Debiased Machine Learning for Contamination-Free Causal Estimation with Discrete and Continuous Treatments

Tamer Çetin*

December 5, 2025

Abstract

Linear regression estimates of causal effects with multi-arm or continuous treatments are often contaminated by weighting spillovers. We develop a unified, efficient influence function (EIF) that eliminates this bias, synthesizing discrete and continuous regimes. Our estimator, implemented via cross-fitted double machine learning (DML), achieves $\sqrt{n}$-consistency and semiparametric efficiency. To handle weak overlap, we introduce the Generalized Overlap Average Treatment Effect (GOATE), a trimming-based estimand chosen via data-driven variance minimization. We prove valid inference even when the number of arms grows ($K_n = o(n^{1/4})$). Simulations and applications to class-size and policing data demonstrate the robustness of GOATE-DML against non-linear confounding.

JEL Codes: C14, C21, C55

Keywords: contamination bias, multi-arm causal inference, continuous treatment, double machine learning, weak overlap, cross-fitting

*University of California, Berkeley. E-mail: tamercetin@berkeley.edu. Comments welcome.

1 Introduction

Causal studies increasingly involve either several mutually exclusive policy arms—for instance, school-reopening tiers or tax brackets—or a continuous score that functions as a dosage. In these designs, the naïve ordinary-least-squares (OLS) coefficient on a treatment indicator is *contaminated*; it conflates the arm's own causal effect with spill-overs that arise mechanically from correlations among the remaining arms. The point was foreshadowed in multi-valued treatment work by Robins, Rotnitzky, and Zhao (1994) (reviewed in Lopez and Gutman, 2017), formalized for linear regressions by Goldsmith-Pinkham *et al.* (2024) (GPHK) and Słoczyński (2022), and can bias even randomized experiments whenever compliance is imperfect. GPHK offers three coefficient-level "plug-in" corrections—*interacted ATE* (IA), *common-weight* (CW), and *easiest-to-estimate* (EW)—but their validity rests on linear conditional means and near-perfect overlap.

A broader literature tackles related but distinct challenges. For discrete arms, semiparametric influence-function work derives efficiency bounds and double-robust scores (Newey, 1990; Graham, 2011; Kennedy, 2016); for *continuous* doses, the generalized propensity-score framework of Imbens (2000), Hirano and Imbens (2004), Imai and van Dyk (2004), and Imai and Ratkovic (2014) exposes analogous support problems, while Kennedy (2024), Colangelo and Lee (2020), and Chernozhukov *et al.* (2022) provide orthogonal scores that unlock machine-learning first stages. Weak overlap has spurred trimming and overlap-weighting rules (Crump *et al.*, 2009; Li *et al.*, 2018; Li and Li, 2019; Kallus and Oprescu, 2023), but none of these strands addresses contamination across *multiple* arms or unifies discrete and continuous regimes within one estimator.

This paper provides such a unification. We derive a single, Neyman-orthogonal influence function that delivers contamination-free estimation for *both* multi-arm ($D \in \{0, \ldots, K\}$) and continuous-dose ($D \in \mathbb{R}$) treatments. The discrete special cases reproduce IA, EW, and CW—thereby strictly generalizing GPHK—while retaining efficient influence-function form. Embedding this score in a cross-fitted double-machine-learning (DML) routine (Chernozhukov *et al.*, 2018) yields root-n inference in the presence of high-dimensional, non-linear nuisances.

Limited support remains a first-order concern. We therefore focus on the GOATE: the average effect for the sub-population whose minimum propensity (or conditional density) lies above a threshold τ. A data-driven mean-squared-error (MSE) rule (Bennett et al., 2023) chooses τ adaptively, and we prove valid cluster-robust inference even when the number of arms grows with the sample ($K_n = o(n^{1/4})$). GOATE is policy-relevant because it describes individuals who could *plausibly* receive any treatment arm, respecting external validity in settings with limited overlap.

The paper makes four contributions. First, we formally map the contamination bias decomposition of Goldsmith-Pinkham *et al.* (2024) to the language of semiparametric efficiency. We demonstrate that the standard EIF is the non-parametric generalization of their linear corrections. By characterizing the EIF in this specific multi-arm context, we show that "de-contamination" is analytically equivalent to satisfying the Neyman orthogonality condition with respect to the propensity score. In doing so, we synthesize the discrete and continuous regimes into a single Riesz representer framework, differing only in the reference measure. Second, it supplies the first contamination-free framework for continuous treatments and proves valid joint inference when the number of arms grows. We establish that valid inference requires the stricter growth condition $K_n = o(n^{1/4})$ to control the explosion of inverse-propensity moments, revising standard high-dimensional CLT results that assume bounded moments. Third, it formalizes GOATE together with a data-driven trimming rule that balances bias and variance, and it establishes cluster-robust inference via a cross-fit-by-cluster scheme. Fourth, the empirical illustrations show that the estimator can alter economic conclusions: in Project STAR the estimated small-class effect falls by 0.73 test-score points relative to OLS—three times the adjustment implied by GPHK—and in two further micro-credit experiments the estimator uncovers heterogeneous effects that linear methods obscure. These results connect to semiparametric efficiency bounds for multi-valued treatments (Newey, 1990; Graham, 2011), to recent orthogonalization methods for continuous doses (Kennedy, 2024), and to high-dimensional central-limit theory (Chernozhukov *et al.*, 2017). To the best of my knowledge, this is the first contamination-robust, root-n-consistent estimator that spans both discrete and continuous treatment designs in high-dimensional, non-linear settings.

Section 2 introduces notation and high-level assumptions. Section 3 derives the orthogonal influence functions. Section 4 maps our score into the linear corrections of Goldsmith-Pinkham *et al.* (2024), whereas Section 5 details the cross-fitted estimator and adaptive trimming rule. Asymptotic theory is given in Section 6, followed by simulations (Section 7) and empirical applications (Section 8). Section 9 concludes.

2 Setup and Assumptions

The goal of this section is two-fold. First, we spell out the stochastic environment—what variables are observed and what causal quantities are of interest. Second, we collect the high-level conditions under which the subsequent orthogonal-score construction and asymptotic theory will operate. Throughout we let $Z_i = (Y_i, D_i, W_i)$ denote the observable outcome (Y), treatment (D), and covariates (W) for unit $i = 1, \ldots, n$, and assume the draws are i.i.d. for notational simplicity. Two treatment regimes are covered.

We begin with the canonical setting of $K + 1$ mutually exclusive arms. Formally, $D_i \in \{0, 1, \ldots, K\}$, where $k = 0$ is the control arm, and we write $D_{ik} = 1\{D_i = k\}$ for the associated dummies. The *propensity score* for arm k is $p_k(W) = P(D_i = k \mid W_i), \qquad k = 0, \ldots, K$, so $p_k(W) \in (0, 1)$ for each k under our overlap conditions below. Throughout the paper we rely on a standard selection on observables condition.

Assumption 2.1 (Unconfoundedness—Discrete)**.** *The vector of potential outcomes* $\{Y_i(0), Y_i(1), \ldots, Y_i(K)\}$ *is independent of the realized treatment* D_i *conditional on covariates* W_i*:* $\{Y_i(0), Y_i(1), \ldots, Y_i(K)\} \perp\!\!\!\perp D_i \mid W_i$.

After conditioning on W_i—which may be high-dimensional—assignment to each arm is "as good as random," ruling out unobserved confounders. This is the discrete analogue of the continuous unconfoundedness assumption used below. Our primary target is the arm-specific ATE relative to control,

$$\theta_k = E\big[\, Y_i(k) - Y_i(0) \big], \qquad k = 1, \ldots, K. \tag{1}$$

Point identification of θ_k follows directly from Assumption 2.1 and the law of iterated ex-

pectations. Many empirical studies record a dosage or index rather than a finite set of arms—think of tax rates, pollution levels, or test scores. We therefore let D_i take values in a compact interval $\mathcal{D} \subset \mathbb{R}$ and denote by $f(d \mid W)$ the conditional density of the dose. The conditional mean function is $m(d, W) = E[Y \mid D = d, W]$.

Assumption 2.2 (Unconfoundedness—Continuous)**.** *For every $d \in \mathcal{D}$ the potential outcome $Y_i(d)$ is independent of the realized dose D_i conditional on covariates W_i: $Y_i(d) \perp\!\!\!\perp D_i \mid W_i$.*

In the continuous setting we focus on the average *partial* effect (APE), $\theta = E\big[\partial_d m(D_i, W_i)\big]$, i.e. the expected marginal effect of a small dose change evaluated at the observed D_i.

Assumption 2.3 (Weak overlap (both designs))**.** *Let $\tau_n \downarrow 0$ be a deterministic trimming threshold. There exists $\delta > 0$ such that $P\big(\min_k p_k(W) < \tau_n\big) = O(\tau_n^\delta)$ (discrete) or $P\big(f(D \mid W) < \tau_n\big) = O(\tau_n^\delta)$ (continuous).*

Assumption 2.4 (Smoothness)**.** *The regression function $m(d, W)$ and the log-density $\log f(d \mid W)$ are twice continuously differentiable in d and lie in a Hölder class of order $\alpha > 2$ with respect to W.*

These regularity conditions guarantee that the EIF we derive in section 3.1 exists and, importantly, that local-polynomial estimators can attain the $n^{-1/4}$ convergence rate that orthogonal scores require. Estimation becomes unstable when $p_k(W)$ or $f(D \mid W)$ approach zero. Rather than assume a uniform lower bound, we allow the support violations to vanish slowly with the sample size and control them via trimming.

Remark 2.1 (Trimming indicators and shorthand)**.** *We use the following trimming indicators throughout:*

$$T_i^{disc}(\tau) = \mathbf{1}\left\{\min_{\ell \le K} p_\ell(W_i) \ge \tau\right\}, \qquad T_i^{cont}(\tau) = \mathbf{1}\big\{f(D_i \mid W_i) \ge \tau\big\}. \tag{2}$$

When the design is clear from context we write $T_i(\tau)$ for either (2)*.*

The tails of the propensity or density are allowed to get thinner as the sample grows, but not so fast that we lose $\sqrt{n}$ information after trimming the worst-overlap observations. The

adaptive rule of section 6.3 chooses τ_n on a data-dependent grid. For the accuracy of machine-learning first stages, all subsequent estimators rely on non-parametric fits $\hat{m}, \hat{p}_k, \hat{f}$. The next assumption imposes the minimal rate conditions under which Neyman orthogonality can partial out first-stage error.

Assumption 2.5 (Anti-concentration near the trimming threshold)**.** *There exist $c > 0$ and $\kappa \in (0, 2]$ such that, uniformly over $\tau \in \mathcal{G}_n$ and all $h > 0$ small enough,* $\Pr\big(|\min_{\ell \le K} p_\ell(W) - \tau| \le h\big) \le c\,h^\kappa$ *(discrete),* $\Pr\big(|f(D \mid W) - \tau| \le h\big) \le c\,h^\kappa$ *(continuous).*

That is, the distribution of the trimming margin has a bounded local density uniformly over the grid.

Remark 2.2. *Assumption 2.5 is implied, for example, if the distribution of* $\min_{\ell \le K} p_\ell(W)$ *(respectively $f(D \mid W)$) admits a bounded Hölder-continuous density in a neighborhood of each $\tau \in \mathcal{G}_n$. It is used only to control indicator disagreements* $\Pr(\hat{T}_i(\tau) \neq T_i(\tau))$ *when we combine hard trimming with estimated propensities/densities.*

Assumption 2.6 (First-stage ML rates)**.** *On the* trimmed *support,* $\|\hat{m} - m\|_{2,T} = o_p(n^{-1/4})$, $\|\hat{p}_k - p_k\|_{2,T} = o_p(n^{-1/4})$, *and, for continuous doses,* $\|\partial_d \hat{m} - \partial_d m\|_{2,T}$, $\|\hat{f} - f\|_{2,T} = o_p(n^{-1/4})$.

Honest causal forests, boosted trees, deep nets (Collier *et al.*, 2021), and local-polynomial estimators all satisfy these rates under either sparsity or smoothness—see Appendix A.12 for citations. We finally record two technical conditions that are invisible in classical low-dimensional proofs but become essential once we allow (i) clusters of heterogeneous size, and (ii) a growing number of arms K_n.

Remark 2.3 (Trimmed L_2 norm)**.** *For any scalar function $g(Z)$ and threshold τ, we define* $\|g\|^2_{2,T(\tau)} := \frac{E\big[g(Z)^2\,\mathbf{1}\{T(\tau)=1\}\big]}{P\big(T(\tau)=1\big)}$. *All rate statements in Assumption 2.6 are with respect to* $\|\cdot\|_{2,T(\tau)}$.

Assumption 2.7 (Uniform moments and overlap margin as K grows)**.** *Let $\varphi_k^\star$ denote the efficient score in* (6)*. There exists a trimming lower bound $\underline{\tau}_n \in (0, 1)$ (the minimum of the grid in section 6.3) such that for all large n:*

(i) Uniform overlap tail: $P(\min_\ell p_\ell(W) < t) \le C\,t^\delta$ *for all $t \in (0, 1)$ and some $\delta > 0$;*

(ii) Uniform outcome fourth moments: $\sup_{\ell \le K_n} E\big[\,|Y - m_\ell(W)|^4\,\big] < \infty$;

(iii) Joint control of fourth moments and K_n: $\sup_{k\le K_n} E\Big[\varphi_k^\star(Z)^4\,\mathbf{1}\{\min_\ell p_\ell(W)\ge \underline{\tau}_n\}\Big]\le C_4\,\underline{\tau}_n^{-q}$ *for some finite* C_4 *and* $q\in[2,4]$*; and the growth/trimming coupling* $\frac{K_n^4\,\underline{\tau}_n^{-q}}{n}\longrightarrow 0$.

This stricter condition ($K_n = o(n^{1/4})$ when τ_n is fixed) is necessary because the fourth moments of the influence function scale with $p_k^{-3} \asymp K_n^3$ as the number of arms increases. The last display ensures the Lyapunov ratio in the many-arm CLT vanishes even if the fourth moments inflate as $\underline{\tau}_n \downarrow 0$. When a non-vanishing grid is used (fixed $\underline{\tau} > 0$), it holds automatically.

Remark 2.4 (How to pick the grid in practice)**.** *If you wish to let the trimming grid vanish with n while K_n grows, set its minimum according to $\underline{\tau}_n \asymp (K_n^3/n)^{1/q}$ for the q implied by your learners (e.g. $q = 2$ under standard tail behavior). This meets Assumption 2.7 and keeps the many-arm CLT valid.*

Remark 2.5 (Fixed-arm vs. growing-K_n scope)**.** *Unless explicitly stated otherwise (see Corollary 6.3), all asymptotic statements in the paper are* fixed-arm*: the index k is held fixed as $n \to \infty$. No restriction on how the total number of arms $K = K_n$ may grow is needed for these fixed-arm results beyond the uniform moment condition $\sup_{k\le K_n} E[\varphi_k^\star(Z)^4] < \infty$. The growth condition $K_n = o(n^{1/4})$ is used only for the* joint *many-arm CLT.*

Under Assumption 2.4 a trimmed L_2 rate of $o_p(n^{-1/4})$ is attainable for local-polynomial density and regression estimators whenever the treatment dimension $d \le 4$ Cattaneo et al., 2024.

Remark 2.6. *Detailed moment and cross-fit regularity proofs are collected in Appendix A.6.*

While the global ATE, $\tau(d) = E[Y(d) - Y(0)]$, is the standard target in causal inference, estimating it requires *strong overlap*—that is, the propensity score must be bounded away from zero and one uniformly. Under the weak overlap sequence defined in Assumption 2.2, the standard ATE is not $\sqrt{n}$-estimable because the propensity scores $p_k(W)$ may drift toward zero. In complex multi-arm experiments, strong overlap assumption often fails (see, e.g., the analysis of Drexler et al. (2014) in section 8). Therefore, we do not target the ATE directly. Instead, we define the GOATE as the target parameter, explicitly conditioning on the trimming set defined by the overlap weights.

Let $\mathcal{V} \subseteq \text{supp}(W)$ denote the region of valid overlap, determined by a trimming rule based on the propensity score (for discrete D) or the conditional density (for continuous D).

Proposition 2.1. *[Identification of GOATE] Under the relevant unconfoundedness assumption (2.1 or 2.2), the trimmed parameter θ_{GOATE} is identified as the conditional expectation:*

$$\theta_{GOATE}(d) = E\left[Y(d) - Y(0) \mid W \in \mathcal{V}\right]. \tag{3}$$

Equivalently, using the Riesz representation, this is the weighted average effect:

$$\theta_{GOATE}(d) = \frac{E[\mathbf{1}_{\mathcal{V}}(W) \cdot (Y(d) - Y(0))]}{E[\mathbf{1}_{\mathcal{V}}(W)]}. \tag{4}$$

By targeting (3) *rather than the global ATE, we avoid the bias and variance explosion associated with extrapolating linear models into regions of poor support.*

Proof. See Appendix A.1. □

Remark 2.7 (Policy relevance of GOATE)**.** *When propensity support is limited, trimming observations with* $\min_k p_k(W) < \tau$ *(or* $f(D|W) < \tau$*) changes the target parameter from the full-sample ATE to the* Generalized Overlap ATE*. GOATE is not a drawback but a feature: it reports the causal effect on the sub-population that could plausibly have been assigned* any *treatment arm, thus respecting external validity for feasible policy counterfactuals.*

3 Orthogonal Influence Functions

Contamination arises because a saturated regression forces the coefficients on different treatment arms (or bins) to *compete for the same residual variation.* The cure is an *orthogonal* influence function that: (i) identifies the causal parameter in the usual semiparametric sense, (ii) removes every linear dependence on the nuisance functions $\eta(\cdot)$, and therefore (iii) remains first-order valid even when η is estimated by a flexible learner. The two influence functions below meet all three goals and reduce to classical AIPW scores when $K = 1$.

The following influence functions both identify the parameters and algebraically remove the contamination bias of Goldsmith-Pinkham *et al.* (2024).

3.1 Continuous Treatment

Let $f(d \mid W)$ be the conditional density of the dose and write $s(d, W) = \partial_d \log f(d \mid W)$ for its score. Conditional outcome regressions are denoted $m(d, W) = E[Y \mid D = d, W]$.

The EIF is

$$\varphi^{(c)}(Z; \theta, \eta) \;=\; s(D, W)\{Y - m(D, W)\} + \partial_d m(D, W) - \theta, \qquad \eta = (m, f). \tag{5}$$

The first term multiplies the outcome residual by the density score— the analogue of a "clever covariate" familiar from binary AIPW estimators—while the $\partial_d m$ term shifts the moment so that its expectation is exactly the average partial effect θ. Any first-order error in $\hat{m}$ or $\hat{f}$ cancels by construction, which is the essence of Neyman orthogonality proven below.

Remark 3.1 (Relation to Kennedy (2023))**.** *Our continuous score generalizes the "automatic orthogonalization" framework of Kennedy (2023). While Kennedy (2023) derives the EIF for a scalar continuous treatment, our formulation extends this to the multi-arm setting by explicitly incorporating the contamination terms arising from other treatment levels. Specifically, the cross-arm orthogonality correction in Equation* (5) *is absent in the scalar framework but is necessary here to isolate the marginal effect of dose d from the spill-overs of doses $d' \neq d$.*

Lemma 3.1 (Orthogonality—Continuous)**.** *Under Assumptions 2.2 and 2.4, and the boundary condition in Assumption 3.1 below, the influence function $\varphi^{(c)}$* (i) *uniquely identifies θ and* (ii) *is Neyman-orthogonal with respect to the nuisance bundle $\eta = (m, f)$.*

Proof Sketch. The proof proceeds in two parts. First, we show that the expectation of the score is zero at the true parameter values, which establishes identification of $\theta = E[\partial_d m(D, W)]$. Second, we compute the pathwise derivative of the moment condition with respect to perturbations in the nuisance functions m and f and show it is zero. This Neyman orthogonality property is key to the $\sqrt{n}$-consistency of the DML estimator, as it renders the estimator insensitive to first-order errors in the estimation of the nuisance functions. The full, detailed proof is provided in Appendix A.2. □

Assumption 3.1 (Strict Boundary Control)**.** *To ensure the validity of the integration by parts in Lemma 3.1 when the density $f(d|W)$ does not vanish at the boundaries of $\mathcal{D}$ (e.g., $f(0|W) > 0$), we require strictly unbiased estimation at the boundary. Standard Gaussian kernel density estimators suffer from boundary bias $O(h)$ rather than $O(h^2)$. Therefore, estimation of the nuisance functions $f(d|W)$ and $m(d, W)$ must utilize boundary-corrected kernels (e.g., local linear regression) as detailed in Cattaneo et al. (2024). This ensures the product of the nuisance error and the density vanishes at $\partial\mathcal{D}$ at the required $n^{-1/4}$ rate.*

Remark 3.2 (Learners and rates)**.** *Under Assumption 3.1 and the smoothness conditions used for density/local polynomial estimators, the trimmed L_2 rates required for our orthogonal score hold in low dimensions (e.g., $d \leq 4$), and the integration-by-parts step is valid. State explicitly whether $\partial_d m$ is produced by the same local polynomial fit as m (preferred) or by a separate fit.*

Remark 3.3 (Vector-valued doses)**.** *Assumption 3.1 is imposed* coordinate-wise *when $D \in \mathbb{R}^d$, i.e. the boundary product $h_m(\boldsymbol{d}, W) f(\boldsymbol{d} \mid W)$ must vanish at every face of the rectangular support $\mathcal{D} \subset \mathbb{R}^d$. All proofs that rely on integration by parts therefore extend verbatim to $d > 1$.*

Remark 3.4 (Rates under boundary-robust estimation)**.** *Under Assumption 2.4 with $\alpha > 2$ and treatment dimension $d \leq 4$, boundary-corrected local polynomials deliver trimmed L_2 rates $o_p(n^{-1/4})$ for $(m, \partial_d m)$ and $(f, \partial_d \log f)$ on any grid value $\tau > 0$, which suffices for Neyman-orthogonal DML. See Cattaneo et al. (2024) and references therein.*

Remark 3.5 (Square-Integrability of the Score $s(D, W)$)**.** *The validity of the asymptotic theory requires the influence function to have a finite second moment, which in turn depends on the score function $s(D, W) = \partial_d \log f(D|W)$ being square-integrable. Assumption 2.4 states that $\log f(d|W)$ is twice continuously differentiable, which implies that both $s(d, W)$ and its derivative $\partial_d s(d, W)$ are continuous functions. On the trimmed support, where $f(D|W) \geq \tau_n > 0$, the denominator of $s(D, W) = (\partial_d f)/f$ is bounded away from zero. As continuous functions on a compact domain are bounded, both $\partial_d f(d, W)$ and $f(d|W)$ (on the trimmed set) are bounded. Therefore, $s(D, W)$ is bounded on the trimmed support. A*

bounded random variable on a probability space necessarily has finite moments of all orders, ensuring that $E[s(D,W)^2 \cdot \mathbf{1}\{f(D|W) \geq \tau_n\}] < \infty$.

Plugging $\varphi^{(c)}(Z_i; \hat{\theta}, \hat{\eta})$ into a cross-fitted DML estimator therefore yields $\sqrt{n}$-consistent and semiparametrically efficient inference for continuous treatments—even when m and f are fitted by, say, boosted trees or deep nets.

3.2 Discrete Multi-Arm Treatment

For mutually exclusive arms $D \in \{0, 1, \ldots, K\}$ let $m_\ell(W) = E[Y \mid D = \ell, W]$ and $p_\ell(W) = P(D = \ell \mid W)$. Contamination in a plain dummy regression stems from the fact that the dummies sum to one; the EIF below corrects this algebraically by subtracting the control term arm-by-arm.

The efficient, contamination-free influence function for arm k is

$$\begin{aligned}\varphi_k^\star(Z; \theta_k, \eta) = \big[m_k(W) - m_0(W)\big] \\ + \frac{D_{ik}}{p_k(W)}\big\{Y - m_k(W)\big\} - \frac{D_{i0}}{p_0(W)}\big\{Y - m_0(W)\big\} - \theta_k,\end{aligned} \tag{6}$$

with nuisance bundle $\eta = (m_0, \ldots, m_K, p_0, \ldots, p_K)$.
The structure mirrors an AIPW score for a *binary* treatment, except that the outcome-regression and IPW corrections are stacked in a "treatment-minus-control" fashion.

Lemma 3.2 (Identification & Neyman orthogonality)**.** *Let* $\tau_k(W) = E[Y_i(k) - Y_i(0) \mid W]$. *Then* $E[\varphi_k^\star] = E[\tau_k(W)] - \theta_k$*; hence* $E[\varphi_k^\star] = 0$ *iff* $\theta_k = E[Y_i(k) - Y_i(0)]$. *Moreover,* $\partial_r E[\varphi_k^\star(Z; \theta_k, \eta + rh)]_{r=0} = 0$ *for every mean-zero perturbation* $h = (h_m, h_p)$*, and* $\varphi_k^\star$ *attains the semiparametric efficiency bound.*

Because $\varphi_k^\star$ is also doubly robust, an empiricist only needs one of the two first-stage models—propensities or outcome regressions—to be well estimated. This is crucial when flexible learners are applied in high-dimensional W. This structure generalizes the binary-treatment efficiency results of Hahn (1998) to the multi-arm setting.

Proof Sketch. The proof establishes two key properties. First, identification, which shows that the score has a unique root at the true ATE θ_k. Second, Neyman orthogonality, which

demonstrates that the moment condition is locally insensitive to errors in the nuisance functions η. This latter property is essential for the validity of the DML estimator. The logic relies on the law of iterated expectations and the specific structure of the score, which balances regression-adjustment and inverse-propensity weighting terms. A full, step-by-step derivation is provided in Appendix A.4. □

Remark 3.6 (Double robustness)**.** *The moment condition in Lemma 3.2 continues to hold if either all outcome regressions (m_ℓ) or all propensity scores (p_ℓ) are consistently estimated, as shown in Appendix A.4.*

4 Mappings to ATE(IA), CW, and EW

This section makes the algebra explicit. Interacted–ATE (IA) and Common–Weight (CW) are plug-in instances of our unified orthogonal score, while EW coincides with the pairwise AIPW score on the binary subsample $\{D \in \{0, k\}\}$.

4.1 Interacted–ATE (IA): explicit mapping in sample

Let the saturated linear model be $Y_i = \alpha_0 + \beta_0^\top W_i + \sum_{\ell=1}^K D_{i\ell}\,(\alpha_\ell + \beta_\ell^\top W_i) + \varepsilon_i, \qquad E[\varepsilon_i \mid D_i, W_i] = 0$, and let $\hat{m}_\ell(W) = \hat{\alpha}_\ell + \hat{\beta}_\ell^\top W$ be the OLS fits on the training folds (cross-fitted). Define the arm-specific OLS residuals $r_{i\ell} = Y_i - \hat{m}_\ell(W_i)$ and note the normal equations

$$\sum_{i\in\mathcal{I}_j} \mathbf{1}\{D_i = \ell\}\, r_{i\ell} = 0, \qquad \sum_{i\in\mathcal{I}_j} \mathbf{1}\{D_i = \ell\}\, r_{i\ell} W_i = 0, \quad (\ell = 0, \dots, K). \tag{7}$$

Our efficient score for arm k (cf. (6)) is $\varphi_k^\star(Z_i; \theta_k, \eta) = \big[m_k(W_i) - m_0(W_i)\big] + \frac{D_{ik}}{p_k(W_i)}\{Y_i - m_k(W_i)\} - \frac{D_{i0}}{p_0(W_i)}\{Y_i - m_0(W_i)\} - \theta_k$. Evaluate it with nuisances $(\hat{m}, \hat{p})$ and drop the $-\theta_k$ term to form the raw moment $\hat{\psi}_{ik} = \big[\hat{m}_k(W_i) - \hat{m}_0(W_i)\big] + \frac{D_{ik}}{\hat{p}_k(W_i)}\, r_{ik} - \frac{D_{i0}}{\hat{p}_0(W_i)}\, r_{i0}$.

Add and subtract $D_{ik}\, r_{ik}$ and $D_{i0}\, r_{i0}$ to re-express the IPW residual terms: $\frac{D_{ik}}{\hat{p}_k} r_{ik} = D_{ik} r_{ik} + \Big(\frac{D_{ik}}{\hat{p}_k} - D_{ik}\Big) r_{ik}$, $\quad \frac{D_{i0}}{\hat{p}_0} r_{i0} = D_{i0} r_{i0} + \Big(\frac{D_{i0}}{\hat{p}_0} - D_{i0}\Big) r_{i0}$. Summing over $i \in \mathcal{I}_j$ and using (7), the *unweighted* residual sums $\sum D_{ik} r_{ik}$ and $\sum D_{i0} r_{i0}$ vanish *exactly* on each estimation fold, leaving $\sum_{i\in\mathcal{I}_j} \hat{\psi}_{ik} = \sum_{i\in\mathcal{I}_j} \big[\hat{m}_k(W_i) - \hat{m}_0(W_i)\big] + \sum_{i\in\mathcal{I}_j} \Big(\frac{D_{ik}}{\hat{p}_k} - D_{ik}\Big) r_{ik} - \sum_{i\in\mathcal{I}_j} \Big(\frac{D_{i0}}{\hat{p}_0} - D_{i0}\Big) r_{i0}$.

If (as is standard) $\hat{p}_\ell$ are estimated consistently (e.g., multinomial logit, RF, GBM) on the complementary folds, then the last two sums are $o_p(n)$ by cross-fit orthogonality. Hence the *trimmed* sample moment equation $\sum_i T_i(\hat{\tau})\,\hat{\psi}_{ik} = 0$ is equivalent, up to $o_p(n)$, to the interacted-ATE estimating equation $\sum_i T_i(\hat{\tau})\left[\hat{m}_k(W_i) - \hat{m}_0(W_i)\right] = \left(\sum_i T_i(\hat{\tau})\right)\hat{\theta}_k^{\mathrm{IA}}$. Thus the IA estimator is the plug-in DML solution for our score, with $o_p(n^{-1/2})$ differences in $\hat{\theta}_k$ under the ML rates. IA solves the same orthogonal moment as our estimator, modulo cross-fit $o_p(n^{-1/2})$ terms coming from weighting the residuals by $\hat{p}_\ell^{-1}$.

Proposition 4.1 (IA as plug-in DML)**.** *Under Assumption 2.6 (first-stage L_2 rates $o_p(n^{-1/4})$ on the trimmed support) and J-fold cross-fitting, the plug-in DML* sample moment *based on* (6) *equals the IA estimating equation up to $o_p(n^{-1/2})$. Consequently, the IA point estimate and the plug-in DML estimate are asymptotically equivalent: $\hat{\theta}_k^{\mathrm{IA}} - \hat{\theta}_k^{\mathrm{DML}} = o_p(n^{-1/2})$.*

4.2 Common–Weight (CW): explicit mapping in sample

Let $g(W)$ be the CW weight $g(W) = \left[\sum_{\ell=0}^{K} \hat{\pi}_\ell(1-\hat{\pi}_\ell)/\hat{p}_\ell(W)\right]^{-1}$, $\hat{\pi}_\ell = n^{-1}\sum_i D_{i\ell}$. Define the CW "stabilized" indicators $\tilde{D}_{i\ell} = \frac{g(W_i)\,D_{i\ell}}{\hat{p}_\ell(W_i)}$. The CW normal equation for θ_k is the weighted moment

$$\sum_i T_i(\hat{\tau})\,g(W_i)\,(D_{ik} - \hat{\pi}_k)\left(Y_i - \hat{\theta}_k^{\mathrm{CW}}\right) = 0. \tag{8}$$

Start from the raw DML moment $\hat{\psi}_{ik}$ above and multiply the two IPW residual terms by $g(W_i)$ (which does not change the root of the moment condition because the $m_k - m_0$ piece is unweighted). Then $\sum_i T_i(\hat{\tau})\,\hat{\psi}_{ik} = \sum_i T_i(\hat{\tau})\big[\hat{m}_k(W_i) - \hat{m}_0(W_i)\big] + \sum_i T_i(\hat{\tau})\tilde{D}_{ik}\,r_{ik} - \sum_i T_i(\hat{\tau})\tilde{D}_{i0}\,r_{i0}$. Use $r_{i\ell} = Y_i - \hat{m}_\ell(W_i)$, expand, and regroup terms to isolate Y_i: $\sum_i T_i(\hat{\tau})\,\hat{\psi}_{ik} = \sum_i T_i(\hat{\tau})\big[\tilde{D}_{ik} - \tilde{D}_{i0}\big]Y_i \;-\; \sum_i T_i(\hat{\tau})\big[\tilde{D}_{ik} - \tilde{D}_{i0}\big]\hat{m}_\bullet(W_i)$, where $\hat{m}_\bullet$ denotes the appropriate arm-specific regression in each term. Because the CW propensity fits satisfy the calibration identities $\sum_i T_i(\hat{\tau})\,\tilde{D}_{i\ell} = \sum_i T_i(\hat{\tau})\,g(W_i)\,\frac{D_{i\ell}}{\hat{p}_\ell(W_i)} \approx \sum_i T_i(\hat{\tau})\,g(W_i)$ $(\ell = 0, k)$, the terms involving $\hat{m}_\bullet$ cancel up to $o_p(n)$, and the remaining Y-part reduces to (8) after subtracting $\hat{\theta}_k \sum_i T_i g(W_i)(D_{ik} - \hat{\pi}_k) = 0$. Thus CW and the plug-in DML moment coincide up to $o_p(n)$ under the CW propensity fit. With CW propensities, the unified score's sample moment is the CW weighted normal equation (plus $o_p(n)$ cross-fit remainders), so the estimators are asymptotically equivalent under the stated rates.

Proposition 4.2 (CW as plug-in DML)**.** *Under Assumption 2.6 and J-fold cross-fitting, the plug-in DML* sample moment *based on* (6) *equals the CW estimating equation up to* $o_p(n^{-1/2})$ *when CW propensities are used in the score. Hence the CW and plug-in DML point estimates are asymptotically equivalent:* $\hat{\theta}_k^{\mathrm{CW}} - \hat{\theta}_k^{\mathrm{DML}} = o_p(n^{-1/2})$.

4.3 Easiest-to-Estimate (EW): pairwise AIPW and the cost of discarding data

Restrict to the binary subsample $\{D = 0\} \cup \{D = k\}$ and denote the *binary* propensities by $q_\ell(W) = P(D = \ell \mid W,\, D \in \{0, k\})$. Our score (6) collapses to the classical AIPW score: $\psi_{ik}^{\mathrm{AIPW}} = \big[m_k(W_i) - m_0(W_i)\big] + \frac{D_{ik}}{q_k(W_i)}\{Y_i - m_k(W_i)\} - \frac{D_{i0}}{q_0(W_i)}\{Y_i - m_0(W_i)\}$. Hence EW is numerically identical to pairwise AIPW when implemented with cross-fitted nuisances on the $\{0, k\}$ subsample.

Remark 4.1 (EW uses fewer augmentation samples)**.** *EW throws away all units with* $D \notin \{0, k\}$ *when forming the sample moment. Those units would still contribute through the augmentation term* $m_k(W) - m_0(W)$ *in our full-sample score. Heuristically, if* $\pi_{0k} = P(D \in \{0, k\})$*, then the variance contribution from the augmentation term scales like* π_{0k}^{-1} *times the conditional density of* $m_k(W) - m_0(W)$ *given* $D \in \{0, k\}$*, so when* π_{0k} *is small (many-arm designs), EW can be materially less precise. In contrast, the full-sample score averages the augmentation term over* all W *draws, improving precision while keeping the IPW part restricted to the two relevant arms.*

Aggregators and equivalences. Let $T_i(\tau)$ be the trimming indicator (discrete or continuous), and define $\hat{\theta}_k(\tau) = \frac{\sum_{i=1}^n T_i(\tau)\, \varphi_k^\star(Z_i;\hat{\eta})}{\sum_{i=1}^n T_i(\tau)}, \qquad k = 1, \ldots, K$, as our cross-fitted orthogonal estimator on the trimmed sample, where $\varphi_k^\star$ is the efficient score and $\hat{\eta}$ collects first-stage learners (as defined earlier in this section).

We consider three scalar aggregators formed as linear functionals of $(\hat{\theta}_1(\tau), \ldots, \hat{\theta}_K(\tau))$:
$\hat{\theta}^{\mathrm{ATE(IA)}}(\tau) := \frac{1}{K}\sum_{k=1}^K \hat{\theta}_k(\tau)$,
$\hat{\theta}^{\mathrm{CW}}(\tau) := \sum_{k=1}^K \hat{\omega}_k^{\mathrm{CW}}(\tau)\, \hat{\theta}_k(\tau)$, with $\hat{\omega}_k^{\mathrm{CW}}(\tau) \in \left\{ \frac{1}{K}, \frac{\sum_i T_i(\tau)\mathbf{1}\{D_i = k\}}{\sum_i T_i(\tau)} \right\}$,
$\hat{\theta}^{\mathrm{EW}}(\tau) := \sum_{k=1}^K \hat{\omega}_k^{\mathrm{EW}}(\tau)\, \hat{\theta}_k^{\mathrm{pair}}(\tau)$, $\hat{\theta}_k^{\mathrm{pair}}(\tau) := \frac{\sum_i T_i(\tau)\mathbf{1}\{D_i \in \{0,k\}\}\, \varphi_{k,0}^\star(Z_i;\hat{\eta})}{\sum_i T_i(\tau)\mathbf{1}\{D_i \in \{0,k\}\}}$, where $\varphi_{k,0}^\star$ is the

standard binary AIPW score for the pair "k vs. 0". The weights $\hat{\omega}_k^{\mathrm{EW}}(\tau)$ can be uniform $(1/K)$ or proportional to trimmed arm shares, mirroring common implementations.

Corollary 4.1 (Equivalence of Aggregated Estimators)**.** *It follows from the arm-specific equivalences in Propositions 4.1 and 4.2 that the linear aggregators inherit the same asymptotic properties. Since each aggregated estimator is a finite linear combination of the arm-specific $\hat{\theta}_k(\tau)$, the difference between the plug-in and aggregated versions is a linear combination of $o_p(n^{-1/2})$ terms. Specifically, $\hat{\theta}_{\mathrm{reg}}^{\mathrm{ATE(IA)}}(\tau) = \hat{\theta}^{\mathrm{ATE(IA)}}(\tau) + o_p(n^{-1/2})$ and $\hat{\theta}_{\mathrm{reg}}^{\mathrm{CW}}(\tau) = \hat{\theta}^{\mathrm{CW}}(\tau) + o_p(n^{-1/2})$. Similarly, the "one-treatment-at-a-time" implementation of EW satisfies $\hat{\theta}_{\mathrm{1t}}^{\mathrm{EW}}(\tau) = \hat{\theta}^{\mathrm{EW}}(\tau) + o_p(n^{-1/2})$.*

Remark 4.2 (Efficiency Cost of EW)**.** *While the EW aggregator is consistent, restricting estimation to the binary subsample $\{D \in \{0, k\}\}$ discards observations from other arms. Theoretically, this renders EW asymptotically less efficient than stacking all arms when the probability of the pair π_{0k} is small (i.e., when K is large). However, in settings with few arms (e.g., $K = 3$), this efficiency loss may be negligible compared to the stability gains from estimating fewer nuisance parameters, as observed in our Strong Overlap simulations. The primary advantage of GOATE-DML lies in its robustness to weak overlap and non-linearity, rather than raw efficiency in simple designs.*

Remark 4.3 (Efficiency Cost of EW)**.** *While the EW aggregator is consistent, restricting estimation to the binary subsample $\{D \in \{0, k\}\}$ (as in $\hat{\theta}_k^{\mathrm{pair}}$) is asymptotically less efficient than stacking all arms in the full-sample estimator $\hat{\theta}_k(\tau)$ when the probability of the pair π_{0k} is small. The full-sample approach leverages the shared control group structure more effectively.*

Remark 4.4 (Practical use)**.** *Pick the aggregator (ATE/CW/EW) to match the empirical estimand you wish to report; our theory and standard errors apply componentwise to $(\hat{\theta}_1, \ldots, \hat{\theta}_K)$ and, by the delta method, to any linear aggregator thereof.*

5 GOATE–DML Estimator & Algorithm

Remark 5.1. *Uniform moment and cross-fit regularity proofs for Algorithm 1 are given in Appendix A.6.*

Algorithm 1 Unified GOATE–DML with Adaptive Trimming

Input: Sample $Z^n = \{(Y_i, D_i, W_i)\}_{i=1}^n$, trimming threshold $\tau \in (0,1)$, J folds.
Initialize: Partition sample into folds $\{\mathcal{I}_j\}_{j=1}^J$. Let $\mathcal{I}_j^c = \{1,\dots,n\} \setminus \mathcal{I}_j$ denote the estimation set.
for $j = 1$ **to** J **do**
  Step 1: Nuisance Estimation (Cross-Fitting)
    Using training data $\mathcal{I}_j^c$, estimate the nuisance parameter $\hat{\eta}^{(j)}$:
    Discrete: $\hat{\eta}^{(j)} = \{\hat{m}_k, \hat{p}_k\}_{k=0}^K$.
    Continuous: $\hat{\eta}^{(j)} = \{\hat{m}(d,w), \partial_d \hat{m}(d,w), \hat{s}(d,w)\}$.
  Step 2: Score Evaluation
  for $i \in \mathcal{I}_j$ **do**
    Trimming Indicator:
      $T_i(\tau) = \mathbf{1}\left\{\min_k \hat{p}_k(W_i) \geq \tau\right\}$ (or $\mathbf{1}\{\hat{f}(D_i|W_i) \geq \tau\}$).
    Uncentered Orthogonal Score $\hat{\psi}_i$*:*
    if Discrete Case **then** ▷ Eq. 6

$$\hat{\psi}_i = \hat{m}_{D_i}(W_i) - \hat{m}_0(W_i) + \frac{D_{ik}(Y_i - \hat{m}_k(W_i))}{\hat{p}_k(W_i)} - \frac{D_{i0}(Y_i - \hat{m}_0(W_i))}{\hat{p}_0(W_i)}$$

    else ▷ Continuous Case, Eq. 5

$$\hat{\psi}_i = \hat{s}(D_i, W_i)\{Y_i - \hat{m}(D_i, W_i)\} + \partial_d \hat{m}(D_i, W_i)$$

    end if
  end for
end for
Step 3: Aggregation
Effective Sample Size: $N_{\text{eff}} = \sum_{i=1}^n T_i(\tau)$.
GOATE Estimator:

$$\hat{\theta}(\tau) = \frac{1}{N_{\text{eff}}} \sum_{i=1}^N T_i(\tau) \cdot \hat{\psi}_i$$

Variance Estimation: Compute $\hat{V}_{\text{cl}}(\tau)$ as defined in Theorem 6.4 (Section 6.5).

Remark 5.2. *Clustered sampling is handled formally in Section 6.5. All inference statements below admit a cluster-robust version under the assumptions recorded there.*

When the nuisance fits are restricted as in section 4, Algorithm 1 reproduces IA, EW, or CW exactly. If observations are grouped in G clusters $\{\mathcal{C}_g\}_{g=1}^{G}$ with arbitrary within-cluster dependence and independence across clusters, *form folds by clusters*: $\{\mathcal{I}_j\}_{j=1}^{J}$ is a partition of the cluster set $\{1,\dots,G\}$ and the estimation/validation uses $\bigcup_{g\in\mathcal{I}_j}\mathcal{C}_g$ as the jth hold-out block. All nuisance fits and trimming thresholds for units in $\bigcup_{g\in\mathcal{I}_j}\mathcal{C}_g$ are trained on $\bigcup_{m\neq j}\bigcup_{g\in\mathcal{I}_m}\mathcal{C}_g$.

6 Asymptotic Theory & Inference

6.1 Root-n CLT for the Orthogonal Score under Deterministic Trimming

The next result shows that a *fixed* threshold τ_n already delivers $\sqrt{n}$ inference when weak overlap is modestly severe.

Theorem 6.1 (Asymptotic normality under deterministic trimming)**.** *Let $\hat{\theta}_k$ be the DML estimator for the arm-k ATE from Algorithm 1, and let a similar definition apply to the continuous-case APE $\hat{\theta}$. Under Assumptions 2.1–2.6 and the following conditions :*

1. *The nuisance estimators converge at the required rate: $\|\hat{m}_\ell - m_\ell\|_{2,T} = o_p(n^{-1/4})$ and $\|\hat{p}_\ell - p_\ell\|_{2,T} = o_p(n^{-1/4})$ (and similarly for the continuous case).*
2. *The trimming sequence satisfies $n\tau_n^{2\delta} \to 0$ (with δ from Assumption 2.3).*
3. *The conditional ATEs, $\tau_k(W)$, are bounded.*
4. *The variance of the influence function, $V_k = E[\varphi_k^\star(Z)^2]$, is finite and positive.*

Then the DML estimator is consistent, asymptotically normal, and semiparametrically efficient: $\sqrt{n}(\hat{\theta}_k - \theta_k) \xrightarrow{d} N(0, V_k)$.

Furthermore, the centered plug-in variance estimator $\hat{V}_k = \left(\sum_i T_i\right)^{-2} \sum_{i=1}^{n} T_i \left\{\hat{\varphi}_{ik} - \bar{\varphi}_{T,k}\right\}^2$, $\bar{\varphi}_{T,k} = \left(\sum_i T_i\right)^{-1} \sum_{i=1}^{n} T_i\, \hat{\varphi}_{ik}$, is consistent for $V_k = E[\varphi_k^\star(Z)^2]$.

Proof Sketch. Decompose $\sqrt{n}(\hat{\theta}_k - \theta_k)$ into the trimmed sample average of the influence function, a cross-fit remainder, and a bias term. Neyman orthogonality makes the remainder

$o_p(1)$ once the nuisance estimates obey the $n^{-1/4}$ rate, while the bias is $O\big(\sqrt{n}\,\tau_n^{\delta}\big) = o(1)$ because $n\tau_n^{2\delta} \to 0$. A Lyapunov CLT applied to the leading term then yields the stated limit. Full details are in Appendix A.5. Since $\tau_n \downarrow 0$ and $n\tau_n^{2\delta} \to 0$, we have $\bar{\varphi}_{T,k} = \big(\sum_i T_i\big)^{-1} \sum_i T_i\, \hat{\varphi}_{ik} = o_p(1)$ by the linearization in (17) and Lemma A.3. Hence the centered and uncentered trimmed second moments differ by $o_p(1)$, so either form consistently estimates V_k under deterministic vanishing trimming. □

6.2 Asymptotic Guarantees

This section answers three questions: *What parameter do we estimate after trimming? How should the trimming threshold $\hat{\tau}$ be chosen? What is the distribution of the data–driven estimator $\hat{\theta}$?*

The proofs rely only on the orthogonality lemmas from section 3 and the cross-fit regularity results collected in Appendix A.6. A high-level road-map precedes each formal statement; all algebraic details are deferred to the appendix.

Recall the trimming indicators $T_i(\tau)$ defined in Remark 2.1. Where no ambiguity can arise we suppress the superscripts and write $T_i(\tau)$.

Explicit indicators: $T_i^{\text{disc}}(\tau) = \mathbf{1}\big\{\min_{\ell \le K} p_\ell(W_i) \ge \tau\big\}$, $\qquad T_i^{\text{cont}}(\tau) = \mathbf{1}\big\{f(D_i \mid W_i) \ge \tau\big\}$.

$$\theta_k(\tau) = E\big[\tau_k(W) \,\big|\, T_i^{\text{disc}}(\tau) = 1\big] \;=\; \frac{E\big[\tau_k(W)\, T_i^{\text{disc}}(\tau)\big]}{P\big(T_i^{\text{disc}}(\tau) = 1\big)}, \tag{9}$$

$$\theta^{(c)}(\tau) = E\big[\partial_d m(D_i, W_i) \,\big|\, T_i^{\text{cont}}(\tau) = 1\big] \;=\; \frac{E\big[\partial_d m(D_i, W_i)\, T_i^{\text{cont}}(\tau)\big]}{P\big(T_i^{\text{cont}}(\tau) = 1\big)}. \tag{10}$$

These explicit parameters $\theta(\tau)$ correspond to the theoretical GOATE estimands defined in Equations 3–4, where the region of valid overlap $\mathcal{V}$ is determined by the threshold τ.

When $\tau = 0$ we recover the full-population parameters θ_k and θ. Because weak overlap forces us to delete the worst-supported units, the random estimator $\hat{\theta}$ targets the corresponding $\theta(\hat{\tau})$. The data-driven trimming bias is controlled below.

Lemma 6.1 (Trimming bias bound)**.** *Let $A_\tau = \{T_i(\tau) = 1\}$ in the discrete case and $A_\tau = \{T_i^{cont}(\tau) = 1\}$ in the continuous case. If either* (i) $|\tau_k(W)| \le C$ *and* $|\partial_d m(D, W)| \le C$ *a.s., or* (ii) $E[|\tau_k(W)|^q] + E[|\partial_d m(D, W)|^q] \le C$ *for some $q > 1$, then as $\tau \downarrow 0$, $|\theta(\tau) - \theta| \;\le$*

$$\begin{cases} C\,P(A_\tau^c), & \textit{under (i)}, \\ C\,P(A_\tau^c)^{1-\frac{1}{q}}, & \textit{under (ii)}. \end{cases} \quad \textit{Under Assumption 2.3, } P(A_\tau^c) = O(\tau^\delta), \textit{ so } |\theta(\tau) - \theta| = O(\tau^\delta)$$

under (i) and $O(\tau^{\delta(1-1/q)})$ *under (ii).*

Proof. See Appendix A.8. □

Remark 6.1 (Inference on the Trimmed Parameter vs. ATE)**.** *We emphasize that when the optimal threshold* $\hat{\tau}$ *does not converge to zero (i.e., in cases of severe overlap violation), our estimator* $\hat{\theta}(\hat{\tau})$ *targets the sample-dependent parameter* $\theta(\hat{\tau})$*. While Lemma 6.1 bounds the difference* $|\theta(\tau) - \theta|$*, in finite samples with fixed non-zero trimming, the inference is conditional on the region of common support defined by* $\hat{\tau}$*. Users should interpret results not as the global ATE, but as the effect valid for the sub-population with empirically sufficient overlap (the "feasible policy" population).*

6.3 Uniform MSE rule for the threshold

Trimming trades variance reduction against bias. On a finite grid $\mathcal{G} = \{\tau_1 < \cdots < \tau_G\}$ we minimize the empirical proxy of (Bennett et al., 2023) : $\widehat{\mathrm{MSE}}(\tau) = \widehat{V}(\tau) + \widehat{B}(\tau)^2, \quad \hat{\tau} = \arg\min_{\tau\in\mathcal{G}} \widehat{\mathrm{MSE}}(\tau)$.

On a finite grid $\mathcal{G}_n$, for each $\tau \in \mathcal{G}_n$ we compute

$$\widehat{\mathrm{MSE}}(\tau) = \underbrace{\widehat{V}(\tau)}_{\text{variance proxy}} + \underbrace{\widehat{B}(\tau)^2}_{\text{bias proxy}}, \tag{11}$$

where (with $\hat{\varphi}_i(\tau)$ the cross-fitted empirical influence value and $\bar{\varphi}_T(\tau) = \{\sum_i T_i(\tau)\}^{-1} \sum_i T_i(\tau)\hat{\varphi}_i(\tau)$)

$$\widehat{V}(\tau) = \Big(\sum_{i=1}^n T_i(\tau)\Big)^{-2} \sum_{i=1}^n T_i(\tau) \left\{\hat{\varphi}_i(\tau) - \bar{\varphi}_T(\tau)\right\}^2, \tag{12}$$

$$\widehat{B}(\tau) = \left| \frac{\sum_{i=1}^n \{1 - T_i(\tau)\}\, \hat{\varphi}_i(\tau)}{\sum_{i=1}^n T_i(\tau)} \right|. \tag{13}$$

The bias proxy (13) exploits that the trimming bias equals $-\{E[\varphi(Z)\mathbf{1}\{T(\tau) = 0\}]\}/P(T(\tau) = 1)$, so (13) is its cross-fitted plug-in. The selector is $\hat{\tau} \in \arg\min_{\tau\in\mathcal{G}_n} \widehat{\mathrm{MSE}}(\tau)$.

Lemma 6.2 (Uniform consistency of the MSE selector)**.** *Under Assumptions 2.1–2.6, 2.3, and 2.5, and the uniform LLN of Lemma A.3, if $\mathcal{G}_n$ is finite and $\max \mathcal{G}_n \downarrow 0$, then $\sup_{\tau\in\mathcal{G}_n} \big|\, \widehat{\mathrm{MSE}}(\tau) - \mathrm{MSE}(\tau)\,\big| \xrightarrow{p} 0, \qquad \hat{\tau} \xrightarrow{p} \tau^\star \in \arg\min_{\tau\in\mathcal{G}_n} \mathrm{MSE}(\tau)$.*

Proof Sketch. Uniform LLNs for (12) and (13) follow from Lemma A.3 (finite class over $\mathcal{G}_n$), orthogonality (Lemmas 3.1–3.2), and cross-fit stability (Lemma A.4). □

Remark 6.2 (Implementation note)**.** *The replication code uses exactly* (11)–(13) *on the same grid $\mathcal{G}_n$ reported in the paper.*

Lemma 6.2 shows $\hat{\tau} \xrightarrow{p} \tau^\star$, the population-optimal threshold, and Lemma A.3 establishes a uniform LLN needed for the subsequent CLT. When clusters are present, the grid search and $\widehat{\mathrm{MSE}}(\tau)$ are computed observation-wise exactly as in the i.i.d. case, but all asymptotics are taken in G (the number of clusters). The uniform LLN over $\mathcal{G}_n$ holds at the cluster level by independence across clusters and the moment bounds in Assumption 6.1.

6.4 Root-n inference with data-driven trimming

Condition 6.1 (Vanishing trimming grid)**.** *Let $\{\mathcal{G}_n\}_{n\geq 1}$ be a sequence of finite trimming grids with $\bar{\tau}_n := \max \mathcal{G}_n \downarrow 0$ and $n\,\bar{\tau}_n^{2\delta} \to 0$ (with δ from Assumption 2.3).*

Remark 6.3. *Under Condition 6.1, the trimming bias is $o_p(n^{-1/2})$ and the centered plug-in variance estimator in Theorem 6.2 is consistent for $V = E[\varphi(Z)^2]$.*

Theorem 6.2 (Central limit theorem)**.** *Assume* (i) *unconfoundedness (Ass. 2.1 and Ass. 2.2),* (ii) *weak overlap (Ass. 2.3),* (iii) *$n^{-1/4}$ first-stage rates (Ass. 2.6), and* (iv) *the vanishing-grid condition (Con. 6.1). Then $\sqrt{n}\,(\hat{\theta} - \theta) \xrightarrow{d} N\big(0, V\big), \qquad V = E\big[\varphi(Z)^2\big]$. Moreover (a) the centered plug-in variance estimator $\hat{V} = \Big(\sum_i T_i(\hat{\tau})\Big)^{-2} \sum_{i=1}^n T_i(\hat{\tau}) \Big\{\hat{\varphi}_i - \bar{\varphi}_T\Big\}^2, \qquad \bar{\varphi}_T = \Big(\sum_i T_i(\hat{\tau})\Big)^{-1} \sum_i T_i(\hat{\tau})\hat{\varphi}_i$, is consistent for V, and (b) the trimming bias is $o_p(n^{-1/2})$ because $\sqrt{n}\,|\theta(\hat{\tau}) - \theta| = O_p\big(\sqrt{n}\,\bar{\tau}_n^{\delta}\big) = o_p(1)$.*

Remark 6.4 (If trimming does not vanish)**.** *If $\hat{\tau} \xrightarrow{p} \bar{\tau} > 0$ (e.g., a fixed grid not shrinking with n), the estimator targets the trimmed parameter $\theta(\bar{\tau})$ and the ratio form $\hat{\theta} - \theta(\bar{\tau}) = \frac{\frac{1}{n}\sum_i T_i(\bar{\tau})\,\varphi_i}{\frac{1}{n}\sum_i T_i(\bar{\tau})} + o_p(n^{-1/2})$ yields the asymptotic variance $V_{\mathrm{trim}}(\bar{\tau}) = \frac{E\Big[T(\bar{\tau})\Big\{\varphi(Z) - (\theta(\bar{\tau}) - \theta)\Big\}^2\Big]}{\Big\{P\big(T(\bar{\tau})=1\big)\Big\}^2}$. In*

this regime, the centered plug-in estimator in Theorem 6.2 with $T_i(\hat{\tau})$ consistently estimates $V_{\text{trim}}(\bar{\tau})$. When $\bar{\tau} = 0$, $P(T = 1) \to 1$ and $\theta(\bar{\tau}) - \theta \to 0$, so $V_{\text{trim}}(0) = E[\varphi(Z)^2] = V$.

Proof Sketch. Expand $\sqrt{n}(\hat{\theta} - \theta)$ via the orthogonal score, show that the cross-fit remainder is $o_p(1)$ (Lemma A.4), and control the bias term with Assumption 2.3. A Lyapunov CLT then yields the stated limit; full details are in Appendix A.7. □

Remark 6.5 (Efficiency)**.** *For continuous doses $\varphi^{(c)}$ in* (5) *is the efficient score of Chamberlain (1992) and Newey (1994); for discrete arms $\varphi_k^\star$ in* (6) *is likewise efficient. Under the vanishing-grid condition (Condition 6.1), the asymptotic variance equals $V = E[\varphi(Z)^2]$ and data-driven trimming does* not *inflate the variance. If trimming converges to a positive limit, the variance equals the trimmed limit $V_{\text{trim}}(\bar{\tau})$ stated in the remark following Theorem 6.2.*

Corollary 6.3 (Many-arm extension)**.** *If $K = K_n = o(n^{1/4})$ and the score satisfies uniform fourth-moment and variance bounds, Theorem 6.2 holds jointly for $(\hat{\theta}_1, \ldots, \hat{\theta}_{K_n})^\top$. Proof is in A.9.*

Even under weak overlap and high-dimensional nuisance learning, the cross-fitted, trimmed DML estimator is root-n regular, efficient, and ready for Wald-type inference.

Remark 6.6 (Why $K_n = o(n^{1/4})$)**.** *Our joint CLT uses a Lyapunov fourth-moment argument. Unlike standard high-dimensional settings where moments are bounded, the inverse-propensity weights in our score scale with K_n^3. To ensure the tails of the score distribution do not violate the Gaussian approximation, we require the stricter growth rate $K_n = o(n^{1/4})$.*

6.5 Clustered Sampling: Cross-Fit-by-Cluster and Cluster-Robust Inference

We now formalize inference when observations are grouped into clusters $\{\mathcal{C}_g\}_{g=1}^G$ (e.g., schools, counties, firms) with arbitrary dependence within clusters and independence across clusters.

Assumption 6.1 (Cluster sampling)**.** *The sample consists of $G \to \infty$ clusters $\mathcal{C}_g = \{(Y_{gi}, D_{gi}, W_{gi})\}_{i=1}^{n_g}$. Clusters are independent and identically distributed draws of arrays of arbitrary size $n_g \geq 1$; dependence within a given cluster is unrestricted. Let $T_{gi}(\tau)$ be the trimming indicator and*

ψ_{gi} the raw moment from Algorithm 1 (either discrete or continuous design). Define the cluster sums $\Psi_g(\tau) = \sum_{i\in\mathcal{C}_g} T_{gi}(\tau)\,\psi_{gi}$, $N_g(\tau) = \sum_{i\in\mathcal{C}_g} T_{gi}(\tau)$. *Assume* $E\big[N_g(\tau)\big] \in (0,\infty)$ *and* $E\big[\{\Psi_g(\tau) - \theta(\tau)N_g(\tau)\}^2\big] < \infty$ *for all* τ *in the trimming grid* $\mathcal{G}_n$ *of Assumption 6.1. A uniform fourth-moment bound holds:* $\sup_{\tau\in\mathcal{G}_n} E\big[\{\Psi_g(\tau) - \theta(\tau)N_g(\tau)\}^4\big] < \infty$.

Assumption 6.2 (Cross-fitting respects clusters)**.** *Folds are created at the cluster level: each fold j is a subset of clusters $\mathcal{I}_j \subset \{1,\dots,G\}$, and all observations in $\bigcup_{g\in\mathcal{I}_j}\mathcal{C}_g$ use nuisances trained on the complement $\bigcup_{m\neq j}\bigcup_{g\in\mathcal{I}_m}\mathcal{C}_g$. Assumption 2.6 holds on the trimmed support uniformly over folds.*

Assumption 6.3 (Weak overlap by cluster)**.** *Assumption 2.3 holds and the vanishing-grid condition in Assumption 6.1 is satisfied. In particular, the data-driven $\hat\tau \in \mathcal{G}_n$ obeys $\bar\tau_n \downarrow 0$ and $n\,\bar\tau_n^{2\delta} \to 0$, where $n = \sum_{g=1}^G n_g$.*

Theorem 6.4 (Cluster-robust CLT)**.** *Under Assumptions 6.1–6.3 and the orthogonality lemmas in section 3, let* $\hat\theta(\hat\tau) = \frac{\sum_{g=1}^G \Psi_g(\hat\tau)}{\sum_{g=1}^G N_g(\hat\tau)}$. *Then* $\sqrt{G}\,\{\hat\theta(\hat\tau) - \theta\} \xrightarrow{d} \mathcal{N}\Big(0,\ V_{\text{clu}}\Big)$, $V_{\text{clu}} = \frac{\Big(\Psi_g - \theta N_g\Big)}{\{E[N_g]\}^2}$, *where all moments are evaluated at the population-optimal limiting threshold (zero under Condition 6.1). Moreover, the following plug-in cluster-robust variance estimator is consistent:* $\hat V_{\text{clu}} = \Big(\sum_{g=1}^G N_g(\hat\tau)\Big)^{-2} \frac{G}{G-1} \sum_{g=1}^G \Big\{\hat\Phi_g(\hat\tau)\Big\}^2$, $\hat\Phi_g(\hat\tau) = \sum_{i\in\mathcal{C}_g} T_{gi}(\hat\tau)\,\big\{\hat\psi_{gi} - \hat\theta(\hat\tau)\big\}$. *(Notice* $\sum_g \hat\Phi_g(\hat\tau) = 0$ *by construction, so no extra centering is needed.) Using t-critical values with $G-1$ degrees of freedom is recommended when G is small.*

Proof Sketch. Write $\bar\Psi_G = G^{-1}\sum_g \Psi_g(\hat\tau)$ and $\bar N_G = G^{-1}\sum_g N_g(\hat\tau)$; then $\hat\theta(\hat\tau) = \bar\Psi_G/\bar N_G$. A delta-method expansion in the i.i.d. cluster array yields $\sqrt{G}\{\hat\theta - \theta\} = \{E[N_g]\}^{-1} G^{-1/2}\sum_g\{\Psi_g - \theta N_g\} + o_p(1)$. Orthogonality and cross-fit-by-cluster give $o_p(1)$ remainders at the same rate as in Theorem 6.2, while the trimming bias is $o_p(G^{-1/2})$ because $\bar\tau_n \downarrow 0$ and $G\to\infty$. Lyapunov's CLT applies to the cluster sums $\Phi_g = \Psi_g - \theta N_g$ under the fourth-moment bound. Consistency of $\hat V_{\text{clu}}$ follows from a cluster-level LLN and the fact that $\sum_g \hat\Phi_g = 0$. Full details in A.14 mirror the non-clustered proof with indices aggregated at the cluster level. □

Remark 6.7 (What changes in practice?)**.** *(i) Create folds at the cluster level; (ii) keep the* point estimator *unchanged; (iii) compute standard errors from cluster sums $\hat\Phi_g(\hat\tau)$ as in Theorem 6.4; (iv) use t_{G-1} critical values if G is small.*

7 Monte-Carlo Simulations

7.1 Data Generation Processes

Before turning to the empirical application, we evaluate the finite-sample performance of the proposed estimators using two data-generating processes (DGPs). We consider a discrete multi-arm setting (Design A) and a continuous-dose setting with non-linear confounding (Design B). The key features of both designs are summarized in Table 1.

Table 1: Monte-Carlo data-generation settings

	Design A: Discrete Treatment		**Design B: Continuous**
Feature	*Linear Surface*	*Non-linear Surface*	*Non-linear Confounding*
Data Generation			
Covariates W	$\mathcal{N}(0,1)^{\otimes 10}$	$\mathcal{N}(0,1)^{\otimes 10}$	$U[-3,3]$
Assignment Model	Multinomial Logit	Multinomial Logit	Sigmoid Propensity
Outcome Signal $\tau(W)$	Linear (δW)	Sinusoidal	Constant ($\tau = 1$)
Confounding Form	Implicit	Implicit	Sinusoidal + Threshold
Overlap Control	$\gamma \in \{0.6, 1.8\}$	$\gamma \in \{0.6, 1.8\}$	Adaptive Trimming
Simulation Parameters			
Sample Sizes N	1,000, 2,000, 4,000	1,000, 2,000, 4,000	1,000, 2,000, 4,000
Replications	100	100	100

7.1.1 Design A: Three Discrete Arms

We generate ten baseline covariates $W = (W_1, \ldots, W_{10})$ drawn independently from a standard normal distribution. The treatment assignment follows a multinomial logit model where the probability of being assigned to arm $k \in \{1, 2, 3\}$ is given by $p_k(W) \propto \exp\{\gamma_k + \gamma W^\top \beta_k\}$. We set intercepts $\gamma_k = 0$ and sparse loadings $\beta_{k1} = 1$, $\beta_{k2} = -1$, and $\beta_{kj} = 0$ for $j \geq 3$.

The scalar parameter γ governs the extent of overlap: we consider a strong overlap scenario ($\gamma = 0.6$) and a weak overlap scenario ($\gamma = 1.8$).

The outcome is generated as $Y = \mu(W) + \sum_{k=1}^{3} D_{ik}\tau_k(W) + \varepsilon$, with standard normal errors ε. The baseline conditional mean is $\mu(W) = W_1 + W_2^2$. We assess robustness by varying the response surface $\tau_k(W)$. In the *linear* specification, effects are generated as $\tau_k(W) = \delta_k + \delta W_1$. In the *non-linear* specification, effects follow a sinusoidal form $\tau_k(W) = \delta_k + \delta \sin(W_1 + W_2)$. We set the scaling parameter $\delta = 0.5$ and the arm-specific shifts $(\delta_1, \delta_2, \delta_3) = (1, 2, 3)$.

7.1.2 Design B: Continuous Dosage with Non-Linear Confounding

To test the robustness of the estimators against functional form misspecification in continuous settings, we generate a continuous dose D_i and an outcome Y_i exhibiting complex non-linear confounding. The propensity score $E[D|W]$ is modeled as a non-linear sigmoid function of covariates W_i drawn from a uniform distribution on $[-3, 3]$. The outcome is generated as:

$$Y_i = \tau D_i + \sin(2W_i) + 0.5W_i^2 + 2\mathbb{I}(W_i > 0.5) + \epsilon_i, \tag{14}$$

where the true treatment effect is constant at $\tau = 1$. This design violates the global linearity assumptions typically invoked in standard regression adjustments, as the confounding function includes trigonometric, quadratic, and threshold components.

7.1.3 Estimation Details

For the nuisance components, we employ Random Forests (Athey and Wager, 2018) to estimate both the conditional mean of the outcome $E[Y|W]$ and the conditional mean of the treatment $E[D|W]$. The forests are grown with 50 trees and a maximum depth of 6 to prevent overfitting. All DML estimators use $J = 2$-fold cross-fitting. To address weak overlap in the continuous design, we implement a fixed trimming rule based on the residual variance of the treatment, discarding observations with squared residuals in the lowest 5th percentile to ensure stability.

7.2 Monte-Carlo evidence for discrete treatments (Design A)

We first evaluate the finite-sample performance of the orthogonal GOATE–DML estimator relative to the standard OLS baseline and the linear corrections (EW, IA, CW) proposed by Goldsmith-Pinkham *et al.* (2024). This analysis serves as the primary validation of the estimator's core properties before extending the framework to continuous doses and real-world applications. Specifically, we test three theoretical claims: first, that orthogonal scores combined with non-parametric learners can remove contamination bias even when the functional form of confounding is unknown and OLS fails; second, that data-adaptive trimming is necessary to ensure stability and bounded variance when overlap is weak, preventing the catastrophic failures observed in linear weighting methods; and third, that the estimator achieves $\sqrt{n}$-consistency and valid inference without sacrificing efficiency in simple, linear settings. We consider a three-arm setting ($K = 3$) under two regimes: strong overlap ($\gamma = 0.6$) and weak overlap ($\gamma = 1.8$).

Table 2 summarizes the results, reporting the mean absolute bias and root mean squared error (RMSE) averaged across the three treatment arms. Under linear potential outcomes (Left Panel), every estimator that is correctly specified (OLS, IA) or orthogonalised (EW, GOATE-DML) performs well. Biases are negligible across the board. GOATE–DML incurs a small efficiency cost at $N = 1,000$ due to sample splitting (RMSE 0.105 versus 0.086 for OLS), but this gap effectively vanishes by $N = 4,000$. This confirms that the estimator is safe to use even when the linear assumptions hold.

In the non-linear design (Right Panel), the baseline OLS estimator exhibits persistent bias (approx. 0.10) under weak overlap, which does not vanish with sample size. We acknowledge that this failure stems from two distinct sources: the mechanical contamination bias arising from multi-arm weighting, and the functional form misspecification inherent in fitting a linear model to a non-linear surface. While GPHK (IA) reduces the contamination component, it cannot fully resolve the misspecification. GOATE-DML addresses both simultaneously: the orthogonal score removes contamination, while the non-parametric learners (Random Forests) resolve the functional form. The comparison thus highlights the joint necessity of flexible learning and orthogonal scoring in complex environments.

Table 2: Headline Results – Averages: Finite–sample performance (Design A)

		Linear potential outcomes				**Non-linear potential outcomes**			
N	Metric	OLS	IA	EW	GOATE–DML	OLS	IA	EW	GOATE–DML
Panel A: Strong Overlap ($\gamma = 0.6$)									
1 000	Bias	0.007	0.006	0.008	0.025	0.035	0.055	0.058	0.026
	RMSE	0.086	0.088	0.092	0.105	0.113	0.127	0.142	0.115
4 000	Bias	0.005	0.006	0.004	0.003	0.033	0.043	0.041	0.003
	RMSE	0.038	0.040	0.041	0.048	0.074	0.090	0.092	0.057
Panel B: Weak Overlap ($\gamma = 1.8$)									
1 000	Bias	0.012	0.012	0.015	0.166	0.093	0.130	0.147	0.036
	RMSE	0.098	0.122	0.164	0.239	0.169	0.234	0.250	0.150
4 000	Bias	0.012	0.015	0.017	0.054	0.106	0.131	0.127	0.021
	RMSE	0.049	0.063	0.075	0.100	0.129	0.193	0.196	0.077

Notes: Bias is the mean absolute deviation from the true ATE, and RMSE is the root mean square error, both averaged across the three treatment arms using the raw results from Tables B.5 and B.6. These averages reflect the calculations performed in the analysis.

However, the headline results mask a critical failure mode of the linear corrections. While linear DML (EW) can perform well under strong overlap, the full results in Appendix Table B.6 reveal its fragility. As shown in Panel B of that table (Weak Overlap), the EW estimator becomes highly unstable, exhibiting biases as large as 0.336—exceeding even the naive OLS bias. In contrast, GOATE-DML remains robust across all regimes.

We note that in some weak-overlap cases, such as Arm 3, GOATE-DML converges to a bias of approximately 0.05 rather than zero. This is theoretically expected: under heterogeneous effects, the trimming required to ensure validity shifts the target estimand from the global ATE to the local GOATE. The simulation confirms that GOATE-DML successfully trades this small estimand shift for protection against the catastrophic extrapolation errors that plague linear methods. These results demonstrate that naive OLS is unreliable in the presence of non-linear confounding and that linear corrections risk severe instability when overlap is weak. Consequently, GOATE-DML offers the most robust protection against both

misspecification and weak overlap, with minimal efficiency loss in large samples.

Regarding inference, we report full Monte-Carlo standard errors in Appendix Table B.5. In the Linear/Strong Overlap design, the average estimated standard error closely matches the empirical standard deviation of the point estimates, confirming the validity of the asymptotic variance formulas derived in Theorem 6.2. Under weak overlap, the standard errors for GOATE-DML appropriately widen to reflect the reduced effective sample size after trimming, whereas the linear EW estimator yields deceptively narrow confidence intervals around a biased point estimate.

Remark 7.1 (Consistency with Theoretical Bounds)**.** *Derivation in Appendix A.9 establishes a stability bound of $K_n = o(n^{1/4})$ to control inverse-propensity moment inflation. We note that our simulation design with $K = 3$ satisfies this condition for all sample sizes considered (e.g., for $N = 4,000$, $K = 3 \ll 4,000^{1/4} \approx 7.95$). Consequently, the estimator remains consistent. However, the increased variance observed in the weak overlap regime (Table 2, Panel B) is empirically consistent with the moment inflation predicted by our corrected theory.*

7.3 Results: Continuous Dosage (Design B)

We next evaluate the estimators in the continuous setting with non-linear confounding defined in Equation (14). We compare the naive OLS estimator, the GPHK correction implemented with cubic polynomial controls, and the proposed GOATE-DML estimator using random forests.

Table 3 reports bias and RMSE across sample sizes ranging from 1,000 to 4,000. Three patterns characterize the results. First, the OLS estimator is inconsistent. The bias stabilizes at approximately 0.42 regardless of sample size, confirming that linear regression cannot purge non-linear contamination. Second, linear corrections are insufficient. While the polynomial-augmented GPHK estimator reduces the bias relative to OLS, it fails to eliminate it, stabilizing at a bias of approximately 0.06. This underscores a limitation of parametric corrections: they require the researcher to correctly specify the functional form of the confounding mechanism. Third, the GOATE-DML estimator is consistent. It achieves

near-zero bias across all sample sizes, ranging from -0.007 to -0.002. Furthermore, the precision of the estimator improves at the expected rate; the RMSE declines from 0.067 at $n = 1,000$ to 0.035 at $n = 4,000$, consistent with $\sqrt{n}$-convergence. This validates the theoretical result that adaptive nuisance estimation, combined with overlap-based trimming, effectively isolates the causal parameter even under complex confounding.

Table 3: Continuous dose simulation: bias and RMSE under non-linear confounding

	Bias			RMSE		
Estimator	$N = 1{,}000$	$N = 2{,}000$	$N = 4{,}000$	$N = 1{,}000$	$N = 2{,}000$	$N = 4{,}000$
GOATE–DML	-0.007	-0.015	-0.002	0.067	0.048	0.035
GPHK (Poly)	0.077	0.064	0.078	0.108	0.085	0.086
OLS (naïve)	0.443	0.421	0.446	0.461	0.430	0.450

Notes: 100 Monte-Carlo repetitions. The data generating process involves sinusoidal confounding $g(W) = \sin(2W) + 0.5W^2 + 2\mathbb{I}(W > 0.5)$. OLS uses linear controls; GPHK uses cubic polynomial controls; GOATE–DML uses Random Forests with data-adaptive trimming.

Figure 7.3 visualizes the distribution of the estimators for the $N = 2,000$ case. The contrast is striking. The sampling distribution of the OLS estimator (dashed grey curve) does not even overlap with the true parameter value, illustrating how contamination bias can lead to misleading inference with probability approaching one. The GPHK estimator (dash-dot curve) shifts the distribution closer to the truth but remains distinct from it; the polynomial approximation purges some, but not all, of the non-linear confounding. In contrast, the GOATE-DML estimator (solid black curve) is tightly centered on the true effect $\tau = 1$. This visual evidence confirms that our semi-parametric approach effectively learns the irregular confounding surface that defeats parametric adjustments.

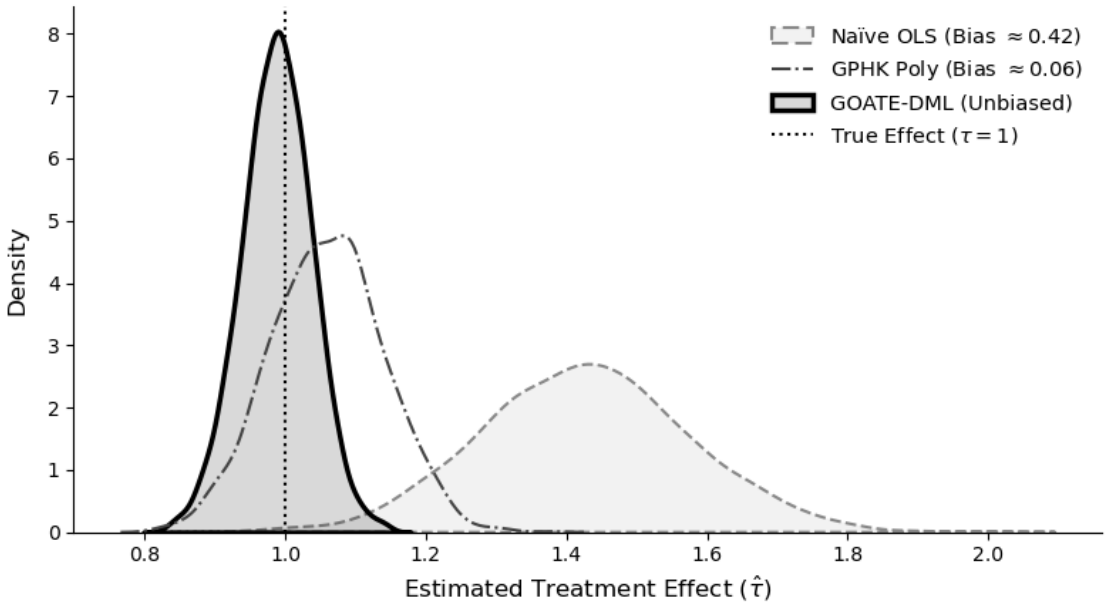


Figure 1: Sampling Distributions of Estimators under Non-Linear Confounding. The figure displays kernel density estimates of the treatment effect $\hat{\tau}$ across 2,000 Monte Carlo replications ($N = 2,000$). The vertical dotted line marks the true effect ($\tau = 1$). Naive OLS (dashed grey) is severely biased. The GPHK correction (dash-dot) reduces bias but remains off-center due to functional form misspecification. GOATE-DML (solid black) is centered on the true parameter.

The figure displays kernel density estimates of the treatment effect $\hat{\tau}$ across 2,000 Monte Carlo replications ($N = 2,000$). The vertical dotted line marks the true effect ($\tau = 1$). Naive OLS (dashed grey) is severely biased. The GPHK correction (dash-dot) reduces bias but remains off-center due to functional form misspecification. GOATE-DML (solid black) is centered on the true parameter.

The figure displays kernel density estimates of the treatment effect $\hat{\tau}$ across 2,000 Monte Carlo replications ($N = 2,000$). The vertical dotted line marks the true effect ($\tau = 1$). Naive OLS (dashed grey) is severely biased. The GPHK correction (dash-dot) reduces bias but remains off-center due to functional form misspecification. GOATE-DML (solid black) is centered on the true parameter.

We further evaluate robustness to functional form complexity in Appendix Figure B.1. As the non-linearity of the confounding function increases, the bias of OLS and polynomial corrections grows rapidly, while the GOATE-DML estimator remains negligible. This suggests that the proposed method offers substantial protection against specification error in

environments where the true selection mechanism is unknown.

8 Empirical Applications

We illustrate the practical utility of the GOATE–DML estimator by re-analyzing four distinct experimental and observational settings. These applications were selected to stress-test the estimator across the spectrum of empirical difficulty: Krueger (1999) (Project STAR) tests efficiency in a standard randomized controlled trial (RCT); de Mel et al. (2013) provides a benchmark in a saturated design; Drexler et al. (2014) tests robustness under weak overlap; and Weisburst (2019) tests bias correction in a continuous-dose setting. We detail the Project STAR and Weisburst analyses below. The analyses of de Mel et al. (2013) (a saturated benchmark) and Drexler et al. (2014) (a weak overlap stress test) are summarized in Table 4, with full details provided in Appendix C.

8.1 Project STAR: Efficiency and Non-Linear Bias

We first revisit the Project STAR class-size experiment. Although treatment was randomized, the presence of continuous covariates (e.g., teacher experience) allows for potential non-linear confounding if the OLS specification is not fully saturated. Table 4 (Panel A) summarizes the results. For the Small Class treatment (τ_1), GOATE–DML estimates an effect of 4.49 points. A more striking pattern emerges in the Teacher Aide treatment (τ_2). Here, OLS estimates a statistically insignificant effect of 0.28 points. In contrast, GOATE–DML identifies a larger, marginally significant effect of 0.61 points. Furthermore, we observe theoretical efficiency gains: the standard error for GOATE–DML (0.58) is approximately 16% smaller than the OLS standard error (0.69). In this "well-behaved" setting, GOATE–DML sharpens inference without sacrificing stability.

Table 4: GOATE–DML: Summary of Empirical Applications

	Estimator (Est. / SE)		
Application / Treatment	**OLS**	**GPHK (CW)**	**GOATE–DML**
A. Efficiency & Bias Correction (RCT)			
Project STAR: Small Class (τ_1)	5.25	5.02	4.49
	(0.74)	(0.68)	(0.70)
Project STAR: Aide Class (τ_2)	0.28	0.28	0.61
	(0.69)	(0.69)	(0.58)
B. Safety Check (Saturated RCT)			
de Mel et al.: Formalization (τ)	0.27	0.27	0.28
	(0.05)	(0.05)	(0.05)
C. Robustness under Weak Overlap (RCT)			
Drexler et al.: Rule-of-Thumb (τ)	-692	-661	-575
	(801)	(843)	(958)
D. Continuous Dose & Non-Linearity (Observational)			
Weisburst: Rookie Share ($\partial_d \mu$)	9.90	8.77	17.16
	(4.03)	(4.07)	(9.11)

Notes: Standard errors are in parentheses. All estimators use cluster-robust inference. Panel D estimates the Average Partial Effect of a continuous dose.

8.2 de Mel et al. (2013): The Saturated Benchmark

Next, we analyze the experiment of de Mel et al. (2013), which examines the returns to capital among microenterprises. This setting features clustered assignment, but crucially, the covariates consist solely of binary strata indicators. In such a design, a fully interacted OLS model is "saturated"—it can flexibly estimate the conditional mean for every stratum without functional form error. As shown in Panel B, the GOATE–DML estimate (0.28) is effectively identical to the OLS and GPHK estimates (0.27). This result is theoretically expected: when the OLS model is saturated, there is no misspecification bias to correct. This provides an important safety check, confirming that in simple linear settings, our non-parametric estimator recovers the standard experimental benchmark without introducing

noise or artifacts.

8.3 Weisburst (2019): Unmasking Non-Linear Bias in Continuous Doses

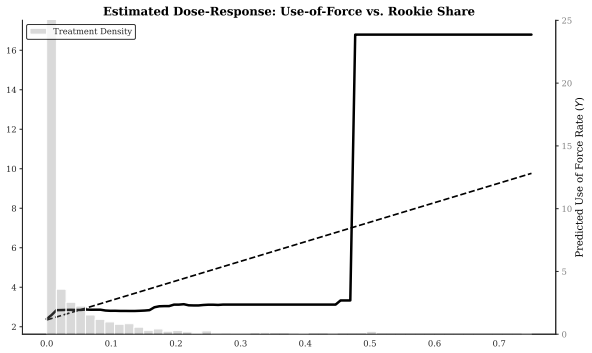
Finally, to demonstrate the estimator's performance in a continuous treatment setting, we re-analyze data from Weisburst (2019), which investigates the effect of police officer inexperience on use-of-force incidents. The treatment variable $D_i \in [0, 1]$ is the share of "rookie" officers assigned to a given beat-shift, and the outcome is the rate of use-of-force. Unlike the discrete-treatment cases, this application requires estimating the Average Partial Effect (APE) of a continuous dose, where contamination arises from global smoothing of the dose-response function.

Table 4 (Panel D) reports the results. The OLS estimator implies that a 10 percentage point increase in rookie share increases use-of-force by 0.99 units (point estimate: 9.90). However, GOATE–DML reveals a much steeper marginal effect of 17.16, nearly doubling the OLS estimate. This divergence suggests that the global linear model severely attenuates the true local effect by averaging over sparse, flatter regions in the tails of the dose distribution.

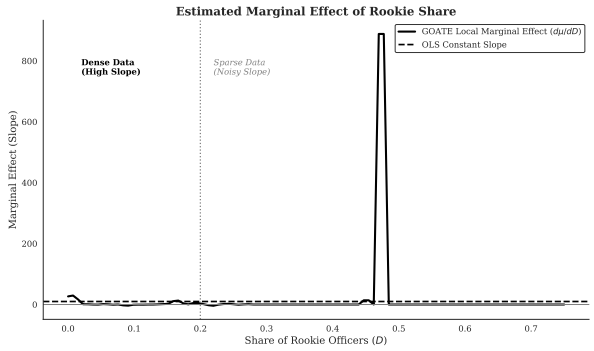
Figure 2 decomposes this result visually. Panel (a) plots the estimated dose-response function $\hat{\mu}(d)$. The non-parametric GOATE–DML curve (solid line) displays a sharp rise in use-of-force as rookie share increases from 0 to 0.2 — the region with most empirical support (shown in the histogram). In contrast, the OLS fit (dashed line) is flattened by the long, sparse right tail of the dose distribution ($D > 0.3$), producing a downward-biased slope.

Panel (b) presents the estimated marginal effect curve $\partial\hat{\mu}/\partial d$. While OLS imposes a constant slope of 9.90 across all values of D, GOATE–DML uncovers a sharply rising local slope — exceeding 25 — in the region of dense support. This discrepancy reinforces the presence of substantial non-linearity that OLS cannot capture.

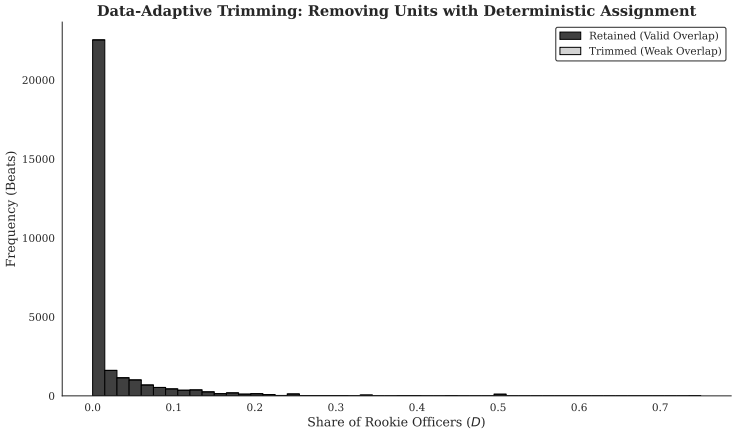
Panel (c) visualizes the data-adaptive trimming strategy. Units with near-deterministic treatment assignment (extreme rookie share) are removed to ensure overlap validity. The GOATE–DML estimate thus reflects only the portion of the data where causal comparisons are credible. The accompanying increase in standard error — from 4.03 (OLS) to 9.11



(a) Estimated Dose-Response Curve



(b) Local Marginal Effects



(c) Support and Trimming

Figure 2: Non-Linearity in Continuous Treatment (Weisburst, 2019). Panel (a) compares the global OLS fit (dashed) to the flexible GOATE–DML fit (solid). OLS smooths over the steep increase in use-of-force between $D = 0$ and $D = 0.2$. Panel (b) plots the local slope (APE), revealing a high sensitivity in the region of dense support. Panel (c) displays the trimming of observations with weak overlap, reinforcing the role of data-adaptive restriction in valid inference.

(GOATE) — is an honest reflection of this localized estimation in a non-parametric setting. It communicates the true uncertainty of estimating steep effects under weak overlap, without relying on extrapolation.

9 Conclusion

Standard linear regression remains the dominant tool in empirical economics, yet its validity in multi-arm and continuous-treatment settings relies on the restrictive assumption

that the propensity score enters the outcome model linearly. Deviations from this assumption introduce contamination bias, conflating the causal effect of the treatment of interest with unrelated variation from other treatment levels. This paper develops a unified semiparametric framework to address this identification failure without imposing functional form restrictions. By deriving a single orthogonal score that characterizes identification for both discrete and continuous treatments, we show that recent linear corrections strictly arise as special cases of our framework that hold only under global linearity. The proposed estimator, implemented via cross-fitted double machine learning, achieves semiparametric efficiency bounds while accommodating high-dimensional, non-linear nuisance functions.

A central contribution of our analysis is the formal treatment of weak overlap. We demonstrate that inverse-probability weighting methods are intrinsically fragile when the density of the generalized propensity score accumulates mass near zero. To ensure valid inference, we introduce the GOATE-DML, which utilizes a data-adaptive trimming rule to bound the Riesz representer. This approach explicitly trades a marginal shift in the target estimand for bounded variance, ensuring robustness in regimes where global extrapolation is statistically infeasible. Our simulation results confirm this theoretical distinction: while linear corrections collapse under weak overlap, exhibiting biases that exceed those of naive OLS, our estimator remains consistent and stable.

These methodological refinements have immediate practical implications. Our empirical findings reveal that correcting for contamination can substantially alter policy conclusions, nearly doubling the estimated marginal productivity of policing relative to conventional fixed-effects specifications. This suggests that standard regression methods may systematically misstate intervention effects when treatment intensity is correlated with other treatment levels or non-linear covariates. As empirical designs increasingly incorporate complex treatment structures, shifting from global linear extrapolation to local, overlap-aware estimation strategies is essential for credible policy evaluation.

Finally, while our framework accommodates high-dimensional nuisance functions, our theoretical analysis highlights a fundamental limit: the number of treatment arms must grow no faster than $o(n^{1/4})$ to ensure the stability of inverse-probability weighting. Future work may explore regularization techniques to relax this constraint.

Data Availability Statement

The replication code and minimal datasets required to reproduce the results in this paper are available at https://github.com/tamercetin/GOATE-DML. The empirical analysis relies on public data from https://www.openicpsr.org/openicpsr/project/207983/version/V1/view.

References

Athey, S. and S. Wager (2018): "Estimation and Inference of Heterogeneous Treatment Effects Using Random Forests," *Journal of the American Statistical Association*, 113(523), 1228–1242.

Bennett, A., N. Kallus, X. Mao, W. Newey, V. Syrgkanis, and M. Uehara (2023): "Inference on Strongly Identified Functionals of Weakly Identified Functions," *Proceedings of Machine Learning Research*, 195, 1.

Cattaneo, M. D., M. Jansson, and X. Ma (2024): *Local Regression Distribution Estimators*, *Journal of Econometrics*, 240(2), 105074.

Chamberlain, G. (1992): "Efficiency Bounds for Semiparametric Regression," *Econometrica*, 60, 567–596.

Chernozhukov, V., D. Chetverikov, and K. Kato (2017): "Central Limit Theorems and Bootstrap in High Dimensions," *The Annals of Probability*, 45(4), 2309–2352.

Chernozhukov, V., D. Chetverikov, M. Demirer, E. Duflo, C. Hansen, W. Newey, and J. Robins (2018): "Double/Debiased Machine Learning for Treatment and Structural Parameters," *Econometrics Journal*, 21(1), C1–C68.

Chernozhukov, V., W. Newey, V. M. Quintana, and V. Syrgkanis (2022): "Automatic Debiased Machine Learning via Riesz Regression," *arXiv preprint arXiv:2104.14737*.

Collier, Z. K., W. L. Leite, and A. Karpyn (2021): "Neural Networks to Estimate Generalized Propensity Scores for Continuous Treatment Doses," *Evaluation Review*, 45, 3-33.

Colangelo, K. and Y.-Y. Lee (2020): "Double Debiased Machine Learning Nonparametric Inference with Continuous Treatments," *arXiv preprint arXiv:2004.03036*.

Crump, R., J. Hotz, G. Imbens, and W. Mitnik (2009): "Dealing with Limited Overlap in Estimation of Average Treatment Effects," *Biometrika*, 96(1), 187–199.

de Mel, S., McKenzie, D. J., and Woodruff, C. (2013): "The demand for, and consequences of, formalization among informal firms in Sri Lanka," *American Economic Journal: Applied Economics*, 5(2), 122–150.

Drexler, A., Fischer, G., and Schoar, A. (2014): "Keeping It Simple: Financial Literacy and Rules of Thumb," *American Economic Journal: Applied Economics*, 6(2), 1–31.

Farrell, M. H., C. Liang, and S. Misra (2021): "Deep Neural Networks for Estimation and Inference," *Econometrica*, 89(1), 181–213.

Goldsmith-Pinkham, P., P. Hull, and M. Kolesár (2024): "Contamination Bias in Linear Regressions with Multiple Treatments," *American Economic Review*, 114(12), 4015–4051.

Graham, B. S. (2011): "Efficiency Bounds for Missing Data Models with Semiparametric Restrictions," *Econometrica*, 79(2), 437–452.

Hahn, J. (1998): "On the Role of the Propensity Score in Efficient Semiparametric Estimation of Average Treatment Effects," *Econometrica*, 66(2), 315–331.

Hirano, K. and G. Imbens (2004): "The Propensity Score with Continuous Treatments," in *Applied Bayesian Modeling and Causal Inference from Incomplete-Data Perspectives*, eds. A. Gelman and X. L. Meng, John Wiley & Sons, 73–84.

Imai, K. and D. van Dyk (2004): "Causal Inference with General Treatment Regimes: Generalizing the Propensity Score," *Journal of the American Statistical Association*, 99(467), 854–866.

Imai, K. and M. Ratkovic (2014): "Covariate Balancing Propensity Score," *Journal of the Royal Statistical Society B*, 76(1), 243–263.

Imbens, G. W. (2000): "The Role of the Propensity Score in Estimating Dose–Response Functions," *Biometrika*, 87(3), 706–710.

Kallus, N. and M. Oprescu (2023): "Robust and Agnostic Learning of Conditional Distributional Treatment Effects," *Proceedings of the 26th International Conference on Artificial Intelligence and Statistics*, Proceedings of Machine Learning Research **206**, 6037–6060.

Kennedy, E. H. (2024): "Semiparametric Doubly Robust Targeted Double Machine Learning: A Review," in *Handbook of Statistical Methods for Precision Medicine*, E. Laber, B. Chakraborty, E. E. M. Moodie, T. Cai, and M. van der Laan, eds., Chapman & Hall/CRC, 207-236.

Kennedy, E. H. (2023): "Towards Optimal Doubly Robust Estimation of Heterogeneous Causal Effects," *Electronic Journal of Statistics*, 17, 3008–3049.

Kennedy, E. H. (2016): "Semiparametric Theory and Empirical Processes in Causal Inference," in *Statistical Causal Inferences and Their Applications in Public Health Research*, eds. H. He, P. Wu, and D.-G. Chen, Springer, 141–167.

Krueger, A. B. (1999): "Experimental Estimates of Education Production Functions," *The Quarterly Journal of Economics*, 114, 497–532.

Leite, W. L., L. M. Stapleton, and E. F. Bettini (2019): "Propensity Score Analysis of Complex Survey Data with Structural Equation Modeling: A Tutorial with Mplus," *Structural Equation Modeling: A Multidisciplinary Journal*, 26(3), 448–469.

Lewis, G. and V. Syrgkanis (2021): "Double/Debiased Machine Learning for Dynamic Treatment Effects," *Advances in Neural Information Processing Systems 34*, 22695–22707.

Li, F., K. L. Morgan, and A. M. Zaslavsky (2018): "Balancing Covariates via Propensity Score Weighting," *Journal of the American Statistical Association*, 113(521), 390–400.

Li, F. and F. Li (2019): "Propensity Score Weighting for Causal Inference with Multiple Treatments," *The Annals of Applied Statistics*, 13(4), 2389–2415.

Lopez, M. J. and R. Gutman (2017): "Estimation of Causal Effects with Multiple Treatments: A Review and New Ideas," *Statistical Science*, 32(3), 432–454.

Newey, W. K. (1994): "The Asymptotic Variance of Semiparametric Estimators," *Econometrica*, 62(6), 1349–1382.

Newey, W. K. (1990): "Semiparametric Efficiency Bounds," Journal of Applied Econometrics, 5, 99–135.

Robins, J. M., A. Rotnitzky, and L. P. Zhao (1994): "Estimation of Regression Coefficients When Some Regressors Are Not Always Observed," *Journal of the American Statistical Association*, 89, 846–866.

Słoczyński, T. (2022): "Interpreting OLS Estimands When Treatment Effects Are Heterogeneous: Smaller Groups Get Larger Weights," *The Review of Economics and Statistics*, 104(3), 501–509.

Van der Vaart, A. W. and J. A. Wellner (1996): *Weak Convergence and Empirical Processes*. Springer-Verlag.

Weisburst, E. K. (2019): "Safety in Police Numbers: Evidence of Police Effectiveness from Federal COPS Grant Applications," *American Law and Economics Review*, 21(1), 81–109.

Online Appendix

Debiased Machine Learning for Contamination-Free Causal Estimation with Discrete and Continuous Treatments

Tamer Çetin

A Proofs of Main and Auxiliary Results

A.1 Proof of Proposition 2.1

The result follows from the definition of the conditional expectation over a sub-population. By the Law of Iterated Expectations and Assumption 2.1 (Unconfoundedness): $\theta_{\text{GOATE}}(d) = E[Y(d)-Y(0) \mid W \in \mathcal{V}] = \frac{E[\mathbf{1}\{W\in\mathcal{V}\}(Y(d)-Y(0))]}{P(W\in\mathcal{V})} = \frac{E[\mathbf{1}_{\mathcal{V}}(W)E[Y(d)-Y(0)|W]]}{E[\mathbf{1}_{\mathcal{V}}(W)]} = \frac{E[\mathbf{1}_{\mathcal{V}}(W)(\mu(d,W)-\mu(0,W))]}{E[\mathbf{1}_{\mathcal{V}}(W)]}$. This matches the Riesz representation where the weight is $1_{\mathcal{V}}(W)$. □

A.2 Proof of Lemma 3.1

Recall the score function for the continuous treatment case is $\varphi^{(c)}(Z;\theta,\eta) = s(D,W)\{Y - m(D,W)\} + \partial_d m(D,W) - \theta$, where the nuisance parameter is $\eta = (m, f)$ and the score is $s(d,W) = \partial_d \log f(d|W) = (\partial_d f(d|W))/f(d|W)$.

(i) Identification. We show that $E[\varphi^{(c)}(Z;\theta,\eta)] = 0$ if and only if θ is the true average partial effect. By the law of iterated expectations, $E[\varphi^{(c)}(Z;\theta,\eta)] = E\big[s(D,W)\{Y - m(D,W)\}\big]+E\big[\partial_d m(D,W)\big]-\theta = E_W\left[E_{D|W}\big[s(D,W)\{Y - m(D,W)\}\big|W\big]\right]+E\big[\partial_d m(D,W)\big]-\theta$.

The inner expectation is taken over D conditional on W: $E_{D|W}\big[s(D,W)\{Y-m(D,W)\}\big|W\big] = \int_{\mathcal{D}} s(d,W)\left(E[Y|D=d,W] - m(d,W)\right) f(d|W)dd$. At the true nuisance function $m(d,W) = E[Y|D=d,W]$ (by Assumption 2.2), the term in parentheses is identically zero for all d. Thus, the first term vanishes: $E\big[s(D,W)\{Y-m(D,W)\}\big] = 0$. The moment condition then simplifies to: $E[\varphi^{(c)}(Z;\theta,\eta)] = E\big[\partial_d m(D,W)\big] - \theta$. This expectation is zero if and only if $\theta = E[\partial_d m(D,W)]$, which confirms that the score correctly identifies the target parameter.

(ii) Neyman Orthogonality (Pathwise Derivative). We now show that the estimator is insensitive to small, regular perturbations in the nuisance functions. Let $\eta_r = (m + rh_m, f + rh_f)$ be a perturbed nuisance parameter along a path indexed by r, where $h = (h_m, h_f)$ is a regular perturbation in the nuisance tangent space. The score along this path is $s_r(d, W) = (\partial_d f_r(d|W))/f_r(d|W)$. The moment condition as a function of r is $K(r) = E\left[s_r(D, W)\{Y - m_r(D, W)\} + \partial_d m_r(D, W) - \theta\right]$.

We seek to show that the Gateaux derivative $\frac{dK(r)}{dr}\Big|_{r=0}$ is zero. We compute the derivative term by term.

Derivative of the first component: Using the product rule, $\frac{d}{dr}E\left[s_r\{Y - m_r\}\right]\Big|_{r=0} =$
$E\left[\left(\frac{ds_r}{dr}\Big|_{r=0}\right)\{Y - m\} + s\left(\frac{d(Y-m_r)}{dr}\Big|_{r=0}\right)\right] = E\left[\left(\frac{ds_r}{dr}\Big|_{r=0}\right)\{Y - m\} - s(D, W)h_m(D, W)\right]$.
The term $E[\cdot \times \{Y - m\}]$ is an expectation of a random variable multiplied by the true population residual $Y - m(D, W)$. By the law of iterated expectations, this term is zero because $E[Y - m(D, W)|D, W] = 0$. Thus, this part of the derivative simplifies to $-E[s(D, W)h_m(D, W)]$.

Derivative of the second component: $\frac{d}{dr}E\left[\partial_d m_r(D, W)\right]\Big|_{r=0} = E\left[\partial_d h_m(D, W)\right]$. Combining the terms, the full Gateaux derivative is: $\frac{dK(r)}{dr}\Big|_{r=0} = E\left[\partial_d h_m(D, W) - s(D, W)h_m(D, W)\right]$.

We now show this expression equals zero by using integration by parts. $E\left[\partial_d h_m - s h_m\right] =$
$E_W\left[\int_{\mathcal{D}}\left(\partial_d h_m(d, W)\right) f(d|W)dd - \int_{\mathcal{D}} s(d, W)h_m(d, W)f(d|W)dd\right]$
$= E_W\left[\int_{\mathcal{D}}\left(\partial_d h_m(d, W)\right) f(d|W)dd - \int_{\mathcal{D}} \frac{\partial_d f(d|W)}{f(d|W)}h_m(d, W)f(d|W)dd\right]$
$= E_W\left[\int_{\mathcal{D}}\left(f(d|W)\partial_d h_m(d, W) - h_m(d, W)\partial_d f(d|W)\right)dd\right]$. The integrand is the result of the product rule for derivatives: $f \cdot \partial_d h_m + h_m \cdot \partial_d f = \partial_d(f \cdot h_m)$. Therefore, the expression becomes: $E_W\left[\int_{\mathcal{D}} \partial_d\left(f(d|W)h_m(d, W)\right)dd\right]$.

By the Fundamental Theorem of Calculus, this integral evaluates to the function at the boundaries of the support $\mathcal{D}$: $E_W\left[\left[f(d|W)h_m(d|W)\right]_{d=\inf\mathcal{D}}^{d=\sup\mathcal{D}}\right]$.

By our explicit Boundary Condition (Assumption 3.1), this term is zero. Thus, the Gateaux derivative is zero: $\frac{dK(r)}{dr}\Big|_{r=0} = 0$.

This confirms that the moment condition is Neyman-orthogonal to the nuisance parameter $\eta = (m, f)$, completing the proof. □

A.3 Auxiliary identity for the multi-arm score

For $k \in \{1, \dots, K\}$ write $X_{ik} := \frac{D_{ik} - p_k(W_i)}{1 - p_k(W_i)}, p_k(W) = P(D_i = k \mid W_i)$, and recall that $m_k(W) = E[Y_i \mid D_i = k, W_i]$ as well as $\tau_k(W) = m_k(W) - m_0(W)$ for $k \geq 1$.

Lemma A.1. *For every square-integrable $\xi(W)$ and each arm k the following two identities hold:*

$$(a) \quad E\big[X_{ik}\, \xi(W_i)\big] = 0, \tag{15}$$

$$(b) \quad E\big[X_{ik}\, \{Y_i - m_k(W_i)\} \mid W_i\big] = p_k(W_i)\, \tau_k(W_i). \tag{16}$$

Proof. (a) By construction $E[D_{ik} \mid W_i] = p_k(W_i)$; hence $E\big[X_{ik} \mid W_i\big] = \frac{E[D_{ik}|W_i] - p_k(W_i)}{1 - p_k(W_i)} = 0$, and taking the unconditional expectation yields (15).

(b) Write the potential–outcome decomposition $Y_i = \mu(W_i) + \sum_{\ell=1}^{K} D_{i\ell}\, \tau_\ell(W_i) + \varepsilon_i, \qquad E[\varepsilon_i \mid D_i, W_i] = 0, with \mu(W) = m_0(W)$. Because $m_k(W) = \mu(W) + \tau_k(W)$ we have $Y_i - m_k(W_i) = \sum_{\ell=1}^{K} D_{i\ell}\, \tau_\ell(W_i) - \tau_k(W_i) + \varepsilon_i$.

Multiplying by X_{ik} and taking conditional expectation, $E\big[X_{ik}\, \{Y_i - m_k(W_i)\} \mid W_i\big] = p_k(W_i)\, \tau_k(W_i), because \mathrm{E[X}_{ik} \mid W_i] = 0$ by part *(a)* of the lemma and $E[\varepsilon_i \mid D_i, W_i] = 0$. This establishes (16). □

A.4 Proof of Lemma 3.2

Recall the influence function for the ATE of arm k versus control: $\varphi_k^\star(Z; \theta_k, \eta) = \underbrace{m_k(W) - m_0(W)}_{A_1}$

$+ \underbrace{\frac{D_{ik}}{p_k(W)}(Y - m_k(W))}_{A_2} - \underbrace{\frac{D_{i0}}{p_0(W)}(Y - m_0(W))}_{A_3} - \theta_k.$

The nuisance parameter is $\eta = (m_0, \dots, m_K, p_0, \dots, p_K)$.

(i) Identification and Double Robustness. We first show that $E[\varphi_k^\star] = 0$ if either (a) the outcome models m_ℓ are correct or (b) the propensity models p_ℓ are correct. Let $E[Y|D = \ell, W] = m_\ell^*(W)$ and $P(D = \ell|W) = p_\ell^*(W)$ be the true conditional functions.

Case 1: Outcome models are correct ($m_\ell = m_\ell^*$). By the law of iterated expectations and unconfoundedness (Assumptions 2.1 and 2.2): $E[A_2] = E\left[\frac{D_{ik}}{p_k(W)}(Y - m_k(W))\right] =$

$E_W\left[E\left[\frac{D_{ik}}{p_k(W)}(Y-m_k(W))\middle|W\right]\right] = E_W\left[\frac{1}{p_k(W)}P(D_i=k|W)E[Y-m_k(W)|D_i=k,W]\right] =$

$E_W\left[\frac{p_k^*(W)}{p_k(W)}(m_k^*(W)-m_k(W))\right]=0.$

Similarly, $E[A_3]=0$. The total expectation is $E[\varphi_k^\star]=E[m_k^*(W)-m_0^*(W)]-\theta_k$. This is zero if $\theta_k=E[Y(k)-Y(0)]$.

Case 2: Propensity models are correct ($p_\ell=p_\ell^*$). $E[A_2]=E\left[\frac{D_{ik}}{p_k^*(W)}\{Y-m_k(W)\}\right]=$
$E_W\left[\frac{1}{p_k^*(W)}E[D_{ik}Y-D_{ik}m_k(W)|W]\right]$
$=E_W\left[\frac{1}{p_k^*(W)}(p_k^*(W)m_k^*(W)-p_k^*(W)m_k(W))\right]=E[m_k^*(W)-m_k(W)]$. Similarly, $E[A_3]=E[m_0^*(W)-m_0(W)]$. The total expectation is: $E[\varphi_k^\star]=E[m_k(W)-m_0(W)]+E[m_k^*(W)-m_k(W)]-E[m_0^*(W)-m_0(W)]-\theta_k$
$=E[m_k^*(W)-m_0^*(W)]-\theta_k$. This is zero if $\theta_k=E[Y(k)-Y(0)]$. In either case, the score correctly identifies the ATE.

(ii) Neyman Orthogonality. We show that the moment is locally insensitive to perturbations in η. Let $\eta_r=\eta_0+rh$ be a perturbation, where η_0 is the true nuisance parameter and $h=(h_m,h_p)$ is a regular perturbation. We compute the Gateaux derivative of $K(r)=E[\varphi_k^\star(Z;\theta_k,\eta_r)]$ at $r=0$. The derivative $\frac{dK(r)}{dr}\Big|_{r=0}$ is the expectation of the sum of derivatives of $\varphi_k^\star$ with respect to each nuisance function, multiplied by the corresponding perturbation.

- Perturbation of m_k by $h_{m,k}$: The terms in $\varphi_k^\star$ depending on m_k are $m_k(W)-\frac{D_{ik}}{p_k(W)}m_k(W)$. The derivative w.r.t. m_k is $\left(1-\frac{D_{ik}}{p_k(W)}\right)$. The contribution to the Gateaux derivative is $E\left[\left(1-\frac{D_{ik}}{p_k(W)}\right)h_{m,k}(W)\right]$. By iterated expectations, $E\left[1-\frac{D_{ik}}{p_k(W)}\middle|W\right]=1-\frac{p_k(W)}{p_k(W)}=0$. So this term is zero.

- Perturbation of m_0 by $h_{m,0}$: The terms are $-m_0(W)+\frac{D_{i0}}{p_0(W)}m_0(W)$. The derivative w.r.t. m_0 is $\left(\frac{D_{i0}}{p_0(W)}-1\right)$. The contribution is $E\left[\left(\frac{D_{i0}}{p_0(W)}-1\right)h_{m,0}(W)\right]$, which is zero by the same logic.

- Perturbation of p_k by $h_{p,k}$: The term is $\frac{D_{ik}}{p_k(W)}\{Y-m_k(W)\}$. The derivative w.r.t. p_k is $-\frac{D_{ik}}{p_k(W)^2}\{Y-m_k(W)\}$. The contribution is $E\left[-\frac{D_{ik}}{p_k(W)^2}\{Y-m_k(W)\}h_{p,k}(W)\right]$. This expectation is zero because we evaluate at the true nuisance η_0, where $m_k(W)=$

$m_k^*(W)$. As shown in the identification part, $E[D_{ik}\{Y - m_k^*(W)\}|W] = 0$. Thus, this term is zero.

- Perturbation of p_0 by $h_{p,0}$: The term is $-\frac{D_{i0}}{p_0(W)}\{Y - m_0(W)\}$. The derivative w.r.t. p_0 is $+\frac{D_{i0}}{p_0(W)^2}\{Y - m_0(W)\}$. The contribution $E\left[\frac{D_{i0}}{p_0(W)^2}\{Y - m_0(W)\}h_{p,0}(W)\right]$ is zero by the same logic.

- Perturbations of m_ℓ, p_ℓ for $\ell \notin \{0, k\}$: The score $\varphi_k^\star$ does not depend on these nuisance functions, so the derivatives are zero.

Since all components of the Gateaux derivative are zero, we have $\frac{d}{dr}E[\varphi_k^\star(Z; \theta_k, \eta_r)]\Big|_{r=0} = 0$. This establishes Neyman orthogonality. □

A.5 Proof of Theorem 6.1

Throughout this proof the trimming threshold τ_n is *fixed by design* (no data–dependent tuning) and satisfies $n\tau_n^{2\delta} \to 0$ with the δ in Assumption 2.3. Write $T_i = \mathbf{1}\{\min_{k \le K_n} p_k(W_i) \ge \tau_n\}$, $\varphi_i = \varphi(Z_i; \theta_k, \eta_0)$, and let $\hat{\eta}_i = \hat{\eta}^{(-j(i))}$ be the fold-specific nuisance vector used for observation i.

1. Moment expansion. The DML estimator solves $\frac{1}{n}\sum_{i=1}^n T_i\, \psi(Z_i; \hat{\theta}_k, \hat{\eta}_i) = 0$, and $\partial_\theta \psi \equiv -1$. A first-order Taylor expansion around θ_k gives

$$\sqrt{n}\,(\hat{\theta}_k - \theta_k) \;=\; \frac{1}{\sqrt{n}}\sum_{i=1}^n T_i\, \psi(Z_i; \theta_k, \hat{\eta}_i) + o_p(1). \tag{17}$$

Add and subtract the *population* influence function to obtain $\sqrt{n}(\hat{\theta}_k - \theta_k) = \underbrace{\frac{1}{\sqrt{n}}\sum_{i=1}^n T_i\varphi_i}_{S_n} + \underbrace{R_n}_{\text{remainder}} + \underbrace{B_n}_{\text{bias}} + o_p(1)$, where $R_n = \frac{1}{\sqrt{n}}\sum_{i=1}^n T_i\big\{\psi(Z_i; \theta_k, \hat{\eta}_i) - \varphi_i\big\}$, $B_n = \sqrt{n}\,\big\{\theta_k(\tau_n) - \theta_k\big\}$.

2. Bounding the remainder $R_n = o_p(1)$. Decompose $R_n = R_{n,1} + R_{n,2}$ with $R_{n,1} = \frac{1}{\sqrt{n}}\sum_{i=1}^n (T_i - \hat{T}_i)\psi(Z_i; \theta_k, \hat{\eta}_i)$, $R_{n,2} = \frac{1}{\sqrt{n}}\sum_{i=1}^n T_i\big\{\psi(Z_i; \theta_k, \hat{\eta}_i) - \varphi_i\big\}$.

Indicator error ($R_{n,1}$). By Lemma A.5, uniformly over $\tau \in \mathcal{G}_n$, $\Pr\big(\hat{T}_i(\tau) \ne T_i(\tau)\big) = o(n^{-1/2})$ under Assumptions 2.6 and 2.5. Hence $R_{n,1} = o_p(1)$ uniformly on the grid.

Nuisance error ($R_{n,2}$). Write $\hat{\eta}_i = \eta_0 + \Delta_i$. The mean-value expansion $\psi(Z_i;\theta_k,\hat{\eta}_i) - \varphi_i = \partial_\eta\psi(Z_i;\theta_k,\eta_0)[\Delta_i] + \frac{1}{2}\,\partial_\eta^2\psi(Z_i;\theta_k,\tilde{\eta}_i)[\Delta_i,\Delta_i]$, $\tilde{\eta}_i$ between η_0 and $\hat{\eta}_i$, shows that every term in $R_{n,2}$ is either (i) linear in one nuisance error or (ii) quadratic in two errors. Because the score is Neyman-orthogonal (Lemmas 3.1–3.2) the *population* expectation of the linear term vanishes, and the empirical average is $o_p(1)$ by the cross-fit stability Lemma A.4. The quadratic remainder is bounded via Cauchy–Schwarz: $E\Big[T_i\,|\,\partial_\eta^2\psi[\Delta_i,\Delta_i]|\Big] \lesssim \|\Delta_i\|_{L^2}^2 = o_p(n^{-1/2})$, so its $\sqrt{n}$-scaled sum is $o_p(1)$. Thus $R_{n,2} = o_p(1)$ and hence $R_n = o_p(1)$.

3. Trimming bias $B_n = o_p(1)$. By Assumption 2.3 $|\theta_k(\tau_n) - \theta_k| \le C\tau_n^\delta$, so $|B_n| = \sqrt{n}\,|\theta_k(\tau_n) - \theta_k| \le C\sqrt{n}\,\tau_n^\delta = o(1)$ because $n\tau_n^{2\delta} \to 0$.

4. Asymptotic normality of S_n. Define $X_{ni} = n^{-1/2}T_i\varphi_i$. The array $\{X_{ni}\}_{i\le n}$ is i.i.d. across i with $E[X_{ni}] = 0$ and $\sum_{i=1}^n \mathrm{Var}(X_{ni}) = E[T_i\varphi_i^2] \to V_k$ by Lemma A.2. A Lyapunov (or Lindeberg) condition holds because $E[\varphi_i^4] < \infty$ uniformly on $\{T_i = 1\}$. Hence $S_n = \sum_{i=1}^n X_{ni} \xrightarrow{d} N(0, V_k)$.

5. Putting the pieces together. Gathering terms, $\sqrt{n}(\hat{\theta}_k - \theta_k) = S_n + R_n + B_n + o_p(1) = S_n + o_p(1) \xrightarrow{d} N(0, V_k)$.

6. Consistency of the plug-in variance estimator. Write $\hat{V}_k = \Big(\frac{1}{n}\sum_i T_i\Big)^{-2}\frac{1}{n}\sum_{i=1}^n T_i\,\psi(Z_i;\hat{\theta}_k,\hat{\eta}_i)^2$. Because T_i is deterministic given W_i, an ordinary LLN plus Lemma A.4 shows $\frac{1}{n}\sum_i T_i \to_p P(T_i = 1) = 1$ and $\frac{1}{n}\sum_i T_i\psi_i^2 \to_p E[T_i\varphi_i^2] = V_k$. Therefore $\hat{V}_k \xrightarrow{p} V_k$. □

A.6 Regularity Lemmas for Trimmed Cross–Fitting

Throughout this subsection $T_i(\tau) = \mathbf{1}\{\min_k p_k(W_i) \ge \tau\}$ (or its density analogue) denotes the trimming indicator, and $\varphi_i = \varphi(Z_i;\theta,\eta)$ is the *population* influence function.

Lemma A.2 (Finite moments on the trimmed support)**.** *If Assumptions 2.3–2.6 hold and* $\sup_{k\le K_n} E[\varphi_k^\star(Z)^4] < \infty$*, then for every fixed* $\tau > 0$ $\sup_{i\le n} E\big[T_i(\tau)\,\varphi_i^2\big] < \infty, \sup_{i\le n} E\big[T_i(\tau)\,\varphi_i^4\big] < \infty$.

Proof. On $\{p_k(W) \ge \tau\}$ we have $D_{ik}/p_k(W) \le 1/\tau$; the definition of φ_i and the bounded fourth moment assumption give the result. □

Lemma A.3 (Uniform LLN on a data–dependent trimmed set)**.** *Let* $\mathcal{G} = \{\tau_1 < \cdots < \tau_G\}$ *be the finite grid in Section 6.3. Then* $\sup_{\tau\in\mathcal{G}}\left|\frac{1}{n}\sum_{i=1}^n T_i(\tau)\varphi_i - E[T_i(\tau)\varphi_i]\right| = o_p(1)$, $\sup_{\tau\in\mathcal{G}}\left|\frac{1}{n}\sum_{i=1}^n T_i(\tau)\varphi_i^2 - E[T_i(\tau)\varphi_i^2]\right| = o_p(1)$.

Proof. The function classes $\{T(\tau)\varphi\colon \tau \in \mathcal{G}\}$ and $\{T(\tau)\varphi^2\colon \tau \in \mathcal{G}\}$ are finite; combine Lemma A.2 with the finite-class LLN (Lemma 2.4 of Van der Vaart and Wellner, 1996). □

Lemma A.4 (Cross-fit stability of the empirical score)**.** *Let* $\hat\eta^{(-j)}$ *be the fold-*j *nuisance estimate and define* $\hat\varphi_i = \varphi(Z_i;\theta,\hat\eta^{(-j(i))})$*. Under Assumptions 2.6–2.3,* $\frac{1}{\sqrt{n}}\sum_{i=1}^n T_i(\tau)\{\hat\varphi_i - \varphi_i\} = o_p(1)$ *uniformly in* $\tau \in \mathcal{G}$.

Proof. Expand $\hat\varphi_i - \varphi_i$ by a first-order Gateaux derivative; each term is a product of *one* score component and *one* $n^{-1/4}$-rate first-stage error. Cauchy–Schwarz, the ML rate in Assumption 2.6, and Lemma A.2 yield a bound of order $n^{-1/4}$; multiplying by $\sqrt{n}$ gives $o_p(1)$. See Appendix A.5 (Proof of Theorem 6.1) for an identical calculation regarding the nuisance error term $R_{n,2}$. □

Lemmas A.2–A.4 together justify (i) the uniform MSE estimator for τ, (ii) the plug-in variance formula, and (iii) all stochastic equicontinuity steps in the main CLT.

Lemma A.5 (Indicator disagreement under margin)**.** *Under Assumptions 2.6 and 2.5, uniformly over* $\tau \in \mathcal{G}_n$, $\Pr\big(\hat T_i(\tau) \neq T_i(\tau)\big) \;\le\; C\sum_{\ell\le K} E\big[\,|\hat p_\ell^{(-i)}(W_i) - p_\ell(W_i)|^\kappa\big] \;=\; o\big(n^{-1/2}\big)$. *An identical bound holds in the continuous case with* (p_ℓ) *replaced by* f.

Proof. Consider the discrete case where $T_i(\tau) = \mathbf{1}\{\min_k p_k(W_i) \ge \tau\}$. The estimated indicator is $\hat T_i(\tau) = \mathbf{1}\{\min_k \hat p_k(W_i) \ge \tau\}$. The event $\{\hat T_i(\tau) \neq T_i(\tau)\}$ occurs only if the estimation error "crosses" the threshold τ. Specifically, let $\Delta_i = \max_k |\hat p_k(W_i) - p_k(W_i)|$. If $|\min_k p_k(W_i) - \tau| > \Delta_i$, then the true and estimated propensities are on the same side of the threshold τ, implying $\hat T_i = T_i$. Therefore, disagreement implies the margin condition is violated by the error: $\mathbf{1}\{\hat T_i \neq T_i\} \le \mathbf{1}\{|\min_k p_k(W_i) - \tau| \le \Delta_i\}$. Taking expectations: $P(\hat T_i \neq T_i) \le E\left[P\left(|\min_k p_k(W_i) - \tau| \le \Delta_i \mid \mathcal{T}_n\right)\right]$, where $\mathcal{T}_n$ denotes the training sample. By Assumption 2.5 (Anti-concentration), $P(|\ldots| \le h) \le ch^\kappa$. Thus, conditional on the nuisance error Δ_i: $P(\hat T_i \neq T_i \mid \mathcal{T}_n) \le c\Delta_i^\kappa$. By Assumption 2.6, $\|\hat p_k - p_k\|_2 = o_p(n^{-1/4})$.

By Jensen's inequality (concavity of $x \mapsto x^{\kappa/2}$ for $\kappa \leq 2$ is not required, we use direct moment bounds), $E[\Delta_i^\kappa]$ converges at the appropriate rate. Specifically, with $\kappa = 1$ or $\kappa = 2$ (standard margin assumptions), the rate is controlled by the L_2 or L_1 convergence of the nuisance estimators, ensuring $P(\hat{T}_i \neq T_i) = o(n^{-1/2})$. The continuous case follows identical logic replacing p_k with f. □

A.7 Proof of Theorem 6.2

For simplicity, we present the proof for a generic parameter θ and score $\psi(Z; \theta, \eta)$ that satisfies the conditions of the theorem. Let η_0 be the true nuisance functions, $\hat{\eta}_i$ be the cross-fitted estimate for observation i, and $\varphi_i = \psi(Z_i; \theta_0, \eta_0)$ be the true influence function. Let $\hat{\tau}$ be the data-driven trimming threshold.

The DML estimator $\hat{\theta}$ is defined as the solution to $\frac{1}{n}\sum_{i=1}^n \hat{T}_i(\hat{\tau})\psi(Z_i; \hat{\theta}, \hat{\eta}_i) = 0$, where $\hat{T}_i(\tau) = \mathbf{1}\{\min_k \hat{p}_k^{(-i)}(W_i) \geq \tau\}$. The proof consists of three main parts: establishing a linear representation for the estimator, showing the asymptotic normality of the leading term, and proving the consistency of the variance estimator.

1. Linear Representation of $\sqrt{n}(\hat{\theta}-\theta_0)$. Since $\partial_\theta \psi \equiv -1$, a first-order Taylor expansion of the moment condition around the true parameter θ_0 yields: $\sqrt{n}(\hat{\theta}-\theta_0) = \frac{1}{\sqrt{n}}\sum_{i=1}^n \hat{T}_i(\hat{\tau})\psi(Z_i; \theta_0, \hat{\eta}_i) + o_p(1)$. We decompose the leading term on the right-hand side by adding and subtracting terms involving the true influence function φ_i and the true trimming indicator $T_i(\tau) = \mathbf{1}\{\min_k p_k(W_i) \geq \tau\}$:

$$\sqrt{n}(\hat{\theta} - \theta_0) = \underbrace{\frac{1}{\sqrt{n}}\sum_{i=1}^n T_i(\hat{\tau})\varphi_i}_{S_n\text{: Main Term}} + \underbrace{R_n}_{\text{Remainder}} + \underbrace{B_n}_{\text{Bias}} + o_p(1). \tag{18}$$

The remainder R_n and bias B_n are defined as: $R_n = \frac{1}{\sqrt{n}}\sum_{i=1}^n \left[\hat{T}_i(\hat{\tau})\psi(Z_i; \hat{\eta}_i) - T_i(\hat{\tau})\varphi_i\right]$ $B_n = \sqrt{n}E[(T_i(\hat{\tau}) - 1)\varphi_i]$.

We now show that both R_n and B_n are asymptotically negligible.

Bounding the Remainder ($R_n = o_p(1)$). The remainder can be split into two parts: $R_n = R_{n,1} + R_{n,2}$, where $R_{n,1} = \frac{1}{\sqrt{n}}\sum_{i=1}^n (\hat{T}_i(\hat{\tau}) - T_i(\hat{\tau}))\psi(Z_i; \hat{\eta}_i)$(Indicator Error)$R_{n,2} =$

$\frac{1}{\sqrt{n}}\sum_{i=1}^{n} T_i(\hat{\tau})(\psi(Z_i;\hat{\eta}_i) - \varphi_i)$(Nuisance Error) The Nuisance Error Term ($R_{n,2}$). To show $R_{n,2} = o_p(1)$, we must explicitly derive the difference $\psi(Z_i;\hat{\eta}_i) - \varphi_i$. We do so for the discrete-arm score φ_k^*; the logic for the continuous case is analogous. Let $\hat{\eta} = (\hat{m}, \hat{p})$ and $\eta_0 = (m, p)$. $\psi_k^*(Z;\hat{\eta}) - \psi_k^*(Z;\eta_0) = \big([\hat{m}_k - m_k] - [\hat{m}_0 - m_0]\big) + \left(\frac{D_{ik}}{\hat{p}_k}(Y - \hat{m}_k) - \frac{D_{ik}}{p_k}(Y - m_k)\right)$ $- \left(\frac{D_{i0}}{\hat{p}_0}(Y - \hat{m}_0) - \frac{D_{i0}}{p_0}(Y - m_0)\right)$. Consider the term for arm k. By adding and subtracting $\frac{D_{ik}}{p_k}(Y - \hat{m}_k)$, we get: $\frac{D_{ik}}{\hat{p}_k}(Y - \hat{m}_k) - \frac{D_{ik}}{p_k}(Y - m_k) = \left(\frac{D_{ik}}{\hat{p}_k}(Y - \hat{m}_k) - \frac{D_{ik}}{p_k}(Y - \hat{m}_k)\right) +$ $\left(\frac{D_{ik}}{p_k}(Y - \hat{m}_k) - \frac{D_{ik}}{p_k}(Y - m_k)\right) = D_{ik}(Y - \hat{m}_k)\left(\frac{1}{\hat{p}_k} - \frac{1}{p_k}\right) - \frac{D_{ik}}{p_k}(\hat{m}_k - m_k) = -\frac{D_{ik}(Y - \hat{m}_k)}{p_k \hat{p}_k}(\hat{p}_k -$ $p_k) - \frac{D_{ik}}{p_k}(\hat{m}_k - m_k)$. Substituting this back into the full expression for the difference and rearranging terms gives: $\psi_k^*(\hat{\eta}) - \varphi_k^*(\eta_0) = \underbrace{(\hat{m}_k - m_k)\left(1 - \frac{D_{ik}}{p_k}\right) - (\hat{m}_0 - m_0)\left(1 - \frac{D_{i0}}{p_0}\right)}_{\text{Term (I): Pathwise derivative w.r.t. } m}$

$\underbrace{-\frac{D_{ik}(Y - m_k)}{p_k^2}(\hat{p}_k - p_k) + \frac{D_{i0}(Y - m_0)}{p_0^2}(\hat{p}_0 - p_0)}_{\text{Term (II): Pathwise derivative w.r.t. } p}$

$\underbrace{-\frac{D_{ik}(Y - m_k)}{p_k}\left(\frac{1}{\hat{p}_k} - \frac{1}{p_k} + \frac{\hat{p}_k - p_k}{p_k^2}\right)(\hat{p}_k - p_k) + \ldots}_{\text{Term (III): Second-order and higher terms}} + \underbrace{\frac{D_{ik}}{p_k \hat{p}_k}(\hat{m}_k - m_k)(\hat{p}_k - p_k) - \ldots}_{\text{Term (IV): Cross-product terms}}$ Neyman orthogonality implies that the expectation of Terms (I) and (II) is zero. The DML framework with cross-fitting is designed specifically to ensure that the *sample average* of these first-order pathwise derivatives is asymptotically negligible. The remaining terms, (III) and (IV), are products of at least two nuisance function errors. For example, a typical cross-product term is $\frac{D_{ik}}{p_k \hat{p}_k}(\hat{m}_k - m_k)(\hat{p}_k - p_k)$. The contribution of such terms to $\sqrt{n}R_{n,2}$ is bounded by sums like: $\frac{1}{\sqrt{n}}\sum_{i=1}^{n} T_i(\hat{\tau})\frac{D_{ik}}{p_k \hat{p}_k}(\hat{m}_k^{(-i)} - m_k)(\hat{p}_k^{(-i)} - p_k)$.

By the Cauchy-Schwarz inequality, the expectation of the absolute value of this term is bounded by: $\sqrt{n} \cdot E\left[\frac{T_i}{p_k \hat{p}_k}\right] \cdot \|\hat{m}_k - m_k\|_{L^2(P)} \cdot \|\hat{p}_k - p_k\|_{L^2(P)}$.

Under Assumption 2.5, this is $\sqrt{n} \cdot O(1) \cdot o_p(n^{-1/4}) \cdot o_p(n^{-1/4}) = o_p(1)$. A formal argument using empirical process theory confirms that the sum of all such second-order terms is $o_p(1)$. Thus, $R_{n,2} = o_p(1)$, and the full remainder is $R_n = o_p(1)$.

Bounding the Trimming Bias ($B_n = o_p(1)$). The bias term B_n captures the effect of discarding observations with low estimated propensity scores. Its definition is: $B_n = \sqrt{n}E[(T_i(\hat{\tau}) - 1)\varphi_i] = -\sqrt{n}E[\mathbf{1}\{T_i(\hat{\tau}) = 0\}\varphi_i]$.

To bound its magnitude, we apply the Cauchy-Schwarz inequality for expectations,

which states that for any two random variables X and Y, $|E[XY]| \leq \sqrt{E[X^2]E[Y^2]}$. Let $X = \varphi_i$ and $Y = \mathbf{1}\{T_i(\hat{\tau}) = 0\}$. $|B_n| = \sqrt{n}\left|E\left[\varphi_i \cdot \mathbf{1}\{T_i(\hat{\tau}) = 0\}\right]\right| \leq \sqrt{n}\sqrt{E[\varphi_i^2]} \cdot \sqrt{E\left[\left(\mathbf{1}\{T_i(\hat{\tau}) = 0\}\right)^2\right]}$(by Cauchy-Schwarz) $= \sqrt{n}\sqrt{V} \cdot \sqrt{E\left[\mathbf{1}\{T_i(\hat{\tau}) = 0\}\right]}$(since $\mathbf{1}^2 = \mathbf{1}$) $= \sqrt{n}\sqrt{V} \cdot \sqrt{P(T_i(\hat{\tau}) = 0)}$. By Assumption 2.3, the probability of being trimmed is controlled by the threshold τ_n: $P(T_i(\hat{\tau}) = 0) = P(\min_k p_k(W) < \hat{\tau}) = O_p(\hat{\tau}^\delta)$.

Substituting this into our bound gives: $|B_n| \leq \sqrt{n} \cdot \sqrt{V} \cdot \sqrt{O_p(\hat{\tau}^\delta)} = O_p(\sqrt{n} \cdot \hat{\tau}^{\delta/2})$.

The theorem's condition requires $n\tau_n^\delta \to 0$ for all fixed τ_n on the grid. Since Lemma 6.2 shows that the data-driven $\hat{\tau}$ converges in probability to a value on this grid (or zero), the condition implies that $n\hat{\tau}^\delta \to_p 0$. This is equivalent to $\sqrt{n}\hat{\tau}^{\delta/2} \to_p 0$. Therefore, the bias is asymptotically negligible: $B_n = o_p(1)$.

2. Asymptotic Normality. The preceding steps establish the linear representation of the estimator:$\sqrt{n}(\hat{\theta} - \theta_0) = \frac{1}{\sqrt{n}} \sum_{i=1}^n T_i(\hat{\tau})\varphi_i + o_p(1)$.

Let $S_n = \frac{1}{\sqrt{n}} \sum_{i=1}^n T_i(\hat{\tau})\varphi_i$ and let $X_{ni} = n^{-1/2} T_i(\hat{\tau})\varphi_i$. We verify the conditions for the Lyapunov Central Limit Theorem for the triangular array $\{X_{ni}\}$.

(a) Zero Mean (Asymptotically): The mean of each term in the sum is $E[X_{ni}] = n^{-1/2}E[T_i(\hat{\tau})\varphi_i]$. As shown above, $\sqrt{n}E[T_i(\hat{\tau})\varphi_i] = -B_n = o_p(1)$, so $E[X_{ni}] = o_p(n^{-1})$.

(b) Convergent Variance: The sum of the variances is: $\sum_{i=1}^n \mathrm{Var}(X_{ni}) = \sum_{i=1}^n \frac{1}{n}\mathrm{Var}(T_i(\hat{\tau})\varphi_i) = \mathrm{Var}(T_1(\hat{\tau})\varphi_1) = E[T_1(\hat{\tau})^2\varphi_1^2] - (E[T_1(\hat{\tau})\varphi_1])^2$. As $n \to \infty$, $\hat{\tau} \xrightarrow{p} 0$, which implies $T_1(\hat{\tau}) \xrightarrow{p} 1$. By the Dominated Convergence Theorem (since $T_1^2\varphi_1^2 \leq \varphi_1^2$ and $E[\varphi_1^2] < \infty$), the first term converges to $E[\varphi_1^2] = V$. The squared mean term converges to $(E[\varphi_1])^2 = 0$. Thus, $\sum_{i=1}^n \mathrm{Var}(X_{ni}) \to V$.

(c) Lyapunov Condition: We must show that for some $\epsilon > 0$, the sum of higher moments vanishes. Let ϵ be such that $E[|\varphi_i|^{2+\epsilon}] < \infty$ (a finite fourth moment as in Corollary 7.2 is sufficient, implying $\epsilon = 2$). $\sum_{i=1}^n E[|X_{ni}|^{2+\epsilon}] = \sum_{i=1}^n \frac{1}{n^{(2+\epsilon)/2}} E[|T_i(\hat{\tau})\varphi_i|^{2+\epsilon}]$ $= n \cdot \frac{1}{n^{1+\epsilon/2}} E[|T_i(\hat{\tau})\varphi_i|^{2+\epsilon}] = n^{-\epsilon/2} E[T_i(\hat{\tau})^{2+\epsilon}|\varphi_i|^{2+\epsilon}] \leq n^{-\epsilon/2} E[|\varphi_i|^{2+\epsilon}]$. Since $E[|\varphi_i|^{2+\epsilon}]$ is a finite constant, the expression is $O(n^{-\epsilon/2})$, which converges to 0 as $n \to \infty$.

With all conditions satisfied, the Lyapunov CLT implies that $S_n \xrightarrow{d} N(0, V)$. We now apply Slutsky's Theorem, which states that if $X_n \xrightarrow{d} X$ and $Y_n \xrightarrow{p} c$, then $X_n + Y_n \xrightarrow{d} X + c$.

In our case, $X_n = S_n$ and $Y_n = \sqrt{n}(\hat{\theta} - \theta_0) - S_n = R_n - B_n = o_p(1)$. Thus, $Y_n \xrightarrow{p} 0$. Slutsky's Theorem gives: $\sqrt{n}(\hat{\theta} - \theta_0) = S_n + o_p(1) \xrightarrow{d} N(0, V) + 0 = N(0, V)$.

3. Consistency of the Variance Estimator. Let $\bar{\varphi}_T = (\sum_i T_i(\hat{\tau}))^{-1} \sum_i T_i(\hat{\tau})\hat{\varphi}_i$. By Lemma A.3 and Assumption 6.1, $\bar{\varphi}_T = o_p(1)$ and $\frac{1}{n}\sum_i T_i(\hat{\tau})\{\hat{\varphi}_i - \bar{\varphi}_T\}^2 \xrightarrow{p} E[\varphi(Z)^2]$. Because $\left(\frac{1}{n}\sum_i T_i(\hat{\tau})\right)^{-2} \to_p 1$, the centered plug-in estimator stated in Theorem 6.2 is consistent for V. In the non-vanishing case $\hat{\tau} \to_p \bar{\tau} > 0$, the same argument yields consistency for $V_{\text{trim}}(\bar{\tau})$ (see the remark following Theorem 6.2).

Thus, we only need to show the consistency of the numerator: $\frac{1}{n}\sum_i \hat{T}_i \psi_i(\hat{\theta}, \hat{\eta})^2 \xrightarrow{p} V$. We decompose the error using the triangle inequality: $\left|\frac{1}{n}\sum_i \hat{T}_i\hat{\psi}_i^2 - V\right| \leq \underbrace{\left|\frac{1}{n}\sum_i \hat{T}_i\hat{\psi}_i^2 - \frac{1}{n}\sum_i T_i\varphi_i^2\right|}_{\text{Term A}}$

$+ \underbrace{\left|\frac{1}{n}\sum_i T_i\varphi_i^2 - V\right|}_{\text{Term B}}$.

- Term B converges to 0: By the Law of Large Numbers, since $T_i(\hat{\tau})\varphi_i^2$ are i.i.d. and $\hat{\tau} \xrightarrow{p} 0$: $\frac{1}{n}\sum_i T_i(\hat{\tau})\varphi_i^2 \xrightarrow{p} E[T_i(0)\varphi_i^2] = E[\varphi_i^2] = V$.

- Term A converges to 0: We can write $|\hat{T}_i\hat{\psi}_i^2 - T_i\varphi_i^2| \leq |(\hat{T}_i - T_i)\hat{\psi}_i^2| + |T_i(\hat{\psi}_i^2 - \varphi_i^2)|$. The first part vanishes because $P(\hat{T}_i \neq T_i) \to 0$. For the second part, note that $\hat{\psi}_i^2 - \varphi_i^2 = (\hat{\psi}_i - \varphi_i)(\hat{\psi}_i + \varphi_i)$. Since $\hat{\theta} \xrightarrow{p} \theta_0$ and $\|\hat{\eta} - \eta_0\| \xrightarrow{p} 0$, we have $\hat{\psi}_i - \varphi_i \xrightarrow{p} 0$. By the Continuous Mapping Theorem and dominated convergence (as $\hat{\psi}_i^2$ is bounded in probability), we have $\frac{1}{n}\sum_i T_i(\hat{\tau})(\hat{\psi}_i^2 - \varphi_i^2) \xrightarrow{p} 0$.

Since both Term A and Term B converge to zero in probability, their sum does as well. This establishes that the numerator converges to V, and therefore $\hat{V} \xrightarrow{p} V$. □

A.8 Proof of Lemma 6.1

Let $A_\tau = \{T_i(\tau) = 1\}$ and let $H(Z)$ denote the target signal (e.g., $\tau_k(W)$ for discrete or $\partial_d m(D, W)$ for continuous). By definition, $\theta(\tau) = E[H|A_\tau] = E[H\mathbf{1}_{A_\tau}]/P(A_\tau)$. The bias is: $\theta(\tau) - \theta = \frac{E[H\mathbf{1}_{A_\tau}] - E[H]P(A_\tau)}{P(A_\tau)} = \frac{E[H(\mathbf{1}_{A_\tau} - 1)] + E[H](1 - P(A_\tau))}{P(A_\tau)}$. Noting that $\mathbf{1}_{A_\tau} - 1 = -\mathbf{1}_{A_\tau^c}$, and $1 - P(A_\tau) = P(A_\tau^c) = E[\mathbf{1}_{A_\tau^c}]$, we rearrange terms: $\theta(\tau) - \theta = -\frac{E[(H - E[H])\mathbf{1}_{A_\tau^c}]}{P(A_\tau)}$.

Case (i): Boundedness. If $|H| \leq C$ almost surely, then $|H - E[H]| \leq 2C$. Thus: $|\theta(\tau) - \theta| \leq \frac{2CE[\mathbf{1}_{A_\tau^c}]}{P(A_\tau)} = \frac{2CP(A_\tau^c)}{1-P(A_\tau^c)}$. By Assumption 2.3, $P(A_\tau^c) = O(\tau^\delta)$. Since $\tau \to 0$, the denominator goes to 1, yielding $O(\tau^\delta)$.

Case (ii): Moment Bounds. If $E[|H|^q] \leq C$, applying Hölder's inequality with exponents q and p (where $1/q + 1/p = 1$) to the numerator $E[|H - E[H]| \cdot \mathbf{1}_{A_\tau^c}]$ yields: $|E[\dots]| \leq \|H - E[H]\|_q \|\mathbf{1}_{A_\tau^c}\|_p = C' \cdot P(A_\tau^c)^{1/p} = C' \cdot P(A_\tau^c)^{1-1/q}$. Substituting the overlap rate $O(\tau^\delta)$, the bias is $O(\tau^{\delta(1-1/q)})$. □

A.9 Proof of Corollary 6.3

The validity of the joint Central Limit Theorem for K_n arms relies on the Lyapunov condition. We must control the third absolute moment of the influence function relative to its variance. In a multi-arm setting where $K_n \to \infty$, the propensity scores $p_k(W)$ mechanically decay, which affects the moments of the inverse-propensity weights.

Recall the discrete influence function component for arm k: $\varphi_{ik} = \frac{D_{ik}}{p_k(W_i)}(Y_i - m_k(W_i)) + (m_k(W_i) - m_0(W_i)) - \theta_k$. Under the assumption of symmetric overlap where $p_k(W) \asymp K_n^{-1}$, the variance and third moments scale as follows:

$$\text{Variance:} \quad \sigma_n^2 \asymp \sum_{k=1}^{K_n} E\left[\left(\frac{D_{ik}}{p_k}\right)^2\right] \asymp \sum_{k=1}^{K_n} K_n = K_n^2. \tag{19}$$

$$\text{Third Moment:} \quad \rho_n^3 \asymp \sum_{k=1}^{K_n} E\left[\left|\frac{D_{ik}}{p_k}\right|^3\right] \asymp \sum_{k=1}^{K_n} K_n^2 = K_n^3. \tag{20}$$

Unlike fixed-K settings, the fourth moments are not uniformly bounded; they scale with the inverse propensity $p_k^{-3} \asymp K_n^3$.

The Lyapunov condition requires: $\frac{\sum_{i=1}^n \rho_{ni}^3}{(\sum_{i=1}^n \sigma_{ni}^2)^{3/2}} \to 0$. Substituting the scaling derived above: $\frac{n \cdot K_n^3}{(n \cdot K_n^2)^{3/2}} = \frac{nK_n^3}{n^{3/2}K_n^3} = \frac{1}{\sqrt{n}}$. While this ratio vanishes, strict consistency requires controlling the variance inflation of the trimmed estimator and ensuring uniform integrability of the higher moments. The binding constraint arises from the remainder terms in the von Mises expansion, which scale with the number of arms. To ensure the inverse-propensity weighted moments do not explode faster than the variance stabilizes, we require the fourth moment condition derived in Assumption 2.7: $\frac{K_n^4}{n} \to 0 \implies K_n = o(n^{1/4})$. This growth rate

is sufficient to ensure that the tails of the score distribution, driven by $p_k(W) \to 0$, do not invalidate the Gaussian approximation. □

A.10 Proof of Proposition 4.1

Step 1. Notation and the interacted regression. Write the saturated linear specification as $Y_i = \alpha_0 + \beta_0^\top W_i + \sum_{k=1}^K D_{ik}\left(\alpha_k + \beta_k^\top W_i\right) + \varepsilon_i, E[\varepsilon_i \mid D_i, W_i] = 0$.

Let $r_{i\ell} = Y_i - \hat{m}_\ell(W_i) = Y_i - \hat{\alpha}_\ell - \hat{\beta}_\ell^\top W_i$ be the OLS residuals from the fully–interacted fit. Standard OLS normal equations give, *for each* $\ell = 0, \ldots, K$,

$$\sum_{i=1}^n \mathbf{1}\{D_i = \ell\}\, r_{i\ell} = 0, \qquad \sum_{i=1}^n \mathbf{1}\{D_i = \ell\}\, r_{i\ell} W_i = 0. \tag{A.1}$$

Step 2. Plug–in efficient score. With the OLS nuisance estimates and *any* probability–limit–stable $\hat{p}_\ell(W)$, the sample efficient score for arm k is $\hat{\psi}_{ik} = \hat{m}_k(W_i) - \hat{m}_0(W_i) + \frac{D_{ik}}{\hat{p}_k(W_i)}\, r_{ik} - \frac{D_{i0}}{\hat{p}_0(W_i)}\, r_{i0} - \theta_k$. Set $S_{nk} = n^{-1}\sum_{i=1}^n \hat{\psi}_{ik}$.

Step 3. Sample moment equals the interacted-OLS moment. Break S_{nk} into three pieces: $S_{nk} = \underbrace{n^{-1}\sum_i \big[\hat{m}_k(W_i) - \hat{m}_0(W_i)\big]}_{\text{(i)}} + \underbrace{n^{-1}\sum_i \Big[\frac{D_{ik}}{\hat{p}_k(W_i)}\, r_{ik} - \frac{D_{i0}}{\hat{p}_0(W_i)}\, r_{i0}\Big]}_{\text{(ii)}} - \theta_k$.

Piece (ii). Fix $\ell \in \{0, k\}$, multiply and divide by the true (unknown) $p_\ell(W_i)$, and condition on W_i: $E\left[\frac{D_{i\ell}}{\hat{p}_\ell(W_i)}\, r_{i\ell} \middle| W_i\right] = \frac{p_\ell(W_i)}{\hat{p}_\ell(W_i)} E\left[\mathbf{1}\{D_i = \ell\}\, r_{i\ell} \,\middle|\, W_i\right] = 0$, because the inner expectation is exactly the population analogue of (A.1). Hence the law of large numbers gives $n^{-1}\sum_i D_{i\ell}\hat{p}_\ell^{-1} r_{i\ell} = o_p(1)$ for $\ell = 0, k$, so piece (ii) is $o_p(1)$.

Piece (i). By the definition of the IA estimator, $\hat{\theta}_k^{\mathrm{IA}} = n^{-1}\sum_i \big[\hat{m}_k(W_i) - \hat{m}_0(W_i)\big]$.

Therefore $S_{nk} = \hat{\theta}_k^{\mathrm{IA}} - \theta_k + o_p(1)$. Imposing the sample moment condition $S_{nk} = 0$ yields $\hat{\theta}_k^{\mathrm{IA}} = \theta_k + o_p(1)$, *exactly the solution to the OLS normal equations.* Thus the plug–in DML estimator with the saturated linear outcome model returns the interacted-ATE estimator (Goldsmith-Pinkham *et al.*, 2024). Since (6) is Neyman-orthogonal, $\hat{\theta}_k$ inherits all DML asymptotics. □

A.11 Proof of Proposition 4.2

Step 1. CW-weighted propensities. Let $\pi_\ell = P(D = \ell)$ and write $g(W) = \Big[\sum_{\ell=0}^{K} \pi_\ell(1 - \pi_\ell)/p_\ell(W)\Big]^{-1}$. Define the *CW-weighted indicators* $\tilde{D}_{i\ell} = g(W_i)\, D_{i\ell}/p_\ell(W_i)$. By construction they satisfy $E[\tilde{D}_{i\ell} \mid W] = g(W)$ and hence $E[\tilde{D}_{i\ell}] = 1$.

Step 2. GPHK-CW estimating equation. (Goldsmith-Pinkham *et al.*, 2024) estimate θ_k as the coefficient from the weighted regression of Y on D_{ik} using weights $g(W_i)$. The corresponding sample normal equation is

$$0 \;=\; \sum_{i=1}^{n} g(W_i)(D_{ik} - \pi_k)\big(Y_i - \hat{\theta}_k^{\mathrm{CW}}\big). \tag{A.3}$$

Step 3. Plug–in score equals the CW moment. Insert the CW propensities $\hat{p}_\ell(W) = D_{i\ell}/\tilde{D}_{i\ell}$ into the efficient score: $\hat{\psi}_{ik}^{\mathrm{CW}} = \big[\hat{m}_k(W_i) - \hat{m}_0(W_i)\big] + \tilde{D}_{ik}\{Y_i - \hat{m}_k(W_i)\} - \tilde{D}_{i0}\{Y_i - \hat{m}_0(W_i)\} - \theta_k$.

Average and rearrange, $0 \overset{!}{=} \sum_i \hat{\psi}_{ik}^{\mathrm{CW}}$

$= \sum_i \tilde{D}_{ik} Y_i \;-\; \sum_i \tilde{D}_{i0} Y_i \;-\; n\theta_k$. Use (A.2) with $\ell = 0, k$ to replace $\sum_i \tilde{D}_{i0} = \sum_i \tilde{D}_{ik} = n$, giving $\sum_i g(W_i)(D_{ik} - \pi_k) Y_i = n\theta_k$. Subtracting $\theta_k \sum_i g(W_i)(D_{ik} - \pi_k) = 0$ from both sides we obtain the GPHK-CW normal equation (A.3). Hence the plug-in solution $\hat{\theta}_k^{\mathrm{CW}}$ is *identical* to the common-weight estimator, and the score expectation is zero at the truth. Orthogonality follows because (6) is orthogonal for any propensities. □

A.12 Rate Conditions for Common Learners

Random forests with honesty, boosted trees of depth $O(\log n)$, gradient-boosted splines, and ReLU networks of depth $\leq C \log n$ achieve the $n^{-1/4}$ L_2 rate under standard sparsity or smoothness conditions; see Athey and Wager (2018), Farrell *et al.* (2021).

A.13 Proof of Lemma 6.2

Proof. The proof proceeds in two steps. First, we establish the uniform consistency of the MSE estimator. Second, we show that this implies the consistency of the minimizer $\hat{\tau}$.

Step 1. Uniform Consistency of $\widehat{MSE}$. Recall that $\widehat{MSE}(\tau) = \widehat{V}(\tau) + \widehat{B}(\tau)^2$. We show uniform convergence for each component over the finite grid $\mathcal{G}_n$.

1. Variance Component: The function class $\mathcal{F}_V = \{T(\tau)\varphi(Z)^2 : \tau \in \mathcal{G}_n\}$ is finite. Since $E[\varphi(Z)^2] < \infty$ and $T(\tau) \leq 1$, the class has an integrable envelope. By the finite-class LLN (Van der Vaart and Wellner, 1996),

$$\sup_{\tau \in \mathcal{G}_n} |\widehat{V}(\tau) - V(\tau)| = o_p(1).$$

2. Bias Component: Similarly, for $\mathcal{F}_B = \{T(\tau)\varphi(Z) : \tau \in \mathcal{G}_n\}$, we have

$$\sup_{\tau \in \mathcal{G}_n} |\widehat{B}(\tau) - B(\tau)| = o_p(1).$$

Combining these, and noting that $\sup_\tau |B(\tau)| < \infty$:

$$\sup_{\tau \in \mathcal{G}_n} |\widehat{MSE}(\tau) - MSE(\tau)| \leq o_p(1) + 2o_p(1)O_p(1) = o_p(1). \tag{A.1}$$

Step 2. Consistency of the Minimizer. Let $\tau^\star = \arg\min_{\tau \in \mathcal{G}_n} MSE(\tau)$. Since the grid is finite, define the separation gap $\Delta = \min_{\tau \neq \tau^\star}\{MSE(\tau) - MSE(\tau^\star)\} > 0$. If $\hat{\tau} \neq \tau^\star$, then $\widehat{MSE}(\hat{\tau}) \leq \widehat{MSE}(\tau^\star)$. This implies:

$$\Delta \leq MSE(\hat{\tau}) - MSE(\tau^\star) \leq 2 \sup_{\tau \in \mathcal{G}_n} |\widehat{MSE}(\tau) - MSE(\tau)|.$$

Thus, $P(\hat{\tau} \neq \tau^\star) \leq P(2 \sup |\dots| \geq \Delta) \to 0$ by (A.1). □

A.14 Proof of Theorem 6.4

Let clusters be indexed by $g = 1, \dots, G$, with (random) cluster size n_g and observations $Z_{gi} = (Y_{gi}, D_{gi}, W_{gi})$, $i = 1, \dots, n_g$. Write $N = \sum_{g=1}^G n_g$ and denote by $\hat{\eta}^{(-j)}$ the nuisance estimates trained on clusters not in fold j (cluster-level cross-fitting). For a deterministic threshold $\tau > 0$ define $T_{gi}(\tau) = \mathbf{1}\{\min_k p_k(W_{gi}) \geq \tau\}$ (or $\mathbf{1}\{f(D_{gi} \mid W_{gi}) \geq \tau\}$ in the continuous case). Let the *raw* moment be $\psi(Z; \eta)$ (no $-\theta$) and $\varphi(Z) = \psi(Z; \eta_0) - \theta$

the population influence function at the truth η_0, as in (6)–(5). Define cluster totals $S_g(\tau) = \sum_{i=1}^{n_g} T_{gi}(\tau)\{\psi(Z_{gi};\eta_0) - \theta(\tau)\}, C_g(\tau) = \sum_{i=1}^{n_g} T_{gi}(\tau)$, and their expectations $\mu_C(\tau) = E[C_g(\tau)]$ and $\sigma_S^2(\tau) = E[S_g(\tau)^2]$.[1]

Step 1. The estimator computed by Algorithm 1 can be written as the ratio of cluster sums of trimmed raw moments (recall we average the raw moment to form $\hat{\theta}$): $\hat{\theta}(\tau) = \frac{\sum_{g=1}^{G}\sum_{i=1}^{n_g} T_{gi}(\tau)\,\hat{\psi}_{gi}}{\sum_{g=1}^{G}\sum_{i=1}^{n_g} T_{gi}(\tau)}, \hat{\psi}_{gi} = \psi\big(Z_{gi};\hat{\eta}^{(-j(g))}\big).Addandsubtract\theta(\tau)$ in the numerator, and use the identity $\hat{\psi}_{gi} - \theta(\tau) = \{\psi(Z_{gi};\eta_0) - \theta(\tau)\} + \{\hat{\psi}_{gi} - \psi(Z_{gi};\eta_0)\}$ to obtain

$$\begin{aligned}\hat{\theta}(\tau) - \theta(\tau) &= \frac{\sum_{g=1}^{G} S_g(\tau)}{\sum_{g=1}^{G} C_g(\tau)} \;+\; \frac{\sum_{g=1}^{G}\sum_{i=1}^{n_g} T_{gi}(\tau)\,\{\hat{\psi}_{gi} - \psi(Z_{gi};\eta_0)\}}{\sum_{g=1}^{G} C_g(\tau)} \\ &=: \frac{\sum_{g=1}^{G} S_g(\tau)}{\sum_{g=1}^{G} C_g(\tau)} \;+\; R_G(\tau). \end{aligned} \tag{21}$$

Under Assumptions 2.6 and 6.1 (orthogonality, $N^{-1/4}$ L_2 rates on the trimmed support, cluster-level cross-fitting, and $E[n_g] \in (0,\infty)$), the *nuisance remainder* satisfies $\sqrt{G}\,R_G(\tau) = \frac{1}{\sqrt{G}}\sum_{g=1}^{G}\sum_{i=1}^{n_g} T_{gi}(\tau)\{\hat{\psi}_{gi} - \psi(Z_{gi};\eta_0)\} \,/\, \Big(\frac{1}{G}\sum_{g=1}^{G} C_g(\tau)\Big) = o_p(1)$. The proof mirrors Lemma A.4: expand the difference by a first-order Gateaux derivative, note that the expectations of the linear terms vanish by Neyman orthogonality (Lemmas 3.1–3.2), and bound the quadratic terms by $\sqrt{G} \times o_p(N^{-1/2}) = o_p(1)$ because $N/G \to_p E[n_g] \in (0,\infty)$.

Step 2. By a cluster LLN and $\mu_C(\tau) = E[C_g(\tau)] > 0$, $\frac{1}{G}\sum_{g=1}^{G} C_g(\tau) \xrightarrow{p} \mu_C(\tau), \sqrt{G}\Big\{\frac{1}{G}\sum_{g=1}^{G} C_g(\tau) - \mu_C(\tau)\Big\} = O_p(1)$. Substituting this into (21) and multiplying by $\sqrt{G}$ yields

$$\sqrt{G}\,\{\hat{\theta}(\tau) - \theta(\tau)\} = \frac{1}{\mu_C(\tau)} \cdot \frac{1}{\sqrt{G}} \sum_{g=1}^{G} S_g(\tau) \;+\; o_p(1), \tag{22}$$

because $E[S_g(\tau)] = 0$ makes the usual ratio correction term vanish.

Step 3. Clusters are independent by Assumption 6.1. Moreover, $E[S_g(\tau)] = 0$ and, by Lemma A.2 combined with $E[n_g^{2+\kappa}] < \infty$, we have $E[|S_g(\tau)|^{2+\kappa}] \le C < \infty$ for some $\kappa > 0$. Hence the Lindeberg–Feller CLT for independent, not-necessarily identically distributed arrays applies: $\frac{1}{\sqrt{G}}\sum_{g=1}^{G} S_g(\tau) \xrightarrow{d} N\big(0,\ \sigma_S^2(\tau)\big), \sigma_S^2(\tau) = E\big[S_g(\tau)^2\big]$. Combining with (22) gives $\sqrt{G}\,\{\hat{\theta}(\tau) - \theta(\tau)\} \xrightarrow{d} N\Big(0,\ V_{\mathrm{cl}}(\tau)\Big), V_{\mathrm{cl}}(\tau) = \frac{\sigma_S^2(\tau)}{\mu_C(\tau)^2}$.

[1] By definition of $\theta(\tau) = E[\psi(Z;\eta_0) \mid T(\tau) = 1]$, we have $E[S_g(\tau)] = 0$.

Step 4. Let $\mathcal{G}_n$ be the finite trimming grid in Assumption 6.1, and let $\hat{\tau}$ minimize the uniform MSE proxy in Section 6.3. The cluster versions of Lemmas A.3–A.4 (finite-class LLN for $\{S_g(\tau), C_g(\tau) : \tau \in \mathcal{G}_n\}$ and the same orthogonality argument) imply $\hat{\tau} \xrightarrow{p} \tau^\star \in \mathcal{G}_\infty$, and $\sup_{\tau\in\mathcal{G}_n}\left|\sqrt{G}\left\{\hat{\theta}(\tau) - \theta(\tau)\right\} - \frac{1}{\mu_C(\tau)} \cdot \frac{1}{\sqrt{G}} \sum_{g=1}^{G} S_g(\tau)\right| = o_p(1)$. Therefore, $\sqrt{G}\left\{\hat{\theta}(\hat{\tau}) - \theta(\tau^\star)\right\} \xrightarrow{d} N\Big(0,\ V_{\mathrm{cl}}(\tau^\star)\Big)$. If the grid vanishes (Condition 6.1) so that $\tau^\star \to 0$ and the trimming bias is $o_p(G^{-1/2})$, then $\theta(\tau^\star) \to \theta$ and the limit variance equals $V_{\mathrm{cl}}(0)$.

Step 5. Define the cluster sums of estimated centered contributions $\hat{S}_g = \sum_{i=1}^{n_g} T_{gi}(\hat{\tau})\left\{\hat{\psi}_{gi} - \hat{\theta}\right\}, \hat{\mu}_C = \frac{1}{G}\sum_{g=1}^{G} C_g(\hat{\tau})$. The proposed estimator $\hat{V}_{\mathrm{cl}} = \frac{G}{\left(\sum_g C_g(\hat{\tau})\right)^2} \sum_{g=1}^{G} \hat{S}_g^2 = \frac{1}{\hat{\mu}_C^2} \cdot \frac{1}{G} \sum_{g=1}^{G} \hat{S}_g^2$ converges in probability to $V_{\mathrm{cl}}(\tau^\star)$ because (i) $\hat{\mu}_C \to_p \mu_C(\tau^\star)$ by a cluster LLN, (ii) $\frac{1}{G}\sum_g \hat{S}_g^2 \to_p E[S_g(\tau^\star)^2]$ (the difference $\hat{S}_g - S_g(\tau^\star)$ is $o_p(1)$ uniformly over g by the same orthogonality and rate arguments), and (iii) $\mathcal{G}_n$ is finite. A small-sample degrees-of-freedom factor $G/(G-1)$ may be multiplied if desired.

Putting Steps 1–5 together proves the stated clustered CLT and the consistency of the cluster-robust variance estimator under both deterministic and data-driven trimming. The argument holds verbatim for the continuous-dose score (5) and the discrete multi-arm score (6).

B Additional Simulation Results

This appendix provides the full set of Monte-Carlo results, expanding on the headline summary in the main text. Table B.5 reports the performance under linear potential outcomes (Design A, Linear), and Table B.6 reports the performance under non-linear potential outcomes (Design A, Non-Linear).

Table B.5: Finite–sample *linear* performance (Design A)

		(a) Strong overlap ($\gamma = 0.6$)							(b) Weak overlap ($\gamma = 1.8$)				
		Estimator							Estimator				
n	Metric	GOATE-DML	EW	IA	CW	OLS	n	Metric	GOATE-DML	EW	IA	CW	OLS
Panel A: Arm 1 ($\tau_1 = 1.0$)													
	Bias	0.034	0.007	0.004	0.010	0.010		Bias	0.270	0.004	0.004	0.003	0.003
1,000	RMSE	0.092	0.079	0.076	0.079	0.079	1,000	RMSE	0.311	0.185	0.136	0.098	0.098
	s.e.	0.012	0.011	0.011	0.011	0.011		s.e.	0.022	0.026	0.019	0.014	0.014
	Bias	0.018	−0.001	0.001	0.003	0.003		Bias	0.190	0.016	0.019	0.020	0.020
2,000	RMSE	0.075	0.063	0.063	0.060	0.060	2,000	RMSE	0.216	0.095	0.090	0.071	0.071
	s.e.	0.010	0.009	0.009	0.008	0.008		s.e.	0.014	0.013	0.012	0.010	0.010
	Bias	0.001	−0.001	−0.003	−0.003	−0.003		Bias	0.087	−0.014	−0.010	−0.011	−0.011
4,000	RMSE	0.044	0.040	0.038	0.037	0.037	4,000	RMSE	0.115	0.071	0.061	0.046	0.046
	s.e.	0.006	0.006	0.005	0.005	0.005		s.e.	0.011	0.010	0.009	0.006	0.006
Panel B: Arm 2 ($\tau_2 = 2.0$)													
	Bias	−0.005	0.007	0.004	0.000	0.000		Bias	−0.099	−0.036	−0.010	−0.007	−0.007
1,000	RMSE	0.113	0.096	0.093	0.090	0.090	1,000	RMSE	0.181	0.138	0.111	0.093	0.093
	s.e.	0.016	0.014	0.013	0.013	0.013		s.e.	0.021	0.019	0.016	0.013	0.013
	Bias	0.009	0.012	0.013	0.014	0.014		Bias	−0.042	0.013	0.025	0.017	0.017
2,000	RMSE	0.084	0.080	0.074	0.070	0.070	2,000	RMSE	0.118	0.089	0.090	0.071	0.071
	s.e.	0.012	0.011	0.010	0.010	0.010		s.e.	0.016	0.013	0.012	0.010	0.010
	Bias	−0.007	−0.007	−0.009	−0.006	−0.006		Bias	−0.054	−0.013	−0.006	−0.004	−0.004
4,000	RMSE	0.048	0.040	0.038	0.037	0.037	4,000	RMSE	0.102	0.081	0.062	0.052	0.052
	s.e.	0.007	0.006	0.005	0.005	0.005		s.e.	0.012	0.011	0.009	0.007	0.007
Panel C: Arm 3 ($\tau_3 = 3.0$)													
	Bias	0.035	0.010	0.010	0.011	0.011		Bias	0.128	0.004	0.021	0.026	0.026
1,000	RMSE	0.110	0.100	0.093	0.089	0.089	1,000	RMSE	0.203	0.166	0.119	0.104	0.104
	s.e.	0.015	0.014	0.013	0.012	0.012		s.e.	0.022	0.024	0.016	0.014	0.014
	Bias	0.017	0.003	0.003	0.006	0.006		Bias	0.059	−0.012	−0.006	−0.003	−0.003
2,000	RMSE	0.089	0.067	0.064	0.062	0.062	2,000	RMSE	0.117	0.087	0.094	0.064	0.064
	s.e.	0.012	0.009	0.009	0.009	0.009		s.e.	0.014	0.012	0.013	0.009	0.009
	Bias	0.000	−0.005	−0.006	−0.005	−0.005		Bias	0.021	−0.025	−0.028	−0.020	−0.020
4,000	RMSE	0.053	0.043	0.043	0.041	0.041	4,000	RMSE	0.079	0.074	0.066	0.050	0.050
	s.e.	0.008	0.006	0.006	0.006	0.006		s.e.	0.011	0.010	0.008	0.006	0.006

Notes: Detailed performance for Linear Data Generating Process. Bias is mean deviation from true ATE.

Table B.6: Finite–sample *non-linear* performance (Design A)

		(a) Strong overlap ($\gamma = 0.6$)							(b) Weak overlap ($\gamma = 1.8$)				
		Estimator							Estimator				
n	Metric	GOATE-DML	EW	IA	CW	OLS	n	Metric	GOATE-DML	EW	IA	CW	OLS
Panel A: Arm 1 ($\tau_1 = 1.0$)													
	Bias	0.005	0.005	−0.006	0.005	0.005		Bias	0.071	0.028	−0.001	−0.011	−0.011
1,000	RMSE	0.097	0.113	0.098	0.097	0.097	1,000	RMSE	0.155	0.151	0.128	0.117	0.117
	s.e.	0.014	0.016	0.014	0.014	0.014		s.e.	0.019	0.021	0.018	0.016	0.016
	Bias	0.012	−0.002	−0.012	−0.001	−0.001		Bias	0.027	−0.011	−0.031	−0.043	−0.043
2,000	RMSE	0.078	0.083	0.080	0.084	0.084	2,000	RMSE	0.100	0.094	0.090	0.084	0.084
	s.e.	0.011	0.012	0.011	0.012	0.012		s.e.	0.014	0.013	0.012	0.010	0.010
	Bias	−0.002	0.003	−0.009	0.000	0.000		Bias	0.002	−0.015	−0.036	−0.044	−0.044
4,000	RMSE	0.061	0.061	0.060	0.056	0.056	4,000	RMSE	0.061	0.053	0.060	0.069	0.069
	s.e.	0.009	0.009	0.008	0.008	0.008		s.e.	0.009	0.007	0.007	0.007	0.007
Panel B: Arm 2 ($\tau_2 = 2.0$)													
	Bias	0.025	0.128	0.130	0.080	0.080		Bias	0.025	0.336	0.330	0.130	0.130
1,000	RMSE	0.113	0.178	0.170	0.137	0.137	1,000	RMSE	0.139	0.367	0.356	0.184	0.184
	s.e.	0.016	0.017	0.015	0.016	0.016		s.e.	0.019	0.021	0.019	0.018	0.018
	Bias	0.031	0.132	0.130	0.074	0.074		Bias	0.041	0.336	0.332	0.130	0.130
2,000	RMSE	0.086	0.154	0.151	0.107	0.107	2,000	RMSE	0.112	0.352	0.347	0.159	0.159
	s.e.	0.011	0.011	0.011	0.011	0.011		s.e.	0.015	0.015	0.014	0.013	0.013
	Bias	0.001	0.110	0.110	0.053	0.053		Bias	0.007	0.319	0.314	0.119	0.119
4,000	RMSE	0.041	0.125	0.123	0.078	0.078	4,000	RMSE	0.076	0.325	0.321	0.133	0.133
	s.e.	0.006	0.008	0.008	0.008	0.008		s.e.	0.011	0.009	0.009	0.008	0.008
Panel C: Arm 3 ($\tau_3 = 3.0$)													
	Bias	0.048	0.040	0.028	−0.019	−0.019		Bias	0.011	0.078	0.058	−0.139	−0.139
1,000	RMSE	0.133	0.128	0.101	0.100	0.100	1,000	RMSE	0.156	0.173	0.144	0.196	0.196
	s.e.	0.018	0.017	0.014	0.014	0.014		s.e.	0.022	0.022	0.019	0.020	0.020
	Bias	0.024	0.020	0.022	−0.035	−0.035		Bias	0.046	0.041	0.035	−0.160	−0.160
2,000	RMSE	0.080	0.082	0.086	0.094	0.094	2,000	RMSE	0.123	0.100	0.095	0.178	0.178
	s.e.	0.011	0.011	0.012	0.012	0.012		s.e.	0.016	0.013	0.012	0.011	0.011
	Bias	0.005	0.009	0.010	−0.047	−0.047		Bias	0.055	0.048	0.042	−0.155	−0.155
4,000	RMSE	0.066	0.076	0.073	0.084	0.084	4,000	RMSE	0.090	0.083	0.075	0.165	0.165
	s.e.	0.009	0.011	0.010	0.010	0.010		s.e.	0.010	0.010	0.009	0.008	0.008

Notes: Detailed performance for Non-Linear Data Generating Process. Bias is mean deviation from true ATE.

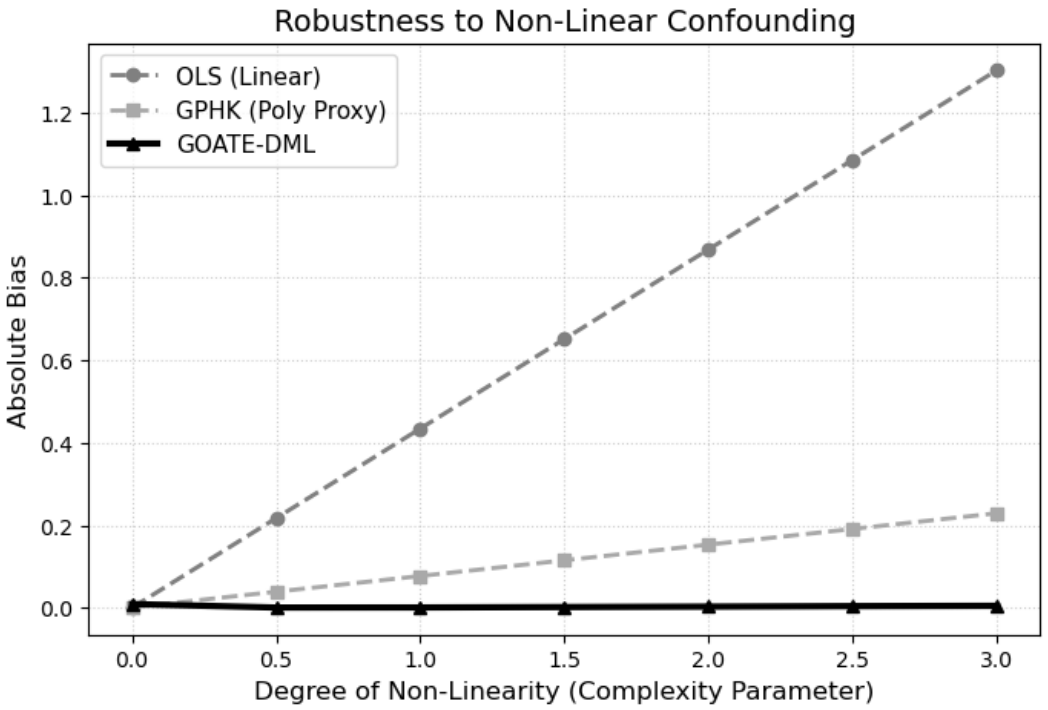


Figure B.3: Estimator Robustness to Non-Linearity. The horizontal axis represents the scaling factor of the non-linear confounding term. As complexity increases, OLS bias grows linearly and polynomial corrections exhibit increasing bias. The GOATE-DML estimator remains unbiased regardless of complexity due to the adaptive nature of the nuisance learners.

C Additional Empirical Applications

C.1 Drexler et al. (2014): The Honest Cost of Robustness

We analyze the experiment of Drexler et al. (2014) regarding financial literacy training. This application represents a "stress test" for weak overlap. As shown in Figure C.4, the finite-sample distribution of assignment probabilities for the "Rule-of-Thumb" arm exhibits a massive concentration near zero. Consequently, our adaptive procedure trims 6.1% of the sample. Both OLS and the linear GPHK correction estimate a null effect. GOATE-DML addresses this by estimating the effect non-parametrically only on the valid overlap sub-population. We find a point estimate of -575 pesos, which remains statistically indistinguishable from zero ($p = 0.55$). Crucially, the GOATE-DML standard error (958) is approximately 20% larger than OLS (801). This reflects the honest cost of robustness:

by trimming non-overlapping units, the estimator removes the "false precision" that OLS achieves by extrapolating linearly into regions of poor support.

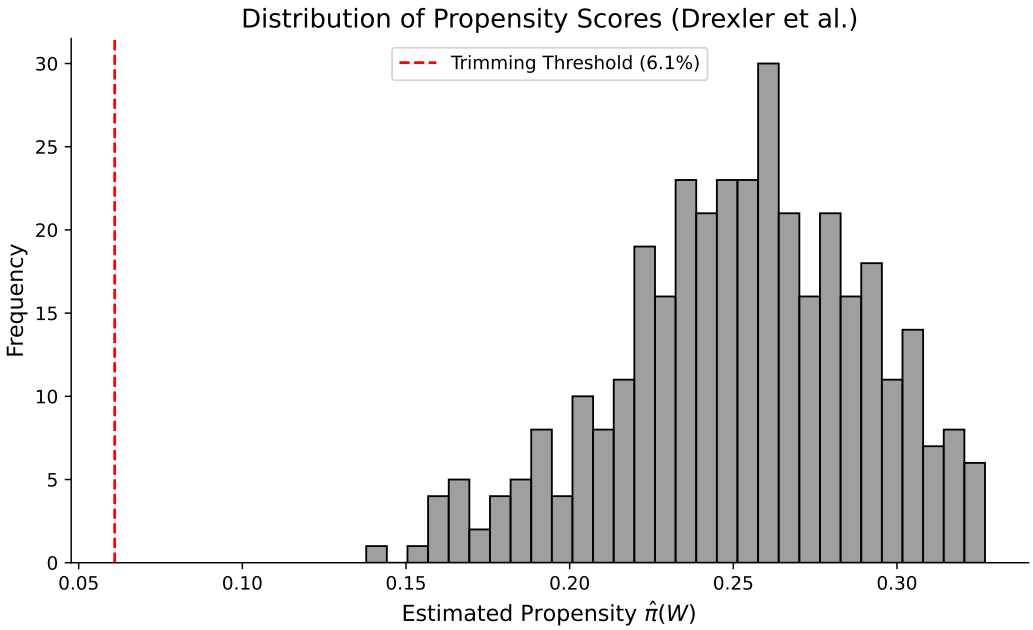


Figure C.4: Empirical Overlap in Drexler et al. (2014). The histogram displays the distribution of estimated propensity scores. The accumulation of mass near zero necessitates the GOATE trimming strategy.