ORTHOGONAL LEARNER FOR ESTIMATING HETEROGENEOUS LONG-TERM TREATMENT EFFECTS

Anonymous authors
Paper under double-blind review

ABSTRACT

Estimation of heterogeneous *long-term* treatment effects (HLTEs) is relevant for personalized decision-making in marketing, economics, and medicine, where short-term datasets are often combined with long-term observational datasets. However, HLTE estimation is challenging due to limited overlap in treatment assignments or in long-term outcomes for certain subpopulations, which can lead to unstable HLTE estimates with large finite-sample variance. To address this challenge, we introduce the **LT-O-learners** (Long-Term Orthogonal Learners), a set of novel orthogonal learners for HLTE estimation in the canonical HLTE setting with surrogacy. The key idea of our LT-O-learners is to *retarget* the loss via custom overlap weights that downweight low-overlap samples. We show that the retargeted loss recovers the true HLTE pointwise and satisfies Neyman-orthogonality. We further prove two key theoretical results: (i) The nuisance error enters the error bound only through higher-order terms, which means our learners are robust to nuisance estimation error. (ii) Under a linear function class, the retargeting effectively controls the asymptotic variance of the HLTE estimator via the overlap weights in low-overlap regimes. We conduct experiments on synthetic and real-world datasets to confirm the theoretical properties of our LT-O-learners, particularly robustness in low-overlap regimes. To our knowledge, ours are the first orthogonal learners for HLTE estimation robust to low overlap in long-term settings.

1 INTRODUCTION

Estimating heterogeneous *long-term* treatment effects (HLTEs) from short-term studies is a common problem in marketing, economics, and medicine (Feuerriegel et al., 2024; Yang et al., 2024b; Athey et al., 2026). In many real-world applications, the primary outcomes of interest materialize only after substantial delays or incur high measurement costs. As a result, short-term studies typically record only surrogate outcomes, and must be combined with auxiliary data sources that capture long-term outcomes in order to estimate HLTEs (Cai et al., 2024; 2025; Chen et al., 2025a).

Example. In job training, policy-makers seek long-term effects on employment (e.g., earnings in 3 years) but only observe employment status or earnings within a short time period (Hotz et al., 2006). Similarly, recommending a new therapy targets long-term benefits (e.g., years of life gained), while only short-term surrogates such as tumor shrinkage are available (Belin et al., 2020). Timely decisions thus require estimating heterogeneous long-term effects from short-term outcomes, aided by historical data that contain long-term outcomes but may lack treatment information.

Formally, the task of estimating long-term treatment effects is typically formalized as follows (Chen & Ritzwoller, 2023; Yang et al., 2024a; Athey et al., 2026): a short-term dataset $\mathcal{D}_1$ (either experimental or observational), in which treatment is observed alongside surrogate outcomes, and a long-term observational dataset $\mathcal{D}_2$, in which long-term outcomes are recorded but treatment assignment may be missing or unobserved (e.g., because it is a new medical treatment or marketing intervention that was not yet applied before). The two datasets have shared covariates, which then enable identification and estimation of long-term treatment effects under suitable assumptions. In this paper, we focus on the canonical setting with surrogacy and comparability assumptions (Athey et al., 2026; Chen & Ritzwoller, 2023; Yang et al., 2024a; Kallus & Mao, 2025). Under identification, classic confounding

Reviewers: please read the Reviewer Guidelines (iclr.cc/Conferences/2027/ReviewerGuidelines) and the AI Policy for Reviewers (iclr.cc/Conferences/2027/AIPolicyForReviewers). If you used AI to expand, edit, or polish your review, please provide the input text to the LLM. Better still, consider skipping the LLM and submitting your original text: we, and the authors, are much more interested in your unedited thoughts than in what an LLM has to say. AI-assisted or not, you are putting your name and reputation behind your review: LLM-generated falsehoods, hallucinations or misrepresentations are subject to disciplinary action, which may include desk-rejecting all papers you have authored.

adjustment strategies can be adapted to derive meta-learners[1] for HLTE estimation, such as the *long-term* T-learner, RA-learner, IPW-learner, and DR-learner (See Section 4, Appendix H.1).

However, the meta-learners mentioned above fail to address challenges due to *low overlap*, which can lead to unstable estimates with large finite-sample variance (Jesson et al., 2020; Fisher, 2023). Broadly, overlap refers to the availability of comparable samples across relevant regimes for individuals with similar covariates. In the HLTE setting, overlap arises in two different forms. (i) ***Limited treatment overlap:*** in the short-term dataset $\mathcal{D}_1$, some subpopulations may have a low probability of receiving one of the treatment arms, leading to limited treatment overlap in finite samples. (ii) ***Limited long-term outcome overlap:*** unique to the long-term setting is that some subpopulations may have a low probability of observing long-term outcomes in the historical dataset $\mathcal{D}_2$. As a result, when overlap is weak in either dimension, HLTE meta-learners are prone to high finite-sample variance and unstable subgroup-level estimates (see Section 4, 5.2 and Appendix F.3).

In this paper, we address these challenges by proposing the **LT-O-learners** (Long-Term Orthogonal Learners), a novel class of orthogonal meta-learners for estimating HLTEs from a short-term dataset $\mathcal{D}_1$ and a long-term observational dataset $\mathcal{D}_2$. The key idea is to *retarget* the learning objective via overlap weights that down-weight samples with weak treatment or long-term outcome overlap, yielding a loss that still recovers the true HLTE pointwise and satisfies Neyman-orthogonality. Because the efficiency-optimal weight is intractable—it depends on all nuisances and breaks orthogonality—we instead build a class of orthogonal learners with customizable weights, instantiated by four canonical members (LT-O-DR-learner, LT-O-TO-learner, LT-O-LO-learner, LT-O-DO-learner) that target different overlap regimes. We further prove the benefits of Neyman-orthogonality and overlap weighting with two respective theorems to highlight the theoretical strength. Importantly, our LT-O-learners are fully model-agnostic and can be instantiated with arbitrary machine learning models, making them broadly applicable in high-dimensional HLTE settings.

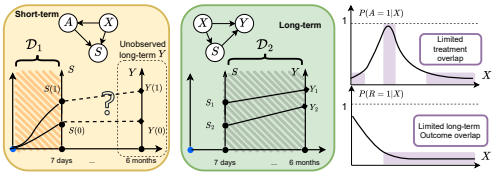


Figure 1: **Our HLTE setting:** The yellow and green box illustrate our two-sample setting. $\mathcal{D}_1$ and $\mathcal{D}_2$ share covariates X and surrogates S. The surrogate index (i.e. $\mathbb{E}_{\mathcal{D}_2}[Y|S,X]$), learned in the long-term dataset, is used as a proxy for long-term outcomes in the short-term dataset. The left plot shows the two overlap issues.

Our **contributions** are:[2] (1) We propose the *LT-O-learners*, a novel class of orthogonal meta-learners for HLTE estimation that are robust to different forms of low overlap. (2) We prove two theoretical results: (i) the nuisance error enters the error bound only through higher-order terms; and (ii) under a linear function class, the weighting scheme controls the asymptotic variance of the HLTE estimator in low-overlap regimes. (3) We establish empirical well-posedness of the loss. (4) We conduct experiments on synthetic and real-world datasets to confirm the theoretical properties of our LT-O-learners, particularly robustness in low-overlap regimes.

2 Related Work

Orthogonal learners for heterogeneous treatment effects (HTEs): Several machine learning methods have been developed for the standard HTE setting with a single dataset (e.g., Shalit et al., 2017; Curth & van der Schaar, 2021), where, unlike our HLTE setting, the outcome of interest is directly available in the randomized (or observational) data. State-of-the-art methods are model-agnostic two-stage estimators, in which nuisance functions are estimated in the first stage and where the target HTE is learned by minimizing an orthogonal loss function (Foster & Syrgkanis, 2023) in the second stage to predict the HTEs. Orthogonality ensures robustness w.r.t. nuisance estimation errors (Chernozhukov et al., 2018) and thus prevents plug-in bias (Kennedy, 2022). Examples include the DR-learner (Kennedy, 2023), and R-learner (Nie & Wager, 2021). However, these learners are *not* directly transferable from the HTE to the HLTE setting.

[1]Meta-learners are model-agnostic algorithms that decompose heterogeneous treatment effect estimation into a collection of sub-regression problems (Künzel et al., 2019; Curth & van der Schaar, 2021).

[2]Code is available at https://anonymous.4open.science/r/Orthogonal-learners-for-Long-term-effects-4BF8.

Average long-term treatment effects (ALTEs): A related but different setting aims at estimating *average* long-term treatment effects (ALTEs). In this setting, ALTEs are typically identified by mediating the treatment effects via surrogate variables under suitable assumptions (e.g., Prentice, 1989; Frangakis & Rubin, 2002; Lauritzen et al., 2004). One approach is to build a "surrogate index" (i.e., a group of short-term endpoints) as the proxy for long-term outcomes under unconfoundedness (Athey et al., 2026; Chen & Ritzwoller, 2023). Subsequent works estimate ALTEs with modified settings and assumptions (e.g., Athey et al., 2020; Ghassami et al., 2022; Goffrier et al., 2023; Obradović, 2024; Cai et al., 2024; Imbens et al., 2025; Huang et al., 2026). Other works also study ALTE estimation in the context of policy optimization tasks (e.g., McDonald et al., 2023; Park & Sasaki, 2024; Yang et al., 2024a;b; Wang et al., 2022; Saito et al., 2024). However, these works all target the ALTE, which only captures *population-level* effects and which thus *ignores* how long-term treatment effects vary across individuals or subpopulations. We summarize the different settings for long-term causal effects in Table 3.

Heterogeneous long-term treatment effects (HLTEs): There are some works studying HLTEs under alternative settings, such as *dose–response* estimation (Singh & Zhou, 2022; Yang et al., 2025) (See Table 3 for an overview of the long-term causal effects literature). These approaches focus on *continuous* treatments and rely on different estimation strategies because of which they are *not* applicable to the canonical HLTE setting with binary treatments considered in this paper. Huang & Ascarza (2024) constructs a noise-reduced surrogate index via separate imputation; however, this requires additional long-term variables other than Y and is thus *not* applicable to our setting. Chen et al. (2025a) proposes an EIF-based orthogonal HLTE estimator in the setting of Conditional Additive Equi-Confounding Bias (Ghassami et al., 2022). However, it requires observed treatment in both datasets and does not adjust for confounding in the canonical setting with the surrogacy assumption. Thus, it is *not* applicable. Finally, these learners are *not* designed to be robust to (i) low treatment overlap and (ii) low long-term outcome overlap in finite samples.

Overlap-weighted estimators for ATEs/HTEs: A separate line of work uses retargeted overlap weighting to stabilize ATE/HTE estimation in low-overlap regimes. Crump et al. (2009) retarget the average treatment effect (ATE) to a weighted ATE on the well-supported subpopulation, thereby reducing finite-sample variance at the price of changing the estimand. For heterogeneous effects, R-learner (Nie & Wager, 2021) implicitly weights its loss by the treatment propensity $\pi(1-\pi)$ (Fisher, 2023), and Morzywołek et al. (2023) generalize this to a general class of treatment-propensity-weighted orthogonal learners; Hess et al. (2026) further extend treatment-overlap weighting to the longitudinal HTE estimation. However, these works are *not* tailored for the HLTE setting, and *do not* address the limited overlap in observing long-term outcomes.

Research gap: Existing methods for HLTE estimation fail to address the challenge due to (i) limited treatment overlap and (ii) limited long-term outcome issue in finite samples (see Table 1). To the best of our knowledge, we are the first to provide a general class of orthogonal learners that is robust in low-overlap regimes[3].

3 Problem formulation

Setup: We follow the canonical long-term outcome setting (Athey et al., 2026; Chen & Ritzwoller, 2023), but adapt it from ALTEs to HLTEs. We thus consider two datasets: a short-term dataset $\mathcal{D}_1$ and a long-term historical observational dataset $\mathcal{D}_2$, with N_1 and N_2 units respectively. $\mathcal{D}_1$ and $\mathcal{D}_2$ differ in terms of which variables are observed (see Figure 1).

Let $A \in \{0,1\}$ be the treatment variable, $X \in \mathcal{X} \subset \mathbb{R}^{d_x}$ be pre-treatment covariates, $S \in \mathcal{S} \subset \mathbb{R}^{d_s}$ be the short-term surrogate variable, $Y \in \mathbb{R}$ be the long-term outcome, and $R \in \{0,1\}$ be the dataset indicator, denoting whether the observation belongs to $\mathcal{D}_1$ ($R=0$) or $\mathcal{D}_2$ ($R=1$).

The covariates X and short-term surrogates S are observed in both $\mathcal{D}_1$ and $\mathcal{D}_2$. The long-term outcome Y is observed only in $\mathcal{D}_2$. The treatment A is only observed in the short-term dataset $\mathcal{D}_1$ but *not* in $\mathcal{D}_2$ (e.g., because this is a new treatment that was not applied previously). Formally, we can combine $\mathcal{D}_1$ and $\mathcal{D}_2$ to build a unified dataset with $N = N_1 + N_2$ units: $\mathcal{D} = \mathcal{D}_1 \cup \mathcal{D}_2 = \{(X_i, (1-R_i)A_i, S_i, R_i, R_iY_i)\}_{i=1}^{N}$.

Nuisance functions: Our method relies on a set of nuisance functions:

[3]See Appendix A for extended related work.

$$\pi_s(s,x) = \Pr(A=1 \mid S=s, X=x, R=0), \quad \pi(x) = \Pr(A=1 \mid X=x, R=0), \tag{1}$$

$$\rho_s(s,x) = \Pr(R=1 \mid S=s, X=x), \quad \rho(x) = \Pr(R=1 \mid X=x), \tag{2}$$

$$h(s,x) = \mathbb{E}[Y \mid S=s, X=x, R=1], \quad \mu(a,x) = \mathbb{E}\big[h(S,X) \mid A=a, X=x, R=0\big]. \tag{3}$$

We let η denote the complete set of nuisance functions, i.e.,

$$\eta = (\pi(x), \pi_s(s,x), \rho(x), \rho_s(s,x), h(s,x), \mu(a,x)). \tag{4}$$

Causal estimand: We build upon the potential outcome framework (Rubin, 1972) and aim to estimate the HLTE in the short-term dataset, given by

$$\tau^0(x) = \mathbb{E}\left[Y(1) - Y(0) \mid X=x, R=0\right]. \tag{5}$$

This estimand captures how long-term treatment effects vary at the individual or subgroup level as a function of covariates X, thus allowing for fine-grained heterogeneity in long-term responses to inform personalized decision-making. In contrast, population-level estimands such as the ALTE as in Athey et al. (2026) target average effects across the entire population and do not characterize which individuals or subpopulations benefit most from the treatment.

Identifiability: We use the following standard assumptions from Athey et al. (2026) to ensure the identification.

Assumption 3.1 (Consistency)**.** *For all $a \in \{0,1\}$, we have $S(a) = S$, $Y(a) = Y$ whenever $A = a$.*

Assumption 3.2 (Positivity)**.** *There exist constants $\alpha_\pi, \alpha_\rho, \alpha_R > 0$ such that, for every $x \in \mathcal{X}$, $\alpha_\pi < \pi(x) < 1 - \alpha_\pi$, and $1 - \alpha_R > \rho(x) > \alpha_\rho$.*

Assumption 3.3 ($\mathcal{D}_1$ unconfoundedness)**.** *For all $a \in \{0,1\}$, we have $A \perp\!\!\!\perp S(a), Y(a) \mid X, R=0$.*

Assumption 3.4 (Surrogacy (Prentice, 1989))**.** *The long-term effect is fully mediated by the surrogate variable in $\mathcal{D}_1$: $A \perp\!\!\!\perp Y \mid S, X, R=0$* and $0 < P(R=1 \mid S=s, X=x) < 1$.

Assumption 3.5 (Long-term outcome comparability)**.** *The distribution of Y is invariant to whether it belongs to $\mathcal{D}_1$ or $\mathcal{D}_2$: $R \perp\!\!\!\perp Y \mid S, X$.*

Assumptions 3.1, 3.2, 3.3 are standard in the observational causal inference literature (Rubin, 1974; Imbens & Angrist, 1994; Kallus & Mao, 2025; Athey et al., 2026). Assumption 3.4 is widely adopted in prior literature on long-term effects (Athey et al., 2026; Chen & Ritzwoller, 2023; Yang et al., 2024a; Kallus & Mao, 2025). Assumption 3.5 requires that the conditional distribution of Y given (S, X) is transportable between $\mathcal{D}_1$ and $\mathcal{D}_2$, which is a mild assumption that is common in most data combination tasks (Athey et al., 2026; Chen & Ritzwoller, 2023; Yang et al., 2024a). When all the above assumptions are satisfied, we call the problem the *standard surrogacy model* (Chen & Ritzwoller, 2023). Under the standard surrogacy model, $h(S,X) = \mathbb{E}[Y \mid S, X, R=1]$ is an unbiased proxy of $\mathbb{E}[Y(1) \mid S, X, R=0]$ and $\mathbb{E}[Y(0) \mid S, X, R=0]$ on $\mathcal{D}_1$[4]. We can then identify the HLTEs via (see Appendix C.1 for a formal proof):

$$\begin{aligned} &\tau^0(x) = \mathbb{E}[Y(1) - Y(0) \mid X=x, R=0] \\ =&\mathbb{E}\big[h(S,X) \mid A=1, X=x, R=0\big] - \mathbb{E}\big[h(S,X) \mid A=0, X=x, R=0\big] = \mu(1,x) - \mu(0,x). \end{aligned} \tag{6}$$

Note on relaxing assumptions: Surrogacy is a strong assumption, which requires S to fully mediate the effect of A on Y, and $A \perp\!\!\!\perp Y \mid S, X, R=0$ cannot be checked from $\mathcal{D}_1$ alone (Prentice, 1989; Freedman et al., 1992); Athey et al. (2026) argue that high-dimensional, multi-modal surrogates – e.g., comprehensive job skills or multi-biomarker oncology panels – make surrogacy substantively more plausible. We adopt both assumptions without relaxation, but make their impact transparent: in Appendix D we derive a pointwise bias decomposition for the T-learner (a direct instantiation of the identification) under joint violation (in analogy with the ALTE analysis of Athey et al. (2026)), and Corollary D.2 turns this into a closed-form sensitivity bound computable from $\mathcal{D}_1$ data alone given a user-supplied bound on the residual direct effect.

[4] Athey et al. (2026) refers to h as the surrogate index.

4 T-, RA-, AND IPW-LEARNERS FOR HLTE ESTIMATION

In this section, we briefly introduce how the T-, RA-, and IPW-learners for the standard HTE estimation setting[5] can be adapted to the HLTE setting (see Appendix H.1 for details). Under the surrogacy model, h is an unbiased proxy of the unobserved Y in $\mathcal{D}_1$. As a result, we can simply replace Y by $h(S, X)$ in the standard HTE estimators to construct the analogous HLTE learners. While new in our setting, these learners serve as natural baselines and provide useful intuition for the challenges of HLTE estimation, and thus motivate our LT-O-learners. We summarize these learners in Table 1.

• **Long-term T-learner (LT-T-learner):** The simplest approach is to estimate the conditional response functions $\mu_a(x) = \mathbb{E}\left[h(S,x)|A=a, X=x, R=0\right],\ a \in \{0,1\}$. Then, we obtain the *plug-in* HLTE estimator: $\hat{\tau}_T^0(x) = \hat{\mu}_1(x) - \hat{\mu}_0(x)$. We name it the long-term T-learner following the standard HTE setting (Künzel et al., 2019).

Table 1: **Overview of meta-learners for HLTE estimation.** Our LT-O-learners are the only approaches that are both Neyman-orthogonal and robust to different types of low-overlap scenarios.

Meta-learner	Analogy in standard HTE	Neyman orthogonal	Robust to low treatment overlap	Robust to low outcome overlap
LT-T-learner	T-learner	✗	✗	✗
LT-RA-learner	RA-learner	✗	✗	✗
LT-IPW-learner	IPW-learner	✗	✗	✗
LT-O-DR-learner (*ours*, $\omega \equiv 1$)	DR-learner	✓	✗	✗
LT-O-TO-learner (*ours*)	R-learner	✓	✓	✗
LT-O-LO-learner (*ours*)	–	✓	✗	✓
LT-O-DO-learner (*ours*)	–	✓	✓	✓

• **Long-term RA-learner (LT-RA-learner):** The RA-learner follows a two-stage pseudo-outcome regression. It fits estimates $(\hat{h}, \hat{\mu}_1, \hat{\mu}_0)$ in the first stage to construct the regression-adjusted pseudo-outcome $\mathcal{T}_{\text{RA}} = A(\hat{h}(S,X) - \hat{\mu}_0(X)) + (1-A)(\hat{\mu}_1(X) - \hat{h}(S,X))$. In the second stage, it regresses the regression-adjusted pseudo-outcome on $\mathcal{D}_1$:

$$\mathcal{L}_{\text{RA}} = \mathbb{E}\left[\mathbb{1}(R=0)(\mathcal{T}_{\text{RA}} - g(X))^2\right]. \tag{7}$$

• **Long-term IPW-learner (LT-IPW-learner):** The IPW-learner uses inverse-propensity-based pseudo-outcome: $\mathcal{T}_{\text{IPW}} = \left(\frac{A}{\hat{\pi}(X)} - \frac{1-A}{1-\hat{\pi}(X)}\right)\hat{h}(S,X)$, and the second stage loss is defined as:

$$\mathcal{L}_{\text{IPW}} = \mathbb{E}\left[\mathbb{1}(R=0)(\mathcal{T}_{\text{IPW}} - g(X))^2\right] \tag{8}$$

Why do we need orthogonal learners: The LT-T-learner is subject to plug-in bias (Kennedy, 2022). While the LT-RA-learner and LT-IPW-learner mitigate plug-in bias with a two-stage procedure, they are still sensitive to the nuisance estimation errors $\hat{\eta} - \eta$. To overcome these limitations, current state-of-the-art two-stage methods for treatment effect estimation use the so-called *orthogonal learners* (Chernozhukov et al., 2018; Foster & Syrgkanis, 2023). A learner with loss function $\mathcal{L}(g, \eta)$ is defined as orthogonal if the Gateaux derivative of the loss w.r.t. the nuisance parameters vanishes at the truth:

$$D_\eta D_g \mathcal{L}(g, \eta)[\hat{g} - g, \hat{\eta} - \eta] = 0. \tag{9}$$

Intuitively, orthogonality ensures that the gradient of the loss function is insensitive to perturbations in the nuisance parameters (Chernozhukov et al., 2018). This allows faster convergence of the target estimator even with a slow convergence of the nuisances. We prove this property for our LT-O-learners in Theorem 5.2.

5 OUR LT-O-LEARNERS

Main idea of our LT-O-learners: Existing studies have shown that orthogonal learners without sample reweighting are usually hurt by low overlaps under finite samples (Fisher, 2023; Frauen et al., 2025a;b) (see Appendix F.3 for the variance analysis of the unweighted LT-O-DR-learner). Our goal is to design learners that are both orthogonal and robust against low-overlap regimes. To do this, the key idea is to downweight the samples with (i) low treatment or (ii) low long-term outcome overlap. Conceptually, this corresponds to optimizing a weighted oracle risk of the form

$$\mathcal{L}_\omega^{\text{oracle}}(g) = \mathbb{E}\big[\omega(\pi(X), \rho(X)) \cdot (\tau^0(X) - g(X))^2 \mid R = 0\big],$$

[5]This refers to the setting where there is only a single dataset that contains fully observed (X, A, Y) with no unobserved confounding (Künzel et al., 2019).

where $\omega(\pi(X), \rho(X)) > 0$ represents a *weighting function* that *retargets* the objective to downweight the samples with low overlap while preserving the target HLTE $\tau^0(x)$. A simple example of such weighting function is $\omega(X) = \pi(x)(1-\pi(x))\rho(x)$, where $\pi(x)(1-\pi(x))$ downweights samples with low treatment overlap and $\rho(x)$ downweights samples with low long-term outcome overlap.

However, the loss $\mathcal{L}_\omega^{\text{oracle}}$ is an oracle objective because it depends on the true HLTE $\tau^0(x)$, which is our target estimand and thus unobserved in practice. Hence, our goal is to construct a loss function $\mathcal{L}_\omega(g;\eta)$ that is equivalent to this oracle objective and can be optimized using observed data. This requires a non-trivial derivation to ensure that the resulting loss (i) shares the same minimizer as the oracle loss and (ii) is Neyman-orthogonal w.r.t. the nuisance functions.

Overview: In Section 5.1, we orthogonalize the oracle weighted loss and describe the training algorithm. In Section 5.2, we motivate the weighted instantiations through optimal efficiency bound. Finally, in Section 5.3, we establish the theoretical benefits of orthogonality and weighting.

5.1 Orthogonalizing the weighted oracle loss

In this section, we derive the orthogonal loss that has the same minimizer as the weighted oracle loss, under true nuisance η and model class $\mathcal{G}$. Following a similar approach as in (Morzywołek et al., 2023), we first derive the EIF for the ω-*weighted* average long-term treatment effect (WALTE), based on which we construct our loss (See Appendix C.3). We denote $\omega(x) = \omega(\pi(x), \rho(x))$.

LT-O-learners: Let Z be (X, A, S) when $R = 0$ and (X, S, Y) when $R = 1$. Based on the derived EIF, we propose our orthogonal loss as follows:

$$\mathcal{L}_\omega(g, \eta) = \mathbb{E}\left[\omega^*(Z;\eta)g(X)^2 - 2\mathcal{T}_{\text{LT}}(Z;\eta)\cdot g(X)\right], \tag{10}$$

where ω^* is the *LT weighting function*, defined as:

$$\omega^*(Z;\eta) = \mathbb{1}(R=0)\omega(X) + \Omega(Z;\eta), \text{ where } \Omega(Z;\eta) = \mathbb{1}(R=0)\frac{\partial\omega}{\partial\pi}(A-\pi) + (1-\rho)\frac{\partial\omega}{\partial\rho}(R-\rho), \tag{11}$$

and the *LT pseudo-outcome* $\mathcal{T}_{\text{LT}}(Z;\eta)$ is

$$\mathcal{T}_{\text{LT}}(Z;\eta) = \mathbb{1}(R=0)\omega(X)\hat{\tau}_{\text{AIPW}}(Z;\eta) + \mathbb{1}(R=1)\omega(X)\psi_{\text{obs}}(Z;\eta) + (\mu(1,X)-\mu(0,X))\Omega(Z;\eta), \tag{12}$$

$$\text{with } \hat{\tau}_{\text{AIPW}}(Z;\eta) = \mu(1,X) - \mu(0,X) + \frac{A-\pi(X)}{\pi(X)(1-\pi(X))}(h(S,X)-\mu(A,X)), \tag{13}$$

$$\psi_{\text{obs}} = \frac{1-\rho_s(S,X)}{\rho_s(S,X)}\left(\frac{\pi_s(S,X)-\pi(X)}{\pi(X)(1-\pi(X))}\right)\cdot\big(Y - h(S,X)\big) \tag{14}$$

Here, $\hat{\tau}_{\text{AIPW}}(Z;\eta)$ is the augmented IPW (AIPW) score, and $\psi_{\text{obs}}(Z;\eta)$ is the bias-correction term.

Intuitively, ω^* consists of a main part $\mathbb{1}(R=0)\omega(x)$ and a residual-based $\Omega(Z;\eta)$ and thus preserves the weight $\omega(x)$ on $\mathcal{D}_1$. Take $\omega = \pi^2(1-\pi)^2\rho$ as an example: When either (a) treatment overlap is low (small $\pi(1-\pi)$) or (b) long-term outcome overlap is low (small ρ), the sample is downweighted. In fact, we have the following (see Appendix F for detailed proofs):

$$\mathbb{E}\left[\omega^*(Z;\eta) \mid X\right] = (1-\rho(X))\omega(X), \quad \mathbb{E}\left[\mathcal{T}_{\text{LT}}(Z;\eta) \mid X\right] = (1-\rho(X))\omega(X)\tau^0(X).$$

Hence, at true nuisances, the unrestricted minimizer is $g^*(X) = \mathbb{E}[\mathcal{T}_{\text{LT}} \mid X]/\mathbb{E}[\omega^* \mid X] = \tau^0(X)$ almost surely. We further state the properties of the LT-O-learners loss in the following theorem.

Theorem 5.1 (Properties of LT-O-learners)**.** *Let η_0 and $\hat{\eta}$ denote the true and estimated nuisances, respectively. Then under regular assumptions (see Appendix F.1):*

1. ***Neyman orthogonality.*** *For arbitrary $g \in \mathcal{G}$ and direction v, $D_\eta D_g \mathcal{L}_\omega(g, \eta_0)[v, \hat{\eta} - \eta_0] = 0$.*
2. ***Oracle equivalence.*** *We have $\arg\min_{g\in\mathcal{G}} \mathcal{L}_\omega(g,\eta_0) = \arg\min_{g\in\mathcal{G}} \mathcal{L}_\omega^{\text{oracle}}(g)$. Consequently, both losses have the same minimizers over $\mathcal{G}$; when $\tau^0 \in \mathcal{G}$, they recover τ^0 almost surely.*

3. ***Empirical well-posedness.*** *Denote the empirical, cross-fitted loss as $\mathcal{L}_{\omega,N}(g,\hat{\eta})$, then we have:*

$$\inf_{g\in\mathcal{G}} \mathcal{L}_{\omega,N}(g,\hat{\eta}) \geq -M, \quad \sup_{g\in\mathcal{G}} |\mathcal{L}_{\omega,N}(g,\hat{\eta}) - \mathcal{L}_{\omega}(g,\eta_0)| \xrightarrow{P} 0, \quad \textit{with } \mathcal{L}_{\omega}(g,\eta_0) \textit{ strongly convex.}$$

Proof. See Appendix F.4, F.5, and F.6 for the respective proofs of properties 1, 2, and 3. □

Point 1 and point 2 together establish that $\mathcal{L}_{\omega}(g,\eta)$ is indeed the orthogonalized version of the oracle weighted loss $\mathcal{L}_{\omega}^{\text{oracle}}(g)$. We point out that the orthogonalized weight $\omega^*(Z;\eta)$ may be negative; hence, the empirical loss may not be convex under finite sample. Point 3 ensures that the empirical loss always has a minimizer and uniformly converges to a strongly convex limit. As a result, the empirical objective is still well-posed and our orthogonalization is not a problem for optimization.

Note on why signed weights arise. Negative weights arise from orthogonalization, which reflects an inherent trade-off between bias and convexity intrinsic to long-term estimation. This reveals a broader phenomenon in orthogonalized overlap-weighted estimators that previous work has not made explicit (see Appendix C.4).

In practice, we minimize an empirical surrogate of $\mathcal{L}_{\omega}(g,\hat{\eta})$ via K-fold cross-fitting (Chernozhukov et al., 2018; Kennedy, 2023; Morzywołek et al., 2023; Foster & Syrgkanis, 2023): we partition $\mathcal{D} = \{Z_i\}_{i=1}^N$ into $K = 5$ folds; for each fold, we fit the nuisances $\hat{\eta} = (\hat{\pi}, \hat{\pi}_s, \hat{\rho}, \hat{\rho}_s, \hat{h}, \hat{\mu})$ on the remaining $K-1$ folds to evaluate $\hat{\omega}_i^* = \omega^*(Z_i;\hat{\eta})$ and $\hat{\mathcal{T}}_i = \mathcal{T}_{\text{LT}}(Z_i;\hat{\eta})$, and then solve

$$\hat{g} = \arg\min_{g\in\mathcal{G}} \sum_{i=1}^{N} \hat{\omega}_i^* g(X_i)^2 - 2\hat{\mathcal{T}}_i g(X_i). \tag{15}$$

5.2 Instantiations

Is there an optimal ω? We derive the optimal weighting function ω_{opt} that minimizes the efficiency bound for the sample version of the WALTE in Eq. (32), by treating the weight as fixed (see Appendix J). However, $\omega_{\text{opt}}(x)$ involves conditional expectations of additional nuisances, so directly plugging it into our loss would require further orthogonal corrections. To obtain a tractable weighting function, we derive a closed-form upper bound to ω_{opt} (see Appendix J.3):

$$\omega_{\text{opt}}(x) \lesssim \left(\frac{C_t(x)^2}{\pi(x)(1-\pi(x))} + \frac{C(x)^2}{\rho(x)\,\pi(x)^2(1-\pi(x))^2} \right)^{-1}. \tag{16}$$

From Eq. (16), we can see that the variance depends on two sources of overlap: (1) treatment overlap, through $\pi(x)(1-\pi(x))$ and $\pi(x)^2(1-\pi(x))^2$; and (2) outcome overlap, through $\rho(x)$. This naturally motivates the choices of ω summarized below (see Appendix E for full derivations).

- ▲ **LT-O-DR-learner** ($\omega \equiv 1$): no overlap weighting; recovers the EIF of ALTE (Athey et al., 2026; Chen & Ritzwoller, 2023).
- ✚ **LT-O-LO-learner** ($\omega = \rho$): targets low long-term outcome overlap.
- ★ **LT-O-TO-learner** ($\omega = \pi^2(1-\pi)^2$): targets low treatment overlap.
- ■ **LT-O-DO-learner** ($\omega = \pi^2(1-\pi)^2\rho$): targets joint treatment and outcome overlap.

Problems of the LT-O-DR-learner: Although orthogonal, the LT-O-DR-learner does not apply any sample-weighting strategy, which leads to large finite-sample variance (Fisher, 2023). We analyze the variance of the DR pseudo-outcome (defined in Appendix E.1) and show why it inflates under low overlap in Appendix F.3. As a result, the LT-O-DR-learner is not robust to low overlap, motivating the weighted instantiations.

Note: We point out that ω_{opt} is only optimal in the sense of minimizing the efficiency bound, *not* the HLTE estimation error. This motivates a class of learners with customizable weights. In practice, we recommend choosing a suitable weight based on the estimated overlap scores $\hat{\pi}$ and $\hat{\rho}$.

5.3 THEORETICAL BENEFITS

In this section, we show the benefits of orthogonality and weighting with two theorems:

Theorem 5.2 (Error bound[6])**.** *Let $\hat{g}$ be the minimizer of the empirical loss $\mathcal{L}_{\omega,n}(g,\hat{\eta})$ and g_0 be the minimizer of $\mathcal{L}_\omega(g,\eta)$ under the true nuisance. Under the assumptions in Appendix F.1, $\|\hat{g}-g_0\|^2 \lesssim R_g + \|\hat{\mu}-\mu\|_4^2(\|\hat{\rho}-\rho\|_4^2+\|\hat{\pi}-\pi\|_4^2) + \|\hat{h}-h\|_4^2(\|\hat{\pi}_s-\pi_s\|_4^2+\|\hat{\rho}_s-\rho_s\|_4^2) + \|\hat{\pi}-\pi\|_4^4 + \|\hat{\rho}-\rho\|_4^4$, where $R_g = \mathcal{L}_\omega(\hat{g},\hat{\eta}) - \mathcal{L}_\omega(g_0,\hat{\eta})$ is the quasi-oracle rate.*

Furthermore, under a linear class $\mathcal{G}=\{g(x)=\theta^\top x : \theta\in\mathbb{R}^d\}$ with $\tau^0\in\mathcal{G}$, $\hat{\theta}$ admits a closed-form solution when the empirical curvature matrix is positive definite (Appendix K). We conduct a novel asymptotic variance analysis for the orthogonal learner to show the stabilization below.

Theorem 5.3 (Variance stabilization)**.** *Under standard Z-estimator regularity conditions and Assumption K.8:* (a) $\sqrt{n}(\hat{\theta}-\theta_0) \xrightarrow{d} \mathcal{N}\big(0,\,\mathcal{V}(\omega)\big)$*, where $\mathcal{V}(\omega)$ is the asymptotic covariance matrix of $\hat{\theta}$ (see Appendix K);* (b) *for any sequence of DGPs with overlap score $0<\epsilon\to 0$, $v^\top\mathcal{V}(\mathbf{1})\,v\to\infty$ for every direction $v\neq 0$, while $v^\top\mathcal{V}\big(\pi(1-\pi)\rho\big)\,v$ stays bounded.*

Proof. See Appendix F.7 and K.3 for proofs of Theorems 5.2 and 5.3, respectively. □

Theorem 5.2 shows that, due to Neyman-orthogonality, nuisance error enters only through higher-order interaction terms, so the HLTE estimator preserves its convergence rate even when nuisances are fit by flexible machine learning at slow nonparametric rates (Chernozhukov et al., 2018; Foster & Syrgkanis, 2023). Theorem 5.3 then shows that, for linear target models, the weight $\pi(1-\pi)\rho$ suffices to absorb the inverse-overlap blow-up of unweighted LT-O-DR-learner's variance; stability of the LT-O-DO-learner with a stronger weight $\pi^2(1-\pi)^2\rho$ follows naturally.

6 EXPERIMENTS

In this section, we conduct experiments to empirically validate our theoretical results and to demonstrate the benefits of our proposed LT-O-learners in low-overlap settings. The experiments serve two main purposes: (1) *to show that the limited overlap is a primary source of estimation instability under finite samples in standard learners*; and (2) *to demonstrate that our proposed learner addresses this challenge through Neyman-orthogonality and our overlap weighting*.

Table 2: **Synthetic experiments.** Mean ± std. dev. of the PEHE [↓] across 5 runs. Improv. denotes the percentage improvement of the best LT-O-learner over the best baseline.

Overlap type	$\mathcal{D}_*$	$\mathcal{D}_\mathbf{t}$	$\mathcal{D}_\mathbf{o}$	$\mathcal{D}_{\mathbf{t}+\mathbf{o}}$
LT-T-learner	0.22 ± 0.02	0.43 ± 0.10	0.20 ± 0.02	0.26 ± 0.03
LT-RA-learner	0.21 ± 0.02	0.56 ± 0.07	0.18 ± 0.03	0.67 ± 0.08
LT-IPW-learner	0.23 ± 0.03	0.70 ± 0.23	0.19 ± 0.04	1.27 ± 0.78
LT-O-DR-learner (*ours*)	0.03 ± 0.01	0.60 ± 0.61	0.23 ± 0.17	1.08 ± 0.55
LT-O-TO-learner (*ours*)	$\mathbf{0.03 \pm 0.01}$	0.09 ± 0.01	0.29 ± 0.29	0.23 ± 0.18
LT-O-LO-learner (*ours*)	0.05 ± 0.03	0.45 ± 0.16	0.08 ± 0.02	0.39 ± 0.19
LT-O-DO-learner (*ours*)	0.03 ± 0.01	$\mathbf{0.07 \pm 0.02}$	$\mathbf{0.07 \pm 0.01}$	$\mathbf{0.10 \pm 0.02}$
Improv.	84.9%	84.1%	62.2%	61.2%

Baselines: We evaluate the four instantiations of our LT-O-learners (see Table 1). There is no direct benchmark specifically designed for our HLTE setting. Therefore, as baselines, we derive and adapt standard meta-learners, namely, a LT-T-learner, LT-RA-learner, and LT-IPW-learner to our setting (see Section 4; details in Appendix H). The LT-O-DR-learner is the unweighted special case of our LT-O-learners; we report it alongside the weighted variants but do not consider it part of our core contribution. In our experiments, all meta-learners are instantiated using the same neural network architecture and identical hyperparameters (see Appendix I), thus ensuring that performance differences reflect the effectiveness of the respective adjustment strategies rather than model architecture. We measure the *precision in estimation of heterogeneous effects (PEHE)*. Each experiment is repeated 5 times with different random seeds.

Experiments with synthetic data: We evaluate the performances across multiple synthetic datasets. We simulate 4 scenarios: $\mathcal{D}_*, \mathcal{D}_t, \mathcal{D}_o, \mathcal{D}_{t+o}$, which correspond to no overlap issue, low treatment overlap only, low long-term outcome overlap only, and mixed overlap issue (see Appendix G.1 for details on the data-generating process). *Results:* The results are in Table 2. We find: **(1)** As

[6]The proof adapts the technique of Morzywołek et al. (2023) to the long-term surrogate setting.

suggested by our theory, our LT-O-learners outperform all baselines on $\mathcal{D}_t, \mathcal{D}_o, \mathcal{D}_{t+o}$, with 84.1%, 62.2%, and 61.2% improvements over the best baseline. Notably, the LT-O-DR-learner performs well on $\mathcal{D}_*$ but worse than baselines under overlap regimes. **(2)** The LT-O-DO-learner dominates every low overlap regime, validating the joint weight $\omega = \pi^2(1-\pi)^2\rho$ derived from the variance lower bound (Appendix J); the LT-O-TO-learner and LT-O-LO-learner remain competitive within their target regimes. **(3)** Overall, our LT-O-learners, based on overlap weighting, consistently outperform baselines, highlighting the benefits of Neyman-orthogonality and sample weighting in low-overlap settings. In Appendix B.2, we further report PEHE stratified by overlap level to confirm that our LT-O-learners do not trade accuracy on low-overlap samples for overall stability. We also conduct bias-variance decompostion and empirical coverage analysis in Appendix B.3, and B.4, respectively.

• **Overlap-induced high variance:** We vary the parameter $(\gamma_\pi, \gamma_\rho) = (\gamma, 0.5\gamma)$ to gradually worsen the overlap in the data-generating process of $\mathcal{D}_{t+o}(\gamma)$. For each γ, we repeat the experiment using Monte Carlo simulation for 5 runs and then calculate the variance of the resulting estimators, i.e., $V(\gamma) = \mathbb{E}_X\left[\mathrm{Var}_{\mathcal{D}}(\hat{\tau}(X)) \mid \text{overlap} = \gamma\right]$. *Results:* Figure 2 shows that the variance of the LT-O-DO-learner remains low even under very-low-overlap regions, indicating it is more stable compared to the LT-O-DR-learner. This result confirms that our novel overlap weighting strategy is effective in reducing the variance of the pseudo-outcomes.

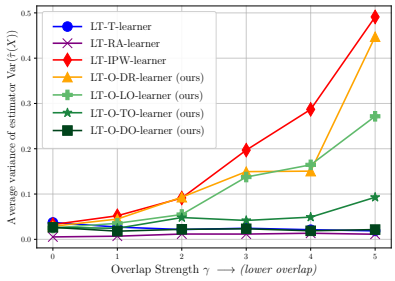


Figure 2: **Variance of estimator across different overlap scenarios (γ):** Low overlap leads to a high finite-sample variance of the DR-learner, while our LT-O-learners maintain a low variance.

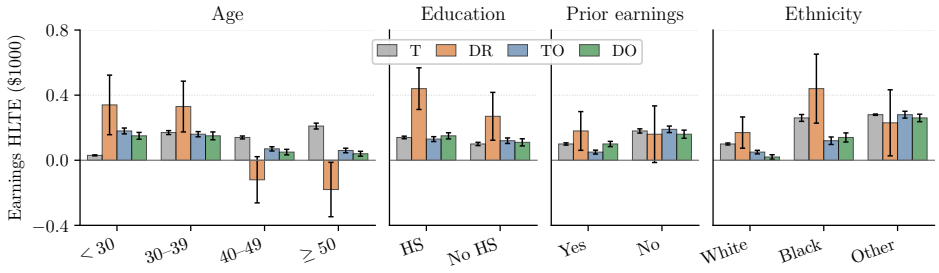


Figure 3: **Subgroup Heterogeneity on GAIN+PSID.** Calibrated subgroup effects (in 1000 USD) with std on earnings over quarters 41–48, averaged across 10 seeds. Our LT-O-DO-learner and LT-O-TO-learner yield HLTEs that are more stable than LT-O-DR-learner and align with domain knowledge. (HS: Highschool)

Experiments with real-world data: We estimate the long-term earnings effects of a California welfare-to-work program (GAIN). GAIN was a randomized experiment with 19,170 participants (Hotz et al., 2006); long-term outcomes are observed through the Panel Study of Income Dynamics (PSID; 6,859 participants), a long-running observational household panel (McGonagle et al., 2012; Pfeffer et al., 2025; Panel Study of Income Dynamics, 2026). GAIN provides baseline covariates X, randomized assignment A, and shorter-run earnings S. PSID provides comparable (X, S) together with later earnings Y in quarters 41–48, beyond the follow-up in GAIN. We observe good overlap in year-9 earnings, which supports comparability across the two samples (Appendix G.2.3). *Results:* We compare subgroup-level HLTE estimates and the std of the LT-O-DO-learner and LT-O-TO-learner against LT-T-learner and LT-O-DR-learner in Figure 3. Across all four subgroups, our LT-O-DO-learner and LT-O-TO-learner achieves stable estimates and recover known heterogeneity patterns of the GAIN program (Hotz et al., 2006): *training effects decline sharply with age, are slightly larger for high-school graduates and individuals without prior earnings, and vary substantially across ethnic subgroups.*

Conclusion: We develop a novel class of orthogonal meta-learners for estimating HLTEs. Through theorectical analysis and empirical experiments, we demonstrate that our method achieves superior robustness in low-overlap regimes.

AI Use Statement: We used generative AI tools to generate synthetic datasets; clean, process, and explore the experimental data; assist in writing and checking mathematical proofs; provide feedback on our methodology and experiments; and implement the proposed method. We did not use generative AI tools to develop the theoretical models, the conceptual framework, or the critical ingredients of the mathematical proofs, nor to formulate mathematical claims, design the research methodology and experiments, or interpret results. Hypothesis generation and translation assistance are not applicable to this work. All AI-assisted code, proofs, and analyses were reviewed and verified by the authors. We take responsibility for the final content of this work, including text, claims, and artifacts produced with the aid of generative AI.

Ethics Statement: This work is methodological and aims to improve the stability of HLTE estimation under limited overlap. We do not foresee direct harms or negative societal impacts arising from this work.

Reproducibility Statement: We release the code and describe the datasets used in our experiments to support reproducibility.

REFERENCES

Susan Athey, Raj Chetty, and Guido W. Imbens. Combining experimental and observational data to estimate treatment effects on long term outcomes. *arXiv preprint arXiv:2006.09676*, 2020.

Susan Athey, Raj Chetty, Guido W. Imbens, and Hyunseung Kang. The surrogate index: Combining short-term proxies to estimate long-term treatment effects more rapidly and precisely. *The Review of Economic Studies*, 93(4):2284–2312, 2026. doi: 10.1093/restud/rdaf087.

Lisa Belin, Aidan Tan, Yann De Rycke, and Agnès Dechartres. Progression-free survival as a surrogate for overall survival in oncology trials: a methodological systematic review. *British Journal of Cancer*, 122(11):1707–1714, 2020. doi: 10.1038/s41416-020-0805-y.

Peter J. Bickel, Chris A. J. Klaassen, Ya'acov Ritov, and Jon A. Wellner. *Efficient and Adaptive Estimation for Semiparametric Models*. Johns Hopkins University Press, 1993.

C. L. Brantner, T. H. Chang, T. Q. Nguyen, H. Hong, Leon Di Stefano, and E. A. Stuart. Methods for integrating trials and non-experimental data to examine treatment effect heterogeneity. *Statistical Science*, 38(4):640–654, Nov 2023. Epub 2023 Nov 6.

C. L. Brantner, T. Q. Nguyen, T. Tang, C. Zhao, H. Hong, and E. A. Stuart. Comparison of methods that combine multiple randomized trials to estimate heterogeneous treatment effects. *Statistics in Medicine*, 43(7):1291–1314, 2024.

Ruichu Cai, Weilin Chen, Zeqin Yang, Shu Wan, Chen Zheng, Xiaoqing Yang, and Jiecheng Guo. Long-term causal effects estimation via latent surrogates representation learning. *Neural Networks*, 176:106336, 2024.

Ruichu Cai, Junjie Wan, Weilin Chen, Zeqin Yang, Zijian Li, Peng Zhen, and Jiecheng Guo. Long-term individual causal effect estimation via identifiable latent representation learning. In *IJCAI*, pp. 4788–4796, 2025.

Jiafeng Chen and David M. Ritzwoller. Semiparametric estimation of long-term treatment effects. *Journal of Econometrics*, 237(2):105545, 2023.

Weilin Chen, Ruichu Cai, Junjie Wan, Zeqin Yang, and José Miguel Hernández-Lobato. Nonparametric heterogeneous long-term causal effect estimation via data combination. *arXiv preprint arXiv:2502.18960*, 2025a.

Weilin Chen, Ruichu Cai, Yuguang Yan, Zhifeng Hao, and José Miguel Hernández-Lobato. Long-term causal inference via modeling sequential latent confounding. *arXiv preprint arXiv:2502.18994*, 2025b.

Victor Chernozhukov, Denis Chetverikov, Mert Demirer, Esther Duflo, Christian Hansen, Whitney Newey, and James Robins. Double/debiased machine learning for treatment and structural parameters. *The Econometrics Journal*, 21(1):C1–C68, 2018.

Bénédicte Colnet, Imke Mayer, Guanhua Chen, Awa Dieng, Ruohong Li, Gaël Varoquaux, Jean-Philippe Vert, Julie Josse, and Shu Yang. Causal inference methods for combining randomized trials and observational studies: A review. *Statistical Science*, 39(1):165–191, Feb 2024.

Richard K. Crump, V. Joseph Hotz, Guido W. Imbens, and Oscar A. Mitnik. Dealing with limited overlap in estimation of average treatment effects. *Biometrika*, 96(1):187–199, 2009.

Alicia Curth and Mihaela van der Schaar. Nonparametric estimation of heterogeneous treatment effects: From theory to learning algorithms. In *AISTATS*, volume 130 of *Proceedings of Machine Learning Research*, pp. 1810–1818. PMLR, 2021. URL https://proceedings.mlr.press/v130/curth21a.html.

Thomas P. A. Debray, Karel G. M. Moons, Gert van Valkenhoef, Orestis Efthimiou, Noemi Hummel, Rolf H. H. Groenwold, Johannes B. Reitsma, and on behalf of the GetReal methods review group. Get real in individual participant data (IPD) meta-analysis: a review of the methodology. *Research Synthesis Methods*, 6(4):293–309, Dec 2015. ISSN 1759-2879.

Stefan Feuerriegel, Dennis Frauen, Valentyn Melnychuk, Jonas Schweisthal, Konstantin Hess, Alicia Curth, Stefan Bauer, Niki Kilbertus, Isaac S. Kohane, and Mihaela van der Schaar. Causal machine learning for predicting treatment outcomes. *Nature Medicine*, 30:958–968, 2024.

Aaron Fisher. The connection between r-learning and inverse-variance weighting for estimation of heterogeneous treatment effects. *arXiv preprint arXiv:2307.09700*, 2023.

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

Constantine E. Frangakis and Donald B. Rubin. Principal stratification in causal inference. *Biometrics*, 58(1):21–29, 2002.

Dennis Frauen, Konstantin Hess, and Stefan Feuerriegel. Model-agnostic meta-learners for estimating heterogeneous treatment effects over time. In *International Conference on Learning Representations (ICLR)*, 2025a.

Dennis Frauen, Maresa Schröder, Konstantin Hess, and Stefan Feuerriegel. Orthogonal survival learners for estimating heterogeneous treatment effects from time-to-event data. In *NeurIPS*, 2025b.

Laurence S. Freedman, Barry I. Graubard, and Arthur Schatzkin. Statistical validation of intermediate endpoints for chronic diseases. *Statistics in Medicine*, 11(2):167–178, 1992.

AmirEmad Ghassami, Chang Liu, Alan Yang, David Richardson, Ilya Shpitser, and Eric Tchetgen Tchetgen. Combining experimental and observational data for identification and estimation of long-term causal effects. *arXiv preprint arXiv:2201.10743*, 2022.

Graham Van Goffrier, Lucas Maystre, and Ciarán Mark Gilligan-Lee. Estimating long-term causal effects from short-term experiments and long-term observational data with unobserved confounding. In *Proceedings of the Second Conference on Causal Learning and Reasoning*, volume 213 of *Proceedings of Machine Learning Research*, pp. 791–813. PMLR, August 2023. URL https://proceedings.mlr.press/v213/goffrier23a.html.

Frank R. Hampel. The influence curve and its role in robust estimation. *Journal of the American Statistical Association*, 69(346):383–393, 1974.

Tobias Hatt, Jeroen Berrevoets, Alicia Curth, Stefan Feuerriegel, and Mihaela van der Schaar. Combining observational and randomized data for estimating heterogeneous treatment effects. *arXiv preprint arXiv:2202.12891*, 2022.

Konstantin Hess, Dennis Frauen, Mihaela van der Schaar, and Stefan Feuerriegel. Overlap-weighted orthogonal meta-learner for treatment effect estimation over time. In *ICLR*, 2026.

V. Joseph Hotz, Guido W. Imbens, and Jacob A. Klerman. Evaluating the differential effects of alternative welfare-to-work training components: A reanalysis of the california gain program. *Journal of Labor Economics*, 24(3):521–566, 2006.

Shan Huang, Chen Wang, Yuan Yuan, Jinglong Zhao, and Brocco (Jingjing) Zhang. Estimating effects of long-term treatments. *Management Science*, 2026. doi: 10.1287/mnsc.2023.02575. Articles in Advance.

Ta-Wei Huang and Eva Ascarza. Doing more with less: Overcoming ineffective long-term targeting using short-term signals. *Marketing Science*, 43(4):863–884, 2024.

Guido Imbens, Nathan Kallus, Xiaojie Mao, and Yuhao Wang. Long-term causal inference under persistent confounding via data combination. *Journal of the Royal Statistical Society Series B: Statistical Methodology*, 87(2):362–388, 2025.

Guido W. Imbens and Joshua D. Angrist. Identification and estimation of local average treatment effects. *Econometrica*, 62(2):467–475, 1994.

Andrew Jesson, Sören Mindermann, Uri Shalit, and Yarin Gal. Identifying causal-effect inference failure with uncertainty-aware models. In *NeurIPS*, 2020.

Nathan Kallus and Xiaojie Mao. On the role of surrogates in the efficient estimation of treatment effects with limited outcome data. *Journal of the Royal Statistical Society Series B: Statistical Methodology*, 87(2):480–509, 2025.

Edward H. Kennedy. Semiparametric doubly robust targeted double machine learning: a review. *arXiv preprint arXiv:2203.06469*, 2022.

Edward H. Kennedy. Towards optimal doubly robust estimation of heterogeneous causal effects. *Electronic Journal of Statistics*, 17(2):3008–3049, 2023. doi: 10.1214/23-EJS2157.

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

Steffen L. Lauritzen, Odd O. Aalen, Donald B. Rubin, and Elja Arjas. Discussion on causality [with reply]. *Scandinavian Journal of Statistics*, 31(2):189–201, 2004.

Thomas M. McDonald, Lucas Maystre, Mounia Lalmas, Daniel Russo, and Kamil Ciosek. Impatient bandits: Optimizing recommendations for the long-term without delay. In *KDD*, 2023.

Katherine A. McGonagle, Robert F. Schoeni, Narayan Sastry, and Vicki A. Freedman. The Panel Study of Income Dynamics: Overview, recent innovations, and potential for life course research. *Longitudinal and Life Course Studies*, 3(2):268–284, 2012.

Paweł Morzywołek, Johan Decruyenaere, and Stijn Vansteelandt. On a general class of orthogonal learners for the estimation of heterogeneous treatment effects. *arXiv preprint arXiv:2303.12687*, 2023. URL `https://arxiv.org/abs/2303.12687v1`.

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

Filip Obradović. Identification of long-term treatment effects via temporal links, observational, and experimental data. *arXiv preprint arXiv:2411.04380*, 2024.

Panel Study of Income Dynamics. Panel Study of Income Dynamics, public-use dataset. Produced and distributed by the Survey Research Center, Institute for Social Research, University of Michigan, Ann Arbor, MI, 2026. URL `https://psidonline.isr.umich.edu/`.

Harsh Parikh, Marco Morucci, Vittorio Orlandi, Sudeepa Roy, Cynthia Rudin, and Alexander Volfovsky. A double machine learning approach for combining experimental and observational studies. *Observational Studies*, 11(3):249–300, 2025.

Yechan Park and Yuya Sasaki. A bracketing relationship for long-term policy evaluation with combined experimental and observational data. *arXiv preprint arXiv:2401.12050*, 2024.

Fabian T. Pfeffer, Davis Daumler, and Esther Friedman. PSID-SHELF, 1968–2021: The PSID's social, health, and economic longitudinal file (PSID-SHELF), beta release. Inter-university Consortium for Political and Social Research [distributor], Ann Arbor, MI, 2025. URL https://doi.org/10.3886/E194322. Last modified 2025-02-24.

Ross L. Prentice. Surrogate endpoints in clinical trials: Definition and operational criteria. *Statistics in Medicine*, 8(4):431–440, 1989.

Donald Rubin. Estimating causal effects of treatments in experimental and observational studies. *ETS Research Bulletin Series*, 1972(2):i–31, 1972. ISSN 2333-8504. doi: 10.1002/j.2333-8504.1972.tb00631.x.

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

Yuta Saito, Himan Abdollahpouri, Jesse Anderton, Ben Carterette, and Mounia Lalmas. Long-term off-policy evaluation and learning. In *Proceedings of the ACM Web Conference 2024*, WWW '24, pp. 3432–3443. ACM, 2024. doi: 10.1145/3589334.3645446.

Uri Shalit, Fredrik D. Johansson, and David Sontag. Estimating individual treatment effect: generalization bounds and algorithms. In *ICML*, volume 70 of *PMLR*, pp. 3076–3085, 06–11 Aug 2017.

Rahul Singh and Hannah Zhou. Kernel methods for long term dose response curves. *arXiv preprint arXiv:2201.05139*, 2022.

Xiaoqing Tan, Chung-Chou H. Chang, Ling Zhou, and Lu Tang. A tree-based model averaging approach for personalized treatment effect estimation from heterogeneous data sources. In *ICML*, volume 162 of *Proceedings of Machine Learning Research*, pp. 21013–21036. PMLR, 2022. URL https://proceedings.mlr.press/v162/tan22a.html.

Anastasios A. Tsiatis. *Semiparametric Theory and Missing Data*. Springer, 2006.

A. W. van der Vaart. *Asymptotic Statistics*. Cambridge University Press, 1998.

Yuxin Wang, Maresa Schröder, Dennis Frauen, Jonas Schweisthal, Konstantin Hess, and Stefan Feuerriegel. Constructing confidence intervals for average treatment effects from multiple datasets. In *ICLR*, 2025.

Yuyan Wang, Mohit Sharma, Can Xu, Sriraj Badam, Qian Sun, Lee Richardson, Lisa Mijung Chung, Ed H. Chi, and Minmin Chen. Surrogate for long-term user experience in recommender systems. In *KDD*, pp. 4100–4109. ACM, 2022.

Lili Wu and Shu Yang. Integrative r-learner of heterogeneous treatment effects combining experimental and observational studies. In Bernhard Schölkopf, Caroline Uhler, and Kun Zhang (eds.), *Proceedings of the First Conference on Causal Learning and Reasoning*, volume 177 of *Proceedings of Machine Learning Research*, pp. 904–926. PMLR, 11–13 Apr 2022.

Jeremy Yang, Dean Eckles, Paramveer Dhillon, and Sinan Aral. Targeting for long-term outcomes. *Management Science*, 70(6):3841–3855, 2024a. Publisher: INFORMS.

Qinwei Yang, Xueqing Liu, Yan Zeng, Ruocheng Guo, Yang Liu, and Peng Wu. Learning the optimal policy for balancing short-term and long-term rewards. In *NeurIPS*, 2024b.

Shu Yang and Peng Ding. Combining multiple observational data sources to estimate causal effects. *Journal of the American Statistical Association*, 115(531):1540–1554, 2020.

Zeqin Yang, Weilin Chen, Ruichu Cai, Yuguang Yan, Zhifeng Hao, Zhipeng Yu, Zhichao Zou, Jixing Xu, Zhen Peng, and Jiecheng Guo. Estimating long-term heterogeneous dose-response curve: generalization bound leveraging optimal transport weights. *arXiv preprint arXiv:2406.19195*, 2025.

Dong Yao, Caizhi Tang, Qing Cui, and Longfei Li. Combining incomplete observational and randomized data for heterogeneous treatment effects. In *CIKM*, 2024.

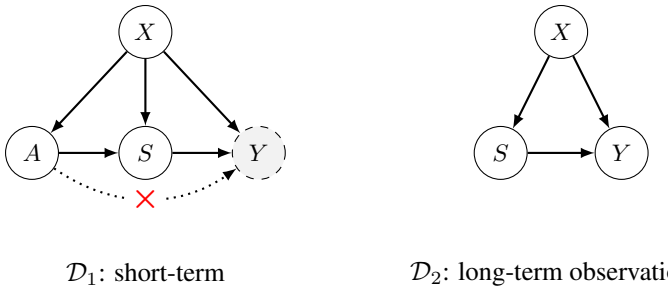
A Extended related work

A.1 A broader overview of related work

Data combination for causal inference: A large body of work studies treatment effect estimation by combining multiple datasets, such as combinations of observational datasets (e.g., Yang & Ding, 2020; Wang et al., 2025), combinations of experimental datasets (e.g., Debray et al., 2015; Tan et al., 2022; Brantner et al., 2024), or mixed combinations of both observational *and* experimental datasets (e.g., Hatt et al., 2022; Wu & Yang, 2022; Brantner et al., 2023; Yao et al., 2024; Wang et al., 2025; Parikh et al., 2025; Chen et al., 2025a; Cai et al., 2025; Chen et al., 2025b) , including for both average and heterogeneous treatment effects. For a detailed review, we refer to Colnet et al. (2024). In these settings, the motivation for data combination is typically to leverage complementary information across datasets while assuming that treatment *and* outcomes are observed in *all* datasets. Hence, the works from this stream are different from our paper and *not* applicable. One paper aims to relax the setting by assuming a combination where long-term outcomes are observed in only one dataset while treatments must still be observed in both (Cai et al., 2025); however, this is still *different* from our setting, where treatments are observed in only one dataset and long-term outcomes in the other dataset.

A.2 Comparison of settings for long-term causal effect estimation

The literature on long-term causal effect estimation differs along three axes: (i) which variables are observed in the short-term ($\mathcal{D}_1$) and long-term ($\mathcal{D}_2$) datasets, (ii) the identification mechanism used to bridge S and Y, and (iii) the target estimand. Figure 4 depicts the canonical surrogacy setting, and Table 3 summarizes representative methods grouped by their high-level identification setting: latent unconfoundedness, equi-confounding, proximal data fusion, and surrogacy. Within the surrogacy setting in which A is unobserved in $\mathcal{D}_2$, ours is the only method targeting the heterogeneous long-term treatment effect while simultaneously achieving Neyman orthogonality and robustness to low overlap.



$\mathcal{D}_1$: short-term $\mathcal{D}_2$: long-term observational

Figure 4: **Causal diagram for the canonical surrogacy setting.** The short-term dataset $\mathcal{D}_1$ observes (X, A, S) while Y is unobserved (dashed node); the long-term observational dataset $\mathcal{D}_2$ observes (X, S, Y) with treatment A absent. The dotted arrow $A \to Y$ marked with $\times$ indicates that the surrogacy assumption rules out a direct effect of A on Y given (S, X).

Table 3: **Comparison of settings for long-term causal effect estimation.** Methods are grouped by their high-level identification setting. ✓/✓ indicates that a variable is observed; for the treatment columns, ✓ denotes binary/categorical treatment and ✓ denotes continuous treatment. ✗ indicates that the variable is not observed. ✓/✗ in the Benefits columns indicate whether each method satisfies the corresponding property.

Setting	**Paper**	**Short-term $\mathcal{D}_1$**			**Long-term $\mathcal{D}_2$**				**Identification through**	**Target**	**Benefits**	
		X	A	S	X	A	S	Y			Neyman orthogonality	Robustness to low overlap
Latent unconfoundedness	Chen & Ritzwoller (2023) (latent unconf.)	✓	✓	✓	✓	✓	✓	✓	Latent unconfoundedness	ALTE	✓	✗
	Yang et al. (2025)	✓	✓	✓	✓	✓	✓	✓	Latent unconfoundedness	HLTE (dose-response)	✗	✗
Equi-confounding	Ghassami et al. (2022)	✓	✓	✓	✓	✓	✓	✓	Additive equi-conf.†	ALTE, ATT	✓	✗
	Chen et al. (2025a)	✓	✓	✓	✓	✓	✓	✓	CAECB	HLTE	✓	✗
Proximal	Imbens et al. (2025)	✓	✓	✓	✓	✓	✓	✓	Outcome + selection bridge functions	ALTE	✓	✗
Surrogacy (canonical setting)	Athey et al. (2026)	✓	✓	✓	✓	✗	✓	✓	Surrogacy + comparability	ALTE	✓	✗
	Chen & Ritzwoller (2023) (stat. surrogacy)	✓	✓	✓	✓	✗	✓	✓	Surrogacy + comparability	ALTE	✓	✗
	Singh & Zhou (2022)	✓	✓	✓	✓	✗	✓	✓	Surrogacy + comparability (RKHS)	Dose-response curve	✗	✗
	Yang et al. (2024a)	✓	✓	✓	✓	✗	✓	✓	Surrogacy + comparability	Policy learning	✓	✗
	Ours	✓	✓	✓	✓	✗	✓	✓	**Surrogacy + comparability**	**HLTE**	✓	✓

†Ghassami et al. (2022) additionally propose bespoke-IV and single-proxy proximal variants under the same data setting.

B ADDITIONAL EXPERIMENTS

B.1 ABLATIONS W.R.T. NEYMAN ORTHOGONALITY

In this experiment, we analyze why Neyman-orthogonality is crucial for robustness. We use the synthetic $\mathcal{D}_{t+o}$ dataset (mixed overlap) and vary the sample size, where a lower sample size implies that the nuisance function estimation is more difficult. In addition to the existing baselines, we add two ablations: **weighted but non-orthogonal** versions of the LT-RA-learner and LT-DR-learner (see Appendix H for details). These ablations also employ the same overlap weighting $\omega = \pi^2(1-\pi)^2\rho$ as the LT-O-DO-learner, allowing us thus to isolate the effect of Neyman-orthogonality.

Results: The left panel of Figure 5 shows that the orthogonal LT-O-DO-learner consistently outperforms non-orthogonal ablations in finite-sample regimes. The right panel further illustrates how nuisance function estimation performance deteriorates as the sample size decreases. Taken together, these results show that our orthogonal LT-O-learners remain robust, thus effectively mitigating the bias induced by nuisance estimation errors.

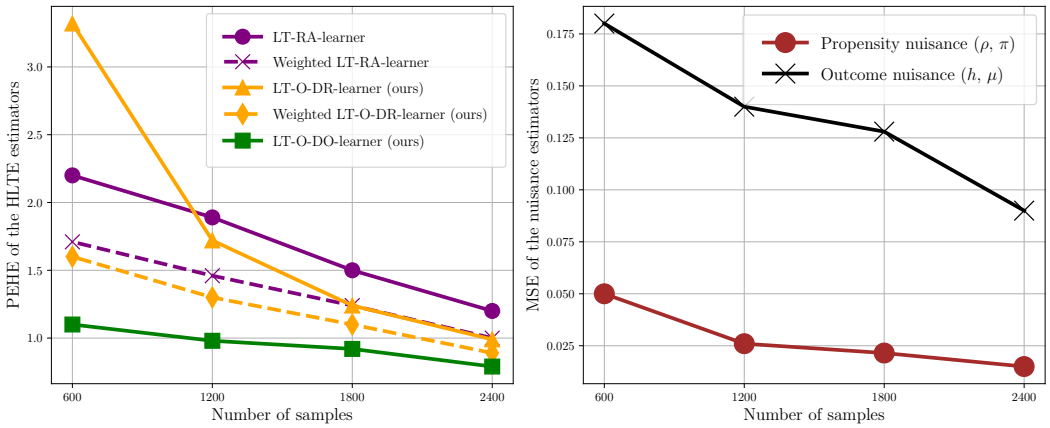


Figure 5: **Benefit of Neyman-orthogonality:** (*Left*) PEHE vs. sample size. The orthogonal LT-O-DO-learner outperforms the weighted (but non-orthogonal) DR/RA-learner. (*Right*) Mean squared error (MSE) of the nuisance estimation. Taken together, high errors in the nuisance functions at small sample sizes drive the instability of non-orthogonal baselines.

B.2 PEHE STRATIFIED BY OVERLAP

In this experiment, we take a granular examination of the performance of our LT-O-learners in the synthetic experiment. We stratify the test samples into five groups based on the overlap score $\text{overlap}(x) = \pi(x)(1-\pi(x))\rho(x)$. We then report the quantiles of the per-sample $\|\hat{\tau}(x) - \tau^*(x)\|$ within each group on the test set. We conduct this analysis for all four overlap scenarios, i.e., healthy-overlap regime ($\mathcal{D}_*$), low treatment overlap ($\mathcal{D}_t$), low outcome overlap ($\mathcal{D}_o$), and low treatment and outcome overlap ($\mathcal{D}_{t+o}$).

Results: We plot the quantiles of the per-sample $\|\hat{\tau}(x) - \tau^*(x)\|$ within each overlap stratum in Figures 6–9. Importantly, our LT-O-learners remain accurate even in low-overlap strata, despite the sample retargeting. For instance, in the low treatment overlap scenario $\mathcal{D}_t$ and joint low overlap scenario $\mathcal{D}_{t+o}$, our LT-O-DO-learner and LT-O-TO-learner achieve the lowest PEHE across all strata. In contrast, non-orthogonal baselines and the LT-O-DR-learner have significantly higher PEHE. This shows that our LT-O-learners do *not* trade accuracy in the low-overlap samples for the overall population stability. On the contrary, it stabilizes the estimation across all the samples, including those with low overlap.

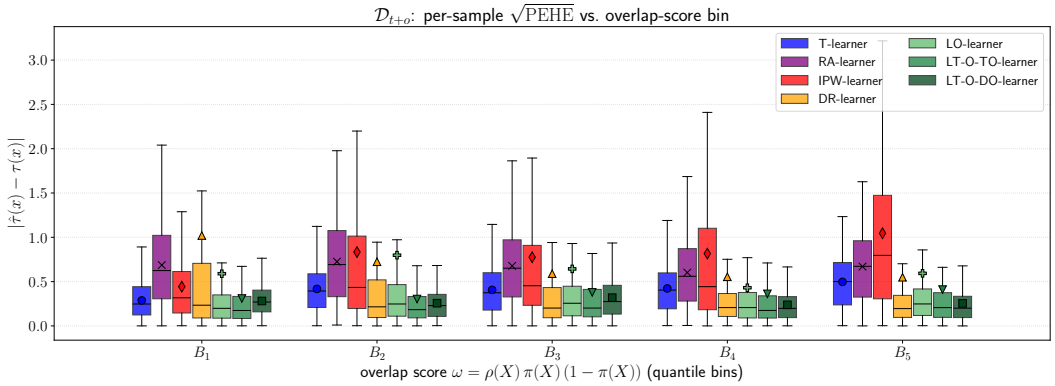


Figure 6: **PEHE stratified by overlap on $\mathcal{D}_{t+o}$:** The boxplot shows the 0%, 25%, 50%, 75%, and 100% quantiles, as well as the mean (marker) of the absolute effect-estimation error within each overlap stratum. The plots show that our LT-O-learners (LT-O-DO-learner, LT-O-TO-learner, LT-O-LO-learner) remain accurate even in low-overlap strata, despite the sample retargeting.

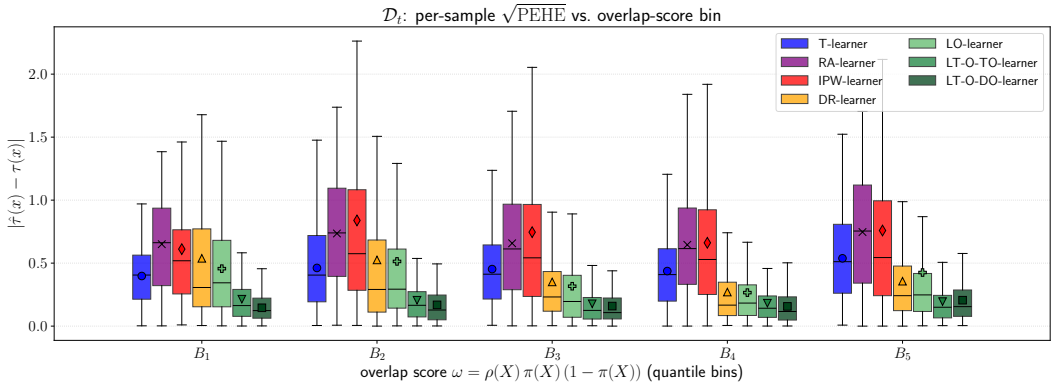


Figure 7: **PEHE stratified by overlap on $\mathcal{D}_t$:** The boxplot shows the 0%, 25%, 50%, 75%, and 100% quantiles, as well as the mean (marker) of the absolute effect-estimation error within each overlap stratum. The plots show that our LT-O-learners (LT-O-DO-learner, LT-O-TO-learner) remain accurate even in low-overlap strata, despite the sample retargeting.

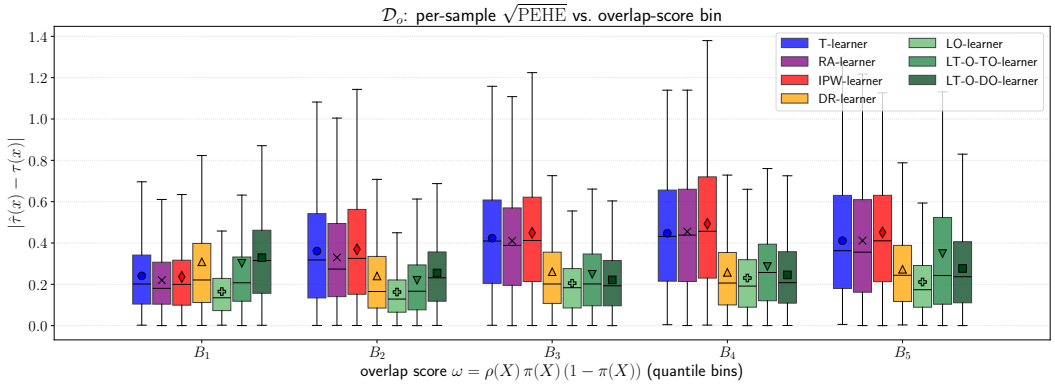


Figure 8: **PEHE stratified by overlap on $\mathcal{D}_o$:** The boxplot shows the 0%, 25%, 50%, 75%, and 100% quantiles, as well as the mean (marker) of the absolute effect-estimation error within each overlap stratum. The plots show that our LT-O-LO-learner remains accurate even in low-overlap strata, despite the sample retargeting.

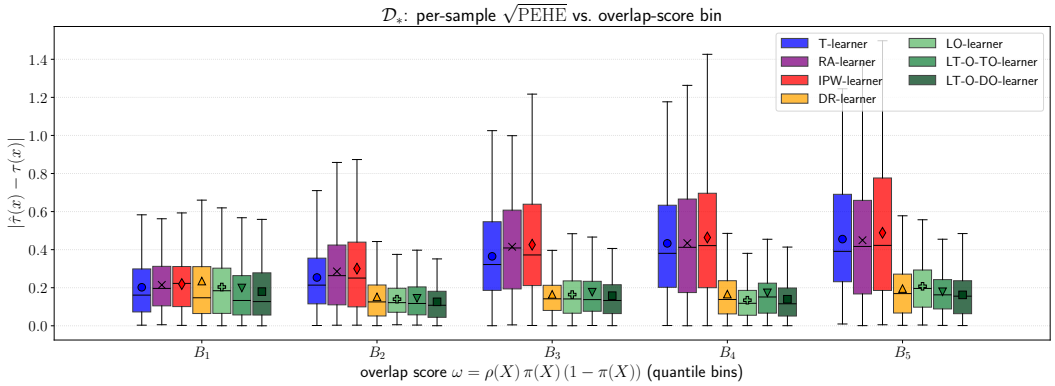


Figure 9: **PEHE stratified by overlap on $\mathcal{D}_*$:** The boxplot shows the 0%, 25%, 50%, 75%, and 100% quantiles, as well as the mean (marker) of the absolute effect-estimation error within each overlap stratum. The plots show that our LT-O-learners (LT-O-DO-learner, LT-O-TO-learner, LT-O-LO-learner) remain accurate even in low-overlap strata, despite the sample retargeting.

B.3 BIAS–VARIANCE DECOMPOSITION

The bias–variance decomposition separates the sources of estimation error. We use the synthetic setting with 20 repeated data draws; this differs from varying training seeds on the fixed GAIN–PSID dataset. For an estimator $\hat{\tau}(x)$ of the true HLTE $\tau^0(x)$ obtained over repeated data draws D, the PEHE decomposes as

$$\mathbb{E}_X\left[\mathbb{E}_D\left[(\hat{\tau}(X)-\tau^0(X))^2\right]\right] = \mathbb{E}_X\left[\left(\mathbb{E}_D[\hat{\tau}(X)]-\tau^0(X)\right)^2\right] + \mathbb{E}_X[\mathrm{Var}_D(\hat{\tau}(X))]\,. \tag{17}$$

The first term is the squared bias (Bias^2), and the second term is the variance (Var). The sum yields the PEHE; it is reported as a square root. Table 4 reports the decomposition for the joint low-overlap scenario $\mathcal{D}_{t+o}$, with the covariates held fixed across repeated data draws.

Table 4: **Bias–variance decomposition on $\mathcal{D}_{t+o}$.** We report squared bias, variance, and the resulting square-root PEHE over 20 runs. Lower is better.

Model	Bias^2	Var	$\sqrt{\mathrm{PEHE}}$
LT-T-learner	0.353	**0.003**	0.596
LT-RA-learner	0.644	0.007	0.806
LT-O-DR-learner	4.913	8.867	3.712
LT-O-TO-learner	**0.035**	0.575	0.781
LT-O-LO-learner	0.106	1.740	1.358
LT-O-DO-learner	0.043	0.021	**0.253**

Results: The LT-O-DO-learner has both low bias and low variance and achieves the lowest overall PEHE. The LT-T-learner and LT-RA-learner have low variance but are subject to plug-in errors. The LT-O-TO-learner has the lowest bias but higher variance because it does not control instability from limited long-term outcome overlap.

B.4 EMPIRICAL COVERAGE UNDER LIMITED OVERLAP

We use a synthetic dataset with a linear target $\tau^0(x) = \theta_0^\top x$, which allows us to evaluate empirical coverage as the overlap score ϵ decreases from 0.5 (healthy overlap) to 0.01 (severely limited overlap). For each ϵ, we use 500 Monte Carlo samples of $N = 10{,}000$ observations with oracle nuisances and form nominal 95% Wald intervals on a fixed grid of 100 covariate points. Coverage and interval widths are averaged over this grid. We also record the estimated asymptotic variance $e_1^\top \widehat{\mathcal{V}}(\omega) e_1$ from Theorem 5.3. Table 5 compares the LT-O-DR-learner against the LT-O-DO-learner.

Table 5: **Empirical coverage and inferential stability across overlap levels.** Coverage is evaluated against the nominal 95% level over 500 Monte Carlo samples. Lower interval width, parameter RMSE, and estimated variance are better.

Metric	Model	$\epsilon = 0.5$	$\epsilon = 0.1$	$\epsilon = 0.05$	$\epsilon = 0.02$	$\epsilon = 0.01$
Coverage	LT-O-DR-learner	0.951	0.952	0.956	0.955	0.956
Coverage	LT-O-DO-learner	0.950	0.951	0.952	0.952	0.945
CI width	LT-O-DR-learner	0.202	0.417	0.588	0.927	1.284
CI width	LT-O-DO-learner	0.201	0.276	0.312	0.322	0.324
$\hat{\theta}$ RMSE	LT-O-DR-learner	0.087	0.183	0.252	0.413	0.586
$\hat{\theta}$ RMSE	LT-O-DO-learner	0.086	0.116	0.129	0.133	0.137
$e_1^\top \widehat{\mathcal{V}} e_1$	LT-O-DR-learner	2.38	7.98	14.9	35.9	69.4
$e_1^\top \widehat{\mathcal{V}} e_1$	LT-O-DO-learner	2.37	8.27	14.9	17.3	17.5

Results: Coverage near 95% is consistent with the stated pointwise target under this correctly specified linear model with oracle nuisances; coverage alone does not establish target preservation

under misspecification. Second, the CI width, $\hat{\theta}$ RMSE, and $e_1^\top \widehat{\mathcal{V}} e_1$ grow much more slowly for the LT-O-DO-learner than for the LT-O-DR-learner. This pattern is consistent with robustness to limited overlap and yields more informative inference through tighter confidence intervals. Hence, the LT-O-DO-learner provides stabilized inference while retaining approximately nominal coverage.

B.5 Representative bias under retargeting

Retargeting can create representative bias in practice for two reasons: model misspecification and finite-sample nuisance estimation. To assess this risk, we perform a bias–variance decomposition with a misspecified second-stage model. Specifically, we use a simple linear regression for the second stage and train and test the learners on the $\mathcal{D}_{t+o}$ synthetic dataset. We report the decomposition separately for the low-overlap stratum, defined by overlap scores below the 25% quantile, and the high-overlap stratum, defined by scores above the 75% quantile.

Table 6: **Representative bias under second-stage misspecification.** Bias–variance decomposition in the low- and high-overlap strata of $\mathcal{D}_{t+o}$. Lower is better.

Model	Stratum	Bias^2	Var	PEHE
LT-T-learner	Low	0.322	**0.003**	0.325
LT-T-learner	High	0.551	**0.002**	0.553
LT-X-learner	Low	0.267	0.003	0.270
LT-X-learner	High	0.463	0.004	0.467
LT-O-DR-learner	Low	0.429	23.62	24.05
LT-O-DR-learner	High	7.104	7.39	14.49
LT-O-DO-learner (ours)	Low	**0.083**	0.168	**0.251**
LT-O-DO-learner (ours)	High	**0.143**	0.239	**0.382**

Results: Importantly, the LT-O-DO-learner achieves the lowest bias in both the low- and high-overlap strata while keeping variance low. The LT-T-learner achieves the lowest variance but is subject to plug-in bias. In contrast, Neyman orthogonality of the LT-O-DO-learner minimizes the influence of nuisance error on bias. The LT-O-DR-learner is also orthogonal but does not control the numerical instability caused by inverse propensities, which leads to high variance. Overall, the LT-O-DO-learner outperforms the baselines in both overlap strata in terms of PEHE.

In short, sample retargeting can introduce bias under model misspecification. Our LT-O-learners mitigate this risk by (i) using retargeting to avoid the extreme inverse propensities that destabilize the LT-O-DR-learner and (ii) reducing bias from other sources, including estimated nuisances. Table 6 indicates that this trade-off is favorable in the evaluated setting, but practitioners should assess subgroup performance when applying retargeted learners to populations with severe overlap limitations.

B.5.1 Subgroup estimates

We estimated subgroup HLTEs with LT-T-learner, LT-RA-learner, LT-O-DR-learner, LT-O-LO-learner, LT-O-TO-learner, and LT-O-DO-learner over 10 random seeds, keeping the dataset fixed. Table 7 reports the mean and standard deviation of the subgroup estimates after applying the comparability correction to each learner. Figure 3 compares LT-T-learner, LT-O-DR-learner, LT-O-TO-learner, and LT-O-DO-learner across age, education, prior earnings, and ethnicity. Variation across seeds provides an empirical evaluation of numerical stability. However, due to absence of ground truth, this should not be seen as a strict metric.

Table 7: GAIN–PSID subgroup HLTE estimates (1000 USD): mean ± standard deviation across 10 seeds on the same dataset, after the calibration correction. Standard deviations describe variation across seeds, not sampling uncertainty or estimation error.

Subgroup	LT-T-learner	LT-RA-learner	LT-O-DR-learner	LT-O-LO-learner	LT-O-TO-learner	LT-O-DO-learner
Age < 30	0.03 ± 0.003	0.13 ± 0.006	0.34 ± 0.183	0.10 ± 0.035	0.18 ± 0.018	0.15 ± 0.021
Age 30–39	0.17 ± 0.012	0.18 ± 0.009	0.33 ± 0.156	0.12 ± 0.041	0.16 ± 0.016	0.15 ± 0.024
Age 40–49	0.14 ± 0.009	0.08 ± 0.007	-0.12 ± 0.142	0.08 ± 0.028	0.07 ± 0.013	0.05 ± 0.017
Age ≥ 50	0.21 ± 0.018	0.03 ± 0.021	-0.48 ± 0.267	0.07 ± 0.032	0.06 ± 0.014	0.04 ± 0.015
HS diploma	0.14 ± 0.008	0.15 ± 0.007	0.44 ± 0.128	0.13 ± 0.030	0.13 ± 0.015	0.15 ± 0.019
No HS diploma	0.10 ± 0.010	0.10 ± 0.008	0.27 ± 0.147	0.20 ± 0.044	0.12 ± 0.017	0.11 ± 0.022
Prior earnings	0.10 ± 0.007	0.14 ± 0.006	0.18 ± 0.119	0.11 ± 0.026	0.05 ± 0.012	0.10 ± 0.016
No prior earnings	0.18 ± 0.013	0.11 ± 0.010	0.16 ± 0.174	0.19 ± 0.048	0.19 ± 0.020	0.16 ± 0.025
White	0.10 ± 0.006	0.12 ± 0.005	0.17 ± 0.096	0.11 ± 0.024	0.05 ± 0.011	0.02 ± 0.014
Black	0.26 ± 0.021	0.27 ± 0.010	0.94 ± 0.412	0.22 ± 0.057	0.12 ± 0.023	0.14 ± 0.028
Other ethnicity	0.28 ± 0.004	-0.02 ± 0.006	0.23 ± 0.203	0.07 ± 0.039	0.28 ± 0.021	0.26 ± 0.023

C Theory

C.1 Identification

In this section, we briefly demonstrate the identifiability of the heterogeneous long-term treatment effects under the surrogacy model. The identification of average long-term treatment effects is first shown by Athey et al. (2026). We make a trivial extension to the heterogeneous setting.

Lemma C.1 (Identification of HLTE)**.** *Under Assumptions 3.1 (Consistency), 3.2 (Positivity), 3.3 (Experimental Unconfoundedness), 3.4 (Surrogacy), and 3.5 (Comparability), the heterogeneous long-term treatment effect $\tau^0(x)$ is identified as:*

$$\tau^0(x) = \mu(1, x) - \mu(0, x), \tag{18}$$

where $\mu(a, x) = \mathbb{E}[h(S, X) \mid A = a, X = x, R = 0]$ and $h(s, x) = \mathbb{E}[Y \mid S = s, X = x, R = 1]$.

Proof. We begin with the definition of the causal estimand of interest:

$$\tau^0(x) = \mathbb{E}[Y(1) - Y(0) \mid X = x, R = 0]. \tag{19}$$

To estimate τ^0, it suffices to identify the conditional expectation of the potential outcome $\mathbb{E}[Y(a) \mid X = x, R = 0]$ for any $a \in \{0, 1\}$.

First, using Assumption 3.1 (Consistency) and Assumption 3.3 (Unconfoundedness), we can relate the potential outcome to the conditional expectation given the treatment assignment $A = a$:

$$\begin{aligned} \mathbb{E}[Y(a) \mid X = x, R = 0] &= \mathbb{E}[Y(a) \mid A = a, X = x, R = 0] \\ &= \mathbb{E}[Y \mid A = a, X = x, R = 0]. \end{aligned} \tag{20}$$

Then, we apply the tower law for expectation by conditioning on the surrogate variable S:

$$\mathbb{E}[Y \mid A = a, X = x, R = 0] = \mathbb{E}_S\Big[\mathbb{E}[Y \mid S, A = a, X = x, R = 0]\Big| A = a, X = x, R = 0\Big]. \tag{21}$$

We invoke Assumption 3.4 (Surrogacy), which states that $A \perp\!\!\!\perp Y \mid S, X, R = 0$. This allows us to remove the conditioning on A in the inner expectation:

$$\mathbb{E}[Y \mid S, A = a, X = x, R = 0] = \mathbb{E}[Y \mid S, X, R = 0]. \tag{22}$$

Since the long-term outcome Y is not observed in the experimental dataset $\mathcal{D}_1$ ($R = 0$), we require Assumption 3.5 (Long-term outcome comparability), specifically $Y \perp\!\!\!\perp R \mid S, X$, to transport the outcome model from the observational dataset $\mathcal{D}_2$:

$$\begin{aligned} \mathbb{E}[Y \mid S, X, R = 0] &= \mathbb{E}[Y \mid S, X, R = 1] \\ &= h(S, X). \end{aligned} \tag{23}$$

Substituting this back into Eq. (21), we get:

$$\mathbb{E}[Y(a) \mid X = x, R = 0] = \mathbb{E}_S\Big[h(S, X)\Big|A = a, X = x, R = 0\Big] = \mu(a, x). \tag{24}$$

Hence, we yield

$$\tau^0(x) = \mu(1, x) - \mu(0, x). \tag{25}$$

□

C.2 Efficient influence function

What is the efficient influence function (EIF)? The EIF is a fundamental concept in semiparametric statistics that describes the sensitivity of an estimator to the underlying data distribution. Formally, consider a functional $\psi(\mathbb{P})$ of the data distribution $\mathbb{P}$. In our case,

$$\psi(\mathbb{P}) = \mathbb{E}_S[\mathbb{E}[Y \mid S, X, R = 1]|A = 1, X = x, R = 0] - \mathbb{E}_S[\mathbb{E}[Y \mid S, X, R = 1]|A = 0, X = x, R = 0]. \tag{26}$$

The influence function $\phi(Z)$ of an estimator $\hat{\psi}$ is defined as the Gateaux derivative of the functional at $\mathbb{P}$ in the direction of a point mass δ_Z (Hampel, 1974):

$$\lim_{\epsilon \to 0} \frac{\psi((1-\epsilon)\mathbb{P} + \epsilon\delta_Z) - \psi(\mathbb{P})}{\epsilon} = \phi(Z). \tag{27}$$

To illustrate this in our context, consider the average long-term treatment effect in the short-term population, $\tau = \mathbb{E}[\tau^0(X) \mid R = 0]$. With $p_0 = P(R = 0)$, the EIF in the standard surrogacy model follows from Eq. (33) at $\omega \equiv 1$:

$$\phi_{\mathrm{ALTE}}(Z) = \frac{\mathcal{T}_{\mathrm{DR}}(Z; \eta_0) - \mathbb{1}(R = 0)\tau}{p_0}. \tag{28}$$

First derived in Chen & Ritzwoller (2023); Athey et al. (2026), the corresponding EIF defines the **efficiency bound** of the ALTE estimation. An estimator is asymptotically efficient (i.e., it achieves the lowest possible asymptotic variance, known as the Cramér-Rao lower bound) if and only if it is asymptotically linear with influence function equal to the EIF (Bickel et al., 1993; Tsiatis, 2006).

Properties of EIF: The EIF has two properties that are critical for constructing robust two-stage learners:

1. **Mean zero:** The expectation of the influence function is zero, i.e., $\mathbb{E}[\phi(Z)] = 0$. Here $\mathbb{E}[\mathcal{T}_{\mathrm{DR}}] = p_0\tau$, so the normalized signal $\mathcal{T}_{\mathrm{DR}}/p_0$ has mean τ.
2. **Neyman-orthogonality:** The most important property is that the EIF is insensitive to perturbations in the nuisance parameters η. Mathematically, the Gâteaux derivative of the expected EIF with respect to the nuisance functions is zero:

$$\frac{\partial}{\partial r}\mathbb{E}\left[\phi(Z; \eta + r(\hat{\eta} - \eta))\right]\Big|_{r=0} = 0. \tag{29}$$

   This *orthogonality* condition implies that the estimation error of the nuisance functions $\hat{\eta} - \eta$ affects the final target estimate τ only at a second-order rate (i.e., $\|\hat{\eta} - \eta\|^2$). This allows the final estimator to be still $\sqrt{n}$-consistent even if the nuisance components are estimated at slower rates (e.g., $n^{-1/4}$), which is common when using machine learning methods (Chernozhukov et al., 2018).

Connection between EIF and DR-pseudo-outcome: In our setting, we estimate a **conditional parameter** $\tau^0(x)$ instead of a scalar parameter τ^0. The connection between the EIF and the *EIF-based pseudo-outcome* (or, equivalently, DR-pseudo-outcome) used in DR-learner (Kennedy, 2023) is that the EIF-based pseudo-outcome is constructed to be the *uncentered* EIF of the corresponding average effect, conditioned on covariates X.

Specifically, $\mathcal{T}_{\mathrm{DR}}$ is the EIF-based signal in Eq. (55), with conditional means

$$\mathbb{E}[\mathcal{T}_{\mathrm{DR}}(Z) \mid X = x, R = 0] = \tau^0(x) \tag{30}$$

$$\mathbb{E}[\mathcal{T}_{\mathrm{DR}}(Z) \mid X = x, R = 1] = 0. \tag{31}$$

By defining a loss function $\mathcal{L}(g) = \mathbb{E}[\mathbb{1}(R=0)g(X)^2 - 2\mathcal{T}_{\mathrm{DR}}g(X)]$, the minimizer is the conditional expectation $\mathbb{E}[\mathcal{T}_{\mathrm{DR}}|X, R = 0]$, which recovers $\tau^0(x)$. Crucially, because $\mathcal{T}_{\mathrm{DR}}(Z)$ inherits the orthogonality property of the EIF, the second stage regression retains robustness against nuisance estimation errors.

Semantics of Neyman-orthogonality: functional vs. loss. In this paper, we frequently use the concept of Neyman-orthogonality. We point out that actually orthogonality may refer to different target objects with different semantic implications:

- **Functional orthogonality:** Primarily used for scalar parameters θ (e.g., ATE) defined by a moment condition $\mathbb{E}[\psi(Z;\theta,\eta)] = 0$. Orthogonality requires $\partial_\eta \mathbb{E}[\psi(Z;\theta_0,\eta_0)] = 0$ (Chernozhukov et al., 2018). It ensures that nuisance estimation error $\hat{\eta} - \eta$ only has second-order impact on the target parameter $\hat{\theta}$.
- **Loss orthogonality:** Used for learning target functions $g(\cdot)$ (e.g., CATE/HLTE) via risk minimization $\min_g \mathcal{L}(g,\eta)$. Orthogonality requires the **cross-derivative** to vanish: $D_\eta D_g \mathcal{L}(g_0,\eta_0) = 0$ (Foster & Syrgkanis, 2023). It thus ensures that distortions in $\hat{\eta}$ may change the shape of the loss landscape but do not shift the *location* of the minimizer g^* to the first order.

C.3 EIF OF THE ω-WEIGHTED AVERAGE LONG-TERM EFFECTS

Recall that our objective is to construct Neyman-orthogonal loss functions. Since the Gateaux derivative of the EIF w.r.t. the nuisances vanishes at the truth (Chernozhukov et al., 2018), an EIF-based pseudo-outcome naturally produces a Neyman-orthogonal loss. Following a similar approach as in (Morzywołek et al., 2023), we first derive the EIF for the *ω-weighted* average long-term treatment effect (WALTE), which is defined as:

$$\tau^0_\omega = \frac{\mathbb{E}\left[\omega(X)\tau^0(X) \mid R=0\right]}{\mathbb{E}\left[\omega(X) \mid R=0\right]} := \frac{N_\omega}{D_\omega}. \tag{32}$$

Here, $\omega(x) = \omega(\pi(x), \rho(x))$.

The derivation generalizes the EIF for the unweighted ALTE established by Chen & Ritzwoller (2023) and Athey et al. (2026) to the ω-weighted case. The efficient influence function for the parameter τ^0_ω in Eq. (32) is given by:

$$\phi(Z;\eta) = \frac{1}{p_0 D_\omega}\Bigg[\mathbb{1}(R=0)\,\omega(X)\big(\hat{\tau}_{\mathrm{AIPW}}(Z;\eta) - \tau^0_\omega\big) + \mathbb{1}(R=1)\,\omega(X)\,\psi_{\mathrm{obs}}(Z;\eta) + \big(\tau^0(X) - \tau^0_\omega\big)\Omega(Z;\eta)\Bigg], \tag{33}$$

where $p_0 = P(R=0)$ and $D_\omega = \mathbb{E}[\omega(X)|R=0]$. For simplicity, we use the shorthand $\omega(X) \equiv \omega(\pi(X), \rho(X))$. The components are defined as:

$$\hat{\tau}_{\mathrm{AIPW}}(Z;\eta) = \mu(1,X) - \mu(0,X) + \frac{A}{\pi(X)}\big(h(S,X) - \mu(1,X)\big) - \frac{1-A}{1-\pi(X)}\big(h(S,X) - \mu(0,X)\big) \tag{34}$$

$$\psi_{\mathrm{obs}}(Z;\eta) = \frac{1-\rho_s(S,X)}{\rho_s(S,X)}\left(\frac{\pi_s(S,X) - \pi(X)}{\pi(X)(1-\pi(X))}\right)\big(Y - h(S,X)\big) \tag{35}$$

$$\Omega(Z;\eta) = \mathbb{1}(R=0)\frac{\partial\omega}{\partial\pi}(A - \pi(X)) + (1-\rho(X))\frac{\partial\omega}{\partial\rho}(R - \rho(X)). \tag{36}$$

Lemma C.2 (Neyman-orthogonality)**.** *$\phi(Z;\eta)$ is the efficient influence function, which satisfies the Neyman-orthogonality condition with respect to the complete set of nuisance functions $\eta = (\pi(x), \pi_s(S,X), \rho(x), \rho_s(S,X), h(s,x), \mu(a,x))$. That is, the Gâteaux derivative of the expected EIF vanishes at the true nuisance parameters:*

$$\partial_r \mathbb{E}\left[\phi(Z;\eta + r\delta)\right]\big|_{r=0} = 0, \tag{37}$$

for any valid direction $\delta = \hat{\eta} - \eta$.

The proof is essentially the same as in Appendix F.4. Hence, we omit the proof.

C.4 Why signed weights arise

The loss in Eq. (10) is quadratic in g, but its sample-wise coefficient $\omega^*(Z;\eta)$ need not be positive. We show why signed coefficients are unavoidable for weights with the boundary behavior of the LT-O-DO-learner weight. Appendix F.6 separately establishes a finite-sample lower bound and minimizer existence under the stated target-model conditions, and uniform convergence to a strongly convex population loss.

Proposition C.3 (Orthogonalization forces negative weights)**.** *Consider the orthogonalization in Eq. (11) with a positive differentiable weight $\omega(\pi,\rho)$. Fix $\rho \in (0,1)$ and suppose, as $\pi \downarrow 0$,*

$$\frac{\pi\,\partial_\pi \omega}{\omega} \longrightarrow 2, \qquad \frac{\rho\,\partial_\rho \omega}{\omega} \longrightarrow 1. \tag{38}$$

Then, on $R=0, A=0$, $\omega^*/\omega \to \rho - 2 < 0$. Consequently, no guarantee of nonnegative sample-wise weights, or convexity of the empirical loss for every sample and target class, holds over this family of weights and data distributions. The LT-O-DO-learner weight satisfies Eq. (38); the analogous obstruction holds as $\pi \uparrow 1$ on $R=0, A=1$.

Proof. Substitution of $R=0, A=0$ in Eq. (11) yields

$$\frac{\omega^*}{\omega} = 1 - \frac{\pi\,\partial_\pi \omega}{\omega} - (1-\rho)\frac{\rho\,\partial_\rho \omega}{\omega} \longrightarrow \rho - 2. \tag{39}$$

For $\omega = \pi^2(1-\pi)^2\rho$, the ratio equals $\rho - 2(1-2\pi)/(1-\pi)$, which is negative whenever $\pi < (2-\rho)/(4-\rho)$. On $R=0, A=1$, it equals $\rho + 2(1-2\pi)/\pi$, which is negative whenever $\pi > 2/(4-\rho)$. □

Hence, preserving this overlap-weight decay together with its orthogonal correction precludes a general finite-sample convexity guarantee. The trade-off concerns first-order nuisance-bias correction for these weights; it does not imply that every orthogonal learner has negative weights or that the LT-O-DO-learner weight minimizes HLTE estimation error. Indeed, the LT-O-DR-learner and LT-O-LO-learner have nonnegative coefficients. Signed weights also arise in broader orthogonalized estimation frameworks, including overlap-weighted CATE estimation (Morzywołek et al., 2023), treatment effects with time-varying confounding (Hess et al., 2026), and heterogeneous treatment effects on censored survival outcomes (Frauen et al., 2025b). We retain the orthogonal correction and the chosen overlap weighting, and control the resulting optimization problem through the conditions in Appendix F.6.

D BIAS ANALYSIS UNDER ASSUMPTION VIOLATION

In this section, we rigorously analyze the pointwise bias of the HLTE identification formulation when the surrogacy assumption (Assumption 3.4) is violated. Note that Athey et al. (2026) analyzed the bias of the ALTE under surrogacy violations, while we adapt their method to extend to the pointwise HLTE setting.

Proposition D.1 (Pointwise bias of the T-learner under surrogacy and comparability violations)**.** *Let Assumptions 3.1 (Consistency), 3.2 (Positivity), and 3.3 ($\mathcal{D}_1$ Unconfoundedness) hold, but allow Assumption 3.4 (Surrogacy) and Assumption 3.5 (Comparability) to be violated. Recall*

$$\pi(s,x) \;:=\; P(A=1 \mid S=s, X=x, R=0), \qquad \pi(x) \;:=\; P(A=1 \mid X=x, R=0), \quad (40)$$

and the conditional outcome regressions, assumed to be well-defined on the relevant conditional supports, i.e.,

$$\begin{aligned}\mu(s,a,x,R=0) &:= E[Y \mid S=s, A=a, X=x, R=0],\\ \mu(s,x,R=0) &:= E[Y \mid S=s, X=x, R=0],\\ \mu(s,x,R=1) &:= E[Y \mid S=s, X=x, R=1].\end{aligned}$$

Then, we have that

$$\mu(s,x,R=0) \;=\; \pi(s,x)\,\mu(s,1,x,R=0) + (1-\pi(s,x))\,\mu(s,0,x,R=0). \quad (41)$$

Let

$$\delta(s,x) \;:=\; \mu(s,1,x,R=0) - \mu(s,0,x,R=0), \qquad \eta(s,x) \;:=\; \mu(s,x,R=1) - \mu(s,x,R=0) \quad (42)$$

denote the residual conditional treatment contrast in the short-term sample and the comparability gap*, respectively. Since S is post-treatment, $\delta(s,x)$ is not generally a causal direct effect without additional assumptions. Assumption 3.4 and Assumption 3.5 imply the corresponding conditional mean restrictions*

$$\delta(s,x) \equiv 0, \qquad \eta(s,x) \equiv 0. \quad (43)$$

Define the T-learner population estimand

$$\bar{\tau}(x) \;:=\; \mu^{\mathrm{TL}}(1,x) - \mu^{\mathrm{TL}}(0,x), \qquad \mu^{\mathrm{TL}}(a,x) \;:=\; E\big[\mu(S,X,R=1) \,\big|\, A=a, X=x, R=0\big]. \quad (44)$$

Then, the bias of $\bar{\tau}$ with respect to $\tau^0(x) = E[Y(1)-Y(0) \mid X=x, R=0]$ is

$$\tau^0(x) - \bar{\tau}(x) \;=\; \frac{1}{\pi(x)\,(1-\pi(x))}\Big\{\mathcal{B}_{\mathrm{surr}}(x) \;-\; \mathcal{B}_{\mathrm{comp}}(x)\Big\}, \quad (45)$$

where

$$\begin{aligned}\mathcal{B}_{\mathrm{surr}}(x) &:= E\big[\,\pi(S,X)\,(1-\pi(S,X))\,\delta(S,X) \,\big|\, X=x, R=0\big],\\ \mathcal{B}_{\mathrm{comp}}(x) &:= E\big[\,(\pi(S,X)-\pi(x))\,\eta(S,X) \,\big|\, X=x, R=0\big].\end{aligned}$$

Marginalizing Eq. (45) over $X \mid R=0$ gives the corresponding ALTE bias decomposition, in analogy with Theorem 4(iv) of Athey et al. (2026).

Proof. We establish Eq. (45) in five steps.

Identification of $\tau^0(x)$ under unconfoundedness. Assumption 3.3 gives $A \perp\!\!\!\perp (Y(0), Y(1)) \mid X, R=0$. For $a \in \{0,1\}$,

$$\begin{aligned}E[Y(a) \mid X=x, R=0] &= E[Y(a) \mid A=a, X=x, R=0] && \text{(Assumption 3.3)}\\ &= E[Y \mid A=a, X=x, R=0] && \text{(Assumption 3.1)}\\ &= E\big[\mu(S,a,X,R=0) \,\big|\, A=a, X=x, R=0\big] && \text{(tower law).}\end{aligned}$$

Subtracting the $a=0$ case from the $a=1$ case,

$$\tau^0(x) \;=\; E[\mu(S,1,X,R=0) \mid A=1, X=x, R=0] - E[\mu(S,0,X,R=0) \mid A=0, X=x, R=0]. \quad (46)$$

Pointwise decomposition. By the definition of $\pi(s,x)$,

$$\mu(s,x,R=0) \;=\; \pi(s,x)\,\mu(s,1,x,R=0) + (1-\pi(s,x))\,\mu(s,0,x,R=0). \tag{47}$$

Inserting $\mu(s,x,R=0)$ as a pivot, for any s,x:

$$\begin{aligned}\mu(s,x,R=1)-\mu(s,1,x,R=0) &= \big[\mu(s,x,R=1)-\mu(s,x,R=0)\big] + \big[\mu(s,x,R=0)-\mu(s,1,x,R=0)\big]\\ &= \eta(s,x) - (1-\pi(s,x))\,\delta(s,x), \qquad (48)\\ \mu(s,x,R=1)-\mu(s,0,x,R=0) &= \eta(s,x) + \pi(s,x)\,\delta(s,x). \qquad (49)\end{aligned}$$

Computing $\bar\tau(x)-\tau^0(x)$. Combining the definition of $\bar\tau$ with Eq. (46), Eq. (48), and Eq. (49),

$$\begin{aligned}\bar\tau(x)-\tau^0(x) &= E[\mu(S,X,R=1)-\mu(S,1,X,R=0)\mid A=1,X=x,R=0]\\ &\quad - E[\mu(S,X,R=1)-\mu(S,0,X,R=0)\mid A=0,X=x,R=0]\\ &= \underbrace{\big\{E[\eta(S,X)\mid A=1,x,R=0]-E[\eta(S,X)\mid A=0,x,R=0]\big\}}_{=:\,(\mathrm{I})}\\ &\quad - \underbrace{\big\{E[(1-\pi(S,X))\delta(S,X)\mid A=1,x,R=0]+E[\pi(S,X)\delta(S,X)\mid A=0,x,R=0]\big\}}_{=:\,(\mathrm{II})}.\end{aligned} \tag{50}$$

Bayes' rule. For any integrable g, since $f(s\mid X=x,A=a,R=0) = f(s\mid X=x,R=0)\,P(A=a\mid s,x,R=0)/P(A=a\mid x,R=0)$,

$$E[g(S,X)\mid A=1,X=x,R=0] = \frac{E[\pi(S,X)\,g(S,X)\mid X=x,R=0]}{\pi(x)}, \tag{51}$$

$$E[g(S,X)\mid A=0,X=x,R=0] = \frac{E[(1-\pi(S,X))\,g(S,X)\mid X=x,R=0]}{1-\pi(x)}. \tag{52}$$

Deriving (I) and (II). Apply Eq. (51) to the surrogacy term:

$$\begin{aligned}(\mathrm{II}) &= \frac{E[\pi(S,X)(1-\pi(S,X))\,\delta(S,X)\mid x,R=0]}{\pi(x)} + \frac{E[\pi(S,X)(1-\pi(S,X))\,\delta(S,X)\mid x,R=0]}{1-\pi(x)}\\ &= E[\pi(S,X)(1-\pi(S,X))\,\delta(S,X)\mid X=x,R=0]\cdot\frac{(1-\pi(x))+\pi(x)}{\pi(x)(1-\pi(x))} \;=\; \frac{\mathcal{B}_{\mathrm{surr}}(x)}{\pi(x)(1-\pi(x))}.\end{aligned}$$

For the comparability term, again by Eq. (51) and the identity $(1-\pi(x))\,\pi(s,x)-\pi(x)\,(1-\pi(s,x)) = \pi(s,x)-\pi(x)$,

$$\begin{aligned}(\mathrm{I}) &= \frac{E[\pi(S,X)\,\eta(S,X)\mid x,R=0]}{\pi(x)} - \frac{E[(1-\pi(S,X))\,\eta(S,X)\mid x,R=0]}{1-\pi(x)}\\ &= \frac{E\big[\,\{(1-\pi(x))\,\pi(S,X)-\pi(x)\,(1-\pi(S,X))\}\,\eta(S,X)\,\big|\,x,R=0\big]}{\pi(x)(1-\pi(x))} \;=\; \frac{\mathcal{B}_{\mathrm{comp}}(x)}{\pi(x)(1-\pi(x))}.\end{aligned}$$

Substituting both into Eq. (50) gives $\bar\tau(x)-\tau^0(x) = \big(\mathcal{B}_{\mathrm{comp}}(x)-\mathcal{B}_{\mathrm{surr}}(x)\big)/\big(\pi(x)(1-\pi(x))\big)$, which is Eq. (45). □

Corollary D.2 (Pointwise sensitivity bound under bounded direct effect)**.** *Suppose Assumption 3.5 (Comparability) holds, and there exists $c:\mathcal{X}\to\mathbb{R}_+$ with $|\delta(s,x)|\le c(x)$ for almost every s in the support of $S\mid X=x,R=0$. Then*

$$\big|\,\tau^0(x)-\bar\tau(x)\,\big| \;\le\; c(x)\cdot\frac{E[\pi(S,X)(1-\pi(S,X))\mid X=x,R=0]}{\pi(x)(1-\pi(x))} \;\le\; c(x). \tag{53}$$

The second inequality follows from concavity of $u\mapsto u(1-u)$ and Jensen's inequality applied to $\pi(x)=E[\pi(S,X)\mid X=x,R=0]$. In particular, the right-hand side is computable from $R=0$ data alone (it does not require Y), making it directly usable as a sensitivity-analysis input.

E INSTANTIATIONS

E.1 DERIVATION OF THE LT-O-DR-LEARNER (SPECIAL CASE $\omega \equiv 1$)

When we trivially set $\omega \equiv 1$, i.e. no overlap weighting in the loss, the LT pseudo-outcome reduces to the EIF of the ALTE from Chen & Ritzwoller (2023); Athey et al. (2026), and we recover the LT-O-DR-learner under the standard surrogacy model. Substituting $\omega \equiv 1$ into Eq. (10) and Eq. (12) yields

$$\mathcal{L}_1(g, \eta) = \mathbb{E}\left[\mathbb{1}(R=0) g(X)^2 - 2\mathcal{T}_{\mathrm{DR}} \cdot g(X)\right], \tag{54}$$

$$\mathcal{T}_{\mathrm{DR}}(Z; \eta) = \mathbb{1}(R=0)\hat{\tau}_{\mathrm{AIPW}}(Z; \eta) + \mathbb{1}(R=1)\psi_{\mathrm{obs}}(Z; \eta). \tag{55}$$

Here, $\hat{\tau}_{\mathrm{AIPW}}(Z; \eta)$ is the AIPW pseudo-outcome with Y replaced by the surrogate index $h(S, X)$, and ψ_{obs} corrects the first-order error of h using the observed long-term outcome Y in $\mathcal{D}_2$.

E.2 DERIVATION OF THE LT-O-TO-LEARNER

To derive the LT-O-TO-learner, we set the weighting function to $\omega(X) = \pi(X)^2(1-\pi(X))^2$ to specifically target treatment overlap. The squared form is motivated by the variance lower bound in Appendix J, whose binding term scales with $1/[\rho\,\pi^2(1-\pi)^2]$. We substitute ω into Eq. (11) and Eq. (12).

First, the residual term $\Omega(Z; \eta)$ becomes:

$$\begin{aligned}\Omega(Z; \eta) &= \mathbb{1}(R=0)\frac{\partial \omega}{\partial \pi}(A - \pi) \\ &= \mathbb{1}(R=0)\, 2\pi(1-\pi)(1-2\pi)(A-\pi).\end{aligned} \tag{56}$$

Substituting back into ω^*:

$$\omega^*(Z; \eta) = \mathbb{1}(R=0)\big[\pi^2(1-\pi)^2 + 2\pi(1-\pi)(1-2\pi)(A-\pi)\big]. \tag{57}$$

Although ω^* might be negative, its conditional expectation $\mathbb{E}[\omega^* \mid X] = (1-\rho)\,\pi^2(1-\pi)^2$ is positive. The LT-pseudo-outcome $\mathcal{T}_{\mathrm{LT}}$ becomes:

$$\begin{aligned}\mathcal{T}_{\mathrm{LT}}(Z; \eta) &= \mathbb{1}(R=0)\Big[\pi^2(1-\pi)^2\,\hat{\tau}_{\mathrm{AIPW}} + (\mu_1 - \mu_0)\,2\pi(1-\pi)(1-2\pi)(A-\pi)\Big] \\ &\quad + \mathbb{1}(R=1)\,\pi^2(1-\pi)^2\,\psi_{\mathrm{obs}} \\ &= (\mu_1 - \mu_0)\,\omega^*(Z; \eta) \\ &\quad + \mathbb{1}(R=0)\,\pi(1-\pi)(A-\pi)\big(h(S, X) - \mu(A, X)\big) \\ &\quad + \mathbb{1}(R=1)\,\pi^2(1-\pi)^2\,\psi_{\mathrm{obs}}.\end{aligned} \tag{58}$$

Plugging ω^* and $\mathcal{T}_{\mathrm{LT}}$ into $\mathcal{L}_\omega = \mathbb{E}_n[\omega^* g^2 - 2\mathcal{T}_{\mathrm{LT}} g]$ yields the LT-O-TO-learner loss

$$\mathcal{L}_{\pi^2(1-\pi)^2}(g; \eta) = \mathbb{E}\Big[\omega^*(Z; \eta)\, g(X)^2 - 2\,\mathcal{T}_{\mathrm{LT}}(Z; \eta)\, g(X)\Big]. \tag{59}$$

Remark: The simpler choice $\omega = \pi(1-\pi)$ yields the closed form $\omega^* = \mathbb{1}(R=0)(A-\pi)^2$ and a loss whose $R=0$ component completes the square to $\big((h(S, X) - m(X)) - (A-\pi)g(X)\big)^2$, closely resembling the standard R-learner (Nie & Wager, 2021). The squared variant $\omega = \pi^2(1-\pi)^2$ adopted here departs from this clean form, but it more closely tracks the binding term of the variance lower bound in long-term settings (Appendix J) and yields better empirical performance under low overlap (cf. Section 6). The simpler $\omega = \pi(1-\pi)$ variant is provided as an alternative instantiation in Appendix E.5.

E.3 DERIVATION OF THE LT-O-LO-LEARNER

The LT-O-LO-learner is intended for settings where only long-term outcome overlap (LO) is low. We choose $\omega(x) = \rho(x)$ to downweight samples with a low probability of observing long-term outcomes.

The corresponding LT weighting function and LT-pseudo-outcome are

$$\begin{aligned}
\Omega(Z;\eta) &= (1-\rho(X))(R-\rho(X)),\\
\omega^*(Z;\eta) &= (R-\rho(X))^2,\\
\mathcal{T}_{\mathrm{LT}}(Z;\eta) &= (R-\rho(X))^2(\mu(1,X)-\mu(0,X))\\
&\quad + \mathbb{1}(R=0)\,\rho(X)\frac{A-\pi(X)}{\pi(X)(1-\pi(X))}\big(h(S,X)-\mu(A,X)\big)\\
&\quad + \mathbb{1}(R=1)\,\rho(X)\psi_{\mathrm{obs}}(Z;\eta).
\end{aligned}$$

E.4 Derivation of the LT-O-DO-learner

The LT-O-DO-learner is designed for settings in which both treatment overlap and long-term outcome overlap are limited. We choose $\omega(x) = \pi(x)^2(1-\pi(x))^2\rho(x)$, which is the inverse of the binding term in the variance lower bound of Appendix J and jointly downweights samples with weak treatment overlap and low probability of observing long-term outcomes.

Substituting into Eq. (11) with $\partial\omega/\partial\pi = 2\pi(1-\pi)(1-2\pi)\rho$ and $\partial\omega/\partial\rho = \pi^2(1-\pi)^2$, the residual term is

$$\begin{aligned}
\Omega(Z;\eta) &= \mathbb{1}(R=0)\,2\pi(1-\pi)(1-2\pi)\rho\,(A-\pi)\\
&\quad + (1-\rho)\,\pi^2(1-\pi)^2\,(R-\rho).
\end{aligned}$$

The resulting LT-weighting function and LT-pseudo-outcome are

$$\begin{aligned}
\omega^*(Z;\eta) &= \mathbb{1}(R=0)\,\rho\big[\rho\,\pi^2(1-\pi)^2 + 2\pi(1-\pi)(1-2\pi)(A-\pi)\big]\\
&\quad + \mathbb{1}(R=1)\,\pi^2(1-\pi)^2(1-\rho)^2,\\
\mathcal{T}_{\mathrm{LT}}(Z;\eta) &= (\mu(1,X)-\mu(0,X))\,\omega^*(Z;\eta)\\
&\quad + \mathbb{1}(R=0)\,\pi(1-\pi)\,\rho\,(A-\pi)\big(h(S,X)-\mu(A,X)\big)\\
&\quad + \mathbb{1}(R=1)\,\pi^2(1-\pi)^2\,\rho\,\psi_{\mathrm{obs}}(Z;\eta).
\end{aligned}$$

As with the LT-O-TO-learner, ω^* is sample-wise sign-indefinite but has non-negative conditional expectation $\mathbb{E}[\omega^* \mid X] = (1-\rho)\,\pi^2(1-\pi)^2\rho \geq 0$.

E.5 Other possible instantiations

The four instantiations above (LT-O-DR-learner, LT-O-TO-learner, LT-O-LO-learner, LT-O-DO-learner) cover the most common overlap regimes. We also list a few additional variants that may be useful in practice.

LT-O-TO-learner (alternative, $\omega = \pi(1-\pi)$): The simpler choice $\omega(X) = \pi(X)(1-\pi(X))$ — the canonical Crump et al. (2009) treatment-overlap weight — yields a particularly clean derivation. With $\partial\omega/\partial\pi = 1-2\pi$, the residual term is $\Omega(Z;\eta) = \mathbb{1}(R=0)(1-2\pi)(A-\pi)$, and

$$\begin{aligned}
\omega^*(Z;\eta) &= \mathbb{1}(R=0)\big[\pi(1-\pi) + (1-2\pi)(A-\pi)\big]\\
&= \mathbb{1}(R=0)(A-\pi)^2,\\
\mathcal{T}_{\mathrm{LT}}(Z;\eta) &= \mathbb{1}(R=0)(A-\pi)\big(h(S,X)-m(X)\big) + \mathbb{1}(R=1)\,\psi_t(Z;\eta),
\end{aligned}$$

where $m(X) := \pi\mu_1 + (1-\pi)\mu_0 = \mathbb{E}[h(S,X) \mid X, R=0]$ and $\psi_t := \frac{1-\rho_s}{\rho_s}(\pi_s-\pi)(Y-h)$. Completing the square for the $R=0$ component yields, up to an additive constant independent of g,

$$\begin{aligned}
\mathcal{L}_{\pi(1-\pi)}(g;\eta) \equiv \mathbb{E}\Big[&\mathbb{1}(R=0)\big((h(S,X)-m(X)) - (A-\pi(X))g(X)\big)^2\\
&- 2\,\mathbb{1}(R=1)\,\psi_t(Z;\eta)\,g(X)\Big],
\end{aligned}$$

which closely resembles the standard R-learner (Nie & Wager, 2021): the $\mathbb{1}(R=0)$ component merely replaces Y with the surrogate index h, and the $\mathbb{1}(R=1)$ correction restores Neyman-orthogonality against the first-order error in $\hat{h}$. We adopt the squared variant $\omega = \pi^2(1-\pi)^2$ as the

main LT-O-TO-learner since it more closely tracks the binding term of the variance lower bound and yields better empirical performance under low overlap (cf. Section 6).

LT-O-DO-learner (alternative, $\omega = \pi(1-\pi)\rho$): Pairing the canonical treatment-overlap weight with ρ also targets joint overlap but with a single $\pi(1-\pi)$ factor rather than the squared one. The resulting LT-weighting function and LT-pseudo-outcome are

$$\begin{aligned}\omega^*(Z;\eta) &= \mathbb{1}(R=0)\,\rho\big[\rho\,\pi(1-\pi) + (1-2\pi)(A-\pi)\big] \\ &\quad + \mathbb{1}(R=1)\,\pi(1-\pi)(1-\rho)^2, \\ \mathcal{T}_{\mathrm{LT}}(Z;\eta) &= (\mu(1,X)-\mu(0,X))\,\omega^*(Z;\eta) \\ &\quad + \mathbb{1}(R=0)\,\rho(A-\pi)\big(h(S,X)-\mu(A,X)\big) \\ &\quad + \mathbb{1}(R=1)\,\rho\,\pi(1-\pi)\,\psi_{\mathrm{obs}}(Z;\eta).\end{aligned}$$

LT-O-LO$\frac{1}{2}$-learner: We consider an alternative weighting function $\omega(x) = \sqrt{\rho(x)}$ to provide a smaller penalty for low long-term outcome overlap. Here $\partial\omega/\partial\rho = 1/(2\sqrt{\rho})$, so the corresponding LT-weighting function and LT-pseudo-outcome are given by:

$$\begin{aligned}\omega^*(Z;\eta) &= \begin{cases} \frac{\sqrt{\rho(X)}(1+\rho(X))}{2}, & \text{if } R=0, \\ \frac{(1-\rho(X))^2}{2\sqrt{\rho(X)}}, & \text{if } R=1. \end{cases} \\ \mathcal{T}_{\mathrm{LT}}(Z;\eta) &= (\mu(1,X)-\mu(0,X))\,\omega^*(Z;\eta) \\ &\quad + \mathbb{1}(R=0)\sqrt{\rho(X)}\,\frac{A-\pi(X)}{\pi(X)(1-\pi(X))}\big(h(S,X)-\mu(A,X)\big) \\ &\quad + \mathbb{1}(R=1)\sqrt{\rho(X)}\,\psi_{\mathrm{obs}}(Z;\eta).\end{aligned}$$

Equivalently, the normalized pseudo-outcome $\mathcal{T}_{\mathrm{LT}}(Z;\eta)/\omega^*(Z;\eta)$ is

$$\mu(1,X)-\mu(0,X) + \begin{cases} \frac{2}{1+\rho(X)}\,\frac{A-\pi(X)}{\pi(X)(1-\pi(X))}\big(h(S,X)-\mu(A,X)\big), & \text{if } R=0, \\ \frac{2\rho(X)}{(1-\rho(X))^2}\,\psi_{\mathrm{obs}}(Z;\eta), & \text{if } R=1. \end{cases}$$

F PROOFS

Overview. We state the assumptions in Section F.1. In Section F.2, we establish the conditional-mean identities for the LT weight and pseudo-outcome. In Section F.3, we derive the variance decomposition of the DR pseudo-outcome. In Section F.4, we prove Neyman orthogonality, and in Section F.5, we prove oracle equivalence. In Section F.6, we prove the empirical lower bound, uniform convergence, population strong convexity, and minimizer existence under the stated target-model conditions. Finally, in Section F.7, we derive positive curvature under estimated nuisances and use it to prove the error bound.

F.1 ASSUMPTIONS

Notation: $\|f\| := \|f\|_2 = \sqrt{\mathbb{E}[f^2]}$ denotes the L_2 norm under the marginal distribution of f's arguments (X for $\hat{g}, \hat{\mu}, \hat{\pi}, \hat{\rho}$, and (S, X) for $\hat{h}, \hat{\pi}_s, \hat{\rho}_s$), and $\|f\|_4 := (\mathbb{E}[f^4])^{1/4}$ the corresponding L_4 norm.

Assumption F.1 (Boundedness)**.** *(i) All nuisance estimators are bounded in L_2, i.e., $\exists M > 0, r > 0$, s.t. $\forall \hat{\eta} \in \mathcal{B}_r(\eta, \mathcal{G})$, $\|\hat{\eta}\| \leq M$. Each component $\hat{f}$ additionally has finite fourth moment, $\|\hat{f}\|_4 < \infty$. (ii) All random variables are bounded. (iii) $\exists \alpha_\omega > 0, r > 0$, s.t. $\forall \hat{\eta} \in \mathcal{B}_r(\eta, \mathcal{G})$, we have $\hat{\omega}(X) > \alpha_\omega$. (iv) The oracle target is uniformly bounded: $\|g_0\|_\infty \leq G_0 < \infty$.*

Assumption F.2 (Smoothness of ω)**.** *The weighting function $\omega(x)$ is lower bounded by a positive $\alpha_\omega > 0$. On a convex neighborhood containing the true and estimated propensity functions, $\omega^*(Z; \eta)$ is twice continuously differentiable in (π, ρ), and its first derivatives and conditional expected Hessian are uniformly bounded. On the same neighborhood, $|\omega|$, $|\partial_\pi \omega|$, and $|\partial_\rho \omega|$ are uniformly bounded as well. In particular,*

$$\sup_{\bar{\eta}} \sup_{x \in \mathcal{X}} \left\| \mathbb{E}\left[\nabla_\eta^2 \omega^*(Z; \bar{\eta}) \mid X = x \right] \right\|_{\text{op}} < \infty, \tag{60}$$

Assumption F.3 (Nuisance overlap and localization)**.** *There exist constants $\epsilon > 0$ and $r > 0$ such that, almost surely,*

$$\epsilon \leq \hat{\pi}(X) \leq 1 - \epsilon, \qquad \epsilon \leq \rho_s(S, X), \hat{\rho}_s(S, X) \leq 1 - \epsilon, \tag{61}$$

and $\|\hat{\pi} - \pi\|_\infty + \|\hat{\rho} - \rho\|_\infty \leq r$.

Assumption F.4 (Target class)**.** *For the error-bound analysis, $\mathcal{G}$ is a nonempty closed convex subset of $L_2(P_X)$, and the oracle loss admits a minimizer $g_0 \in \mathcal{G}$.*

Assumption F.5 (Cross-fitting)**.** Each nuisance estimate used to evaluate an observation is trained on data independent of that observation. When taking expectations involving estimated nuisances, we hold their training data fixed and average over independent evaluation data.

Assumption 3.2 and Assumptions F.1–F.5 collect the regularity conditions used below; overlap, boundedness, smoothness, and cross-fitting are standard in orthogonal-learning analyses (Foster & Syrgkanis, 2023; Kennedy, 2023).

Assumption F.6 (Additional regularity for empirical well-posedness)**.** Let $Z_1, \ldots, Z_N$ be independent draws from the pooled law, with a fixed number K of folds whose sizes tend to infinity, and let $\hat{\eta}_{-k}$ denote the nuisance estimate trained without fold k as in Assumption F.5. The target class $\mathcal{G}$ consists of measurable functions, satisfies $\sup_{g \in \mathcal{G}} \|g\|_\infty \leq B < \infty$, and has finite covering number $\mathcal{N}(\epsilon, \mathcal{G}, \|\cdot\|_\infty)$ for every $\epsilon > 0$. For the true and fitted nuisance, almost surely, $|\omega^*(Z; \eta)| \leq W$ and $|\mathcal{T}_{\text{LT}}(Z; \eta)| \leq H$, where $W, H < \infty$ do not depend on N. Finally, the fitted LT weight and pseudo-outcome are consistent in the following conditional mean sense:

$$d_{N,k} := \mathbb{E}\left[B^2 |\omega^*(Z; \hat{\eta}_{-k}) - \omega^*(Z; \eta_0)| + 2B |\mathcal{T}_{\text{LT}}(Z; \hat{\eta}_{-k}) - \mathcal{T}_{\text{LT}}(Z; \eta_0)| \;\middle|\; \hat{\eta}_{-k} \right] \xrightarrow{P} 0 \tag{62}$$

for each fold.

F.2 PROPERTIES OF THE LT WEIGHTING FUNCTION AND PSEUDO-OUTCOME

In the following section, we make use of the following properties:

(1) $\mathbb{E}\left[\Omega(Z;\eta) \mid X\right] = 0$:

$$\begin{aligned}\mathbb{E}\left[\Omega(Z;\eta) \mid X\right] &= \mathbb{E}\left[\mathbb{1}(R=0)\frac{\partial\omega}{\partial\pi}(A-\pi(X)) + (1-\rho(X))\frac{\partial\omega}{\partial\rho}(R-\rho(X)) \mid X\right] \\ &= (1-\rho(X))\mathbb{E}\left[\frac{\partial\omega}{\partial\pi}(A-\pi(X)) \mid X, R=0\right] \\ &= 0. \qquad \text{(By definition, } \pi(X) = \mathbb{E}\left[A \mid X, R=0\right]\text{)}\end{aligned} \tag{63}$$

(2) $\mathbb{E}\left[\omega^*(Z;\eta) \mid X\right] = (1-\rho(X))\omega(X)$:

Since $\mathbb{E}\left[\Omega(Z;\eta) \mid X\right] = 0$ and

$$\omega^*(Z;\eta) = \mathbb{1}(R=0)\omega(X) + \Omega(Z;\eta), \tag{64}$$

we have

$$\mathbb{E}\left[\omega^*(Z;\eta) \mid X\right] = \mathbb{E}\left[\mathbb{1}(R=0)\omega(X) \mid X\right] = (1-\rho(X))\omega(X) \tag{65}$$

(3) $\mathbb{E}\left[\mathbb{1}(R=1)\psi_{\text{obs}}(Z;\eta) \mid X\right] = 0$:

$$\begin{aligned}\mathbb{E}\left[\mathbb{1}(R=1)\psi_{\text{obs}}(Z;\eta) \mid X\right] &= \rho(X)\mathbb{E}\left[\frac{1-\rho_s(S,X)}{\rho_s(S,X)}\left(\frac{\pi_s(S,X)-\pi(X)}{\pi(X)(1-\pi(X))}\right)\big(Y - h(S,X) \mid X, R=1\big]\right] \\ &= \rho(X)\mathbb{E}_S\left[\frac{1-\rho_s(S,X)}{\rho_s(S,X)}\left(\frac{\pi_s(S,X)-\pi(X)}{\pi(X)(1-\pi(X))}\right)\mathbb{E}_Y\left[\big(Y - h(S,X) \mid S, X, R=1\big] \mid X, R=1\right]\right]\end{aligned} \tag{66}$$

$$= 0. \tag{67}$$

(4) $\mathbb{E}\left[\hat{\tau}_{\text{AIPW}}(Z;\eta) \mid X, R=0\right] = \tau^0(X)$:

$$\begin{aligned}&\mathbb{E}[\hat{\tau}_{\text{AIPW}}(Z;\eta) \mid X, R=0] \\ &= \mathbb{E}\Big[\mu(1,X) - \mu(0,X) + \frac{A}{\pi(X)}\big(h(S,X) - \mu(1,X)\big) - \frac{1-A}{1-\pi(X)}\big(h(S,X) - \mu(0,X)\big) \;\Big|\; X, R=0\Big] \\ &= \mu(1,X) - \mu(0,X) + \mathbb{E}\Big[\mathbb{E}\Big[\frac{A}{\pi(X)}\big(h(S,X) - \mu(1,X)\big) \;\Big|\; X, A, R=0\Big] \;\Big|\; X, R=0\Big] \\ &\quad - \mathbb{E}\Big[\mathbb{E}\Big[\frac{1-A}{1-\pi(X)}\big(h(S,X) - \mu(0,X)\big) \;\Big|\; X, A, R=0\Big] \;\Big|\; X, R=0\Big].\end{aligned} \tag{68}$$

For the first augmentation term, using that $\pi(X)$ and $\mu(1,X)$ are X-measurable and that $A \in \{0,1\}$,

$$\mathbb{E}\Big[\frac{A}{\pi(X)}\big(h(S,X) - \mu(1,X)\big) \;\Big|\; X, A, R=0\Big] = \frac{A}{\pi(X)}\Big(\mathbb{E}[h(S,X) \mid X, A, R=0] - \mu(1,X)\Big). \tag{69}$$

When $A=1$, by definition of $\mu(1,X)$,

$$\mathbb{E}[h(S,X) \mid X, A=1, R=0] = \mu(1,X), \tag{70}$$

so the bracket in Eq. (69) is zero; when $A=0$, $A/\pi(X) = 0$. Hence,

$$\mathbb{E}\Big[\frac{A}{\pi(X)}\big(h(S,X) - \mu(1,X)\big) \;\Big|\; X, R=0\Big] = 0. \tag{71}$$

And with the same derivation, we have $\mathbb{E}\Big[\frac{1-A}{1-\pi(X)}\big(h(S,X) - \mu(0,X)\big) \;\Big|\; X, A, R=0\Big] = 0$. Therefore,

$$\mathbb{E}[\hat{\tau}_{\text{AIPW}}(Z;\eta) \mid X, R=0] = \mu(1,X) - \mu(0,X) = \tau^0(X). \tag{72}$$

(5) $\mathbb{E}\left[\mathcal{T}_{\text{LT}}(Z;\eta) \mid X\right] = (1-\rho(X))\omega(X)\tau^0(X)$**:**

By definition,

$\mathcal{T}_{\text{LT}}(Z;\eta) = \mathbb{1}(R=0)\omega(X)\hat{\tau}_{\text{AIPW}}(Z;\eta) + \mathbb{1}(R=1)\cdot\omega(X)\psi_{\text{obs}}(Z;\eta) + (\mu(1,X) - \mu(0,X))\Omega(Z;\eta)$,
Using properties (1), (3) and (4) from above, we have immediately that $\mathbb{E}\left[\mathcal{T}_{\text{LT}}(Z;\eta) \mid X\right] = (1-\rho(X))\omega(X)\tau^0(X)$.

F.3 VARIANCE ANALYSIS

Here, we formally state the variance of the long-term DR-pseudo-outcome:

Lemma F.7 (Variance of DR-pseudo-outcome)**.** *The variance of DR-pseudo-outcome is*

$$\mathrm{Var}(\mathcal{T}_{\mathrm{DR}} \mid X) \geq \underbrace{(1-\rho(X)) \cdot \Sigma_t(X)}_{\textit{treatment overlap}} + \underbrace{\rho(X) \cdot \Sigma_o(X)}_{\textit{long-term outcome overlap}} \tag{73}$$

$$\Sigma_t(X) = \mathbb{E}\left[\frac{(A-\pi(X))^2 \sigma_h(A,X)}{\pi(X)^2(1-\pi(X))^2} \,\middle|\, X, R=0\right] = \frac{\sigma_h(1,X)}{\pi(X)} + \frac{\sigma_h(0,X)}{1-\pi(X)}, \tag{74}$$

$$\Sigma_o(X) = \mathbb{E}\left[\left(\frac{1-\rho_s(S,X)}{\rho_s(S,X)} \frac{\pi_s(S,X)-\pi(X)}{\pi(X)(1-\pi(X))}\right)^2 \sigma_y(S,X) \,\middle|\, X, R=1\right], \tag{75}$$

where $\sigma_h(A,X) = \mathrm{Var}(h(S,X)|X,A,R=0)$*,* $\sigma_y(S,X) = \mathrm{Var}(Y|S,X,R=1)$*. All nuisances are evaluated at truth, with finite conditional second moments.*

The inverse propensity factors can amplify variance under limited overlap, as in Fisher (2023). If the arm-specific surrogate-index variances remain bounded above and away from zero, $\Sigma_t(X)$ scales as $1/[\pi(X)(1-\pi(X))]$. The outcome term also depends on surrogate informativeness and residual outcome variance; low $\rho(X)$ alone does not imply a pointwise rate for $\rho_s(S,X)$. Assumption K.8 states sufficient conditions for variance divergence.

Proof. By the Law of Total Variance, conditioning on the dataset indicator R, we have:

$$\begin{aligned}\mathrm{Var}(\mathcal{T}_{\mathrm{DR}} \mid X) &= \mathbb{E}_R[\mathrm{Var}(\mathcal{T}_{\mathrm{DR}} \mid X, R) \mid X] + \mathrm{Var}_R(\mathbb{E}[\mathcal{T}_{\mathrm{DR}} \mid X, R] \mid X) &(76)\\ &\geq \mathbb{E}_R[\mathrm{Var}(\mathcal{T}_{\mathrm{DR}} \mid X, R) \mid X] &(77)\end{aligned}$$

where the inequality holds because variance is non-negative. Expanding the expectation over $R \in \{0,1\}$ with $\rho(X) = \mathbb{P}(R=1|X)$, we obtain:

$$\mathrm{Var}(\mathcal{T}_{\mathrm{DR}} \mid X) \geq (1-\rho(X)) \underbrace{\mathrm{Var}(\mathcal{T}_{\mathrm{DR}} \mid X, R=0)}_{\text{(A)}} + \rho(X) \underbrace{\mathrm{Var}(\mathcal{T}_{\mathrm{DR}} \mid X, R=1)}_{\text{(B)}} \tag{78}$$

Part (A): Treatment overlap instability ($R=0$). In the experimental dataset ($R=0$), $\mathcal{T}_{\mathrm{DR}}$ reduces to the AIPW estimator:

$$\hat{\tau}_{\mathrm{AIPW}} = Z_0 + \mu(1,X) - \mu(0,X), \tag{79}$$

$$Z_0 = \frac{A-\pi(X)}{\pi(X)(1-\pi(X))}(h(S,X) - \mu(A,X)). \tag{80}$$

Hence, we have

$$\begin{aligned}\mathrm{Var}(\mathcal{T}_{\mathrm{DR}} \mid X, R=0) &= \mathrm{Var}(\hat{\tau}_{\mathrm{AIPW}} \mid X, R=0)\\ &= \mathrm{Var}\left(\frac{A-\pi(X)}{\pi(X)(1-\pi(X))}(h(S,X) - \mu(A,X)) \mid X, R=0\right)\\ &= \mathbb{E}_A[\mathrm{Var}(Z_0 \mid A, X, R=0)] + \mathrm{Var}_A(\underbrace{\mathbb{E}[Z_0 \mid A, X, R=0]}_{=0}) \qquad (81)\end{aligned}$$

The inner expectation is zero because $\mu(A,X) = \mathbb{E}[h(S,X)|A,X,R=0]$. Thus, we are left with the expected variance:

$$\begin{aligned}\Sigma_t(X) &= \mathrm{Var}(Z_0 \mid X, R=0)\\ &= \mathbb{E}\left[\frac{(A-\pi(X))^2 \sigma_h(A,X)}{\pi(X)^2(1-\pi(X))^2} \,\middle|\, X, R=0\right]. \qquad (82)\end{aligned}$$

Averaging $(A-\pi)^2$ over the treatment arm yields $\sigma_h(1,X)/\pi(X) + \sigma_h(0,X)/(1-\pi(X))$, rather than a squared inverse-overlap rate.

Part (B): Long-term outcome overlap instability ($R = 1$). In the observational dataset ($R = 1$), $\mathcal{T}_{\mathrm{DR}}$ uses the observation proxy ψ_{obs}. Let:

$$Z_1 = \frac{1 - \rho_s(S, X)}{\rho_s(S, X)} W_\pi(S, X)(Y - h(S, X)),$$

where $W_\pi(S, X) = \left(\frac{\pi_s(S,X) - \pi(X)}{\pi(X)(1-\pi(X))}\right)$. Conditioning on the surrogate S, we get:

$$\mathrm{Var}(Z_1 \mid X, R = 1) = \mathbb{E}_S[\mathrm{Var}(Z_1 \mid S, X, R = 1)] + \mathrm{Var}_S(\underbrace{\mathbb{E}[Z_1 \mid S, X, R = 1]}_{=0}). \tag{83}$$

The inner expectation is zero because $h(S, X) = \mathbb{E}[Y|S, X, R = 1]$. Thus:

$$\begin{aligned}\Sigma_o(X) &= \mathrm{Var}(Z_1 \mid X, R = 1) \\ &= \mathbb{E}\left[\left(\frac{1 - \rho_s(S, X)}{\rho_s(S, X)}\right)^2 W_\pi(S, X)^2 \sigma_y(S, X) \,\middle|\, X, R = 1\right].\end{aligned} \tag{84}$$

This term scales with $\frac{1}{\rho_s(S,X)^2}$, causing high variance under limited long-term outcome overlap.

Combining (A) and (B) together yields the lower bound:

$$\mathrm{Var}(\mathcal{T}_{\mathrm{DR}} \mid X) \geq (1 - \rho(X))\Sigma_t(X) + \rho(X)\Sigma_o(X) \tag{85}$$

F.4 Neyman-orthogonality

For fixed candidate nuisances $\bar{\eta}$, write $q_{\bar{\eta}}(X) = \mathbb{E}[\omega^*(Z; \bar{\eta}) \mid X]$ and $b_{\bar{\eta}}(X) = \mathbb{E}[\mathcal{T}_{\mathrm{LT}}(Z; \bar{\eta}) \mid X]$. Then, for every fixed g and admissible direction v,

$$D_g \mathcal{L}_\omega(g, \bar{\eta})[v] = 2\mathbb{E}[(q_{\bar{\eta}}(X)g(X) - b_{\bar{\eta}}(X))v(X)]. \tag{86}$$

It suffices to show that both conditional coefficients have zero first derivative in every nuisance direction at η_0. All derivatives below are evaluated at truth, with integrability permitting differentiation under the expectation.

Treatment and sampling propensities (π, ρ). Holding the other nuisances at truth, Eq. (11) gives

$$\begin{aligned}q_{\bar{\eta}} = (1 - \rho)\omega(\bar{\pi}, \bar{\rho}) + (1 - \rho)\partial_\pi \omega(\bar{\pi}, \bar{\rho})(\pi - \bar{\pi}) \\ + (1 - \bar{\rho})\partial_\rho \omega(\bar{\pi}, \bar{\rho})(\rho - \bar{\rho}).\end{aligned} \tag{87}$$

The derivatives of the first term cancel those of the residual corrections at $(\bar{\pi}, \bar{\rho}) = (\pi, \rho)$, so $D_{(\pi,\rho)} q_{\eta_0} = 0$. Moreover, the true outcome nuisances satisfy $\mathbb{E}[h - \mu_a \mid A = a, X, R = 0] = 0$ and $\mathbb{E}[Y - h \mid S, X, R = 1] = 0$. Consequently, even with candidate $(\bar{\pi}, \bar{\rho})$, $b_{\bar{\eta}} = (\mu_1 - \mu_0)q_{\bar{\eta}}$, giving $D_{(\pi,\rho)} b_{\eta_0} = 0$.

Outcome regressions μ_a. The coefficient q does not involve μ_a. For a perturbation δ_μ,

$$\begin{aligned}D_\mu b_{\eta_0}[\delta_\mu] = (1 - \rho)\omega\mathbb{E}\left[\delta_{\mu_1} - \delta_{\mu_0} - \frac{A}{\pi}\delta_{\mu_1} + \frac{1 - A}{1 - \pi}\delta_{\mu_0} \,\middle|\, X, R = 0\right] \\ + (\delta_{\mu_1} - \delta_{\mu_0})\mathbb{E}[\Omega \mid X] = 0.\end{aligned} \tag{88}$$

This uses $\mathbb{E}[A \mid X, R = 0] = \pi$ and $\mathbb{E}[\Omega \mid X] = 0$.

Surrogate index h. Again $D_h q = 0$. The two contributions to $D_h b$ are

$$(1 - \rho)\omega\mathbb{E}\left[\frac{\pi_s - \pi}{\pi(1 - \pi)}\delta_h \,\middle|\, X, R = 0\right], \tag{89}$$

$$-\rho\omega\mathbb{E}\left[\frac{1 - \rho_s}{\rho_s}\frac{\pi_s - \pi}{\pi(1 - \pi)}\delta_h \,\middle|\, X, R = 1\right]. \tag{90}$$

They cancel by the change-of-measure identity $\rho\mathbb{E}[(1-\rho_s)f/\rho_s \mid X, R = 1] = (1-\rho)\mathbb{E}[f \mid X, R = 0]$.

Surrogate propensities (π_s, ρ_s). These nuisances enter only through ψ_{obs}. Their derivative terms multiply $Y - h(S, X)$, whose conditional mean given $(S, X, R = 1)$ is zero. Thus their derivatives of b vanish, and they do not enter q.

By linearity in the perturbation direction, $D_\eta q_{\eta_0} = D_\eta b_{\eta_0} = 0$. Substitution into the loss derivative proves $D_\eta D_g \mathcal{L}_\omega(g, \eta_0)[v, \delta] = 0$ for every fixed g and admissible v, δ. The argument uses observed-data nuisance identities; identification is needed only to interpret $\mu_1 - \mu_0$ as the causal HLTE.

F.5 Proof of Theorem 5.1

At true nuisances $\eta = \eta_0$, the loss $\mathcal{L}_\omega(g;\eta)$ differs by a constant independent of g from:

$$\mathcal{R}_\omega(g;\eta) = \mathbb{E}\left[\mathbb{1}(R=0)\omega(X)\big(\hat{\tau}_{\mathrm{AIPW}} - g\big)^2 - 2\mathbb{1}(R=1)\omega(X)\psi_{\mathrm{obs}}\cdot g + \Omega\big((\mu(1,X)-\mu(0,X)) - g\big)^2\right]. \tag{91}$$

Recall that

$$\psi_{\mathrm{obs}}(Z;\eta) = \frac{1-\rho_s(S,X)}{\rho_s(S,X)}\left(\frac{\pi_s(S,X)-\pi(X)}{\pi(X)(1-\pi(X))}\right)\big(Y - h(S,X)\big). \tag{92}$$

Hence, the conditional expectation of $\mathbb{1}(R=1)\psi_{\mathrm{obs}}(Z;\eta)$ given X equals zero:

$$\mathbb{E}\left[\mathbb{1}(R=1)\psi_{\mathrm{obs}}(Z;\eta) \mid X\right] \tag{93}$$

$$= \mathbb{E}\left[\mathbb{1}(R=1)\frac{1-\rho_s(S,X)}{\rho_s(S,X)}\left(\frac{\pi_s(S,X)-\pi(X)}{\pi(X)(1-\pi(X))}\right)\big(Y-h(S,X)\big)\middle| X\right] \tag{94}$$

$$= \mathbb{E}\left[\mathbb{1}(R=1)\frac{1-\rho_s(S,X)}{\rho_s(S,X)}\left(\frac{\pi_s(S,X)-\pi(X)}{\pi(X)(1-\pi(X))}\right)\underbrace{\mathbb{E}[Y-h(S,X)\mid S,X,R=1]}_{=0}\middle| X\right] \tag{95}$$

$$= 0. \tag{96}$$

Note that the conditional expectation of $\Omega(Z;\eta)$ given X is also zero:

$$\mathbb{E}\left[\Omega(Z;\eta)\mid X\right] = \mathbb{E}\left[\mathbb{1}(R=0)\frac{\partial\omega}{\partial\pi}(A-\pi(X)) + (1-\rho(X))\frac{\partial\omega}{\partial\rho}(R-\rho(X))\middle| X\right] \tag{97}$$

$$= \frac{\partial\omega}{\partial\pi}\mathbb{E}\left[\mathbb{1}(R=0)(A-\pi(X))\mid X\right] + (1-\rho(X))\frac{\partial\omega}{\partial\rho}\underbrace{\mathbb{E}\left[R-\rho(X)\mid X\right]}_{=0} \tag{98}$$

$$= \frac{\partial\omega}{\partial\pi}\Pr(R=0|X)\underbrace{\mathbb{E}\left[A-\pi(X)\mid X, R=0\right]}_{=0} = 0. \tag{99}$$

Hence, the risk function simplifies to

$$\mathcal{R}_\omega(g;\eta) = \mathbb{E}\left[\mathbb{1}(R=0)\omega(X)\big(\hat{\tau}_{\mathrm{AIPW}}(Z;\eta) - g(X)\big)^2\right], \tag{100}$$

where

$$\hat{\tau}_{\mathrm{AIPW}}(Z;\eta) = \mu(1,X) - \mu(0,X) + \frac{A}{\pi(X)}(h(S,X)-\mu(1,X)) - \frac{1-A}{1-\pi(X)}(h(S,X)-\mu(0,X)). \tag{101}$$

We further write the Eq. (100) as:

$$\mathcal{R}_\omega(g;\eta) = \mathbb{E}\left[\mathbb{1}(R=0)\omega(X)\big(\hat{\tau}_{\mathrm{AIPW}}(Z;\eta) - g(X)\big)^2\right] \tag{102}$$

$$= \mathbb{E}\left[\mathbb{1}(R=0)\omega(X)\big(\hat{\tau}_{\mathrm{AIPW}}(Z;\eta) - \tau^0(X) + \tau^0(X) - g(X)\big)^2\right] \tag{103}$$

$$= \mathbb{E}\left[\mathbb{1}(R=0)\omega(X)\big(\hat{\tau}_{\mathrm{AIPW}}(Z;\eta) - \tau^0(X)\big)^2\right] + \mathbb{E}\left[\mathbb{1}(R=0)\omega(X)\big(\tau^0(X) - g(X)\big)^2\right]$$
$$+ 2\mathbb{E}\left[\mathbb{1}(R=0)\omega(X)\big(\hat{\tau}_{\mathrm{AIPW}}(Z;\eta) - \tau^0(X)\big)\big(\tau^0(X) - g(X)\big)\right] \tag{104}$$

$$= \mathbb{E}\left[\mathbb{1}(R=0)\omega(X)\big(\hat{\tau}_{\mathrm{AIPW}}(Z;\eta) - \tau^0(X)\big)^2\right] + \mathbb{E}\left[\mathbb{1}(R=0)\omega(X)\big(\tau^0(X) - g(X)\big)^2\right] \tag{105}$$

$$= p_0\mathbb{E}\left[\omega(X)\big(\tau^0(X) - g(X)\big)^2 \mid R=0\right] + \mathrm{Const}_g \tag{106}$$

$$= p_0\mathcal{L}_\omega^{\mathrm{oracle}}(g) + \mathrm{Const}_g. \tag{107}$$

The reason the cross term in the decomposition is zero is because

$$\begin{aligned}
&\mathbb{E}\left[\mathbb{1}(R=0)\omega(X)\big(\hat{\tau}_{\mathrm{AIPW}}(Z;\eta)-\tau^0(X)\big)\big(\tau^0(X)-g(X)\big)\right] \\
&=\mathbb{E}\left[\mathbb{E}\left[\mathbb{1}(R=0)\omega(X)\big(\hat{\tau}_{\mathrm{AIPW}}(Z;\eta)-\tau^0(X)\big)\big(\tau^0(X)-g(X)\big)\mid X\right]\right] &(108)\\
&=\mathbb{E}\left[\mathbb{1}(R=0)\omega(X)\big(\tau^0(X)-g(X)\big)\underbrace{\mathbb{E}\left[\hat{\tau}_{\mathrm{AIPW}}(Z;\eta)-\tau^0(X)\mid X,R=0\right]}_{=0}\right] &(109)\\
&=0. &(110)
\end{aligned}$$

Therefore we successfully show that $\mathcal{R}_\omega(g;\eta)$ has the same minimizer as the oracle loss $\mathcal{L}_\omega^{\text{oracle}}(g;\eta)$. Since $\mathcal{R}_\omega(g;\eta)$ is also equivalent with $\mathcal{L}_\omega(g;\eta)$, we finish the proof.

F.6 Empirical well-posedness

We use Assumption F.6 from Appendix F.1.

Proof of Theorem 5.1, part (iii). Write $a_0(Z)=\omega^*(Z;\eta_0)$, $b_0(Z)=\mathcal{T}_{\mathrm{LT}}(Z;\eta_0)$, and $\ell_0(g;Z)=a_0(Z)g(X)^2-2b_0(Z)g(X)$. For i in fold $k(i)$, let $\hat{a}_i=\omega^*(Z_i;\hat{\eta}_{-k(i)})$ and $\hat{b}_i=\mathcal{T}_{\mathrm{LT}}(Z_i;\hat{\eta}_{-k(i)})$. Normalize Eq. (15) by N, which leaves its minimizers unchanged:

$$\mathcal{L}_{\omega,N}(g,\hat{\eta})=\frac{1}{N}\sum_{i=1}^{N}\{\hat{a}_i g(X_i)^2-2\hat{b}_i g(X_i)\}. \tag{111}$$

For every $g\in\mathcal{G}$ and every sample,

$$\mathcal{L}_{\omega,N}(g,\hat{\eta})\geq-\frac{1}{N}\sum_{i=1}^{N}\{|\hat{a}_i|B^2+2|\hat{b}_i|B\}\geq-(WB^2+2HB)=:-M. \tag{112}$$

Thus, the lower bound is uniform in N and g, regardless of coefficient signs.

For uniform convergence, define the empirical coefficient error

$$\Delta_N=\frac{1}{N}\sum_{i=1}^{N}\{B^2|\hat{a}_i-a_0(Z_i)|+2B|\hat{b}_i-b_0(Z_i)|\}. \tag{113}$$

Conditional on the training data for each fold, the mean of its average summand is $d_{N,k}$. Conditional Markov's inequality and Eq. (62) imply that this average is $o_P(1)$. Explicitly, for any $t,\gamma>0$, its probability of exceeding t is at most $P(d_{N,k}>\gamma)+\gamma/t$. Since K is fixed, $\Delta_N=o_P(1)$ without requiring independence between fitted models across folds.

Let $J=2WB+2H$. The oracle losses obey $|\ell_0(g;Z)|\leq M$ and $|\ell_0(g;Z)-\ell_0(\tilde{g};Z)|\leq J\|g-\tilde{g}\|_\infty$. An ϵ-net, Hoeffding's inequality at each net point, and a union bound therefore give, with probability at least $1-\delta$,

$$\sup_{g\in\mathcal{G}}|\mathcal{L}_{\omega,N}(g,\hat{\eta})-\mathcal{L}_\omega(g,\eta_0)|\leq\Delta_N+2J\epsilon+M\sqrt{\frac{2\log\{2\mathcal{N}(\epsilon,\mathcal{G},\|\cdot\|_\infty)/\delta\}}{N}}. \tag{114}$$

First fixing ϵ and then taking it arbitrarily small proves uniform convergence in probability.

The conditional identities $\mathbb{E}[a_0(Z)\mid X]=q_0(X)$ and $\mathbb{E}[b_0(Z)\mid X]=q_0(X)\tau^0(X)$ with $q_0(X)=(1-\rho(X))\omega(X)$ and $\kappa_0=\operatorname{ess\,inf}_x q_0(x)\geq\alpha_R\alpha_\omega>0$ by Assumptions 3.2 and F.2, give

$$\mathcal{L}_\omega(g,\eta_0)-\mathcal{L}_\omega(\tau^0,\eta_0)=\mathbb{E}[q_0(X)(g(X)-\tau^0(X))^2]. \tag{115}$$

For any direction $v\in L_2(P_X)$,

$$D_g^2\mathcal{L}_\omega(g,\eta_0)[v,v]=2\mathbb{E}[q_0(X)v(X)^2]\geq2\kappa_0\|v\|_{L_2(P_X)}^2. \tag{116}$$

This proves the asserted strong convexity when $\kappa_0>0$. Positivity of q_0 almost surely alone gives strict convexity on distinct $L_2(P_X)$ equivalence classes, but need not give a positive strong-convexity constant in that norm.

Finally, if $\mathcal{G} = \{g_\theta : \theta \in \Theta\}$ with compact $\Theta \subset \mathbb{R}^p$ and pointwise continuous $\theta \mapsto g_\theta(x)$, the finite sum in Eq. (111) is continuous and is guaranteed to have a minimizer. More generally, on a nonempty closed Θ with the same continuity and bounded outputs, for any $\lambda > 0$,

$$\mathcal{L}_{\omega,N}(g_\theta, \hat{\eta}) + \lambda\|\theta\|_2^2 \geq -M + \lambda\|\theta\|_2^2 \longrightarrow +\infty \quad \text{as } \|\theta\|_2 \longrightarrow \infty. \tag{117}$$

Notably, this existence of minimizer does not require convexity at all. □

F.7 Proof of the error bound

F.7.1 Positive curvature

We show that the population loss has positive curvature under the given assumptions.

Lemma F.8 (Positive curvature under estimated nuisances)**.** *Suppose Assumptions 3.2, F.2, and F.3 hold. Define*

$$Q_0(x) = (1 - \rho(x))\omega(\pi(x), \rho(x)), \qquad \kappa_0 = \operatorname*{ess\,inf}_x Q_0(x), \tag{118}$$

$$q_{\hat{\eta}}(x) = \mathbb{E}[\omega^*(Z; \hat{\eta}) \mid X = x]. \tag{119}$$

Then $\kappa_0 > 0$, and there exists a constant $C < \infty$, depending only on the overlap and smoothness constants, such that

$$q_{\hat{\eta}}(x) \geq \kappa_0 - Cr^2 \qquad \textit{almost surely}. \tag{120}$$

Consequently, for every $v \in L_2(P_X)$,

$$\mathbb{E}\left[\omega^*(Z; \hat{\eta})v(X)^2\right] \geq (\kappa_0 - Cr^2)\|v\|_{L_2(P_X)}^2. \tag{121}$$

In particular, if $r^2 \leq \kappa_0/(2C)$, this curvature is at least $\kappa_0/2 > 0$.

Proof. For arbitrary candidate propensities $\bar{\eta} = (\bar{\pi}, \bar{\rho})$, define

$$\begin{aligned} q_{\bar{\eta}}(x) &:= \mathbb{E}[\omega^*(Z; \bar{\eta}) \mid X = x] \\ &= (1 - \rho)\omega(\bar{\pi}, \bar{\rho}) + (1 - \rho)\partial_\pi\omega(\bar{\pi}, \bar{\rho})(\pi - \bar{\pi}) + (1 - \bar{\rho})\partial_\rho\omega(\bar{\pi}, \bar{\rho})(\rho - \bar{\rho}), \end{aligned} \tag{122}$$

where all functions are evaluated at x. At $\bar{\eta} = \eta_0 = (\pi, \rho)$, Eq. (122) gives $q_{\eta_0}(x) = Q_0(x)$. Moreover, direct differentiation shows that both first derivatives of $q_{\bar{\eta}}(x)$ with respect to $\bar{\pi}$ and $\bar{\rho}$ vanish at $\bar{\eta} = \eta_0$: the derivative of the base weight is canceled by the corresponding residual term in Ω. Assumption F.2 and nondegenerate overlap uniformly bound the Hessian of $q_{\bar{\eta}}(x)$ in a neighborhood of η_0. Taylor's theorem therefore yields

$$q_{\hat{\eta}}(x) \geq Q_0(x) - Cr^2 \geq \kappa_0 - Cr^2 \tag{123}$$

almost surely, after absorbing the norm-equivalence constant into C. Since $v(X)^2$ is nonnegative and measurable with respect to X, iterated expectation gives Eq. (121). □

Corollary F.9 (Strong convexity of the LT-O-learners population loss)**.** *Suppose that $\mathcal{G} \subset L_2(P_X)$ is convex and that the conditions of Lemma F.8 hold, with Assumption F.1 ensuring finite loss and first derivative. Let $\alpha_r = \kappa_0 - Cr^2 > 0$. Then, for every $g_1, g_2 \in \mathcal{G}$,*

$$\mathcal{L}_\omega(g_2, \hat{\eta}) \geq \mathcal{L}_\omega(g_1, \hat{\eta}) + D_g\mathcal{L}_\omega(g_1, \hat{\eta})[g_2 - g_1] + \alpha_r\|g_2 - g_1\|_{L_2(P_X)}^2. \tag{124}$$

Hence, under the convention that a μ-strongly convex functional has remainder $(\mu/2)\|g_2 - g_1\|_{L_2(P_X)}^2$, the LT-O-learners population loss evaluated at $\hat{\eta}$ is $2\alpha_r$-strongly convex. In particular, if $r^2 \leq \kappa_0/(2C)$, it is at least κ_0-strongly convex.

Proof. For $v = g_2 - g_1$, the quadratic form of the loss gives the exact identity

$$\mathcal{L}_\omega(g_2, \hat{\eta}) = \mathcal{L}_\omega(g_1, \hat{\eta}) + D_g\mathcal{L}_\omega(g_1, \hat{\eta})[v] + \mathbb{E}[\omega^*(Z; \hat{\eta})v(X)^2]. \tag{125}$$

Lemma F.8 establishes $\mathbb{E}[\omega^*(Z; \hat{\eta}) \mid X] \geq \alpha_r$ almost surely. Applying iterated expectation to the final term yields Eq. (124). □

F.7.2 PROOF

We adapt the argument of Morzywołek et al. (2023). The proof first uses positive curvature to reduce the estimation error to a directional derivative of the loss. We then express this derivative through a score, separate its oracle component from the nuisance errors, and bound those errors to obtain the claimed rate. Throughout, we hold the fitted nuisance functions fixed and take expectations over a fresh observation. Write $v = \hat{g} - g_0$, $\delta_f = \hat{f} - f$ for each nuisance, $\tau = \mu_1 - \mu_0 = \tau^0$, and $q_{\hat{\eta}}(X) = \mathbb{E}[\hat{\omega}^* \mid X]$.

Step 1: Reduce the estimation error to a score bound. The quadratic loss in Eq. (10) gives the exact expansion

$$\mathcal{L}_\omega(\hat{g}, \hat{\eta}) - \mathcal{L}_\omega(g_0, \hat{\eta}) = D_g \mathcal{L}_\omega(g_0, \hat{\eta})[v] + \mathbb{E}[q_{\hat{\eta}}(X) v(X)^2], \tag{126}$$

$$D_g^2 \mathcal{L}_\omega(g, \hat{\eta})[v, v] = 2\mathbb{E}[q_{\hat{\eta}}(X) v(X)^2]. \tag{127}$$

For $r^2 \leq \kappa_0/(2C)$, Lemma F.8 bounds $q_{\hat{\eta}}$ below by $\alpha = \kappa_0/2$. On the event defining R_g in Theorem 5.2, the loss difference on the left of Eq. (126) is at most R_g, since $g_0 \in \mathcal{G}$. Substituting these two bounds into Eq. (126) yields

$$\alpha \|v\|_2^2 \leq R_g - D_g \mathcal{L}_\omega(g_0, \hat{\eta})[v]. \tag{128}$$

Thus, Eq. (128) shows that it remains to bound $-D_g \mathcal{L}_\omega(g_0, \hat{\eta})[v]$ from above. For this purpose, define the observation-level loss score

$$s_{\hat{\eta}}(Z; g) := \widehat{\mathcal{T}}_{\mathrm{LT}}(Z) - \hat{\omega}^*(Z) g(X). \tag{129}$$

This is the negative one-half derivative of the observation-level loss with respect to its prediction $g(X)$. The fourth-moment, overlap, and weight bounds in Assumptions F.1, F.2, and F.3 ensure that $s_{\hat{\eta}}(Z; g_0) \in L_2(P)$. Since $v \in L_2(P_X)$, its product with the score is integrable by Cauchy–Schwarz. Differentiating the quadratic population loss and then conditioning on X therefore gives

$$-\tfrac{1}{2} D_g \mathcal{L}_\omega(g_0, \hat{\eta})[v] = \mathbb{E}[s_{\hat{\eta}}(Z; g_0) v(X)] = \mathbb{E}[\mathbb{E}[s_{\hat{\eta}}(Z; g_0) \mid X] v(X)]\,. \tag{130}$$

Equation (130) shows that the remaining term in Eq. (128) can be controlled via bounding this conditional mean of the score in direction v.

Step 2: Decompose the conditional score into oracle and nuisance terms. Recall from Eqs. (12) and (11) that the fitted pseudo-outcome and weight are

$$\widehat{\mathcal{T}}_{\mathrm{LT}}(Z) = \mathbb{1}(R=0)\hat{\omega}(X)\hat{\tau}_{\mathrm{AIPW}}(Z) + \mathbb{1}(R=1)\hat{\omega}(X)\hat{\psi}_{\mathrm{obs}}(Z) + (\hat{\mu}_1(X) - \hat{\mu}_0(X))\hat{\Omega}(Z),$$
$$\hat{\omega}^*(Z) = \mathbb{1}(R=0)\hat{\omega}(X) + \hat{\Omega}(Z),$$

where $\hat{\omega}(X) = \omega(\hat{\pi}(X), \hat{\rho}(X))$ and $\hat{\Omega}(Z) = \Omega(Z; \hat{\eta})$; the AIPW and observation-correction terms are evaluated at the fitted nuisances as well. Substituting these definitions into Eq. (129) shows that the conditional score mean contains the conditional mean of the first two terms of $\widehat{\mathcal{T}}_{\mathrm{LT}}$, plus the weight-correction and target terms. We therefore first compute the combined mean in Eq. (134) below. Separating its oracle part from its nuisance errors will provide the B_μ and B_h terms in the score decomposition, Eq. (137), and hence in the derivative bound required by Eq. (128).

To expose the cancellation of first-order error in h between these first two terms, define $\hat{\gamma} = (1 - \hat{\rho}_s)/\hat{\rho}_s$ and $\gamma = (1 - \rho_s)/\rho_s$. Conditioning the AIPW term on $X, R = 0$ gives

$$\mathbb{E}[\hat{\tau}_{\mathrm{AIPW}} \mid X, R=0] = \tau + B_\mu(X) + \mathbb{E}\left[\frac{\pi_s - \hat{\pi}}{\hat{\pi}(1-\hat{\pi})}\delta_h \mid X, R=0\right], \tag{131}$$

$$B_\mu(X) = \delta_\pi \left(\frac{\delta_{\mu_1}}{\hat{\pi}} + \frac{\delta_{\mu_0}}{1-\hat{\pi}}\right). \tag{132}$$

The remaining term in Eq. (131) is first order in δ_h. To combine it with the long-term outcome correction, use $\mathbb{E}[Y - \hat{h} \mid S, X, R=1] = -\delta_h$ and the change of measure between the two samples:

$$\mathbb{E}[\mathbb{1}(R=1)\hat{\omega}\hat{\psi}_{\mathrm{obs}} \mid X] = -(1-\rho)\hat{\omega}\mathbb{E}\left[\frac{\hat{\gamma}}{\gamma}\frac{\hat{\pi}_s - \hat{\pi}}{\hat{\pi}(1-\hat{\pi})}\delta_h \mid X, R=0\right]. \tag{133}$$

Multiplying Eq. (131) by $(1-\rho)\hat{\omega}$ and adding Eq. (133) gives

$$\mathbb{E}[\mathbb{1}(R=0)\hat{\omega}\hat{\tau}_{\mathrm{AIPW}} + \mathbb{1}(R=1)\hat{\omega}\hat{\psi}_{\mathrm{obs}} \mid X] = (1-\rho)\hat{\omega}\{\tau + B_\mu(X) + B_h(X)\}, \tag{134}$$

$$B_h(X) = \mathbb{E}[\delta_h \mathcal{R}_h \mid X, R=0], \quad \text{with } \mathcal{R}_h = \frac{\pi_s - \hat{\pi} - (\hat{\gamma}/\gamma)(\hat{\pi}_s - \hat{\pi})}{\hat{\pi}(1-\hat{\pi})}. \tag{135}$$

To bound $\mathcal{R}_h$ in Eq. (135), use the exact identities

$$1 - \frac{\hat{\gamma}}{\gamma} = \frac{\delta_{\rho_s}}{\hat{\rho}_s(1-\rho_s)},$$
$$\mathcal{R}_h = \frac{-(\hat{\gamma}/\gamma)\delta_{\pi_s} + \{\delta_{\rho_s}/[\hat{\rho}_s(1-\rho_s)]\}(\pi_s - \hat{\pi})}{\hat{\pi}(1-\hat{\pi})}.$$

Assumption F.3 bounds $\hat{\gamma}/\gamma$ and the inverse denominators, while $|\pi_s - \hat{\pi}| \le 1$. The triangle inequality then gives

$$|\mathcal{R}_h| \le C(|\delta_{\pi_s}| + |\delta_{\rho_s}|). \tag{136}$$

Together, Eqs. (135) and (136) show that the surviving error in h is multiplied by another nuisance error. We use this product structure in Step 3.

We now insert the combined mean in Eq. (134) into the score definition in Eq. (129). Let $\bar{\Omega}(X) = \mathbb{E}[\hat{\Omega} \mid X]$ and $q_0(X) = (1-\rho)\omega$. The recalled definitions of $\widehat{\mathcal{T}}_{\mathrm{LT}}$ and $\hat{\omega}^*$ give the intermediate identity

$$\mathbb{E}[s_{\hat{\eta}}(Z; g_0) \mid X] = (1-\rho)\hat{\omega}(\tau + B_\mu + B_h) + (\hat{\mu}_1 - \hat{\mu}_0)\bar{\Omega} - q_{\hat{\eta}} g_0.$$

Using $q_{\hat{\eta}} = (1-\rho)\hat{\omega} + \bar{\Omega}$ and $\hat{\mu}_1 - \hat{\mu}_0 = \tau + \delta_{\mu_1} - \delta_{\mu_0}$, and then adding and subtracting $q_0(\tau - g_0)$, we obtain

$$\begin{aligned}\mathbb{E}[s_{\hat{\eta}}(Z; g_0) \mid X] = {} & q_0(\tau - g_0) + (q_{\hat{\eta}} - q_0)(\tau - g_0) \\ & + (\delta_{\mu_1} - \delta_{\mu_0})\bar{\Omega} + (1-\rho)\hat{\omega}(B_\mu + B_h).\end{aligned} \tag{137}$$

The first term in Eq. (137) is the oracle score mean. The other three terms are nuisance errors: a curvature error multiplied by the approximation residual, an outcome-regression error multiplied by the weight correction, and the two bias terms from Eq. (134). To bound the first two nuisance terms, note that

$$\bar{\Omega} = (1-\rho)\partial_\pi \omega(\hat{\pi}, \hat{\rho})(\pi - \hat{\pi}) + (1-\hat{\rho})\partial_\rho \omega(\hat{\pi}, \hat{\rho})(\rho - \hat{\rho}) \le C(|\delta_\pi| + |\delta_\rho|), \tag{138}$$

$$|q_{\hat{\eta}} - q_0| \le C(|\delta_\pi|^2 + |\delta_\rho|^2). \tag{139}$$

Equation (138) follows by bounding the weight derivatives. Equation (139) follows from the same conditional Taylor expansion as Lemma F.8. Bounded outcomes and Assumption F.1(iv) bound $|\tau - g_0|$, so the curvature-error term in Eq. (137) is also second order.

Step 3: Bound the score contribution to the directional derivative. Before bounding the nuisance terms in Eq. (137), we show that its oracle term has a nonpositive contribution to Eq. (130). Assumption F.4 gives the variational inequality

$$D_g \mathcal{L}_\omega^{\mathrm{oracle}}(g_0)[g - g_0] \ge 0, \qquad g \in \mathcal{G}. \tag{140}$$

Since $g_0 + t(g - g_0)$ is feasible for $t \in [0,1]$, its right derivative at the minimum $t = 0$ is nonnegative. Taking $g = \hat{g}$ in Eq. (140) and using $q_0 = (1-\rho)\omega$ gives

$$\mathbb{E}[q_0(\tau - g_0)v] = -\frac{P(R=0)}{2} D_g \mathcal{L}_\omega^{\mathrm{oracle}}(g_0)[v] \le 0. \tag{141}$$

Thus, we need only bound the three nuisance terms in Eq. (137). Denote their sum by

$$\mathcal{E}_\eta(X) := \mathbb{E}[s_{\hat{\eta}}(Z; g_0) \mid X] - q_0(X)(\tau(X) - g_0(X)). \tag{142}$$

Substituting Eq. (142) into Eq. (130), applying Eq. (141) to the oracle term, and then using Cauchy–Schwarz yields

$$\begin{aligned}-\tfrac{1}{2} D_g \mathcal{L}_\omega(g_0, \hat{\eta})[v] &= \mathbb{E}[q_0(\tau - g_0)v] + \mathbb{E}[\mathcal{E}_\eta v] \\ &\le \mathbb{E}[\mathcal{E}_\eta v] \le |\mathbb{E}[\mathcal{E}_\eta v]| \le \|\mathcal{E}_\eta\|_2 \|v\|_2.\end{aligned} \tag{143}$$

The oracle sign follows from convexity of $\mathcal{G}$ and optimality of g_0; neither v nor $\mathcal{E}_\eta$ needs to have a fixed sign. The final inequality is Cauchy–Schwarz under the pooled marginal law P_X.

It remains to control $\|\mathcal{E}_\eta\|_2$. Define $\|\delta_\mu\|_4 = \max_{a\in\{0,1\}} \|\hat{\mu}_a - \mu_a\|_4$. We use Hölder's product inequality in the form

$$\|fg\|_2^2 = \mathbb{E}[|f|^2|g|^2] \le (\mathbb{E}|f|^4)^{1/2}(\mathbb{E}|g|^4)^{1/2}, \qquad \|fg\|_2 \le \|f\|_4\|g\|_4,$$

which is Cauchy–Schwarz applied to $|f|^2$ and $|g|^2$. We also use $\|bf\|_2 \le \|b\|_\infty\|f\|_2$ for a bounded multiplier b. Applying these to Eqs. (139), (138), and (132), respectively, gives

$$\begin{aligned} \|(q_{\hat{\eta}} - q_0)(\tau - g_0)\|_2 &\le C(\|\delta_\pi\|_4^2 + \|\delta_\rho\|_4^2), \\ \|(\delta_{\mu_1} - \delta_{\mu_0})\bar{\Omega}\|_2 &\le C\|\delta_\mu\|_4(\|\delta_\pi\|_4 + \|\delta_\rho\|_4), \\ \|B_\mu\|_2 &\le C\|\delta_\mu\|_4\|\delta_\pi\|_4. \end{aligned} \tag{144}$$

Here, the first row uses the bounded multiplier $\tau - g_0$; the other two rows use Hölder's product bound for each treatment arm and sum the resulting bounds. No independence between nuisance errors is required.

For B_h in Eq. (135), conditional Jensen, the sampling-overlap bound in Assumption 3.2, and Hölder's inequality applied to the residual bound in Eq. (136) give

$$\begin{aligned} \|B_h\|_2^2 &\le \mathbb{E}_X\left[\mathbb{E}[\delta_h^2\mathcal{R}_h^2 \mid X, R=0]\right] \le \alpha_R^{-1}\mathbb{E}[\delta_h^2\mathcal{R}_h^2] \\ &\le C\mathbb{E}\left[\delta_h^2(\delta_{\pi_s}^2 + \delta_{\rho_s}^2)\right] \\ &\le C\|\delta_h\|_4^2\left(\|\delta_{\pi_s}\|_4^2 + \|\delta_{\rho_s}\|_4^2\right). \end{aligned} \tag{145}$$

Here and below, C may change between bounds but depends only on the stated regularity constants. Since $(1-\rho)\hat{\omega}$ is bounded, the decomposition in Eq. (137), the triangle inequality, and Eqs. (144)–(145) imply

$$\|\mathcal{E}_\eta\|_2 \le CA_\eta, \tag{146}$$

$$A_\eta := \|\delta_\mu\|_4(\|\delta_\pi\|_4 + \|\delta_\rho\|_4) + \|\delta_h\|_4(\|\delta_{\pi_s}\|_4 + \|\delta_{\rho_s}\|_4) + \|\delta_\pi\|_4^2 + \|\delta_\rho\|_4^2. \tag{147}$$

Substituting Eq. (146) into Eq. (143) now gives the bound on the directional derivative required by Step 1:

$$-D_g\mathcal{L}_\omega(g_0, \hat{\eta})[v] \le 2CA_\eta\|v\|_2. \tag{148}$$

Step 4: Complete the estimation-error bound. Insert Eq. (148) into the initial curvature bound, Eq. (128). Apply Young's inequality $2ab \le ta^2 + t^{-1}b^2$ with $a = \|v\|_2$, $b = CA_\eta$, and $t = \alpha/2 > 0$:

$$\alpha\|v\|_2^2 \le R_g + 2CA_\eta\|v\|_2 \le R_g + \frac{\alpha}{2}\|v\|_2^2 + \frac{2C^2}{\alpha}A_\eta^2. \tag{149}$$

Moving the squared-error term to the left gives $\|v\|_2^2 \le (2/\alpha)R_g + (4C^2/\alpha^2)A_\eta^2$. Finally, apply $(\sum_{j=1}^4 a_j)^2 \le 4\sum_{j=1}^4 a_j^2$ to the four nonnegative terms defining A_η in Eq. (147), then use $(a+b)^2 \le 2(a^2+b^2)$ for the two inner sums. Absorbing the resulting constants gives the final bound:

$$\|\hat{g} - g_0\|_2^2 \lesssim R_g + \|\delta_\mu\|_4^2(\|\delta_\pi\|_4^2 + \|\delta_\rho\|_4^2) + \|\delta_h\|_4^2(\|\delta_{\pi_s}\|_4^2 + \|\delta_{\rho_s}\|_4^2) + \|\delta_\pi\|_4^4 + \|\delta_\rho\|_4^4. \tag{150}$$

G DATASETS

G.1 SYNTHETIC DATA

Here, we take inspiration from Nie & Wager (2021); Morzywołek et al. (2023), and extend the data-generating process to a two-sample setting. The generic simulation steps are:

1. Sample belongingness R: First, we use the pre-defined true nuisance $\rho(x)$ to decide the belonging of the sample with covariate $X \sim P_X$. $R \in \{0, 1\}$ is randomly drawn from $\text{Ber}(\rho(X))$.
2. Treatment in $\mathcal{D}_1$: $A \mid X \sim \text{Ber}(\pi(X))$. Here π is the true treatment propensity score function.
3. Implicit treatment in observational data: $A \mid X \sim \text{Ber}(e(X))$. Here e is a pre-defined implicit treatment propensity score function.
4. Short-term outcome: $S = a(X) + (A - 0.5)\tau_S(X) + \delta$, with $\delta \sim \mathcal{N}(0, \sigma_S)$.
5. Long-term outcome in $\mathcal{D}_2$: $Y = b(X, S) + (A - 0.5)\tau_Y(X) + \epsilon$. Here ϵ is exogenous noise variable $\epsilon \sim \mathcal{N}(0, \sigma_Y)$. Under our surrogacy assumption ($A \perp\!\!\!\perp Y \mid S, X$), $\tau_Y^* = 0$. Then:

$$Y = b(S, X) + \epsilon. \tag{151}$$

The true heterogeneous effect of A on the long-term outcome Y is:

$$\tau(X) = \mathbb{E}\left[Y(1) - Y(0) \mid X, R = 0\right] = \mathbb{E}\left[b(X, S(1)) - b(X, S(0)) \mid X, R = 0\right]. \tag{152}$$

We now instantiate the above generic framework in analogy with Setup A of Nie & Wager (2021) (with surrogacy). Let $X \sim \text{Unif}([-1, 1]^{10})$, and define the sample belonging propensity score

$$\rho(X) = \eta_\rho + (1 - \eta_\rho)\,\sigma\big(X_1 + X_2 - \gamma_\rho\big), \tag{153}$$

where $\sigma(\cdot)$ is the logistic sigmoid and $\eta_\rho \in [0, 1)$ enforces a lower bound $\rho(X) \geq \eta_\rho$ (we use $\eta_\rho = 10^{-3}$). The parameter γ_ρ controls long-term outcome overlap: larger γ_ρ shifts ρ toward η_ρ and shrinks the observational sample.

The treatment propensity on the short-term dataset uses a standardized logit. Let

$$h(X) := X_1X_2 + X_3 + X_4 + X_8^2, \qquad \mu_h = \mathbb{E}[h(X)], \qquad \sigma_h = \sqrt{\text{Var}[h(X)]}, \tag{154}$$

under $X \sim \text{Unif}([-1, 1]^{10})$. Then

$$\pi(X) = \eta_\pi + (1 - 2\eta_\pi)\,\sigma\Big(\gamma_\pi \cdot \tfrac{h(X) - \mu_h}{\sigma_h}\Big), \tag{155}$$

with $\eta_\pi = 10^{-2}$ enforcing $\pi(X) \in [\eta_\pi, 1 - \eta_\pi]$. The parameter γ_π controls treatment overlap: larger γ_π pushes π toward the bounds. Standardizing $h(X)$ keeps the effective overlap difficulty comparable across choices of h.

For the observational dataset, the implicit treatment propensity is

$$e(X) = \text{trim}_{0.1}\big(\sigma(X_2 + X_3 + X_4)\big), \qquad \text{trim}_\eta(u) := \max\{\eta, \min(u, 1 - \eta)\}. \tag{156}$$

For the short-term outcome S, we set

$$a(X) = \sin(\pi X_1X_2) + 2(X_3 - 0.5)^2 + X_4 + 0.5X_5 + X_6, \tag{157}$$

and the short-term treatment effect $\tau_S(X) = 1 + \frac{X_1 + X_2 + X_3 + X_4}{4}$, so

$$S = a(X) + (A - 0.5)\,\tau_S(X) + \delta, \qquad \delta \sim \mathcal{N}(0, \sigma_S^2),\ \sigma_S = 0.2. \tag{158}$$

The long-term outcome regression is quadratic in S:

$$Y = b(X, S) + \epsilon = \sin(X_1X_2) + X_7^2 + X_8 + \tfrac{S^2}{4} + \epsilon, \qquad \epsilon \sim \mathcal{N}(0, 0.5^2). \tag{159}$$

Note: Outcomes are normalized by the sample mean and standard deviation.

Experimental regimes. We instantiate four regimes by varying $(\gamma_\pi, \gamma_\rho)$ at a fixed total sample size $n = 10000$ (drawn before the R-split):

Regime	γ_π	γ_ρ	n
$\mathcal{D}_*$ (healthy)	0	0	10000
$\mathcal{D}_t$ (low treatment overlap)	5	0	10000
$\mathcal{D}_o$ (low outcome overlap)	0	4	10000
$\mathcal{D}_{t+o}$ (dual overlap)	5	3	10000

G.2 Long-term earnings effects with GAIN and PSID

G.2.1 Basic information of GAIN job training dataset

The Greater Avenues for Independence (GAIN) study (Hotz et al., 2006) records randomized welfare-to-work assignment and subsequent earnings and welfare receipt. We use the four-site sample from Riverside, Alameda, Los Angeles, and San Diego in the GAIN–PSID application below. Access to the GAIN microdata can be obtained via ICPSR (study 38125, `https://www.icpsr.umich.edu/web/ICPSR/studies/38125`).

G.2.2 Basic information of PSID dataset

The Panel Study of Income Dynamics (PSID) (Panel Study of Income Dynamics, 2026; McGonagle et al., 2012) is a longitudinal study of U.S. families that began in 1968 and records demographics, earnings, employment, and welfare receipt across generations. We use observations from 1980–1993 from PSID-SHELF (Pfeffer et al., 2025), a harmonized longitudinal file covering 1968–2021, in the GAIN–PSID application below. The public-use PSID data are available through the PSID Data Center (`https://psidonline.isr.umich.edu/`).

G.2.3 Dataset setting and assumption diagnostics

We combine the randomized Greater Avenues for Independence (GAIN) data (Hotz et al., 2006) with the observational Panel Study of Income Dynamics (PSID) (McGonagle et al., 2012; Pfeffer et al., 2025; Panel Study of Income Dynamics, 2026). GAIN studies a randomized welfare-to-work program for welfare recipients in California, with data on treatment assignment, demographics, early earnings, employment, and welfare receipt. PSID is a longitudinal study of U.S. families that repeatedly measures demographics, earnings, employment, and welfare receipt, with long-term outcomes. The overlap of measures enables us to combine the two studies to estimate how GAIN assignment affects subsequent earnings in the GAIN population.

Experimental setup. We treat GAIN as $\mathcal{D}_1$ ($R = 0$), retaining (X, A, S), and PSID as $\mathcal{D}_2$ ($R = 1$), retaining (X, S, Y). Here, A denotes GAIN assignment, X contains pretreatment characteristics, and S contains early earnings, employment, and welfare receipt. We restricted PSID to working-age people comparable to GAIN participants and harmonized the shared variables at the annual level. The resulting datasets contain $19{,}170$ GAIN participants and $6{,}859$ PSID participants. The target outcome Y is average earnings over quarters 41–48 (years 11–12), beyond GAIN's nine-year follow-up, expressed in 1000 USD. The GAIN treatment is absent from PSID, whereas PSID records earnings beyond the GAIN observation horizon. This combination therefore supplies the complementary observed variables required by the two-sample setting without constructing a pseudo-ground-truth outcome.

Why the combination is plausible. Early earnings, employment, and welfare receipt capture pathways through which welfare-to-work training may affect subsequent earnings. Shared measurements and restriction to comparable working-age participants make the surrogate relationship a plausible bridge between the studies. We therefore use an earlier outcome available in both studies, Y_{cal}, defined as average earnings over quarters 33–36 (year 9), for assumption diagnostics and calibration.

Surrogacy and comparability diagnostics. We conducted a cross-fitted generalized covariance test of $A \perp\!\!\!\perp Y_{\mathrm{cal}} \mid S, X, R = 0$ in GAIN with three random seeds. The test did not reject the null for any seed (Table 8). This checks a residual-covariance implication of surrogacy at the calibration horizon; nonrejection does not establish conditional independence or surrogacy for the later target outcome.

In contrast, the comparability diagnostic for $R \perp\!\!\!\perp Y_{\mathrm{cal}} \mid S, X$ rejected the null for all three seeds at the 5% level.

Table 8: Assumption diagnostics using year-9 earnings Y_{cal}. Surrogacy is not rejected, whereas comparability is rejected at the 5% level for each seed.

	Surrogacy		Comparability	
Seed	z	p-value	z	p-value
42	-0.386	0.699	2.33	0.020
7	-0.301	0.764	2.59	0.0095
2026	0.232	0.816	2.38	0.017

Calibration for comparability violations. Following the bias decomposition in Eq. (45), we estimated the conditional mean gap

$$\widehat{\eta}_{\mathrm{cal}}(s, x) = \widehat{\mu}_{\mathrm{cal}}(s, x, R = 1) - \widehat{\mu}_{\mathrm{cal}}(s, x, R = 0), \tag{160}$$

where $\mu_{\mathrm{cal}}(s, x, R = r) = \mathbb{E}[Y_{\mathrm{cal}} \mid S = s, X = x, R = r]$, and applied the resulting comparability correction post hoc to every learner. GAIN calibration outcomes were used only for diagnostics and calibration, not to train the uncorrected learners. In terms of Eq. (45), the comparability correction subtracts the estimated contrast

$$c_{\mathrm{cal}}(x) = \mathbb{E}[\eta_{\mathrm{cal}}(S, X) \mid A = 1, X = x, R = 0] - \mathbb{E}[\eta_{\mathrm{cal}}(S, X) \mid A = 0, X = x, R = 0]$$

from each learner's prediction. Transferring this correction from year 9 to years 11–12 requires that the corresponding conditional mean gap remains approximately stable. It does not correct a residual surrogacy violation. Thus, the observed variable pattern matches the canonical setting, while causal interpretation still depends on these transport and surrogacy conditions.

H BASELINES

H.1 ANALOGY BETWEEN HTE LEARNERS AND HLTE LEARNERS

In this section, we briefly explain the analogy between standard HTE learners and their counterparts for HLTE estimation. We show how do we adapt the classic confounding adjustment strategies for HTE estimation to the long-term outcome setting.

HTE setting: In the standard HTE estimation setting, we have a single dataset $\mathcal{D} = \{(X_i, A_i, Y_i)\}_{i=1}^N$ where X are pre-treatment covariates, $A \in \{0, 1\}$ is the treatment, and Y is the outcome. The goal is to estimate the conditional average treatment effect (CATE)

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

Under *unconfoundedness* ($A \perp\!\!\!\perp \{Y(0), Y(1)\} \mid X$) and positivity, the CATE is identified from observed data via two equivalent expressions:

$$\tau(x) = \mu_1(x) - \mu_0(x) = \mathbb{E}\left[Y \cdot \frac{A - \pi(X)}{\pi(X)(1 - \pi(X))} \,\middle|\, X = x \right],$$

where $\mu_a(x) := \mathbb{E}[Y \mid A = a, X = x]$ and $\pi(x) := \Pr(A = 1 \mid X = x)$. The four standard meta-learners below correspond to different ways of plugging estimated nuisances $(\hat{\mu}_a, \hat{\pi})$ into these identities.

Plug-in estimator (T-learner (Künzel et al., 2019)): The T-learner directly plugs the estimated response surfaces $\hat{\mu}_a(x)$ into the first identification formula:

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

The estimator is simple but inherits the full estimation error of $\hat{\mu}_a$ (*plug-in bias* (Kennedy, 2022)).

RA-learner (Curth & van der Schaar, 2021): The RA-learner constructs a regression-adjusted pseudo-outcome

$$\mathcal{T}_{\mathrm{RA}} = A\left(Y - \hat{\mu}_0(X)\right) + (1 - A)\left(\hat{\mu}_1(X) - Y\right),$$

which has conditional mean $\tau(x)$ under unconfoundedness at the true response surfaces, and regresses $\mathcal{T}_{\mathrm{RA}}$ on X in the second stage.

IPW-learner (Curth & van der Schaar, 2021): The IPW-learner instantiates the second identification formula via the inverse-propensity-weighted pseudo-outcome

$$\mathcal{T}_{\mathrm{IPW}} = \left(\frac{A}{\hat{\pi}(X)} - \frac{1 - A}{1 - \hat{\pi}(X)} \right) Y,$$

and regresses $\mathcal{T}_{\mathrm{IPW}}$ on X. The estimator is unstable when $\hat{\pi}$ is close to 0 or 1 (low treatment overlap).

DR-learner (Kennedy, 2023): The DR-learner is based on the *efficient influence function* (EIF) of the ATE, which yields the AIPW pseudo-outcome

$$\mathcal{T}_{\mathrm{AIPW}} = \hat{\mu}_1(X) - \hat{\mu}_0(X) + \frac{A - \hat{\pi}(X)}{\hat{\pi}(X)(1 - \hat{\pi}(X))}\left(Y - \hat{\mu}_A(X)\right).$$

Because the Gateaux derivative of the EIF w.r.t. the nuisances vanishes at the truth (Chernozhukov et al., 2018), the resulting second-stage loss is Neyman-orthogonal in $(\hat{\mu}, \hat{\pi})$: first-order errors in either nuisance contribute only at second order to the estimator.

How we adapt to the HLTE setting: In the HLTE setting, the long-term outcome Y is unobserved in $\mathcal{D}_1$ (where the treatment A is observed). The bridge is the *surrogate index* $h(S, X) = \mathbb{E}[Y \mid S, X, R = 1]$, which under surrogacy and comparability (Assumptions 3.4–3.5) is an unbiased proxy for Y on $\mathcal{D}_1$ (Athey et al., 2026; Chen & Ritzwoller, 2023). Concretely:

- For the **non-orthogonal learners** (LT-T-learner, LT-RA-learner, LT-IPW-learner), we fit $\hat{h}$ on $\mathcal{D}_2$, substitute every occurrence of Y with $\hat{h}(S, X)$ in the pseudo-outcomes above, and run the second-stage regression on $\mathcal{D}_1$. The remaining nuisances $\hat{\mu}_a$ and $\hat{\pi}$ are fit on $\mathcal{D}_1$.

- For the **orthogonal LT-O-DR-learner**, we compute the EIF for the long-term ATE of Athey et al. (2026); Chen & Ritzwoller (2023) to get the pseudo-outcome $\mathcal{T}_{\mathrm{DR}}$. We recover LT-O-DR-learner as the special case $\omega \equiv 1$ of our LT-O-learners.

The non-orthogonal learners thus inherit *both* the plug-in/IPW limitations of the standard HTE estimators *and* the additional first-order error from $\hat{h}$. The LT-O-DR-learner removes the latter via orthogonality but remains sensitive to low treatment or outcome overlap, motivating our weighted variants in Section 5.2.

I IMPLEMENTATION

All experiments are run on Intel Core Ultra 7 155U (1.70 GHz) CPU and 32GB RAM. Data processing and model training are all implemented with Python. We provide the source code at https://anonymous.4open.science/r/Orthogonal-learners-for-Long-term-effects-4BF8.

Model architectures: We instantiate a separate multi-layer perceptron (MLP) for each sub-regression task. Neural networks that estimate propensity nuisance functions $(\rho, \rho_s, \pi, \pi_s)$ have two middle layers and a sigmoid output transformation at the end. Networks used to estimate outcome-related nuisances, as well as the second-stage predictor optimized under the meta-learner objective, consist of four hidden layers. All hidden layers use ReLU activations. Each learner instantiates the subset of these networks required by its algorithm.

Training: Hidden-layer widths are set to either $(20, 10)$ (two-layer MLPs) or $(20, 20, 10, 10)$ (four-layer MLPs). Nuisance networks are trained for 20 epochs with batch size 64 and learning rate 10^{-3}. The second-stage network is trained for 40 epochs with learning rate 10^{-3}.

Cross-fitting: We partition the unified dataset $\mathcal{D} = \{Z_i\}_{i=1}^N$ into $K = 5$ folds. For each held-out fold k, we fit the nuisances on the remaining $K - 1$ folds in a sequential order dictated by the definitions in Eq. (4). Specifically, $(\hat{\pi}, \hat{\pi}_s, \hat{\rho}, \hat{\rho}_s)$ are fit by direct regression, and $\hat{h}$ is fit on the $R = 1$ subsample as a regression of Y on (S, X). Since $\mu(a, x) = \mathbb{E}[h(S, X) \mid A = a, X = x, R = 0]$, we then form the pseudo-targets $\hat{h}(S_i, X_i)$ on the $R = 0$ subsample and fit $\hat{\mu}$ by regressing them on (A, X). We finally evaluate $\hat{\omega}_i^*$ and $\hat{\mathcal{T}}_i$ on the held-out fold k using the cross-fitted nuisances and aggregate across folds to solve Eq. (15).

J Optimality of the Overlap Weighting Scheme

In this section, we provide a semiparametric efficiency-based justification for the overlap weighting scheme $\omega(x) = \pi(x)^2(1-\pi(x))^2\,\rho(x)$ used in the LT-O-DO-learner. Our argument proceeds in three steps: (i) we derive the efficiency bound for the ω-weighted average long-term treatment effect (WALTE), which we denote V_ω^{eff}; (ii) we characterize the weight $\omega_{\text{opt}}(x)$ that minimizes V_ω^{eff} for the *sample* estimand; (iii) we show that the LT-O-DO weight $\omega(x) = \pi^2(1-\pi)^2\rho$ arises naturally as a principled, orthogonality-preserving relaxation of ω_{opt} via a structural lower bound on the variance functional $V(x)$.

J.1 Efficiency bound for the weighted long-term ATE

Recall the efficient influence function (EIF) for the WALTE, under the simplification that $\omega(\cdot)$ does not depend on the nuisances ρ or π:

$$\phi(Z;\eta) = \frac{1}{p_0\,D_\omega}\left[\mathbb{1}(R=0)\,\omega(X)\,\big(\hat\tau_{\text{AIPW}}(Z;\eta) - \tau_\omega^0\big) + \mathbb{1}(R=1)\,\omega(X)\,\psi_{\text{obs}}(Z;\eta)\right], \quad (161)$$

where $p_0 = P(R=0)$, $D_\omega = \mathbb{E}[\omega(X) \mid R=0]$, and the components are

$$\hat\tau_{\text{AIPW}}(Z;\eta) = \mu(1,X) - \mu(0,X) + \frac{A}{\pi(X)}\big(h(S,X) - \mu(1,X)\big) - \frac{1-A}{1-\pi(X)}\big(h(S,X) - \mu(0,X)\big), \quad (162)$$

$$\psi_{\text{obs}}(Z;\eta) = \frac{1-\rho_s(S,X)}{\rho_s(S,X)}\left(\frac{\pi_s(S,X)-\pi(X)}{\pi(X)(1-\pi(X))}\right)\big(Y - h(S,X)\big). \quad (163)$$

Proposition J.1 (Efficiency bound for the WALTE)**.** *Let* $\sigma^2_{h|a}(x) := \text{Var}(h(S,X) \mid A=a, X=x, R=0)$ *and* $\sigma^2_Y(s,x) := \text{Var}(Y \mid S=s, X=x, R=1)$*. Then, the semiparametric efficiency bound for the* ω*-weighted long-term ATE is*

$$V_\omega^{\text{eff}} = \frac{1}{p_0 D_\omega^2}\,\mathbb{E}\left[\omega(X)^2\left(\frac{\sigma^2_{h|1}(X)}{\pi(X)} + \frac{\sigma^2_{h|0}(X)}{1-\pi(X)} + \big(\tau^0(X) - \tau_\omega^0\big)^2\right)\,\Big|\,R=0\right]$$
$$+ \frac{1-p_0}{p_0^2 D_\omega^2}\,\mathbb{E}\left[\omega(X)^2\left(\frac{1-\rho_s(S,X)}{\rho_s(S,X)}\right)^2\left(\frac{\pi_s(S,X)-\pi(X)}{\pi(X)(1-\pi(X))}\right)^2\sigma^2_Y(S,X)\,\Big|\,R=1\right]. \quad (164)$$

Proof. The efficiency bound is $V_\omega^{\text{eff}} = \mathbb{E}[\phi(Z;\eta)^2]$. By squaring Eq. (161), the cross-term vanishes since $\mathbb{1}(R=0)\cdot\mathbb{1}(R=1) = 0$, thus leaving two terms.

Experimental term ($R=0$). Decompose $\hat\tau_{\text{AIPW}} - \tau_\omega^0 = (\hat\tau_{\text{AIPW}} - \tau^0(X)) + (\tau^0(X) - \tau_\omega^0)$. At true nuisances, $\mathbb{E}[\hat\tau_{\text{AIPW}} \mid X, R=0] = \tau^0(X)$, so the cross-term vanishes upon conditioning. The conditional variance is computed using $A \mid X, R=0 \sim \text{Ber}(\pi(X))$, yielding

$$\text{Var}(\hat\tau_{\text{AIPW}} \mid X, R=0) = \frac{\sigma^2_{h|1}(X)}{\pi(X)} + \frac{\sigma^2_{h|0}(X)}{1-\pi(X)}. \quad (165)$$

Observational term ($R=1$). Since $h(S,X) = \mathbb{E}[Y \mid S, X, R=1]$, we have $\mathbb{E}[Y - h(S,X) \mid S, X, R=1] = 0$, so conditioning on $(S, X, R=1)$:

$$\mathbb{E}[\psi^2_{\text{obs}} \mid S, X, R=1] = \left(\frac{1-\rho_s}{\rho_s}\right)^2\left(\frac{\pi_s - \pi}{\pi(1-\pi)}\right)^2\sigma^2_Y(S,X). \quad (166)$$

Combining and accounting for the prefactors $1/(p_0 D_\omega)^2$ and the marginal probabilities $P(R=0) = p_0$, $P(R=1) = 1-p_0$ yields Eq. (164). □

J.2 Optimal weight for the sample WALTE

Following Crump et al. (2006), we target the *sample* WALTE $\tau_{\omega,S} = \sum_i \tau^0(X_i)\omega(X_i)/\sum_i \omega(X_i)$ rather than its population counterpart. Conditioning on the target-sample covariates removes the heterogeneity contribution $(\tau^0(X) - \tau^0_\omega)^2$ from Eq. (164), since the sample estimand absorbs τ^0_ω as a sample average rather than a population functional.

To unify the two terms, we use the density ratio identity $(1 - p_0)\, f_{X|R=1}(x) = p_0\, f_{X|R=0}(x)\, \rho(x)/(1 - \rho(x))$ to rewrite the observational expectation under the experimental measure. For this, we define

$$\Sigma_o(x) := \mathbb{E}\left[\left(\frac{1-\rho_s}{\rho_s}\right)^2 \left(\frac{\pi_s - \pi}{\pi(1-\pi)}\right)^2 \sigma^2_Y(S, X) \,\middle|\, X, R = 1\right]. \tag{167}$$

The objective becomes

$$J[\omega] = p_0^{-1} \cdot \mathbb{E}\big[\omega(X)^2 V(X) \mid R = 0\big], \qquad V(x) := \underbrace{\frac{\sigma^2_{h|1}(x)}{\pi(x)} + \frac{\sigma^2_{h|0}(x)}{1-\pi(x)}}_{\Sigma_t(x)} + \underbrace{\frac{\rho(x)}{1-\rho(x)}\,\Sigma_o(x)}_{\text{outcome overlap cost}}, \tag{168}$$

This is the variance criterion with the covariate-dependent weight held fixed, rather than the full efficiency bound for a nuisance-dependent weight in Eq. (33). It is minimized subject to the normalization $D_\omega = \mathbb{E}[\omega(X) \mid R = 0] = 1$.

Proposition J.2 (Efficiency-optimal weight)**.** *Suppose $V(X) > 0$ almost surely and $0 < \mathbb{E}[V(X)^{-1} \mid R = 0] < \infty$. The weight minimizing $J[\omega]$ subject to $D_\omega = 1$ is*

$$\omega_{\mathrm{opt}}(x) = \frac{V(x)^{-1}}{\mathbb{E}[V(X)^{-1} \mid R = 0]} \;\propto\; V(x)^{-1}, \tag{169}$$

and the corresponding minimum efficiency bound is $V^{\mathit{eff}}_{\omega_{\mathrm{opt}}} = \big(p_0\, \mathbb{E}[V(X)^{-1} \mid R = 0]\big)^{-1}$.

Proof. Form the Lagrangian $\mathcal{L}[\omega, \lambda] = p_0^{-1}\, \mathbb{E}[\omega(X)^2 V(X) \mid R = 0] - \lambda\big(\mathbb{E}[\omega(X) \mid R = 0] - 1\big)$. Taking the functional derivative under perturbation $\omega \to \omega + \varepsilon\eta$,

$$\left.\frac{\mathrm{d}\mathcal{L}}{\mathrm{d}\varepsilon}\right|_{\varepsilon=0} = 2p_0^{-1}\, \mathbb{E}[\eta(X)\omega(X)V(X) \mid R = 0] - \lambda\, \mathbb{E}[\eta(X) \mid R = 0] = 0 \quad \forall\, \eta. \tag{170}$$

By the fundamental lemma of the calculus of variations, $2p_0^{-1}\, \omega(x)V(x) = \lambda$ a.s., which yields $\omega_{\mathrm{opt}}(x) \propto V(x)^{-1}$. The constant is fixed by $\mathbb{E}[\omega_{\mathrm{opt}}(X) \mid R = 0] = 1$. □

J.3 Lower bound on $V(x)$ and the LT-O-DO weight

The optimal weight ω_{opt} depends on *all* nuisance functions $(\sigma^2_{h|a}, \pi, \rho, \rho_s, \pi_s, \sigma^2_Y)$ through $V(x)$, so its plug-in estimation would require additional corrections beyond those for (π, ρ) in the present loss. We now derive a structural lower bound on $V(x)$ that depends on π and ρ in a closed form, motivating the LT-O-DO weight $\omega(x) = \pi^2(1-\pi)^2\rho$ as a principled approximation to ω_{opt}.

Treatment overlap term. Since $\sigma^2_{h|1}(1-\pi) + \sigma^2_{h|0}\pi \geq \min\{\sigma^2_{h|1}, \sigma^2_{h|0}\}$,

$$\Sigma_t(x) = \frac{\sigma^2_{h|1}(1-\pi) + \sigma^2_{h|0}\pi}{\pi(1-\pi)} \;\geq\; \frac{C_t(x)^2}{\pi(x)(1-\pi(x))}, \qquad C_t(x)^2 := \min\big\{\sigma^2_{h|1}(x), \sigma^2_{h|0}(x)\big\}. \tag{171}$$

Outcome overlap term. We separate the π-dependent factor from the surrogate-dependent factor by writing

$$\Sigma_o(x) = \frac{1}{\pi(x)^2(1-\pi(x))^2}\, \mathbb{E}\left[\left(\tfrac{1-\rho_s}{\rho_s}\right)^2 \tilde{m}(S, X) \,\middle|\, X, R = 1\right], \qquad \tilde{m}(s, x) := (\pi_s(s,x) - \pi(x))^2\, \sigma^2_Y(s, x) \geq 0. \tag{172}$$

Define $\bar{m}(x)$ as the ratio of the weighted expectation below to $\mathbb{E}[((1-\rho_s)/\rho_s)^2 \mid X=x, R=1]$, assumed finite and positive. It is a nonnegative weighted average of $\tilde{m}$, so

$$\mathbb{E}\left[\left(\tfrac{1-\rho_s}{\rho_s}\right)^2 \tilde{m}(S,X) \,\middle|\, X, R=1\right] = \bar{m}(x) \cdot \mathbb{E}\left[\left(\tfrac{1-\rho_s}{\rho_s}\right)^2 \,\middle|\, X, R=1\right]. \tag{173}$$

Applying Jensen's inequality to $t \mapsto t^2$,

$$\mathbb{E}\left[\left(\tfrac{1-\rho_s}{\rho_s}\right)^2 \,\middle|\, X, R=1\right] \geq \left(\mathbb{E}\left[\tfrac{1-\rho_s}{\rho_s} \,\middle|\, X, R=1\right]\right)^2. \tag{174}$$

The first moment admits an exact identity. Using $f(s|x, R=1) = \rho_s(s,x)\, f(s|x) \,/\, \rho(x)$ and the tower law $\mathbb{E}[\rho_s(S,X) \mid X] = \mathbb{E}[R \mid X] = \rho(x)$, we yield

$$\mathbb{E}\left[\tfrac{1-\rho_s}{\rho_s} \,\middle|\, X, R=1\right] = \frac{1}{\rho(x)} \int (1-\rho_s(s,x))\, f(s|x)\, \mathrm{d}s = \frac{1-\rho(x)}{\rho(x)}. \tag{175}$$

Combining Eq. (173)–Eq. (175), we arrive at

$$\frac{\rho(x)}{1-\rho(x)}\, \Sigma_o(x) \geq \frac{\rho}{1-\rho} \cdot \frac{1}{\pi^2(1-\pi)^2} \cdot \frac{(1-\rho)^2}{\rho^2}\, \bar{m}(x) = \frac{(1-\rho(x))\, \bar{m}(x)}{\rho(x)\, \pi(x)^2 (1-\pi(x))^2} = \frac{C(x)^2}{\rho(x)\, \pi(x)^2 (1-\pi(x))^2}, \tag{176}$$

where we define the surrogate informativeness factor $C(x)^2 := (1-\rho(x))\, \bar{m}(x)$.

Combined lower bound. Adding Eq. (171) and Eq. (176),

$$\boxed{V(x) \;\geq\; \frac{C_t(x)^2}{\pi(x)(1-\pi(x))} + \frac{C(x)^2}{\rho(x)\, \pi(x)^2 (1-\pi(x))^2}.} \tag{177}$$

Justification of $\omega(x) = \pi^2(1-\pi)^2\rho$. With $c_V = \{\mathbb{E}[V(X)^{-1} \mid R=0]\}^{-1}$, Eq. (177) gives

$$\omega_{\mathrm{opt}}(x) \;\leq\; c_V \left(\frac{C_t(x)^2}{\pi(1-\pi)} + \frac{C(x)^2}{\rho\, \pi^2 (1-\pi)^2} \right)^{-1}. \tag{178}$$

If $C(x)^2$ remains bounded away from zero and $C_t(x)^2$ remains bounded in a joint low-overlap regime, the outcome-overlap term dominates this lower bound. Its inverse contributes the factor $\rho\pi^2(1-\pi)^2$. Omitting the unknown variance multiplier motivates the structural choice for the LT-O-DO weight:

$$\omega(x) = \pi(x)^2 (1-\pi(x))^2\, \rho(x). \tag{179}$$

This weight depends only on (π, ρ) and therefore preserves Neyman-orthogonality of the LT-O-learner framework. The squared treatment-overlap factor $\pi^2(1-\pi)^2$ arises because poor treatment overlap amplifies the outcome overlap cost via the $(\pi_s - \pi)^2/[\pi(1-\pi)]^2$ factor in Σ_o.

K LINEAR PARAMETRIC INSTANTIATION: CLOSED FORM AND ASYMPTOTIC VARIANCE

We instantiate the LT-O-learner with the linear parametric class $\mathcal{G} = \{g(x) = \theta^\top x : \theta \in \mathbb{R}^d\}$. This case admits a closed-form solution and yields an explicit asymptotic-variance comparison with the LT-O-DR-learner. Throughout, $n = N = N_1 + N_2$ is the pooled sample size, $\widehat{\omega}_i^* := \omega^*(Z_i; \hat{\eta})$ and $\widehat{\mathcal{T}}_i := \mathcal{T}_{\mathrm{LT}}(Z_i; \hat{\eta})$ are the plug-in versions of Eq. (11) and Eq. (12).

K.1 CLOSED-FORM ESTIMATOR

Proposition K.1 (Closed-form estimator)**.** *Let* $\mathbf{X} \in \mathbb{R}^{n\times d}$ *stack* $X_i^\top$ *as rows,* $\widehat{W} = \mathrm{diag}(\widehat{\omega}_1^*, \ldots, \widehat{\omega}_n^*)$*, and* $\widehat{\mathcal{T}} = (\widehat{\mathcal{T}}_1, \ldots, \widehat{\mathcal{T}}_n)^\top$*. If* $\mathbf{X}^\top \widehat{W}\, \mathbf{X} \succ 0$*, the empirical loss in Eq. (15) is strictly convex, and its unique minimizer is*

$$\widehat{\theta} = \big(\mathbf{X}^\top \widehat{W}\, \mathbf{X}\big)^{-1} \mathbf{X}^\top \widehat{\mathcal{T}} = \Big(\sum_{i=1}^n \widehat{\omega}_i^* X_i X_i^\top\Big)^{-1} \Big(\sum_{i=1}^n \widehat{\mathcal{T}}_i X_i\Big). \tag{180}$$

Proof. Substituting $g(X_i) = \theta^\top X_i$ into Eq. (15), we yield $\widehat{L}_\omega(\theta) = \theta^\top \widehat{H}\theta - 2\theta^\top \widehat{b}$ with $\widehat{H} = n^{-1}\sum_i \widehat{\omega}_i^* X_i X_i^\top$ and $\widehat{b} = n^{-1}\sum_i \widehat{\mathcal{T}}_i X_i$. The Hessian $\nabla_\theta^2 \widehat{L}_\omega = 2\widehat{H}$ is positive definite by assumption, so the first-order condition $\widehat{H}\theta = \widehat{b}$ has a unique solution that is the global minimizer. □

Remark K.2 (Sample-level positivity). For LT-O-LO ($\omega^* = (R-\rho)^2$) and LT-O-DR ($\omega^* = \mathbf{1}(R = 0)$), $\widehat{\omega}_i^* \geq 0$ a.s., so $\mathbf{X}^\top \widehat{W} \mathbf{X} \succeq 0$, and strict positivity follows from the usual full-rank condition on $\{X_i\}_{i:\widehat{\omega}_i^* > 0}$. For LT-O-TO and LT-O-DO, however, the $R = 0$ component carries the residual term Ω and is not sample-wise sign-definite: LT-O-TO has $\mathbb{1}(R=0)\big[\pi^2(1-\pi)^2 + 2\pi(1-\pi)(1-2\pi)(A-\pi)\big]$, and the $R = 0$ component of LT-O-DO is $\mathbb{1}(R=0)\,\rho\big[\rho\,\pi^2(1-\pi)^2 + 2\pi(1-\pi)(1-2\pi)(A-\pi)\big]$. Assumption K.4(i),(v) ensures positive empirical curvature with probability tending to one. In finite samples, a ridge penalty $\lambda\|\theta\|^2$ ensures positive curvature when $\lambda > -\lambda_{\min}(\widehat{H})$.

Remark K.3 (WLS interpretation). When $\widehat{\omega}_i^* > 0$ for every i, the loss can be completed into a square:

$$\widehat{L}_\omega(\theta) = \frac{1}{n}\sum_{i=1}^n \widehat{\omega}_i^* \big(\widetilde{Y}_i - X_i^\top \theta\big)^2 - \frac{1}{n}\sum_{i=1}^n \frac{\widehat{\mathcal{T}}_i^2}{\widehat{\omega}_i^*}, \qquad \widetilde{Y}_i := \widehat{\mathcal{T}}_i / \widehat{\omega}_i^*, \tag{181}$$

and $\widehat{\theta}$ is the WLS fit of $\widetilde{Y}_i$ on X_i with weights $\widehat{\omega}_i^*$. This applies to LT-O-DO when $\widehat{\omega}_i^* > 0$ for all i. For LT-O-DR and LT-O-TO, however, $R = 1$ observations have $\widehat{\omega}_i^* = 0$ but $\widehat{\mathcal{T}}_i \neq 0$, so $\widetilde{Y}_i$ is undefined and the WLS reduction is invalid. The correct unified statement is that $\widehat{\theta}$ solves the linear system $\widehat{H}\theta = \widehat{b}$, which reduces to WLS only in the strictly positive-weight regime.

K.2 ASYMPTOTIC VARIANCE

We invoke the standard Z-estimator framework (van der Vaart, 1998) with cross-fitting (Chernozhukov et al., 2018). Define the score

$$m(Z; \theta, \eta) = \omega^*(Z; \eta)\, XX^\top \theta - \mathcal{T}_{\mathrm{LT}}(Z; \eta)\, X. \tag{182}$$

Assumption K.4 (Z-estimation regularity)**.** Suppose $\tau_0(x) = \theta_0^\top x$ with $\theta_0 \in \mathbb{R}^d$, d fixed, and that the nuisance estimator $\hat{\eta}$ is constructed via K-fold cross-fitting. For independent draws from the pooled law, under the identification assumptions in Section 3, Assumptions F.1 and F.2, and the nuisance-neighborhood condition in Lemma F.8, assume:

(i) *Invertibility:* $J(\omega) := \mathbb{E}[\omega^*(Z; \eta_0)\, XX^\top] \succ 0$.

(ii) *Bounded second moments:* $\mathbb{E}[\|X\|^4] < \infty$ and $\mathbb{E}[\xi(Z; \eta_0)^2 \|X\|^2] < \infty$, where $\xi(Z; \eta_0) := \mathcal{T}_{\mathrm{LT}}(Z; \eta_0) - \omega^*(Z; \eta_0)\, \tau_0(X)$.

(iii) *Stochastic equicontinuity:* $\big\|m(\cdot; \theta_0, \hat{\eta}) - m(\cdot; \theta_0, \eta_0)\big\|_{P,2} = o_p(1)$.

(iv) *Orthogonality remainder;* $\sqrt{n}\,\big\|\mathbb{E}\big[m(Z;\theta_0,\hat{\eta}) - m(Z;\theta_0,\eta_0) \mid \hat{\eta}\big]\big\| = o_p(1)$.

(v) *Convergence of* $\widehat{J}$*:* Let $\widehat{J} := \mathbb{P}_n[\omega^*(Z;\hat{\eta})XX^\top]$. Then $\|\widehat{J} - J(\omega)\|_{\mathrm{op}} = o_p(1)$.

Conditions (iii)–(v) are the standard Z-estimator conditions; (iv) is implied by Theorem 5.1 (Neyman-orthogonality) when nuisance products are $o_p(n^{-1/2})$. We state them explicitly because the quasi-oracle rate alone does not control the empirical process and bias remainders of the moment equation.

Lemma K.5 (Asymptotic distribution)**.** *Under Assumption K.4,*

$$\sqrt{n}\big(\widehat{\theta} - \theta_0\big) \xrightarrow{d} \mathcal{N}\big(0,\ \mathcal{V}(\omega)\big), \qquad \mathcal{V}(\omega) = J(\omega)^{-1}\Sigma(\omega)J(\omega)^{-1}, \tag{183}$$

with

$$J(\omega) = \mathbb{E}[(1-\rho(X))\,\omega(X)\,XX^\top], \qquad \Sigma(\omega) = \mathbb{E}[V_\omega^\star(X)\,XX^\top], \tag{184}$$

and

$$V_\omega^\star(X) = \mathrm{Var}\big(\mathcal{T}_{\mathrm{LT}}(Z;\eta_0) - \omega^*(Z;\eta_0)\tau_0(X) \mid X\big). \tag{185}$$

Proof. At the truth, $m(Z;\theta_0,\eta_0) = -X\,\xi(Z;\eta_0)$. The two identities $\mathbb{E}[\omega^* \mid X] = (1-\rho)\omega$ and $\mathbb{E}[\mathcal{T}_{\mathrm{LT}} \mid X] = (1-\rho)\omega\,\tau_0$ give $\mathbb{E}[\xi \mid X] = 0$, hence $\mathbb{E}[m(Z;\theta_0,\eta_0)] = 0$ and $\Sigma(\omega) = \mathbb{E}[\xi^2\,XX^\top] = \mathbb{E}[V_\omega^\star(X)\,XX^\top]$. Linearity of m in θ gives $\nabla_\theta m = \omega^* XX^\top$, and by the tower property, $\mathbb{E}[\nabla_\theta m] = J(\omega)$. The estimating equation $n^{-1}\sum_i m(Z_i;\widehat{\theta},\hat{\eta}) = 0$, combined with linearity and Assumption K.4(iii)–(iv), yields the asymptotic linearization

$$\sqrt{n}(\widehat{\theta} - \theta_0) = J(\omega)^{-1}\,\frac{1}{\sqrt{n}}\sum_{i=1}^{n} X_i\,\xi(Z_i;\eta_0) + o_p(1), \tag{186}$$

and the central-limit theorem under Assumption K.4(ii) gives the stated normal limit. □

K.3 Variance comparison in low-overlap regimes

In this section, we prove Theorem 5.3.

Direct algebra yields explicit forms for $V_\omega^\star(X)$ in two canonical cases.

Lemma K.6 (Conditional residual variances)**.** *Define* $\Psi(X) := \mathbb{E}\Big[\dfrac{(1-\rho_s)^2(\pi_s-\pi)^2(Y-h)^2}{\rho_s^2} \,\Big|\, X, R=1\Big]$ *and* $\sigma_h^2(a,X) := \mathrm{Var}(h(S,X) \mid A=a, X, R=0)$*. Then*

$$V_\omega^\star(X) \;:=\; \mathrm{Var}\big(\xi(Z;\eta_0) \,\big|\, X\big) \;=\; \omega(X)^2\big[(1-\rho(X))\,\Sigma_t(X) + \rho(X)\,\Sigma_o(X)\big], \tag{187}$$

where

$$\Sigma_t(X) := \frac{\sigma_h^2(1,X)}{\pi(X)} + \frac{\sigma_h^2(0,X)}{1-\pi(X)},$$

$$\Sigma_o(X) := \mathbb{E}\left[\left(\tfrac{1-\rho_s(S,X)}{\rho_s(S,X)}\right)^2\left(\tfrac{\pi_s(S,X)-\pi(X)}{\pi(X)(1-\pi(X))}\right)^2\sigma_y(S,X)\,\middle|\, X, R=1\right].$$

Specifically,

$$V_1^\star(X) = (1-\rho)\left[\frac{\sigma_h^2(1,X)}{\pi} + \frac{\sigma_h^2(0,X)}{1-\pi}\right] + \frac{\rho\,\Psi(X)}{\pi^2(1-\pi)^2}, \tag{188}$$

$$V_\omega^\star(X) = (1-\rho)\rho^2\pi(1-\pi)\big[(1-\pi)\sigma_h^2(1,X) + \pi\sigma_h^2(0,X)\big] + \rho^3\Psi(X), \tag{189}$$

the second formula corresponding to $\omega(x) = \pi(x)(1-\pi(x))\rho(x)$*. The theorem shows that the weight* $\pi(x)(1-\pi(x))\rho(x)$ *is already sufficient to resolve the overlap issue; stability of the LT-O-DO-learner with the stronger weight* $\omega = \pi^2(1-\pi)^2\rho$ *used in the main text follows naturally.*

Proof. Substituting the definitions

$$\mathcal{T}_{\mathrm{LT}}(Z;\eta_0) = \mathbf{1}(R{=}0)\,\omega(X)\,\hat{\tau}_{\mathrm{AIPW}}(Z;\eta_0) + \mathbf{1}(R{=}1)\,\omega(X)\,\psi_{\mathrm{obs}}(Z;\eta_0) + \tau_0(X)\,\Omega(Z;\eta_0), \tag{190}$$

$$\omega^*(Z;\eta_0) = \mathbf{1}(R{=}0)\,\omega(X) + \Omega(Z;\eta_0), \tag{191}$$

into ξ, and using $\tau_0(X) = \mu(1,X) - \mu(0,X)$,

$$\begin{aligned}\xi(Z;\eta_0) &= \mathbf{1}(R{=}0)\,\omega(X)\,\hat{\tau}_{\mathrm{AIPW}} + \mathbf{1}(R{=}1)\,\omega(X)\,\psi_{\mathrm{obs}} + \tau_0(X)\,\Omega \\ &\quad - \mathbf{1}(R{=}0)\,\omega(X)\,\tau_0(X) - \Omega\,\tau_0(X) \\ &= \mathbf{1}(R{=}0)\,\omega(X)\big[\hat{\tau}_{\mathrm{AIPW}}(Z;\eta_0) - \tau_0(X)\big] \;+\; \mathbf{1}(R{=}1)\,\omega(X)\,\psi_{\mathrm{obs}}(Z;\eta_0). \end{aligned} \tag{192}$$

The two Ω terms cancel exactly. Then, the derivation is similar to the standard DR variance decomposition in Appendix F.3. For simplicity, we omit the details. □

Remark K.7 (Tower-law identity for Ψ). Using $f(s \mid X, R=1) = f(s \mid X)\rho_s/\rho$ and Jensen, $\mathbb{E}[\rho_s^{-1} \mid X, R=1] = \rho^{-1}$ and $\mathbb{E}[\rho_s^{-2} \mid X, R=1] \geq \rho^{-2}$. Thus $\Psi(X) \asymp \rho^{-2}$ (here and below, $A \asymp B$ means $cB \leq A \leq CB$ for some constants $0 < c \leq C < \infty$) holds whenever $|\pi_s - \pi|$ and $\mathbb{E}[(Y-h)^2 \mid X, S, R=1]$ are bounded above and below and ρ_s degenerates uniformly with ρ (i.e. $\rho_s \asymp \rho$). We assume this explicitly below.

Assumption K.8 (Low-overlap sequence)**.** There exist constants $c_0, c_1, c_3, c_4, p_0, p_1, \bar{\sigma}, \bar{\Psi} > 0$, $\lambda_R, \lambda_B > 0$, $\varepsilon_0 > 0$, and a covariate bound $C_X < \infty$ such that, for all $\varepsilon \leq \varepsilon_0$:

(a) *Bounded covariates and nonvanishing experimental-sample mass:* $\|X\| \leq C_X$ and $\rho_\varepsilon \leq 1 - c_3$ a.s. under P_ε.

(b) *Bulk overlap:* there exists $\mathcal{B}$ with $P_\varepsilon(\mathcal{B}) \geq p_1$ on which $\pi_\varepsilon(1-\pi_\varepsilon)\rho_\varepsilon \geq c_0$, and $\mathbb{E}_\varepsilon[XX^\top \mathbf{1}\{\mathcal{B}\}] \succeq \lambda_B I$ uniformly.

(c) *Low-overlap region:* there exists $\mathcal{R}_\varepsilon$ with $P_\varepsilon(\mathcal{R}_\varepsilon) \geq p_0$, $\mathbb{E}_\varepsilon[XX^\top \mathbf{1}\{\mathcal{R}_\varepsilon\}] \succeq \lambda_R I$ uniformly, and on $\mathcal{R}_\varepsilon$ either $\pi_\varepsilon(1-\pi_\varepsilon) \asymp \varepsilon$ (low-TO scenario), or $\rho_\varepsilon \asymp \varepsilon$ with $\pi_\varepsilon(1-\pi_\varepsilon) \geq c_4$ (low-LO scenario)

(d) *Bounded conditional variances:* $c_1 \leq \sigma^2_{h,\varepsilon}(a,X) \leq \bar{\sigma}^2$ uniformly.

(e) *Surrogate-propensity scaling:* in the low-LO scenario, $\Psi_\varepsilon(X) \asymp \rho_\varepsilon(X)^{-2}$ on $\mathcal{R}_\varepsilon$; in the low-TO scenario, Ψ_ε is bounded above on $\mathcal{R}_\varepsilon$. In both scenarios, $\rho_\varepsilon(X)^3\,\Psi_\varepsilon(X) \leq \bar{\Psi}$ a.s. under P_ε.

We call ε in condition (c) the **overlap score** as it describes the severity of the overlap issue. We first take $n \to \infty$ for each fixed ε, then compare the asymptotic variances as $\varepsilon \to 0^+$.

Theorem 5.3 is a direct consequence of the following more explicit statement.

Theorem K.9 (Variance stabilization, detailed)**.** *Under Assumption K.8, with $\omega_\varepsilon = \pi_\varepsilon(1-\pi_\varepsilon)\rho_\varepsilon$, for every fixed $v \in \mathbb{R}^d \setminus \{0\}$,*

$$v^\top \mathcal{V}_\varepsilon(\mathbf{1})\,v \;\underset{\varepsilon\downarrow 0}{\longrightarrow}\; \infty, \qquad \limsup_{\varepsilon\downarrow 0}\; v^\top \mathcal{V}_\varepsilon(\omega_\varepsilon)\,v \;<\; \infty. \tag{193}$$

In particular, $v^\top \mathcal{V}_\varepsilon(\mathbf{1})\,v \,/\, v^\top \mathcal{V}_\varepsilon(\omega_\varepsilon)\,v \to \infty$ as $\varepsilon \downarrow 0$.

Proof. Substituting $V_\varepsilon(\omega) = J_\varepsilon(\omega)^{-1}\Sigma_\varepsilon(\omega)J_\varepsilon(\omega)^{-1}$:

$$v^\top \mathcal{V}_\varepsilon(\omega)\,v = \big(J_\varepsilon(\omega)^{-1}v\big)^\top \Sigma_\varepsilon(\omega)\big(J_\varepsilon(\omega)^{-1}v\big) = \mathbb{E}_\varepsilon\big[V^\star_{\omega,\varepsilon}(X)\,(X^\top u_\varepsilon)^2\big], \tag{194}$$

with $u_\varepsilon := J_\varepsilon(\omega)^{-1}v$. We bound this quantity for $\omega = \mathbf{1}$ and $\omega = \omega_\varepsilon$ separately.

Step 1 (Regularity of J_ε). For $\omega \in \{\mathbf{1}, \omega_\varepsilon\}$ we have $0 \leq (1-\rho_\varepsilon)\omega \leq 1$, so $\|J_\varepsilon(\omega)\|_{\mathrm{op}} \leq C_X^2$ by (a). Hence $u_\varepsilon := J_\varepsilon(\omega)^{-1}v$ satisfies the lower bound $\|u_\varepsilon\| \geq \|v\|/C_X^2$ (used in Step 2). Conversely, (a)–(b) give

$$J_\varepsilon(\mathbf{1}) \succeq c_3\,\mathbb{E}_\varepsilon[XX^\top \mathbf{1}\{\mathcal{B}\}] \succeq c_3\lambda_B I, \qquad J_\varepsilon(\omega_\varepsilon) \succeq c_0 c_3\,\mathbb{E}_\varepsilon[XX^\top \mathbf{1}\{\mathcal{B}\}] \succeq c_0 c_3 \lambda_B I,$$

so $\|J_\varepsilon(\omega)^{-1}\|_{\mathrm{op}} \le 1/(c_0c_3\lambda_B)$ and $\|u_\varepsilon\| \le \|v\|/(c_0c_3\lambda_B)$ uniformly in ε (used in Step 3).

Step 2 (DR diverges). Restrict the integral in Eq. (194) to $\mathcal{R}_\varepsilon$. By Lemma K.6 and (a),(c)–(e): in the low-TO scenario, using $\sigma^2_{h,\varepsilon}(a, X) \ge c_1$ from (d) and the algebraic identity $1/\pi + 1/(1-\pi) = 1/[\pi(1-\pi)]$, the first term of Eq. (188) satisfies

$$(1-\rho_\varepsilon)\Big[\tfrac{\sigma^2_{h,\varepsilon}(1,X)}{\pi_\varepsilon} + \tfrac{\sigma^2_{h,\varepsilon}(0,X)}{1-\pi_\varepsilon}\Big] \;\ge\; \tfrac{c_3c_1}{\pi_\varepsilon(1-\pi_\varepsilon)} \;\asymp\; \tfrac{1}{\varepsilon}, \tag{195}$$

giving $V^\star_{\mathbf{1},\varepsilon} \gtrsim 1/\varepsilon$ on $\mathcal{R}_\varepsilon$; in the low-LO scenario, since $\pi_\varepsilon(1-\pi_\varepsilon) \ge c_4$ on $\mathcal{R}_\varepsilon$, the second term of Eq. (188) satisfies $\rho_\varepsilon\Psi_\varepsilon/[\pi_\varepsilon(1-\pi_\varepsilon)]^2 \asymp \rho_\varepsilon^{-1} \asymp 1/\varepsilon$. Either way,

$$v^\top \mathcal{V}_\varepsilon(\mathbf{1})\, v \ge \frac{c}{\varepsilon}\,\mathbb{E}_\varepsilon\big[(X^\top u_\varepsilon)^2 \mathbf{1}\{\mathcal{R}_\varepsilon\}\big] \ge \frac{c}{\varepsilon}\,\lambda_R\|u_\varepsilon\|^2 \ge \frac{c'}{\varepsilon}\|v\|^2, \tag{196}$$

where the second inequality uses (c) and the third uses Step 1. Taking $\varepsilon \downarrow 0$ gives divergence.

Step 3 (DO bounded). By Lemma K.6, (a) and (d)–(e), $V^\star_{\omega_\varepsilon,\varepsilon}(X)$ is uniformly bounded on $\mathcal{X}$: the first term of Eq. (189) is at most $\bar\sigma^2/4$ via $(1-\rho_\varepsilon)\rho_\varepsilon^2\pi_\varepsilon(1-\pi_\varepsilon) \le 1/4$ and (d), while the second term $\rho_\varepsilon^3\Psi_\varepsilon \le \bar\Psi$ by (e). Setting $\bar V := \bar\sigma^2/4 + \bar\Psi$,

$$v^\top \mathcal{V}_\varepsilon(\omega_\varepsilon)\, v \le \bar V \cdot \mathbb{E}_\varepsilon\big[(X^\top u_\varepsilon)^2\big] = \bar V \cdot u_\varepsilon^\top \mathbb{E}_\varepsilon[XX^\top] u_\varepsilon \le \bar V C_X^2 \|u_\varepsilon\|^2, \tag{197}$$

which is uniformly bounded by Step 1.

Step 4 (Ratio well-defined). On $\mathcal{B}$ we have $(1-\rho_\varepsilon)\omega_\varepsilon = (1-\rho_\varepsilon)\pi_\varepsilon(1-\pi_\varepsilon)\rho_\varepsilon \ge c_3c_0$, and from Eq. (189) together with (d), $V^\star_{\omega_\varepsilon,\varepsilon} \ge c_3c_0^2c_1$ on $\mathcal{B}$. Combined with $\mathbb{E}_\varepsilon[XX^\top\mathbf{1}\{\mathcal{B}\}] \succeq \lambda_B I$ from (b), $v^\top\mathcal{V}_\varepsilon(\omega_\varepsilon)v \ge c_3c_0^2c_1\,\lambda_B\|u_\varepsilon\|^2 > 0$ for $v \ne 0$. Combining Steps 2–4 gives the claim. □

Remark K.10 (Squared overlap weight). For the squared LT-O-DO-learner weight $\widetilde\omega_\varepsilon = \pi_\varepsilon^2(1-\pi_\varepsilon)^2\rho_\varepsilon$, Eq. (187) and Assumption K.8(a),(b) give

$$V^\star_{\widetilde\omega_\varepsilon} \le V^\star_{\omega_\varepsilon}/16, \qquad J_\varepsilon(\widetilde\omega_\varepsilon) \succeq c_3c_0^2\lambda_B I.$$

The same variance-stabilization argument therefore applies to $\widetilde\omega_\varepsilon$.

Remark K.11 (Scope of the comparison). Theorem 5.3 (detailed in Theorem K.9) establishes *directional robustness in the low-overlap regime*, not uniform PSD (positive semi-definite) dominance: in the high-overlap bulk, the smaller $J(\omega)$ relative to $J(\mathbf{1})$ can tilt the comparison in the opposite direction. The non-degeneracy condition $\mathbb{E}[XX^\top\mathbf{1}\{\mathcal{R}_\varepsilon\}] \succeq \lambda_R I$ is what makes the divergence hold for every direction v; without it, one obtains divergence only along v for which $X^\top J(\mathbf{1})^{-1}v$ has non-vanishing L^2 mass on the low-overlap region.