 2018.

[37] Kun Kuang, Peng Cui, Bo Li, Meng Jiang, Shiqiang Yang, and Fei Wang. Treatment effect estimation with data-driven variable decomposition. In *Proceedings of the AAAI Conference on Artificial Intelligence*, volume 31, 2017.

[38] Kun Kuang, Peng Cui, Hao Zou, Bo Li, Jianrong Tao, Fei Wu, and Shiqiang Yang. Data-driven variable decomposition for treatment effect estimation. *IEEE Transactions on Knowledge and Data Engineering*, 34(5):2120–2134, 2020.

[39] Daniel Kuhn, Peyman Mohajerin Esfahani, Viet Anh Nguyen, and Soroosh Shafieezadeh-Abadeh. Wasserstein distributionally robust optimization: Theory and applications in machine learning. In *Operations research & management science in the age of analytics*, pages 130–166. Informs, 2019.

[40] Sören R Künzel, Jasjeet S Sekhon, Peter J Bickel, and Bin Yu. Metalearners for estimating heterogeneous treatment effects using machine learning. *Proceedings of the national academy of sciences*, 116(10):4156–4165, 2019.

[41] Daniel Levy, Yair Carmon, John C Duchi, and Aaron Sidford. Large-scale methods for distributionally robust optimization. *Advances in Neural Information Processing Systems*, 33:8847–8860, 2020.

[42] Shuangning Li and Stefan Wager. Random graph asymptotics for treatment effect estimation under network interference. *The Annals of Statistics*, 50(4):2334–2358, 2022.

[43] Christos Louizos, Uri Shalit, Joris M Mooij, David Sontag, Richard Zemel, and Max Welling. Causal effect inference with deep latent-variable models. *Advances in neural information processing systems*, 30, 2017.

[44] Shumin Ma, Cheuk Hang Leung, Qi Wu, Wei Liu, and Nanbo Peng. Understanding distributional ambiguity via non-robust chance constraint. In *Proceedings of the First ACM International Conference on AI in Finance*, pages 1–8, 2020.

[45] Divyat Mahajan, Ioannis Mitliagkas, Brady Neal, and Vasilis Syrgkanis. Empirical analysis of model selection for heterogenous causal effect estimation. *International Conference on Learning Representations*, 2024.

[46] Peyman Mohajerin Esfahani and Daniel Kuhn. Data-driven distributionally robust optimization using the wasserstein metric: performance guarantees and tractable reformulations. *Mathematical Programming*, 171(1-2):115–166, 2018.

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

[48] Ana Rita Nogueira, Andrea Pugnana, Salvatore Ruggieri, Dino Pedreschi, and João Gama. Methods and tools for causal discovery and causal inference. *Wiley interdisciplinary reviews: data mining and knowledge discovery*, 12(2):e1449, 2022.

[49] Yung-Kyun Noh, Masashi Sugiyama, Song Liu, Marthinus C Plessis, Frank Chongwoo Park, and Daniel D Lee. Bias reduction and metric learning for nearest-neighbor estimation of kullback-leibler divergence. In *Artificial Intelligence and Statistics*, pages 669–677. PMLR, 2014.

[50] Miruna Oprescu, Vasilis Syrgkanis, and Zhiwei Steven Wu. Orthogonal random forest for causal inference. In *International Conference on Machine Learning*, pages 4932–4941. PMLR, 2019.

[51] Harsh Parikh, Carlos Varjao, Louise Xu, and Eric Tchetgen Tchetgen. Validating causal inference methods. In *International Conference on Machine Learning*, pages 17346–17358. PMLR, 2022.

[52] Georg Ch Pflug. Multistage stochastic decision problems: Approximation by recursive structures and ambiguity modeling. *European Journal of Operational Research*, 306(3):1027–1039, 2023.

[53] Zhaozhi Qian, Yao Zhang, Ioana Bica, Angela Wood, and Mihaela van der Schaar. Synctwin: Treatment effect estimation with longitudinal outcomes. *Advances in Neural Information Processing Systems*, 34:3178–3190, 2021.

[54] Craig A Rolling and Yuhong Yang. Model selection for estimating treatment effects. *Journal of the Royal Statistical Society Series B: Statistical Methodology*, 76(4):749–769, 2014.

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

[56] Donald B Rubin. Causal inference using potential outcomes: Design, modeling, decisions. *Journal of the American Statistical Association*, 100(469):322–331, 2005.

[57] Yuta Saito and Shota Yasui. Counterfactual cross-validation: Stable model selection procedure for causal inference models. In *International Conference on Machine Learning*, pages 8398–8407. PMLR, 2020.

[58] Alejandro Schuler, Michael Baiocchi, Robert Tibshirani, and Nigam Shah. A comparison of methods for model selection when estimating individual treatment effects. *arXiv preprint arXiv:1804.05146*, 2018.

[59] Alejandro Schuler, Ken Jung, Robert Tibshirani, Trevor Hastie, and Nigam Shah. Synth-validation: Selecting the best causal inference method for a given dataset. *arXiv preprint arXiv:1711.00083*, 2017.

[60] Uri Shalit, Fredrik D Johansson, and David Sontag. Estimating individual treatment effect: generalization bounds and algorithms. In *International conference on machine learning*, pages 3076–3085. PMLR, 2017.

[61] Claudia Shi, David Blei, and Victor Veitch. Adapting neural networks for the estimation of treatment effects. *Advances in neural information processing systems*, 32, 2019.

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

[63] Jie Wang, Rui Gao, and Yao Xie. Sinkhorn distributionally robust optimization. *arXiv preprint arXiv:2109.11926*, 2021.

[64] Qing Wang, Sanjeev R Kulkarni, and Sergio Verdú. A nearest-neighbor approach to estimating divergence between continuous random vectors. In *2006 IEEE International Symposium on Information Theory*, pages 242–246. IEEE, 2006.

[65] Anpeng Wu, Junkun Yuan, Kun Kuang, Bo Li, Runze Wu, Qiang Zhu, Yueting Zhuang, and Fei Wu. Learning decomposed representations for treatment effect estimation. *IEEE Transactions on Knowledge and Data Engineering*, 35(5):4989–5001, 2022.

[66] Liuyi Yao, Zhixuan Chu, Sheng Li, Yaliang Li, Jing Gao, and Aidong Zhang. A survey on causal inference. *ACM Transactions on Knowledge Discovery from Data (TKDD)*, 15(5):1–46, 2021.

[67] Liuyi Yao, Sheng Li, Yaliang Li, Mengdi Huai, Jing Gao, and Aidong Zhang. Representation learning for treatment effect estimation from observational data. *Advances in neural information processing systems*, 31, 2018.

[68] Jinsung Yoon, James Jordon, and Mihaela Van Der Schaar. Ganite: Estimation of individualized treatment effects using generative adversarial nets. In *International conference on learning representations*, 2018.

[69] Linying Zhang, Yixin Wang, Anna Ostropolets, Jami J Mulgrave, David M Blei, and George Hripcsak. The medical deconfounder: assessing treatment effects with electronic health records. In *Machine Learning for Healthcare Conference*, pages 490–512. PMLR, 2019.

[70] Yao Zhang, Alexis Bellot, and Mihaela Schaar. Learning overlapping representations for the estimation of individualized treatment effects. In *International Conference on Artificial Intelligence and Statistics*, pages 1005–1014. PMLR, 2020.

Appendix

A CATE Estimation Strategies

A.1 CATE Learners

We now detail how to construct CATE learners using the observed samples $\{(X_i, T_i, Y_i)\}_{i=1}^n$. Note that CATE learners are learned on the training set, so the sample size n here equals the training sample size. Denote n_t by the sample size in the treat group, and n_c by the sample size in the control group such that $n = n_t + n_c$.

- S-learner: Let predictors=(X, T), response=Y. Train a model $\hat{\mu}(X, T)$. Then we obtain $\hat{\tau}_S(X)$:

$$\hat{\tau}_S(X) = \hat{\mu}(X, 1) - \hat{\mu}(X, 0).$$

- T-learner: Let predictors=X^T (covariates in the treat), response=Y^T (outcome in the treat). Train a model $\hat{\mu}_1(X)$. Let predictors=X^C (covariates in the control), response=Y^C (outcome in the control). Train a model $\hat{\mu}_0(X)$. Then we obtain $\hat{\tau}_T(X)$:

$$\hat{\tau}_T(X) = \hat{\mu}_1(X) - \hat{\mu}_0(X).$$

- PS-learner: Fisrt-step: Train $\hat{\tau}_S(X)$ using the above-mentioned step in S-learner. Second-step: Let predictors=X, response=$\hat{\tau}_S(X)$. Train a model $\hat{\tau}_{PS}(X)$ from the following objective:

$$\hat{\tau}_{PS} = \arg\min_{\tau} \ \frac{1}{n}\sum_{i=1}^{n}(\tau(X_i) - \hat{\tau}_S(X_i))^2.$$

- IPW-learner: First-step: let predictors=X, response=T. Train a propensity score model $\hat{\pi}(X)$. Construct surrogate of CATE using pseudo-outcomes with inverse propensity weighting (IPW) formula: $Y_{IPW}^{1,0} = Y_{IPW}^1 - Y_{IPW}^0$, where $Y_{IPW}^1 = \frac{TY}{\hat{\pi}(X)}$ and $Y_{IPW}^0 = \frac{(1-T)Y}{1-\hat{\pi}(X)}$. Train a model $\hat{\tau}_{IPW}(X)$ from the following objective:

$$\hat{\tau}_{IPW} = \arg\min_{\tau} \ \frac{1}{n}\sum_{i=1}^{n}(\tau(X_i) - Y_{i,IPW}^{1,0})^2.$$

- X-learner [40]: First-step: Train $\hat{\mu}_1(X)$ and $\hat{\mu}_0(X)$ using the the above-mentioned procedure in T-learner. Train a propensity score model $\hat{\pi}(X)$ using the the above-mentioned procedure in IPW-learner. Second-step: Let predictors=X^T, response=$\hat{\mu}_1(X^T) - Y^T$, and predictors=X^C, response=$\hat{\mu}_0(X^C) - Y^C$. Obtain a model $\hat{\tau}_X(X)$ by learning two separate functions $\hat{\tau}_X^1(X)$ and $\hat{\tau}_X^0(X)$:

$$\begin{aligned}
\hat{\tau}_X(X) &= (1 - \hat{\pi}(X))\hat{\tau}_X^1(X) + \hat{\pi}(X)\hat{\tau}_X^0(X),\\
\hat{\tau}_X^1 &= \arg\min_{\tau} \ \frac{1}{n_t}\sum_{i=1}^{n_t}(\tau(X_i) - (Y_i - \hat{\mu}_0(X_i)))^2,\\
\hat{\tau}_X^0 &= \arg\min_{\tau} \ \frac{1}{n_c}\sum_{i=1}^{n_c}(\tau(X_i) - (\hat{\mu}_1(X_i) - Y_i))^2.
\end{aligned}$$

- DR-learner [34, 21]: First-step: Train $\hat{\mu}_1(X)$ and $\hat{\mu}_0(X)$ using the the above-mentioned procedure in T-learner. Train a propensity score model $\hat{\pi}(X)$ using the the above-mentioned procedure in IPW-learner. Second-step: Construct surrogate of CATE using pseudo-outcomes with doubly robust (DR) formula: $Y_{DR}^{1,0} = Y_{DR}^1 - Y_{DR}^0$, where $Y_{DR}^1 = \hat{\mu}_1(X) + \frac{T}{\hat{\pi}(X)}(Y - \hat{\mu}_1(X))$ and $Y_{DR}^0 = \hat{\mu}_0(X) + \frac{1-T}{1-\hat{\pi}(X)}(Y - \hat{\mu}_0(X))$. Train a model $\hat{\tau}_{DR}(X)$ from the following objective:

$$\hat{\tau}_{DR} = \arg\min_{\tau} \frac{1}{n}\sum_{i=1}^{n}(\tau(X_i) - Y_{i,DR}^{1,0})^2.$$

- R-learner [47]: First-step: Let predictors=X, response=Y. Train a model $\hat{\mu}(X)$ to approximate the conditional mean outcome $\mathbb{E}[Y|X]$. Train a propensity score model $\hat{\pi}(X)$ using the the above-mentioned procedure in IPW-learner. Second-step: Compute the outcome residual $\xi = Y - \hat{\mu}(X)$ and treatment residual $\nu = T - \hat{\pi}(X)$. Train a model $\hat{\tau}_R(X)$ from the following objective:

$$\hat{\tau}_{R} = \arg\min_{\tau} \frac{1}{n}\sum_{i=1}^{n}(\xi_i - \nu_i\tau(X_i))^2.$$

- RA-learner [14]: First-step: Train $\hat{\mu}_1(X)$ and $\hat{\mu}_0(X)$ using the the above-mentioned procedure in T-learner. Second-step: Construct surrogate of CATE using pseudo-outcomes with regression adjustment (RA) formula: $Y_{RA} = T(Y - \hat{\mu}_0(X)) + (1-T)(\hat{\mu}_1(X) - Y)$. Train a model $\hat{\tau}_{RA}(X)$ from the following objective:

$$\hat{\tau}_{RA} = \arg\min_{\tau} \frac{1}{n}\sum_{i=1}^{n}(\tau(X_i) - Y_{i,RA})^2.$$

A.2 CATE Selectors

We now detail how to construct CATE selectors using the observed samples $\{(X_i, T_i, Y_i)\}_{i=1}^n$. Note that CATE selectors are constructed on the validation set, so the sample size n here equals the validation sample size.

- Plug-in selector: Obtain any CATE learners $\tilde{\tau}$ using the observational validation data. Then plug-in $\tilde{\tau}$ into the following metric $\mathcal{R}_{\tilde{\tau}}^{plug}(\hat{\tau})$:

$$\mathcal{R}_{\tilde{\tau}}^{plug}(\hat{\tau}) = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(\hat{\tau}(X_i) - \tilde{\tau}(X_i))^2}.$$

  For each plug-in selector $\tilde{\tau}$, the selected j^*-th CATE estimator is $\hat{\tau}_{j^*}$, where $j^* = \arg\min_{j\in\{1,\dots,J\}} \mathcal{R}_{\tilde{\tau}}^{plug}(\hat{\tau}_j)$.
- Pseudo-outcome selector:
  1. Pseudo-DR: Utilize validation data to estimate nuisance parameters $(\tilde{\mu}_1, \tilde{\mu}_0, \tilde{\pi})$, following the procedure described in Section A.1. $\tilde{Y}_{DR} = \tilde{Y}_{DR}^1 - \tilde{Y}_{DR}^0$, where $\tilde{Y}_{DR}^1 = \tilde{\mu}_1(X) + \frac{T}{\tilde{\pi}(X)}(Y - \tilde{\mu}_1(X))$ and $\tilde{Y}_{DR}^0 = \tilde{\mu}_0(X) + \frac{1-T}{1-\tilde{\pi}(X)}(Y - \tilde{\mu}_0(X))$. Then the pseudo-DR metric is

$$\mathcal{R}_{DR}^{pseudo}(\hat{\tau}) = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(\hat{\tau}(X_i) - \tilde{Y}_{i,DR})^2}.$$

     For pseudo-DR selector, the selected j^*-th CATE estimator is $\hat{\tau}_{j^*}$, where $j^* = \arg\min_{j\in\{1,\dots,J\}} \mathcal{R}_{DR}^{pseudo}(\hat{\tau}_j)$.
  2. Pseudo-R: Utilize validation data to estimate nuisance parameters $(\tilde{\mu}, \tilde{\pi})$, following the procedure described in Section A.1. Then the pseudo-R metric is

$$\mathcal{R}_{R}^{pseudo}(\hat{\tau}) = \sqrt{\frac{1}{n}\sum_{i=1}^{n}((Y_i - \tilde{\mu}(X_i)) - \hat{\tau}(X_i)(T_i - \tilde{\pi}(X_i)))^2}.$$

For pseudo-R selector, the selected j^*-th CATE estimator is $\hat{\tau}_{j^*}$, where $j^* = \arg\min_{j\in\{1,\dots,J\}} \mathcal{R}_R^{pseudo}(\hat{\tau}_j)$.

3. Pseudo-IF [3]: Utilize validation data to estimate nuisance parameters $(\tilde{\mu}_1, \tilde{\mu}_0, \tilde{\pi})$, following the procedure described in Section A.1. Let $\tilde{\tau}(X) = (\tilde{\mu}_1(X) - \tilde{\mu}_0(X))$. Then the pseudo-IF metric is

$$\mathcal{R}_{IF}^{pseudo}(\hat{\tau}) = \sqrt{\frac{1}{n}\sum_{i=1}^{n}((1-B_i)\tilde{\tau}^2(X_i) + B_iY_i(\tilde{\tau}(X_i) - \hat{\tau}(X_i)) - A_i(\tilde{\tau}(X_i) - \hat{\tau}(X_i))^2 + \hat{\tau}^2(X_i))},$$

$$where\ A_i = T_i - \tilde{\pi}(X_i),\ B_i = 2T_i(T_i - \tilde{\pi}(X_i))C_i^{-1},\ C_i = \tilde{\pi}(X_i)(1 - \tilde{\pi}(X_i)).$$

For pseudo-IF selector, the selected j^*-th CATE estimator is $\hat{\tau}_{j^*}$, where $j^* = \arg\min_{j\in\{1,\dots,J\}} \mathcal{R}_{IF}^{pseudo}(\hat{\tau}_j)$.

4. Other pseudo-outcome selector: By manipulating the formula of $\tilde{Y}$, it is possible to create additional pseudo-outcome selectors, such as the pseudo-IPW selector. In our paper, we choose pseudo-DR as the baseline because it is representative in the causal inference literature and it often demonstrates superior performance, owing to its doubly robust property.

B Proofs

B.1 Proof of Proposition 4.1

Proof.

$$\begin{aligned}
&\mathbb{E}[(\hat{\tau}(X) - \tau_{true}(X))^2] \\
&= \mathbb{E}[(\hat{\tau}(X) - (\mu_1(X) - \mu_0(X)))^2] \\
&= \mathbb{E}[(\hat{\tau}(X) - \mu_1(X) + \mu_0(X))^2] \\
&= \mathbb{E}[(\hat{\tau}(X) - \mu_1(X))^2] + \mathbb{E}[\mu_0(X)^2] + 2\mathbb{E}[(\hat{\tau}(X) - \mu_1(X))\mu_0(X)] \\
&= \mathbb{E}[\hat{\tau}(X)^2] + \mathbb{E}[\mu_1(X)^2] - 2\mathbb{E}[\hat{\tau}(X)\mu_1(X)] + \mathbb{E}[\mu_0(X)^2] + 2\mathbb{E}[\hat{\tau}(X)\mu_0(X)] - 2\mathbb{E}[\mu_1(X)\mu_0(X)] \\
&= \mathbb{E}[\hat{\tau}(X)^2] - 2\mathbb{E}[\hat{\tau}(X)(\mu_1(X) - Y^1 + Y^1)] + 2\mathbb{E}[\hat{\tau}(X)(\mu_0(X) - Y^0 + Y^0)] \\
&\quad + \mathbb{E}[\mu_1(X)^2] + \mathbb{E}[\mu_0(X)^2] - 2\mathbb{E}[\mu_1(X)\mu_0(X)] \\
&= \mathbb{E}[\hat{\tau}(X)^2] - 2\mathbb{E}[\hat{\tau}(X)Y^1] - 2\mathbb{E}[\hat{\tau}(X)(\mu_1(X) - Y^1)] + 2\mathbb{E}[\hat{\tau}(X)Y^0] + 2\mathbb{E}[\hat{\tau}(X)(\mu_0(X) - Y^0)] \\
&\quad + \mathbb{E}[\mu_1(X)^2] + \mathbb{E}[\mu_0(X)^2] - 2\mathbb{E}[\mu_1(X)\mu_0(X)] \\
&= \mathbb{E}[\hat{\tau}(X)^2] - 2\mathbb{E}[\hat{\tau}(X)Y^1] - 2\mathbb{E}[\mathbb{E}[\hat{\tau}(X)\mu_1(X) - \hat{\tau}(X)Y^1|X]] + 2\mathbb{E}[\hat{\tau}(X)Y^0] \\
&\quad + 2\mathbb{E}[\mathbb{E}[\hat{\tau}(X)\mu_0(X) - \hat{\tau}(X)Y^0|X]] + \mathbb{E}[\mu_1(X)^2] + \mathbb{E}[\mu_0(X)^2] - 2\mathbb{E}[\mu_1(X)\mu_0(X)] \\
&= \mathbb{E}[\hat{\tau}(X)^2] - 2\mathbb{E}[\hat{\tau}(X)Y^1] - 2\mathbb{E}[\hat{\tau}(X)\mu_1(X) - \hat{\tau}(X)\mathbb{E}[Y^1|X]] + 2\mathbb{E}[\hat{\tau}(X)Y^0] \\
&\quad + 2\mathbb{E}[\hat{\tau}(X)\mu_0(X) - \hat{\tau}(X)\mathbb{E}[Y^0|X]] + \mathbb{E}[\mu_1(X)^2] + \mathbb{E}[\mu_0(X)^2] - 2\mathbb{E}[\mu_1(X)\mu_0(X)] \\
&= \mathbb{E}[\hat{\tau}(X)^2] - 2\mathbb{E}[\hat{\tau}(X)Y^1] - 2\mathbb{E}[\hat{\tau}(X)\mu_1(X) - \hat{\tau}(X)\mu_1(X)] + 2\mathbb{E}[\hat{\tau}(X)Y^0] \\
&\quad + 2\mathbb{E}[\hat{\tau}(X)\mu_0(X) - \hat{\tau}(X)\mu_0(X)] + \mathbb{E}[\mu_1(X)^2] + \mathbb{E}[\mu_0(X)^2] - 2\mathbb{E}[\mu_1(X)\mu_0(X)] \\
&= \mathbb{E}[\hat{\tau}(X)^2] + 2\mathbb{E}[\hat{\tau}(X)Y^0] - 2\mathbb{E}[\hat{\tau}(X)Y^1] + \mathbb{E}[\mu_1(X)^2] + \mathbb{E}[\mu_0(X)^2] - 2\mathbb{E}[\mu_1(X)\mu_0(X)] \\
&= \mathbb{E}[\hat{\tau}(X)^2] + 2\mathbb{E}[\hat{\tau}(X)Y^0] - 2\mathbb{E}[\hat{\tau}(X)Y^1] + \zeta.
\end{aligned}$$

□

B.2 Proof of Proposition 4.6

The following Proposition B.1 is useful in proving Proposition 4.6.

Proposition B.1. *Assuming the random variable tuple (X, T, Y^1, Y^0) satisfies Assumption 3.1, we have*

$$\begin{aligned}
p(X, Y^0, Y^1|T = 0) = p(Y^0, Y^1|X)p(X|T = 0); \\
p(X, Y^0, Y^1|T = 1) = p(Y^0, Y^1|X)p(X|T = 1).
\end{aligned} \tag{16}$$

Proof.

$$\begin{aligned}
&p(X, Y^0, Y^1|T=0)\\
=&p(Y^0, Y^1|X, T=0)p(X|T=0)\\
=&p(Y^0, Y^1|X)p(X|T=0). \quad \text{(Unconfoundedness)}\\
&p(X, Y^0, Y^1|T=1)\\
=&p(Y^0, Y^1|X, T=1)p(X|T=1)\\
=&p(Y^0, Y^1|X)p(X|T=1). \quad \text{(Unconfoundedness)}
\end{aligned}$$

□

Now we can prove Proposition 4.6.

Proof.

$$\begin{aligned}
&D_{KL}(P_C||P_T)\\
=&D_{KL}(P(X, Y^0, Y^1|T=0)||P(X, Y^0, Y^1|T=1))\\
=&\int_{\mathcal{X}}\int_{\mathcal{Y}^0}\int_{\mathcal{Y}^1} p(x, y^0, y^1|T=0)\log\frac{p(x, y^0, y^1|T=0)}{p(x, y^0, y^1|T=1)}dy^1dy^0dx\\
=&\int_{\mathcal{X}}\int_{\mathcal{Y}^0}\int_{\mathcal{Y}^1} p(y^0, y^1|x)p(x|T=0)\log\frac{p(y^0, y^1|x)p(x|T=0)}{p(y^0, y^1|x)p(x|T=1)}dy^1dy^0dx \quad \text{(By Proposition B.1)}\\
=&\int_{\mathcal{X}}\int_{\mathcal{Y}^0}\int_{\mathcal{Y}^1} p(y^0, y^1|x)p(x|T=0)\log\frac{p(x|T=0)}{p(x|T=1)}dy^1dy^0dx\\
=&\int_{\mathcal{X}}\left(\int_{\mathcal{Y}^0}\int_{\mathcal{Y}^1} p(y^0, y^1|x)dy^1dy^0\right)p(x|T=0)\log\frac{p(x|T=0)}{p(x|T=1)}dx\\
=&\int_{\mathcal{X}} p(x|T=0)\log\frac{p(x|T=0)}{p(x|T=1)}dx\\
=&D_{KL}(P(X|T=0)||P(X|T=1))\\
=&D_{KL}(P_X^C||P_X^T).
\end{aligned}$$

Similarly, it is easy to show $D_{KL}(P_T||P_C) = D_{KL}(P_X^T||P_X^C)$ □

B.3 Proof of Theorem 4.4

Lemma B.2 (Theorem 1 in [29]). *Let $f_\theta(X)$ denote the loss function of X and it is bounded almost surely. $\theta \in \Theta$ represents the model parameters of the function $f_\theta(X)$. Let $\mathcal{B}_\epsilon(P)$ be the uncertainty ball centered at distribution P with ambiguity radius ϵ. Define κ as the mass of the distribution P on its essential supremum (Proposition 2 in [29]). Assume $f_\theta(X)$ is bounded and* $\log\kappa + \epsilon < 0$*, then we have*

$$\mathcal{V} := \sup_{Q\in\mathcal{B}_\epsilon(P)} \mathbb{E}^Q[f_\theta(X)] = \min_{\lambda>0} \lambda\epsilon + \lambda\log\mathbb{E}^P[\exp(f_\theta(X)/\lambda)].$$

Our Theorem 4.4 follows by directly applying the above Lemma B.2.

B.4 Proof of Theorem 4.5

For notational simplicity, we denote $W = (X, T, Y) \in \mathcal{W}$ and $Z = \hat{\tau}(X)Y$. Assume Z is bounded within the range $\underline{M}$ and $\bar{M}$. Define the following functions:

$$\begin{aligned}
G_0(\lambda_0; W) = \mathbb{E}[g_0(\lambda_0; W)], \quad \hat{G}_0(\lambda_0; W) = \frac{1}{n}\sum_{i=1}^n g_0(\lambda_0; W_i),&\\
\text{where } \ g_0(\lambda_0; W) = (1-T)\exp\left(Z/\lambda_0\right);&\\
G_1(\lambda_1; W) = \mathbb{E}[g_1(\lambda_1; W)], \quad \hat{G}_1(\lambda_1; W) = \frac{1}{n}\sum_{i=1}^n g_1(\lambda_1; W_i),&\\
\text{where } \ g_1(\lambda_1; W) = T\exp\left(-Z/\lambda_1\right).&
\end{aligned}$$

Then we have the following lemma that guarantees the convergence for $\hat{G}_0(\lambda_0; W)$ and $\hat{G}_1(\lambda_1; W)$.

Lemma B.3. *Assume $0 < \underline{\lambda} \leq \lambda_0, \lambda_1 \leq \bar{\lambda}$, and $\hat{\tau}(X)Y$ is bounded within the range of $\underline{M}$ to $\bar{M}$. Then with probability $1-\delta$, we have*

$$
\begin{aligned}
&\textit{If } \underline{M} \leq \bar{M} \leq 0: \\
&|\hat{G}_0(\lambda_0; W) - G_0(\lambda_0; W)| \leq \mathcal{O}\left(\sqrt{\frac{2\log\frac{2}{\delta}\left(\exp\left(\bar{M}/\bar{\lambda}\right)\right)^2}{n}}\right); \\
&|\hat{G}_1(\lambda_1; W) - G_1(\lambda_1; W)| \leq \mathcal{O}\left(\sqrt{\frac{2\log\frac{2}{\delta}\left(\exp\left(-\underline{M}/\underline{\lambda}\right)\right)^2}{n}}\right). \\
&\textit{If } \underline{M} \leq 0, \bar{M} \geq 0: \\
&|\hat{G}_0(\lambda_0; W) - G_0(\lambda_0; W)| \leq \mathcal{O}\left(\sqrt{\frac{2\log\frac{2}{\delta}\left(\exp\left(\bar{M}/\underline{\lambda}\right)\right)^2}{n}}\right); \\
&|\hat{G}_1(\lambda_1; W) - G_1(\lambda_1; W)| \leq \mathcal{O}\left(\sqrt{\frac{2\log\frac{2}{\delta}\left(\exp\left(-\underline{M}/\underline{\lambda}\right)\right)^2}{n}}\right). \\
&\textit{If } 0 \leq \underline{M} \leq \bar{M}: \\
&|\hat{G}_0(\lambda_0; W) - G_0(\lambda_0; W)| \leq \mathcal{O}\left(\sqrt{\frac{2\log\frac{2}{\delta}\left(\exp\left(\bar{M}/\underline{\lambda}\right)\right)^2}{n}}\right); \\
&|\hat{G}_1(\lambda_1; W) - G_1(\lambda_1; W)| \leq \mathcal{O}\left(\sqrt{\frac{2\log\frac{2}{\delta}\left(\exp\left(-\underline{M}/\bar{\lambda}\right)\right)^2}{n}}\right).
\end{aligned}
\tag{17}
$$

Proof. Denote $h_0(W_1, W_2, \ldots, W_n) = \frac{1}{n}\sum_{i=1}^n g_0(\lambda_0; W_i)$. We notice that $h_0(W_1, W_2, \ldots, W_n)$ satisfies the bounded difference inequality:

$$
\begin{aligned}
&\sup_{W_1,\ldots,W_n,W_i'\in\mathcal{W}} |h_0(W_1,\ldots,W_i,\cdots,W_n) - h_0(W_1,\ldots,W_i',\cdots,W_n)| \\
&= \sup_{W_i,W_i'\in\mathcal{W}} \frac{|g_0(\lambda_0;W_i) - g_0(\lambda_0;W_i')|}{n} \\
&\leq 2\sup_{W_i\in\mathcal{W}} \frac{|g_0(\lambda_0;W_i)|}{n} \leq \frac{2\exp\left(\bar{M}/\lambda_0\right)}{n}.
\end{aligned}
$$

Note that $|\hat{G}_0(\lambda_0; W) - G_0(\lambda_0; W)| = |h_0(W_1, W_2, \ldots, W_n) - \mathbb{E}[h_0(W_1, W_2, \ldots, W_n)]|$. Then using McDiarmid's inequality, for any $\epsilon > 0$, we have

$$
\begin{aligned}
&P\left(\left|\hat{G}_0(\lambda_0;W) - G_0(\lambda_0;W)\right| \geq \epsilon\right) \\
&= P\left(|h_0(W_1,W_2,\ldots,W_n) - \mathbb{E}[h_0(W_1,W_2,\ldots,W_n)]| \geq \epsilon\right) \\
&\leq 2\exp\left(-\frac{2\epsilon^2}{n(\frac{2\exp(\bar{M}/\lambda_0)}{n})^2}\right) = 2\exp\left(\frac{-n\epsilon^2}{2\left(\exp\left(\bar{M}/\lambda_0\right)\right)^2}\right).
\end{aligned}
$$

For some $\delta > 0$, we have

$$
P\left(\left|\hat{G}_0(\lambda_0;W) - G_0(\lambda_0;W)\right| \geq \epsilon\right) \leq 2\exp\left(\frac{-n\epsilon^2}{2\left(\exp\left(\bar{M}/\lambda_0\right)\right)^2}\right) \leq \delta.
$$

This solves ϵ such that

$$
\epsilon \geq \sqrt{\frac{2\log\frac{2}{\delta}\left(\exp\left(\bar{M}/\lambda_0\right)\right)^2}{n}}.
$$

The above inequality should hold for any λ_0 such that $0 < \underline{\lambda} \leq \lambda_0 \leq \bar{\lambda}$. Therefore, we have

$$\text{If } \bar{M} \geq 0: \quad \epsilon \geq \sqrt{\frac{2\log\frac{2}{\delta}\left(\exp\left(\bar{M}/\underline{\lambda}\right)\right)^2}{n}};$$
$$\text{If } \bar{M} \leq 0: \quad \epsilon \geq \sqrt{\frac{2\log\frac{2}{\delta}\left(\exp\left(\bar{M}/\bar{\lambda}\right)\right)^2}{n}}.$$

Similarly, denote $h_1(W_1, W_2, \dots, W_n) = \frac{1}{n}\sum_{i=1}^{n} g_1(\lambda_1; W_i)$. We note that $h_1(W_1, W_2, \dots, W_n)$ satisfies the bounded difference inequality:

$$\begin{aligned}
&\sup_{W_1,\dots,W_n,W_i' \in \mathcal{W}} |h_1(W_1, \dots, W_i, \cdots, W_n) - h_1(W_1, \dots, W_i', \cdots, W_n)| \\
&= \sup_{W_i, W_i' \in \mathcal{W}} \frac{|g_1(\lambda_1; W_i) - g_1(\lambda_1; W_i')|}{n} \\
&\leq 2 \sup_{W_i \in \mathcal{W}} \frac{|g_1(\lambda_1; W_i)|}{n} \leq \frac{2\exp\left(-\underline{M}/\lambda_1\right)}{n}.
\end{aligned}$$

Then using McDiarmid's inequality, for any $\epsilon > 0$, we have

$$\begin{aligned}
&P\left(\left|\hat{G}_1(\lambda_1; W) - G_1(\lambda_1; W)\right| \geq \epsilon\right) \\
&= P\left(|h_1(W_1, W_2, \dots, W_n) - \mathbb{E}[h_1(W_1, W_2, \dots, W_n)]| \geq \epsilon\right) \\
&\leq 2\exp\left(-\frac{2\epsilon^2}{n(\frac{2\exp(-\underline{M}/\lambda_1)}{n})^2}\right) = 2\exp\left(\frac{-n\epsilon^2}{2\left(\exp\left(-\underline{M}/\lambda_1\right)\right)^2}\right).
\end{aligned}$$

For some $\delta > 0$, we have

$$P\left(\left|\hat{G}_1(\lambda_1; W) - G_1(\lambda_1; W)\right| \geq \epsilon\right) \leq 2\exp\left(\frac{-n\epsilon^2}{2\left(\exp\left(-\underline{M}/\lambda_1\right)\right)^2}\right) \leq \delta.$$

This solves ϵ such that

$$\epsilon \geq \sqrt{\frac{2\log\frac{2}{\delta}\left(\exp\left(-\underline{M}/\lambda_1\right)\right)^2}{n}}.$$

The above inequality should hold for any λ_1 such that $0 < \underline{\lambda} \leq \lambda_1 \leq \bar{\lambda}$. Therefore, we have

$$\text{If } \underline{M} \geq 0: \quad \epsilon \geq \sqrt{\frac{2\log\frac{2}{\delta}\left(\exp\left(-\underline{M}/\bar{\lambda}\right)\right)^2}{n}};$$
$$\text{If } \underline{M} \leq 0: \quad \epsilon \geq \sqrt{\frac{2\log\frac{2}{\delta}\left(\exp\left(-\underline{M}/\underline{\lambda}\right)\right)^2}{n}}.$$

□

In the following content, we will bound terms $\left|\log(\hat{G}_0(\lambda_0; W)) - \log\left(G_0(\lambda_0; W)\right)\right|$ and $\left|\log(\hat{G}_1(\lambda_1; W)) - \log\left(G_1(\lambda_1; W)\right)\right|$. Lemma B.4 is useful for bounding these two terms.

Lemma B.4. *Let c be a constant. For any x_1, x_2 such that $x_1, x_2 \geq c > 0$, we have*

$$|\log(x_1) - \log(x_2)| \leq \frac{1}{c}|x_1 - x_2| \tag{18}$$

Proof. Without loss of generality, assume $0 < c \leq x_1 \leq x_2$. We then have

$$\log(x_2) - \log(x_1) = \log(\frac{x_2}{x_1}) = \log(1 + \frac{x_2}{x_1} - 1) \leq \frac{x_2}{x_1} - 1 = \frac{x_2 - x_1}{x_1} \leq \frac{x_2 - x_1}{c}.$$

Taking the absolute value of both the left-hand side and the right-hand side, we have

$$|\log(x_1) - \log(x_2)| \leq \frac{1}{c}|x_1 - x_2|.$$

□

Next, we introduce Lemma B.5 that bounds terms $\left|\log(\hat{G}_0(\lambda_0;W)) - \log\left(G_0(\lambda_0;W)\right)\right|$ and $\left|\log(\hat{G}_1(\lambda_1;W)) - \log\left(G_1(\lambda_1;W)\right)\right|$.

Lemma B.5. *Let u denote the probability of treat, i.e., $u = P(T=1)$. Assume that $\lambda_0, \lambda_1 \in \Lambda := [\underline{\lambda}, \bar{\lambda}]$ and $\hat{\tau}(X)Y$ is bounded within $\underline{M}$ and $\bar{M}$. Then for $n \geq \max\{\frac{2}{u^2}\log\left(\frac{2}{\delta}\right), \frac{2}{(1-u)^2}\log\left(\frac{2}{\delta}\right)\}$, with probability $1-\delta$, we have*

$$\begin{aligned}
&\textit{If } \underline{M} \leq \bar{M} \leq 0: \\
&\left|\log(\hat{G}_0(\lambda_0;W)) - \log\left(G_0(\lambda_0;W)\right)\right| \leq \frac{2}{\exp(\underline{M}/\underline{\lambda})(1-u)}\left|\hat{G}_0(\lambda_0;W) - G_0(\lambda_0;W)\right|; \\
&\left|\log(\hat{G}_1(\lambda_1;W)) - \log\left(G_1(\lambda_1;W)\right)\right| \leq \frac{2}{\exp(-\bar{M}/\bar{\lambda})u}\left|\hat{G}_1(\lambda_1;W) - G_1(\lambda_1;W)\right|. \\
&\textit{If } \underline{M} \leq 0, \bar{M} \geq 0: \\
&\left|\log(\hat{G}_0(\lambda_0;W)) - \log\left(G_0(\lambda_0;W)\right)\right| \leq \frac{2}{\exp(\underline{M}/\underline{\lambda})(1-u)}\left|\hat{G}_0(\lambda_0;W) - G_0(\lambda_0;W)\right|; \\
&\left|\log(\hat{G}_1(\lambda_1;W)) - \log\left(G_1(\lambda_1;W)\right)\right| \leq \frac{2}{\exp(-\bar{M}/\underline{\lambda})u}\left|\hat{G}_1(\lambda_1;W) - G_1(\lambda_1;W)\right|. \\
&\textit{If } 0 \leq \underline{M} \leq \bar{M}: \\
&\left|\log(\hat{G}_0(\lambda_0;W)) - \log\left(G_0(\lambda_0;W)\right)\right| \leq \frac{2}{\exp(\underline{M}/\bar{\lambda})(1-u)}\left|\hat{G}_0(\lambda_0;W) - G_0(\lambda_0;W)\right|; \\
&\left|\log(\hat{G}_1(\lambda_1;W)) - \log\left(G_1(\lambda_1;W)\right)\right| \leq \frac{2}{\exp(-\bar{M}/\underline{\lambda})u}\left|\hat{G}_1(\lambda_1;W) - G_1(\lambda_1;W)\right|.
\end{aligned} \tag{19}$$

Proof. First, we bound the term $\left|\log(\hat{G}_0(\lambda_0;W)) - \log\left(G_0(\lambda_0;W)\right)\right|$.

$G_0(\lambda_0;W)$ and $\hat{G}_0(\lambda_0;W)$ are greater than 0 and bounded because $Z = \hat{\tau}(X)Y$ is bounded within the range $\underline{M}$ and $\bar{M}$. Therefore, applying Lemma B.4, we have

$$\left|\log(\hat{G}_0(\lambda_0;W)) - \log\left(G_0(\lambda_0;W)\right)\right| \leq \frac{1}{c}\left|\hat{G}_0(\lambda_0;W) - G_0(\lambda_0;W)\right|,$$
$$\text{where } c = \min\left\{\inf_{\lambda_0\in\Lambda, W\in\mathcal{W}} \hat{G}_0(\lambda_0;W), \inf_{\lambda_0\in\Lambda, W\in\mathcal{W}} G_0(\lambda_0;W)\right\}.$$

Moreover, for any $\lambda_0 \in \Lambda$, we have

$$\begin{aligned}
\text{If } \underline{M} \geq 0: \quad G_0(\lambda_0;W) &= \mathbb{E}[(1-T)\exp(Z/\lambda_0)] = \mathbb{E}[\exp(Z/\lambda_0)|T=0]P(T=0) \\
&\geq \mathbb{E}[\exp(\underline{M}/\bar{\lambda})|T=0](1-u) = \exp(\underline{M}/\bar{\lambda})(1-u); \\
\hat{G}_0(\lambda_0;W) &= \frac{1}{n}\sum_{i=1}^{n}(1-T_i)\exp(Z_i/\lambda_0) \\
&\geq \frac{1}{n}\sum_{i=1}^{n}(1-T_i)\exp(\underline{M}/\bar{\lambda}) = \exp(\underline{M}/\bar{\lambda})(1-\hat{u}).
\end{aligned} \tag{20}$$

$$\begin{aligned}
\text{If } \underline{M} \leq 0: \quad G_0(\lambda_0;W) &= \mathbb{E}[(1-T)\exp(Z/\lambda_0)] = \mathbb{E}[\exp(Z/\lambda_0)|T=0]P(T=0) \\
&\geq \mathbb{E}[\exp(\underline{M}/\underline{\lambda})|T=0](1-u) = \exp(\underline{M}/\underline{\lambda})(1-u); \\
\hat{G}_0(\lambda_0;W) &= \frac{1}{n}\sum_{i=1}^{n}(1-T_i)\exp(Z_i/\lambda_0) \\
&\geq \frac{1}{n}\sum_{i=1}^{n}(1-T_i)\exp(\underline{M}/\underline{\lambda}) = \exp(\underline{M}/\underline{\lambda})(1-\hat{u}).
\end{aligned} \tag{21}$$

Given $\hat{u} = \frac{1}{n}\sum_{i=1}^{n} T_i$ and $u = \mathbb{E}[\frac{1}{n}\sum_{i=1}^{n} T_i]$, using Hoeffding's inequality, we have

$$P\left(\left|\frac{1}{n}\sum_{i=1}^{n}(1-T_i) - \mathbb{E}[\frac{1}{n}\sum_{i=1}^{n}(1-T_i)]\right| \geq \frac{\mathbb{E}[\frac{1}{n}\sum_{i=1}^{n}(1-T_i)]}{2}\right) \leq 2\exp\left(-\frac{2(\frac{1-u}{2})^2}{n(\frac{1}{n})^2}\right) \leq \delta.$$

We can solve n by

$$2\exp\left(-\frac{n(1-u)^2}{2}\right) \leq \delta \Rightarrow n \geq \frac{2}{(1-u)^2}\log\left(\frac{2}{\delta}\right).$$

This indicates that $(1-\hat{u}) \geq (1-u)/2$ with probability $1-\delta$ when $n \geq \frac{2}{(1-u)^2}\log\left(\frac{2}{\delta}\right)$. Combining this with equations (20) and (21), with probability $1-\delta$, when $n \geq \frac{2}{(1-u)^2}\log\left(\frac{2}{\delta}\right)$, we have

$$\begin{aligned}
\text{If } \ \underline{M} \geq 0: \quad & \inf_{\lambda_0\in\Lambda, W\in\mathcal{W}} G_0(\lambda_0; W) \geq \exp(\underline{M}/\bar{\lambda})(1-u); \\
& \inf_{\lambda_0\in\Lambda, W\in\mathcal{W}} \hat{G}_0(\lambda_0; W) \geq \exp(\underline{M}/\bar{\lambda})(1-\hat{u}) \geq \exp(\underline{M}/\bar{\lambda})(1-u)/2. \\
\text{If } \ \underline{M} \leq 0: \quad & \inf_{\lambda_0\in\Lambda, W\in\mathcal{W}} G_0(\lambda_0; W) \geq \exp(\underline{M}/\underline{\lambda})(1-u); \\
& \inf_{\lambda_0\in\Lambda, W\in\mathcal{W}} \hat{G}_0(\lambda_0; W) \geq \exp(\underline{M}/\underline{\lambda})(1-\hat{u}) \geq \exp(\underline{M}/\underline{\lambda})(1-u)/2.
\end{aligned}$$

Therefore, with probability $1-\delta$, when $n \geq \frac{2}{(1-u)^2}\log\left(\frac{2}{\delta}\right)$, we have

$$\begin{aligned}
& \text{If } \ \underline{M} \geq 0: \\
& \left|\log(\hat{G}_0(\lambda_0; W)) - \log\left(G_0(\lambda_0; W)\right)\right| \leq \frac{2}{\exp(\underline{M}/\bar{\lambda})(1-u)}\left|\hat{G}_0(\lambda_0; W) - G_0(\lambda_0; W)\right|; \\
& \text{If } \ \underline{M} \leq 0: \\
& \left|\log(\hat{G}_0(\lambda_0; W)) - \log\left(G_0(\lambda_0; W)\right)\right| \leq \frac{2}{\exp(\underline{M}/\underline{\lambda})(1-u)}\left|\hat{G}_0(\lambda_0; W) - G_0(\lambda_0; W)\right|.
\end{aligned}$$

Next, we bound the term $\left|\log(\hat{G}_1(\lambda_1; W)) - \log\left(G_1(\lambda_1; W)\right)\right|$. $G_1(\lambda_1; W)$ and $\hat{G}_1(\lambda_1; W)$ are greater than 0 and bounded above. Therefore, applying Lemma B.4, we have

$$\begin{aligned}
& \left|\log(\hat{G}_1(\lambda_1; W)) - \log\left(G_1(\lambda_1; W)\right)\right| \leq \frac{1}{c}\left|\hat{G}_1(\lambda_1; W) - G_1(\lambda_1; W)\right|, \\
& \text{where } c = \min\left\{\inf_{\lambda_1\in\Lambda, W\in\mathcal{W}} \hat{G}_1(\lambda_1; W), \inf_{\lambda_1\in\Lambda, W\in\mathcal{W}} G_1(\lambda_1; W)\right\}.
\end{aligned}$$

Moreover, for any $\lambda_1 \in \Lambda$, we have

$$\begin{aligned}
\text{If } \ \bar{M} \geq 0: \quad G_1(\lambda_1; W) &= \mathbb{E}[T\exp(-Z/\lambda_1)] = \mathbb{E}[\exp(-Z/\lambda_1)|T=1]P(T=1) \\
&\geq \mathbb{E}[\exp(-\bar{M}/\underline{\lambda})|T=1]u = \exp(-\bar{M}/\underline{\lambda})u; \\
\hat{G}_1(\lambda_1; W) &= \frac{1}{n}\sum_{i=1}^{n} T_i\exp(-Z_i/\lambda_1) \qquad\qquad (22) \\
&\geq \frac{1}{n}\sum_{i=1}^{n} T_i\exp(-\bar{M}/\underline{\lambda}) = \exp(-\bar{M}/\underline{\lambda})\hat{u}. \\
\text{If } \ \bar{M} \leq 0: \quad G_1(\lambda_1; W) &= \mathbb{E}[T\exp(-Z/\lambda_1)] = \mathbb{E}[\exp(-Z/\lambda_1)|T=1]P(T=1) \\
&\geq \mathbb{E}[\exp(-\bar{M}/\bar{\lambda})|T=1]u = \exp(-\bar{M}/\bar{\lambda})u; \\
\hat{G}_1(\lambda_1; W) &= \frac{1}{n}\sum_{i=1}^{n} T_i\exp(-Z_i/\lambda_1) \qquad\qquad (23) \\
&\geq \frac{1}{n}\sum_{i=1}^{n} T_i\exp(-\bar{M}/\bar{\lambda}) = \exp(-\bar{M}/\bar{\lambda})\hat{u}.
\end{aligned}$$

Given $\hat{u} = \frac{1}{n}\sum_{i=1}^{n} T_i$ and $u = \mathbb{E}[\frac{1}{n}\sum_{i=1}^{n} T_i]$, using Hoeffding's inequality, we have

$$P\left(\left|\frac{1}{n}\sum_{i=1}^{n} T_i - \mathbb{E}[\frac{1}{n}\sum_{i=1}^{n} T_i]\right| \geq \frac{\mathbb{E}[\frac{1}{n}\sum_{i=1}^{n} T_i]}{2}\right) \leq 2\exp\left(-\frac{2(\frac{u}{2})^2}{n(\frac{1}{n})^2}\right) \leq \delta.$$

We can solve n by

$$2\exp\left(-\frac{nu^2}{2}\right) \leq \delta \Rightarrow n \geq \frac{2}{u^2}\log\left(\frac{2}{\delta}\right).$$

This indicates that $hatu \geq u/2$ with probability $1-\delta$ when $n \geq \frac{2}{u^2}\log\left(\frac{2}{\delta}\right)$. Combining this with equations (22) and (23), with probability $1-\delta$, when $n \geq \frac{2}{u^2}\log\left(\frac{2}{\delta}\right)$, we have

$$\begin{aligned}
\text{If } \bar{M} \geq 0: \quad & \inf_{\lambda_1 \in \Lambda, W \in \mathcal{W}} G_1(\lambda_1; W) \geq \exp(-\bar{M}/\underline{\lambda})u; \\
& \inf_{\lambda_1 \in \Lambda, W \in \mathcal{W}} \hat{G}_1(\lambda_1; W) \geq \exp(-\bar{M}/\underline{\lambda})\hat{u} \geq \exp(-\bar{M}/\underline{\lambda})u/2.
\end{aligned}$$

$$\begin{aligned}
\text{If } \bar{M} \leq 0: \quad & \inf_{\lambda_1 \in \Lambda, W \in \mathcal{W}} G_1(\lambda_1; W) \geq \exp(-\bar{M}/\bar{\lambda})u; \\
& \inf_{\lambda_1 \in \Lambda, W \in \mathcal{W}} \hat{G}_1(\lambda_1; W) \geq \exp(-\bar{M}/\bar{\lambda})\hat{u} \geq \exp(\bar{M}/\bar{\lambda})u/2.
\end{aligned}$$

Therefore, with probability $1-\delta$, when $n \geq \frac{2}{u^2}\log\left(\frac{2}{\delta}\right)$, we have

$$\begin{aligned}
&\text{If } \bar{M} \geq 0: \\
&\left|\log(\hat{G}_1(\lambda_1; W)) - \log\left(G_1(\lambda_1; W)\right)\right| \leq \frac{2}{\exp(-\bar{M}/\underline{\lambda})u}\left|\hat{G}_1(\lambda_1; W) - G_1(\lambda_1; W)\right|;
\end{aligned}$$

$$\begin{aligned}
&\text{If } \bar{M} \leq 0: \\
&\left|\log(\hat{G}_1(\lambda_1; W)) - \log\left(G_1(\lambda_1; W)\right)\right| \leq \frac{2}{\exp(-\bar{M}/\bar{\lambda})u}\left|\hat{G}_1(\lambda_1; W) - G_1(\lambda_1; W)\right|.
\end{aligned}$$

This completes the proof of Lemma B.5. □

Additionally, the following Lemma B.6 provides the bound of $|\log(\hat{u}) - \log(u)|$.

Lemma B.6. *Let $\hat{u} = \frac{1}{n}\sum_{i=1}^n T_i$ and $u = \mathbb{E}[\frac{1}{n}\sum_{i=1}^n T_i]$. For $n \geq \frac{2}{u^2}\log\left(\frac{2}{\delta}\right)$, with probability $1-\delta$, we have*

$$|\log(\hat{u}) - \log(u)| \leq \mathcal{O}\left(\sqrt{\frac{2\log(\frac{2}{\delta})}{nu^2}}\right). \tag{24}$$

Proof. Using Hoeffding's inequality, we have

$$\begin{aligned}
& P(|\hat{u} - u| \geq \epsilon) = P\left(\left|\frac{1}{n}\sum_{i=1}^n T_i - \mathbb{E}[\frac{1}{n}\sum_{i=1}^n T_i]\right| \geq \epsilon\right) \leq 2\exp\left(-2n\epsilon^2\right), \\
& 2\exp\left(-2n\epsilon^2\right) \leq \delta \quad \text{solves} \quad \epsilon \geq \sqrt{\frac{\log(\frac{2}{\delta})}{2n}}.
\end{aligned}$$

Notably, using the results in the previous lemma, we know for $n \geq \frac{2}{u^2}\log\left(\frac{2}{\delta}\right)$, $\hat{u} \geq u/2$. Therefore, we have

$$\begin{aligned}
|\log(\hat{u}) - \log(u)| &\leq \frac{1}{\min\{\hat{u}, u\}}|\hat{u} - u|. \quad \text{(By Lemma B.4)} \\
&\leq \frac{2}{u}|\hat{u} - u| \leq \frac{2}{u}\mathcal{O}\left(\sqrt{\frac{\log(\frac{2}{\delta})}{2n}}\right) = \mathcal{O}\left(\sqrt{\frac{2\log(\frac{2}{\delta})}{nu^2}}\right).
\end{aligned}$$

□

In the following, we will bound the term $|\hat{\mathcal{V}}(\hat{\tau})-\mathcal{V}(\hat{\tau})|$ using above lemmas. We first define functions $F_0(\lambda_0)$, $\hat{F}_0(\lambda_0)$, $F_1(\lambda_1)$, and $\hat{F}_1(\lambda_1)$:

$$\begin{aligned}
F_0(\lambda_0) &= \lambda_0\epsilon_0 + \lambda_0 \log(\mathbb{E}^{P_C}[\exp(\hat{\tau}(X)Y/\lambda_0)]) \\
&= \lambda_0\epsilon_0 + \lambda_0 \log\left(\frac{1}{1-u}\mathbb{E}[(1-T)\exp(\hat{\tau}(X)Y/\lambda_0)]\right); \\
\hat{F}_0(\lambda_0) &= \lambda_0\epsilon_0 + \lambda_0 \log(\frac{1}{n_c}\sum_{i=1}^{n}(1-T_i)\exp(\hat{\tau}(X_i)Y_i/\lambda_0)) \\
&= \lambda_0\epsilon_0 + \lambda_0 \log\left(\frac{1}{n(1-\hat{u})}\sum_{i=1}^{n}(1-T_i)\exp(\hat{\tau}(X_i)Y_i/\lambda_0)\right). \\
F_1(\lambda_1) &= \lambda_1\epsilon_1 + \lambda_1 \log(\mathbb{E}^{P_T}[\exp(-\hat{\tau}(X)Y/\lambda_1)]) \\
&= \lambda_1\epsilon_1 + \lambda_1 \log\left(\frac{1}{u}\mathbb{E}[T\exp(-\hat{\tau}(X)Y/\lambda_1)]\right); \\
\hat{F}_1(\lambda_1) &= \lambda_1\epsilon_1 + \lambda_1 \log(\frac{1}{n_t}\sum_{i=1}^{n}T_i\exp(-\hat{\tau}(X_i)Y_i/\lambda_1)) \\
&= \lambda_1\epsilon_1 + \lambda_1 \log\left(\frac{1}{n\hat{u}}\sum_{i=1}^{n}T_i\exp(-\hat{\tau}(X_i)Y_i/\lambda_1)\right).
\end{aligned}$$

The following Lemma B.7 bounds the term $|\hat{F}(\lambda)-F(\lambda)|$.

Lemma B.7. *Let $u := P(T=1)$. Assuming that $0 < \underline{\lambda} \le \lambda \le \bar{\lambda}$ and $\hat{\tau}(X)Y$ is bounded within the range of $\underline{M}$ to $\bar{M}$. Define $C_{exp} = \mathbf{1}_{\{\underline{M}\le\bar{M}\le 0\}}\exp\left(\bar{M}/\bar{\lambda}-\underline{M}/\underline{\lambda}\right) + \mathbf{1}_{\{\underline{M}\le 0,\bar{M}\ge 0\}}\exp\left(\bar{M}/\underline{\lambda}-\underline{M}/\underline{\lambda}\right) + \mathbf{1}_{\{0\le\underline{M}\le\bar{M}\}}\exp\left(\bar{M}/\underline{\lambda}-\underline{M}/\bar{\lambda}\right)$. For $n \ge 2/u^2\log(2/\delta)$, with probability $1-\delta$, we have*

$$\begin{aligned}
|\hat{F}_0(\lambda_0)-F_0(\lambda_0)| &\le \mathcal{O}\left(\sqrt{\frac{8\lambda_0^2\log\frac{2}{\delta}}{n(1-u)^2}C_{exp}^2}\right) + \mathcal{O}\left(\sqrt{\frac{2\lambda_0^2\log(\frac{2}{\delta})}{n(1-u)^2}}\right); \\
|\hat{F}_1(\lambda_1)-F_1(\lambda_1)| &\le \mathcal{O}\left(\sqrt{\frac{8\lambda_1^2\log\frac{2}{\delta}}{nu^2}C_{exp}^2}\right) + \mathcal{O}\left(\sqrt{\frac{2\lambda_1^2\log(\frac{2}{\delta})}{nu^2}}\right).
\end{aligned} \tag{25}$$

Proof.

$$\begin{aligned}
&|\hat{F}_0(\lambda_0)-F_0(\lambda_0)| \\
&= \left|\lambda_0\left(\log\left(\frac{1}{1-u}\mathbb{E}[(1-T)\exp(\hat{\tau}(X)Y/\lambda_0)]\right) - \log\left(\frac{1}{n(1-\hat{u})}\sum_{i=1}^{n}(1-T_i)\exp(\hat{\tau}(X_i)Y_i/\lambda_0)\right)\right)\right| \\
&= \lambda_0\left|\log\left(\mathbb{E}[(1-T)\exp(\hat{\tau}(X)Y/\lambda_0)]\right) - \log\left(\frac{1}{n}\sum_{i=1}^{n}(1-T_i)\exp(\hat{\tau}(X_i)Y_i/\lambda_0)\right) + \log(1-\hat{u}) - \log(1-u)\right| \\
&\le \lambda_0\left|\log\left(\mathbb{E}[(1-T)\exp(\hat{\tau}(X)Y/\lambda_0)]\right) - \log\left(\frac{1}{n}\sum_{i=1}^{n}(1-T_i)\exp(\hat{\tau}(X_i)Y_i/\lambda_0)\right)\right| + \lambda_0\left|\log(1-\hat{u})-\log(1-u)\right|.
\end{aligned}$$

$$
\begin{aligned}
&\text{If } \ \underline{M} \leq \bar{M} \leq 0: \\
&|\hat{F}_0(\lambda_0) - F_0(\lambda_0)| \\
&\leq \frac{2\lambda_0}{\exp(\underline{M}/\underline{\lambda})(1-u)} \left|\hat{G}_0(\lambda_0; W) - G_0(\lambda_0; W)\right| + \lambda_0 \left|\log(1-\hat{u}) - \log(1-u)\right| \quad \text{(By Lemma B.5)} \\
&\leq \mathcal{O}\left(\sqrt{\frac{8\lambda_0^2 \log\frac{2}{\delta}}{n(1-u)^2}\left(\exp\left(\bar{M}/\bar{\lambda} - \underline{M}/\underline{\lambda}\right)\right)^2}\right) + \mathcal{O}\left(\sqrt{\frac{2\lambda_0^2 \log(\frac{2}{\delta})}{n(1-u)^2}}\right) \quad \text{(By Lemma B.3 and Lemma B.6)} \\
&\text{If } \ \underline{M} \leq 0, \bar{M} \geq 0: \\
&|\hat{F}_0(\lambda_0) - F_0(\lambda_0)| \\
&\leq \frac{2\lambda_0}{\exp(\underline{M}/\underline{\lambda})(1-u)} \left|\hat{G}_0(\lambda_0; W) - G_0(\lambda_0; W)\right| + \lambda_0 \left|\log(1-\hat{u}) - \log(1-u)\right| \quad \text{(By Lemma B.5)} \\
&\leq \mathcal{O}\left(\sqrt{\frac{8\lambda_0^2 \log\frac{2}{\delta}}{n(1-u)^2}\left(\exp\left(\bar{M}/\underline{\lambda} - \underline{M}/\underline{\lambda}\right)\right)^2}\right) + \mathcal{O}\left(\sqrt{\frac{2\lambda_0^2 \log(\frac{2}{\delta})}{n(1-u)^2}}\right) \quad \text{(By Lemma B.3 and Lemma B.6)} \\
&\text{If } \ 0 \leq \underline{M} \leq \bar{M}: \\
&|\hat{F}_0(\lambda_0) - F_0(\lambda_0)| \\
&\leq \frac{2\lambda_0}{\exp(\underline{M}/\bar{\lambda})(1-u)} \left|\hat{G}_0(\lambda_0; W) - G_0(\lambda_0; W)\right| + \lambda_0 \left|\log(1-\hat{u}) - \log(1-u)\right| \quad \text{(By Lemma B.5)} \\
&\leq \mathcal{O}\left(\sqrt{\frac{8\lambda_0^2 \log\frac{2}{\delta}}{n(1-u)^2}\left(\exp\left(\bar{M}/\underline{\lambda} - \underline{M}/\bar{\lambda}\right)\right)^2}\right) + \mathcal{O}\left(\sqrt{\frac{2\lambda_0^2 \log(\frac{2}{\delta})}{n(1-u)^2}}\right) \quad \text{(By Lemma B.3 and Lemma B.6)}
\end{aligned}
$$

$$
\begin{aligned}
&|\hat{F}_1(\lambda_1) - F_1(\lambda_1)| \\
&= \left|\lambda_1 \left(\log\left(\frac{1}{u}\mathbb{E}[T\exp(-\hat{\tau}(X)Y/\lambda_1)]\right) - \log\left(\frac{1}{n\hat{u}}\sum_{i=1}^{n} T_i \exp(\hat{\tau}(X_i)Y_i/\lambda_0)\right)\right)\right| \\
&= \lambda_1 \left|\log\left(\mathbb{E}[T\exp(\hat{\tau}(X)Y/\lambda_1)]\right) - \log\left(\frac{1}{n}\sum_{i=1}^{n} T_i \exp(\hat{\tau}(X_i)Y_i/\lambda_1)\right) + \log(\hat{u}) - \log(u)\right| \\
&\leq \lambda_1 \left|\log\left(\mathbb{E}[T\exp(\hat{\tau}(X)Y/\lambda_1)]\right) - \log\left(\frac{1}{n}\sum_{i=1}^{n} T_i \exp(\hat{\tau}(X_i)Y_i/\lambda_1)\right)\right| + \lambda_1 \left|\log(\hat{u}) - \log(u)\right|.
\end{aligned}
$$

$$
\begin{aligned}
&\text{If } \underline{M} \leq \bar{M} \leq 0: \\
&|\hat{F}_1(\lambda_1) - F_1(\lambda_1)| \\
&\leq \frac{2\lambda_1}{\exp(-\bar{M}/\bar{\lambda})u} \left|\hat{G}_0(\lambda_0; W) - G_0(\lambda_0; W)\right| + \lambda_1 \left|\log(\hat{u}) - \log(u)\right| \quad \text{(By Lemma B.5)} \\
&\leq \mathcal{O}\left(\sqrt{\frac{8\lambda_1^2 \log\frac{2}{\delta}}{nu^2}\left(\exp\left(\bar{M}/\bar{\lambda} - \underline{M}/\underline{\lambda}\right)\right)^2}\right) + \mathcal{O}\left(\sqrt{\frac{2\lambda_1^2 \log(\frac{2}{\delta})}{nu^2}}\right) \quad \text{(By Lemma B.3 and Lemma B.6)} \\
&\text{If } \underline{M} \leq 0, \bar{M} \geq 0: \\
&|\hat{F}_1(\lambda_1) - F_1(\lambda_1)| \\
&\leq \frac{2\lambda_1}{\exp(-\bar{M}/\underline{\lambda})u} \left|\hat{G}_1(\lambda_1; W) - G_1(\lambda_1; W)\right| + \lambda_1 \left|\log(\hat{u}) - \log(u)\right| \quad \text{(By Lemma B.5)} \\
&\leq \mathcal{O}\left(\sqrt{\frac{8\lambda_1^2 \log\frac{2}{\delta}}{nu^2}\left(\exp\left(\bar{M}/\underline{\lambda} - \underline{M}/\underline{\lambda}\right)\right)^2}\right) + \mathcal{O}\left(\sqrt{\frac{2\lambda_1^2 \log(\frac{2}{\delta})}{nu^2}}\right) \quad \text{(By Lemma B.3 and Lemma B.6)} \\
&\text{If } 0 \leq \underline{M} \leq \bar{M}: \\
&|\hat{F}_1(\lambda_1) - F_1(\lambda_1)| \\
&\leq \frac{2\lambda_1}{\exp(-\bar{M}/\underline{\lambda})u} \left|\hat{G}_1(\lambda_1; W) - G_1(\lambda_1; W)\right| + \lambda_1 \left|\log(\hat{u}) - \log(u)\right| \quad \text{(By Lemma B.5)} \\
&\leq \mathcal{O}\left(\sqrt{\frac{8\lambda_1^2 \log\frac{2}{\delta}}{nu^2}\left(\exp\left(\bar{M}/\underline{\lambda} - \underline{M}/\bar{\lambda}\right)\right)^2}\right) + \mathcal{O}\left(\sqrt{\frac{2\lambda_1^2 \log(\frac{2}{\delta})}{nu^2}}\right) \quad \text{(By Lemma B.3 and Lemma B.6)}
\end{aligned}
$$

□

Now, we can prove the result in Theorem 4.5.

Proof. Let $\hat{\lambda}_0 = \arg\min_\lambda \hat{F}_0(\lambda_0)$, $\lambda_0^* = \arg\min_{\lambda_0} F_0(\lambda_0)$, $\hat{\lambda}_1 = \arg\min_\lambda \hat{F}_1(\lambda_1)$ and $\lambda_1^* = \arg\min_{\lambda_1} F_1(\lambda_1)$. Then we have

$$
\begin{aligned}
\mathcal{V}^0(\hat{\tau}) - \hat{\mathcal{V}}^0(\hat{\tau}) &= F_0(\lambda_0^*) - \hat{F}_0(\hat{\lambda}_0) \\
&= F_0(\lambda_0^*) - \hat{F}_0(\hat{\lambda}_0) + F_0(\hat{\lambda}_0) - F_0(\hat{\lambda}_0) \\
&= F_0(\hat{\lambda}_0) - \hat{F}_0(\hat{\lambda}_0) + F_0(\lambda_0^*) - F_0(\hat{\lambda}_0) \\
&\leq |F_0(\hat{\lambda}_0) - \hat{F}_0(\hat{\lambda}_0)| + 0 \\
&\leq \sup_{\lambda_0} |F_0(\lambda_0) - \hat{F}_0(\lambda_0)|.
\end{aligned}
$$

$$
\begin{aligned}
\hat{\mathcal{V}}^0(\hat{\tau}) - \mathcal{V}^0(\hat{\tau}) &= \hat{F}_0(\hat{\lambda}_0) - F_0(\lambda_0^*) \\
&= \hat{F}_0(\hat{\lambda}_0) - F_0(\lambda_0^*) + \hat{F}_0(\lambda_0^*) - \hat{F}_0(\lambda_0^*) \\
&= \hat{F}_0(\lambda_0^*) - F_0(\lambda_0^*) + \hat{F}_0(\hat{\lambda}_0) - \hat{F}_0(\lambda_0^*) \\
&\leq |\hat{F}_0(\lambda_0^*) - F_0(\lambda_0^*)| + 0 \\
&\leq \sup_{\lambda_0} |\hat{F}_0(\lambda_0) - F_0(\lambda_0)|.
\end{aligned}
$$

$$
\begin{aligned}
\mathcal{V}^1(\hat{\tau}) - \hat{\mathcal{V}}^1(\hat{\tau}) &= F_1(\lambda_1^*) - \hat{F}_1(\hat{\lambda}_1) \\
&= F_1(\lambda_1^*) - \hat{F}_1(\hat{\lambda}_1) + F_1(\hat{\lambda}_1) - F_1(\hat{\lambda}_1) \\
&= F_1(\hat{\lambda}_1) - \hat{F}_1(\hat{\lambda}_1) + F_1(\lambda_1^*) - F_1(\hat{\lambda}_1) \\
&\leq |F_1(\hat{\lambda}_1) - \hat{F}_1(\hat{\lambda}_1)| + 0 \\
&\leq \sup_{\lambda_1} |F_1(\lambda_1) - \hat{F}_1(\lambda_1)|.
\end{aligned}
$$

$$
\begin{aligned}
\hat{\mathcal{V}}^1(\hat{\tau}) - \mathcal{V}^1(\hat{\tau}) &= \hat{F}_1(\hat{\lambda}_1) - F_1(\lambda_1^*) \\
&= \hat{F}_1(\hat{\lambda}_1) - F_1(\lambda_1^*) + \hat{F}_1(\lambda_1^*) - \hat{F}_1(\lambda_1^*) \\
&= \hat{F}_1(\lambda_1^*) - F_1(\lambda_1^*) + \hat{F}_1(\hat{\lambda}_1) - \hat{F}_1(\lambda_1^*) \\
&\le |\hat{F}_1(\lambda_1^*) - F_1(\lambda_1^*)| + 0 \\
&\le \sup_{\lambda_1} |\hat{F}_1(\lambda_1) - F_1(\lambda_1)|.
\end{aligned}
$$

Therefore, we have

If $\underline{M} \le \bar{M} \le 0$:

$$
|\hat{\mathcal{V}}^0(\hat{\tau}) - \mathcal{V}^0(\hat{\tau})| \le \sup_{\lambda} |\hat{F}(\lambda) - F(\lambda)| \le \mathcal{O}\left(\sqrt{\frac{8\bar{\lambda}^2 \log \frac{2}{\delta}}{n(1-u)^2} \left(\exp\left(\bar{M}/\bar{\lambda} - \underline{M}/\underline{\lambda}\right)\right)^2}\right) + \mathcal{O}\left(\sqrt{\frac{2\bar{\lambda}^2 \log(\frac{2}{\delta})}{n(1-u)^2}}\right);
$$

$$
|\hat{\mathcal{V}}^1(\hat{\tau}) - \mathcal{V}^1(\hat{\tau})| \le \sup_{\lambda} |\hat{F}(\lambda) - F(\lambda)| \le \mathcal{O}\left(\sqrt{\frac{8\bar{\lambda}^2 \log \frac{2}{\delta}}{nu^2} \left(\exp\left(\bar{M}/\bar{\lambda} - \underline{M}/\underline{\lambda}\right)\right)^2}\right) + \mathcal{O}\left(\sqrt{\frac{2\bar{\lambda}^2 \log(\frac{2}{\delta})}{nu^2}}\right).
$$

If $\underline{M} \le 0, \bar{M} \ge 0$:

$$
|\hat{\mathcal{V}}^0(\hat{\tau}) - \mathcal{V}^0(\hat{\tau})| \le \sup_{\lambda} |\hat{F}(\lambda) - F(\lambda)| \le \mathcal{O}\left(\sqrt{\frac{8\bar{\lambda}^2 \log \frac{2}{\delta}}{n(1-u)^2} \left(\exp\left(\bar{M}/\underline{\lambda} - \underline{M}/\underline{\lambda}\right)\right)^2}\right) + \mathcal{O}\left(\sqrt{\frac{2\bar{\lambda}^2 \log(\frac{2}{\delta})}{n(1-u)^2}}\right);
$$

$$
|\hat{\mathcal{V}}^1(\hat{\tau}) - \mathcal{V}^1(\hat{\tau})| \le \sup_{\lambda} |\hat{F}(\lambda) - F(\lambda)| \le \mathcal{O}\left(\sqrt{\frac{8\bar{\lambda}^2 \log \frac{2}{\delta}}{nu^2} \left(\exp\left(\bar{M}/\underline{\lambda} - \underline{M}/\underline{\lambda}\right)\right)^2}\right) + \mathcal{O}\left(\sqrt{\frac{2\bar{\lambda}^2 \log(\frac{2}{\delta})}{nu^2}}\right).
$$

If $0 \le \underline{M} \le \bar{M}$:

$$
|\hat{\mathcal{V}}^0(\hat{\tau}) - \mathcal{V}^0(\hat{\tau})| \le \sup_{\lambda} |\hat{F}(\lambda) - F(\lambda)| \le \mathcal{O}\left(\sqrt{\frac{8\bar{\lambda}^2 \log \frac{2}{\delta}}{n(1-u)^2} \left(\exp\left(\bar{M}/\underline{\lambda} - \underline{M}/\bar{\lambda}\right)\right)^2}\right) + \mathcal{O}\left(\sqrt{\frac{2\bar{\lambda}^2 \log(\frac{2}{\delta})}{n(1-u)^2}}\right);
$$

$$
|\hat{\mathcal{V}}^1(\hat{\tau}) - \mathcal{V}^1(\hat{\tau})| \le \sup_{\lambda} |\hat{F}(\lambda) - F(\lambda)| \le \mathcal{O}\left(\sqrt{\frac{8\bar{\lambda}^2 \log \frac{2}{\delta}}{nu^2} \left(\exp\left(\bar{M}/\underline{\lambda} - \underline{M}/\bar{\lambda}\right)\right)^2}\right) + \mathcal{O}\left(\sqrt{\frac{2\bar{\lambda}^2 \log(\frac{2}{\delta})}{nu^2}}\right).
$$

Finally, we have

$$
|\hat{\mathcal{V}}_t(\hat{\tau}) - \mathcal{V}_t(\hat{\tau})| \le \mathcal{O}\left(\sqrt{\frac{8\bar{\lambda}^2 \log \frac{2}{\delta}}{nu_t^2} C_{exp}^2}\right) + \mathcal{O}\left(\sqrt{\frac{2\bar{\lambda}^2 \log(\frac{2}{\delta})}{nu_t^2}}\right).
$$

Note that $u_1 = P(T = 1)$ and $u_0 = P(T = 0)$. $C_{exp} = \mathbf{1}_{\{\underline{M} \le \bar{M} \le 0\}} \exp\left(\bar{M}/\bar{\lambda} - \underline{M}/\underline{\lambda}\right) + \mathbf{1}_{\{\underline{M} \le 0, \bar{M} \ge 0\}} \exp\left(\bar{M}/\underline{\lambda} - \underline{M}/\underline{\lambda}\right) + \mathbf{1}_{\{0 \le \underline{M} \le \bar{M}\}} \exp\left(\bar{M}/\underline{\lambda} - \underline{M}/\bar{\lambda}\right)$.

□

C Additional Experimental Results

C.1 The Complementary Results with 24 Candidate Pool

This Section reports the complementary results for 24 candidate CATE estimators, where the candidate pool contains 3 ML models (LR, SVM, and RF) $\times$ 8 learners (S-, T-, PS-, IPW-, X-, DR-, R-, RA-).

Table 3: Comparison of PEHE for different selectors across Settings A, B, and C (Note that B ($\xi = 1$) matches A ($\rho = 0.1$)), with base model for CATE estimator being {LR, SVM, RF}. Reported values (mean $\pm$ standard deviation) are computed over 100 experiments. Smaller is better.

	A ($\rho = 0$)	A ($\rho = 0.1$)	A ($\rho = 0.3$)	B ($\xi = 0$)	B ($\xi = 2$)	C ($m = 0.1$)	C ($m = 0.5$)	C ($m = 0.9$)
Plug-S	4.28±6.07	4.74±5.44	5.62±5.17	**3.65**±7.32	10.88±13.94	**5.53**±5.76	8.75±8.71	14.42±12.59
Plug-PS	4.28±6.07	4.74±5.44	5.62±5.17	**3.44**±7.09	10.88±13.94	**5.53**±5.76	**8.74**±8.71	**14.41**±12.59
Plug-T	39.00±21.96	38.25±19.89	37.67±19.38	16.80±22.80	47.91±22.90	39.78±20.17	38.09±20.65	35.63±14.86
Plug-X	9.36±7.96	9.24±7.02	9.86±6.70	9.19±12.54	17.17±16.44	12.80±13.43	16.93±18.07	20.24±13.05
Plug-IPW	30.34±23.39	28.43±22.83	27.08±21.12	13.94±25.47	35.64±19.05	35.53±27.32	28.02±21.95	30.49±18.67
Plug-DR	36.36±22.13	36.70±21.53	36.21±19.67	14.64±22.95	46.58±23.95	38.95±20.56	36.44±21.72	32.59±16.27
Plug-R	**2.74**±6.25	**3.13**±5.96	**4.47**±5.71	**2.31**±4.18	**6.42**±8.35	**4.08**±5.53	**5.45**±7.99	**7.59**±8.84
Plug-RA	39.71±21.77	38.75±19.88	37.76±19.29	15.47±22.97	48.50±23.06	39.77±19.90	38.16±20.35	36.03±15.19
Pseudo-DR	38.94±21.83	37.62±20.80	36.64±20.34	17.00±22.37	47.96±23.19	39.05±20.04	37.43±20.56	35.46±15.99
Pseudo-R	**2.08**±3.51	**4.46**±9.92	**4.51**±4.10	6.39±16.07	**9.78**±17.53	6.68±13.35	11.47±12.87	19.37±13.37
Pseudo-IF	32.88±10.56	33.05±11.50	33.37±10.79	19.00±18.27	36.05±13.35	34.40±13.23	31.71±8.88	27.03±6.51
Random	38.32±50.26	39.33±30.85	32.00±31.27	16.06±15.91	44.45±48.36	38.97±33.22	35.84±36.39	33.55±28.59
Fact	40.13±31.41	40.50±30.97	39.90±30.99	6.65±10.14	60.58±40.12	39.91±28.49	39.12±28.86	42.67±27.81
Matching	36.67±19.00	35.75±18.53	35.55±16.88	15.46±20.39	42.95±21.24	36.43±17.57	35.42±15.76	32.83±13.76
DRM	**1.11**±1.08	**1.97**±0.70	**3.42**±0.82	4.34±16.36	**3.24**±7.95	**2.70**±6.34	**2.90**±1.44	**4.68**±2.62

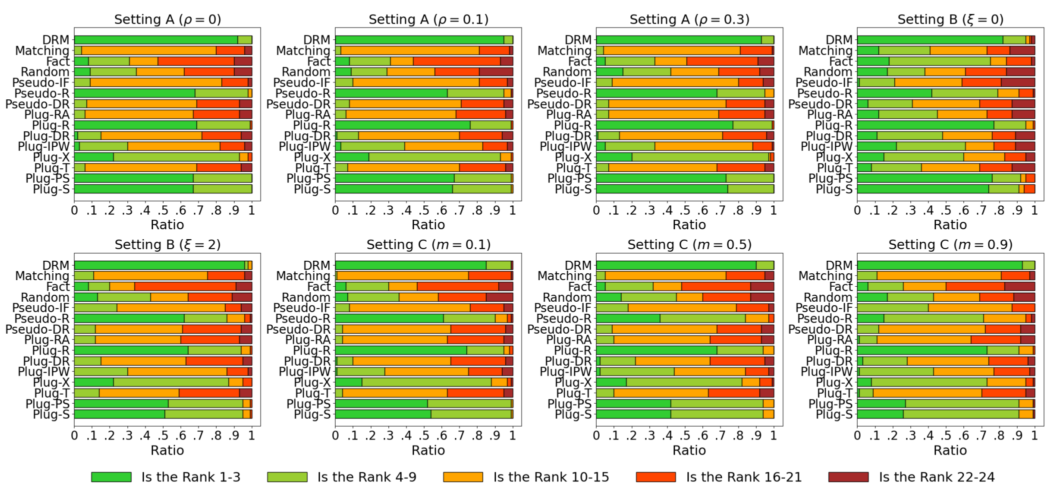

Figure 1: The stacked bar chart showing the distribution of the selected estimator's rank for each evaluation metric across rank intervals: [1-3], [4-9], [10-15], [16-21], and [22-24]. The greener (or redder) color indicates that the selected estimator ranks higher (or lower). For example, the **dark red** (or **green**) indicates the percentage of cases (out of 100 experiments) where the selected estimator ranks among the worst 3 estimators, specifically as ranks 22, 23, or 24. (or among the best 3 estimators, specifically as ranks 1, 2, or 3).

C.2 The Complementary Results with 8 Candidate Pool (LR)

This Section reports the complementary results for 8 candidate CATE estimators, where the candidate pool contains 1 ML model (LR) $\times$ 8 learners (S-, T-, PS-, IPW-, X-, DR-, R-, RA-).

Table 4: Comparison of Regret for different selectors across Settings A, B, and C (Note that B ($\xi = 1$) matches A ($\rho = 0.1$)), with base model for CATE estimator fixed as LR. Reported values (mean $\pm$ standard deviation) are computed over 100 experiments. Smaller is better.

	A ($\rho = 0$)	A ($\rho = 0.1$)	A ($\rho = 0.3$)	B ($\xi = 0$)	B ($\xi = 2$)	C ($m = 0.1$)	C ($m = 0.5$)	C ($m = 0.9$)
Plug-S	0.00±0.01	0.16±1.62	0.35±2.76	**0.03**±0.13	**0.57**±3.16	**0.11**±0.80	0.98±5.15	3.52±7.73
Plug-PS	**0.00**±0.01	0.16±1.62	0.35±2.76	0.03±0.13	**0.57**±3.16	**0.11**±0.80	0.98±5.15	3.52±7.73
Plug-T	15.21±17.73	15.16±17.42	15.83±17.03	1.71±3.72	26.67±21.77	14.09±16.90	13.68±17.51	15.15±14.34
Plug-X	0.64±3.21	0.86±3.89	0.72±3.55	0.67±1.54	0.83±4.55	0.61±3.56	1.63±6.71	6.29±10.60
Plug-IPW	8.39±14.21	8.79±14.28	8.70±14.43	1.04±3.41	23.26±22.11	11.08±16.51	8.45±14.47	10.17±12.11
Plug-DR	12.76±16.90	13.15±16.82	14.40±16.83	0.94±2.30	24.37±21.74	13.50±16.94	11.52±16.45	13.32±14.04
Plug-R	0.65±5.36	0.65±5.29	0.51±5.08	0.10±0.34	1.01±4.82	0.23±2.28	0.36±2.13	**1.98**±6.63
Plug-RA	15.21±17.73	15.16±17.42	15.83±17.03	1.04±2.43	26.15±21.79	14.37±17.17	14.40±17.69	15.10±14.40
Pseudo-DR	14.93±17.62	14.96±17.35	15.98±17.00	1.83±4.50	25.35±21.82	14.79±17.32	12.76±16.75	14.85±14.70
Pseudo-R	0.05±0.47	0.82±8.15	**0.22**±2.15	0.27±0.74	4.75±15.17	1.60±7.22	0.95±4.47	6.07±10.32
Pseudo-IF	28.85±16.33	28.68±15.88	28.74±15.72	7.92±10.35	37.09±17.92	26.72±19.22	22.02±18.11	14.11±15.25
Random	11.36±22.65	15.05±24.73	8.96±20.28	3.62±8.84	13.12±26.79	12.51±23.95	11.43±21.10	11.61±18.74
Fact	37.16±32.63	36.58±32.72	35.06±32.83	4.61±8.86	57.10±43.03	35.28±29.27	28.59±25.26	26.67±21.32
Matching	17.19±18.08	17.40±17.69	17.23±16.94	0.78±1.65	29.25±21.09	17.56±17.74	17.18±16.69	15.98±15.14
DRM	0.00±0.01	**0.00**±0.00	0.59±4.16	0.09±0.37	3.55±16.56	0.60±3.15	**0.36**±2.23	3.46±7.89

Table 5: Comparison of PEHE for different selectors across Settings A, B, and C (Note that B ($\xi = 1$) matches A ($\rho = 0.1$)), with base model for CATE estimator fixed as LR. Reported values (mean $\pm$ standard deviation) are computed over 100 experiments. Smaller is better.

	A ($\rho = 0$)	A ($\rho = 0.1$)	A ($\rho = 0.3$)	B ($\xi = 0$)	B ($\xi = 2$)	C ($m = 0.1$)	C ($m = 0.5$)	C ($m = 0.9$)
Plug-S	**5.48**±4.51	6.09±4.98	7.33±6.20	**2.94**±1.46	**8.26**±9.08	**6.44**±5.60	9.07±7.53	17.03±13.16
Plug-PS	**5.48**±4.51	6.09±4.98	7.33±6.20	**2.94**±1.46	**8.26**±9.08	**6.44**±5.60	9.07±7.53	17.03±13.16
Plug-T	20.69±19.21	21.08±18.95	22.81±18.64	4.62±3.93	34.35±21.87	20.42±17.22	21.77±16.83	28.66±14.53
Plug-X	6.12±5.40	6.78±5.66	7.70±5.51	3.58±1.84	8.52±8.14	6.95±6.39	9.72±8.96	19.81±14.24
Plug-IPW	13.87±15.61	14.72±15.73	15.68±15.78	3.95±3.78	30.94±23.10	17.42±17.14	16.54±14.44	23.68±14.13
Plug-DR	18.24±18.65	19.08±18.51	21.39±18.46	3.85±2.63	32.06±21.93	19.83±17.17	19.61±16.05	26.84±14.42
Plug-R	6.13±7.03	6.57±6.94	7.50±6.83	3.01±1.50	8.70±8.55	6.56±5.38	8.45±5.51	**15.49**±12.03
Plug-RA	20.69±19.21	21.08±18.95	22.81±18.64	3.95±2.70	33.84±21.78	20.71±17.34	22.48±16.71	28.61±14.65
Pseudo-DR	20.41±19.16	20.89±18.91	22.96±18.58	4.74±4.77	33.04±21.91	21.12±17.46	20.85±16.05	28.36±14.81
Pseudo-R	5.53±4.57	6.74±8.92	**7.20**±4.94	3.18±1.47	12.43±15.65	7.94±8.57	9.03±7.30	19.59±14.23
Pseudo-IF	34.33±15.99	34.61±15.52	35.73±15.79	10.83±10.32	44.77±15.83	33.05±18.87	30.10±15.78	27.63±14.61
Random	16.84±23.01	20.98±25.35	15.94±21.31	6.53±9.03	20.80±28.56	18.85±24.32	19.52±21.91	25.12±23.17
Fact	42.64±34.02	42.50±34.03	42.04±34.10	7.52±9.22	64.79±44.48	41.62±30.19	36.68±25.97	40.19±24.50
Matching	22.67±19.39	23.32±19.06	24.21±18.42	3.68±2.19	36.93±20.47	23.90±17.76	25.26±15.18	29.50±15.23
DRM	**5.48**±4.51	**5.93**±4.44	7.57±6.82	3.00±1.51	11.23±19.27	6.94±7.36	**8.44**±5.91	16.98±13.26

Table 6: Comparison of rank correlation for different selectors across Settings A, B, and C (Note that B ($\xi = 1$) matches A ($\rho = 0.1$)), with base model for CATE estimator fixed as LR. Reported values (mean $\pm$ standard deviation) are computed over 100 experiments. Smaller is better.

	A ($\rho = 0$)	A ($\rho = 0.1$)	A ($\rho = 0.3$)	B ($\xi = 0$)	B ($\xi = 2$)	C ($m = 0.1$)	C ($m = 0.5$)	C ($m = 0.9$)
Plug-S	0.97±0.04	0.97±0.04	**0.97**±0.06	0.97±0.03	0.96±0.05	**0.97**±0.05	**0.95**±0.09	0.89±0.19
Plug-PS	**0.98**±0.04	**0.97**±0.04	0.97±0.05	**0.98**±0.04	**0.97**±0.05	0.96±0.05	0.95±0.09	0.89±0.19
Plug-T	0.65±0.49	0.65±0.50	0.63±0.50	0.87±0.20	0.63±0.47	0.70±0.37	0.55±0.51	0.49±0.54
Plug-X	0.91±0.12	0.91±0.16	0.90±0.16	0.93±0.11	0.91±0.12	0.91±0.11	0.84±0.22	0.78±0.31
Plug-IPW	0.83±0.23	0.84±0.22	0.83±0.25	0.93±0.12	0.74±0.40	0.82±0.25	0.73±0.40	0.68±0.37
Plug-DR	0.71±0.41	0.71±0.42	0.68±0.44	0.93±0.11	0.67±0.42	0.72±0.36	0.63±0.47	0.55±0.49
Plug-R	0.96±0.17	0.95±0.17	0.95±0.18	0.97±0.04	0.94±0.16	0.95±0.10	0.95±0.07	**0.92**±0.17
Plug-RA	0.65±0.48	0.66±0.48	0.63±0.50	0.90±0.15	0.64±0.45	0.69±0.37	0.56±0.50	0.49±0.54
Pseudo-DR	0.67±0.47	0.66±0.48	0.65±0.48	0.87±0.20	0.65±0.44	0.69±0.37	0.59±0.49	0.48±0.54
Pseudo-R	0.91±0.12	0.90±0.18	0.92±0.11	0.96±0.05	0.87±0.22	0.90±0.15	0.89±0.14	0.77±0.31
Pseudo-IF	0.53±0.51	0.52±0.51	0.50±0.51	0.44±0.62	0.60±0.47	0.55±0.46	0.54±0.52	0.59±0.52
Random	0.44±0.13	0.41±0.16	0.42±0.13	0.42±0.16	0.41±0.17	0.42±0.18	0.40±0.18	0.37±0.21
Fact	0.27±0.13	0.28±0.13	0.28±0.13	0.35±0.14	0.26±0.13	0.31±0.14	0.28±0.15	0.22±0.20
Matching	0.66±0.46	0.66±0.46	0.63±0.48	0.92±0.09	0.66±0.41	0.68±0.38	0.60±0.48	0.51±0.50
DRM	0.64±0.20	0.64±0.20	0.62±0.25	0.89±0.11	0.61±0.25	0.65±0.24	0.71±0.22	0.70±0.24

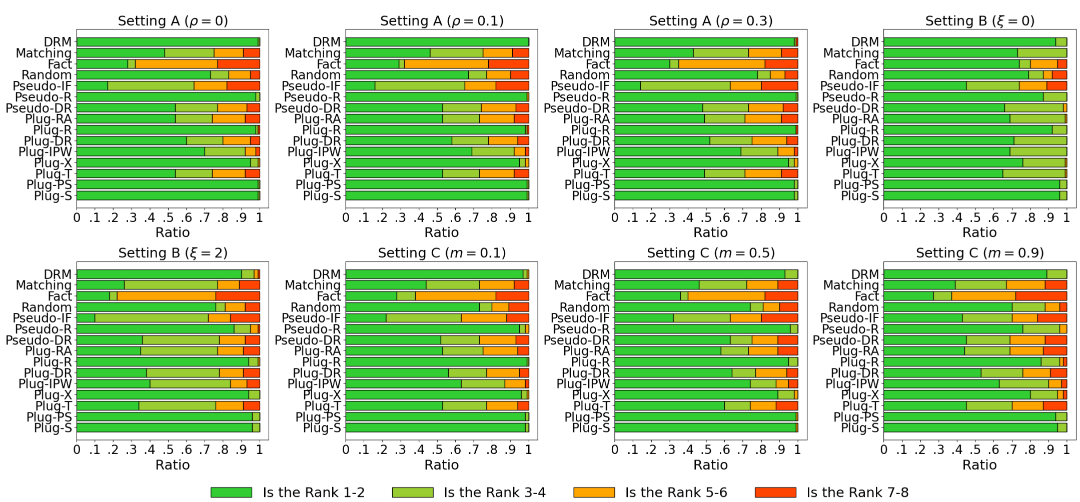

Figure 2: The stacked bar chart showing the distribution of the selected estimator's rank for each evaluation metric across rank intervals: [1-2], [3-4], [5-6], [7-8]. The estimator selection is over 8 candidate estimators, with the underlying ML model for CATE estimator fixed as LR.

C.3 The Complementary Results with 8 Candidate Pool (SVM)

This Section reports the complementary results for 8 candidate CATE estimators, where the candidate pool contains 1 ML model (SVM) $\times$ 8 learners (S-, T-, PS-, IPW-, X-, DR-, R-, RA-).

Table 7: Comparison of Regret for different selectors across Settings A, B, and C (Note that B ($\xi = 1$) matches A ($\rho = 0.1$)), with base model for CATE estimator fixed as SVM. Reported values (mean $\pm$ standard deviation) are computed over 100 experiments. Smaller is better.

	A ($\rho = 0$)	A ($\rho = 0.1$)	A ($\rho = 0.3$)	B ($\xi = 0$)	B ($\xi = 2$)	C ($m = 0.1$)	C ($m = 0.5$)	C ($m = 0.9$)
Plug-S	3.08±6.14	2.35±5.53	2.21±5.00	0.03±0.07	7.39±10.03	2.61±5.44	4.52±6.94	9.44±9.71
Plug-PS	3.08±6.14	2.35±5.53	2.21±5.00	0.03±0.07	7.34±10.05	2.61±5.44	4.42±6.92	9.44±9.71
Plug-T	26.12±7.91	25.25±7.75	23.62±7.92	2.48±2.58	27.09±8.35	24.44±8.29	21.84±8.91	23.00±8.67
Plug-X	7.58±6.94	6.81±6.46	5.88±5.79	1.35±1.98	12.42±9.40	8.65±6.36	9.37±7.72	13.25±8.49
Plug-IPW	19.15±10.72	18.23±10.88	16.64±10.66	1.01±1.90	25.64±7.58	18.70±9.63	16.72±10.31	18.47±9.08
Plug-DR	24.15±9.77	23.14±9.87	21.79±9.41	1.35±2.23	26.43±9.32	23.35±9.36	20.75±9.42	19.89±9.36
Plug-R	1.28±3.15	0.75±2.56	0.66±2.14	0.14±0.50	3.61±6.82	1.97±5.50	1.79±4.70	2.76±6.27
Plug-RA	25.89±7.98	24.97±7.98	23.61±7.91	1.91±2.49	27.08±8.36	24.70±8.28	22.44±8.59	22.81±8.84
Pseudo-DR	25.07±8.31	24.15±8.41	22.64±8.55	2.72±2.66	26.80±8.80	24.27±8.34	22.21±8.95	23.10±8.27
Pseudo-R	0.76±3.38	1.80±6.27	1.07±4.03	0.72±1.98	4.93±9.48	3.03±7.94	7.02±8.56	13.31±10.34
Pseudo-IF	28.87±3.29	27.87±3.38	26.41±3.64	4.53±3.15	27.46±3.30	28.07±3.32	25.71±4.22	21.71±4.27
Random	3.81±8.40	3.36±7.77	3.80±7.20	1.24±2.55	5.73±10.37	4.41±8.36	3.76±7.73	6.21±9.22
Fact	11.80±11.16	10.79±11.06	9.27±10.46	0.17±0.40	17.13±13.31	11.27±11.03	13.95±12.09	17.14±12.43
Matching	27.62±7.04	26.72±7.07	24.98±7.29	3.05±2.90	27.83±7.03	26.58±6.56	25.22±7.34	24.20±7.55
DRM	**0.04**±0.22	**0.01**±0.07	**0.02**±0.10	**0.03**±0.07	**0.91**±6.72	**0.04**±0.35	**0.16**±0.46	**1.04**±1.83

Table 8: Comparison of PEHE for different selectors across Settings A, B, and C (Note that B ($\xi = 1$) matches A ($\rho = 0.1$)), with base model for CATE estimator fixed as SVM. Reported values (mean $\pm$ standard deviation) are computed over 100 experiments. Smaller is better.

	A ($\rho = 0$)	A ($\rho = 0.1$)	A ($\rho = 0.3$)	B ($\xi = 0$)	B ($\xi = 2$)	C ($m = 0.1$)	C ($m = 0.5$)	C ($m = 0.9$)
Plug-S	4.13±6.29	4.34±5.52	5.55±4.99	1.69±0.63	9.53±10.18	4.63±5.56	7.37±7.40	13.13±10.22
Plug-PS	4.13±6.29	4.34±5.52	5.55±4.99	1.69±0.63	9.48±10.20	4.63±5.56	7.27±7.39	13.13±10.22
Plug-T	27.17±7.98	27.24±7.79	26.95±7.90	4.15±2.64	29.23±8.33	26.47±8.39	24.70±9.15	26.70±9.00
Plug-X	8.64±7.04	8.80±6.44	9.22±5.66	3.01±1.97	14.56±9.52	10.68±6.45	12.23±8.14	16.95±9.05
Plug-IPW	20.21±10.81	20.22±10.89	19.97±10.65	2.68±2.00	27.78±7.53	20.73±9.68	19.57±10.50	22.17±9.44
Plug-DR	25.20±9.88	25.13±9.93	25.13±9.40	3.01±2.31	28.58±9.30	25.37±9.39	23.60±9.67	23.58±9.77
Plug-R	2.33±3.31	2.74±2.70	3.99±2.22	1.80±0.74	5.75±6.72	4.00±5.51	4.64±4.99	6.45±6.61
Plug-RA	26.94±8.04	26.96±8.03	26.95±7.91	3.57±2.55	29.23±8.35	26.72±8.38	25.30±8.79	26.51±9.19
Pseudo-DR	26.13±8.41	26.14±8.43	25.98±8.51	4.38±2.72	28.95±8.79	26.29±8.43	25.07±9.18	26.79±8.65
Pseudo-R	1.81±3.40	3.79±6.31	4.41±4.11	2.39±2.02	7.07±9.52	5.05±7.99	9.88±8.93	17.01±10.59
Pseudo-IF	29.93±3.32	29.86±3.34	29.75±3.52	6.19±3.11	29.61±3.26	30.10±3.32	28.56±4.21	25.41±4.34
Random	4.86±8.35	5.35±7.70	7.14±7.31	2.90±2.55	7.88±10.43	6.44±8.52	6.61±7.78	9.90±9.46
Fact	12.85±11.33	12.78±11.11	12.60±10.41	1.83±0.73	19.28±13.44	13.30±11.14	16.81±12.40	20.84±12.88
Matching	28.68±7.15	28.72±7.12	28.32±7.25	4.71±2.95	29.98±7.04	28.61±6.54	28.08±7.50	27.90±7.94
DRM	**1.10**±0.49	**2.00**±0.56	**3.36**±0.69	**1.69**±0.62	**3.06**±6.72	**2.07**±0.67	**3.01**±1.25	**4.74**±2.41

Table 9: Comparison of rank correlation for different selectors across Settings A, B, and C (Note that B ($\xi = 1$) matches A ($\rho = 0.1$)), with base model for CATE estimator fixed as SVM. Reported values (mean $\pm$ standard deviation) are computed over 100 experiments. Smaller is better.

	A ($\rho = 0$)	A ($\rho = 0.1$)	A ($\rho = 0.3$)	B ($\xi = 0$)	B ($\xi = 2$)	C ($m = 0.1$)	C ($m = 0.5$)	C ($m = 0.9$)
Plug-S	0.86±0.27	0.87±0.25	0.86±0.26	0.96±0.04	0.63±0.47	0.84±0.25	0.74±0.42	0.45±0.60
Plug-PS	0.86±0.27	0.86±0.25	0.86±0.26	**0.96**±0.04	0.63±0.47	0.85±0.26	0.74±0.42	0.46±0.59
Plug-T	-0.61±0.30	-0.62±0.32	-0.63±0.29	0.50±0.43	-0.53±0.35	-0.64±0.27	-0.55±0.46	-0.70±0.37
Plug-X	0.60±0.37	0.60±0.39	0.61±0.40	0.72±0.30	0.45±0.45	0.49±0.44	0.44±0.48	0.10±0.64
Plug-IPW	-0.18±0.57	-0.19±0.58	-0.21±0.59	0.77±0.27	-0.38±0.43	-0.28±0.53	-0.23±0.56	-0.36±0.55
Plug-DR	-0.48±0.44	-0.50±0.44	-0.54±0.39	0.67±0.39	-0.48±0.42	-0.58±0.34	-0.44±0.52	-0.53±0.52
Plug-R	**0.94**±0.13	**0.93**±0.15	**0.93**±0.13	0.95±0.07	**0.87**±0.27	**0.89**±0.30	**0.88**±0.28	0.78±0.45
Plug-RA	-0.61±0.29	-0.62±0.28	-0.63±0.28	0.59±0.42	-0.54±0.35	-0.64±0.28	-0.58±0.43	-0.70±0.39
Pseudo-DR	-0.57±0.33	-0.59±0.32	-0.60±0.31	0.49±0.42	-0.52±0.38	-0.65±0.25	-0.59±0.44	-0.73±0.32
Pseudo-R	0.84±0.20	0.81±0.30	0.85±0.21	0.83±0.25	0.61±0.41	0.71±0.39	0.44±0.56	-0.08±0.66
Pseudo-IF	-0.49±0.36	-0.52±0.35	-0.55±0.33	0.06±0.52	-0.38±0.43	-0.55±0.33	-0.55±0.36	-0.53±0.38
Random	0.50±0.18	0.49±0.17	0.47±0.16	0.31±0.19	0.47±0.20	0.49±0.17	0.47±0.18	0.36±0.25
Fact	0.38±0.15	0.38±0.14	0.37±0.14	0.23±0.20	0.35±0.17	0.38±0.14	0.35±0.16	0.23±0.18
Matching	-0.63±0.26	-0.64±0.25	-0.66±0.24	0.47±0.47	-0.55±0.33	-0.67±0.24	-0.70±0.31	-0.77±0.29
DRM	0.75±0.18	0.75±0.19	0.76±0.17	0.91±0.07	0.61±0.24	0.76±0.18	0.81±0.20	**0.79**±0.17

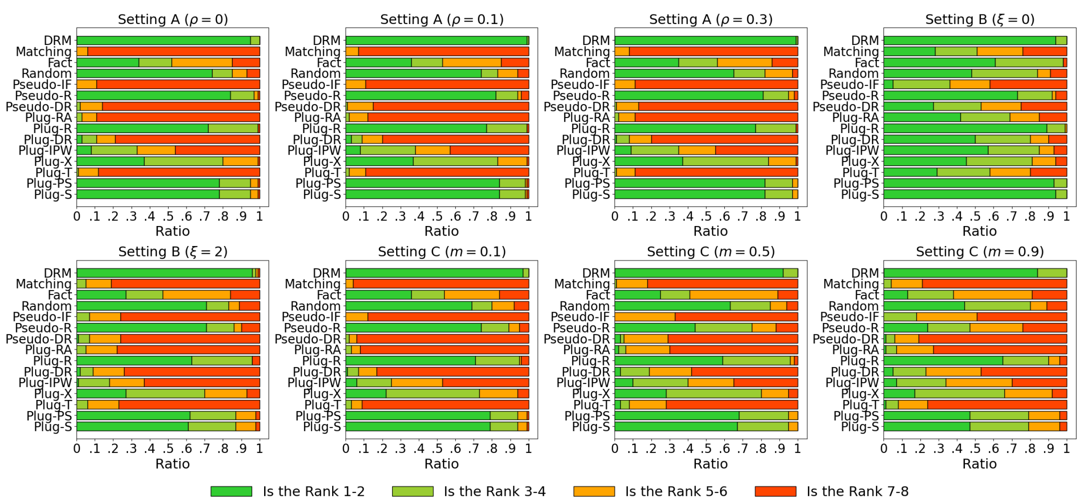

Figure 3: The stacked bar chart showing the distribution of the selected estimator's rank for each evaluation metric across rank intervals: [1-2], [3-4], [5-6], [7-8]. The estimator selection is over 8 candidate estimators, with the underlying ML model for CATE estimator fixed as SVM.

C.4 The Complementary Results with 8 Candidate Pool (RF)

This Section reports the complementary results for 8 candidate CATE estimators, where the candidate pool contains 1 ML models (RF) $\times$ 8 learners (S-, T-, PS-, IPW-, X-, DR-, R-, RA-).

Table 10: Comparison of Regret for different selectors across Settings A, B, and C (Note that B ($\xi = 1$) matches A ($\rho = 0.1$)), with base model for CATE estimator fixed as RF. Reported values (mean $\pm$ standard deviation) are computed over 100 experiments. Smaller is better.

	A ($\rho = 0$)	A ($\rho = 0.1$)	A ($\rho = 0.3$)	B ($\xi = 0$)	B ($\xi = 2$)	C ($m = 0.1$)	C ($m = 0.5$)	C ($m = 0.9$)
Plug-S	1.90±9.49	**0.60**±3.11	1.20±6.81	2.60±8.27	3.56±15.59	**0.20**±0.93	2.62±10.58	6.68±14.93
Plug-PS	1.90±9.49	**0.60**±3.11	**1.20**±6.81	2.46±8.08	3.55±15.59	**0.20**±0.93	2.62±10.58	6.68±14.93
Plug-T	38.66±29.66	39.03±27.58	38.74±25.87	19.97±31.76	36.64±30.38	43.99±33.40	34.41±26.65	32.36±24.71
Plug-X	4.21±15.47	3.47±13.31	3.34±13.84	8.75±14.76	6.37±19.04	6.12±18.12	8.05±18.44	11.64±19.31
Plug-IPW	24.79±28.74	24.62±29.12	21.84±27.40	14.86±29.66	27.83±29.06	29.80±31.69	18.54±24.80	21.40±23.38
Plug-DR	35.45±30.28	37.25±29.08	35.24±27.52	17.67±31.89	34.98±30.96	38.93±24.27	31.07±27.55	25.94±26.01
Plug-R	**1.54**±8.05	1.79±13.03	2.98±21.18	**1.79**±8.91	**3.27**±9.29	0.45±2.93	2.93±13.39	**3.39**±13.54
Plug-RA	39.01±29.53	39.75±27.18	38.83±26.01	18.39±31.81	36.40±30.41	44.30±33.03	34.44±26.88	32.66±24.92
Pseudo-DR	39.12±30.37	38.99±27.69	37.70±27.20	20.18±31.52	36.40±30.55	45.96±32.71	33.75±26.91	32.68±24.73
Pseudo-R	3.29±16.11	3.90±16.84	2.88±21.17	5.44±16.35	19.28±91.50	3.71±11.50	7.08±17.83	11.22±21.11
Pseudo-IF	33.88±27.42	35.58±28.96	34.16±25.83	24.98±26.82	31.98±28.50	34.60±30.16	35.63±27.64	20.77±25.06
Random	18.71±49.28	14.81±28.40	13.16±27.63	9.54±16.13	16.83±41.52	15.81±29.80	16.17±34.19	15.48±23.93
Fact	43.51±31.31	41.58±31.02	41.09±30.15	11.15±18.29	39.60±31.81	35.37±29.91	35.29±32.17	38.58±32.67
Matching	40.21±28.40	41.88±25.64	39.68±25.96	19.95±30.82	36.69±27.51	38.61±25.34	35.39±24.92	29.01±23.52
DRM	2.56±9.32	2.97±14.11	1.42±7.01	3.95±16.96	4.43±11.12	2.12±7.40	**1.98**±9.22	4.00±9.93

Table 11: Comparison of PEHE for different selectors across Settings A, B, and C (Note that B ($\xi = 1$) matches A ($\rho = 0.1$)), with base model for CATE estimator fixed as RF. Reported values (mean $\pm$ standard deviation) are computed over 100 experiments. Smaller is better.

	A ($\rho = 0$)	A ($\rho = 0.1$)	A ($\rho = 0.3$)	B ($\xi = 0$)	B ($\xi = 2$)	C ($m = 0.1$)	C ($m = 0.5$)	C ($m = 0.9$)
Plug-S	11.75±17.60	**11.58**±15.79	11.62±15.75	4.36±8.41	28.74±26.74	**10.98**±13.84	14.76±19.66	21.52±21.67
Plug-PS	11.75±17.60	**11.58**±15.79	**11.62**±15.75	4.22±8.23	28.73±26.73	**10.98**±13.84	14.76±19.66	21.52±21.67
Plug-T	48.50±27.93	50.00±25.45	49.16±25.59	21.73±31.80	61.83±24.43	54.76±31.74	46.55±27.02	47.20±25.50
Plug-X	14.05±20.03	14.44±19.47	13.76±18.17	10.51±14.67	31.55±27.11	16.89±21.18	20.19±24.61	26.48±25.28
Plug-IPW	34.63±30.01	35.59±30.48	32.26±29.88	16.62±29.67	53.01±28.23	40.57±31.34	30.67±27.64	36.24±27.13
Plug-DR	45.30±29.36	48.23±27.52	45.66±27.56	19.43±31.93	60.16±26.42	49.70±23.62	43.20±28.33	40.77±27.88
Plug-R	**11.38**±17.37	12.76±20.59	13.40±27.38	**3.55**±9.02	**28.45**±24.90	11.22±14.33	15.07±22.44	**18.23**±21.59
Plug-RA	48.85±27.90	50.72±25.18	49.25±25.72	20.15±31.84	61.59±24.64	55.07±31.29	46.58±27.09	47.50±25.77
Pseudo-DR	48.96±28.75	49.96±25.93	48.12±27.23	21.94±31.57	61.58±24.59	56.74±30.67	45.89±27.41	47.51±26.19
Pseudo-R	13.14±24.19	14.87±25.57	13.30±27.24	7.20±16.40	44.46±97.86	14.48±19.71	19.22±25.26	26.06±26.53
Pseudo-IF	43.73±25.42	46.55±26.65	44.58±23.95	26.74±26.81	57.16±21.46	45.37±29.03	47.76±27.14	35.61±26.79
Random	28.55±50.55	25.79±31.08	23.58±30.03	11.30±16.09	42.01±47.38	26.59±33.53	28.31±36.36	30.32±27.97
Fact	53.35±33.60	52.56±31.51	51.51±32.29	12.91±18.26	64.78±33.96	46.14±33.02	47.42±35.97	53.42±33.90
Matching	50.06±27.25	52.85±24.51	50.10±26.10	21.71±30.77	61.87±22.23	49.38±23.58	47.53±25.97	43.85±25.60
DRM	12.41±19.23	13.94±21.69	11.84±15.52	5.71±17.04	29.61±26.37	12.89±17.57	**14.12**±18.47	18.84±19.16

Table 12: Comparison of rank correlation for different selectors across Settings A, B, and C (Note that B ($\xi = 1$) matches A ($\rho = 0.1$)), with base model for CATE estimator fixed as RF. Reported values (mean $\pm$ standard deviation) are computed over 100 experiments. Smaller is better.

	A ($\rho = 0$)	A ($\rho = 0.1$)	A ($\rho = 0.3$)	B ($\xi = 0$)	B ($\xi = 2$)	C ($m = 0.1$)	C ($m = 0.5$)	C ($m = 0.9$)
Plug-S	**0.90**±0.13	**0.88**±0.16	0.90±0.16	0.80±0.16	0.90±0.15	**0.90**±0.14	0.87±0.15	0.84±0.20
Plug-PS	**0.90**±0.13	**0.88**±0.16	0.90±0.16	0.80±0.16	0.90±0.15	**0.90**±0.14	0.87±0.15	0.84±0.20
Plug-T	0.33±0.47	0.34±0.45	0.30±0.47	0.57±0.35	0.32±0.46	0.27±0.43	0.34±0.50	0.32±0.46
Plug-X	0.89±0.14	0.87±0.19	0.88±0.19	0.72±0.23	0.89±0.15	0.87±0.16	0.82±0.23	0.78±0.25
Plug-IPW	0.60±0.42	0.61±0.44	0.63±0.41	0.64±0.33	0.53±0.41	0.52±0.43	0.66±0.37	0.64±0.35
Plug-DR	0.40±0.47	0.39±0.46	0.39±0.46	0.63±0.31	0.35±0.47	0.34±0.43	0.44±0.50	0.50±0.43
Plug-R	0.89±0.15	0.88±0.16	**0.90**±0.15	**0.81**±0.16	**0.90**±0.14	0.90±0.13	**0.88**±0.14	**0.87**±0.19
Plug-RA	0.33±0.47	0.34±0.45	0.31±0.47	0.59±0.34	0.32±0.46	0.26±0.42	0.34±0.50	0.32±0.48
Pseudo-DR	0.32±0.47	0.33±0.45	0.31±0.47	0.56±0.36	0.32±0.46	0.27±0.41	0.34±0.50	0.32±0.48
Pseudo-R	0.80±0.21	0.79±0.24	0.81±0.21	0.76±0.19	0.74±0.30	0.77±0.23	0.75±0.23	0.72±0.27
Pseudo-IF	0.48±0.36	0.44±0.39	0.48±0.36	0.41±0.42	0.54±0.33	0.42±0.34	0.41±0.38	0.53±0.41
Random	0.56±0.20	0.57±0.18	0.58±0.16	0.28±0.17	0.56±0.18	0.55±0.20	0.51±0.18	0.43±0.26
Fact	0.39±0.14	0.41±0.13	0.39±0.14	0.14±0.12	0.40±0.15	0.41±0.15	0.35±0.17	0.26±0.23
Matching	0.30±0.44	0.29±0.42	0.29±0.45	0.59±0.35	0.35±0.41	0.30±0.40	0.34±0.46	0.37±0.46
DRM	0.72±0.23	0.75±0.21	0.77±0.18	0.74±0.22	0.73±0.22	0.74±0.18	0.74±0.20	0.70±0.20

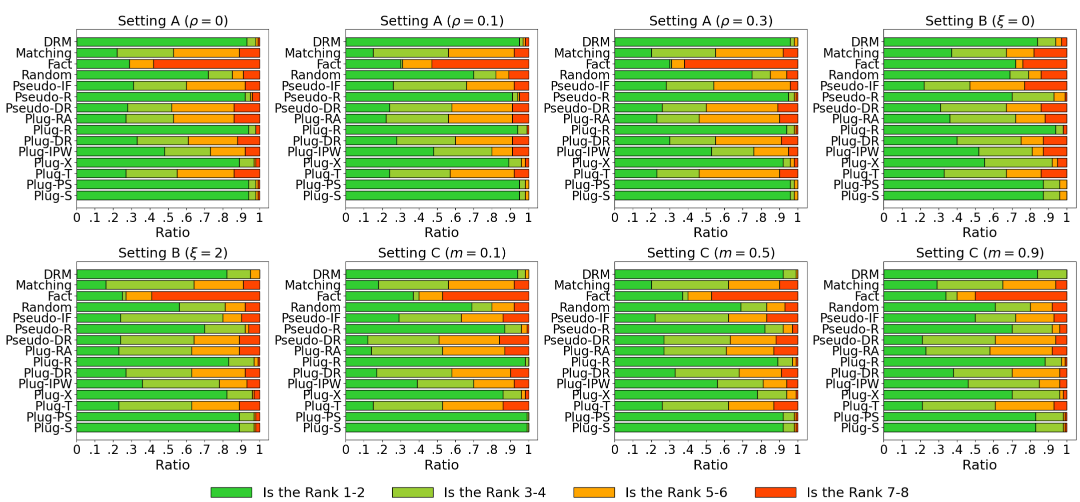

Figure 4: The stacked bar chart showing the distribution of the selected estimator's rank for each evaluation metric across rank intervals: [1-2], [3-4], [5-6], [7-8]. The estimator selection is over 8 candidate estimators, with the underlying ML model for CATE estimator fixed as RF.

NeurIPS Paper Checklist

1. **Claims**

   Question: Do the main claims made in the abstract and introduction accurately reflect the paper's contributions and scope?

   Answer: [Yes]

   Justification: [Yes]

   Guidelines:

   - The answer NA means that the abstract and introduction do not include the claims made in the paper.
   - The abstract and/or introduction should clearly state the claims made, including the contributions made in the paper and important assumptions and limitations. A No or NA answer to this question will not be perceived well by the reviewers.
   - The claims made should match theoretical and experimental results, and reflect how much the results can be expected to generalize to other settings.
   - It is fine to include aspirational goals as motivation as long as it is clear that these goals are not attained by the paper.

2. **Limitations**

   Question: Does the paper discuss the limitations of the work performed by the authors?

   Answer: [Yes]

   Justification: [Yes]

   Guidelines:

   - The answer NA means that the paper has no limitation while the answer No means that the paper has limitations, but those are not discussed in the paper.
   - The authors are encouraged to create a separate "Limitations" section in their paper.
   - The paper should point out any strong assumptions and how robust the results are to violations of these assumptions (e.g., independence assumptions, noiseless settings, model well-specification, asymptotic approximations only holding locally). The authors should reflect on how these assumptions might be violated in practice and what the implications would be.
   - The authors should reflect on the scope of the claims made, e.g., if the approach was only tested on a few datasets or with a few runs. In general, empirical results often depend on implicit assumptions, which should be articulated.
   - The authors should reflect on the factors that influence the performance of the approach. For example, a facial recognition algorithm may perform poorly when image resolution is low or images are taken in low lighting. Or a speech-to-text system might not be used reliably to provide closed captions for online lectures because it fails to handle technical jargon.
   - The authors should discuss the computational efficiency of the proposed algorithms and how they scale with dataset size.
   - If applicable, the authors should discuss possible limitations of their approach to address problems of privacy and fairness.
   - While the authors might fear that complete honesty about limitations might be used by reviewers as grounds for rejection, a worse outcome might be that reviewers discover limitations that aren't acknowledged in the paper. The authors should use their best judgment and recognize that individual actions in favor of transparency play an important role in developing norms that preserve the integrity of the community. Reviewers will be specifically instructed to not penalize honesty concerning limitations.

3. **Theory Assumptions and Proofs**

   Question: For each theoretical result, does the paper provide the full set of assumptions and a complete (and correct) proof?

   Answer: [Yes]

Justification: [Yes]

Guidelines:

- The answer NA means that the paper does not include theoretical results.
- All the theorems, formulas, and proofs in the paper should be numbered and cross-referenced.
- All assumptions should be clearly stated or referenced in the statement of any theorems.
- The proofs can either appear in the main paper or the supplemental material, but if they appear in the supplemental material, the authors are encouraged to provide a short proof sketch to provide intuition.
- Inversely, any informal proof provided in the core of the paper should be complemented by formal proofs provided in appendix or supplemental material.
- Theorems and Lemmas that the proof relies upon should be properly referenced.

4. **Experimental Result Reproducibility**

   Question: Does the paper fully disclose all the information needed to reproduce the main experimental results of the paper to the extent that it affects the main claims and/or conclusions of the paper (regardless of whether the code and data are provided or not)?

   Answer: [Yes]

   Justification: [Yes]

   Guidelines:

   - The answer NA means that the paper does not include experiments.
   - If the paper includes experiments, a No answer to this question will not be perceived well by the reviewers: Making the paper reproducible is important, regardless of whether the code and data are provided or not.
   - If the contribution is a dataset and/or model, the authors should describe the steps taken to make their results reproducible or verifiable.
   - Depending on the contribution, reproducibility can be accomplished in various ways. For example, if the contribution is a novel architecture, describing the architecture fully might suffice, or if the contribution is a specific model and empirical evaluation, it may be necessary to either make it possible for others to replicate the model with the same dataset, or provide access to the model. In general. releasing code and data is often one good way to accomplish this, but reproducibility can also be provided via detailed instructions for how to replicate the results, access to a hosted model (e.g., in the case of a large language model), releasing of a model checkpoint, or other means that are appropriate to the research performed.
   - While NeurIPS does not require releasing code, the conference does require all submissions to provide some reasonable avenue for reproducibility, which may depend on the nature of the contribution. For example
     (a) If the contribution is primarily a new algorithm, the paper should make it clear how to reproduce that algorithm.
     (b) If the contribution is primarily a new model architecture, the paper should describe the architecture clearly and fully.
     (c) If the contribution is a new model (e.g., a large language model), then there should either be a way to access this model for reproducing the results or a way to reproduce the model (e.g., with an open-source dataset or instructions for how to construct the dataset).
     (d) We recognize that reproducibility may be tricky in some cases, in which case authors are welcome to describe the particular way they provide for reproducibility. In the case of closed-source models, it may be that access to the model is limited in some way (e.g., to registered users), but it should be possible for other researchers to have some path to reproducing or verifying the results.

5. **Open access to data and code**

   Question: Does the paper provide open access to the data and code, with sufficient instructions to faithfully reproduce the main experimental results, as described in supplemental material?

Answer: [Yes]

Justification: [Yes]

Guidelines:

- The answer NA means that paper does not include experiments requiring code.
- Please see the NeurIPS code and data submission guidelines (`https://nips.cc/public/guides/CodeSubmissionPolicy`) for more details.
- While we encourage the release of code and data, we understand that this might not be possible, so "No" is an acceptable answer. Papers cannot be rejected simply for not including code, unless this is central to the contribution (e.g., for a new open-source benchmark).
- The instructions should contain the exact command and environment needed to run to reproduce the results. See the NeurIPS code and data submission guidelines (`https://nips.cc/public/guides/CodeSubmissionPolicy`) for more details.
- The authors should provide instructions on data access and preparation, including how to access the raw data, preprocessed data, intermediate data, and generated data, etc.
- The authors should provide scripts to reproduce all experimental results for the new proposed method and baselines. If only a subset of experiments are reproducible, they should state which ones are omitted from the script and why.
- At submission time, to preserve anonymity, the authors should release anonymized versions (if applicable).
- Providing as much information as possible in supplemental material (appended to the paper) is recommended, but including URLs to data and code is permitted.

6. **Experimental Setting/Details**

Question: Does the paper specify all the training and test details (e.g., data splits, hyper-parameters, how they were chosen, type of optimizer, etc.) necessary to understand the results?

Answer: [Yes]

Justification: [Yes]

Guidelines:

- The answer NA means that the paper does not include experiments.
- The experimental setting should be presented in the core of the paper to a level of detail that is necessary to appreciate the results and make sense of them.
- The full details can be provided either with the code, in appendix, or as supplemental material.

7. **Experiment Statistical Significance**

Question: Does the paper report error bars suitably and correctly defined or other appropriate information about the statistical significance of the experiments?

Answer: [Yes]

Justification: [Yes]

Guidelines:

- The answer NA means that the paper does not include experiments.
- The authors should answer "Yes" if the results are accompanied by error bars, confidence intervals, or statistical significance tests, at least for the experiments that support the main claims of the paper.
- The factors of variability that the error bars are capturing should be clearly stated (for example, train/test split, initialization, random drawing of some parameter, or overall run with given experimental conditions).
- The method for calculating the error bars should be explained (closed form formula, call to a library function, bootstrap, etc.)
- The assumptions made should be given (e.g., Normally distributed errors).
- It should be clear whether the error bar is the standard deviation or the standard error of the mean.

- It is OK to report 1-sigma error bars, but one should state it. The authors should preferably report a 2-sigma error bar than state that they have a 96% CI, if the hypothesis of Normality of errors is not verified.
- For asymmetric distributions, the authors should be careful not to show in tables or figures symmetric error bars that would yield results that are out of range (e.g. negative error rates).
- If error bars are reported in tables or plots, The authors should explain in the text how they were calculated and reference the corresponding figures or tables in the text.

8. **Experiments Compute Resources**

   Question: For each experiment, does the paper provide sufficient information on the computer resources (type of compute workers, memory, time of execution) needed to reproduce the experiments?

   Answer: [Yes]

   Justification: [Yes]

   Guidelines:

   - The answer NA means that the paper does not include experiments.
   - The paper should indicate the type of compute workers CPU or GPU, internal cluster, or cloud provider, including relevant memory and storage.
   - The paper should provide the amount of compute required for each of the individual experimental runs as well as estimate the total compute.
   - The paper should disclose whether the full research project required more compute than the experiments reported in the paper (e.g., preliminary or failed experiments that didn't make it into the paper).

9. **Code Of Ethics**

   Question: Does the research conducted in the paper conform, in every respect, with the NeurIPS Code of Ethics `https://neurips.cc/public/EthicsGuidelines`?

   Answer: [Yes]

   Justification: [Yes]

   Guidelines:

   - The answer NA means that the authors have not reviewed the NeurIPS Code of Ethics.
   - If the authors answer No, they should explain the special circumstances that require a deviation from the Code of Ethics.
   - The authors should make sure to preserve anonymity (e.g., if there is a special consideration due to laws or regulations in their jurisdiction).

10. **Broader Impacts**

    Question: Does the paper discuss both potential positive societal impacts and negative societal impacts of the work performed?

    Answer: [NA]

    Justification: [NA]

    Guidelines:

    - The answer NA means that there is no societal impact of the work performed.
    - If the authors answer NA or No, they should explain why their work has no societal impact or why the paper does not address societal impact.
    - Examples of negative societal impacts include potential malicious or unintended uses (e.g., disinformation, generating fake profiles, surveillance), fairness considerations (e.g., deployment of technologies that could make decisions that unfairly impact specific groups), privacy considerations, and security considerations.
    - The conference expects that many papers will be foundational research and not tied to particular applications, let alone deployments. However, if there is a direct path to any negative applications, the authors should point it out. For example, it is legitimate to point out that an improvement in the quality of generative models could be used to

generate deepfakes for disinformation. On the other hand, it is not needed to point out that a generic algorithm for optimizing neural networks could enable people to train models that generate Deepfakes faster.

- The authors should consider possible harms that could arise when the technology is being used as intended and functioning correctly, harms that could arise when the technology is being used as intended but gives incorrect results, and harms following from (intentional or unintentional) misuse of the technology.
- If there are negative societal impacts, the authors could also discuss possible mitigation strategies (e.g., gated release of models, providing defenses in addition to attacks, mechanisms for monitoring misuse, mechanisms to monitor how a system learns from feedback over time, improving the efficiency and accessibility of ML).

11. **Safeguards**

Question: Does the paper describe safeguards that have been put in place for responsible release of data or models that have a high risk for misuse (e.g., pretrained language models, image generators, or scraped datasets)?

Answer: [NA]

Justification: [NA]

Guidelines:

- The answer NA means that the paper poses no such risks.
- Released models that have a high risk for misuse or dual-use should be released with necessary safeguards to allow for controlled use of the model, for example by requiring that users adhere to usage guidelines or restrictions to access the model or implementing safety filters.
- Datasets that have been scraped from the Internet could pose safety risks. The authors should describe how they avoided releasing unsafe images.
- We recognize that providing effective safeguards is challenging, and many papers do not require this, but we encourage authors to take this into account and make a best faith effort.

12. **Licenses for existing assets**

Question: Are the creators or original owners of assets (e.g., code, data, models), used in the paper, properly credited and are the license and terms of use explicitly mentioned and properly respected?

Answer: [NA]

Justification: [NA]

Guidelines:

- The answer NA means that the paper does not use existing assets.
- The authors should cite the original paper that produced the code package or dataset.
- The authors should state which version of the asset is used and, if possible, include a URL.
- The name of the license (e.g., CC-BY 4.0) should be included for each asset.
- For scraped data from a particular source (e.g., website), the copyright and terms of service of that source should be provided.
- If assets are released, the license, copyright information, and terms of use in the package should be provided. For popular datasets, `paperswithcode.com/datasets` has curated licenses for some datasets. Their licensing guide can help determine the license of a dataset.
- For existing datasets that are re-packaged, both the original license and the license of the derived asset (if it has changed) should be provided.
- If this information is not available online, the authors are encouraged to reach out to the asset's creators.

13. **New Assets**

Question: Are new assets introduced in the paper well documented and is the documentation provided alongside the assets?

Answer: [NA]

Justification: [NA]

Guidelines:

- The answer NA means that the paper does not release new assets.
- Researchers should communicate the details of the dataset/code/model as part of their submissions via structured templates. This includes details about training, license, limitations, etc.
- The paper should discuss whether and how consent was obtained from people whose asset is used.
- At submission time, remember to anonymize your assets (if applicable). You can either create an anonymized URL or include an anonymized zip file.

14. **Crowdsourcing and Research with Human Subjects**

Question: For crowdsourcing experiments and research with human subjects, does the paper include the full text of instructions given to participants and screenshots, if applicable, as well as details about compensation (if any)?

Answer: [NA]

Justification: [NA]

Guidelines:

- The answer NA means that the paper does not involve crowdsourcing nor research with human subjects.
- Including this information in the supplemental material is fine, but if the main contribution of the paper involves human subjects, then as much detail as possible should be included in the main paper.
- According to the NeurIPS Code of Ethics, workers involved in data collection, curation, or other labor should be paid at least the minimum wage in the country of the data collector.

15. **Institutional Review Board (IRB) Approvals or Equivalent for Research with Human Subjects**

Question: Does the paper describe potential risks incurred by study participants, whether such risks were disclosed to the subjects, and whether Institutional Review Board (IRB) approvals (or an equivalent approval/review based on the requirements of your country or institution) were obtained?

Answer: [NA]

Justification: [NA]

Guidelines:

- The answer NA means that the paper does not involve crowdsourcing nor research with human subjects.
- Depending on the country in which research is conducted, IRB approval (or equivalent) may be required for any human subjects research. If you obtained IRB approval, you should clearly state this in the paper.
- We recognize that the procedures for this may vary significantly between institutions and locations, and we expect authors to adhere to the NeurIPS Code of Ethics and the guidelines for their institution.
- For initial submissions, do not include any information that would break anonymity (if applicable), such as the institution conducting the review.