Robust estimation of heterogeneous treatment effects in randomized trials leveraging external data

Rickard Karlsson [1] **Piersilvio De Bartolomeis** [2] **Issa J. Dahabreh** [3] **Jesse H. Krijthe** [1]

Abstract

Randomized trials are typically designed to detect average treatment effects but often lack the statistical power to uncover individual-level treatment effect heterogeneity, limiting their value for personalized decision-making. To address this, we propose the QR-learner, a model-agnostic learner that estimates conditional average treatment effects (CATE) within the trial population by leveraging external data from other trials or observational studies. The method is robust: it has the potential to reduce the CATE prediction mean squared error while maintaining consistency, even when the external data is not aligned with the trial. We examine the performance of our approach in simulation studies, and find that it is robust and reduces CATE estimation mean-squared error.

1. Introduction

By randomly assigning the interventions of interest, randomized trials are unique in their ability to estimate causal effects with little reliance on untestable assumptions. However, their high cost often limits sample size, which in turn limits the precision of statistical inferences that can be drawn from trial data. This limitation becomes particularly problematic when the goal extends beyond estimating an average treatment effect to understanding treatment effect heterogeneity (Lagakos et al., 2006), which is an essential step toward personalized decision-making for the underlying population represented by the trial. A central quantity for this purpose is the conditional average treatment effect (CATE), which captures how treatment effects depend on individual-level covariates (Künzel et al., 2019). However, the estimation of CATEs for different subgroups is more challenging than the estimation of average effects; therefore, the data from trials powered to detect average treatment effects are typically not adequate for the precise estimation of CATEs (Dahabreh et al., 2016). As a result, accurately estimating CATEs within a trial population remains a difficult yet important challenge.

In recent years, there has been growing interest in augmenting randomized trials with external data, such as from another randomized trial or an observational study, mainly in the context of improving average treatment effect estimation (van Rosmalen et al., 2018; Jahanshahi et al., 2021). A key challenge in this setting is to properly account for systematic differences and distribution shifts between the trial population and the population underlying the external data (Ung et al., 2024). Central to this challenge is the question of transportability – whether causal quantities such as the CATE remain invariant across populations (Bareinboim and Pearl, 2016; Dahabreh and Hernán, 2019). In this work, we study the analogous problem of augmenting CATE estimation in a randomized trial using external data, focusing on settings where the trial and external populations may be misaligned, due to both transportability violations and unmeasured confounding in the external data.

Contributions We propose the QR-learner, a model-agnostic learner that improves estimation of the CATE in the population underlying a trial by leveraging external data from other trials or observational studies. We prove this learner is robust even when the external data are not aligned with the trial data: it recovers the true CATE even when external data come from a population which is not transportable with the trial population or are affected by uncontrolled confounding, while at the same time it can reduce the estimated CATE mean-squared error compared to using trial data alone if the external data is sufficiently aligned. Our approach combines insights from domain adaptation – specifically, how to improve performance in a target domain (here, the trial population) by transferring knowledge from a separate source domain – with known robustness properties of certain existing CATE learners. We evaluate the QR-learner in simulations and find that our method is robust and improves CATE estimation mean-squared error.

[1]Delft University of Technology [2]ETH Zürich [3]Harvard T.H. Chan School of Public Health. Correspondence to: Rickard Karlsson <r.k.a.karlsson@tudelft.nl>.

Proceedings of the 42^{nd} International Conference on Machine Learning, Vancouver, Canada. PMLR 267, 2025.

2. Related works

Our work builds upon a rich and growing literature on CATE learners. Our proposed learner is most closely related to model-agnostic "meta-learners" which allow for estimating the CATE and any nuisance models using any supervised learning algorithm (Künzel et al., 2019; Nie and Wager, 2021; Kennedy, 2023). However, most existing model-agnostic learners are tailored to settings where data is drawn from a single source, such as a single randomized trial or an observational study. More recently, several CATE learners have been proposed for multi-source settings, often relying on the assumption that the CATE is transportable across the underlying populations (Kallus et al., 2018; Hatt et al., 2022; Wu and Yang, 2022; Shyr et al., 2023; Schweisthal et al., 2024). Although some of these approaches might be adapted to specifically learn the CATE in the trial under non-transportability, to our knowledge only Asiaee et al. (2023) explicitly address this setting.

Our work draws on recent developments in the data integration literature for average treatment effect estimation in trials using data from an external population. This work has emphasized robustness to integrating external data misaligned with the trial data (Schuler et al., 2022; Huang et al., 2023; Liao et al., 2023; Karlsson et al., 2024; De Bartolomeis et al., 2025). In particular, we will adapt ideas from the randomization-aware estimator framework proposed by Karlsson et al. (2024) to construct CATE learners that are also robust to misaligned external data.

3. Problem setting

Notation Let $X \in \mathcal{X}$ denote baseline (pre-treatment) covariates; S the binary indicator of data source ($S = 1$ for trial participants; $S = 0$ for individuals in the external data); A the binary indicator for treatment assignment ($A = 1$ for the experimental treatment; $A = 0$ denotes the control); and $Y \in \mathcal{Y}$ the outcome (continuous, binary, or count) collected at the end of follow-up. Throughout, we use italic capital letters to denote random variables and lowercase letters for their specific values. We write $f(\cdot)$ to denote the density functions of random variables.

Study design and data structure We consider a non-nested trial design where the trial and external data are separately obtained and modeled as simple random samples from different populations, obtained with unknown and possibly unequal sampling fractions (Dahabreh et al., 2021). For observation i with $S_i = s$, the data are modeled as i.i.d., conditional on study source, with the random tuple $O_i = (X_i, S_i = s, A_i, Y_i)$ for $i = 1, \ldots, n_s$, where n_s is the number of observations from source $S = s$. The composite dataset has total sample size $n = n_1 + n_0$, where the proportions of trial and external participants in the composite dataset may not reflect the size of their underlying populations. In the trial, treatment is randomly assigned according to the propensity score $e(X) = \Pr(A = 1 \mid X, S = 1)$ which is assumed to be known (Rosenbaum and Rubin, 1983). As $n \to \infty$, we assume the ratios of the trial and external data sample sizes to the total sample size converge to positive constants, i.e., $n_s/n \to q_s > 0$.

3.1. Identification of causal effects

To define the causal quantity of interest, we use potential (counterfactual) outcomes (Rubin, 1974). For individual i and for $a \in \{0, 1\}$, the potential outcome Y_i^a denotes the outcome under intervention to set treatment A to a, possibly contrary to fact. Our goal is to estimate the conditional average treatment effect (CATE) in the population underlying the trial, $\tau(x) = \mathbb{E}[Y^1 - Y^0 \mid X = x, S = 1]$. Under standard conditions, the CATE function $\tau(x)$ is identifiable from data in the trial population.

Condition 3.1 (Consistency)**.** If $A_i = a$, then $Y_i^a = Y_i$ for every individual i and treatment $a \in \{0, 1\}$.

Condition 3.2 (Strong ignorability in the trial population)**.** *Positivity in trial:* for each treatment $a \in \{0, 1\}$, if $f(x, S = 1) \neq 0$, then $\Pr(A = a|X = x, S = 1) > 0$. *Conditional exchangeability in trial:* for each $a \in \{0, 1\}$, $Y^a \perp\!\!\!\perp A|(X, S = 1)$.

Conditions 3.1 and 3.2 are typically supported by a well-designed randomized trial and together suffice to identify the CATE function $\tau(x) = g_1(x) - g_0(x)$, where $g_a(x) = \mathbb{E}[Y \mid X = x, A = a, S = 1]$. However, it is common to assume additional conditions to enable identification and estimation of $\tau(x)$ using both the trial and external data.

Condition 3.3 (Strong ignorability in the external population)**.** *Positivity in external population:* for each treatment $a \in \{0, 1\}$, if $f(x, S = 0) \neq 0$, then $\Pr(A = a|X = x, S = 0) > 0$. *Conditional exchangeability in external population:* for each $a \in \{0, 1\}$, $Y^a \perp\!\!\!\perp A|(X, S = 0)$.

Condition 3.4 (Transportability)**.** *Conditional exchangeability between populations*: For each $a \in \{0, 1\}$, $Y^a \perp\!\!\!\perp S \mid X$. *Positivity of selection*: $f(X = x, S = 0) > 0 \Rightarrow f(X = x, S = 1) > 0$.

The above two conditions can be controversial, especially if the external data originate from an observational study, because these conditions are uncertain and typically require substantial domain expertise to justify. Notably, Condition 3.1 to 3.4 together have testable implications that can be empirically assessed to falsify them, see e.g. Hussain

et al. (2023); De Bartolomeis et al. (2024); Dahabreh et al. (2024). This can be used in particular to evaluate Conditions 3.3 and 3.4 because Condition 3.1 and 3.2 are supported by the trial's experimental design. Nonetheless, performing such falsification tests remains an inherently difficult task (Fawkes et al., 2025).

4. Leveraging external data to learn the CATE in the randomized trial

4.1. A class of robust pseudo-outcomes

Our goal is to learn a CATE function from a class of candidates $\mathcal{F}$ that minimizes the risk relative to the true CATE function, namely $\arg\min_{\tilde{\tau}\in\mathcal{F}} R^*(\tilde{\tau})$ where $R^*(\tilde{\tau}) = \mathbb{E}[(\tau(X) - \tilde{\tau}(X))^2 \mid S = 1]$. However, as we cannot minimize $R^*(\tilde{\tau})$ directly because the true CATE $\tau(X)$ is unknown, we study a class of CATE learners obtained by minimizing a pseudo-risk (Foster and Syrgkanis, 2023),

$$R(\tilde{\tau};\eta) = \mathbb{E}[(\psi(O;\eta) - \tilde{\tau}(X))^2 \mid S = 1]\,, \qquad (1)$$

where we introduce an auxiliary random variable, sometimes referred to as a pseudo-outcome:

$$\psi(O_i;\eta) = w(O_i)\cdot\big(Y_i - h_{A_i}(X_i)\big) + h_1(X_i) - h_0(X_i) \quad (2)$$

with $w(O_i) = \frac{A_i - e(X_i)}{e(X_i)(1-e(X_i))}$, which is indexed by some nuisance models $\eta = \{h_1, h_0\}$, where $h_1 : \mathcal{X} \to \mathbb{R}$ and $h_0 : \mathcal{X} \to \mathbb{R}$ are functions defined on the covariate space $\mathcal{X}$.

Depending on our choice of η, we obtain different CATE learners when minimizing (1). For instance, if $\eta = \{0, 0\}$, we obtain the (inverse) propensity weighted learner, refereed to as the PW-learner by Curth and Van der Schaar (2021). Meanwhile, if $\eta = \{g_1, g_0\}$, where $g_a = \mathbb{E}[Y \mid X, A = a, S = 1]$, we obtain the DR-learner (Kennedy, 2023). More generally, for any choice of η, we can prove an important robustness property of $R(\tilde{\tau};\eta)$ guaranteed by the study design of the randomized trial.

Theorem 4.1. *Under Condition 3.1 and 3.2 with the propensity score $e(X)$ known, for any choice of η_{fixed} to compute the pseudo-outcomes $\psi(O;\eta_{\text{fixed}})$ where η_{fixed} is held fixed, minimizing (1) always yields the true CATE as its unique solution if $\tau \in \mathcal{F}$; that is, $\tau = \arg\min_{\tilde{\tau}\in\mathcal{F}} R(\tilde{\tau};\eta_{\text{fixed}})$.*

While the above result have appeared in the literature before (see e.g. Morzywolek et al. (2023) and references therein), we derive them for completeness in Appendix B.2. Notably, we denoted the nuisance models η_{fixed} as fixed to emphasize that these results requires that the nuisance models are chosen separately from the dataset used to compute the pseudo-outcome. This can be achieved via cross-fitting, which we provide further details on in the following subsection.

The above theorem guarantees that pseudo-risk $R(\tilde{\tau};\eta)$ is a proper model selection criterion for the CATE regardless of the choice of η. However, the choice of η ultimately still plays a crucial role in selecting the best CATE based on the observed data, due to the uncertainty when minimizing the finite sample analog of $R(\tilde{\tau};\eta)$. In the next section, we present an algorithm that leverages external data to choose η, enhancing the finite-sample performance of $R(\tilde{\tau};\eta)$ as a model selection criterion while preserving its robustness.

4.2. The QR-learner algorithm

In this section, we introduce a novel learner for estimating the CATE, which we call the *Quasi-optimized Randomization-aware* learner, or QR-learner. This method is model-agnostic, allowing it to use any supervised learning algorithm for estimating the CATE function and the nuisance models. To avoid overfitting and fulfill the condition that the estimated nuisance model is obtained separately from the dataset used to compute the pseudo-outcomes, we employ a cross-fitting procedure, where we partition the data as $\mathcal{D} = \mathcal{D}^1 \cup \mathcal{D}^2$, stratified by the study indicator S.

In the first stage, using both trial and external data from $\mathcal{D}^1$, we aim to find the nuisance models $\eta^* = \{h_1^*, h_0^*\}$ such that $h_a^* = \arg\min_{h_a\in\mathcal{H}} L_a(h_a)$ for each $a \in \{0, 1\}$, where the objective is defined as

$$\begin{aligned} \mathcal{L}_a(h_a) &= \mathbb{E}\left[\nu_a(X)\big(Y - h_a(X_i)\big)^2 \mid A = a\right] \\ \nu_a(X) &= \pi_a(X)\left\{\frac{1 - e(X)}{e(X)}\right\}^{2a-1} \end{aligned} \qquad (3)$$

where $\pi(X) = \Pr(S = 1 \mid X, A = a)$ is the conditional probability of participating in the trial, which is a quantity that also needs to be estimated. Notably, under the conditions of strong ignorability in the external population (Condition 3.3) and transportability (Condition 3.4), in addition to those ensured by the randomized trial (Conditions 3.1 and 3.2), Karlsson et al. (2024) showed that a quantity similar to (3) is proportional to the variance of a class of robust average treatment effect estimators, which motivated them to minimize it in order to improve statistical precision. In the context of CATE estimation, we instead show that $L_a(h_a)$ serves as an upper bound to a quantity related to the finite-sample model selection performance of $R(\tilde{\tau};\eta)$. Further details and proofs of these results are provided in Appendix A. Since our procedure aims to minimize this upper bound, we refer to the first stage as a quasi-optimized procedure.

Moreover, the construction of the objective $\mathcal{L}_a(h_a)$ is closely related to ideas from domain adaptation. It can be viewed as the mean-squared error of h_a for predicting Y within the treatment arm $\{A = a\}$, where ν_a re-weights ob-

Table 1: Average root mean-squared error with standard errors reported over 500 repeated runs with a trial dataset size $n_1 = 250$ under different scenarios. Lowest number is marked in bold.

External sample size	100		1000		10000	
Condition 3 and 4 violated?	No	Yes	No	Yes	No	Yes
Predict ATE	0.31 (1e-3)	**0.31** (1e-3)	0.31 (1e-3)	0.31 (1e-3)	0.31 (1e-3)	0.31 (1e-3)
DR-learner	0.28 (4e-3)	0.32 (4e-3)	0.28 (4e-3)	0.32 (4e-3)	0.27 (4e-3)	0.32 (4e-3)
T-learner	0.55 (3e-3)	0.55 (3e-3)	0.55(3e-3)	0.55 (3e-3)	0.55 (3e-3)	0.55 (3e-3)
Pooled T-learner	0.52 (2e-3)	0.60 (3e-3)	0.47 (9e-4)	0.60 (1e-3)	0.33 (7e-4)	0.48 (1e-3)
Asiaee et al. (2023)	0.34 (5e-3)	0.36 (4e-3)	0.24 (4e-3)	0.30 (3e-3)	**0.19** (3e-3)	0.28 (3e-3)
Kallus et al. (2018)	0.71 (1e-2)	0.76 (1e-2)	0.72 (1e-2)	0.77 (1e-2)	0.71 (1e-2)	0.77 (1e-2)
QR-learner	**0.28** (4e-3)	0.32 (4e-3)	**0.23** (3e-3)	**0.29** (3e-3)	**0.19** (3e-3)	**0.27** (3e-3)

servations so that the error reflects the trial population, even though the expectation is taken over both the trial and external data, akin to importance-weighting strategies in transfer learning that correct for distributional shifts between source and target populations (Shimodaira, 2000).

In the second stage, we leverage the robustness results presented in the previous subsection and use only the trial data from the split $\mathcal{D}^2$ to regress on the pseudo-outcomes, to specifically learn the CATE in the trial. We call this step randomization-aware, since it critically relies on using the known propensity score when computing the pseudo-outcomes. We estimate the CATE function by solving $\hat{\tau} = \arg\min_{\tilde{\tau}\in\mathcal{F}} \sum_{i\in D^2:S_i=1}(\psi(O_i;\hat{\eta}^*) - \tilde{\tau}(X_i))^2$, where $\hat{\eta}^*$ is the sample analog estimator of η^*. To efficiently use all available data, we reverse the roles of the splits to obtain a second CATE estimator, and then take the average of the predictions from the two resulting estimators; this procedure naturally extends to more than two data splits if desired.

5. Simulation study

We evaluate the performance of our proposed QR-learner against several baselines. We apply the DR-learner with known propensity scores using only trial data (Kennedy, 2023); the T-learner, which computes CATE estimates as $\hat{g}_1(x) - \hat{g}_0(x)$ with the nuisance models used in the DR-learner; a pooled variant of the T-learner obtained by estimating $\mathbb{E}[Y \mid X, A = 1] - \mathbb{E}[Y \mid X, A = 0]$ using both the trial and external data; the method proposed by Asiaee et al. (2023); and the linear additive bias correction method of Kallus et al. (2018). We assess performance using the root mean squared error (RMSE) relative to the true CATE function $\tau(x)$ in the trial population.

In our simulations, we have a fixed trial size while varying the size of the external dataset. We consider two scenarios: (i) an idealized setting in which both Conditions 3.3 and 3.4 hold – i.e., the external data is unconfounded and transportability holds – and (ii) a more realistic, challenging setting where these assumptions are violated. We use gradient boosting regressors to estimate the nuisance components and a linear regression model to estimate the final conditional average treatment effect; this modeling choice aligns with the data-generating process, where the baseline outcome is a highly nonlinear function of the covariates, while the underlying CATE function is linear. Full implementation details are in Appendix C.

Table 1 shows that the QR-learner consistently achieves the lowest or near-lowest RMSE across all settings, with performance improving as the external sample size increase. The trial-only DR- and T-learners performs no better or worse than simply predicting an estimated average treatment effect (ATE) for all individuals, indicating the difficulty of this task based on trial data alone. While all integrative methods performs best when Conditions 3.3 and 3.4 hold, it is noteworthy when these conditions are violated that our method and the approach of Asiaee et al. (2023) exhibited RMSEs comparable to that of the trial-only DR-learner. Interestingly, the bias correction method of Kallus et al. (2018) underperforms in all settings.

6. Discussion

Our experimental findings demonstrate that the proposed learners for estimating the CATE using external data can effectively reduce the mean squared error of CATE estimates, while remaining robust in scenarios where the external data has unmeasured confounders or transportability is violated. Notably, in cases where the DR-learner failed to outperform a simple baseline that predicts the average treatment effect – thus providing limited value – our proposed learners were still able to achieve better accuracy. While further work is needed to explore the limits of our proposed method, our results highlights the potential value of incorporating external data into analyses of heterogeneous treatment effects in randomized trials.

References

A. Asiaee, C. Di Gravio, Y. Mei, and J. D. Huling. Leveraging observational data for efficient cate estimation in randomized controlled trials. *arXiv preprint arXiv:2306.17478*, 2023.

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

W. Cao, A. A. Tsiatis, and M. Davidian. Improving efficiency and robustness of the doubly robust estimator for a population mean with incomplete data. *Biometrika*, 96 (3):723–734, 2009.

A. Curth and M. 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.

I. J. Dahabreh and M. A. Hernán. Extending inferences from a randomized trial to a target population. *European journal of epidemiology*, 34:719–722, 2019.

I. J. Dahabreh, R. Hayward, and D. M. Kent. Using group data to treat individuals: understanding heterogeneous treatment effects in the age of precision medicine and patient-centred evidence. *International journal of epidemiology*, 45(6):2184–2193, 2016.

I. J. Dahabreh, S. J. A. Haneuse, J. M. Robins, S. E. Robertson, A. L. Buchanan, E. A. Stuart, and M. A. Hernán. Study designs for extending causal inferences from a randomized trial to a target population. *American journal of epidemiology*, 190(8):1632–1642, 2021.

I. J. Dahabreh, A. Matthews, J. A. Steingrimsson, D. O. Scharfstein, and E. A. Stuart. Using trial and observational data to assess effectiveness: trial emulation, transportability, benchmarking, and joint analysis. *Epidemiologic reviews*, 46(1):1–16, 2024.

P. De Bartolomeis, J. Abad, K. Donhauser, and F. Yang. Detecting critical treatment effect bias in small subgroups. In *Uncertainty in Artificial Intelligence*, pages 943–965. PMLR, 2024.

P. De Bartolomeis, J. Abad, G. Wang, K. Donhauser, R. M. Duch, F. Yang, and I. J. Dahabreh. Efficient randomized experiments using foundation models. *arXiv preprint arXiv:2502.04262*, 2025.

J. Fawkes, M. O'Riordan, A. Vlontzos, O. Corcoll, and C. M. Gilligan-Lee. The hardness of validating observational studies with experimental data. *arXiv preprint arXiv:2503.14795*, 2025.

D. J. Foster and V. Syrgkanis. Orthogonal statistical learning. *The Annals of Statistics*, 51(3):879–908, 2023.

T. Hatt, J. Berrevoets, A. Curth, S. Feuerriegel, and M. van der Schaar. Combining observational and randomized data for estimating heterogeneous treatment effects. *arXiv preprint arXiv:2202.12891*, 2022.

M. Huang, N. Egami, E. Hartman, and L. Miratrix. Leveraging population outcomes to improve the generalization of experimental results: Application to the jtpa study. *The Annals of Applied Statistics*, 17(3):2139–2164, 2023.

Z. Hussain, M.-C. Shih, M. Oberst, I. Demirel, and D. Sontag. Falsification of internal and external validity in observational studies via conditional moment restrictions. In *International Conference on Artificial Intelligence and Statistics*, pages 5869–5898. PMLR, 2023.

M. Jahanshahi, K. Gregg, G. Davis, A. Ndu, V. Miller, J. Vockley, C. Ollivier, T. Franolic, and S. Sakai. The use of external controls in fda regulatory decision making. *Therapeutic Innovation & Regulatory Science*, 55 (5):1019–1035, 2021.

N. Kallus, A. M. Puli, and U. Shalit. Removing hidden confounding by experimental grounding. *Advances in neural information processing systems*, 31, 2018.

R. Karlsson, G. Wang, P. De Bartolomeis, J. H. Krijthe, and I. J. Dahabreh. Robust integration of external control data in randomized trials. *arXiv preprint arXiv:2406.17971*, 2024.

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

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

S. W. Lagakos et al. The challenge of subgroup analyses-reporting without distorting. *New England Journal of Medicine*, 354(16):1667, 2006.

L. D. Liao, E. Højbjerre-Frandsen, A. E. Hubbard, and A. Schuler. Prognostic adjustment with efficient estimators to unbiasedly leverage historical data in randomized trials. *arXiv preprint arXiv:2305.19180*, 2023.

P. Morzywolek, J. Decruyenaere, and S. Vansteelandt. On weighted orthogonal learners for heterogeneous treatment effects. *arXiv preprint arXiv:2303.12687*, 2023.

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

F. Pedregosa, G. Varoquaux, A. Gramfort, V. Michel, B. Thirion, O. Grisel, M. Blondel, P. Prettenhofer, R. Weiss, V. Dubourg, J. Vanderplas, A. Passos, D. Cournapeau, M. Brucher, M. Perrot, and E. Duchesnay. Scikit-learn: Machine learning in Python. *Journal of Machine Learning Research*, 12:2825–2830, 2011.

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

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

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

A. Schuler, D. Walsh, D. Hall, J. Walsh, C. Fisher, C. P. for Alzheimer's Disease, A. D. N. Initiative, and A. D. C. Study. Increasing the efficiency of randomized trial estimates via linear adjustment for a prognostic score. *The International Journal of Biostatistics*, 18(2):329–356, 2022.

J. Schweisthal, D. Frauen, M. Van Der Schaar, and S. Feuerriegel. Meta-learners for partially-identified treatment effects across multiple environments. In *Forty-first International Conference on Machine Learning*, 2024.

H. Shimodaira. Improving predictive inference under covariate shift by weighting the log-likelihood function. *Journal of statistical planning and inference*, 90(2):227–244, 2000.

C. Shyr, B. Ren, P. Patil, and G. Parmigiani. Multi-study r-learner for estimating heterogeneous treatment effects across studies using statistical machine learning. *arXiv e-prints*, pages arXiv–2306, 2023.

L. Ung, G. Wang, S. Haneuse, M. A. Hernan, and I. J. Dahabreh. Combining an experimental study with external data: study designs and identification strategies. *arXiv preprint arXiv:2406.03302*, 2024.

J. van Rosmalen, D. Dejardin, Y. van Norden, B. Löwenberg, and E. Lesaffre. Including historical data in the analysis of clinical trials: Is it worth the effort? *Statistical methods in medical research*, 27(10):3167–3182, 2018.

L. Wu and S. Yang. Integrative r-learner of heterogeneous treatment effects combining experimental and observational studies. In *Conference on Causal Learning and Reasoning*, pages 904–926. PMLR, 2022.

A. Using external data to improve CATE model selection in the trial population

To estimate the CATE from observed data, we must consider the sample analog of $R(\tilde{\tau}; \eta)$, defined as

$$\widehat{R}(\tilde{\tau}; \hat{\eta}) = \frac{1}{n_1} \sum_{i:S_i=1} \left(\psi\left(O_i; \hat{\eta}\right) - \tilde{\tau}\left(X_i\right)\right)^2 . \tag{4}$$

We then obtain the CATE estimate by solving $\hat{\tau} = \arg\min_{\tilde{\tau} \in \mathcal{F}} \widehat{R}(\tilde{\tau}; \hat{\eta})$.

Although the robustness property of randomization-aware pseudo-outcomes discussed earlier might suggest that the choice of nuisance models η is inconsequential, we will show that this is not the case. Because the model selection criterion $\widehat{R}(\tilde{\tau}; \hat{\eta})$ is a sample average, in finite samples, this criterion can choose suboptimal CATE models and, importantly, this behavior is influenced by the choice of nuisance models η. To understand this, we first note that we can decompose $\widehat{R}(\tilde{\tau}; \hat{\eta})$ as

$$\frac{1}{n_1} \sum_{i:S_i=1} \left\{ \underbrace{\left(\tau(X_i) - \tilde{\tau}\left(X_i\right)\right)^2}_{\text{True error}} - 2 \underbrace{\left(\tau(X_i) - \tilde{\tau}\left(X_i\right)\right)\left(\widehat{\psi}_i - \tau(X_i)\right)}_{\text{Finite-sample error}} + \underbrace{\left(\widehat{\psi}_i - \tau(X_i)\right)^2}_{\text{Independent of } \tilde{\tau}} \right\} . \tag{5}$$

From the above decomposition, we see that $\widehat{R}(\tilde{\tau}; \hat{\eta})$ consists of three components: a term that in expectation equals the true risk $R^*(\tilde{\tau}) = E[(\tau(X) - \tilde{\tau}(X))^2 \mid S = 1]$; a second term that introduces finite-sample uncertainty; and a third term which is independent of the candidate model $\tilde{\tau}$. Therefore, we may still end up selecting a suboptimal $\tilde{\tau}$, mainly due to the finite-sample error influencing the model selection based on $\widehat{R}(\tilde{\tau}; \hat{\eta})$. To improve the performance of $\widehat{R}(\tilde{\tau}; \hat{\eta})$ as a model selection criterion, a natural strategy is to choose η to minimize the variance of the finite-sample errors. The next result provides insight into how this can be achieved (see proof in Appendix B.3).

Theorem A.1. *Under Condition 3.1 and 3.2 where the propensity score $e(X)$ is known and $\hat{\eta} = \{\hat{h}_1, \hat{h}_0\}$ is estimated on a different dataset than the one used to compute the pseudo-outcomes $\widehat{\psi}$, it holds for the finite-sample errors terms $\epsilon := (\tau(X) - \tilde{\tau}\left(X\right))\left(\widehat{\psi} - \tau(X)\right)$ that $\mathbb{E}[\epsilon \mid S = 1] = 0$ and $\mathrm{Var}(\epsilon \mid S = 1) \lesssim 2\{L_1(\hat{h}_1) + L_0(\hat{h}_0)\} + \tilde{\sigma}^2$ where*

$$L_a(\hat{h}_a) = \mathbb{E}\left[\left\{\frac{1 - e(X)}{e(X)}\right\}^{2a-1} \left(Y - \hat{h}_a\left(X\right)\right)^2 \mid A = a, S = 1\right], \quad a \in \{0, 1\}, \tag{6}$$

and $\tilde{\sigma}^2 = \mathrm{Var}\left(Y^1 - Y^0 \mid S = 1\right) - \mathrm{Var}\left(\tau\left(X\right) \mid S = 1\right)$.

The above theorem shows that a weighted mean squared error for each treatment of the components in $\hat{\eta} = \{\hat{h}_1, \hat{h}_0\}$ appears in an upper bound on the variance of the finite-sample error terms that appeared in the decomposition of the model selection criterion $\widehat{R}(\tilde{\tau}; \hat{\eta})$. This motivates our approach for selecting η: choose it to directly minimize both L_1 and L_0. Similar approaches to selecting η has appeared in prior work in single-source settings. Cao et al. (2009) addressed a related problem of mean estimation with missing data, while Saito and Yasui (2020) considered conditional average treatment effect estimation.

In the next result, we show how external data can be used to minimize the identified upper-bound in Theorem A.1. To achieve this, we invoke the assumptions of strong ignorability in the external data (Condition 3.3) and transportability (Condition 3.4). However, because these conditions are uncertain and may not hold in many settings, we recall that these conditions are not needed to ensure the robustness properties of the randomization-aware pseudo-outcomes.

Lemma A.2. *Under Condition 3.3 and 3.4 in addition to the conditions assumed in Theorem A.1, we can express $L_a(h_a)$ in (6) as*

$$L_a(\hat{h}_a) = \mathbb{E}\left[\frac{\Pr(S = 1 \mid X = x, A = a)}{\Pr(S = 1 \mid A = a)} \left\{\frac{1 - e(X)}{e(X)}\right\}^{2a-1} \left(Y - \hat{h}_a\left(X\right)\right)^2 \mid A = a\right] . \tag{7}$$

The above lemma, proven in Appendix B.4, shows that the functions L_1 and L_0, which appeared in the upper-bound of the variance $\mathrm{Var}(\epsilon \mid S = 1)$, can be rewritten in terms of all observed data from both the trial and external populations. The quantity in Lemma 7 is proportional to the objective in (3), up to a constant $1/\Pr(S = 1 \mid A = a)$.

B. Proofs

B.1. Decomposition of $\widehat{R}(\tilde{\tau}; \hat{\eta})$

We have that

$$\begin{aligned}
\widehat{R}(\tilde{\tau}; \hat{\eta}) &= \frac{1}{n_1} \sum_{i:S_i=1} \left(\widehat{\psi}_i - \tilde{\tau}(X_i) \right)^2 \\
&= \frac{1}{n_1} \sum_{i:S_i=1} \left(\widehat{\psi}_i - \tau(X_i) + \tau(X_i) - \tilde{\tau}(X_i) \right)^2 \\
&= \frac{1}{n_1} \sum_{i:S_i=1} \left\{ \left(\tau(X_i) - \tilde{\tau}\left(X_i\right)\right)^2 - 2\left(\tau(X_i) - \tilde{\tau}\left(X_i\right)\right)\left(\widehat{\psi}_i - \tau(X_i) \right) + \left(\widehat{\psi}_i - \tau(X_i) \right)^2 \right\}.
\end{aligned}$$

B.2. Proof of Theorem 4.1

Proof. We write $\psi_{\text{fixed}} = \psi(O; \eta_{\text{fixed}})$.

First, we need to prove conditional unbiasedness, $\mathbb{E}[\psi_{\text{fixed}} \mid X = x, S = 1] = \tau(x)$, as follows:

$$\begin{aligned}
\mathbb{E}[\psi_{\text{fixed}} \mid X = x] &= \mathbb{E}\left[\frac{A}{e(X)}(Y - h_1(X)) - \frac{1-A}{1-e(X)}(Y - h_0(X)) + h_1(X) - h_0(X) \mid X = x, S = 1 \right] \\
&= \mathbb{E}\left[\frac{A}{e(X)}(Y^1 - h_1(X)) - \frac{1-A}{1-e(X)}(Y^0 - h_0(X)) \mid X = x, S = 1 \right] + h_1(x) - h_0(x)
\end{aligned}$$

where the second equality follows from consistency in Condition 3.1. Next, we inspect the first term inside the above expectation, which can be rewritten as follows

$$\begin{aligned}
\mathbb{E}\left[\frac{A}{e(X)}(Y^1 - h_1(X)) \mid X = x, S = 1 \right] &= \mathbb{E}\left[\frac{A}{e(X)} \mid X = x, S = 1 \right] \left(\mathbb{E}\Big[Y^1 \mid X = x, S = 1 \Big] - h_1(x) \right) \\
&= \frac{e(x)}{e(x)} \left(\mathbb{E}\Big[Y^1 \mid X = x, S = 1 \Big] - h_1(X) \right) \\
&= \mathbb{E}\Big[Y^1 \mid X = x, S = 1 \Big] - h_1(X)
\end{aligned}$$

where the first equality follows from conditional exchangeability in the trial population, $Y^a \perp\!\!\!\perp A \mid X, S = 1$, in Condition 3.2 and the second equality follows from that $\mathbb{E}[A \mid X = x, S = 1] = e(x)$. Similarly, we can show that

$$\mathbb{E}\left[\frac{1-A}{1-e(X)}(Y^0 - h_0(X)) \mid X = x, S = 1 \right] = \mathbb{E}\Big[Y^0 \mid X = x, S = 1 \Big] - h_0(X)\,.$$

Putting all of the above together, we see that

$$\mathbb{E}[\psi_{\text{fixed}} \mid X = x] = \mathbb{E}\Big[Y^1 - Y^0 \mid X = x, S = 1 \Big] = \tau(x)$$

Next, we show that $\tau = \arg\min_{\tilde{\tau} \in \mathcal{F}} R(\tilde{\tau}; \eta_{\text{fixed}})$ when $\tau \in \mathcal{F}$. By adding and subtracting $\tau(X)$ inside $R(\tilde{\tau}; \eta_{\text{fixed}})$, we can decompose it as

$$\underbrace{\mathbb{E}\left[(\tau(X) - \tilde{\tau}(X))^2 \mid S = 1 \right]}_{(a)} - \underbrace{\mathbb{E}\left[\left(\tau(X) - \tilde{\tau}\left(X\right)\right)\left(\psi_{\text{fixed}} - \tau(X)\right) \mid S = 1 \right]}_{(b)} + \underbrace{\mathbb{E}\left[\left(\psi_{\text{fixed}} - \tau(X)\right)^2 \mid S = 1 \right]}_{(c)}$$

First, we see that $(a) = R^*(\tilde{\tau})$. Next, we have that $(b) = 0$ because

$$\begin{aligned}
&\mathbb{E}\left[\left(\tau(X) - \tilde{\tau}\left(X\right)\right)\left(\psi_{\text{fixed}} - \tau(X)\right) \mid S = 1\right] = \\
&\quad = \mathbb{E}\left[\mathbb{E}\left[\left(\tau(X) - \tilde{\tau}\left(X\right)\right)\left(\psi_{\text{fixed}} - \tau(X)\right) \mid X, S = 1\right] \mid S = 1\right] \\
&\quad = \mathbb{E}\left[\left(\tau(X) - \tilde{\tau}\left(X\right)\right)\mathbb{E}\left[\left(\psi_{\text{fixed}} - \tau(X)\right) \mid X, S = 1\right] \mid S = 1\right] \\
&\quad = 0
\end{aligned}$$

where the last equality follows from conditional unbiasedness such that $\mathbb{E}\left[\left(\psi_{\text{fixed}} - \tau(X)\right) \mid X, S = 1\right] = 0$. Finally, $(c) = C$ is a real-valued constant $C \geq 0$ independent of $\tilde{\tau}$. Thus, we can write that

$$R(\tilde{\tau}; \eta_{\text{fixed}}) = R^*(\tilde{\tau}) + C$$

which implies that

$$\arg\min_{\tilde{\tau} \in \mathcal{F}} R(\tilde{\tau}; \eta_{\text{fixed}}) = \arg\min_{\tilde{\tau} \in \mathcal{F}} \{R^*(\tilde{\tau}) + C\} = \arg\min_{\tilde{\tau} \in \mathcal{F}} R^*(\tilde{\tau}) = \tau\ .$$

□

B.3. Proof of Theorem A.1

Proof. To make it more explicit that the estimated nuisance models $\hat{\eta}$ are obtained independently of the observations used to compute the pseudo-outcomes, we denote it as $\eta_{\text{fixed}} = \{h_1, h_0\}$. Moreover, we write $\psi_{\text{fixed}} = \psi(O; \eta_{\text{fixed}})$.

Defining $\epsilon := \left(\tau(X) - \tilde{\tau}\left(X\right)\right)\left(\psi_{\text{fixed}} - \tau(X)\right)$, we then have that $\mathbb{E}[\epsilon \mid S = 1] = 0$ which we showed in the proof of Theorem 4.1. Next, we note that

$$\begin{aligned}
\text{Var}(\epsilon \mid S = 1) &= \mathbb{E}\left[\epsilon^2 \mid S = 1\right] \\
&= \mathbb{E}\left[\left(\tau(X) - \tilde{\tau}\left(X\right)\right)^2\left(\psi_{\text{fixed}} - \tau(X)\right)^2 \mid S = 1\right] \\
&= \mathbb{E}\left[\mathbb{E}\left[\left(\tau(X) - \tilde{\tau}\left(X\right)\right)^2\left(\psi_{\text{fixed}} - \tau(X)\right)^2 \mid X, S = 1\right] \mid S = 1\right] \\
&= \mathbb{E}\left[\left(\tau(X) - \tilde{\tau}\left(X\right)\right)^2\mathbb{E}\left[\left(\psi_{\text{fixed}} - \tau(X)\right)^2 \mid X, S = 1\right] \mid S = 1\right] \\
&= \mathbb{E}\left[\left(\tau(X) - \tilde{\tau}\left(X\right)\right)^2 \text{Var}\left(\psi_{\text{fixed}} \mid X, S = 1\right) \mid S = 1\right]
\end{aligned}$$

where the first equality follows from that $\mathbb{E}[\epsilon \mid S = 1] = 0$ and the last from the conditional unbiasedness of the pseudo-outcome, $\mathbb{E}[\psi_{\text{fixed}} \mid X = x, S = 1] = \tau(x)$, which we derived in the proof of Theorem 4.1.

Next, we show how to upper-bound $\text{Var}(\epsilon \mid S = 1)$ as follows:

$$\begin{aligned}
\text{Var}(\epsilon \mid S = 1) &= \mathbb{E}\left[\left(\tau(X) - \tilde{\tau}\left(X\right)\right)^2 \text{Var}\left(\widehat{\psi} \mid X, S = 1\right) \mid S = 1\right] \\
&\leq C_{\tilde{\tau}} \cdot \mathbb{E}\left[\text{Var}\left(\psi_{\text{fixed}} \mid X, S = 1\right) \mid S = 1\right]
\end{aligned}$$

where the inequality holds if define the constant $\tilde{C} = \max_{x \in \mathcal{X}}(\tau(x) - \tilde{\tau}(x))^2 \geq 0$. Next, we have from the law of total variance that

$$\begin{aligned}
\mathbb{E}\left[\text{Var}\left(\psi_{\text{fixed}} \mid X, S = 1\right) \mid S = 1\right] &= \text{Var}(\psi_{\text{fixed}} \mid S = 1) - \text{Var}\left(\mathbb{E}\left[\psi_{\text{fixed}} \mid X, S = 1\right] \mid S = 1\right) \\
&= \text{Var}(\psi_{\text{fixed}} \mid S = 1) - \text{Var}\left(\tau\left(X\right) \mid S = 1\right)\ .
\end{aligned}$$

where the second equality follows from the conditional unbiasedness of the pseudo-outcomes. Thus, so far, we have $\text{Var}(\epsilon \mid S = 1) \lesssim \text{Var}(\psi_{\text{fixed}} \mid S = 1) - \text{Var}\left(\tau\left(X\right) \mid S = 1\right)$.

Next, we inspect the variance $\text{Var}(\psi_{\text{fixed}} \mid S = 1)$, which we rewrite using the law of total variance:

$$\begin{aligned}
\text{Var}(\psi_{\text{fixed}} \mid S = 1) &= \mathbb{E}\left[\text{Var}\left(\psi_{\text{fixed}} \mid Y^1, Y^0, X, S = 1\right) \mid S = 1\right] + \text{Var}\left(\mathbb{E}\left[\psi_{\text{fixed}} \mid Y^1, Y^0, X, S = 1\right] \mid S = 1\right) \\
&= \mathbb{E}\left[\text{Var}\left(\psi_{\text{fixed}} \mid Y^1, Y^0, X, S = 1\right) \mid S = 1\right] + \text{Var}\left(Y^1 - Y^0 \mid S = 1\right)
\end{aligned}$$

where the second inequality follows from that

$$\mathbb{E}\left[\psi_{\text{fixed}} \mid Y^1, Y^0, X, S=1\right] = \mathbb{E}\left[Y^1 - Y^0 \mid Y^1, Y^0, X, S=1\right] = Y^1 - Y^0$$

where the first equality stems from the conditional unbiasedness of the pseudo-outcomes.

We next write $\psi_{\text{fixed}} = \psi_{1,\text{fixed}} - \psi_{0,\text{fixed}}$ where $\psi_{a,\text{fixed}} = \frac{\mathbf{1}(A=a)}{\mathbf{1}(A=1)e(X)+\mathbf{1}(A=0)(1-e(X))}(Y - h_a(X)) + h_a(X)$. This will help us simplify the expression for the above inner conditional variance as follows,

$$\begin{aligned}
&\text{Var}\left(\psi_{\text{fixed}} \mid Y^1, Y^0, X, S=1\right) = \text{Var}\left(\psi_{1,\text{fixed}} - \psi_{0,\text{fixed}} \mid Y^1, Y^0, X, S=1\right) \\
&= \mathbb{E}\left[\left\{\psi_{1,\text{fixed}} - \psi_{0,\text{fixed}} - \underbrace{\mathbb{E}\left[\psi_{1,\text{fixed}} - \psi_{0,\text{fixed}} \mid Y^1, Y^0, X, S=1\right]}_{=Y^1-Y^0}\right\}^2 \mid Y^1, Y^0, X, S=1\right] \\
&= \mathbb{E}\left[\left\{\left(\psi_{1,\text{fixed}} - Y^1\right) - \left(\psi_{0,\text{fixed}} - Y^0\right)\right\}^2 \mid Y^1, Y^0, X, S=1\right] \\
&\leq 2\mathbb{E}\left[\left(\psi_{1,\text{fixed}} - Y^1\right)^2 + \left(\psi_{0,\text{fixed}} - Y^0\right)^2 \mid Y^1, Y^0, X, S=1\right]
\end{aligned}$$

where the third inequality follows again from the conditional unbiasedness of the pseudo-outcomes and the final inequality from that $(a-b)^2 \leq 2(a^2+b^2)$ for any real numbers a and b. At last, we note that

$$\begin{aligned}
&\mathbb{E}\left[\left(\psi_{1,\text{fixed}} - Y^1\right)^2 \mid Y^1, Y^0, X, S=1\right] = \\
&= \mathbb{E}\left[\left(\frac{A}{e(X)}(Y - h_1(X)) + h_1(X) - Y^1\right)^2 \mid Y^1, Y^0, X, S=1\right] \\
&= \mathbb{E}\left[\left(\frac{A}{e(X)}(Y^1 - h_1(X)) + h_1(X) - Y^1\right)^2 \mid Y^1, Y^0, X, S=1\right] \\
&= \mathbb{E}\left[\left(\frac{A}{e(X)} - 1\right)^2 \mid X, S=1\right](Y^1 - h_1(X))^2 \\
&= \frac{1-e(X)}{e(X)}(Y^1 - h_1(X))^2
\end{aligned}$$

where the second equality follows from consistency (Condition 3.1) and the third equality from conditional exchangeability in the trial population (Condition 3.2). Similarly, we have that

$$\mathbb{E}\left[\left(\psi_{0,\text{fixed}} - Y^0\right)^2 \mid Y^1, Y^0, X, S=1\right] = \frac{e(X)}{1-e(X)}(Y^0 - h_0(X))^2 .$$

At last, plugging the above expressions back into our original expression for $\text{Var}(\psi_{\text{fixed}} \mid S=1)$, we obtain the inequality

$$\begin{aligned}
\text{Var}(\epsilon \mid S=1) \lesssim\ & 2\mathbb{E}\left[\frac{1-e(X)}{e(X)}(Y^1 - h_1(X))^2 \mid S=1\right] + 2\mathbb{E}\left[\frac{e(X)}{1-e(X)}(Y^0 - h_0(X))^2 \mid S=1\right] \\
&+ \text{Var}\left(Y^1 - Y^0 \mid X, S=1\right) - \text{Var}\left(\tau(X) \mid S=1\right) \\
=\ & 2\mathbb{E}\left[\frac{1-e(X)}{e(X)}(Y - h_1(X))^2 \mid X, A=1, S=1\right] + 2\mathbb{E}\left[\frac{e(X)}{1-e(X)}(Y - h_0(X))^2 \mid X, A=0, S=1\right] \\
&+ \text{Var}\left(Y^1 - Y^0 \mid S=1\right) - \text{Var}\left(\tau(X) \mid S=1\right)
\end{aligned}$$

where the equality follows from consistency and conditional exchangeability in the trial population again (Condition 3.1 and 3.2). □

B.4. Proof of Lemma A.2

Proof. We begin by writing

$$\begin{aligned}
L_a(h) &= \mathbb{E}\left[\left\{\frac{1-e(X)}{e(X)}\right\}^{2a-1}(Y-h(X))^2 \mid A=a, S=1\right] \\
&= \mathbb{E}\left[\frac{S}{\Pr(S=1 \mid A=a)}\left\{\frac{1-e(X)}{e(X)}\right\}^{2a-1}(Y-h(X))^2 \mid A=a\right] \\
&= \mathbb{E}\left[\mathbb{E}\left[\frac{S}{\Pr(S=1 \mid A=a)}\left\{\frac{1-e(X)}{e(X)}\right\}^{2a-1}(Y-h(X))^2 \mid X, A=a\right] \mid A=a\right] \\
&= \mathbb{E}\left[\frac{1}{\Pr(S=1 \mid A=a)}\left\{\frac{1-e(X)}{e(X)}\right\}^{2a-1}\mathbb{E}\left[S(Y-h(X))^2 \mid X, A=a\right] \mid A=a\right].
\end{aligned}$$

We inspect the inner conditional expectation, $\mathbb{E}\left[S(Y-h(X))^2 \mid X, A=a\right]$, and note that

$$\begin{aligned}
\mathbb{E}\left[S(Y-h(X))^2 \mid X, A=a\right] &= \mathbb{E}[S \mid X, A=a]\mathbb{E}[(Y-h(X))^2 \mid X, A=a] \\
&= \Pr(S=1 \mid X, A=a)\mathbb{E}[(Y-h(X))^2 \mid X, A=a]
\end{aligned}$$

where the first equality follows from that $Y \perp\!\!\!\perp S \mid (X, A)$ holds under Conditions 3.1-3.4. To show this, we note that the conditional independencies $Y^a \perp\!\!\!\perp A \mid (X, S=1)$ (Condition 3.2) and $Y^a \perp\!\!\!\perp A \mid (X, S=0)$ (Condition 3.3) jointly imply that $Y^a \perp\!\!\!\perp A \mid (X, S)$. Combining this conditional independence statement with $Y^a \perp\!\!\!\perp S \mid X$ from Condition 3.4, they together imply $Y^a \perp\!\!\!\perp (A, S) \mid X$. Thus, from the weak union of conditional independence, we have that $Y^a \perp\!\!\!\perp S \mid (X, A) \Rightarrow Y \perp\!\!\!\perp S \mid (X, A)$ where the final implication follows from consistency (Condition 3.1).

Combining all of the above, we finally obtain the following expression,

$$L_a(h) = \mathbb{E}\left[\frac{\Pr(S=1 \mid X, A=a)}{\Pr(S=1 \mid A=a)}\left\{\frac{1-e(X)}{e(X)}\right\}^{2a-1}(Y-h(X))^2 \mid A=a\right]$$

□

C. Experimental details

C.1. Data-generating process in simulation study

We simulate data as follows: We set $S_i = 1$ for $i = 1, \ldots, n_1$, and for $S_i = 0$ for $i = n_1 + 1, \ldots, n_1 + n_0$, and sampled a Normal d-dimensional covariate according to $X_i \sim N(\mu_{S_i}, \frac{1}{\sqrt{d}}\Sigma)$ with the mean $\mu_1 = \mathbf{0}$ or $\mu_0 = 0.2 \cdot \mathbf{1}$ and the covariance matrix Σ of shape $d \times d$ had its diagonal elements set to 1 and its off-diagonal elements set to 0.1. Thereafter, we sampled the treatment $T_i \sim \text{Bern}(e(X_i, S_i))$ according to the Bernoulli probability

$$e(X_i, S_i) = \begin{cases} 0.5, & \text{if } S_i = 1 \\ \frac{1}{1+\exp\{-(\alpha_0 + \alpha^\top X_i)\}}, & \text{otherwise} \end{cases}$$

Finally, we computed outcomes $Y_i = b(X_i) + A_i \cdot \tau(X_i) + \varepsilon_i$ where the noise variables was sampled according to $\varepsilon_i \sim N(0, \sigma^2 = 1/4)$.

For the experiment, we modeled a highly non-linear baseline risk together with a linear CATE, which as done by defining:

$$b(X_i) = \sum_{j=1}^{d} \frac{3}{d} \cos\left(\frac{3}{2} X_{ij}\right) + \sum_{j=1}^{d} \sum_{j'=1}^{d} \frac{1}{d} X_{ij} X_{ij'},$$
$$\tau(X_i) = \sum_{j=1}^{d} \frac{1}{d} X_{ij}.$$

In the scenario where Conditions 3.3 and 3.4 held, we set the covariate dimension to $d = 5$. To simulate violations of these assumptions, we increased the covariate dimension to $d = 7$ but masked the last two dimensions, so that only 5 covariates remained observed.

C.2. Implementation details for CATE learners

For the estimators used inside the CATE learners we used implementations from the *scikit-learn* Python package (Pedregosa et al., 2011) . For the DR-learner, T-learner, pooled T-learner, the method from Asiaee et al. (2023), and QR-learner, we used histogram-based gradient boosting regression tree using default hyperparameter. As the final CATE regressor in the two-stage CATE learners (all of the above except the T-learner variants), we used a linear regression model. For estimating $\pi_a(X) = \Pr(S = 1 \mid X, A = a)$, we used a cross-validated logistic regression with ridge penalty. For the additive correction model in Kallus et al. (2018), we fitted the DR-learner on the external dataset and then used a linear regression to estimate the bias model. We applied crossfitting to all two-stage CATE learners using two folds consistently. For cross-fold validation in the combined QR- and DR-learner, we used three folds.

To predict with the average treatment effect (ATE), we used the difference-in-means estimate

$$\hat{\tau}_{DM} = \frac{\sum_{i=1}^{n_1} A_i Y_i}{\sum_{i=1}^{n_1} A_i} - \frac{\sum_{i=1}^{n_1} (1 - A_i) Y_i}{\sum_{i=1}^{n_1} 1 - A_i}$$

as a constant CATE prediction $\hat{\tau}(x) = \hat{\tau}_{DM}$ for all x.