FINITE-SAMPLE CERTIFICATION OF CAPACITY-CONSTRAINED TREATMENT ALLOCATIONS IN PRESPECIFIED COHORTS

Anonymous authors
Paper under double-blind review

ABSTRACT

Capacity-constrained treatment allocation uses experimental evidence to select a fixed number of individuals from a prespecified cohort for treatment. We study a complementary problem: certifying the welfare loss of a selected allocation relative to the best feasible alternative. Under a prespecified treatment effect model linking the randomized trial and target cohort, we characterize which allocation welfare differences are identified. These differences can remain identifiable even when some individual target treatment effects are not. We then construct simultaneous finite-sample upper bounds directly for the welfare differences that determine regret. The resulting regret certificates remain valid when the same trial is used for both allocation selection and certification. The construction applies to a prespecified candidate set and extends to all exact-k allocations through exchange bounds and maximum-weight partial matching. We further bound the excess of the certificate over true cohort-specific regret and show that, under a common confidence region, direct comparison gives no larger uncertainty radius than propagating separate treatment effect or allocation welfare intervals. Experimental results confirm post-selection coverage and show smaller mean certificates from direct comparison in the evaluated three-feature settings.

1 INTRODUCTION

Treatment allocation under limited capacity is a common causal decision problem. A health system may be able to provide intensive follow-up to only a fixed number of eligible patients, and a public program may have funding for only part of an eligible population. Statistical treatment rules and individualized treatment policies formalize how heterogeneous treatment effects can guide treatment decisions (Manski, 2004; Qian & Murphy, 2011; Kitagawa & Tetenov, 2018; Athey & Wager, 2021). Resource-constrained formulations incorporate treatment-capacity or budget constraints (Bhattacharya & Dupas, 2012; Luedtke & van der Laan, 2016; Pellatt, 2022; Kraus et al., 2024). In many deployments, the individuals are already known when the decision is made. For a prespecified cohort with m members and exactly k available treatment slots, the decision is to choose an exact-k treatment allocation for that realized cohort.

Existing work provides several ways to construct such an allocation. Resource-constrained policy learning maximizes expected welfare under treatment or cost limits (Bhattacharya & Dupas, 2012; Luedtke & van der Laan, 2016; Pellatt, 2022), ranking methods learn treatment priorities (Kamran et al., 2024), and recent work characterizes the sample complexity of selecting a near-optimal fixed-size allocation when treatment effects must be estimated (Casacuberta & Hardt, 2026). A separate literature studies how randomized trial evidence can be generalized to target populations whose covariate distributions differ from those in the trial (Cole & Stuart, 2010; Dahabreh et al., 2019; Degtiar & Rose, 2023; Robertson et al., 2023), including policy learning under sample-selection bias (Hatt et al., 2022). We study a different inferential problem: providing a finite-sample regret certificate for a selected allocation in an observed cohort using evidence from a randomized trial.

Certification in this setting has two requirements. First, the trial must identify the welfare differences between the target allocations being compared, even though those allocations are defined on covariates in the target cohort. Second, coverage must remain valid after data-dependent allocation

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

selection when the same trial outcomes are used for both selection and certification (Berk et al., 2013; Benjamini et al., 2019; Kuzborskij et al., 2021; Andrews & Chen, 2025; Chernozhukov et al., 2025). For a selected allocation, we define regret as its welfare gap to the best allocation in a declared comparison set, and a regret certificate as an observable finite-sample upper confidence bound on this loss. This leads to the central question of this paper: *Given a selected exact-k treatment allocation for a prespecified target cohort, how can randomized trial evidence provide a finite-sample upper bound on its regret that remains valid when the same trial is also used to select the allocation?*

We address this question by directly bounding welfare differences between the selected allocation and feasible alternatives. Under a prespecified treatment effect model linking the trial and target covariates, we characterize exactly when each required allocation welfare difference is identified from the randomized trial. We then construct simultaneous one-sided bounds that remain valid when allocation selection and certification use the same trial. Direct comparison removes cohort members whose treatment status is unchanged between two allocations before uncertainty is propagated. The certificate can also incorporate prespecified bounds on treatment effect model approximation error. For certification against all exact-k allocations, exchange bounds and maximum-weight partial matching avoid enumerating the full allocation space. We prove finite-sample validity and bound how much the certificate can exceed true cohort-specific regret.

Our **contributions** are threefold: **(i)** We formulate finite-sample certification for a selected exact-k treatment allocation in a prespecified cohort using randomized trial evidence, and characterize when the welfare comparisons required for certification are identified. **(ii)** We develop simultaneous finite-sample bounds that remain valid when allocation selection and certification use the same trial, and accommodate prespecified bounds on treatment effect model approximation error. **(iii)** We show that certification against all exact-k allocations can be computed through maximum-weight partial matching, bound how much the resulting certificate can exceed true cohort-specific regret, and evaluate the certificates using synthetic, semi-synthetic, and randomized real data.

2 Related Work

Treatment allocation under resource constraints. Statistical treatment rules and empirical welfare maximization choose treatment policies to maximize expected welfare (Manski, 2004; Qian & Murphy, 2011; Kitagawa & Tetenov, 2018; Athey & Wager, 2021). Resource-constrained formulations add treatment or cost limits (Bhattacharya & Dupas, 2012; Luedtke & van der Laan, 2016; Pellatt, 2022), and ranking methods learn treatment priorities when treatment order determines the allocation (Kamran et al., 2024). Casacuberta & Hardt (2026) study fixed-size allocation across groups using per-group effect estimates, with sample-complexity guarantees and an instance-dependent certification condition. We bound additive regret for an observed individual cohort using a separate randomized trial and an explicit trial-to-target effect model.

Generalizing randomized evidence to target populations. Generalizability and transportability methods characterize when randomized trial evidence identifies causal effects in a target population (Cole & Stuart, 2010; Dahabreh et al., 2019; Degtiar & Rose, 2023; Robertson et al., 2023). Related work studies treatment effect heterogeneity and policy learning under sample-selection bias (Hatt et al., 2022), as well as sensitivity to limited support and outcome distribution shift (Asiaee et al., 2026; Arbour et al., 2026). Our target is one realized cohort whose pretreatment covariates are observed. We impose a prespecified treatment effect model linking treatment effects at the trial and target covariates and study which allocation welfare differences are identified under that model. We allow prespecified deviations from the model through trial and target residuals.

Decision certification and inference after selection. Post-selection inference constructs simultaneous statements that remain valid for data-dependent reported quantities (Berk et al., 2013; Benjamini et al., 2019). Certified Decisions formalizes high-probability loss certificates for data-dependent actions (Andrews & Chen, 2025). Confidence-aware off-policy evaluation and policy choice similarly incorporate uncertainty when actions or policies are selected from data (Kuzborskij et al., 2021; Chernozhukov et al., 2025; Mate et al., 2023). Pairwise off-policy evaluation uses covariance between policy-value estimates to improve precision when estimating their difference (Jeunen & Ustimenko, 2024). We use these ideas for cohort-specific treatment allocations. Candidate-set certification controls all relevant pairwise allocation welfare differences simultaneously, and full al-

location certification combines simultaneous exchange bounds through matching. Appendix F gives a more detailed comparison of the estimands, information structures, and guarantees.

3 Problem Setup

We study finite-sample certification of capacity-constrained treatment allocations in prespecified cohorts. Inference is based on data from a randomized trial, while deployment concerns one realized cohort in which exactly k members must be selected for treatment.

Randomized trial and prespecified cohort. Outcomes are expressed on a known range normalized to $[0,1]$. The randomized trial is $D_n = \{(X_i, A_i, Y_i, e_i)\}_{i=1}^n$, where $A_i \in \{0,1\}$, $Y_i = A_i Y_i(1) + (1-A_i)Y_i(0)$, and e_i is the known treatment probability. Let $B = (x_1^B, \dots, x_m^B)$ be the target cohort. Independent training data D_{train} determine a feature map $\phi : \mathcal{X} \to \mathbb{R}^p$ and a candidate set

$$\mathcal{Z}_B = \{z_1, \dots, z_M\} \subset \{0,1\}^m, \qquad \mathbf{1}^\top z_r = k.$$

Here $M \geq 1$, $m \geq 0$, and $0 \leq k \leq m$. The equality constraint reflects the deployment requirement that all k treatment slots are filled. When allowing approximation error in the treatment effect model below, $\mathcal{U}$ denotes a prespecified class of trial and target residual pairs. Define

$$\mathcal{F}_0 = \sigma(B, D_{\text{train}}, X_{1:n}, e_{1:n}, \phi, \mathcal{Z}_B, \mathcal{U}, k).$$

Conditional on $\mathcal{F}_0$, trial units are independent, $A_i \sim \text{Bern}(e_i)$ is independent of the potential outcomes, and $0 < \eta \leq e_i \leq 1 - \eta$. We assume consistency, no interference, and

$$\mathbb{E}\{Y_i(a) \mid \mathcal{F}_0\} = \mu_a(X_i), \qquad a \in \{0,1\}. \tag{1}$$

The cohort, feature map, residual class, and candidate set are fixed before the trial assignments and outcomes used for certification are observed. The selected allocation may use those trial outcomes because the confidence event constructed below is simultaneous over the comparisons available to the selector.

Allocation welfare, pairwise differences, and regret. For trial units, let $\tau(X_i) = \mu_1(X_i) - \mu_0(X_i)$. For target members, $\tau(x_j^B)$ denotes the conditional average treatment effect at their covariates, and $\tau_B = (\tau(x_j^B))_{j=1}^m$. The treatment effect model below links these target effects to the trial. For a candidate allocation z_r, define

$$V_B(r) = z_r^\top \tau_B, \qquad \Delta_{rs} = V_B(r) - V_B(s), \qquad R_B(s) = \max_{1 \leq r \leq M} \Delta_{rs}. \tag{2}$$

Here $V_B(r)$ is total incremental welfare, Δ_{rs} is the pairwise allocation welfare difference, and $R_B(s)$ is regret relative to the best candidate allocation. Dividing welfare and certificate values by m gives the cohort-average convention when $m \geq 1$.

The full feasible allocation set is

$$\mathcal{A}_k = \{z \in \{0,1\}^m : \mathbf{1}^\top z = k\}, \qquad R_B^{\text{all}}(z) = \max_{a \in \mathcal{A}_k} (a - z)^\top \tau_B.$$

For a candidate $\widehat{s}$ selected using the same randomized trial, the candidate-set goal is an observable $C(\widehat{s}) \geq 0$ satisfying

$$\mathbb{P}\{R_B(\widehat{s}) \leq C(\widehat{s}) \mid \mathcal{F}_0\} \geq 1 - \delta. \tag{3}$$

The prespecified failure probability satisfies $\delta \in (0,1)$. The full allocation goal replaces R_B by R_B^{all}. When the true target cohort treatment effects are available for evaluation, define the candidate set gap

$$\Gamma_B = \max_{a \in \mathcal{A}_k} a^\top \tau_B - \max_r V_B(r),$$

the welfare difference between the best feasible allocation and the best allocation in the candidate set. For every candidate allocation,

$$R_B^{\text{all}}(z_s) = \Gamma_B + R_B(s). \tag{4}$$

Thus the two scopes differ only in the comparison set used to define regret. Boundary conventions and candidate-set augmentation are given in Appendix A.4.

Identification from the randomized trial. Let $\Phi \in \mathbb{R}^{n\times p}$ have row i equal to $\phi(X_i)^\top$, let $\Phi_B \in \mathbb{R}^{m\times p}$ have row j equal to $\phi(x_j^B)^\top$, and set $G = \Phi^\top\Phi$. We also consider singular G when the welfare differences being certified are supported by the trial.

Assumption 1 (Common treatment effect model)**.** *The feature map $\phi : \mathcal{X} \to \mathbb{R}^p$ is fixed before the trial assignments and outcomes used for certification are observed. There is a common coefficient vector $\theta \in \mathbb{R}^p$ such that $\tau(X_i) = \phi(X_i)^\top\theta$ for trial units and $\tau(x_j^B) = \phi(x_j^B)^\top\theta$ for target cohort members.*

The feature map may contain nonlinear transformations or interactions of the original covariates. The common coefficient vector is the restriction linking treatment effects at the trial and target covariates. If Φ has full column rank, the trial effect vector $\tau_{\mathrm{tr}} = \Phi\theta$ determines θ and hence every target allocation comparison. Under rank deficiency, a particular comparison may still be identified.

Proposition 1 (Identification of allocation welfare differences)**.** *Under Assumption 1, fix two allocations r, s and define $g_{rs} = \Phi_B^\top(z_r - z_s)$. Across the stated model class without further restrictions on θ, the welfare difference Δ_{rs} is identified from the conditional distribution of the observed randomized trial data if and only if $g_{rs} \in \mathrm{Range}(\Phi^\top)$. When this holds, $w_{rs} = \Phi G^\dagger g_{rs}$ satisfies $\Delta_{rs} = w_{rs}^\top\tau_{\mathrm{tr}}$, where $\tau_{\mathrm{tr}} = (\tau(X_i))_{i=1}^n$ and $G^\dagger$ is the Moore–Penrose pseudoinverse.*

Appendix B.1 proves the result. The constructions below assume that every candidate direction g_{rs}, or every exchange direction for full-allocation certification, lies in the trial row space. Figure 1(a) illustrates a one-person exchange: allocation r replaces treated member i in allocation s with member j. The shared welfare contributions cancel, giving $\Delta_{rs} = \tau(x_j^B) - \tau(x_i^B)$. Under Assumption 1, this difference is identified precisely when $\phi(x_j^B) - \phi(x_i^B)$ lies in the trial row space, even when the two target effects are not individually identified, as illustrated in Figure 1(b).

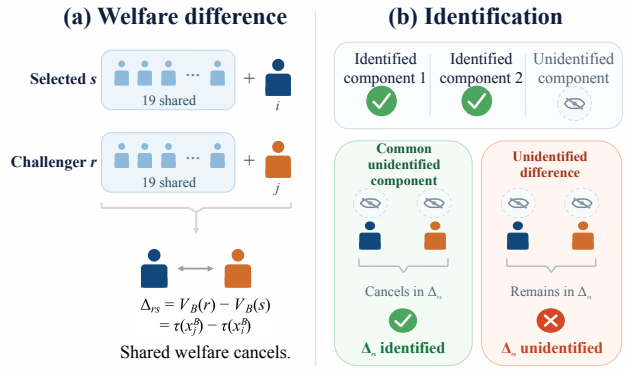

Figure 1: **A one-person allocation swap.** (a) The 19 shared members contribute equally to both allocations, leaving the effect difference between the exchanged members. (b) Under the common treatment effect model, a shared unidentified component cancels from this difference. Identification depends on whether the difference contains a component unidentified by the trial.

4 Method

We construct a finite-sample upper bound on the selected allocation's regret by estimating the welfare advantages of feasible alternative allocations. The method applies uncertainty directly to these pairwise welfare differences before the maximization that defines regret.

4.1 Candidate-set certification

Direct estimation of pairwise welfare differences. For a center ζ fixed before trial assignments and outcomes, define the centered inverse-probability weighted (IPW) score

$$T_i^{(\zeta)} = \left\{\frac{A_i}{e_i} - \frac{1-A_i}{1-e_i}\right\}(Y_i - \zeta), \qquad \widehat{\theta}_\zeta = G^\dagger\Phi^\top T^{(\zeta)}. \tag{5}$$

The common center cancels in expectation, giving $\mathbb{E}(T_i^{(\zeta)} \mid \mathcal{F}_0) = \tau(X_i)$. Since outcomes lie in $[0,1]$, we set $\zeta_{\mathrm{mid}} = 0.5$, the midpoint of this known range. It minimizes the worst-case score support length used to calibrate the bound. Write $T = T^{(\zeta_{\mathrm{mid}})}$ and $\widehat{\theta} = \widehat{\theta}_{\zeta_{\mathrm{mid}}}$. Appendix A.1 derives the concentration radii.

For two candidate allocations, let $d_{rs} = z_r - z_s$ and $g_{rs} = \Phi_B^\top d_{rs}$. Their estimated welfare difference is $\widehat{\Delta}_{rs} = g_{rs}^\top \widehat{\theta}$, while Assumption 1 gives $\Delta_{rs} = g_{rs}^\top \theta$. Cohort members with the same treatment status in both allocations have zero entries in d_{rs}, so their contributions disappear before uncertainty is propagated.

Simultaneous candidate set certificate. For a supported comparison, let $w_{rs} = \Phi G^\dagger g_{rs}$, so $\widehat{\Delta}_{rs} = w_{rs}^\top T$. For a family of K comparisons, let $a_K(w)$ be a one-sided sampling radius with failure probability δ/K for each comparison. We use the smaller of the design-based Hoeffding and Bernstein radii given in Appendix A.1. For the $K = M(M-1)$ ordered candidate comparisons, set $a_{rs} = a_K(w_{rs})$ and define

$$U_{rs} = \widehat{\Delta}_{rs} + a_{rs}, \qquad C_{\mathrm{pair}}(s) = \max_r U_{rs}, \qquad U_{ss} = a_{ss} = 0. \tag{6}$$

The maximum is a finite-sample upper bound on the largest welfare advantage among candidate alternatives. Simultaneous calibration is performed before the selected allocation is known, which permits the same trial to be used for allocation selection and certification.

Treatment effect model approximation error. We next allow the treatment effect specification in Assumption 1 to be approximate: $\tau_{\mathrm{tr}} = \Phi\theta + u_{\mathrm{tr}}$ and $\tau_B = \Phi_B\theta + u_B$. The two residuals enter the pairwise difference bias as

$$\mathbb{E}(\widehat{\Delta}_{rs} \mid \mathcal{F}_0) - \Delta_{rs} = w_{rs}^\top u_{\mathrm{tr}} - d_{rs}^\top u_B. \tag{7}$$

Writing $\xi = T - \mathbb{E}(T \mid \mathcal{F}_0)$, the realized error is $\Delta_{rs} - \widehat{\Delta}_{rs} = -w_{rs}^\top \xi + d_{rs}^\top u_B - w_{rs}^\top u_{\mathrm{tr}}$. The sampling radius a_{rs} controls the first term. For a prespecified residual class $\mathcal{U}$ containing the true pair, b_{rs} bounds the absolute bias in Equation (7), with $b_{ss} = 0$. The certificate adds both radii. A constant target residual cancels at fixed capacity, while a trial constant cancels only if the weights sum to zero (Appendix B.3).

For comparison, the treatment effect interval benchmark C_{eff} first constructs simultaneous intervals for target conditional treatment effects and propagates them through an allocation welfare difference. The allocation welfare interval benchmark C_{welf} first constructs simultaneous intervals for allocation welfare and then combines their endpoints. In comparisons using a common confidence region, all three constructions start from the same effect estimator and joint confidence region.

4.2 Full-allocation certification

The full feasible allocation set $\mathcal{A}_k$ contains $\binom{m}{k}$ allocations. Every alternative allocation can be represented by exchanges between treated and untreated cohort members. For a selected unit i and an unselected unit j, let d_{ji} remove i and add j. Write $g_{ji} = \Phi_B^\top d_{ji}$, $w_{ji} = \Phi G^\dagger g_{ji}$, and $\widehat{\Delta}_{ji} = g_{ji}^\top \widehat{\theta}$. Define

$$U_{ji} = \widehat{\Delta}_{ji} + a_{ji} + b_{ji}, \qquad a_{ji} = a_{K_{\mathrm{ex}}}(w_{ji}), \quad K_{\mathrm{ex}} = m(m-1). \tag{8}$$

The residual support b_{ji} is defined in Equation (14). All ordered exchanges are calibrated before the selected allocation is known.

For a selected allocation $\mathbf{1}_{\mathcal{I}}$, the indicator of treated set $\mathcal{I}$, let $\mathfrak{M}(\mathcal{I})$ be the partial one-to-one matchings from $\mathcal{I}$ to its complement, including the empty matching. Define

$$C_{\mathrm{pair}}^{\mathrm{all}}(\mathcal{I}) = \max_{\mathcal{M} \in \mathfrak{M}(\mathcal{I})} \sum_{(i,j)\in\mathcal{M}} U_{ji}. \tag{9}$$

Each feasible alternative allocation corresponds to a partial matching. Maximizing the summed exchange bounds therefore certifies regret over all exact-k allocations with a polynomial-time maximum-weight matching algorithm. Computation and boundary cases are given in Appendices A.4 and B.5.

5 THEORETICAL ANALYSIS

We establish validity after selection using the same trial for both candidate-set and full-allocation certification, then bound how far the certificates can exceed true regret.

5.1 VALIDITY AFTER SELECTION USING THE SAME TRIAL

Theorem 2 (Candidate-set certification after selection using the same trial)**.** *Under Section 3's conditions, assume every candidate direction g_{rs} lies in $\mathrm{Range}(\Phi^\top)$. Conditional on $\mathcal{F}_0$, with probability at least $1-\delta$, every ordered candidate pair satisfies $\Delta_{rs} \le U_{rs}$. Hence every measurable candidate selector based on the same randomized trial obeys*

$$R_B(\widehat{s}) \le C_{\mathrm{pair}}(\widehat{s}), \qquad R_B(s) \le C_{\mathrm{pair}}(s) \le R_B(s) + 2\max_r a_{rs} \quad \forall s.$$

For $M = 1$, both quantities are zero.

The common event covers every ordered candidate comparison before an allocation is selected. It therefore covers any candidate allocation chosen as a measurable function of the same trial. The second inequality controls the amount by which the certificate can exceed true candidate set regret. Appendix B.2 gives the proof. The residual-corrected version is Theorem 4 in Appendix B.3.

Theorem 3 (Full-allocation certification after selection using the same trial)**.** *Under the conditions in Section 3 except Assumption 1, assume every ordered exchange direction g_{ji} lies in $\mathrm{Range}(\Phi^\top)$ and suppose $\tau_{\mathrm{tr}} = \Phi\theta + u_{\mathrm{tr}}$, $\tau_B = \Phi_B\theta + u_B$, and $(u_{\mathrm{tr}}, u_B) \in \mathcal{U}$. Assume all exchange residual bounds b_{ji} are finite and observable. Conditional on $\mathcal{F}_0$, with probability at least $1-\delta$, simultaneously for every size-k treated index set $\mathcal{I}$,*

$$R_B^{\mathrm{all}}(\mathbf{1}_{\mathcal{I}}) \le C_{\mathrm{pair}}^{\mathrm{all}}(\mathcal{I}).$$

The guarantee includes any measurable allocation $\widehat{\mathcal{I}}$ selected using the same trial.

The exchange event is simultaneous over all ordered one-person exchanges. Every feasible alternative allocation decomposes into a partial matching of these exchanges, which yields the stated bound for every selected allocation. Appendix B.4 proves the matching reduction.

5.2 CERTIFICATE PRECISION

For $m \ge 2$ and $1 \le k \le m-1$, write $h_{\max} = \min(k, m-k)$ and $r_{\max} = \max_{i \ne j}(a_{ji} + b_{ji})$. On the validity event, Corollary 5 gives

$$C_{\mathrm{pair}}^{\mathrm{all}}(\mathcal{I}) - R_B^{\mathrm{all}}(\mathbf{1}_{\mathcal{I}}) \le 2h_{\max}r_{\max}.$$

The number of exchanges needed by an alternative allocation and the largest exchange uncertainty determine this worst-case excess. Appendix B.5 gives the proof and a bound using the realized matching weights.

Write $\|v\|_A = (v^\top A v)^{1/2}$ for a positive-semidefinite matrix A. When $G \succ 0$, Proposition 8 compares candidate-pair radii under one common ellipsoidal confidence region of radius β:

$$a_{rs}^{\mathrm{pair}} = \beta\|g_{rs}\|_{G^{-1}} \le \min\{a_{rs}^{\mathrm{eff}}, a_{rs}^{\mathrm{welf}}\}.$$

The treatment effect interval benchmark carries uncertainty through target treatment effect intervals, while the allocation welfare interval benchmark carries uncertainty through allocation welfare intervals. The pairwise difference construction applies the common confidence region to the feature difference after cohort members with unchanged treatment status cancel. Appendix B.6 proves the inequality and gives a sequence for which the radius ratio diverges.

6 EXPERIMENTS

We evaluate three aspects of certification. **Q1. Validity:** Do the certificates attain nominal coverage when allocation selection and certification use the same randomized trial? **Q2. Informativeness:** How far are the three certificates above true regret, how often do they certify a specified loss tolerance, and how does this change with trial size? **Q3. Sensitivity:** How do treatment effect model approximation error, calibration, and trial feature support affect certification?

Table 1: **Regret and certificates under two comparison scopes.** Mean true regret $\bar{R}$ and mean certificate $\bar{C}$ at $n = 6000, m = 500, k = 20$. All three constructions evaluate the same selected allocations using the same 95% confidence ellipsoid. Hillstrom true regret is unobserved. Lower $\bar{C}$ is tighter.

		Candidate-set				Full-allocation			
Data type	Setting	Regret $\bar{R}$	Pairwise $\bar{C}$	Effect $\bar{C}$	Welfare $\bar{C}$	Regret $\bar{R}$	Pairwise $\bar{C}$	Effect $\bar{C}$	Welfare $\bar{C}$
Synthetic	Raw-feature	0.007	**1.091**	3.510	4.367	0.287	**2.560**	5.293	5.830
	Nonlinear-feature	0.071	**2.590**	5.005	4.975	0.634	**4.144**	6.367	6.349
Semi-synthetic	SUPPORT2	0.024	**1.698**	4.007	4.328	0.366	**2.899**	5.155	5.250
Real data	Hillstrom-V	N/A	**3.482**	5.120	4.387	N/A	**4.672**	5.970	5.262
	Hillstrom-C	N/A	**3.467**	5.383	4.652	N/A	**5.029**	6.030	5.325

6.1 Experimental setup

Evaluation settings. Raw-feature and Nonlinear-feature use synthetic covariates and outcomes. Raw-feature uses raw covariates in the feature map, while Nonlinear-feature uses nonlinear transformations and an interaction. Both exactly satisfy Assumption 1. SUPPORT2 uses observed covariates with simulated treatments and outcomes (Asuncion et al., 2007). These three settings provide the true target cohort treatment effects needed to evaluate regret and coverage. Hillstrom-V and Hillstrom-C use the Visit and Conversion outcomes from a randomized email experiment (Hillstrom, 2008). Target conditional treatment effects are unobserved in Hillstrom, so these settings evaluate certificate size and sensitivity without true regret. Their common-ellipsoid displays assume zero model residual. Appendix D.1 examines declared nonzero residual limits, and Appendix C gives the data-generating processes and preprocessing details.

Certificate constructions. **Pairwise (ours)** bounds allocation welfare differences before maximizing over alternatives. Its full allocation version aggregates simultaneous exchange bounds through maximum-weight partial matching. **Treatment effect intervals (Effect)** propagate simultaneous intervals for target conditional treatment effects. **Allocation welfare intervals (Welfare)** combine simultaneous upper and lower bounds on allocation welfare. Within each trial, all three constructions use the same fitted feature map and selected allocation.

Protocol and calibration. Disjoint samples supply candidate training, target cohort covariates, and randomized trial inference. Nineteen prespecified scoring rules generate exact-k candidate allocations, with duplicates removed. Within each trial, $\widehat{s}$ maximizes $\widehat{V}_B(r) = z_r^\top \Phi_B \widehat{\theta}$ over this candidate set. Defaults are $n = 6000, m = 500, k = 20, p = 3$ and $\delta = 0.05$. Results with known treatment effects average five outer replicates, each with 200 randomized trial realizations conditional on fixed pretrial objects. Hillstrom results average three splits. Table 1 and Figures 2–3 use one finite-sample 95% confidence ellipsoid for all three constructions. Appendix D.4 reports method-specific calibration, and Appendix C gives the experimental protocols.

Evaluation metrics. For either comparison scope, let R be true regret and C its upper bound. We report their means $\bar{R}$ and $\bar{C}$, empirical coverage $\widehat{\mathrm{Cov}} = \widehat{\mathbb{E}}[\mathbf{1}\{R \le C\}]$, and the certification rate $\widehat{\mathrm{Cert}}(\epsilon) = \widehat{\mathbb{E}}[\mathbf{1}\{C \le \epsilon\}]$ for an allowed total welfare loss ϵ. Rates are percentages. Appendix C.10 specifies aggregation.

6.2 Experimental results

Coverage after same-trial selection. Across the three settings and four trial sizes in Figure 2, all three constructions cover true regret in all 12,000 evaluations under both comparison scopes. Selection and certification use the same trial. Appendix A.3 connects the common-ellipsoid event to the simultaneous candidate and exchange guarantees.

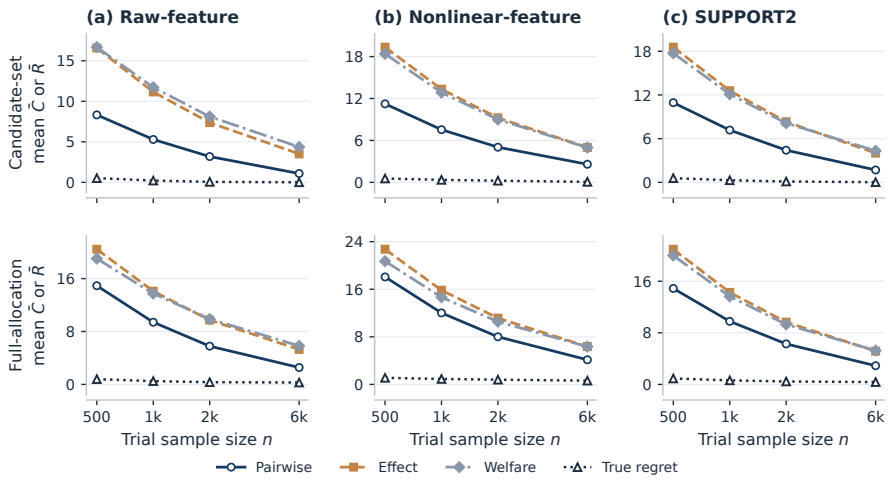

Figure 2: **Regret and certificates versus randomized trial size.** Candidate-set (top) and full-allocation (bottom) mean true regret $\bar{R}$ and mean certificate $\bar{C}$. All three constructions use the same selected allocation and the same 95% confidence ellipsoid at each trial size. Empirical coverage is 100% for each construction and scope.

Smaller mean certificates in the evaluated settings. Table 1 evaluates the same selected allocations under one confidence ellipsoid. Pairwise has the smallest mean certificate in every setting and scope. For Raw-feature, expanding the comparison set from the candidates to all exact-k allocations increases mean true regret from 0.007 to 0.287 and the Pairwise bound from 1.091 to 2.560. The regret increase measures the welfare gap between the best candidate and best feasible allocation. Full-allocation certification therefore also accounts for improvements omitted by the candidate set. Figure 2 shows all three bounds decreasing with trial size. Under method-specific Hoeffding calibration, the Raw-feature full-allocation means are 2.693, 4.310, and 13.784 for Pairwise, Effect, and Welfare (Appendix Tables 5 and 6).

Certificate size determines which tolerances can be certified. Figure 3 translates certificate size into the percentage of selected allocations certified at an allowed total welfare loss. Since welfare sums effects across k treated members, ϵ/k expresses the allowed loss per treatment slot. At $\epsilon = 2$, Pairwise certifies 87.3% of Raw-feature candidate-set trials, while Effect and Welfare certify none. At $\epsilon = 5$, the corresponding rates are 100%, 97.5%, and 71.9%. All Raw-feature true candidate regrets are below five, so these differences reflect bound conservatism.

Residual assumptions and calibration affect informativeness. In the partial-cancellation setting at $n = 6000$, Pairwise gives mean candidate-set and full-allocation bounds of 0.224 and 4.209 under a shared-coefficient residual class. The corresponding certification rates at tolerance five are 100% and 98.8%. An independent box class containing the same true residual gives bounds of 13.630 and 22.785, with no certifications at that tolerance. The shared class preserves cancellation between trial and target residuals along the allocation contrast. Across all four residual mechanisms, the lowest coverage averaged across the five designs is 99.8% when the declared class contains the true residual (Appendix Table 2). Calibration also affects the comparison: under weak overlap, Bernstein radii can reverse the ordering of Pairwise and Effect (Appendix Table 7).

Supported differences remain certifiable under rank deficiency. Figure 4 illustrates Proposition 1. Although the trial leaves a common effect component unidentified, the candidates differ in welfare by 0.10. At $n = 6000$, mean Pairwise, Effect, and Welfare bounds are 0.006, 0.194, and 2.000. At $\epsilon = 0.05$, Pairwise certifies 99.3% of selected allocations and the interval benchmarks certify none. The common unidentified component cancels in the welfare difference but remains in the marginal intervals. Appendix D.2 specifies the parameter region, range cap, and nonidentified comparison.

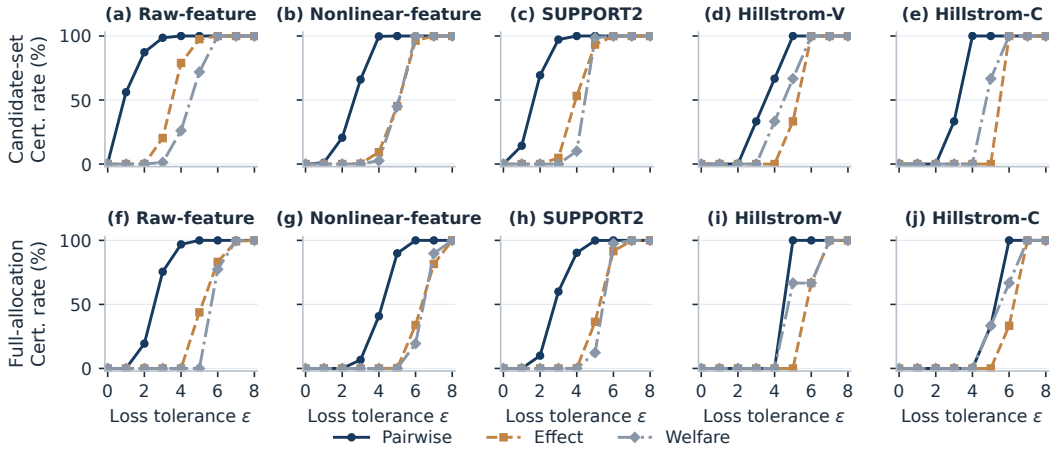

Figure 3: **Certification versus allowed total welfare loss.** Certification rate $\widehat{\mathrm{Cert}}(\epsilon)$, in percent, with certificate $C \leq \epsilon$ for candidate-set (top) and full-allocation (bottom) certification. All three constructions use the same selected allocation and common 95% confidence ellipsoid. Larger values mean more allocations are certified at a fixed tolerance.

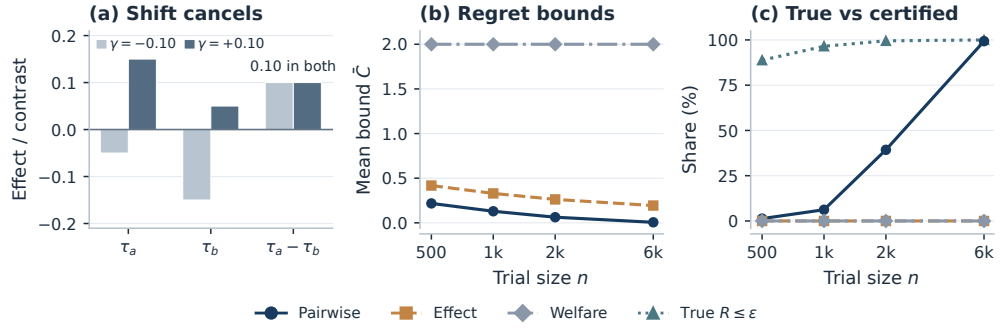

Figure 4: **Certification with a rank-deficient trial.** (a) The effects τ_a and τ_b belong to the two exchanged members. A trial-unidentified component shifts both effects but cancels from $\tau_a - \tau_b$. (b) Mean candidate-set bounds by trial size. (c) Shares with true regret or certificate at most 0.05. All methods use one parameter region. Each point averages five designs and 200 trials per design.

7 Conclusion

This paper studies how to certify the welfare loss of a capacity-constrained treatment allocation for a realized cohort using randomized trial evidence. Our analysis shows that certification can be based directly on the welfare differences between competing allocations, which may be identified even when some individual target treatment effects are not. Building on this observation, we develop finite-sample regret certificates that remain valid when the same trial is used for allocation selection and certification. The framework scales from prespecified candidate sets to all exact-k allocations through exchange bounds and maximum-weight partial matching, and our theory quantifies the conservativeness of the resulting certificates. These results provide a way to accompany a cohort-specific treatment allocation with a finite-sample guarantee on its welfare loss relative to the best allocation in the declared comparison set.

Limitations and future work. The current analysis uses a prespecified feature map and prespecified bounds on treatment effect model approximation error. An interesting next step is to learn these components from data while retaining finite-sample certification after accounting for the additional data dependence. Extending the framework beyond binary treatment and a single cardinality constraint is another natural direction, including settings with multiple treatments, interacting treatment effects, or multiple resource constraints.

AI USE STATEMENT

Generative AI tools assisted with manuscript editing and content polishing.

REFERENCES

Isaiah Andrews and Jiafeng Chen. Certified decisions. *arXiv preprint arXiv:2502.17830*, 2025.

David Arbour, Harsh Parikh, Bijan A. Niknam, Elizabeth Stuart, Kara E. Rudolph, and Avi Feller. Regularizing extrapolation in causal inference. *Proceedings of machine learning research*, 300: 784, 2026.

Amir Asiaee, Samhita Pal, Cole Beck, and Jared Davis Huling. Sharp bounds for treatment effect generalization under outcome distribution shift. *arXiv preprint arXiv:2602.09595*, 2026.

Arthur Asuncion, David Newman, et al. UCI machine learning repository, 2007.

Susan Athey and Stefan Wager. Policy learning with observational data. *Econometrica*, 89(1): 133–161, 2021.

Susan Athey, Julie Tibshirani, and Stefan Wager. Generalized random forests. *The Annals of Statistics*, 2019.

Yoav Benjamini, Yotam Hechtlinger, and Philip B. Stark. Confidence intervals for selected parameters. *arXiv preprint arXiv:1906.00505*, 2019.

Richard Berk, Lawrence Brown, Andreas Buja, Kai Zhang, and Linda Zhao. Valid post-selection inference. *The Annals of Statistics*, pp. 802–837, 2013.

Debopam Bhattacharya and Pascaline Dupas. Inferring welfare maximizing treatment assignment under budget constraints. *Journal of Econometrics*, 167(1):168–196, 2012.

Sílvia Casacuberta and Moritz Hardt. Good allocations from bad estimates. *arXiv preprint arXiv:2601.05597*, 2026.

Victor Chernozhukov, Sokbae Lee, Adam M. Rosen, and Liyang Sun. Policy learning with confidence. *arXiv preprint arXiv:2502.10653*, 2025.

Stephen R. Cole and Elizabeth A. Stuart. Generalizing evidence from randomized clinical trials to target populations: the ACTG 320 trial. *American journal of epidemiology*, 172(1):107–115, 2010.

Santiago Cortes-Gomez, Naveen Janaki Raman, Aarti Singh, and Bryan Wilder. Data-driven design of randomized control trials with guaranteed treatment effects. *arXiv preprint arXiv:2410.11212*, 2025.

Issa J. Dahabreh, Sarah E. Robertson, Eric J. Tchetgen Tchetgen, Elizabeth A. Stuart, and Miguel A. Hernán. Generalizing causal inferences from individuals in randomized trials to all trial-eligible individuals. *Biometrics*, 75(2):685–694, 2019.

Irina Degtiar and Sherri Rose. A review of generalizability and transportability. *Annual Review of Statistics and Its Application*, 10(1):501–524, 2023.

Tanner Fiez, Lalit Jain, Kevin G. Jamieson, and Lillian Ratliff. Sequential experimental design for transductive linear bandits. *Advances in neural information processing systems*, 32, 2019.

Tobias Hatt, Daniel Tschernutter, and Stefan Feuerriegel. Generalizing off-policy learning under sample selection bias. In *Uncertainty in Artificial Intelligence*, pp. 769–779. PMLR, 2022.

Kevin Hillstrom. The minethatdata e-mail analytics and data mining challenge. *MineThatData blog*, 2008.

Olivier Jeunen and Aleksei Ustimenko. δ-OPE: Off-policy estimation with pairs of policies. In *Proceedings of the 18th ACM Conference on Recommender Systems*, pp. 878–883, 2024.

Fahad Kamran, Maggie Makar, and Jenna Wiens. Learning to rank for optimal treatment allocation under resource constraints. In *International Conference on Artificial Intelligence and Statistics*, pp. 3727–3735. PMLR, 2024.

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

Mathias Kraus, Stefan Feuerriegel, and Maytal Saar-Tsechansky. Data-driven allocation of preventive care with application to diabetes mellitus type II. *Manufacturing & Service Operations Management*, 26(1):137–153, 2024.

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

Ilja Kuzborskij, Claire Vernade, Andras Gyorgy, and Csaba Szepesvari. Confident off-policy evaluation and selection through self-normalized importance weighting. In *International Conference on Artificial Intelligence and Statistics*, pp. 640–648. PMLR, 2021.

Alexander R. Luedtke and Mark J. van der Laan. Optimal individualized treatments in resource-limited settings. *The international journal of biostatistics*, 12(1):283–303, 2016.

Arnab Maiti, Yunbei Xu, and Kevin Jamieson. On the power of adaptivity for ε-best arm identification in linear bandits. *arXiv preprint arXiv:2605.15663*, 2026.

Charles F. Manski. Statistical treatment rules for heterogeneous populations. *Econometrica*, 72(4): 1221–1246, 2004.

Aditya Mate, Bryan Wilder, Aparna Taneja, and Milind Tambe. Improved policy evaluation for randomized trials of algorithmic resource allocation. In *International Conference on Machine Learning*, pp. 24198–24213. PMLR, 2023.

Giacomo Opocher. Learning where to look: Delaunay matching for policy choice and data collection. *arXiv e-prints*, pp. arXiv–2607, 2026.

Daniel F. Pellatt. PAC-bayesian treatment allocation under budget constraints. *arXiv preprint arXiv:2212.09007*, 2022.

Min Qian and Susan A. Murphy. Performance guarantees for individualized treatment rules. *Annals of statistics*, 39(2):1180, 2011.

Sarah E. Robertson, Jon A. Steingrimsson, and Issa J. Dahabreh. Regression-based estimation of heterogeneous treatment effects when extending inferences from a randomized trial to a target population. *European journal of epidemiology*, 38(2):123–133, 2023.

A IMPLEMENTATION AND CALIBRATION

A.1 CENTERED IPW SCORE AND SIMULTANEOUS CALIBRATION

The normalized outcome range is $[0, 1]$, giving $\zeta_{\mathrm{mid}} = 1/2$. Any pretrial fixed center preserves the IPW score's treatment effect mean. The midpoint minimizes its range-based envelopes. Appendix B.2 derives these statements for a general center. For each ordered pairwise comparison, write $d_{rs} = z_r - z_s$, $h_r = \Phi_B^\top z_r$, $g_{rs} = h_r - h_s$, and $w_{rs} = \Phi G^\dagger g_{rs}$ for a supported direction.

Each of K signed linear functionals receives failure probability δ/K. Superscripts H and B denote Hoeffding and Bernstein radii.

For a prespecified family of K signed linear functionals and a fixed weight vector w, let $t_K = \log(K/\delta)$ and $L_i = 1/\min(e_i, 1 - e_i)$. The support-length radius is

$$a_K^{\mathrm{H}}(w) = \left\{ \frac{t_K}{2} \sum_{i=1}^{n} w_i^2 L_i^2 \right\}^{1/2}. \tag{10}$$

For the variance-based radius, define

$$\overline{V}(w) = \sum_{i=1}^{n} \frac{w_i^2}{4e_i(1 - e_i)}, \qquad c(w) = \max_i \frac{|w_i|}{\min(e_i, 1 - e_i)}.$$

Then

$$a_K^{\mathrm{B}}(w) = \frac{c(w)t_K}{3} + \sqrt{2\overline{V}(w)t_K + \left\{ \frac{c(w)t_K}{3} \right\}^2}, \qquad a_K(w) = \min\{a_K^{\mathrm{H}}(w), a_K^{\mathrm{B}}(w)\}. \tag{11}$$

Here a_K^{H} is the propensity-specific Hoeffding radius and a_K^{B} is a Bernstein radius for the midpoint IPW score. Both depend only on information in $\mathcal{F}_0$. The smaller radius is therefore chosen before assignments and outcomes are observed, separately for each linear functional. In (10), L_i bounds the midpoint IPW score's support length. In (11), $\overline{V}(w)$ bounds its weighted variance and $c(w)$ bounds its largest centered summand. The proof is in Appendix B.2.

Under balanced assignment, the Bernstein radius is no smaller than the Hoeffding radius. Under heterogeneous assignment, either may be selected. The uniform-overlap specialization replaces each support length by $1/\eta$, as in (12).

A.2 METHOD-SPECIFIC CALIBRATION FOR THE THREE CONSTRUCTIONS

For a comparison family, Equation (10) uses the same total failure probability δ and a method-specific count K. Candidate-set Pairwise calibrates $M(M-1)$ ordered contrasts. Effect calibrates both signs of m individual effects, so $K = 2m$. Welfare calibrates upper and lower bounds for each of M allocations, so $K = 2M$. For full allocation, the corresponding counts are $m(m-1)$, $2m$, and $2\binom{m}{k}$. Each construction separately attains conditional coverage at least $1 - \delta$ under the exact treatment effect model when every queried functional is supported by the trial. The Effect and Welfare formulas in this subsection assume $G \succ 0$. Prespecified residual support terms give the corresponding approximate-model guarantees.

For Welfare, write $h_A = \Phi_B^\top z_A$ and $w_A = \Phi G^{-1} h_A$ when $G \succ 0$. Its simultaneous interval is centered at $h_A^\top \widehat{\theta}$, with radius $a_K^{\mathrm{H}}(w_A)$ or $a_K^{\mathrm{B}}(w_A)$. For a prespecified residual class, add the support term

$$b_A^{\mathrm{welf}} = \sup_{(v_{\mathrm{tr}}, v_B) \in \mathcal{U}} |z_A^\top v_B - w_A^\top v_{\mathrm{tr}}|$$

to either radius. This term is zero under exact specification. The certificate compares each alternative's upper endpoint with the selected allocation's lower endpoint, with a zero self-comparison. The full-allocation version maximizes over every exact-k allocation. For the chosen allocation family, K counts both signs of every allocation welfare. Each endpoint receives a Hoeffding or Bernstein bound with failure probability δ/K. A union bound covers all upper and lower endpoints on one event. On this event, $V_B(A) - V_B(S)$ is no larger than the upper endpoint for A minus the lower endpoint for S, simultaneously for every alternative A and selected allocation S. The same guarantee

therefore holds when S is selected using the trial. In the three-dimensional experiment of Table 7, both radii are convex in h_A. The maximum upper endpoint occurs at a vertex of the feasible feature-sum hull. Exact top-k sorting exposes these vertices. Lower and upper numerical endpoints differ by at most 10^{-4} welfare units. The common-ellipsoid comparisons below use a different calibration.

A.3 Common confidence ellipsoid for method comparisons

The comparisons based on one common confidence ellipsoid use a different calibration from the finite family of pairwise difference comparisons above. Under the exact treatment effect specification in Assumption 1 and $G \succ 0$, define

$$\mathcal{E}_\delta = \{\vartheta : \|\vartheta - \widehat{\theta}\|_G \le \beta_\delta\}, \qquad \beta_\delta = \frac{1}{2\eta}\sqrt{2p\log(2p/\delta)}.$$

For its conditional finite-sample coverage, let $q = G^{-1/2}\Phi^\top(T - \Phi\theta)$. Each coordinate of q is a sum of independent, conditionally mean-zero terms. Its coefficients have squared sum one because $G^{-1/2}\Phi^\top\Phi G^{-1/2} = I_p$, while the midpoint IPW score has support length at most $1/\eta$. Hoeffding's inequality and a union bound over both signs of p coordinates give

$$\mathbb{P}\left\{\max_{j\le p}|q_j| > \frac{1}{2\eta}\sqrt{2\log(2p/\delta)} \;\middle|\; \mathcal{F}_0\right\} \le \delta.$$

Since $\|\widehat{\theta}-\theta\|_G = \|q\|_2 \le \sqrt{p}\max_j|q_j|$, it follows that $\mathbb{P}\{\theta \in \mathcal{E}_\delta \mid \mathcal{F}_0\} \ge 1-\delta$. On $\{\theta \in \mathcal{E}_\delta\}$, every candidate and exchange direction g has absolute estimation error at most $\beta_\delta\|g\|_{G^{-1}}$. The candidate maximization and matching arguments in Theorems 2 and 3 therefore apply with these radii on this single event. The candidate and matching excess inequalities follow with these substituted radii from the same two-sided bounds. Projection onto individual effects or allocation welfares gives simultaneous intervals for the two benchmarks. This coverage requires Assumption 1. For Hillstrom, the region is evaluated under this zero-residual assumption, whose validity cannot be checked from the observed target covariates.

A.4 Boundary cases and comparison sets

For $M = 1$, the candidate set regret and certificate are zero, and no pairwise comparison radius is evaluated. For $m \le 1$ or $k \in \{0, m\}$, the full allocation regret and certificate are zero. The matching algorithm operates on computed finite edge bounds. Polynomial aggregation counts arithmetic and comparison operations on these bounds. Appendix B.4 gives the reduction. When $m = 0$, the unique feasible allocation is the empty allocation and the candidate set contains only that allocation. For $\mathcal{I} \subseteq [m]$, $\mathbf{1}_\mathcal{I} \in \{0,1\}^m$ denotes its indicator vector.

For a fixed allocation $z \in \mathcal{A}_k$ and a nonempty candidate set $\mathcal{Q} \subseteq \mathcal{A}_k$ that may exclude it, use $\mathcal{Q}_z = \mathcal{Q} \cup \{z\}$. Define regret and the candidate set gap against the same augmented candidate set, so both are nonnegative and still sum to $R_B^{\text{all}}(z)$. Calibration must cover a prespecified family containing all possible selected and alternative allocations. A candidate set augmented after selection cannot determine the calibration family.

A.5 Reference bound from the treatment effect range

If target effects also lie in $[-1, 1]$, then $R_B^{\text{all}} \le H_{\text{all}} = 2\min(k, m-k)$ and $R_B(s) \le H_s = \max_r \|z_r - z_s\|_1$. The target-range condition is not implied by trial randomization. Replacing a bound C by $\min(C, H)$ requires no extra failure probability, but cannot change certification at a tolerance below H. At $m = 500, k = 20$, $H_{\text{all}} = 40$.

B Theoretical details and proofs

All probability statements in this section condition on $\mathcal{F}_0$.

B.1 Identification with rank-deficient trial features

Proof of Proposition 1. Condition on $\mathcal{F}_0$, fixing both feature matrices. Randomization, overlap, consistency, and the conditional mean restriction in Equation (1) make $\tau(X_i) = \mathbb{E}(Y_i \mid A_i =$

$1, \mathcal{F}_0) - \mathbb{E}(Y_i \mid A_i = 0, \mathcal{F}_0)$ a functional of the complete trial observed distribution for each i. Identification here concerns that conditional distribution. If $g_{rs} \in \mathrm{Range}(\Phi^\top)$, choose w with $\Phi^\top w = g_{rs}$. Assumption 1 gives

$$w^\top \tau_{\mathrm{tr}} = w^\top \Phi\theta = g_{rs}^\top \theta = \Delta_{rs},$$

establishing sufficiency. The matrix identity $\mathrm{Range}(G) = \mathrm{Range}(\Phi^\top)$ implies $GG^\dagger g_{rs} = g_{rs}$, so $w = \Phi G^\dagger g_{rs}$ is one such choice. In the full-rank case it reduces to the main construction $\Phi G^{-1} g_{rs}$.

For necessity over the model class defined by Assumption 1, suppose $g_{rs} \notin \mathrm{Range}(\Phi^\top)$. Linear algebra gives $v \in \ker(\Phi)$ with $g_{rs}^\top v \neq 0$. Let $h = \max_j |\phi(x_j^B)^\top v| > 0$ and choose a nonzero t with $|t|h \leq 1/4$. Construct two models with parameters $\theta_0 = 0$ and $\theta_1 = tv$. In both, trial control and treatment potential outcomes are Bernoulli$(1/2)$, with identical independent randomization. Their entire conditional trial observed distributions therefore agree, since $\Phi v = 0$. For target members, give both models control mean $1/2$. Set treatment means to $1/2$ in model zero and $1/2 + t\phi(x_j^B)^\top v$ in model one. These are within $[1/4, 3/4]$, satisfy Assumption 1, and can be extended to compatible conditional mean functions. Yet their pairwise welfare differences differ by $tg_{rs}^\top v \neq 0$. Hence the trial law does not identify the welfare difference throughout this model class. Extra coefficient restrictions at a particular boundary law can change pointwise identification. The proposition concerns identification throughout the stated model class.

Supported comparison families. For any supported g, the pretrial weight $w = \Phi G^\dagger g$ satisfies $\Phi^\top w = g$. Thus the sampling argument in Appendix B.2 applies without full rank. Support must be checked before using the pseudoinverse: under the exact model, an unsupported g would give expected estimate $g^\top G^\dagger G\theta$, which can differ from the target $g^\top\theta$. For full-allocation certification with $0 < k < m$, every ordered exchange is supported exactly when the projections of all target feature vectors onto $\ker(\Phi)$ are equal. Indeed, an exchange direction $\phi(x_j^B) - \phi(x_i^B)$ is orthogonal to $\ker(\Phi)$ exactly when those projections agree. Under this condition all exact-k allocation contrasts are supported, and the matching proof in Appendix B.4 applies to the pseudoinverse weights. For $k = 0$ or $k = m$, there is only one feasible allocation and its regret is zero. The individual treatment effect interval benchmark continues to require identifiable individual effects or a separately declared finite parameter range.

B.2 Candidate-set validity: IPW score, concentration, and simultaneous coverage

IPW score and centering. If only the overlap envelope is used, the propensity-specific radius in Equation (10) is bounded by

$$\bar{a}_{rs} = \frac{1}{2\eta}\sqrt{2\log(K/\delta)}\,\underbrace{\sqrt{g_{rs}^\top G^\dagger g_{rs}}}_{\ell_{rs}}. \tag{12}$$

Let $T_i^{(\zeta)}$ be the IPW score in (5) for a fixed $\zeta \in [0,1]$, measurable with respect to $\mathcal{F}_0$, and use the corresponding $\widehat{\theta}_\zeta$, estimated welfare differences, and radii throughout. On the normalized outcome range, the default is $\zeta_{\mathrm{mid}} = 1/2$. The case $\zeta = 0$ verifies the uncentered ablation. Randomization, consistency, and the conditional mean restriction in Equation (1) imply

$$\mathbb{E}\left[\frac{A_i(Y_i - \zeta)}{e_i} - \frac{(1-A_i)(Y_i - \zeta)}{1-e_i}\,\middle|\,\mathcal{F}_0\right] = \mathbb{E}[Y_i(1) - \zeta \mid \mathcal{F}_0] - \mathbb{E}[Y_i(0) - \zeta \mid \mathcal{F}_0] = \tau(X_i).$$

Conditional on $\mathcal{F}_0$, the IPW score belongs to the interval with endpoints

$$-\max\left\{\frac{\zeta}{e_i}, \frac{1-\zeta}{1-e_i}\right\}, \qquad \max\left\{\frac{1-\zeta}{e_i}, \frac{\zeta}{1-e_i}\right\}.$$

Thus the tight enclosing interval based only on $Y_i \in [0,1]$ has length

$$L_i^{(\zeta)} = \max\left\{\frac{1-\zeta}{e_i}, \frac{\zeta}{1-e_i}\right\} + \max\left\{\frac{\zeta}{e_i}, \frac{1-\zeta}{1-e_i}\right\}. \tag{13}$$

Particular outcome distributions can have shorter support. In particular,

$$L_i^{(0)} = \frac{1}{e_i} + \frac{1}{1-e_i}, \qquad L_i^{(\zeta_{\mathrm{mid}})} = \frac{1}{\min(e_i, 1-e_i)}.$$

The function $L_i^{(\zeta)}$ is convex and invariant under $\zeta \mapsto 1-\zeta$, so it is minimized at ζ_{mid}. Under unbalanced assignment, other centers can attain the same minimum. Also,

$$\mathrm{Var}(T_i^{(\zeta)} \mid \mathcal{F}_0) \le \mathbb{E}\{(T_i^{(\zeta)})^2 \mid \mathcal{F}_0\} \le \frac{\max\{\zeta, 1-\zeta\}^2}{e_i(1-e_i)}.$$

The midpoint uniquely minimizes $\max\{\zeta, 1-\zeta\}$, the worst-case absolute deviation of a normalized outcome from its center. Actual variances and certificates also depend on the outcome distribution and estimated welfare difference.

Concentration for one fixed comparison. For any fixed supported direction $g \in \mathrm{Range}(\Phi^\top)$, put $w = \Phi G^\dagger g$ and $\xi_i = T_i^{(\zeta)} - \mathbb{E}(T_i^{(\zeta)} \mid \mathcal{F}_0)$. Then

$$g^\top(\widehat{\theta}_\zeta - \theta) = w^\top \xi, \qquad \|w\|_2^2 = g^\top G^\dagger G G^\dagger g = g^\top G^\dagger g.$$

Hoeffding's lemma and conditional independence give

$$\mathbb{E}\{\exp(t w^\top \xi) \mid \mathcal{F}_0\} \le \exp\left\{\frac{t^2}{8}\sum_i w_i^2 (L_i^{(\zeta)})^2\right\}.$$

Consequently, with

$$a_K^{\mathrm{H},(\zeta)}(w) = \left\{\frac{1}{2}\log(K/\delta)\sum_i w_i^2 (L_i^{(\zeta)})^2\right\}^{1/2},$$
$$\mathbb{P}\{-w^\top \xi > a_K^{\mathrm{H},(\zeta)}(w) \mid \mathcal{F}_0\} \le \delta/K.$$

For the remainder of the main-theorem proof set $\zeta = \zeta_{\mathrm{mid}}$, so $T = T^{(\zeta_{\mathrm{mid}})}$ and ξ is its mean-zero noise. The preceding second-moment bound specializes to

$$\mathrm{Var}(T_i \mid \mathcal{F}_0) \le \mathbb{E}(T_i^2 \mid \mathcal{F}_0) \le \frac{1}{4e_i} + \frac{1}{4(1-e_i)} = \frac{1}{4e_i(1-e_i)}.$$

$|w_i\{T_i - \mathbb{E}(T_i \mid \mathcal{F}_0)\}| \le |w_i|/\min(e_i, 1-e_i)$. Hence the sum has conditional variance at most $\overline{V}(w)$ and every centered summand is bounded in absolute value by $c(w)$. Conditional Bernstein's inequality yields

$$\mathbb{P}\{-w^\top \xi > x \mid \mathcal{F}_0\} \le \exp\left\{-\frac{x^2}{2\{\overline{V}(w) + c(w)x/3\}}\right\}.$$

The positive solution of $x^2 = 2\log(K/\delta)\{\overline{V}(w) + c(w)x/3\}$ is exactly $a_K^{\mathrm{B}}(w)$ in Equation (11). Thus both $a_K^{\mathrm{H}}(w)$ and $a_K^{\mathrm{B}}(w)$ have conditional failure probability at most δ/K. Their values and the choice of the smaller one are $\mathcal{F}_0$-measurable. The chosen radius $a_K(w)$ therefore has the same failure bound for each fixed direction. Its selection uses pretrial information.

One event for every candidate. A union bound over the $K = M(M-1)$ ordered candidate comparisons proves the simultaneous one-sided statement for $a_{rs} = a_K(w_{rs})$. Since each reversed pair is also present and all three radii are invariant to replacing w by $-w$, the event also gives the corresponding two-sided error bound. For zero directions, the welfare-difference error is exactly zero and no tail division is needed. On the event,

$$R_B(s) = \max_r \Delta_{rs} \le \max_r U_{rs} = C_{\mathrm{pair}}(s) \le R_B(s) + 2\max_r a_{rs}$$

for every s simultaneously. Substituting any selected index proves validity. No independence between selection and certification outcomes is required once this event is available. Independence is instead required when freezing the comparison directions and feature map. The uncentered ablation uses only $a_K^{\mathrm{H},(0)}$. The Bernstein calculation above is specific to the fixed midpoint IPW score.

Uniform overlap corollary and fixed centering. For the midpoint IPW score, $L_i^{(\zeta_{\mathrm{mid}})} \le 1/\eta$ and $\sum_i w_{rs,i}^2 = \ell_{rs}^2$, so $a_K(w_{rs}) \le a_K^{\mathrm{H}}(w_{rs}) \le \bar{a}_{rs}$ in (12). Under balanced assignment, $\overline{V}(w) = \|w\|_2^2$, $c(w) = 2\|w\|_\infty$, and

$$a_K^{\mathrm{B}}(w) = \frac{2\|w\|_\infty t_K}{3} + \sqrt{2\|w\|_2^2 t_K + \left\{\frac{2\|w\|_\infty t_K}{3}\right\}^2} \ge \sqrt{2t_K}\|w\|_2 = a_K^{\mathrm{H}}(w).$$

Thus the selected radius equals the propensity-specific Hoeffding radius under balanced assignment. It also equals the uniform-envelope radius when the sharp envelope $\eta = 1/2$ is used. At a known e, the uncentered-to-midpoint support-length ratio is $1/\max(e, 1-e)$, equal to two under balanced assignment. The final certificate also depends on the estimated welfare difference. A center fitted on trial outcomes would require a separate dependence analysis.

B.3 Joint trial-to-target residual correction

The worst absolute welfare-difference bias over the prespecified residual class is

$$b_{rs} = \sup_{(v_{\mathrm{tr}}, v_B) \in \mathcal{U}} |d_{rs}^\top v_B - w_{rs}^\top v_{\mathrm{tr}}|. \tag{14}$$

Theorem 4 (Joint trial-to-target residual correction)**.** *Under the conditions in Section 3 except Assumption 1, assume every candidate direction g_{rs} lies in $\mathrm{Range}(\Phi^\top)$ and the existence of $\theta \in \mathbb{R}^p$, $u_{\mathrm{tr}} \in \mathbb{R}^n$, and $u_B \in \mathbb{R}^m$ such that $\tau_{\mathrm{tr}} = \Phi\theta + u_{\mathrm{tr}}$, $\tau_B = \Phi_B\theta + u_B$, and $(u_{\mathrm{tr}}, u_B) \in \mathcal{U}$, where $\mathcal{U}$ is $\mathcal{F}_0$-measurable. If every b_{rs} is finite and observable, define $U_{rs} = \widehat{\Delta}_{rs} + a_{rs} + b_{rs}$. Then, with conditional probability at least $1-\delta$, simultaneously for every s,*

$$R_B(s) \le C_{\mathrm{pair}}(s) \le R_B(s) + 2\max_r(a_{rs} + b_{rs}),$$

and the same inequalities hold after substituting any measurable candidate selector based on the same trial.

We prove Theorem 4. Conditional on $\mathcal{F}_0$, the residual class and all b_{rs} are fixed. Write the IPW score vector as $T = \Phi\theta + u_{\mathrm{tr}} + \xi$, where ξ has mean zero and the same bounded-score concentration. Then

$$\widehat{\Delta}_{rs} - \Delta_{rs} = w_{rs}^\top \xi + w_{rs}^\top u_{\mathrm{tr}} - d_{rs}^\top u_B.$$

If the true residual pair lies in the class used to define b_{rs}, then on the midpoint concentration event

$$\Delta_{rs} - \widehat{\Delta}_{rs} \le a_{rs} + |d_{rs}^\top u_B - w_{rs}^\top u_{\mathrm{tr}}| \le a_{rs} + b_{rs}.$$

The reversed ordered pair gives the corresponding opposite inequality. Thus replacing every a_{rs} by $a_{rs} + b_{rs}$ in the simultaneous, selection, and excess-bound arguments is valid. The residual class and its bound must be fixed before certification assignments and outcomes, conditional on the objects listed in Section 3. When $M = 1$, there is no distinct ordered pair, so regret and certificate are exactly zero without evaluating a logarithm. For $\mathcal{U} = \{(0,0)\}$, every support vanishes and the complete argument reduces exactly to Theorem 2.

Three standard residual classes give closed forms. For fixed $\Psi_{\mathrm{tr}} \in \mathbb{R}^{n\times h}$, $\Psi_B \in \mathbb{R}^{m\times h}$, $Q \succeq 0$, and $\kappa \ge 0$, define the joint ellipsoidal class as $\mathcal{U}_{\mathrm{ell}} = \{(\Psi_{\mathrm{tr}}Q^{1/2}t, \Psi_B Q^{1/2}t) : \|t\|_2 \le \kappa\}$. For fixed width vectors $b_{\mathrm{tr}} \in \mathbb{R}_+^n$ and $b_B \in \mathbb{R}_+^m$, define the independent box by $\mathcal{U}_{\mathrm{box}} = \{(v_{\mathrm{tr}}, v_B) : |v_{\mathrm{tr},i}| \le b_{\mathrm{tr},i},\ |v_{B,j}| \le b_{B,j}\}$. For a common function f in a fixed reproducing kernel Hilbert space (RKHS) $\mathcal{H}$, define the joint kernel class as exactly its trial and target evaluation pairs for $\|f\|_{\mathcal{H}} \le \kappa$.

For the ellipsoidal class, write $v_{\mathrm{tr}} = \Psi_{\mathrm{tr}}Q^{1/2}t$ and $v_B = \Psi_B Q^{1/2}t$. Then the residual functional is

$$\left\langle Q^{1/2}(\Psi_B^\top d_{rs} - \Psi_{\mathrm{tr}}^\top w_{rs}), t \right\rangle,$$

whose supremum over $\|t\|_2 \le \kappa$ is exactly $b_{rs}^{\mathrm{ell}} = \kappa\|Q^{1/2}(\Psi_B^\top d_{rs} - \Psi_{\mathrm{tr}}^\top w_{rs})\|_2$. For the independent box, each coordinate can choose the sign of its coefficient independently, so the triangle upper bound is attained and equals

$$b_{rs}^{\mathrm{box}} = \sum_j |d_{rs,j}| b_{B,j} + \sum_i |w_{rs,i}| b_{\mathrm{tr},i}.$$

For the common RKHS class, use the same positive-semidefinite kernel for all trial and target blocks. With block matrices $K_{BB}, K_{\mathrm{tr,tr}}, K_{B,\mathrm{tr}}$, the Riesz representer of the residual functional has squared norm

$$d_{rs}^\top K_{BB} d_{rs} + w_{rs}^\top K_{\mathrm{tr,tr}} w_{rs} - 2 d_{rs}^\top K_{B,\mathrm{tr}} w_{rs}.$$

Cauchy–Schwarz gives

$$b_{rs}^{\mathcal{H}} = \kappa \sqrt{d_{rs}^\top K_{BB} d_{rs} + w_{rs}^\top K_{\mathrm{tr,tr}} w_{rs} - 2 d_{rs}^\top K_{B,\mathrm{tr}} w_{rs}},$$

and equality is attained by the normalized representer when it is nonzero (both sides are zero otherwise). This norm retains dependence between the trial and target evaluation functionals. Here $K_{BB} \in \mathbb{R}^{m\times m}$, $K_{\mathrm{tr,tr}} \in \mathbb{R}^{n\times n}$, and $K_{B,\mathrm{tr}} \in \mathbb{R}^{m\times n}$ with $K_{B,\mathrm{tr}}[j,i] = K(x_j^B, X_i)$.

To see why the trial term matters within the original RKHS premise, use the finite covariate domain $\{s, b_1, b_2\}$ and let every trial covariate equal s. Set $\phi(s) = \phi(b_1) = 1$, $\phi(b_2) = 0$, and $\theta = 0$. Let $q(s) = 1$, $q(b_1) = q(b_2) = 0$, use the rank-one kernel $K_r(x, x') = q(x)q(x')$, and take the common residual function $u = -q$, whose K_r-RKHS norm is one. In the trial, let the deterministic potential outcomes be $Y(1) = 0, Y(0) = 1$ and use $e = 1/2$. At both target points let the treatment effect be zero. Outcomes remain in $[0, 1]$, $u_{\mathrm{tr}} = -\mathbf{1}_n$, and $u_B = 0$.

For the pairwise comparison that treats b_1 in place of b_2, the true difference is zero. The midpoint IPW score equals -1 under either assignment, so $\widehat{\theta} = -1$ and $\widehat{\Delta} = -1$ deterministically. Its sampling radius is $a = \sqrt{2\log(2/\delta)/n}$. Because $K_{BB} = 0$, a target-only radius is zero. For every $n > 2\log(2/\delta)$ its claimed upper bound $-1 + a$ is negative and fails to cover the true zero welfare difference with probability one. The joint trial-to-target correction has $b_{rs} = 1$ and restores the bound. The example isolates trial-side projection bias under randomized treatment.

B.4 Full-allocation certificate proof

Proof of Theorem 3. Here d_{ji} has entry -1 at i, entry 1 at j, and zero elsewhere, so $\Delta_{ji} = d_{ji}^\top \tau_B$ and $\widehat{\Delta}_{ji} = w_{ji}^\top T$. Let $\xi = T - \mathbb{E}(T \mid \mathcal{F}_0)$. For each ordered pair $i \neq j$, the least-squares identity and the approximate residual decomposition give

$$\Delta_{ji} - \widehat{\Delta}_{ji} = -w_{ji}^\top \xi + d_{ji}^\top u_B - w_{ji}^\top u_{\mathrm{tr}}. \tag{15}$$

Indeed, $\widehat{\Delta}_{ji} = w_{ji}^\top T$, $w_{ji}^\top \Phi\theta = g_{ji}^\top \theta$, and $\Delta_{ji} = g_{ji}^\top \theta + d_{ji}^\top u_B$. Conditional on $\mathcal{F}_0$, the centered IPW score coordinates are independent and the exchange weights and propensities are fixed. The conditional Hoeffding and Bernstein calculations in Appendix B.2 each give failure probability at most δ/K_{ex} for their respective radius. Because the smaller radius in Equation (8) is selected using only $\mathcal{F}_0$, the selected inequality has the same conditional failure bound. If the direction is zero, every weighted centered term is zero and the sampling inequality holds surely. Membership of the true residual pair in $\mathcal{U}$ bounds the last two terms of Equation (15) above by b_{ji}. A union bound over the K_{ex} ordered pairs therefore yields, with conditional probability at least $1 - \delta$, the single event

$$\mathcal{E}_{\mathrm{ex}} = \bigcap_{i\neq j} \{\Delta_{ji} \leq U_{ji}\}. \tag{16}$$

No independence among the exchange events is needed. For $m \leq 1$, this is an empty intersection and hence the sure event, without evaluating K_{ex} or its logarithm.

Fix a feasible treated set $\mathcal{I}$ and an alternative allocation $\mathcal{J}$ of the same size. The removed indices $D = \mathcal{I} \setminus \mathcal{J}$ and added indices $A = \mathcal{J} \setminus \mathcal{I}$ have the same cardinality. Any bijection $\pi : D \to A$ gives a partial matching $\mathcal{M} = \{(i, \pi(i)) : i \in D\}$ and the exact identity

$$(\mathbf{1}_{\mathcal{J}} - \mathbf{1}_{\mathcal{I}})^\top \tau_B = \sum_{(i,j)\in\mathcal{M}} \Delta_{ji}.$$

Conversely, a partial matching $\mathcal{M}$ removes its distinct left endpoints from $\mathcal{I}$ and adds its distinct right endpoints, giving a feasible alternative allocation with the same identity. The empty matching corresponds to $\mathcal{J} = \mathcal{I}$. Consequently,

$$R_B^{\mathrm{all}}(\mathbf{1}_{\mathcal{I}}) = \max_{\mathcal{M}\in\mathfrak{M}(\mathcal{I})} \sum_{(i,j)\in\mathcal{M}} \Delta_{ji}.$$

On $\mathcal{E}_{\mathrm{ex}}$, each summand is at most its corresponding U_{ji}. The resulting inequality holds for every feasible $\mathcal{I}$ on that one event, so substituting a measurable $\widehat{\mathcal{I}}$ selected from the same assignments and outcomes does not consume another confidence budget. If $k = 0$ or $k = m$, the empty matching is the only one, and both regret and certificate equal zero.

For exact aggregation from computed edge bounds, replace real edge weights by $(U_{ji})_+ = \max(U_{ji}, 0)$, add $m - k$ zero-weight dummy vertices to the left and k to the right, and give every edge incident to a dummy weight zero. A real partial matching extends to a perfect matching in this m-by-m graph by pairing its unmatched real vertices with dummies and then pairing the remaining dummies. Conversely, the positive real–real edges of any perfect matching form a feasible partial matching with the same weight. Removing a nonpositive real edge never lowers the partial optimum, so the two optima are equal. A standard square-assignment algorithm gives a polynomial number of exact arithmetic and comparison operations on explicitly computed weights. The cost of residual-class support functions and bit complexity is separate.

B.5 Full-allocation precision and treatment effect interval benchmark

The same event also quantifies the conservatism introduced by replacing true exchange gains with their upper confidence bounds. For $m \geq 2$ and $1 \leq k \leq m - 1$, define

$$r_{ji} = a_{ji} + b_{ji}, \qquad r_{\max} = \max_{i \neq j} r_{ji}, \qquad h_{\max} = \min(k, m - k).$$

For later use, write $\phi_j^B = \phi(x_j^B)$ and $\widehat{\tau}_j = (\phi_j^B)^\top \widehat{\theta}$.

Corollary 5 (Full-allocation certificate excess)**.** *Under Theorem 3's conditions, for $m \geq 2$ and $1 \leq k \leq m - 1$, with conditional probability at least $1 - \delta$, simultaneously for every size-k set $\mathcal{I}$,*

$$R_B^{\mathrm{all}}(\mathbf{1}_{\mathcal{I}}) \leq C_{\mathrm{pair}}^{\mathrm{all}}(\mathcal{I}) \leq R_B^{\mathrm{all}}(\mathbf{1}_{\mathcal{I}}) + 2 \max_{\mathcal{M} \in \mathfrak{M}(\mathcal{I})} \sum_{(i,j) \in \mathcal{M}} r_{ji} \leq R_B^{\mathrm{all}}(\mathbf{1}_{\mathcal{I}}) + 2 h_{\max} r_{\max}. \tag{17}$$

The same inequality holds for a size-k set selected using the trial.

With $\ell_{ji} = (g_{ji}^\top G^\dagger g_{ji})^{1/2}$, the uniform-overlap radius makes the terms explicit:

$$r_{\max} \leq \frac{1}{2\eta} \sqrt{2 \log \frac{m(m-1)}{\delta}} \max_{i \neq j} \ell_{ji} + \max_{i \neq j} b_{ji}.$$

Thus the trial feature matrix, overlap, and the residual class each contribute to the excess above true regret.

Proof of Corollary 5. Reversing an ordered pair negates its welfare difference and weight. The sampling radius is symmetric under this reversal, and the absolute residual support in Equation (14) gives $b_{ji} = b_{ij}$, so $r_{ji} = a_{ji} + b_{ji} = r_{ij}$. Since the event $\mathcal{E}_{\mathrm{ex}}$ covers both orientations, $|\widehat{\Delta}_{ji} - \Delta_{ji}| \leq r_{ji}$ for every pair. For any feasible matching $\mathcal{M}$,

$$\sum_{(i,j) \in \mathcal{M}} U_{ji} \leq \sum_{(i,j) \in \mathcal{M}} \Delta_{ji} + 2 \sum_{(i,j) \in \mathcal{M}} r_{ji}.$$

Maximizing over matchings and using the exact matching identity for R_B^{all} proves the middle inequality in (17). Each matching has at most $h_{\max}$ edges, so its radius sum is at most $h_{\max} r_{\max}$. Theorem 3 gives the first inequality. Because one event covers every feasible $\mathcal{I}$, it also covers a set selected using the same trial.

Treatment effect interval benchmark with simultaneous sampling radii. The Effect benchmark also gives a valid full-allocation certificate. For target unit j, set

$$w_j^{\mathrm{eff}} = \Phi G^{-1} \phi_j^B, \qquad \widehat{\tau}_j = (\phi_j^B)^\top \widehat{\theta}, \qquad r_j = a_{2m}(w_j^{\mathrm{eff}}).$$

The multiplicity $2m$ covers both signs of every treatment effect coordinate. As with the exchange bounds, the choice between the Hoeffding and Bernstein radii is made from $\mathcal{F}_0$ alone. For $m \geq 1$, define

$$b_j = \sup_{(v_{\mathrm{tr}}, v_B) \in \mathcal{U}} |v_{B,j} - (w_j^{\mathrm{eff}})^\top v_{\mathrm{tr}}|, \qquad \ell_j = \widehat{\tau}_j - r_j - b_j, \qquad u_j = \widehat{\tau}_j + r_j + b_j. \tag{18}$$

If these residual bounds are finite and observable, Theorem 7 gives simultaneous coverage of $[\ell_j, u_j]$. Sorting the upper endpoints outside $\mathcal{I}$ in descending order and the lower endpoints inside $\mathcal{I}$ in ascending order gives

$$C_{\text{eff}}^{\text{all}}(\mathcal{I}) = \max_{0 \le h \le \min(k, m-k)} \left\{ \sum_{r=1}^{h} u_{(r)}^{\text{out}} - \sum_{r=1}^{h} \ell_{[r]}^{\text{in}} \right\}. \tag{19}$$

The no-change term $h = 0$ is zero. Theorem 6 shows that $C_{\text{eff}}^{\text{all}}$ covers the same full allocation regret but uses a different, two-sided calibration and discards dependence across target effects.

Theorem 6 (Certificate from simultaneous treatment effect intervals)**.** *Under the conditions in Section 3 except Assumption 1, assume $G \succ 0$ and the existence of $\theta \in \mathbb{R}^p$, $u_{\text{tr}} \in \mathbb{R}^n$, and $u_B \in \mathbb{R}^m$ such that $\tau_{\text{tr}} = \Phi\theta + u_{\text{tr}}$ and $\tau_B = \Phi_B\theta + u_B$. If $(u_{\text{tr}}, u_B) \in \mathcal{U}$ and every b_j is finite and observable, then, conditional on $\mathcal{F}_0$, with probability at least $1-\delta$, $R_B^{\text{all}}(\mathbf{1}_{\mathcal{I}}) \le C_{\text{eff}}^{\text{all}}(\mathcal{I})$ simultaneously for all size-k sets $\mathcal{I}$, including any set selected using the same trial data. The effect-interval bound also satisfies*

$$C_{\text{eff}}^{\text{all}}(\mathcal{I}) = \sup_{t \in \prod_{j=1}^{m}[\ell_j, u_j]} \max_{a \in \mathcal{A}_k} (a - \mathbf{1}_{\mathcal{I}})^\top t,$$

and Equation (19) *costs $O(m \log m)$. For $m = 0$, the unique feasible allocation has regret and certificate zero and no interval radius is evaluated.*

The sorting formula maximizes every admissible exchange gain over the simultaneous effect intervals. Its proof follows the interval-coverage result below.

Treatment effect interval benchmark.

Theorem 7 (Simultaneous treatment effect intervals)**.** *Under the conditions in Section 3 except Assumption 1, assume $G \succ 0$ and the existence of $\theta \in \mathbb{R}^p$, $u_{\text{tr}} \in \mathbb{R}^n$, and $u_B \in \mathbb{R}^m$ such that $\tau_{\text{tr}} = \Phi\theta + u_{\text{tr}}$, $\tau_B = \Phi_B\theta + u_B$, and $(u_{\text{tr}}, u_B) \in \mathcal{U}$. If every b_j is finite and observable, then, conditional on $\mathcal{F}_0$,*

$$\mathbb{P}\{\ell_j \le \tau(x_j^B) \le u_j \text{ for every } j \in [m] \mid \mathcal{F}_0\} \ge 1 - \delta.$$

For exact specification, choosing $\mathcal{U} = \{(0,0)\}$ gives $b_j = 0$. For $m = 0$, the event is the sure event and no interval radius is evaluated.

We first prove Theorem 7. For target unit j, the least-squares identity gives

$$\widehat{\tau}_j - \tau(x_j^B) = (w_j^{\text{eff}})^\top \xi + (w_j^{\text{eff}})^\top u_{\text{tr}} - u_{B,j}, \qquad w_j^{\text{eff}} = \Phi G^{-1} \phi_j^B,$$

where $\xi_i = T_i - \mathbb{E}(T_i \mid \mathcal{F}_0)$. Conditional on $\mathcal{F}_0$, the $2m$ signed treatment effect coordinates are fixed. Applying the design-based choice of Hoeffding or Bernstein radius proved in Appendix B.2 to this signed family gives

$$\mathbb{P}\{|(w_j^{\text{eff}})^\top \xi| > r_j \mid \mathcal{F}_0\} \le \frac{\delta}{m}.$$

If $(u_{\text{tr}}, u_B) \in \mathcal{U}$, Equation (18) implies $|(w_j^{\text{eff}})^\top u_{\text{tr}} - u_{B,j}| \le b_j$. A union bound over the m two-sided events proves simultaneous inclusion. Zero-length directions satisfy the same statement by continuity, with a deterministic stochastic term of zero.

We next prove Theorem 6. Fix a size-k target-index set $\mathcal{I}$ and any size-k alternative allocation $\mathcal{J}$. Let $A = \mathcal{J} \setminus \mathcal{I}$ and $D = \mathcal{I} \setminus \mathcal{J}$. Then $|A| = |D| = h \le \min(k, m-k)$, and on the simultaneous interval event,

$$(\mathbf{1}_{\mathcal{J}} - \mathbf{1}_{\mathcal{I}})^\top \tau_B = \sum_{j \in A} \tau(x_j^B) - \sum_{j \in D} \tau(x_j^B) \le \sum_{j \in A} u_j - \sum_{j \in D} \ell_j.$$

For fixed h, this expression is maximized by choosing the h largest upper endpoints outside $\mathcal{I}$ and the h smallest lower endpoints inside $\mathcal{I}$. Maximizing over h, including the no-change case $h = 0$, gives Equation (19) and proves the regret inequality simultaneously for every $\mathcal{I}$. Any measurable data-selected $\widehat{\mathcal{I}}$ can therefore be substituted without an independence condition.

For the box-set equality, the previous display proves that the supremum is no larger than $C_{\text{eff}}^{\text{all}}(\mathcal{I})$. For a maximizing h, take $\mathcal{J}$ to add the selected outside indices and drop the selected inside indices.

Within the box set $t_j = u_j$ on added indices and $t_j = \ell_j$ on dropped indices. Other coordinates are arbitrary. This attains the displayed value. Two sorts and cumulative sums establish the $O(m \log m)$ complexity. Ties may be broken by target index and do not change the value.

For constants $b_{\mathrm{tr}}, b_B \geq 0$, define the independent box by $\|u_{\mathrm{tr}}\|_\infty \leq b_{\mathrm{tr}}$ and $\|u_B\|_\infty \leq b_B$. Its target support for $\sum_{j\in A} u_{B,j} - \sum_{j\in D} u_{B,j}$ is $2hb_B$, attained by opposite signs on the added and dropped coordinates. The source functional is $(\sum_{j\in A} w_j^{\mathrm{eff}} - \sum_{j\in D} w_j^{\mathrm{eff}})^\top u_{\mathrm{tr}}$, whose support under $\|u_{\mathrm{tr}}\|_\infty \leq b_{\mathrm{tr}}$ is $b_{\mathrm{tr}}\| \sum_{j\in A} w_j^{\mathrm{eff}} - \sum_{j\in D} w_j^{\mathrm{eff}}\|_1$. These are exact support functions for the named boxes. A common-function residual class can give smaller support.

B.6 Common confidence set and calibration

To isolate the effect of projecting the same confidence set, assume $G \succ 0$, fix $\beta \geq 0$, and give every construction the same ellipsoid $\mathcal{E}_\beta = \{\vartheta : \|\vartheta - \widehat{\theta}\|_G \leq \beta\}$. The pairwise difference projection for a candidate pair has radius $a_{rs}^{\mathrm{pair}} = \beta\|g_{rs}\|_{G^{-1}}$. Propagating simultaneous unit-effect intervals or allocation welfare intervals gives

$$a_{rs}^{\mathrm{eff}} = \beta \sum_j |d_{rs,j}| \|\phi(x_j^B)\|_{G^{-1}}, \qquad a_{rs}^{\mathrm{welf}} = \beta(\|h_r\|_{G^{-1}} + \|h_s\|_{G^{-1}}).$$

Proposition 8 (Radii from a common confidence ellipsoid)**.** *For every pair, $a_{rs}^{\mathrm{pair}} \leq a_{rs}^{\mathrm{eff}}$ and $a_{rs}^{\mathrm{pair}} \leq a_{rs}^{\mathrm{welf}}$. Maximizing the same joint ellipsoid directly gives exactly the pairwise difference radius. For every $0 < \rho \leq 1$ and $\beta > 0$, there is a one-swap construction with $G = I_2$ and target feature norms at most one for which*

$$\frac{a_{rs}^{\mathrm{eff}}}{a_{rs}^{\mathrm{pair}}} = \frac{1 + \sqrt{1-\rho^2}}{\rho}.$$

Thus the ratio diverges as $\rho \downarrow 0$.

Common treated features cancel in g_{rs} before projection. The treatment effect interval benchmark also cancels common members but discards dependence between the remaining individual effects. The allocation welfare interval benchmark can additionally lose cancellation of common members. The comparison uses one common region and concerns direct candidate contrasts. Full-allocation matching sums separate exchange supports and can lose dependence across exchanges, so this proposition does not order its certificate against Welfare.

Proof of Proposition 8. The support function of $\mathcal{E}_\beta$ in direction g is $g^\top\widehat{\theta} + \beta\,\|g\|_{G^{-1}}$. The triangle inequality yields

$$\left\|\sum_j d_{rs,j}\phi_j^B\right\|_{G^{-1}} \leq \sum_j |d_{rs,j}| \left\|\phi_j^B\right\|_{G^{-1}}, \qquad \|h_r - h_s\|_{G^{-1}} \leq \|h_r\|_{G^{-1}} + \|h_s\|_{G^{-1}}.$$

These prove Proposition 8. If the joint welfare image is $\{(h_r^\top\vartheta)_{r=1}^M : \vartheta \in \mathcal{E}_\beta\}$, its support for the coordinate difference $r - s$ is the same support function. Maximizing over finitely many alternative allocations commutes with taking the supremum over this set, establishing equality at the certificate level as well.

For the separation construction, take $G = I_2$, $0 < \rho \leq 1$, and

$$\phi(a) = (\sqrt{1-\rho^2}, \rho)^\top, \qquad \phi(b) = (\sqrt{1-\rho^2}, 0)^\top.$$

Let two exact-k candidates differ only by treating a in place of b. The same $k-1$ other treated individuals, if any, cancel from the welfare difference. Then $\|\phi(a) - \phi(b)\|_2 = \rho$, while $\|\phi(a)\|_2 + \|\phi(b)\|_2 = 1 + \sqrt{1-\rho^2}$. For $\beta > 0$, dividing the treatment effect and pairwise radii gives $(1 + \sqrt{1-\rho^2})/\rho$. All feature norms are at most one. At $\rho = 1$ the ratio equals one, so strict domination is not universal.

C Detailed experimental design

C.1 Data roles and repeated sampling

Every experiment with learned candidate allocations separates three data roles. The training sample D_{train} is used for preprocessing, feature-map construction, and candidate scoring rules. The target cohort B supplies covariates for the exact-k decision. The randomized trial supplies covariates, randomized treatment, and bounded observed outcomes for the candidate set certificate. Candidate allocations are fixed before trial outcomes are observed.

In the main evaluation, each outer replicate fixes D_{train}, the feature map, the target cohort, the candidate set, the trial covariates, and the treatment probabilities. We then generate 200 treatment assignments and outcome samples while holding these objects fixed. Five replicates describe variation across independently generated samples. The 200 trial realizations estimate coverage conditional on each replicate. Simulations use fixed random streams for training, treatment assignment, and outcomes. Within a replicate, the constructions share the training sample, target cohort, and trial draws. Sample sizes use nested prefixes of the same 6,000-observation trial path.

For simulated binary outcomes, independent uniforms generate

$$A_i = \mathbf{1}\{U_i^A \le e_i\}, \qquad Y_i = \mathbf{1}\{U_i^Y \le \mu_{A_i}(X_i)\}.$$

The conditional means in all settings lie strictly between zero and one.

C.2 Synthetic data-generating settings

Training, target, and trial covariates are independently sampled from $X \sim \mathrm{Unif}([-1,1]^{10})$. The two feature maps are

$$\phi_{\text{raw}}(X) = (X_1, X_2, X_4)^\top, \qquad \phi_{\text{nl}}(X) = (|X_1| - 1/2, X_2X_3, X_4)^\top.$$

For each synthetic setting, ϕ denotes the corresponding feature map. For the main coefficient $\theta = (0.10, 0.10, 0.10)^\top$,

$$\tau(X) = \phi(X)^\top\theta, \quad \mu_0(X) = 0.5 + 0.05\{\sin(\pi X_5) + X_6X_7\}, \quad \mu_1(X) = \mu_0(X) + \tau(X),$$

with treatment probability $e(X) = 1/2$. The Raw-feature setting uses raw covariates in ϕ, while the Nonlinear-feature setting uses nonlinear transformations and an interaction. In both settings, Assumption 1 is exactly specified. The baseline outcome regression is unrestricted by the feature map.

C.3 SUPPORT2 settings

We use age, number of comorbidities, hospital day, diabetes, dementia, sex, disease class, and cancer status from SUPPORT2 (Asuncion et al., 2007). The first three variables are standardized using the 2,000 training observations. Binary variables are retained and categorical variables are one-hot encoded. A principal component analysis fitted on the training data supplies three signed directions. Coordinatewise application of $\tanh$ gives $\phi(x) \in [-1,1]^3$.

The setting with an exact treatment effect specification uses

$$\tau(x) = \phi(x)^\top\theta, \qquad \mu_0(x) = 0.5, \qquad \mu_1(x) = 0.5 + \tau(x), \qquad e(x) = 0.5.$$

Treatment effects and outcomes in this SUPPORT2 setting are simulated.

C.4 Candidate set

All candidate methods use the same training observations. No method uses counterfactual outcomes, the true conditional average treatment effect, or any trial outcome used for inference. A score is converted to an allocation by treating its top k target cohort members with deterministic tie breaking. Duplicate allocations are removed before certification.

The main candidate set uses 19 scores. Three generalized random forests (Athey et al., 2019) fit conditional treatment effects, three T-learners (Künzel et al., 2019) take differences of separately fitted

outcome regressions, and three doubly robust learners regress a cross-fitted doubly robust pseudo-outcome. Three ranking learners use cross-fitted doubly robust pseudo-outcomes and two ranking learners use plug-in treatment effect estimates in pairwise ranking objectives. Five simple scores use three feature-map coordinates, an inverse probability weighted linear fit, or no fitted model. All scores use training observations and target covariates. The ranking learners share nuisance estimates within each training replicate.

C.5 Evaluation grid

The main comparisons use $n = 6000, m = 500, k = 20, p = 3$ and $\delta = 0.05$. The sample-size curves use nested trial prefixes $n \in \{500, 1000, 2000, 6000\}$. Each synthetic or SUPPORT2 point averages five outer replicates with 200 conditional trial draws per replicate. The Hillstrom results use three fixed splits, one observed outcome vector per split. The treatment effect approximation error experiment fixes a true residual coefficient of 0.2 and varies its declared bound over $\{0, 0.1, 0.2, 0.4\}$.

Additional displays use a 12-person cohort with capacities $k = 1, \ldots, 11$ and heterogeneous assignment probabilities with minimum $\eta \in \{0.10, 0.25, 0.50\}$ at $n = 2000$ or 6000.

C.6 Hillstrom application

The application retains the Mens E-Mail and No E-Mail arms of Hillstrom (Hillstrom, 2008). Visit and conversion are analyzed separately. Covariates include recency, purchase history, previous men's and women's purchases, new customer status, history segment, zip code, and channel. Training-only standardization, one-hot encoding, three principal directions, and coordinatewise $\tanh$ match the SUPPORT2 preprocessing. In the main application, each outcome uses three fixed splits. Within each split, the 2,000 training records, 500 target cohort records, and up to 6,000 certification records are disjoint.

The two-arm treatment probability is set to $1/2$. The finite-sample interpretation requires independent unit-level randomization. Since target conditional effects are not observed, the application reports only selected allocations, candidate-set and full-allocation certificates, and threshold decisions. For the assumed pointwise treatment effect residual bound $b \in \{0, 0.01, 0.02, 0.05, 0.10\}$, the correction is

$$b_{rs}^{\mathrm{box}} = b\{\|z_r - z_s\|_1 + \|w_{rs}\|_1\}.$$

The assumed residual bound b is a sensitivity parameter.

C.7 Full-allocation certificate protocol

For full allocation regret, the pairwise difference construction covers every ordered target-unit exchange before the selected allocation is known, then takes a maximum-weight partial matching. The treatment effect interval benchmark covers both signs of each target unit's effect and sorts the resulting endpoints. The common-ellipsoid comparisons additionally compute the allocation welfare interval benchmark for the same allocation. Sample sizes use nested prefixes. Five outer replicates and 200 conditional trials per replicate determine each synthetic or SUPPORT2 mean.

The Hillstrom full-allocation sensitivity analysis uses three fixed splits for each observed outcome and varies an assumed residual limit over $\{0, 0.01, 0.02, 0.05, 0.10\}$. It reports bounds and threshold decisions. True allocation regret is unobserved because target conditional treatment effects are unavailable.

C.8 Capacity and calibration protocol

The capacity study fixes $m = 12, n = 6000$, varies $k = 1, \ldots, 11$, and uses the same estimated-effect allocation and trial for all three constructions at each capacity. In the common-ellipsoid display, Pairwise uses matching, Effect uses treatment effect intervals, and Welfare is computed by enumeration. The common ellipsoid and method-specific Hoeffding comparisons use five outer replicates with 200 conditional trials each. The large-cohort calibration study fixes $m = 500, k = 20$, uses $n \in \{2000, 6000\}$ and $\eta \in \{0.10, 0.25, 0.50\}$. It selects the k largest trial-estimated effects

from all m cohort members. The six construction-radius combinations then evaluate this same allocation. This differs from selection within the main experiment's prespecified candidate set.

C.9 Residual sensitivity protocol

The synthetic residual diagnostics use ϕ_{raw}, $\theta = (0.05, -0.04, 0.03)^\top$, and balanced randomization. Draw the raw covariates as in the main synthetic experiment, then set $X_5 = \lambda_{\text{tr}} X_1$ in training and trial samples and $X_5 = \lambda_B X_1$ in the target cohort. Trial-side, target-side, partial-cancellation, and reinforcement mechanisms use $(\lambda_{\text{tr}}, \lambda_B) = (1, 0), (0, 1), (1, 0.7), (1, -1)$, respectively. Set $\tau(x) = \phi_{\text{raw}}(x)^\top \theta + 0.2x_5$, $\mu_0(x) = 0.5 - \tau(x)/2$, and $\mu_1(x) = 0.5 + \tau(x)/2$. The five candidate allocations treat the top k members ranked by $\phi_1, -\phi_1, \phi_2, \phi_3, \phi_1 + \phi_2$, using the coordinates of ϕ_{raw}. Within each trial, estimated welfare selects one candidate, which is held fixed across all residual classes and declared limits.

Figure 5 uses the trial-side and target-side mechanisms. For the declared class cx_5, $|c| \le b_{\text{coef}}$, the assumed coefficient limit varies while observations and allocations stay fixed. Target-only correction bounds only the target term in Equation (7). Correcting both populations bounds their difference, while no error adjustment sets the residual bound to zero.

Table 2 compares the three constructions under four residual mechanisms: trial-side, target-side, partial cancellation, and reinforcement, with five candidate allocations and $n \in \{512, 2000, 6000\}$. For each mechanism, let ψ_{tr} and ψ_B be its trial and target X_5 residual directions, and set $w_d = \Phi G^{-1} \Phi_B^\top d$. The shared-coefficient class contains pairs $(c\psi_{\text{tr}}, c\psi_B)$ with $|c| \le b_{\text{coef}}$. Its exact allowance for a contrast d is $b_{\text{coef}} |d^\top \psi_B - w_d^\top \psi_{\text{tr}}|$. The independent-box class instead assumes $\|u_{\text{tr}}\|_\infty, \|u_B\|_\infty \le b_{\text{coef}}$, giving $b_{\text{coef}}(\|d\|_1 + \|w_d\|_1)$. Each class is supplied unchanged to Pairwise, Effect, and Welfare, with their method-specific simultaneous Hoeffding radii from Appendix A.2. The box-class full-allocation Welfare computation uses $\| \sum_{j \in A} w_j^{\text{eff}} \|_1 \le \sum_{j \in A} \|w_j^{\text{eff}}\|_1$ for its numerical upper bound. At $b_{\text{coef}} \ge 0.2$, either class contains the true residual pair in these four mechanisms. Smaller declared limits are sensitivity diagnostics. The trial draws, fitted estimates, and selected allocations remain fixed across limits and classes.

For the observed Hillstrom outcomes, the both-population class is $\|u_{\text{tr}}\|_\infty, \|u_B\|_\infty \le b_{\text{res}}$. Target-only correction omits the trial term. The main sensitivity curves vary b_{res} from zero to 0.01. These are assumed limits, not estimates of unobserved model error. These curves use the simultaneous pairwise comparison radii. Table 1 uses the common ellipsoid.

C.10 Evaluation metrics and aggregation

For known treatment effects, false certification is $C \le \epsilon < R$. Hillstrom has no observed true regret, so coverage and false certification cannot be calculated. Under exact-k allocation, ϵ/k is the allowed welfare loss per filled treatment slot.

We calculate means, rates, and 95th percentiles over trial draws within each outer replicate, then average the five replicates equally. Each replicate has a separate cohort. Hillstrom averages three fixed splits.

D Additional certification results

Unless noted, methods evaluate the same allocation within a trial at 95% confidence.

D.1 Treatment effect approximation error sensitivity

Table 2 compares two residual classes that contain the true residual across four synthetic mechanisms.

Table 2: **Certificates under two valid residual classes.** Mean $\bar{C}$, coverage, and certification rate at $n = 6000$, $b_{\text{coef}} = 0.2$, with method-specific 95% Hoeffding radii. Shared uses one residual coefficient. Box allows independent trial and target deviations. Full-allocation Welfare under the box uses the numerical relaxation in Appendix C.9. Bold marks the smallest mean certificate within each mechanism, class, and scope.

			Candidate set			Full allocation		
Mechanism	Class	Method	$\bar{C}$	$\widehat{\text{Cov}}$	$\widehat{\text{Cert}}(5)$	$\bar{C}$	$\widehat{\text{Cov}}$	$\widehat{\text{Cert}}(5)$
Trial side	Shared	Pairwise	**1.984**	100.0	100.0	**5.469**	100.0	24.5
		Effect	7.523	100.0	0.0	12.087	100.0	0.0
		Welfare	8.523	100.0	0.0	20.325	100.0	0.0
	Box	Pairwise	**14.111**	100.0	0.0	**22.844**	100.0	0.0
		Effect	22.024	100.0	0.0	27.424	100.0	0.0
		Welfare	21.797	100.0	0.0	34.370	100.0	0.0
Target side	Shared	Pairwise	**7.795**	99.8	1.1	**11.521**	100.0	0.0
		Effect	10.214	100.0	0.0	11.945	100.0	0.0
		Welfare	8.425	99.9	0.1	19.145	100.0	0.0
	Box	Pairwise	**19.799**	100.0	0.0	**26.982**	100.0	0.0
		Effect	25.950	100.0	0.0	27.816	100.0	0.0
		Welfare	22.247	100.0	0.0	33.995	100.0	0.0
Partial cancellation	Shared	Pairwise	**0.224**	100.0	100.0	**4.209**	100.0	98.8
		Effect	3.798	100.0	99.9	7.065	100.0	0.0
		Welfare	3.809	100.0	99.9	14.816	100.0	0.0
	Box	Pairwise	**13.630**	100.0	0.0	**22.785**	100.0	0.0
		Effect	22.236	100.0	0.0	27.463	100.0	0.0
		Welfare	21.980	100.0	0.0	33.751	100.0	0.0
Reinforcement	Shared	Pairwise	**8.760**	99.9	0.0	**12.304**	100.0	0.0
		Effect	12.479	100.0	0.0	19.554	100.0	0.0
		Welfare	15.733	100.0	0.0	27.700	100.0	0.0
	Box	Pairwise	**13.915**	100.0	0.0	**23.059**	100.0	0.0
		Effect	22.621	100.0	0.0	27.770	100.0	0.0
		Welfare	21.971	100.0	0.0	34.282	100.0	0.0

At $n = 6000$, mean candidate-set and full-allocation true regrets are $(0.015, 0.780)$, $(1.262, 1.595)$, $(0, 0.389)$, and $(5.817, 6.233)$ for trial-side, target-side, partial-cancellation, and reinforcement mechanisms. The shared-coefficient class retains cancellation and yields smaller certificates. The independent box gives no certification at tolerance five. The smallest coverage averaged across the five designs in Table 2 is 99.8%. Under reinforcement, mean true candidate regret is 5.817 and no method certifies at tolerance five. This contrasts with partial cancellation, where small regret can be certified under the shared-coefficient class.

Figure 5 varies declared trial and target approximation error limits for fixed models and allocations. True regret is observed only in the synthetic settings.

D.2 Rank-deficient trial demonstration

This synthetic example uses balanced randomization and the Raw-feature data-generating process. The trial has $x_4 = 0$, while every target member has $x_4 = 1$, so $\phi(x) = (x_1, x_2, x_4)$ has trial rank two. Two size-20 allocations in a 500-person cohort share 19 members. The exchanged members have features $(1/2, 0, 1)$ and $(-1/2, 0, 1)$, giving $g_{rs} = (1, 0, 0)$ and an identifiable welfare difference of 0.10 for every unobserved coefficient $\gamma \in [-0.10, 0.10]$.

Figure 4 uses four trial sizes from 500 to 6000, five outer replicates, and 200 conditional trials per replicate. The first candidate has welfare 0.10 above the second. The selected candidate therefore has regret 0 when the first is selected and 0.10 when the second is selected.



Figure 5: **Candidate-set bounds under residual assumptions.** (a,b) Synthetic trial-only and target-only errors, with true regret shown by the dark dash-dotted line on logarithmic vertical scales. (c,d) Hillstrom-V/C, with linear scales starting at 2.4. Solid, dashed, and dotted certificate curves correct both trial and target errors, only target error, or neither. Allocations stay fixed.

Table 3: **Rank-deficient results by fixed cohort design at $n = 6000$.** Mean true regret and the three mean certificate values are shown directly. $\widehat{\mathrm{Cert}}(0.05)$ is the Pairwise certification rate. Effect and Welfare certify 0% in every design.

Design	Regret $\bar{R}$	Pairwise $\bar{C}$	Effect $\bar{C}$	Welfare $\bar{C}$	Pairwise $\widehat{\mathrm{Cert}}(0.05)$
1	0.000	**0.006**	0.193	2.000	99.5%
2	0.000	**0.005**	0.193	2.000	100.0%
3	0.000	**0.007**	0.195	2.000	98.0%
4	0.000	**0.006**	0.193	2.000	100.0%
5	0.000	**0.006**	0.195	2.000	99.0%
Mean	0.000	**0.006**	0.194	2.000	99.3%

All three bounds use the same rank-two confidence region for the identified coefficients α and the same prespecified range $|\gamma| \le \kappa = 0.10$ for the unidentified coefficient. Let $\Phi_{1:2}$ contain the first two trial feature columns, $H = \Phi_{1:2}^\top \Phi_{1:2}$, $\beta = \sqrt{4\log(4/\delta)}$, and $a = \beta\sqrt{(1,0)H^{-1}(1,0)^\top}$. With $\widehat{\Delta}$ the estimated advantage of the first candidate, the uncapped Pairwise and Effect bounds are

$$(a - |\widehat{\Delta}|)_+, \qquad (a + 2\kappa - |\widehat{\Delta}|)_+.$$

Welfare uses the upper endpoint for one complete allocation and the lower endpoint for the other. Its uncertainty radius is the sum of their two identified-coefficient radii plus $2k\kappa$, before subtracting $|\widehat{\Delta}|$. Every method is then capped at 2, the known maximum regret for a one-person exchange when effects lie in $[-1, 1]$. At $n = 6000$, Welfare's uncapped mean is 4.598, while its displayed capped mean is 2.000. The Effect mean is 0.194 and Pairwise mean is 0.006.

If the exchanged second member instead has $x_4 = -1$, then $g_{rs} = (1, 0, 2)$ does not lie in $\mathrm{Range}(\Phi^\top)$. The same trial law allows a welfare difference of -0.10 or 0.30 at the two endpoints of the γ range. This is the nonidentified case in Proposition 1.

D.3 Why center the IPW score?

The outcome scale $[0, 1]$ fixes its midpoint at $\zeta = 0.5$ before trial outcomes are observed. The IPW score has the same conditional treatment effect mean for every fixed ζ. Appendix B.2 shows that $\zeta = 0.5$ minimizes the score's worst-case support length. At balanced assignment, that length is 2 at $\zeta = 0.5$ and 4 at $\zeta = 0$, so the Hoeffding radius for a fixed welfare difference is halved.

Table 4 recomputes both bounds for exactly the same allocation in every trial. The allocation is obtained once from the midpoint welfare estimates and then held fixed. Each center changes both the estimated welfare differences and their calibrated radii. The reported reductions therefore describe the complete certificates.

Table 4: **Midpoint centering for a common allocation.** Mean Pairwise candidate set regret upper bounds at $n = 2000, m = 500, k = 20$. $\zeta = 0.5$ centers the score at the outcome-range midpoint. $\zeta = 0$ leaves it uncentered. Reduction is one minus the ratio of means. Smaller certificates and larger reductions are preferable. Bold marks the smaller certificate in each setting.

Data type	Setting	$\bar{C}(\zeta = 0.5)$	$\bar{C}(\zeta = 0)$	Reduction
Synthetic	Raw-feature	**2.154**	6.123	64.8%
	Nonlinear-feature	**3.675**	8.889	58.7%
Semi-synthetic	SUPPORT2	**3.017**	8.449	64.3%

D.4 Paired comparison of certificate constructions

Table 1 uses one 95% ellipsoid to compare the three constructions for each selected allocation. Here we compare method-specific calibration and describe the numerical evaluation of Welfare over all feasible allocations.

Method-specific simultaneous calibration. The construction and query counts are in Appendix A.2. Table 5 evaluates the same selected allocations under those radii. Empirical coverage is 100% for every method, scope, setting, and trial size in the 12,000 trial realizations. Table 6 reports certification rates at $n = 6000$.

Table 5: **Separately calibrated certificates.** Mean $\bar{C}$ at $n = 6000$ for the same selected allocations, with a 95% simultaneous Hoeffding bound for each construction.

Data type	Setting	Scope	Pairwise	Effect	Welfare
Synthetic	Raw-feature	Candidate set	**0.550**	2.654	2.747
		Full allocation	**2.693**	4.310	13.784
	Nonlinear-feature	Candidate set	**1.848**	4.007	3.160
		Full allocation	**4.323**	5.285	14.794
Semi-synthetic	SUPPORT2	Candidate set	**1.004**	3.075	2.810
		Full allocation	**3.049**	4.198	13.165

Table 6: **Certification rates with method-specific Hoeffding radii.** Values are $\widehat{\mathrm{Cert}}(\epsilon)$ at $n = 6000$, in percent. Bold marks a unique highest nonzero rate within each setting, scope, and tolerance.

		$\epsilon = 2$			$\epsilon = 5$		
Setting	Scope	Pairwise	Effect	Welfare	Pairwise	Effect	Welfare
Raw-feature	Candidate set	**97.7**	8.5	15.1	100.0	100.0	100.0
	Full allocation	**10.3**	0.0	0.0	**100.0**	83.1	0.0
Nonlinear-feature	Candidate set	**54.8**	0.3	0.5	100.0	97.5	100.0
	Full allocation	0.0	0.0	0.0	**77.4**	38.2	0.0
SUPPORT2	Candidate set	**95.3**	2.2	0.0	100.0	100.0	100.0
	Full allocation	**5.7**	0.0	0.0	**99.5**	91.5	0.0

No false certifications $C \leq \epsilon < R$ occur at the reported tolerances in these exact-model trials. Table 6 also shows the remaining conservatism: at $\epsilon = 5$, full-allocation Pairwise certifies 77.4% of Nonlinear-feature trials, while Welfare certifies none in any setting.

Certificate from allocation welfare intervals. Write $h_A = \sum_{i \in A} \phi_i^B$, where A is a size-k allocation, and let S be the selected allocation. With ellipsoid radius β, define

$$U_A = h_A^\top \widehat{\theta} + \beta \|h_A\|_{G^{-1}}, \qquad L_S = h_S^\top \widehat{\theta} - \beta \|h_S\|_{G^{-1}}.$$

The allocation welfare interval construction is $\max\{0, \max_{A \neq S}(U_A - L_S)\}$. The inner maximum ranges over the candidate allocations or over all exact-k allocations, as indicated in the table. Excluding S makes its self-comparison exactly zero.

Numerical evaluation. Set $\widehat{\tau}_i = (\phi_i^B)^\top \widehat{\theta}$ and choose vectors v_i with $v_i^\top v_j = \beta^2 (\phi_i^B)^\top G^{-1} \phi_j^B$. Maximizing U_A equals maximizing $F(u) = \max_{A \neq S, |A|=k} \sum_{i \in A} (\widehat{\tau}_i + v_i^\top u)$ over $\|u\|_2 = 1$. At a fixed u, sorting gives the best feasible allocation. If it is S, one best single exchange gives the second-best allocation. For three features we partition the unit sphere into spherical triangles. If a triangle's flat face is at distance d from the origin and $L_{\mathrm{sph}} \geq \max_{|A|=k} \| \sum_{i \in A} v_i \|_2$, convexity bounds its maximum by the largest vertex value plus $L_{\mathrm{sph}}(1 - d)$. Subdivision retains lower and upper endpoints until their difference is at most 10^{-6}, up to floating-point error. The largest recorded gap in the 12,000-trial sample-size analysis is 1.01×10^{-6}. We report upper endpoints.

D.5 Treatment capacity and uncertainty propagation

For a 12-person cohort, we use the same trial realizations at $n = 6000$, vary $k = 1, \ldots, 11$, and evaluate all three bounds on the same estimated-effect top-k allocation at each capacity.

One common confidence region. Figure 6 uses one confidence ellipsoid for every construction. Enumeration computes the Welfare benchmark, Pairwise aggregates exchange bounds, and Effect propagates target treatment effect intervals.

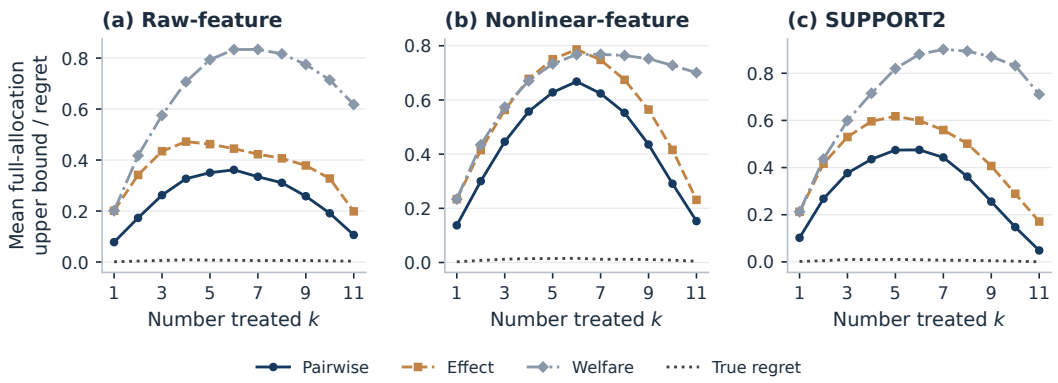

Figure 6: **Full-allocation certification as treatment capacity changes.** All three constructions use the same confidence ellipsoid and allocation. The dotted reference is true regret for $m = 12, n = 6000$.

Separate confidence radii. Figure 7 instead gives each method a Hoeffding radius for its own comparison family: $m(m-1)$ ordered exchanges for Pairwise, both signs of m individual treatment effects for Effect, and both signs of $\binom{m}{k}$ allocation welfares for Welfare. Each family receives the same total failure probability δ, so the bounds need not preserve the ordering obtained under a common confidence region.

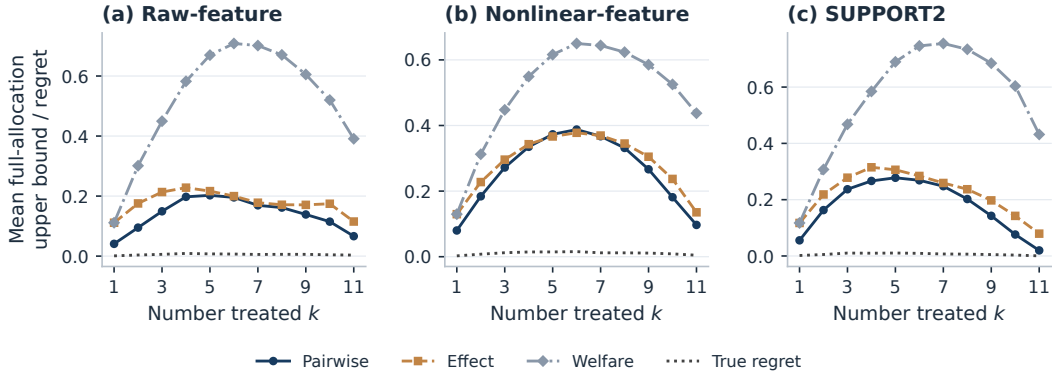

Figure 7: **Full-allocation bounds with method-specific Hoeffding radii.** The allocation and trials match Figure 6. Only the uncertainty construction changes. The dotted curve is the same true regret for all three constructions.

D.6 Hoeffding and Bernstein calibration

Hoeffding uses the bounded range of the IPW score. Bernstein also uses a variance bound (Appendix A.1). Table 7 compares their full-allocation certificates for the same selected allocation. Both radii depend only on known features and treatment probabilities.

Table 7: **Full-allocation certificates under Hoeffding and Bernstein calibration.** Mean raw regret upper bounds for $m = 500, k = 20$, before the known 40-unit range cap. $e_i = \Pr(A_i = 1 \mid X_i)$ is the trial treatment probability. In heterogeneous designs, $e_i = 0.50$ when the first covariate is nonnegative and $e_i = \eta$ otherwise. H and B denote Hoeffding and Bernstein. Lower is better. Bold marks the smallest of the six means in each row.

			Pairwise		Effect		Welfare	
Setting	n	Trial e_i	H	B	H	B	H	B
Raw-feature	2000	0.50	**5.375**	6.228	7.375	8.153	23.463	30.614
Raw-feature	2000	0.50/0.10	31.784	**14.663**	30.524	14.993	85.508	75.910
Raw-feature	2000	0.50/0.25	10.609	**7.946**	12.426	9.658	37.297	40.525
Raw-feature	6000	0.50	**2.133**	2.352	3.782	4.033	13.586	15.947
Raw-feature	6000	0.50/0.10	15.607	**5.174**	16.750	6.977	49.009	32.162
Raw-feature	6000	0.50/0.25	4.597	**2.920**	6.598	4.679	21.506	19.578
Nonlinear-feature	2000	0.50	**7.262**	8.664	8.292	9.302	25.343	34.633
Nonlinear-feature	2000	0.50/0.10	36.592	19.716	33.573	**17.484**	92.133	91.079
Nonlinear-feature	2000	0.50/0.25	13.490	**11.074**	13.802	11.103	40.198	46.949
Nonlinear-feature	6000	0.50	**3.248**	3.640	4.337	4.660	14.638	17.642
Nonlinear-feature	6000	0.50/0.10	19.277	**7.654**	18.659	8.103	53.439	37.939
Nonlinear-feature	6000	0.50/0.25	6.404	**4.469**	7.437	5.417	23.279	22.191

At $\eta = 0.10$, Bernstein reduces the mean bound for all three constructions. When every $e_i = 0.50$, Hoeffding is smaller throughout. Welfare remains the widest because it propagates separately calibrated intervals for complete allocation welfares. All six construction-radius pairs cover true regret in all 12,000 trial realizations.

D.7 Hillstrom full-allocation sensitivity

For full allocation certification, keep each Hillstrom split's empirical-welfare allocation, fitted coefficients, and sampling radii fixed, and vary only the residual class

$$\mathcal{U}(b_{\text{res}}) = \{(u_{\text{tr}}, u_B) : \|u_{\text{tr}}\|_\infty, \|u_B\|_\infty \le b_{\text{res}}\}.$$

The exchange bias bound is $b_{\text{res}}(2 + \|w_{ji}\|_1)$, so the resulting $C^{\text{all}}_{\text{pair}}$ is nondecreasing in b_{res}.

With method-specific simultaneous sampling radii and tolerance $\epsilon = 5$, only one of three fixed splits for each outcome is certified at $b_{\text{res}} = 0$, and that split remains certified up to $b_{\text{res}} = 0.00659$ for Hillstrom-V and 0.00614 for Hillstrom-C. These limits describe sensitivity to the assumed residual class. True regret and residual coverage are unobserved in Hillstrom.

E Optional downstream use: selecting an allocation

The simultaneous guarantee also permits choosing the allocation with the smallest candidate set certificate. This is optional: the certification method accepts an allocation supplied by any covered data-dependent rule. On the validity event, such a choice obeys the oracle inequality

$$R_B(\widehat{s}_C) \le C_{\text{pair}}(\widehat{s}_C) \le \min_s \{R_B(s) + 2 \max_r a_{rs}\}, \qquad \widehat{s}_C \in \arg\min_s C_{\text{pair}}(s). \tag{20}$$

With treatment effect model approximation error, replace each a_{rs} by $a_{rs} + b_{rs}$.

F Detailed related work

Post-selection certification. Certified Decisions studies an observable high-probability bound on the loss of an action selected from data (Andrews & Chen, 2025). We instantiate this general objective for cohort-specific allocation regret, with the trial-to-target identification conditions and exchange aggregation made explicit. Simultaneous inference protects a reported quantity selected using the same observations (Berk et al., 2013; Benjamini et al., 2019). Policy Learning with Confidence maximizes a simultaneous welfare lower bound computed from policy-value estimates and

their joint uncertainty. Its reporting guarantee includes Gaussian and bootstrap approximation errors (Chernozhukov et al., 2025). Confidence-aware off-policy evaluation uses logged actions and outcomes to support policy choice (Kuzborskij et al., 2021). Our bounds instead use known trial treatment probabilities and bounded outcomes to control conditional finite-sample error.

Capacity-constrained allocation. Budgeted policy learning and treatment ranking choose whom to treat under resource limits (Pellatt, 2022; Kamran et al., 2024). Good Allocations studies choosing a fixed number of population groups using an interface that estimates each group's treatment effect to a requested precision (Casacuberta & Hardt, 2026). It gives sample-complexity guarantees for near-optimal selection and an instance-dependent certification condition. Our target is additive regret for one observed individual cohort, inferred from a separate randomized trial through a specified treatment effect model. Related work evaluates randomized allocation policies or designs sequential trials (Mate et al., 2023; Cortes-Gomez et al., 2025). Those objectives differ from bounding the regret of an already selected cohort allocation.

Trial-to-target comparisons. Population-policy generalization, distribution-shift bounds, and regularized causal extrapolation address different trial-to-target links (Hatt et al., 2022; Asiaee et al., 2026; Arbour et al., 2026). Groupwise transfer decisions (Opocher, 2026) also differ from an exact-k cohort allocation. Pairwise off-policy evaluation estimates policy-value differences from logged contextual-bandit data and retains covariance between their value estimates (Jeunen & Ustimenko, 2024). Work on structured action sets also uses comparison directions (Fiez et al., 2019; Maiti et al., 2026). Our construction makes the trial-to-target treatment effect model and its approximation error explicit and computes full allocation bounds by matching.