Robust Estimation of Heterogeneous Treatment Effects in Randomized Trials Leveraging External Data

Rickard Karlsson
TU Delft

Piersilvio De Bartolomeis
ETH Zürich

Issa J. Dahabreh
Harvard University

Jesse H. Krijthe
TU Delft

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 proposed method is robust: it can reduce the mean squared error relative to a trial-only CATE learner, and is guaranteed to recover the true CATE even when the external data are not aligned with the trial. Moreover, we introduce a procedure that combines the QR-learner with a trial-only CATE learner and show that it asymptotically matches or exceeds both component learners in terms of mean squared error. We examine the performance of our approach in simulation studies and apply the methods to a real-world dataset, demonstrating improvements in both CATE estimation and statistical power for detecting heterogeneous effects.

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. This strength has made trials the preferred approach for evaluating interventions across many scientific domains. However, their high cost often limits sample size, which in turn limits the precision of statistical inferences that can be drawn from trial data. The problem is especially acute when the aim is not only to estimate an average treatment effect but also to characterize treatment effect heterogeneity (Lagakos et al., 2006), a key step toward personalized decision-making for the 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 trials with external data, 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 differences between the trial population and the population underlying the external data (Ung et al., 2024). These differences raise a fundamental concern: whether causal quantities such as the CATE remain stable across the two populations – a property known as transportability (Bareinboim and Pearl, 2016; Dahabreh and Hernán, 2019). In this paper, we investigate the analogous problem of using external data, such as from another trial or an observational study, to augment the estimation of CATEs in the trial population, in settings where the populations underlying the trial and external data may be misaligned. Here, we focus on the setting where transportability does not necessarily hold and the external data may be subject to unmeasured confounding. Our objective is to leverage external data for CATE estimation while ensuring that using the external data does not harm the estimation compared to if we had used trial data alone in cases of misalignment between the underlying populations.

Proceedings of the 29th International Conference on Artificial Intelligence and Statistics (AISTATS) 2026, Tangier, Morocco. PMLR: Volume 300. Copyright 2026 by the author(s).

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 when the external data are sufficiently aligned (Sections 4.1-4.3). As another safeguard against potential harm from misaligned external data, we propose a procedure that combines the QR-learner with a trial-only CATE learner and prove that the combined learner asymptotically achieves a mean squared error that is no worse and potentially better than its component learners (Section 4.4). Using simulations and real-world data from the Student/Teacher Achievement Ratio (STAR) project (Word et al., 1990; Krueger, 1999) we find that our method is robust when integrating external data that are not aligned with the trial data, and that it improves both CATE estimation mean squared error and the statistical power to detect treatment effect heterogeneity (Section 5 and 6).

2 RELATED WORKS

Our work builds upon a rich and growing literature on CATE learners. Examples of approaches include adaptations of decision trees (Athey and Imbens, 2016), random forests (Wager and Athey, 2018), and neural networks (Shalit et al., 2017). Our proposed learner is most closely related to model-agnostic "meta-learners" which allow for estimating the CATE and 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 are 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 (Hatt et al., 2022; Wu and Yang, 2022; Shyr et al., 2023; Wu and Yang, 2023). For example, Schweisthal et al. (2024) propose a learner that constructs bounds on the transportable CATE under unmeasured confounding across different populations, while Kallus et al. (2018) rely on the transportability assumption by estimating the CATE from a large, potentially confounded observational dataset and then apply a linear bias correction using data from a small randomized trial. Although some of these approaches might be used to estimate the trial-specific CATE under non-transportability, to our knowledge only Asiaee et al. (2023, 2025) explicitly address this setting; we discuss these methods in more detail later. Yang et al. (2023) studied the related problem of using multi-source data to perform valid statistical inference when estimating treatment effect heterogeneity under violations of transportability. However, their approach is restricted to a parametric linear working model for the CATE, whereas we focus specifically on optimizing the predictive performance (i.e., minimizing mean squared error) for a model-agnostic CATE learner which also allows for more flexible and nonparametric working models.

Finally, our work draws on recent developments in the trial augmentation literature for average treatment effect estimation in trials using data from an external population. These developments have emphasized robustness to integrating external data misaligned with the trial data (Schuler et al., 2022; Huang et al., 2023; Liao et al., 2023; De Bartolomeis et al., 2025; Karlsson et al., 2026). In particular, we will adapt ideas from the randomization-aware estimator framework proposed by Karlsson et al. (2026) to construct CATE learners that are 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). 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 some constants, i.e., $n_s/n \to q_s \in (0, 1)$.

3.1 Identification of Causal Effects

To define the causal quantity of interest, we use potential 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 CATE in the population underlying the trial,

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

Under standard conditions, the CATE $\tau(x)$ is identifiable from data in the trial.

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

Condition 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 1 and 2 are typically supported by a well-designed randomized trial and together suffice to identify the CATE as $\tau(x) = g_1(x) - g_0(x)$, 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 (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 4 (Transportability)**.** *For each $a \in \{0, 1\}$, $Y^a \perp\!\!\!\perp S \mid X$.*

The above two conditions can be controversial, especially when the external data originate from an observational study, because these conditions are uncertain and typically require substantial domain expertise to justify. Notably, Conditions 1 to 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 and 4 because Condition 1 and 2 are supported by the trial's experimental design. Nonetheless, performing such falsification tests remains an inherently difficult task (Fawkes et al., 2025).

4 AUGMENTING TRIALS WITH EXTERNAL DATA

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 population 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 the class of CATE learners obtained by minimizing a pseudo-risk (Foster and Syrgkanis, 2023),

$$\arg\min_{\tilde{\tau} \in \mathcal{F}} R(\tilde{\tau}; \eta) \tag{2}$$

where $R(\tilde{\tau}; \eta) = \mathbb{E}[(\psi(O; \eta) - \tilde{\tau}(X))^2 \mid S = 1]$. Here, we introduce an auxiliary random variable, sometimes referred to as a pseudo-outcome:

$$\begin{aligned}\psi(O_i; \eta) = &\frac{A_i - e(X_i)}{e(X_i)(1 - e(X_i))}\big(Y_i - h_{A_i}(X_i)\big) \\ &+ h_1(X_i) - h_0(X_i)\end{aligned} \tag{3}$$

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 real-valued functions defined on the covariate space $\mathcal{X}$. When clear from context, we omit the arguments and write $\psi_i = \psi(O_i; \eta)$ and $\widehat{\psi}_i = \psi(O_i; \hat{\eta})$ to denote the pseudo-outcomes computed using either the nuisance models η or replacing the nuisance models with their estimates $\hat{\eta}$.

Depending on our choice of η, we obtain different CATE learners when solving (2). 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 trial's randomized design.

Theorem 1. *Under Conditions 1 and 2 where the propensity score $e(X)$ is known, for any fixed specification of the nuisance models η_{fixed}, the minimization problem in (2) always yields the true CATE as its unique solution provided that $\tau \in \mathcal{F}$; that is, $\tau = \arg\min_{\tilde{\tau} \in \mathcal{F}} R(\tilde{\tau}; \eta_{fixed})$.*

Although this result has appeared in the literature before (see, for example, Morzywolek et al. (2023)

and references therein), we provide a derivation in Appendix A.1 for completeness. We denote the nuisance models η_{fixed} with a fixed specification to emphasize that they must be chosen independently of the dataset used to compute the pseudo-outcome; this requirement can be satisfied using cross-fitting, which we describe in the following subsections.

Recognizing the central role of using the known propensity score in the above theorem, we refer to the pseudo-outcomes in (3) constructed using the known propensity score as *randomization-aware* pseudo-outcomes (Karlsson et al., 2026). This framing allows us to consider a broader class of pseudo-outcomes for robust CATE estimation, distinguished by varying specifications of the nuisance models η. While Theorem 1 guarantees that $R(\tilde{\tau};\eta)$ is a proper model selection criterion for the CATE regardless of the choice of η, as we will discuss next, the choice of η ultimately still plays a crucial role in estimating the CATE based on the observed data. This happens as we transition to the sample analog version of the minimization problem in (2).

Remark 1. While our main focus is on the pseudo-risk and pseudo-outcomes in (2) and (3), another relevant pseudo-risk that may share similar robustness properties when having access to the true propensity score $e(X)$ is the one used by the R-learner (Nie and Wager, 2021). In Appendix B, we outline potential connections between our results and the R-learner pseudo-risk, particularly in the special case where the propensity score $e(X)$ is constant.

4.2 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 $\widehat{R}(\tilde{\tau};\hat{\eta}) = \frac{1}{n_1}\sum_{i:S_i=1}\left(\psi\left(O_i;\hat{\eta}\right)-\tilde{\tau}\left(X_i\right)\right)^2$. 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 the sample analog pseudo-risk as

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

From the above decomposition, which we derive in Appendix A.2, we see that $\widehat{R}(\tilde{\tau};\hat{\eta})$ consists of three parts: a first term that in expectation equals the population risk $R^*(\tilde{\tau}) = E[(\tau(X)-\tilde{\tau}(X))^2 \mid S=1]$; a second term that introduces model selection uncertainty; and a third term which is independent of the candidate model $\tilde{\tau}$. Therefore, we may still end up selecting a suboptimal $\tilde{\tau}$ due to the second term influencing $\widehat{R}(\tilde{\tau};\hat{\eta})$.

To see why the second term is problematic for model selection, consider comparing two candidate models $\tilde{\tau}_1$ and $\tilde{\tau}_2$. If we observe $\widehat{\Delta}_{12} = \widehat{R}(\tilde{\tau}_1;\hat{\eta}) - \widehat{R}(\tilde{\tau}_2;\hat{\eta}) > 0$, we would conclude that $\tilde{\tau}_2$ is better than $\tilde{\tau}_1$; we would make the opposite decision when $\widehat{\Delta}_{12} < 0$, or remain inconclusive when $\widehat{\Delta}_{12} = 0$. However, note that only the third term from the decomposition cancels out in the risk difference $\widehat{\Delta}_{12}$. This means that in addition to the first term, which captures the true error relative to the true CATE, the second term can also influence our decision about which candidate model performs better.

To improve $\widehat{R}(\tilde{\tau};\hat{\eta})$ as a model selection criterion, a natural strategy is to choose η to minimize the variance of the problematic second term. The next result provides insight into how this can be achieved (see proof in Appendix A.3).

Lemma 1. *Under Conditions 1 and 2, assume the propensity score $e(X)$ is known and the nuisance functions $\hat{\eta} = \{\hat{h}_1, \hat{h}_0\}$ are estimated on a dataset independent of that used to compute the pseudo-outcomes $\widehat{\psi}$. Define $\epsilon := (\tau(X)-\tilde{\tau}(X))(\widehat{\psi}-\tau(X))$, then we have that $\mathbb{E}[\epsilon \mid S=1] = 0$ and $\mathrm{Var}(\epsilon \mid S=1) \leq \tilde{C}\Big[2\{L_1(\hat{h}_1)+L_0(\hat{h}_0)\}+\tilde{\sigma}^2\Big]$ where*

$$\begin{aligned}L_a(\hat{h}_a) &= \mathbb{E}\left[w_a(X)\left(Y-\hat{h}_a\left(X\right)\right)^2 \mid A=a, S=1\right]\\ \tilde{\sigma}^2 &= \mathrm{Var}(Y^1-Y^0 \mid S=1) - \mathrm{Var}(\tau(X) \mid S=1)\\ \tilde{C} &= \max_{x\in\mathcal{X}}(\tau(x)-\tilde{\tau}(x))^2\end{aligned}$$

with $w_a(X) = \left(\frac{1-e(X)}{e(X)}\right)^{2a-1}$ for $a \in \{0,1\}$.

The above result shows that a weighted mean squared error of the nuisance models $\hat{\eta} = \{\hat{h}_1, \hat{h}_0\}$ appears in an upper bound on the variance of the terms responsible for model selection uncertainty when using $\widehat{R}(\tilde{\tau};\hat{\eta})$.

This motivates our approach for selecting η: choose it to directly minimize both L_1 and L_0. Similar approaches to selecting η have 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.

The tightness of the bound in the above lemma depends on the candidate model $\tilde{\tau}$ for the CATE. Specifically, it is proportional to the non-negative constant $\tilde{C} = \max_{x\in\mathcal{X}}(\tau(x) - \tilde{\tau}(x))^2$. Therefore, when the class of candidate models $\mathcal{F}$ contains good approximations of the true CATE τ and τ itself is bounded, $\tilde{C}$ is expected to be small, resulting in a relatively tight upper bound.

In the next result, we show how external data can be used to minimize the identified upper-bound in Lemma 1. To achieve this, we invoke the assumptions of strong ignorability in the external data (Condition 3) and transportability (Condition 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. We later discuss a strategy for further improving robustness against these violations.

Theorem 2. *Under Conditions 3 and 4, in addition to those of Lemma 1, we can express*

$$L_a(\hat{h}_a) = \mathbb{E}\Big[\tilde{w}_a(X)\,(Y - \hat{h}_a(X))^2 \mid A = a\Big] \quad (4)$$

where $\tilde{w}_a(X) = \frac{\Pr(S=1|X,A=a)}{\Pr(S=1|A=a)}\left(\frac{1-e(X)}{e(X)}\right)^{2a-1}$.

The above theorem, proven in Appendix A.4, shows that the function $L_a(\hat{h}_a)$, 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. Unlike the expression for $L_a(\hat{h}_a)$ in Lemma 1, we no longer condition on S. With this result, we can proceed with proposing a novel CATE learner that chooses the nuisance models η to minimize this upper-bound with the help of external data.

Remark 2. The results in Theorem 2 are reminiscent of results from the transfer learning literature, where one express the expected mean squared error of a prediction function for data within some domain by reweighting data from another domain (Shimodaira, 2000; Weiss et al., 2016). To obtain this type of result in transfer learning, one typically assumes that the conditional distribution of labels remains constant across the two domains; this is analogous to how the transportability condition put certain requirements on the conditional distribution of the potential outcomes in our setting. However, unlike this literature, our goal is to be explicitly robust to violations of this condition, for which we propose a solution to this problem later in Section 4.4.

4.3 The QR-learner Algorithm

In this section, we introduce a novel method 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 and nuisance models. It follows a two-stage procedure: In the first stage, it solves an optimization problem to select η using both trial and external data that minimizes the identified upper bound related to the model selection uncertainty in finite samples. We call this step quasi-optimized because it targets an upper bound rather than the variance directly. In the second stage, the method regresses randomization-aware pseudo-outcomes on the covariates using only trial data to estimate the CATE. The first stage reduces finite-sample uncertainty for model selection, while the second stage leverages the robustness of the randomization-aware pseudo-outcomes to target the true CATE in the trial population.

To prevent overfitting and ensure that $\hat{\eta}$ is estimated independently of the data used for pseudo-outcomes, we employ a cross-fitting procedure that partitions the data into $\mathcal{D}^1 \cup \mathcal{D}^2$, stratified by treatment A and study population S. In the first stage, we use both the trial and external data in $\mathcal{D}^1$ to estimate $\hat{\eta}^* = \{\hat{h}_1^*, \hat{h}_0^*\}$ where each component $\hat{h}_a^*$ is obtained by solving the optimization problem

$$\min_{h_a\in\mathcal{H}} \sum_{i\in\mathcal{D}_a^1} \widehat{\pi}_a(X_i)\Big(\frac{1-e(X_i)}{e(X_i)}\Big)^{2a-1}\big(Y_i - h_a(X_i)\big)^2 \quad (5)$$

where the sum is taken over $\mathcal{D}_a^1 = \{i \in \mathcal{D}^1 : A_i = a\}$, $\mathcal{H}$ is the model class under consideration for the nuisance models $\eta = \{h_1, h_0\}$, and $\widehat{\pi}_a$ is an estimator from the model class $\mathcal{G}$ for the probability of trial participation $\Pr(S = 1 \mid X = x, A = a)$ which is also estimated using $\mathcal{D}_a^1$. Then, in the second stage, the pseudo-outcomes $\psi(O_i; \hat{\eta}^*)$ are computed on $\mathcal{D}^2$ using the estimated nuisance models $\hat{\eta}^*$. We estimate the CATE using $\mathcal{D}^2$ by solving $\hat{\tau} = \arg\min_{\tilde{\tau}\in\mathcal{F}} \sum_{i\in\mathcal{D}^2:S_i=1}(\psi(O_i;\hat{\eta}^*) - \tilde{\tau}(X_i))^2$. 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. Pseudo-code

for the QR-learner is also provided in Appendix C.

We now discuss several noteworthy remarks about the QR-learner. First, we recommend using a linear logistic regression with a cross-validated ridge penalty for estimating $\widehat{\pi}_a$. The reason we use a linear model is that, although the external sample size is large, the trial sample size may be much smaller, so a more flexible model could still overfit or have poor calibration. It is also known that maximum likelihood estimation of logistic regression can be biased in small-sample settings or when the number of observations from one of the events is rare. Adding a shrinkage penalty, as we do here, can reduce this type of bias; see, for example, Firth (1993) and Leitgöb (2020) for more detailed discussion.

Second, while our primary focus is to predict the CATE well with respect to the risk $R^*(\hat{\tau}) = \mathbb{E}[(\tau(X) - \hat{\tau}(X))^2 \mid S = 1]$, we note that in some cases, it may also be possible to use the estimated function $\hat{\tau}(X)$ obtained from the QR-learner for inference about the CATE. Although a full theoretical treatment is beyond the scope of this work, we provide supporting arguments in Appendix D to explain why this may be justified, and later show in a simulation study that this is feasible.

Third, our motivation for proposing the QR-learner stems from Theorem 2, which shows that minimizing the objective in (5) can improve CATE model selection. Here we provide another argument for why the QR-learner may outperform a trial-only learner, specifically the DR-learner. Both learners differ only in how their outcome nuisance models are estimated, and it is known that the DR-learner's convergence rate depends on errors in both estimating the CATE function in the second stage and the outcome nuisance models in the first stage (Kennedy, 2023). The QR-learner can improve upon the DR-learner by estimating these outcome nuisance models more accurately using both trial and external data when populations are well-aligned. With sufficiently large external data, outcome nuisance estimation errors become negligible, and we could expect the QR-learner to outperform the DR-learner, which is also what we observe in our simulations.

Finally, the CFACE learner proposed by Asiaee et al. (2023) shares close similarities with the QR-learner. Their method differs from ours in the first stage by estimating nuisance models $\eta = \{m^*, m^*\}$, where $m^*(x) = e(x) \cdot \mu_0(x) + (1 - e(x)) \cdot \mu_1(x)$ with $\mu_a(x) = \mathbb{E}[Y \mid X = x, A = a, S = 0]$. This formulation comes from solving $\arg\min_m \mathrm{Var}\left(\psi(O; \eta = \{m, m\}) \mid X = x, S = 1\right)$, with nuisance models estimated entirely from external data. Their approach works well when external data are abundant and aligned with the trial population, but can fail if the populations differ substantially or external data are limited. This pitfall was noted by the same authors in Asiaee et al. (2025), where they propose another learner called R-OSCAR. This learner differs more substantially from ours and instead resembles the bias correction approach of Kallus et al. (2018), but tailored specifically for estimating the CATE in the underlying population of a randomized trial. We further discuss the procedures of both CFACE and R-OSCAR in Appendix E.

4.4 Combining CATE Learners

The success of the optimization in (5) relies on two conditions that are not necessary for the identification of the CATE in the trial population. First, it requires that the external data are aligned with the trial data: that is, Conditions 3 and 4 hold. Second, for each $a \in \{0, 1\}$, the estimator $\widehat{\pi}_a(x)$ needs to be a correctly specified model of the probability $\Pr(S = 1 \mid X = x, A = a)$. As a result, while the QR-learner always targets the correct CATE in the trial population, it may perform worse in finite samples, in terms of mean squared error, than a CATE learner based on the trial data alone. To address this issue, we therefore propose a variant which combines the potential benefits from using external data via the QR-learner with the additional robustness from a trial-only learner. Specifically, we use the DR-learner (Kennedy, 2023) obtained by regressing randomization-aware pseudo-outcomes $\psi(O; \hat{\eta} = \{\hat{g}_1, \hat{g}_0\})$ on the covariates X where $\hat{g}_a$ estimates $\mathbb{E}[Y \mid X, A = a, S = 1]$ using only trial data.

Our proposed combined learner is defined as

$$\hat{\tau}(x; \lambda) := \lambda \cdot \hat{\tau}_{\mathrm{QR}}(x) + (1 - \lambda) \cdot \hat{\tau}_{\mathrm{DR}}(x),\ \lambda \in [0, 1]\ ,$$

where $\hat{\tau}_{\mathrm{QR}}$ is the estimator from the QR-learner using both trial and external data, and $\hat{\tau}_{\mathrm{DR}}$ is the estimator from the DR-learner using trial data alone; this combined learner can be viewed as an instance of stacked regression (Breiman, 1996) for CATE estimation. We recommend using the DR-learner as the trial-only learner here because it is also robust to misspecification of the outcome models, even though the theoretical guarantees in this section also hold if we replace it with another CATE learner fitted using only the trial data.

Recall the population risk introduced in Section 4.1, $R^*(\tilde{\tau}) = \mathbb{E}\big[(\tau(X) - \tilde{\tau}(X))^2 \mid S = 1\big]$. For the combined learner we write $R^*(\lambda) := R^*\big(\hat{\tau}(\,\cdot\,; \lambda)\big)$ with $R^*_{\mathrm{DR}} = R^*(0)$ and $R^*_{\mathrm{QR}} = R^*(1)$. Our goal in this section is to choose λ such that $R^*(\lambda)$ is no larger than $\min\{R^*_{\mathrm{DR}}, R^*_{\mathrm{QR}}\}$. We first show that for the oracle weights the mean squared error of the combined

estimator is no worse than the individual components (see proof in Appendix A.5).

Lemma 2. *Let $\lambda^\star = \arg\min_{\lambda\in\Lambda} R^*(\lambda)$ be the oracle weight then $R^*(\lambda^\star) \leq \min\{R^*_{QR}, R^*_{DR}\}$. Furthermore, if $\mathbb{E}[(\hat{\tau}_{QR}(X) - \hat{\tau}_{DR}(X))^2 \mid S = 1] > 0$, then the oracle weight $\lambda^\star$ is unique.*

In practice, the population risk cannot be computed directly, so we minimize the sample analog pseudo-risk $\widehat{R}(\lambda) = \frac{1}{n_1}\sum_{i:S_i=1}(\widehat{\psi}_i - \hat{\tau}(X_i;\lambda))^2$. We show that an empirical estimate of the oracle weight $\lambda^\star$ can be obtained via cross-validation and achieves the same risk asymptotically as the oracle. Let the trial data be partitioned into K mutually exclusive folds. For each observation i, let $\hat{\tau}^{(-i)}_{\text{DR}}$ and $\hat{\tau}^{(-i)}_{\text{QR}}$ denote the learners trained on data excluding the fold containing i, and define $\hat{\tau}^{(-i)}(x;\lambda) := \lambda \cdot \hat{\tau}^{(-i)}_{QR} + (1-\lambda)\cdot\hat{\tau}^{(-i)}_{DR}$. We consider computing the pseudo-outcomes with $\eta = \{0, 0\}$, yielding a model-free combination approach. The K-fold cross-validated pseudo-risk estimate is $\widehat{R}_{CV}(\lambda) = \frac{1}{n_1}\sum_{i=1}^{n_1}\big(\psi_i - \hat{\tau}^{(-i)}(x;\lambda)\big)^2$. The cross-validated weight is then obtained as $\hat{\lambda}^\star = \arg\min_{\lambda\in\Lambda}\widehat{R}_{CV}(\lambda)$.

Theorem 3. *Assume that the pseudo-outcome $\psi(O;\eta = \{0,0\})$ and the base learners $\hat{\tau}_{QR}(X)$ and $\hat{\tau}_{DR}(X)$ are uniformly bounded. Then, we have $R^*(\hat{\lambda}^\star) \leq R^*(\lambda^\star) + o_p(1) \leq \min\{R^*_{DR}, R^*_{QR}\} + o_p(1)$ where $o_p(1)$ converges to 0 as $n \to \infty$.*

The above result describes the oracle excess risk of the combined learner and guarantees that selecting the parameter λ via cross-validation yields a combined learner whose risk differs from the best possible risk by a vanishing term. In other words, the cross-validated choice is asymptotically as good as the oracle-optimal choice and, in large samples, performs at least as well as the better of the two candidate learners. Note that our setting does not require the two component learners to converge to the same CATE function. If they use different model classes and converge to different functions, a linear combination may be strictly better than either component learner. In the special case where both learners converge to the same function such that $R^*_{DR} = R^*_{QR}$, the combined learner also converges to this function.

As a final remark, we note that an alternative approach to combine the QR- and DR-learner would be to take the pseudo-outcomes from their respective first stages and fit a CATE model to a linear combination of these pseudo-outcomes. However, this is a less general solution since our formulation allows the theoretical guarantees in this section to hold for any trial-only learner, including those that do not rely on learning the CATE using pseudo-outcomes such as the S- or T-learner (Künzel et al., 2019).

5 SIMULATION STUDY

We conduct a series of simulations to evaluate the performance of our proposed method against several baselines, focusing on CATE prediction mean squared error. We also assess statistical power to detect treatment effect heterogeneity through interaction tests between the treatment and a hypothesized effect modifier, an analysis commonly performed in real-world trials. For this second evaluation, we consider only methods that support statistical inference. The code to reproduce our experiments is available in the public Github repository `https://github.com/RickardKarl/robust-trial-CATE-augmentation`.

Baselines In addition to our proposed QR-learner and combined learner, 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 same nuisance models that are 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. (2025) called R-OSCAR, in addition to a method from a previous version of their manuscript called CFACE (see Asiaee et al. (2023); an extended discussion on their differences is found in Appendix E); and the linear additive bias correction method of Kallus et al. (2018) which we refer to as KSP (after the authors' initials). Implementation details are provided in Appendix F.2.

5.1 Influence of External Data Sample Size and Population Misalignment

First, we assess the root mean squared error (RMSE) of CATE predictions in a setting with 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 and 4 hold – i.e., the external data are unconfounded and transportability holds – and (ii) a more realistic, challenging setting where these assumptions are violated. We use gradient boosting regressors to estimate the outcome nuisance components and a linear model to estimate both the trial participation probability and the CATE; this modeling choice aligns with the data-generating process, where the baseline outcome is a highly nonlinear function of the covariates, while the CATE is linear.

Table 1 shows that our proposed methods consistently

achieve the lowest or near-lowest RMSE across all settings, with performance improving as the external sample size increases. The trial-only DR- and T-learners performs no better 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 perform best when Conditions 3 and 4 hold, it is noteworthy when these conditions are violated that our methods exhibit RMSEs comparable to that of the trial-only DR-learner. CFACE shows a similar pattern, except when the external dataset is small ($n_0 = 100$), where it performs worse than both the trial-only DR-learner and the ATE predictor. This suggests that their nuisance model fitting is somewhat less robust than ours, as it can underperform relative to the best trial-only CATE learner. Interestingly, R-OSCAR and KSP underperform in all settings; both of these methods rely on fitting separate functions to model the bias between the trial and external population. The difficulty of this bias modeling when the trial size is small could likely explain why both these methods fail.

When examining average prediction bias (full results in Appendix G), we find that all methods are largely unbiased except the pooled T-learner. This is expected since we fit linear CATE models that match the true CATE function. Thus, the observed RMSE gains among the integrative methods using external data primarily stem from variance reduction rather than bias reduction, provided the model is correctly specified.

5.2 Assessing Statistical Power to Detect Interaction Effects

We next assess statistical power to detect treatment effect heterogeneity via interaction tests between the treatment and a hypothesized effect modifier when integrating external data while transportability (Condition 4) is violated. This analysis is conducted in a setting with five observed covariates X, where one is a potential effect modifier $Z \subseteq X$. We consider two scenarios: one in which the effect modifier is present, and another where it is absent. Methods based on the T-learner, R-OSCAR, KSP and the combined learner are excluded, as they do not support inference on the estimated CATE model $\hat{\tau}$. For the remaining methods, we regress $\hat{\tau}$ on Z and perform a two-sided test on the Z coefficient. As a baseline, we also include a simpler test aligned with standard practice in trial analyses, fitting a linear regression of Y on $(A, Z, A \cdot Z)$, testing the $A \cdot Z$ interaction using either only the trial data or the pooled data. We use a significance level $\alpha = 0.05$. Full implementation details of the tests are provided in Appendix F.3.

The results in Figure 1a show that when varying the trial sample size with the external dataset size fixed at $n_0 = 1000$, all methods maintain nominal type I error in the absence of an effect modifier except for the pooled covariate adjustment which can be explained by transportability being violated. Among the methods with nominal type 1 error, the QR-learner and CFACE consistently improve the power by about 10 to 20 percentage points to detect the effect modifier when it is present.

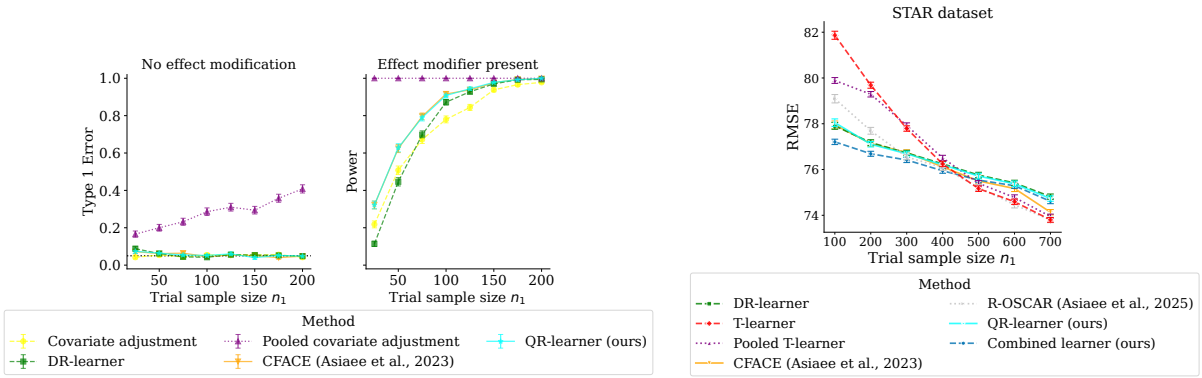
6 CASE STUDY: STAR DATASET

We use data from the Tennessee Student/Teacher Achievement Ratio (STAR) project (Word et al., 1990; Krueger, 1999), a large-scale randomized trial on the effects of class size on student performance. The dataset is divided into two populations based on school location: rural and urban schools. Since both originate from a randomized study, strong ignorability is expected to hold within each population. However, outcome distributions differ across rural and urban schools, and by deliberately omitting school location from the observed covariates, we create a setting where transportability between the two populations is violated. Additional details on the data are provided in Appendix F.4. Performance is measured using RMSE on a held-out test set from the target population, where the pseudo-outcome $\psi(O_i; \eta = \{0, 0\})$ serves as a proxy for the true CATE. Although this proxy has high variance, the use of the known trial propensity score ensures that the estimate aligns in expectation with the true population risk, as discussed in Section 4. Gradient boosting is used for estimating the outcome nuisance models, and ridge-penalized (logistic) linear models to estimate both the trial participation probability and the CATE.

We evaluate the RMSE of each method as the trial sample size increases, while fixing the external data size at $n_0 = 1000$. In Figure 1b, urban schools are treated as the target trial population ($S = 1$) and rural schools as the external population ($S = 0$); the flipped setting is reported in Appendix G. KSP performed significantly worse than the other methods and is therefore omitted from the main figure for ease of visualization, though it is included in the appendix. The integrative methods mostly benefit from external data when the trial sample is small, with gains diminishing as the trial grows. The combined learner consistently outperforms its component learners, the DR- and QR-learners, as predicted by our theory, and also outperforms CFACE and R-OSCAR at smaller sample sizes. The QR-learner offers no improvement over the DR-learner in this setting, though in the flipped setting we observe that it

Table 1: Average root mean squared error reported over 500 repeated runs from the simulation study with a trial sample size $n_1 = 250$ under different scenarios. Lowest number for each scenario is marked with 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)
CFACE (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)
R-OSCAR (Asiaee et al., 2025)	0.60 (4e-03)	0.60 (4e-03)	0.52 (2e-03)	0.58 (2e-03)	0.37 (1e-03)	0.41 (2e-03)
KSP (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 (ours)	**0.28** (4e-3)	0.32 (4e-3)	**0.23** (3e-3)	**0.29** (3e-3)	**0.19** (3e-3)	**0.27** (3e-3)
Combined learner (ours)	0.29 (4e-3)	0.32 (4e-3)	**0.23** (3e-3)	**0.29** (4e-3)	**0.19** (3e-3)	**0.27** (3e-3)



(a) (b)

Figure 1: **(a)**: Evaluating type 1 error (lower better) and power (higher better) in the simulation study with the methods applicable for statistically testing for the presence of an effect modifier as sample size in trial increases, reported over 500 repeated runs. **(b)** We evaluate the RMSE on the STAR dataset when increasing the trial sample size with a fixed external sample size of $n_0 = 1000$. We report the average RMSE and standard error over 200 repeated runs.

improves over the DR-learner. When the trial size becomes large, the trial-only and pooled T-learners, along with CFACE and R-OSCAR, achieve the lowest RMSE. The margin between these approaches and the DR-learner also grows, although this margin depends on whether the target population is urban or rural. This suggests that the advantage may reflect asymmetries between the two populations rather than a general property of the methods. Overall, the results show that our proposed CATE learners can effectively leverage external data to improve prediction accuracy.

7 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. This highlights the potential value of incorporating external data into analyses of heterogeneous treatment effects in randomized trials.

While our proposed methodology performs as intended, it is important to acknowledge its limitations, particularly given its potential impact on real-world decision-making in public health and policy. First, we observe an upper limit to the incremental gain of incorporating additional external data. This is because the external data primarily improves the first-stage estimation of the QR-learner, whereas the second stage relies solely on the randomized trial and is therefore constrained by the trial sample size. Second, although we observe that minimizing the upper bound in Lemma 1 improves CATE accuracy empirically and our method appears robust, it remains of interest to explore under what scenarios tighter bounds can be obtained or more direct strategies for improving CATE estimation can be developed.

Acknowledgments

Research reported in this work was facilitated by the computational resources and support of the Delft AI Cluster (DAIC) at TU Delft. Furthermore, this work was supported by Patient-Centered Outcomes Research Institute (PCORI) award ME-2021C2-22365, National Library of Medicine (NLM) award R01LM013616, and National Heart, Lung, and Blood Institute (NHLBI) award R01HL136708. The content is solely the responsibility of the authors and does not necessarily represent the official views of PCORI, PCORI's Board of Governors, PCORI's Methodology Committee, NLM, or NHLBI.

References

Asiaee, A., Di Gravio, C., Beck, C., Mei, Y., Pal, S., and Huling, J. D. (2025). Improving precision of rct-based cate estimation using data borrowing with double calibration. *arXiv preprint arXiv:2306.17478*.

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

Athey, S. and Imbens, G. (2016). Recursive partitioning for heterogeneous causal effects. *Proceedings of the National Academy of Sciences*, 113(27):7353–7360.

Atienza, D., Bielza, C., and Larrañaga, P. (2022). Pybnesian: An extensible python package for bayesian networks. *Neurocomputing*, 504:204–209.

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

Breiman, L. (1996). Stacked regressions. *Mach. Learn.*, 24(1):49–64.

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

Chernozhukov, V., Hansen, C., Kallus, N., Spindler, M., and Syrgkanis, V. (2024). *Applied Causal Inference Powered by ML and AI*. arXiv:2403.02467.

Curth, A. and Van der Schaar, M. (2021). Nonparametric estimation of heterogeneous treatment effects: From theory to learning algorithms. In *International Conference on Artificial Intelligence and Statistics*, pages 1810–1818. PMLR.

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

Dahabreh, I. J., Hayward, R., and Kent, D. M. (2016). 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.

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

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

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

De Bartolomeis, P., Abad, J., Wang, G., Donhauser, K., Duch, R. M., Yang, F., and Dahabreh, I. (2025). Efficient randomized experiments using foundation models. In *The Thirty-ninth Annual Conference on Neural Information Processing Systems*.

Fawkes, J., O'Riordan, M., Vlontzos, A., Corcoll, O., and Gilligan-Lee, C. M. (2025). The hardness of validating observational studies with experimental data. In *International Conference on Artificial Intelligence and Statistics*, pages 1819–1827. PMLR.

Firth, D. (1993). Bias reduction of maximum likelihood estimates. *Biometrika*, 80(1):27–38.

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

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

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

Hussain, Z., Shih, M.-C., Oberst, M., Demirel, I., and Sontag, D. (2023). 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.

Jahanshahi, M., Gregg, K., Davis, G., Ndu, A., Miller, V., Vockley, J., Ollivier, C., Franolic, T., and Sakai,

S. (2021). The use of external controls in fda regulatory decision making. *Therapeutic Innovation & Regulatory Science*, 55(5):1019–1035.

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

Karlsson, R., Wang, G., De Bartolomeis, P., Krijthe, J. H., and Dahabreh, I. J. (2026). Robust integration of external control data in randomized trials. *Biometrics (forthcoming)*.

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

Krueger, A. B. (1999). Experimental estimates of education production functions. *The quarterly journal of economics*, 114(2):497–532.

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

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

Leitgöb, H. (2020). *Analysis of rare events*. SAGE Publications Limited.

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

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

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

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

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

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

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

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

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

Shalit, U., Johansson, F. D., and Sontag, D. (2017). Estimating individual treatment effect: generalization bounds and algorithms. In *International conference on machine learning*, pages 3076–3085. PMLR.

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

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

Strobl, E. V., Zhang, K., and Visweswaran, S. (2019). Approximate kernel-based conditional independence tests for fast non-parametric causal discovery. *Journal of Causal Inference*, 7(1):20180017.

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

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

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

Weiss, K., Khoshgoftaar, T. M., and Wang, D. (2016). A survey of transfer learning. *Journal of Big data*, 3(1):9.

Word, E., Johnston, J., Bain, H. P., Fulton, B. D., Zaharias, J. B., Achilles, C. M., Lintz, M. N., Folger, J., and Breda, C. (1990). The state of tennessee's student/teacher achievement ratio (star) project. *Tennessee Board of Education*.

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

Wu, L. and Yang, S. (2023). Transfer learning of individualized treatment rules from experimental to real-world data. *Journal of Computational and Graphical Statistics*, 32(3):1036–1045.

Yang, S., Gao, C., Zeng, D., and Wang, X. (2023). Elastic integrative analysis of randomised trial and real-world data for treatment heterogeneity estimation. *Journal of the Royal Statistical Society Series B: Statistical Methodology*, 85(3):575–596.

Checklist

1. For all models and algorithms presented, check if you include:
   (a) A clear description of the mathematical setting, assumptions, algorithm, and/or model. **Yes**: we provide full details on the problem setting and theory in Sections 3 and 4.
   (b) An analysis of the properties and complexity (time, space, sample size) of any algorithm. **Yes**: we theoretically analyze some asymptotic behavior of the combined learner in Section 4.4 and further analyze finite-sample behaviors of all algorithms in our experiments.
   (c) (Optional) Anonymized source code, with specification of all dependencies, including external libraries. **Yes**: we provide code to reproduce our experiments in the supplementary materials.

2. For any theoretical claim, check if you include:
   (a) Statements of the full set of assumptions of all theoretical results. **Yes**: Assumptions are provided in Sections 3 and 4.
   (b) Complete proofs of all theoretical results. **Yes**: Proofs for all theoretical results are provided in Appendix A.
   (c) Clear explanations of any assumptions. **Yes**: We discuss the plausibility of our identifiability assumptions in Section 3 and additional assumptions, such as the need for cross-fitting, in Section 4.

3. For all figures and tables that present empirical results, check if you include:
   (a) The code, data, and instructions needed to reproduce the main experimental results (either in the supplemental material or as a URL). **Yes**: Provided in supplmental materials.
   (b) All the training details (e.g., data splits, hyperparameters, how they were chosen). **Yes**: See Appendix F.2 and F.3.
   (c) A clear definition of the specific measure or statistics and error bars (e.g., with respect to the random seed after running experiments multiple times). **Yes**: See captions of each table and figure.
   (d) A description of the computing infrastructure used. (e.g., type of GPUs, internal cluster, or cloud provider). **Yes**: See Appendix F.5.

4. If you are using existing assets (e.g., code, data, models) or curating/releasing new assets, check if you include:
   (a) Citations of the creator If your work uses existing assets. **Yes**
   (b) The license information of the assets, if applicable. **Not Applicable**
   (c) New assets either in the supplemental material or as a URL, if applicable. **Not Applicable**: No new assets are created.
   (d) Information about consent from data providers/curators. **No**: STAR dataset is a public dataset.
   (e) Discussion of sensible content if applicable, e.g., personally identifiable information or offensive content. **Not Applicable**: The STAR dataset is completely anonymized.

5. If you used crowdsourcing or conducted research with human subjects, check if you include:
   (a) The full text of instructions given to participants and screenshots. **Not Applicable**
   (b) Descriptions of potential participant risks, with links to Institutional Review Board (IRB) approvals if applicable. **Not Applicable**
   (c) The estimated hourly wage paid to participants and the total amount spent on participant compensation. **Not Applicable**

Robust Estimation of Heterogeneous Treatment Effects in Randomized Trials Leveraging External Data: Supplementary Materials

A PROOFS AND DERIVATIONS

A.1 Proof of Theorem 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 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 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

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

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

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

where the last equality follows from conditional unbiasedness such that $\mathbb{E}\left[(\psi_{\text{fixed}} - \tau(X)) \mid X, S=1\right] = 0$. 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

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

□

A.2 Decomposition of the Sample Analog Pseudo-risk

We have that

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

A.3 Proof of Lemma 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 1. Next, we note that

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

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 1.

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

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

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

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

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

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

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

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,

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

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 1) and the third equality from conditional exchangeability in the trial population (Condition 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 $\mathrm{Var}(\psi_{\text{fixed}} \mid S=1)$, we obtain the

inequality

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

where the equality follows from consistency and conditional exchangeability in the trial population again (Condition 1 and 2). □

A.4 Proof of Theorem 2

Proof. We begin by writing

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

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

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

where the first equality follows from that $Y \perp\!\!\!\perp S \mid (X, A)$ holds under Conditions 1-4. To show this, we note that the conditional independencies $Y^a \perp\!\!\!\perp A \mid (X, S=1)$ (Condition 2) and $Y^a \perp\!\!\!\perp A \mid (X, S=0)$ (Condition 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 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 1).

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

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

□

A.5 Proof of Lemma 2

Proof. The inequality $R^*(\lambda^\star) \leq \min\{R^*_{QR}, R^*_{DR}\}$ follows directly from the definition of $\lambda^\star$ as the minimizer of $R^*(\lambda)$ over the set Λ. Since both $\lambda = 0$ and $\lambda = 1$ are in this set, the minimum value $R^*(\lambda^\star)$ cannot be

greater than the value at either endpoint: $R^*(\lambda^\star) \leq R^*(0) = R^*_{DR}$ and $R^*(\lambda^\star) \leq R^*(1) = R^*_{QR}$. Therefore, $R^*(\lambda^\star) \leq \min\{R^*_{DR}, R^*_{QR}\}$.

For uniqueness, we note that the risk $R^*(\lambda)$ can be expanded as

$$\begin{aligned} R^*(\lambda) &= \mathbb{E}\big[(V + \lambda(U - V))^2 \mid S = 1\big] \\ &= \mathbb{E}[(U - V)^2]\lambda^2 + 2\mathbb{E}[V(U - V)]\lambda + \mathbb{E}[V^2] \\ &= A\lambda^2 + B\lambda + C, \end{aligned}$$

where $U = \hat{\tau}_{QR}(X) - \tau(X)$, $V = \hat{\tau}_{DR}(X) - \tau(X)$, $A = \mathbb{E}[(U - V)^2]$, $B = 2\mathbb{E}[V(U - V)]$, and $C = \mathbb{E}[V^2]$. If $A > 0$, the risk function $R^*(\lambda)$ is a strictly convex quadratic function. A strictly convex function has a unique minimum over any convex set, including the set Λ. Therefore, if $A > 0$, the minimizing weight $\lambda^\star$ is unique. □

A.6 Proof of Theorem 3

Proof. Let $\widetilde{R}(\lambda)$ be the expectation of the cross-validated loss term. Since the squared loss is bounded in $[0, B]$, $\Lambda := [0, 1]$ is a compact subset of $\mathbb{R}$, the uniform law of large numbers gives

$$\sup_{\lambda \in \Lambda} \big|\widehat{R}_{CV}(\lambda) - \widetilde{R}(\lambda)\big| \;=\; o_p(1).$$

Further, by definition $\widehat{R}_{CV}(\hat{\lambda}^\star) \leq \widehat{R}_{CV}(\lambda^\star)$, hence

$$\widetilde{R}(\hat{\lambda}^\star) = \widehat{R}_{CV}(\hat{\lambda}^\star) + o_p(1) \;\leq\; \widehat{R}_{CV}(\lambda^\star) + o_p(1) = \widetilde{R}(\lambda^\star) + o_p(1).$$

Finally, it follows from the proof in Lemma 1 that $\widetilde{R}(\lambda) = R^*(\lambda) + C$ for some constant C

$$R^*(\hat{\lambda}^\star) + C \;\leq\; R^*(\lambda^\star) + C + o_p(1) \quad \Longrightarrow \quad R^*(\hat{\lambda}^\star) \;\leq\; R^*(\lambda^\star) + o_p(1).$$

Finally, Lemma 2 yields

$$R^*(\lambda^\star) \;\leq\; \min\{R^*_{DR}, R^*_{QR}\},$$

completing the chain:

$$R^*(\hat{\lambda}^\star) \;\leq\; R^*(\lambda^\star) \;+\; o_p(1) \;\leq\; \min\{R^*_{DR}, R^*_{QR}\} \;+\; o_p(1).$$

□

B CONNECTIONS TO R-LEARNER

It can be shown that the DR-learner (Kennedy, 2023) and the R-learner (Nie and Wager, 2021) both minimize two closely related loss functions when estimating the CATE. Below, we outline this connection and discuss its implications for connections between our proposed QR-learner and the R-learner.

Morzywolek et al. (2023) introduced a general formulation of a loss function,

$$R_o(\tilde{\tau}; \eta, \lambda) = \frac{1}{\mathbb{E}[\lambda\{e(X)\}]} \mathbb{E}\left[\rho\{A, e(X); \lambda\}\{\phi(O; \eta, \lambda) - \tilde{\tau}\}^2 \mid S = 1\right] \tag{6}$$

which, when minimized, yield a broad class of CATE learners. Here, the function $\rho\{A, e(X); \lambda\}$ and pseudo-outcome $\phi(O; \eta, \lambda)$ are defined as

$$\begin{aligned} \rho\{A, e(X); \lambda\} &= \{A - e(X)\}\lambda'\{e(X)\} + \lambda\{e(X)\} \\ \phi(O; \eta, \lambda) &= \frac{\lambda\{e(X)\}}{\rho\{A, e(X)\}} \frac{A - e(X)}{e(X)(1 - e(X))}(Y - g_A(X)) + g_1(X) - g_0(X) \end{aligned}$$

where the nuisance models are $\eta = \{e(X), g_1, g_0\}$ and $g_a = \mathbb{E}[Y \mid X, A = a, S = 1]$ for $a \in 0, 1$.

The loss $R_o(\tilde{\tau}; \eta, \lambda)$ is indexed by a weighting function $\lambda : [0,1] \to \mathbb{R}$ and different CATE learners are obtained depending on this weighting function. For example, the DR-learner and R-learner can both be recovered as special cases with different functions λ.

If we let $\lambda\{e(X)\} \equiv c$ for some constant $c \in \mathbb{R}$, we obtain $\rho\{A, e(X); \lambda\} = c$ and $\phi(O; \eta, \lambda) = \frac{A - e(X)}{e(X)(1-e(X))}(Y - g_A) + g_1(X) - g_0(X)$ which is equivalent to the pseudo-outcome $\psi(O; \eta = \{g_1, g_0\})$ from (3). Consequently, the loss function $R_0(\tilde{\tau}; \eta, \lambda)$ then equals the pseudo-risk $R(\tilde{\tau}; \eta = \{g_1, g_0\})$ of the DR-learner.

Meanwhile, when letting $\lambda\{e(X)\} = e(X)(1 - e(X))$ and following some algebraic steps, we arrive at another expression for the loss function $R_0(\tilde{\tau}; \eta, \lambda)$, given by

$$\frac{1}{\mathbb{E}[e(X)(1 - e(X)) \mid S = 1]} \mathbb{E}\left[(\{Y - q(X)\} - \{A - e(X)\}\, \tilde{\tau}(X))^2 \mid S = 1\right] \ ,$$

where $q(X) = \mathbb{E}[Y \mid X, S = 1] = e(X) \cdot g_1(X) + (1 - e(X)) \cdot g_0(X)$. This expression can be recognized as the loss function of the R-learner (Nie and Wager, 2021).

The above observations highlight an interesting point: the difference between the DR-learner and R-learner depends on the propensity score $e(X)$. If $e(X)$ is constant, then $\lambda\{e(X)\} = e(X)(1 - e(X))$ is also constant, and the two loss functions coincide. Meanwhile, if $e(X)$ is non-constant, the DR-learner and R-learner lead to different CATE learners.

This connection has direct implications for our proposed QR-learner. As discussed in the main paper, the QR-learner has natural links to the DR-learner, since it uses the same loss function but fits the nuisance models differently with the help of external data. When the propensity score $e(X)$ is constant, a scenario realistic in randomized trials which often have fixed treatment probabilities, the R-learner and DR-learner coincide. Hence, in these cases, we argue that the QR-learner can be viewed as minimizing an analogous loss function to either the DR-learner or the R-learner. On the other hand, if the treatment probabilities are covariate-dependent and not fixed, the QR-learner does not minimize the same loss as the R-learner, but rather that of the DR-learner.

C PSEUDO-CODE FOR QR-LEARNER

Algorithm 1 QR-learner: Quasi-optimized Randomization-aware Learner

Require: Data $\mathcal{D} = \{(X_i, S_i, A_i, Y_i)\}_{i=1}^n$, treatment propensity score $e(X)$, model classes $\mathcal{H}$ (for outcome models), $\mathcal{G}$ (for treatment participation probability) and $\mathcal{F}$ (for CATE), number of folds K

Partition $\mathcal{D}$ into K folds $\{\mathcal{D}^k\}_{k=1}^K$, stratified by treatment A and study indicator S.
for $k = 1$ to K **do**
  Stage 1: Estimate nuisance models
  for $a \in \{0, 1\}$ **do**
    Define $\mathcal{D}_a^k = \{i \in \mathcal{D}^k : A_i = a\}$
    Estimate $\Pr(S = 1 \mid X, A = a)$ using $\mathcal{D}_a^k$ using an estimator $\widehat{\pi}_a(X)$ from model class $\mathcal{G}$
    Solve optimization problem in (5)
  end for
  Set $\hat{\eta}^{(k)} = \{\hat{h}_1^{(k)}, \hat{h}_0^{(k)}\}$
  Stage 2: Estimate CATE model
  Let $\mathcal{D}^{-k} = \mathcal{D} \setminus \mathcal{D}^k$
  Compute pseudo-outcome $\psi(O_i; \hat{\eta}^{(k)})$ for each $i \in \mathcal{D}^{-k}$ according to (3)
  Estimate CATE on trial data by solving $\hat{\tau}^{(k)} = \arg\min_{\tilde{\tau} \in \mathcal{F}} \sum_{i \in \mathcal{D}^{-k}: S_i = 1} \left(\psi(O_i; \hat{\eta}^{(k)}) - \tilde{\tau}(X_i)\right)^2$
end for
return $\hat{\tau}(X) = \frac{1}{K} \sum_{k=1}^K \hat{\tau}^{(k)}(X)$

D STATISTICAL INFERENCE WITH QR-LEARNER

In this section, we discuss when statistical inference is feasible for the estimated CATE function $\hat{\tau}(x)$ produced by the QR-learner. Although a full theoretical analysis is beyond the scope of this work, we outline key conditions under which inference is expected to be valid.

Under standard regularity conditions and with the use of cross-fitting, inference on $\hat{\tau}(x)$ becomes possible when the regression of pseudo-outcomes $\widehat{\psi}$ on covariates X is performed using a low-dimensional model (Chernozhukov et al., 2024, Chapter 14). This is enabled by the Neyman orthogonality of the pseudo-outcomes in (3), which ensures that the impact of nuisance estimation errors on $\hat{\tau}(x)$ is second-order (Foster and Syrgkanis, 2023).

Moreover, Kennedy (2023) show that if a "stable regressor" is used to estimate $\tau(x)$ in the DR-learner (we refer to their paper for the formal definition), statistical inference may also be guaranteed. Examples of such regressors include linear regression, smoothing splines, and kernel ridge regression. Importantly, this remains feasible even if the nuisance components $\hat{\eta}$ converge at slower rates than $\hat{\tau}(x)$, since their influence appears as a product of estimation errors – specifically, those from $\hat{\eta}$ and the estimated propensity score (if one were to estimate it). In randomized trials, where the true propensity score is known, we could expect this product to vanish rapidly, making the impact of nuisance estimation asymptotically negligible on the final CATE estimate $\hat{\tau}(x)$. In this case, asymptotically valid inference for $\hat{\tau}(x)$ may be warranted as long as the regressor used for the final stage regressor has inferential guarantees (Kennedy, 2023).

E DISCUSSION ON METHODS FROM ASIAEE ET AL.

In this section, we explain the method R-OSCAR (Robust Observational Studies for CMO-Augmented RCT) from Asiaee et al. (2025) as well as another variant proposed in an earlier version of their manuscript (see Asiaee et al. (2023)), refered to as CFACE (CounterFactual Average Covariate Effect). We start with explaining CFACE since it came out first.

CFACE The central idea of CFACE is to estimate a trial-specific CATE learner leveraging pseudo-outcomes which fits into our randomization-aware framework, but using a different procedure for estimating the outcome nuisance models. Specifically, we instead consider randomization-aware pseudo-outcomes on the restricted form:

$$\psi(O; \eta = \{m, m\}) = \frac{A - e(X)}{e(X)(1 - e(X))}(Y - m(X)), \tag{7}$$

where the propensity score $e(X)$ is known and the nuisance models are constrained to be identical, $h_1 = h_0 = m$. Then, Asiaee et al. (2023) show that m can be chosen to minimize the conditional variance,

$$\begin{aligned} m^*(x) &= \arg\min_m \operatorname{Var}\left(\psi(O; \eta = \{m, m\}) \mid X = x, S = 1\right) \\ &= e(x) \cdot \mu_0^{S=1}(x) + (1 - e(x)) \cdot \mu_1^{S=1}(x), \end{aligned}$$

with $\mu_a^{S=1}(x) = \mathbb{E}[Y \mid X = x, A = a, S = 1]$. Under transportability (Condition 4), $\mu_a^{S=1}(x)$ also equals $\mu_a^{S=0}(x) = \mathbb{E}[Y \mid X = x, A = a, S = 0]$ and can thus be estimated from external data. Thus, we can fit $\mu_a^{S=0}(x)$ on only the external data for $a \in \{0, 1\}$ and then compute pseudo-outcomes using m^* following a two-stage procedure similar to the DR- and QR-learners. However, in their updated manuscript, the authors acknowledge that this procedure is possible but caution that it may fail when trial and external populations are misaligned (i.e., when Condition 4 does not hold). For this reason, they now recommend against this and instead focus on a new CATE learner, called R-OSCAR.

R-OSCAR R-OSCAR departs from CFACE by proposing to fit $\mu_a^{S=0}(x)$ on external data and correct for possible transportability violations by modeling the difference

$$\delta_a(x) = \mathbb{E}[Y \mid X = x, A = a, S = 1] - \mathbb{E}[Y \mid X = x, A = a, S = 0]. \tag{8}$$

Since their focus is on the CATE itself, they further propose modeling the direct discrepancy in the CATE between the populations:

$$\delta(x) = \mathbb{E}[Y^1 - Y^0 \mid X = x, S = 1] - \mathbb{E}[Y^1 - Y^0 \mid X = x, S = 0]. \tag{9}$$

Using the optimality of m^*, computed using estimates of $\mu_a^{S=0}$ and δ_a, and the unbiasedness of $\psi(O; \eta = \{m^*, m^*\})$ for the true CATE, they derive a procedure to also estimate $\delta(x)$. Their final CATE learner, R-OSCAR, is then

$$\hat{\tau}_{\text{R-OSCAR}}(x) = (\hat{\mu}_1^{S=0}(x) + \hat{\delta}_1(x)) - (\hat{\mu}_0^{S=0}(x) + \hat{\delta}_0(x)) + \hat{\delta}(x), \tag{10}$$

where $\hat{\mu}_a^{S=0}$, $\hat{\delta}_a$, and $\hat{\delta}$ are estimated using trial and external data. Modeling the differences δ_a and δ resembles the additive bias correction of Kallus et al. (2018) more than our approach. In fact, R-OSCAR could be seen as a T-learner with additive bias correction, and diverges substantially from the two-stage procedures of the DR- and QR-learners. By contrast, the implementation of CFACE follows a two-stage design and aligns more closely with our framework.

F EXPERIMENTAL DETAILS

F.1 Data-Generating Process in Simulation Studies

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 were sampled according to $\varepsilon_i \sim N(0, \sigma^2 = 1/4)$.

For the experiment in Section 5.1, 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 and 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.

For the experiment in Section 5.2, we considered a setting with a linear baseline outcome and a sparse linear CATE that depended only on a single covariate. To violate transportability (Condition 4), this time we encoded different functions $\tau(X)$ for the trial and external population. Specifically, we defined:

$$b(X_i) = \sum_{j=1}^{d} \frac{1}{d} X_{ij},$$
$$\tau(X_i) = \begin{cases} \beta \cdot d \cdot X_{i1}, & \text{if } S_i = 1 \\ \left(\beta + \frac{1}{20}\right) \cdot d \cdot X_{i1}, & \text{otherwise} \end{cases}$$

where β was a tunable parameter that controlled the size of the interaction effect between the treatment A_i and the first covariate X_{i1}. As before, we set $d = 5$.

F.2 Implementation Details For CATE Learners

For the experiments in Section 5.1 and 6, we implemented the CATE learners as follows.

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, CFACE, and QR-learner, we used histogram-based gradient boosting regression tree using default hyperparameters. 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 KSP, we fitted the DR-learner on the external dataset and then used a linear regression to estimate the bias model. For R-OSCAR, we also used histogram-based gradient boosting regression for the outcome nuisance models and a cross-validated linear regression with ridge penalty for estimating both its bias functions. We applied cross-fitting to all two-stage CATE learners using two folds consistently. For cross-fold validation in the combined 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.

For the case study Tennessee STAR dataset, we made a few changes to the implementation of all learners. First, due to having a large number of features relative to the sample size, we use ridge-penalized linear regression for the CATE model, and further changed from the default hyperparameters of the histogram-based gradient boosting regression to a max depth of 3 for the decision trees (default is unconstrained depth) and a minimum sample size per leaf of 5 (default is 20). Finally, we used 10 folds for the cross-validation in the combined learner, to prevent the training splits from becoming very small in size.

F.3 Statistical Tests for Treatment Effect Heterogeneity

Below we describe our implementations for the statistical tests used to detect treatment effect modification in the experiment in Section 5.2.

For covariate adjustment, we fit a linear regression model of Y on the covariates $(A, X_1, A \cdot X_1)$ using the trial data. We then estimated 95% confidence intervals for the coefficient of $A \cdot X_1$ to assess whether it was significantly different from zero. For the pooled covariate adjustment, we followed the same approach but fit the linear regression model on the combined dataset consisting of both the trial data and the external data.

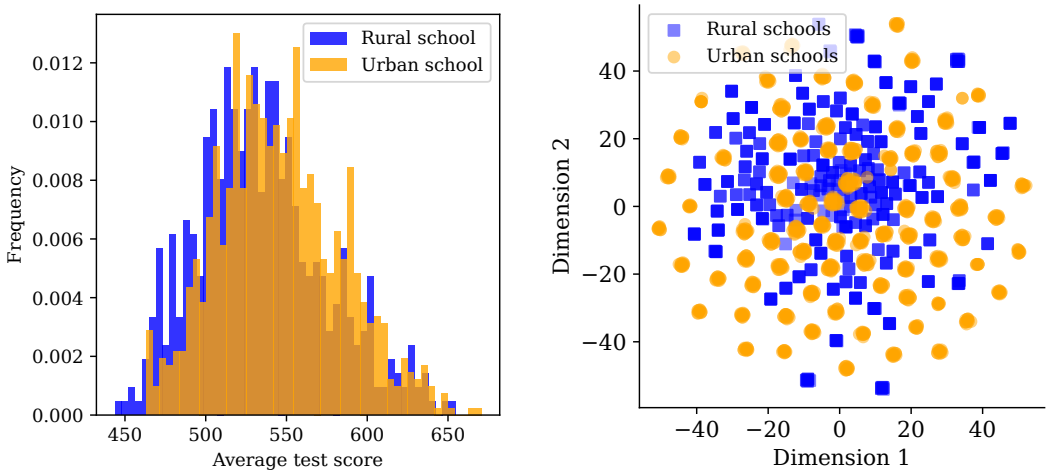
For the DR-learner, QR-learner, and CFACE, we split the data into two folds (stratified by the treatment A and the study population S). We used the first fold to estimate the nuisance components in η via linear regression using their respective strategies as outlined in the main paper, then computed pseudo-outcomes on the second fold. These pseudo-outcomes were regressed on the feature X_1 using another linear regression model, resulting in the first CATE estimate $\hat{\tau}(X_1) = \hat{\alpha}^{(1)} \cdot X_1 + \hat{\beta}^{(1)}$. To utilize the entire dataset efficiently, we repeated the process by swapping the folds, obtaining a second CATE estimate $\hat{\tau}(X_1) = \hat{\alpha}^{(2)} \cdot X_1 + \hat{\beta}^{(2)}$. We then computed 95% confidence intervals to test the null hypothesis that $\alpha = 0$ as follows:

$$\left[\bar{\alpha} - 1.96 \cdot \mathrm{se}(\bar{\alpha}),\ \bar{\alpha} + 1.96 \cdot \mathrm{se}(\bar{\alpha})\right],$$

where $\bar{\alpha} = \frac{1}{2}(\hat{\alpha}^{(1)} + \hat{\alpha}^{(2)})$ and using the individual standard errors $\mathrm{se}(\hat{\alpha}^{(1)})$ and $\mathrm{se}(\hat{\alpha}^{(2)})$ to compute

$$\mathrm{se}(\bar{\alpha}) = \sqrt{\frac{1}{4}\left(\mathrm{se}(\hat{\alpha}^{(1)})^2 + \mathrm{se}(\hat{\alpha}^{(2)})^2\right)},$$

assuming normality and that the covariance between $\hat{\alpha}^{(1)}$ and $\hat{\alpha}^{(2)}$ to be negligible.



(a) (b)

Figure 2: **(a)**: Distributions of the outcome (average test scores) for rural and urban schools in the STAR dataset. We observe a slight shift in the mean between the two groups, suggesting potential violations of transportability, as the primary difference between the trial and the external population lies in school location. **(b)**: t-SNE plot over features colored by study population. We observe some lack of overlap between populations.

F.4 Tennessee STAR Dataset

The data from the STAR study consists of 4218 students: 2811 from rural schools and 1407 from urban schools. The treatment is class size: small ($A = 0$) versus regular ($A = 1$), and the continuous outcome Y is the student's average test score. Our observed covariates include gender, race, birth date, teacher ID, and free lunch eligibility. After applying one-hot encoding to these variables, which are either binary or categorical, we obtain a 310-dimensional feature vector.

To construct two datasets from different populations, we split the dataset into rural and urban school. One school location is assigned as the target trial population and the other as the external population. In the experiment, we sample without replacement half of the target trial population as a held-out test set, and then sample without replacement n_1 and n_0 observations from the target trial and external populations, respectively. We kept $n_0 = 1000$ fixed while varying n_1 from 100 to 700 (about half of the number of urban schools).

We further evaluate whether transportability (Condition 4) is violated between the two populations in this dataset. As shown in Figure 2a, we can see that the distribution of outcomes (average test scores) differs between rural and urban schools. This may indicate that transportability also is violated if the school location is omitted from the observed covariates X. We performed a conditional independence test of $Y \perp\!\!\!\perp S \mid (X, A)$, which is a testable implication for Conditions 1 to 4 jointly (Dahabreh et al., 2024). We applied both a partial linear correlation test and the randomized conditional correlation test (RCoT) (Strobl et al., 2019), implemented in the Python package *pybnesian* (Atienza et al., 2022). Both tests indicate that the conditional independence is violated at the 5% significance level, suggesting that one of Conditions 1 to 4 may not hold. Since consistency (Condition 1) and strong ignorability in both populations (Conditions 2 and 3) are expected to hold, it is most likely that transportability (Condition 4) is violated.

We further used t-SNE for dimensionality reduction to visualize the distribution of covariates across the two populations. These visualizations, shown in Figure 2b, reveal limited overlap in certain regions between rural and urban schools.

Table 2: Average prediction bias reported over 500 repeated runs from the simulation study with a trial sample size $n_1 = 250$ under different scenarios.

External sample size	100		1000		10000	
Condition 3 and 4 violated?	No	Yes	No	Yes	No	Yes
DR-learner	-0.00 (5e-03)	-0.00 (5e-03)	-0.00 (5e-03)	-0.01 (5e-03)	0.00 (4e-03)	0.01 (5e-03)
T-learner	0.01 (4e-03)	0.01 (5e-03)	0.01 (5e-03)	0.00 (5e-03)	0.01 (4e-03)	0.00 (5e-03)
Pooled T-learner	0.06 (3e-03)	0.12 (4e-03)	0.05 (2e-03)	0.20 (3e-03)	-0.03 (1e-03)	0.11 (1e-03)
CFACE (Asiaee et al., 2023)	0.01 (5e-03)	0.00 (5e-03)	0.01 (4e-03)	-0.00 (5e-03)	0.00 (3e-03)	0.00 (4e-03)
R-OSCAR (Asiaee et al., 2025)	0.00 (6e-03)	0.00 (5e-03)	0.00 (4e-03)	0.01 (5e-03)	-0.00 (3e-03)	0.00 (4e-03)
KSP (Kallus et al., 2018)	0.01 (5e-03)	-0.01 (5e-03)	-0.01 (6e-03)	-0.01 (6e-03)	-0.01 (5e-03)	0.01 (6e-03)
QR-learner (ours)	0.01 (5e-03)	-0.00 (5e-03)	0.01 (4e-03)	-0.00 (4e-03)	0.00 (3e-03)	0.00 (4e-03)
Combined learner (ours)	0.00 (5e-03)	0.00 (5e-03)	0.00 (4e-03)	0.00 (4e-03)	-0.00 (3e-03)	0.00 (4e-03)

F.5 Compute Resources Used for Experiments

All experiments were run on a CPU machine. The simulations in Sections 5 and 6 each completed in under 72 hours on a laptop with a 2 GHz Quad-Core Intel Core i5 processor and 16GB of RAM. Including preliminary experiments not shown in the paper, no single run exceeded this runtime.

G ADDITIONAL EXPERIMENTAL RESULTS

G.1 Analysis of Prediction Bias

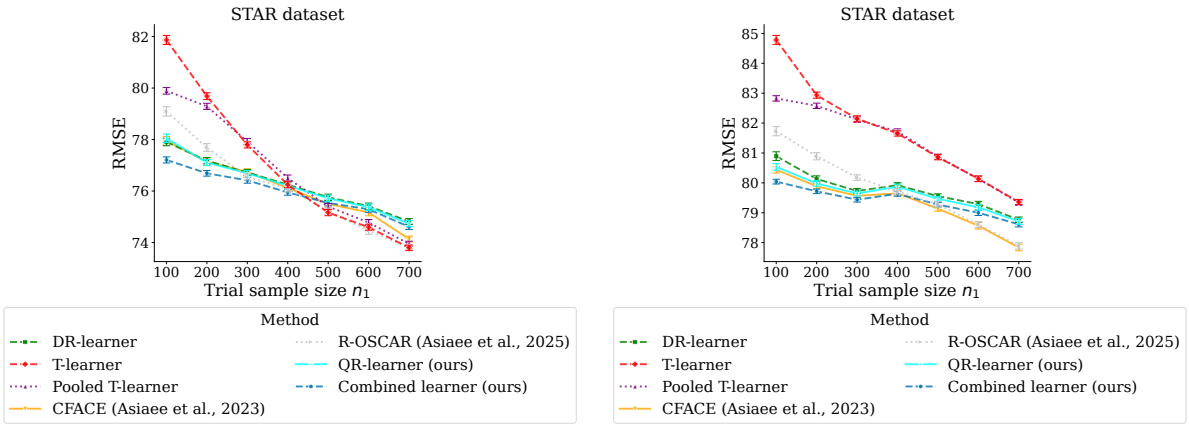
As a complement to Table 1, which reports the average root mean squared error of each method in our simulation study, Table 2 presents the corresponding average prediction bias, defined as $\text{bias}(\hat{\tau}) = \mathbb{E}[\hat{\tau}(X) - \tau(X) \mid S = 1]$. We find that all methods, except the pooled T-learner, were largely unbiased. This aligns with expectations, as we fitted linear models for the CATE which matches with the true underlying linear CATE function. The pooled T-learner, however, exhibited increasing bias when conditions 3 and 4 were violated, which is consistent with its reliance on these conditions for identification of the CATE.

G.2 Tennessee STAR Dataset

We include both settings from the Tennessee STAR dataset case study. In the first setting, presented in the main paper, we treat the urban schools as the target trial population and the rural schools as the external population. We then flipped the populations, treating the rural schools as the target trial population. Overall, we observe similar trends in both cases, but we highlight below the differences that arise when flipping the two populations. The results from both settings are shown in Figure 3.

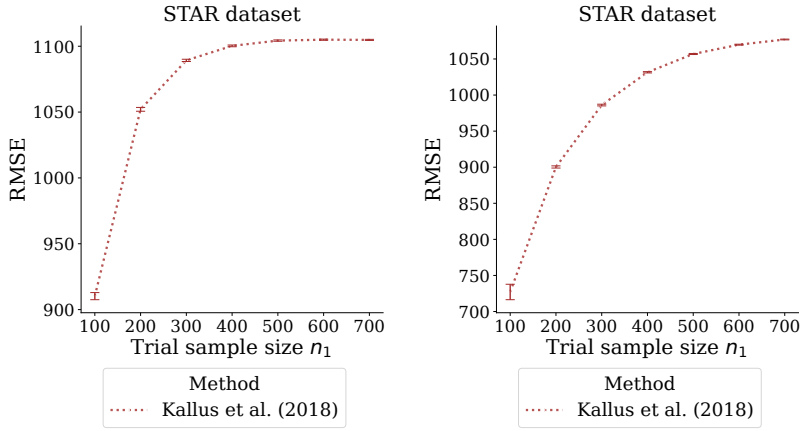
First, we observe that the trial-only and pooled T-learner improve faster when the urban schools are the target population compared to when the rural schools are the target. Second, CFACE and R-OSCAR perform better when the rural schools are the target population. Finally, the QR-learner shows little improvement over the DR-learner when the urban schools are the target, but it exhibits noticeable improvement when the rural schools are the target.

We also include the performance of KSP, the additive bias correction method of Kallus et al. (2018), in Figure 4. This method exhibits significantly higher RMSE than the other methods and, unusually, its performance worsens as the number of trial samples increases. While we cannot fully explain this behavior, one possible reason is that, unlike the other learners, this approach is not specifically designed for estimating a trial-specific CATE.



(a) Urban school as target population (b) Rural schools as target population

Figure 3: We evaluate the RMSE on the STAR dataset when increasing the trial sample size with a fixed external sample size of $n_0 = 1000$. We report the average RMSE and standard error over 200 repeated runs.



(a) Urban school as target population (b) Rural schools as target population

Figure 4: Separate plot for additive bias correction method of Kallus et al. (2018) because its large RMSE values make it difficult to display alongside the other methods. We evaluate the RMSE on the STAR dataset when increasing the trial sample size with a fixed external sample size of $n_0 = 1000$. We report the average RMSE and standard error over 200 repeated runs.