DESIGNING EXPERIMENTAL SELECTION CORRECTION WITH MULTIPLE SECONDARY OUTCOMES

Anonymous authors
Paper under double-blind review

ABSTRACT

Researchers increasingly have access to observational data containing primary and multiple secondary outcomes, together with randomized experiments in which only secondary outcomes are measured. Experimental selection correction (ESC) leverages experimental evidence to correct selection bias in observational estimation of the primary average treatment effect (ATE). Existing ESC methods, however, typically treat the set of measured secondary outcomes and the treatment-assignment rule as fixed, resulting in potential efficiency loss. We address this limitation by developing a novel experimental design framework that jointly optimizes secondary-outcome selection and treatment assignment under resource constraints, accounting for the tradeoff between the informativeness and measurement costs of secondary outcomes. We characterize the oracle design that minimizes the semiparametric efficiency bound for estimation of the ATE. To learn the design, we propose EFFICIENTESC, a two-stage procedure that uses observational data and a randomized pilot to determine the design for the main experiment. We show that the learned design converges to the oracle design, and the proposed ATE estimator is asymptotically normal and attains the minimized semiparametric efficiency bound. Extensive simulations demonstrate substantial efficiency gains over competing methods in finite samples.

1 INTRODUCTION

In causal inference, researchers increasingly have access to both observational datasets and randomized experiments. Large observational datasets often contain information on treatment assignments and primary outcomes of interest, but nonrandom treatment assignment can introduce selection bias into estimates of causal effects. Randomized experiments can avoid this bias (Rubin, 1980), yet the primary outcomes of interest may take years to materialize, be expensive to collect, or become unavailable through attrition (Hogan et al., 2004; Shi et al., 2024). Here is a motivating example.

Example 1 (Class size on test scores and high-school graduation)**.** The application study in Athey et al. (2025) combines data from the randomized Tennessee Project STAR experiment (Krueger, 1999) with observational administrative records from New York City public schools. The treatment is assignment to a small class versus a regular class, and the primary outcome is high-school graduation. Both datasets contain multiple post-treatment test scores, which provide a natural set of candidate secondary outcomes (see, Biasi et al., 2025). Pre-treatment covariates include gender, race, and school-level characteristics. This setting motivates the practical need to jointly determine which secondary outcomes to measure and how to allocate treatment under limited resources.

Experimental Selection Correction (ESC, Athey et al., 2025) uses experimental information on secondary outcomes to correct observational selection bias under latent unconfoundedness. Related work characterizes semiparametric efficiency bounds and constructs efficient estimators (Chen & Ritzwoller, 2023), while alternative identification strategies exploit sequential short-term outcomes to accommodate persistent unmeasured confounding (Imbens et al., 2025). These studies focus on identification and estimation with given secondary outcomes and fixed experimental treatment assignment, without addressing which secondary outcomes to measure or how to assign treatment. Recently, Epanomeritakis & Viviano (2025) jointly choose which experiments to conduct, and how to allocate samples using a minimax proportional-regret criterion. This criterion balances estimation variance against unknown bias in the observational sample. However, their work does not aim

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.

to minimize the variance of the treatment effect estimator, as is standard in the classical response-adaptive design literature (e.g., Hahn et al., 2011; Hu & Rosenberger, 2006; Zhang et al., 2007), and therefore does not attain semiparametric efficiency.

To fill this gap, our objective is instead to jointly select the secondary-outcome set and treatment-assignment rule to minimize the semiparametric efficiency bound of the ATE estimator under a budget constraint. The two design choices are intrinsically coupled through their effects on statistical efficiency. In particular, secondary-outcome selection entails a cost-efficiency tradeoff, where collecting a richer set of secondary outcomes can reduce estimation variance but incurs higher costs, thereby reducing the number of units that can be recruited under a fixed budget. Meanwhile, the treatment-assignment rule directly affects the information contributed by each treatment arm. To achieve this objective, we propose a novel experimental design method, termed EFFICIENTESC, that bridges observational causal inference and adaptive experimental design.

The main contributions of this paper are as follows. *First*, we develop a novel experimental design framework for jointly optimizing secondary-outcome selection and treatment assignment under an expected-budget constraint. We characterize the oracle design by minimizing the semiparametric efficiency bound of the ATE estimator over both design choices.

Second, we propose EFFICIENTESC, a two-stage adaptive experiment that learns the oracle design from observational data and a randomized pilot. We further construct an ATE estimator that combines observational and experimental data for efficient experimental selection correction.

Third, we establish theoretical guarantees for the proposed method. We show that EFFICIENTESC asymptotically recovers the oracle design, and the proposed estimator is asymptotically normal and attains the minimized semiparametric efficiency bound, enabling asymptotically valid inference.

Notation. We write $[K] = \{1, \dots, K\}$. For real numbers a, b, let $a \vee b = \max\{a, b\}$ and $a \wedge b = \min\{a, b\}$. Let $\mathbb{1}\{\cdot\}$ denote the indicator function. For a finite set I, $|I|$ denotes its cardinality. $\lfloor \cdot \rfloor$ denotes the floor function. For a probability distribution P and a measurable function f, write $\|f\|_{P,q} = \{\mathbb{E}_P(|f|^q)\}^{1/q}$ and $L_q(P) = \{f : \|f\|_{P,q} < \infty\}$ for $q \geq 1$, and $\|f\|_\infty = \sup_x |f(x)|$. P-a.s. means almost surely under P. $P \ll Q$ means that P is absolutely continuous with respect to Q. $U \perp\!\!\!\perp V$ denotes random variables U and V are independent. We use $\to_p$ and $\to_d$ for convergence in probability and in distribution, respectively. For a positive deterministic sequence a_n, $X_n = o_p(a_n)$ means that X_n/a_n converges to zero in probability, and $X_n = O_p(a_n)$ means that X_n/a_n is bounded in probability. For positive sequences a_n, b_n, $a_n \asymp b_n$ means that their ratio is bounded above and bounded away from zero.

2 RELATED WORK

Combining experimental and observational data. The surrogate index (Athey et al., 2026) uses secondary outcomes to predict the primary outcome and estimates the experimental effect on that prediction. ESC instead uses experimental information to correct selection bias in observational estimates (Athey et al., 2025). Semiparametric theory characterizes efficiency bounds and provides efficient estimators for combining these data (Chen & Ritzwoller, 2023). Alternative identification strategies exploit the sequential structure of multiple secondary outcomes (Imbens et al., 2025), equiconfounding or proxy restrictions (Ghassami et al., 2022), and historical experiments with imperfect surrogates (Bibaut et al., 2023). A related but distinct setting concerns primary outcomes that are only partially observed. Here, surrogate outcomes can improve estimation of average treatment effects (Kallus & Mao, 2025) and continuous treatment effects (Zeng et al., 2024). For a broader review of methods combining randomized and observational evidence, see Colnet et al. (2024).

Adaptive experimental design. Adaptive experiments improve estimation efficiency through treatment assignment, recruitment, or both (e.g., Hahn et al., 2011; Addanki et al., 2022). For sequential experiments, related work on Neyman allocation and adaptive designs establishes regret guarantees (Dai et al., 2023; Neopane et al., 2025a;b; Noarov et al., 2025). Other work develops efficient estimation and valid inference under adaptive covariate sampling (Cook et al., 2024) or improves efficiency through adaptive instrument assignment under noncompliance (Oprescu et al., 2025). Recent studies leverage surrogates in adaptive experiments for response-adaptive randomization with delayed outcomes (Shi et al., 2024), efficient treatment allocation and inference (Fan et al., 2025), and online evaluation of surrogate-guided designs (Zhang et al., 2026).

Feature selection and acquisition. Feature selection identifies informative variables for prediction (Yamada et al., 2020; Yoon et al., 2019; Blumberg et al., 2024), while active acquisition chooses which variables to observe based on available measurements (Ma et al., 2019; Li & Oliva, 2021; Covert et al., 2023; Valancius et al., 2024). In our setting, the selected variables are post-treatment outcomes used for causal identification and estimation, and the criterion is treatment-effect estimation variance rather than predictive accuracy.

3 THE ORACLE PROBLEM

3.1 PROBLEM SETUP

For each unit i, let $G_i \in \{O, E\}$ indicate the sample from which the unit is drawn, with $G_i = O$ for the observational sample and $G_i = E$ for the experimental sample. Let $W_i \in \{0, 1\}$ denote the treatment indicator, with $W_i = 1$ for treatment and $W_i = 0$ for control, and let $\mathbf{X}_i \in \mathbb{R}^p$ denote pre-treatment covariates. For $w \in \{0, 1\}$, let $Y_i(w)$ denote the potential primary outcome and $\mathbf{S}_i(w) = \big(S_{1,i}(w), \ldots, S_{K,i}(w)\big)^{\mathrm{T}}$ denote the vector of potential secondary outcomes. The observed outcomes satisfy $Y_i = Y_i(W_i)$ and $\mathbf{S}_i = (S_{1,i}, \ldots, S_{K,i})^{\mathrm{T}} = \mathbf{S}_i(W_i)$.

Let $\mathcal{I}_O$, $\mathcal{I}_P$, and $\mathcal{I}_M$ denote the disjoint index sets for the observational, pilot, and main experimental sample, with sample sizes $n = |\mathcal{I}_O|$, $n_P = |\mathcal{I}_P|$, and $n_M = |\mathcal{I}_M|$, respectively, where n_P and n_M are determined by the budget in Section 3.3. We assume that the full data satisfy $(Y_i(0), Y_i(1), W_i, \mathbf{X}_i, \mathbf{S}_i(0), \mathbf{S}_i(1), G_i) \sim P_*$, where the observational data $\mathcal{D}_O = \{(Y_i, W_i, \mathbf{X}_i, \mathbf{S}_i)\}_{i \in \mathcal{I}_O}$ consists of i.i.d. observations from P_O, whose covariate marginal is denoted by P_O^X. Our target is the primary ATE in the observational population, defined as

$$\tau = \mathbb{E}_{P_*}\{Y_i(1) - Y_i(0) \mid G_i = O\}.$$

For $A \subseteq [K]$, let $\mathbf{S}_{i,A}(w)$ and $\mathbf{S}_{i,A}$ denote the subvectors corresponding to A. We consider a prespecified finite collection $\mathcal{A}$ of candidate secondary-outcome sets and assume that their union, $A_P := \bigcup_{A \in \mathcal{A}} A$, also belongs to $\mathcal{A}$. For each $A \in \mathcal{A}$, let

$$\varrho_A(\mathbf{x}) = \mathbb{P}(W_i = 1 \mid \mathbf{X}_i = \mathbf{x}, G_i = E, A)$$

denote the probability of assigning treatment to an experimental unit with covariates $\mathbf{x}$ when the secondary outcomes in A are measured. Let P_E^X denote the covariate distribution in the experimental population, and let $\mathcal{R}_A$ denote a prespecified class of measurable treatment-assignment rules satisfying $\epsilon_\varrho \le \varrho_A(\mathbf{x}) \le 1 - \epsilon_\varrho$, P_E^X-a.s., for a fixed $\epsilon_\varrho \in (0, 1/2)$. We define the design choice space as $\mathfrak{D} := \{(A, \varrho_A) : A \in \mathcal{A}, \varrho_A \in \mathcal{R}_A\} = \bigcup_{A \in \mathcal{A}} \{A\} \times \mathcal{R}_A$. A candidate design is an element $(A, \varrho_A) \in \mathfrak{D}$. Write $P_{E,A}^{\varrho_A}$ for the resulting joint distribution of $(W_i, \mathbf{X}_i, \mathbf{S}_{i,A})$. Following Chen & Ritzwoller (2023); Athey et al. (2025), we impose the following identification conditions.

Assumption 1 (ESC identification conditions)**.** *For each $A \in \mathcal{A}$ and $w \in \{0, 1\}$, we have:*
(i) experimental unconfoundedness: $W_i \perp\!\!\!\perp (Y_i(w), \mathbf{S}_{i,A}(w)) \mid \mathbf{X}_i, G_i = E$;
(ii) conditional external validity: $G_i \perp\!\!\!\perp (Y_i(w), \mathbf{S}_{i,A}(w)) \mid \mathbf{X}_i$;
(iii) latent unconfoundedness: $W_i \perp\!\!\!\perp Y_i(w) \mid \mathbf{S}_{i,A}(w), \mathbf{X}_i, G_i = O$;
(iv) overlap and support: $P_O^X \ll P_E^X$ and $0 < \mathbb{P}\{W_i = w \mid \mathbf{S}_{i,A}(w), \mathbf{X}_i, G_i = O\} < 1$, P_*-a.s.

The restriction $P_O^X \ll P_E^X$ in part (iv) allows us to define the covariate density ratio $r(\mathbf{x}) = (dP_O^X / dP_E^X)(\mathbf{x})$, P_E^X-a.s. We assume that the candidate secondary-outcome collection $\mathcal{A}$ is prespecified to contain only the secondary-outcome sets for which Assumption 1 is valid. For example, the candidate secondary-outcome collection may be constructed around some substantively justified secondary outcomes, with additional outcomes included to improve efficiency. If only one secondary-outcome set A satisfies the identification conditions, as in the case of a single secondary outcome, the pilot experiment introduced in Section 4.1 below is used only to learn the optimal treatment-assignment rule within $\mathcal{R}_A$. This setting is a special case of our proposed framework.

3.2 THE EFFICIENT INFLUENCE FUNCTION

In this section, we present the efficient influence function of the ATE estimation for a fixed candidate design and decompose it into observational and experimental components. For each $A \in \mathcal{A}$

and $w \in \{0,1\}$, define the observational outcome regression $\mu_{w,A}(\mathbf{s},\mathbf{x}) = \mathbb{E}(Y_i \mid W_i = w, \mathbf{S}_{i,A} = \mathbf{s}, \mathbf{X}_i = \mathbf{x}, G_i = O)$. Its conditional mean in the experimental population is $\overline{\mu}_{w,A}(\mathbf{x}) = \mathbb{E}\{\mu_{w,A}(\mathbf{S}_{i,A},\mathbf{x}) \mid W_i = w, \mathbf{X}_i = \mathbf{x}, G_i = E\}$. Under Assumption 1, these regressions identify the target effect through $\tau = \mathbb{E}_{P_O}\{\overline{\mu}_{1,A}(\mathbf{X}) - \overline{\mu}_{0,A}(\mathbf{X})\}$. The influence function also depends on the observational latent propensity $\rho_{w,A}(\mathbf{s},\mathbf{x}) = \mathbb{P}\{W_i = w \mid \mathbf{S}_{i,A}(w) = \mathbf{s}, \mathbf{X}_i = \mathbf{x}, G_i = O\}$ and the covariate density ratio $r(\mathbf{x})$ defined in Section 3.1. Denote $\boldsymbol{\mu}_A = (\mu_{1,A}, \mu_{0,A}, \overline{\mu}_{1,A}, \overline{\mu}_{0,A})^{\mathrm{T}}$, $\boldsymbol{\kappa}_A = (\rho_{1,A}, \rho_{0,A}, r)^{\mathrm{T}}$, and write $\boldsymbol{\eta}_A = (\boldsymbol{\mu}_A^{\mathrm{T}}, \boldsymbol{\kappa}_A^{\mathrm{T}})^{\mathrm{T}}$.

Under Assumption 1, Theorem 3.1 of Chen & Ritzwoller (2023) gives the efficient influence function (EIF) for our investigated model. We first introduce an observational and an experimental component for the formal definition of EIF. The observational component is

$$\psi_{O,A}(Y_i, W_i, \mathbf{S}_{i,A}, \mathbf{X}_i; \tau, \boldsymbol{\eta}_A) = \frac{W_i\{Y_i - \mu_{1,A}(\mathbf{S}_{i,A}, \mathbf{X}_i)\}}{\rho_{1,A}(\mathbf{S}_{i,A}, \mathbf{X}_i)} - \frac{(1-W_i)\{Y_i - \mu_{0,A}(\mathbf{S}_{i,A}, \mathbf{X}_i)\}}{\rho_{0,A}(\mathbf{S}_{i,A}, \mathbf{X}_i)} + \overline{\mu}_{1,A}(\mathbf{X}_i) - \overline{\mu}_{0,A}(\mathbf{X}_i) - \tau. \tag{1}$$

The experimental component under secondary-outcome set A and treatment-assignment rule ϱ_A is

$$\psi_{E,A}(W_i, \mathbf{S}_{i,A}, \mathbf{X}_i; \boldsymbol{\eta}_A, \varrho_A) = r(\mathbf{X}_i)\left[\frac{W_i}{\varrho_A(\mathbf{X}_i)}\{\mu_{1,A}(\mathbf{S}_{i,A}, \mathbf{X}_i) - \overline{\mu}_{1,A}(\mathbf{X}_i)\} - \frac{1-W_i}{1-\varrho_A(\mathbf{X}_i)}\{\mu_{0,A}(\mathbf{S}_{i,A}, \mathbf{X}_i) - \overline{\mu}_{0,A}(\mathbf{X}_i)\}\right]. \tag{2}$$

Given $(A, \varrho_A) \in \mathfrak{D}$, it follows from Chen & Ritzwoller (2023) that the EIF for τ is

$$\psi(A, \varrho_A) = \frac{\mathbb{1}\{G = O\}}{\pi}\psi_{O,A}(Y, W, \mathbf{S}_A, \mathbf{X}; \tau, \boldsymbol{\eta}_A) + \frac{\mathbb{1}\{G = E\}}{1-\pi}\psi_{E,A}(W, \mathbf{S}_A, \mathbf{X}; \boldsymbol{\eta}_A, \varrho_A),$$

where $\pi := \mathbb{P}(G = O)$. Both components in equation 1 and equation 2 have mean zero under their respective sampling distributions.

3.3 Experimental budget

We conduct a pilot experiment followed by a main experiment under a total experimental budget b. The pilot measures the secondary-outcome set A_P under the balanced randomization $\varrho_{\mathrm{pilot}}(\mathbf{x}) = 1/2$. Measuring A_P in the pilot experiment allows us to evaluate every candidate secondary-outcome set using the same pilot. We allocate a proportion $\lambda_{P,b} \in (0,1)$ of the budget b to the pilot, with the remainder reserved for the main experiment. We consider $b \to \infty$, with $n/b \to h \in (0,\infty)$, $\lambda_{P,b} \to 0$ and $\lambda_{P,b} b \to \infty$, i.e., the ratio of the observational sample size to the budget converges to a constant, and the pilot budget diverges while its share of the total budget vanishes.

Let $c_w(A, \mathbf{x})$ denote the known (conditional) per-unit cost of measuring secondary-outcome set A for a unit with covariates $\mathbf{x}$ assigned to arm w. Such cost heterogeneity naturally arises in practice. In Example 1, for instance, collecting test scores across multiple grades may incur higher costs than collecting scores from a single grade, while administrative costs may also vary across schools when $\mathbf{x}$ includes school-level characteristics. The expected per-unit cost under (A, ϱ_A) is

$$\overline{c}_A(\varrho_A) = \mathbb{E}_{P_E^X}[c_1(A, \mathbf{X})\varrho_A(\mathbf{X}) + c_0(A, \mathbf{X})\{1 - \varrho_A(\mathbf{X})\}].$$

Let $\mathcal{I}_E$ index an experimental sample whose covariates are observed before the pilot. We estimate $\overline{c}_A(\varrho_A)$ by $\widehat{\overline{c}}_A(\varrho_A) = |\mathcal{I}_E|^{-1}\sum_{j\in\mathcal{I}_E}[c_1(A, \mathbf{X}_j)\varrho_A(\mathbf{X}_j) + c_0(A, \mathbf{X}_j)\{1 - \varrho_A(\mathbf{X}_j)\}]$. Assume $|I_E| \to \infty$ as $n \to \infty$, and thus $\widehat{\overline{c}}_A(\varrho_A) \to_p \overline{c}_A(\varrho_A)$. The pilot sample size is $n_P = \lfloor \lambda_{P,b} b / \widehat{\overline{c}}_{A_P}(\varrho_{\mathrm{pilot}}) \rfloor$, and for a fixed candidate design (A, ϱ_A), the corresponding main experimental sample size is $n_M(A, \varrho_A) = \lfloor (1 - \lambda_{P,b}) b / \widehat{\overline{c}}_A(\varrho_A) \rfloor$. The sample sizes are determined by the estimated expected per-unit costs, so the realized total experimental expenditure need not equal b in finite samples. In applications requiring a strict finite-sample budget cap, the sample size must be determined by an upper bound for the realized costs. Expected-budget constraints are also used in related active data-collection methods (Zrnic & Candès, 2024; Gligorić et al., 2025). The pilot sample is $\mathcal{D}_P = \{(W_j, \mathbf{X}_j, \mathbf{S}_{j,A_P})\}_{j\in\mathcal{I}_P}$, which are i.i.d. draws from $P_{E,A_P}^{\varrho_{\mathrm{pilot}}}$, independent of $\mathcal{D}_O$.

After determining (A, ϱ_A) and the implemented sample size n_M, we recruit n_M units for the main experiment. We observe $\mathcal{D}_M = \{(W_j, \mathbf{X}_j, \mathbf{S}_{j,A})\}_{j\in\mathcal{I}_M}$, where $|\mathcal{I}_M| = n_M$. Conditional on the information available before the main experiment, these observations are i.i.d. draws from $P_{E,A}^{\varrho_A}$.

Remark 1 (Choice of the pilot budget proportion)**.** *A larger pilot budget can improve the estimation of the optimal design but leaves less budget for the main experiment. Let $N_b := b/\bar{c}_{A_P}(\varrho_{\text{pilot}})$. Under additional conditions in Appendix C.1, the estimation error of the design criterion, which is defined in equation 5 below, follows $O_p(n_P^{-\gamma_P})$ for some $\gamma_P > 0$. Since $n_P/\lambda_{P,b}N_b \to_p 1$ by equation B.14, the estimation error is of order $O_p((\lambda_{P,b}N_b)^{-\gamma_P})$. Allocating a proportion $\lambda_{P,b}$ of the total experimental budget to the pilot results in an $O(\lambda_{P,b})$ increase in the asymptotic variance of the ATE estimator. Balancing these two terms gives $\lambda_{P,b} \asymp (\lambda_{P,b}N_b)^{-\gamma_P}$, and hence $\lambda_{P,b} \asymp N_b^{-\gamma_P/(1+\gamma_P)}$, $n_P \asymp N_b^{1/(1+\gamma_P)}$. For $\gamma_P = 1/2$, we have $\lambda_{P,b} \asymp N_b^{-1/3}$, $n_P \asymp N_b^{2/3}$.*

3.4 The design objective

This section first formulates the design objective, based on which we define an oracle design under the expected-budget constraint. Define $V_O(A) = \mathbb{E}_{P_O}\{\psi_{O,A}(Y, W, \mathbf{S}_A, \mathbf{X}; \tau, \boldsymbol{\eta}_A)^2\}$ and $V_E(A, \varrho_A) = \mathbb{E}_{P_{E,A}^{\varrho_A}}\{\psi_{E,A}(W, \mathbf{S}_A, \mathbf{X}; \boldsymbol{\eta}_A, \varrho_A)^2\}$, with their explicit expressions provided in Appendix C.3. For each fixed design (A, ϱ_A), let $\widehat{\tau}^*(A, \varrho_A)$ denote the infeasible ATE estimator that uses the true nuisance functions, defined in equation A.1. Our design objective is to minimize the variance of $\widehat{\tau}^*(A, \varrho_A)$ over $(A, \varrho_A) \in \mathfrak{D}$, subject to the constraint $n_M\,\widehat{\bar{c}}_A(\varrho_A) \le (1 - \lambda_{P,b})b$. Corollary A.1 establishes that $\sqrt{n}\{\widehat{\tau}^*(A, \varrho_A) - \tau\} \to_d \mathcal{N}\big(0, \sigma_h^2(A, \varrho_A)\big)$ as $n \to \infty$, where

$$\sigma_h^2(A, \varrho_A) := V_O(A) + h\,\bar{c}_A(\varrho_A)V_E(A, \varrho_A), \tag{3}$$

which corresponds to the semiparametric efficiency bound (Hahn, 1998), since $\widehat{\tau}^*(A, \varrho_A)$ solves the estimating equation induced by the EIF. The asymptotic variance in equation 3 also corresponds to the (asymptotic) design criterion hereafter. We therefore define the oracle design by

$$(A^*, \varrho_{A^*}^*) = \arg\min_{(A,\varrho_A)\in\mathfrak{D}} \sigma_h^2(A, \varrho_A). \tag{4}$$

Since $V_O(A)$ is unaffected by treatment assignment, the oracle design can be equivalently defined as: (i) $\varrho_A^* = \arg\min_{\varrho_A\in\mathcal{R}_A} \bar{c}_A(\varrho_A)V_E(A, \varrho_A)$ for each A, and (ii) $A^* = \arg\min_{A\in\mathcal{A}} \sigma_h^2(A, \varrho_A^*)$. This optimality is restricted to the prespecified family of estimators and treatment-assignment rules.

4 EfficientESC

4.1 Learning the oracle design using a pilot experiment

In this section, we use the observational and pilot samples to estimate the design criterion and learn the oracle design for the main experiment. Ignoring integer rounding in the sample size of the main experiment n_M (see Appendix C.4), the normalized variance under the available main experimental budget gives the finite-sample design criterion

$$V_n(A, \varrho_A) := V_O(A) + h_n\,\bar{c}_A(\varrho_A)V_E(A, \varrho_A), \tag{5}$$

where

$$h_n := \frac{n}{(1 - \lambda_{P,b})b} = \frac{n/b}{1 - \lambda_{P,b}} \to h, \quad \text{as} \quad n \to \infty. \tag{6}$$

Hence, for every fixed (A, ϱ_A), $V_n(A, \varrho_A) \to \sigma_h^2(A, \varrho_A)$ as $n \to \infty$. Using K_P-fold cross-fitting with fixed $K_P \ge 2$, we obtain $\widehat{V}_O^{\text{pilot}}(A)$ and $\widehat{V}_E^{\text{pilot}}(A, \varrho_A)$ from cross-fitted estimates of the EIF components; see details in Appendix C.2.1. We estimate expected costs by averaging the conditional costs over the pre-treatment covariates of the experimental units, as defined in Section 3.3.

The estimated design criterion is then defined as

$$\widehat{V}_n^{\text{pilot}}(A, \varrho_A) = \widehat{V}_O^{\text{pilot}}(A) + h_n\widehat{\bar{c}}_A(\varrho_A)\widehat{V}_E^{\text{pilot}}(A, \varrho_A). \tag{7}$$

For each secondary-outcome set, we first optimize the treatment-assignment rule, and then compare the resulting objective values across sets. The learned design is defined as:

$$\widehat{\varrho}_A = \arg\min_{\varrho_A\in\mathcal{R}_A} \widehat{\bar{c}}_A(\varrho_A)\widehat{V}_E^{\text{pilot}}(A, \varrho_A), \qquad \widehat{A} = \arg\min_{A\in\mathcal{A}} \widehat{V}_n^{\text{pilot}}(A, \widehat{\varrho}_A). \tag{8}$$

Assumption 2(iii) below allows an approximate solution whose optimization error is $o_p(1)$.

4.2 MAIN EXPERIMENT AND TREATMENT EFFECT ESTIMATION

In this section, we construct the treatment effect estimator after the main experiment, its variance estimator, and the corresponding confidence interval. Before recruitment, we fix $(\widehat{A}, \widehat{\varrho}_{\widehat{A}})$ obtained from equation 8 and set $n_M = \lfloor (1-\lambda_{P,b}) b / \widehat{c}_{\widehat{A}}(\widehat{\varrho}_{\widehat{A}}) \rfloor$ (see Appendix C.4). Conditional on the pilot-stage information, we collect n_M i.i.d. observations $(W_j, \mathbf{X}_j, \mathbf{S}_{j,\widehat{A}})$, $j \in \mathcal{I}_M$, from $P_{E,\widehat{A}}^{\widehat{\varrho}_{\widehat{A}}}$.

For the treatment effect estimation, we apply cross-fitted learning (Chernozhukov et al., 2018) by independently partitioning $\mathcal{D}_O$ and $\mathcal{D}_M$ into fixed $K_M \geq 2$ approximately equal-sized folds, denoted by $\{I_{O,k}\}_{k=1}^{K_M}$ and $\{I_{E,k}\}_{k=1}^{K_M}$, respectively. For each evaluation fold k, estimate $\widehat{\boldsymbol{\eta}}_{\widehat{A}}^{\text{main},(-k)}$ using the observations indexed by $\bigcup_{k' \neq k} I_{O,k'}$ and $\bigcup_{k' \neq k} I_{E,k'}$. Our proposed ATE estimator $\widehat{\tau}$ solves

$$0 = \frac{1}{K_M} \sum_{k=1}^{K_M} \left\{ \frac{1}{|I_{O,k}|} \sum_{i \in I_{O,k}} \psi_{O,\widehat{A}}\big(Y_i, W_i, \mathbf{S}_{i,\widehat{A}}, \mathbf{X}_i; \widehat{\tau}, \widehat{\boldsymbol{\eta}}_{\widehat{A}}^{\text{main},(-k)}\big) + \frac{1}{|I_{E,k}|} \sum_{j \in I_{E,k}} \psi_{E,\widehat{A}}\big(W_j, \mathbf{S}_{j,\widehat{A}}, \mathbf{X}_j; \widehat{\boldsymbol{\eta}}_{\widehat{A}}^{\text{main},(-k)}, \widehat{\varrho}_{\widehat{A}}\big) \right\}. \tag{9}$$

Let $\widehat{V}_O^{\text{main}}$ and $\widehat{V}_E^{\text{main}}$ denote the estimated variances of the cross-fitted observational and main experimental components, respectively, as defined in Appendix C.2.2. Then, combining equation 3, equation 6 and the definition of n_M, the estimated asymptotic variance of $\widehat{\tau}$ is

$$\widehat{\sigma}_h^2(\widehat{A}, \widehat{\varrho}_{\widehat{A}}) := \widehat{V}_O^{\text{main}} + \frac{n}{n_M} \widehat{V}_E^{\text{main}}. \tag{10}$$

Under the conditions in Section 5, an asymptotically valid $(1-\alpha)$-confidence interval is

$$\widehat{\tau} \pm z_{1-\alpha/2} \sqrt{\widehat{\sigma}_h^2(\widehat{A}, \widehat{\varrho}_{\widehat{A}})/n}, \tag{11}$$

where $z_{1-\alpha/2}$ is the $(1-\alpha/2)$-quantile of the standard normal distribution. Algorithm 1 summarizes the procedures of our method.

Algorithm 1 EfficientESC

1: **Input**: Observational data $\mathcal{D}_O$, candidate secondary-outcome collection $\mathcal{A}$, cost functions c_0, c_1, budget b, pilot budget proportion $\lambda_{P,b}$, the numbers of folds K_P and K_M, and significance level α.
Stage 1: Pilot experiment
2: Estimate the expected cost $\widehat{c}_A(\varrho_A)$ for a given design $(A, \varrho_A) \in \mathfrak{D}$.
3: Compute the pilot sample size $n_P = \left\lfloor \lambda_{P,b} b / \widehat{c}_{A_P}(\varrho_{\text{pilot}}) \right\rfloor$. Recruit n_P units from the experimental sample, assign treatment according to $\varrho_{\text{pilot}} = 1/2$, and measure $\mathbf{S}_{A_P}$ to obtain the pilot sample $\mathcal{D}_P$.
4: **for** $A \in \mathcal{A}$ **do**
5: Estimate the pilot-stage nuisance functions $\widehat{\boldsymbol{\eta}}_A^{\text{pilot},(-k)}$ by K_P-fold cross-fitting.
6: Compute $\widehat{V}_O^{\text{pilot}}(A)$, $\widehat{V}_E^{\text{pilot}}(A, \varrho_A)$, and $\widehat{c}_A(\varrho_A)$.
7: Obtain $\widehat{\varrho}_A = \arg\min_{\varrho_A \in \mathcal{R}_A} \widehat{c}_A(\varrho_A) \widehat{V}_E^{\text{pilot}}(A, \varrho_A)$.
8: **end for**
9: Select $\widehat{A} = \arg\min_{A \in \mathcal{A}} \widehat{V}_n^{\text{pilot}}(A, \widehat{\varrho}_A)$.
Stage 2: Main experiment
10: Compute the main experimental sample size $n_M = \left\lfloor (1-\lambda_{P,b}) b / \widehat{c}_{\widehat{A}}(\widehat{\varrho}_{\widehat{A}}) \right\rfloor$. Recruit n_M units from the experimental sample, assign treatment according to $\widehat{\varrho}_{\widehat{A}}$, and measure $\mathbf{S}_{\widehat{A}}$.
11: Estimate the main-experiment nuisance functions $\widehat{\boldsymbol{\eta}}_{\widehat{A}}^{\text{main},(-k)}$ by K_M-fold cross-fitting.
12: **Output**: $(\widehat{A}, \widehat{\varrho}_{\widehat{A}})$ from equation 8, $\widehat{\tau}$ from equation 9, and the confidence interval from equation 11.

Remark 2. *The pilot is used to learn the oracle design, but the observations from the pilot sample are not used for the estimation of $\widehat{\tau}$ in equation 9. Using a pilot sample for experimental design and a main sample for effect estimation has precedents in two-stage experimental design (e.g., Cai & Rafi, 2024; Viviano, 2020). This choice is made for theoretical convenience. Specifically, conditional on the pilot-stage information, the main observations are i.i.d. under the learned design. However, excluding the pilot in the effect estimation may worsen the finite-sample efficiency. Related two-stage designs also study extensions combining pilot and main-sample estimators (see, Bai, 2022; Tabord-Meehan, 2023). Our method can also be naturally extended to use this technique.*

5 THEORETICAL RESULTS

5.1 ORACLE DESIGN RECOVERY

In this section, we establish the asymptotic recovery of the oracle design using our method. Notice that the norms used below are defined in equation A.5 in Appendix A.1. For each (A, ϱ_A), write $Q_A(\varrho_A) = \overline{c}_A(\varrho_A)V_E(A, \varrho_A)$ and $\widehat{Q}_A(\varrho_A) = \widehat{\overline{c}}_A(\varrho_A)\widehat{V}_E^{\text{pilot}}(A, \varrho_A)$. The following conditions ensure uniform estimation of these objectives and identify their minimizers.

Assumption 2. *(i) For each $A \in \mathcal{A}$ and pilot evaluation fold $k \in [K_P]$, let $\widehat{\boldsymbol{\eta}}_A^{\text{pilot},(-k)}$ denote the nuisance estimates. Then, $\max_{A\in\mathcal{A}} \max_{k\le K_P} \|\widehat{\boldsymbol{\eta}}_A^{\text{pilot},(-k)} - \boldsymbol{\eta}_A\|_{\eta,2} = o_p(1)$, where the norm is defined in equation A.5*
(ii) $\mathcal{R}_A \subset L_2(P_E^X)$ is a measurable and compact VC-subgraph class. For some constants $q > 2$ and $0 < \underline{c} < C < \infty$, $\inf_{(A,\varrho_A)\in\mathfrak{D}} \overline{c}_A(\varrho_A) \ge \underline{c}$, $\max_{A\in\mathcal{A},w\in\{0,1\}} \|c_w(A,\cdot)\|_\infty \le C$, and $\max_{A\in\mathcal{A}} \|\psi_{O,A}(\,\cdot\,;\tau,\boldsymbol{\eta}_A)\|_{P_O,q} + \max_{A\in\mathcal{A}} \sup_{\varrho_A\in\mathcal{R}_A} \|\psi_{E,A}(\,\cdot\,;\boldsymbol{\eta}_A,\varrho_A)\|_{P_{E,A}^{\varrho_A},q} \le C$. Moreover, Q_A is continuous with respect to the $\|\cdot\|_{P_E^X,2}$-norm on $\mathcal{R}_A$, and has the unique minimizer ϱ_A^* defined in Section 3.4.
(iii) For every $A \in \mathcal{A}$, $\widehat{\varrho}_A \in \mathcal{R}_A$ satisfies $\widehat{Q}_A(\widehat{\varrho}_A) \le \inf_{\varrho_A\in\mathcal{R}_A} \widehat{Q}_A(\varrho_A) + \xi_{A,b}$, with $\max_{A\in\mathcal{A}} \xi_{A,b} = o_p(1)$,.
(iv) If $\mathcal{A} \setminus \{A^*\} \neq \varnothing$, then $\min_{A\in\mathcal{A}\setminus\{A^*\}}\{\sigma_h^2(A,\varrho_A^*) - \sigma_h^2(A^*,\varrho_{A^*}^*)\} > 0$.
(v) There exist constants $\epsilon \in (0, 1/2)$ and $0 < \underline{r} < \overline{r} < \infty$, such that for every $A \in \mathcal{A}$ and $w \in \{0,1\}$, $\epsilon < \rho_{w,A}(\mathbf{S}_A, \mathbf{X}) < 1-\epsilon$, P_*-a.s., and $\underline{r} < r(\mathbf{X}) < \overline{r}$, P_E^X-a.s. Estimates of $\rho_{w,A}$ and r are clipped to $[\epsilon, 1-\epsilon]$ and $[\underline{r}, \overline{r}]$, respectively, as in equation C.37 and equation C.38.

Part (ii) in Assumption 2 imposes regularity conditions on the assignment-rule class and uniform bounds on the costs and influence-function moments. For example, rules that assign a different probability to each of a fixed number of prespecified covariate strata form a compact VC-subgraph class. Such restrictions are commonly used in the policy-learning literature (Kitagawa & Tetenov, 2018; Athey & Wager, 2021). The optimization tolerance in part (iii) permits numerical solutions whose error vanishes, while part (iv) distinguishes a unique oracle secondary-outcome set. Together, these conditions support the usual consistency argument for extremum estimators (Newey & McFadden, 1994; van der Vaart, 1998). Part (v) imposes standard overlap and clipping conditions.

Theorem 1 (Recovery of the oracle design)**.** *Suppose that Assumptions 1 and 2 hold, $\lambda_{P,b} \to 0$, and $\lambda_{P,b} b \to \infty$. Then, as $n \to \infty$,*

$$\mathbb{P}(\widehat{A} = A^*) \to 1, \qquad \|\widehat{\varrho}_{\widehat{A}} - \varrho_{A^*}^*\|_{P_E^X,2} = o_p(1).$$

Moreover, $\widehat{V}_n^{\text{pilot}}(\widehat{A}, \widehat{\varrho}_{\widehat{A}}) \to_p \sigma_h^2(A^*, \varrho_{A^*}^*)$ as $n \to \infty$.

Theorem 1 implies that the proposed EFFICIENTESC asymptotically recovers the oracle design.

5.2 ASYMPTOTIC PROPERTIES OF THE PROPOSED ESTIMATOR

This section establishes the asymptotic normality and consistent variance estimation for the proposed estimator. Define $\mathcal{D}_E^X = \{\mathbf{X}_j : j \in \mathcal{I}_E\}$ and $\mathcal{U}_{\text{pilot}} = (\{\widehat{\varrho}_A : A \in \mathcal{A}\}, \widehat{A}, n_M)$, and let $\mathcal{F}_{\text{pilot}} = \sigma(\mathcal{D}_O, \mathcal{D}_E^X, \mathcal{D}_P, \mathcal{U}_{\text{pilot}})$ denote the information in the observational and pilot samples and the corresponding estimated quantities. Conditional on $\mathcal{F}_{\text{pilot}}$, the observations in the main experiment are i.i.d. under $P_{E,\widehat{A}}^{\widehat{\varrho}_{\widehat{A}}}$.

Assumption 3. *(i) For some constant $c > 0$, $V_O(A^*) \wedge V_E(A^*, \varrho_{A^*}^*) \ge c$.*
(ii) On the event $\{\widehat{A} = A^*\}$, the nuisance estimates satisfy $\max_{k\le K_M} \|\widehat{\boldsymbol{\eta}}_{A^*}^{\text{main},(-k)} - \boldsymbol{\eta}_{A^*}\|_{\eta,2} = o_p(1)$, $\max_{k\le K_M} \|\widehat{\boldsymbol{\mu}}_{A^*}^{\text{main},(-k)} - \boldsymbol{\mu}_{A^*}\|_{\mu,2} \times \|\widehat{\boldsymbol{\kappa}}_{A^*}^{\text{main},(-k)} - \boldsymbol{\kappa}_{A^*}\|_{\kappa,2} = o_p(n^{-1/2})$, where the norms are defined in equation A.5.
(iii) There exists a fixed measurable P_O-Donsker class $\mathcal{F}_O$ with a square-integrable envelope containing $\psi_{O,A^*}(\,\cdot\,;\tau,\boldsymbol{\eta}_{A^*})$. On the event $\{\widehat{A} = A^*\}$, with probability tending to one, $\psi_{O,A^*}(\,\cdot\,;\tau,\widehat{\boldsymbol{\eta}}_{A^*}^{\text{main},(-k)}) \in \mathcal{F}_O$ for every $k \le K_M$.

Part (i) ensures a nondegenerate asymptotic variance. In part (ii), the first condition requires consistency of the main-stage nuisance estimates, while the second condition controls the second-order bias remaining after orthogonal correction. In particular, neither nuisance function is required to converge at the parametric $n^{-1/2}$ rate (Chernozhukov et al., 2018; Chen & Ritzwoller, 2023). Under mild assumptions, the required rates can be obtained with sparse estimators (Belloni et al., 2014; Bickel et al., 2009), boosting and forests (Luo & Spindler, 2017; Wager & Walther, 2015), or neural networks (Chen & White, 1999; Farrell et al., 2021). Reusing the observational sample for both design learning and effect estimation induces post-selection dependence that is not removed by the nuisance-function cross-fitting above (Hubbard et al., 2016), and part (iii) controls this dependence through stochastic equicontinuity over a fixed Donsker class (van der Vaart & Wellner, 1996). This condition can be relaxed by additionally cross-fitting the design-learning stage. However, we refrain from doing so, as sample splitting may yield different treatment-assignment rules across experimental units, raising fairness concerns and potentially reducing the efficiency of treatment assignment.

Theorem 2 (Asymptotic properties of our estimator)**.** *Suppose that the assumptions of Theorem 1 and Assumption 3 hold. Then, as $n \to \infty$,*

$$\sqrt{n}(\widehat{\tau} - \tau) \to_d \mathcal{N}\big(0, \sigma_h^2(A^*, \varrho^*_{A^*})\big), \qquad \widehat{\sigma}_h^2(\widehat{A}, \widehat{\varrho}_{\widehat{A}}) \to_p \sigma_h^2(A^*, \varrho^*_{A^*}).$$

Moreover, the confidence interval defined in equation 11 has asymptotic $(1-\alpha)$ coverage for any fixed $\alpha \in (0,1)$.

Theorem 2 demonstrates that the proposed estimator attains the minimized semiparametric efficiency bound $\sigma_h^2(A^*, \varrho^*_{A^*})$.

6 SIMULATIONS

The motivating application in Example 1 relies on New York City administrative data that are not publicly available. We therefore use synthetic data to evaluate the proposed experimental design.

6.1 SETUP

The detailed data generating process is relegated to Appendix D.1.1. We consider three secondary outcomes: the first captures observational confounding, the second provides additional prognostic information about the primary outcome, and the third is redundant once the second is observed. Thus, the candidate secondary-outcome collection is $\mathcal{A} = \{\{1\}, \{1,2\}, \{1,2,3\}\}$, with expected per-unit costs 1.00, 1.15, and 1.30, respectively. All candidate secondary-outcome sets satisfy the identification conditions in Assumption 1. A binary covariate divides the population into two equally likely strata with different optimal treatment probabilities. Each candidate rule assigns different treatment probabilities in $[0.1, 0.9]$ for $X = 0$ and $X = 1$. The observational sample size varies over $n \in \{500, 1000, \ldots, 10000\}$, with $b = n/0.8$. The balanced pilot measures all three secondary outcomes and receives the budget proportion $\lambda_{P,b} = 3(b/1.30)^{-1/3}$. The remaining budget determines the main experimental sample size. All simulation results are based on 5000 replications.

We compare nine competing methods under a common pilot budget proportion, including (i) balanced ($\varrho_A(x) = 1/2$) and adaptive assignment for each of the three fixed secondary-outcome sets, (ii) EFFICIENTESC, (iii) a regret-optimal method proposed by Epanomeritakis & Viviano (2025), (iv) an oracle-rule benchmark. The fixed-set adaptive designs estimate only the treatment-assignment rule, whereas EFFICIENTESC also selects the secondary-outcome set. The oracle-rule benchmark uses the oracle design, while estimating the nuisance functions and treatment effect from the observations. For a fair comparison of efficiency in the main experiment, the same pilot expenditure is deducted for all nine methods, including the prespecified balanced designs, even though these designs do not require a pilot. Appendix D.3.5 reports the additional comparison in which prespecified balanced designs use the full experimental budget.

6.2 RESULTS

Figure 1 reports the normalized variance across observational sample sizes for the nine methods under the same pilot budget. A few trends are observable. *First*, EFFICIENTESC closely tracks the oracle-rule benchmark as n increases and approaches the optimal asymptotic variance, consistent

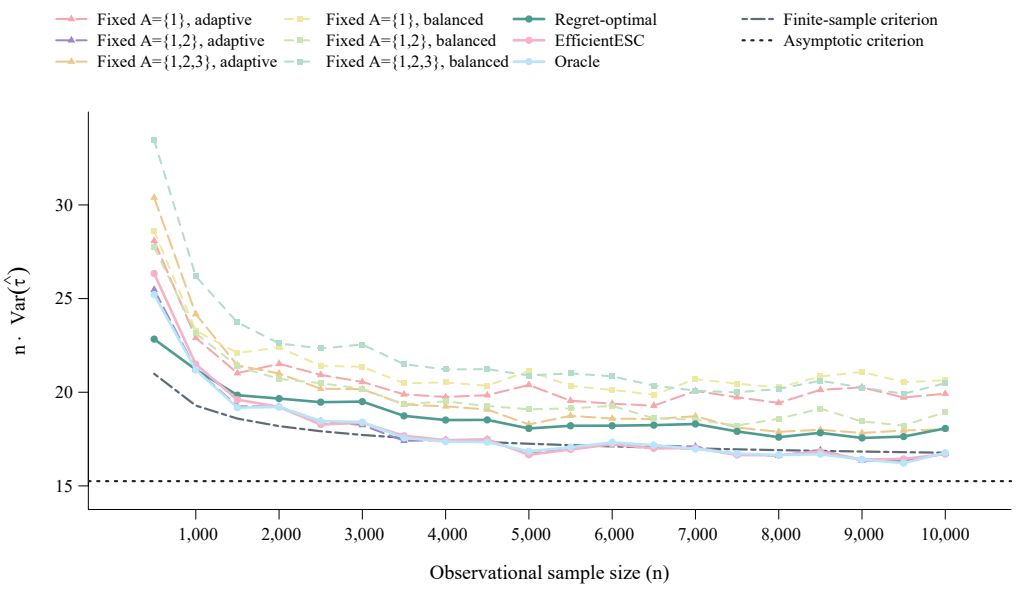

Figure 1: Normalized variance across nine methods under a common pilot budget proportion. The gray dashed curve shows the optimal finite-sample criterion $V^* = \inf_{(A,\varrho_A)\in\mathfrak{D}} V_n(A, \varrho_A)$. The black dotted line shows the asymptotic criterion $\sigma^2_{0.8}(A^*, \varrho^*_{A^*})$.

with Theorem 2. At the budgets considered, its curve also lies close to the finite-sample criterion. For large n, EFFICIENTESC consistently achieves smaller normalized variances than the regret-optimal method. This aligns with their respective design objectives, as the regret-optimal method balances variance against sensitivity to observational bias, whereas our method aims at minimizing the semiparametric efficiency bound.

Second, for each fixed secondary-outcome set, adaptive assignment yields smaller normalized variance than balanced assignment throughout the sample-size range. The improvement therefore illustrates the value of adjusting treatment probabilities to covariates.

Third, under both adaptive and balanced assignment, the fixed set $A = \{1, 2\}$ yields the smallest normalized variance among the three candidate sets. This result highlights the tradeoff between statistical information and measurement cost in secondary-outcome selection. Extensive additional simulations in Appendix D.3 further support the performance of EFFICIENTESC. The empirical size of the proposed test remains close to the nominal level, while its power is significantly higher than that of the competing methods. The learned design converges to the oracle design, consistent with Theorem 1. In the sensitivity analyses, the proposed method exhibits stable performance across different pilot budget proportions and observation-to-budget ratios.

7 DISCUSSION

This paper develops a novel framework that learns the optimal design for using experimentation to correct selection bias in observational studies. Several limitations suggest directions for future work. *First*, excluding pilot observations from final effect estimation may sacrifice finite-sample efficiency. Incorporating them would require accounting for the post-selection dependence induced by using the same pilot sample for both design learning and effect estimation. *Second*, all candidate secondary-outcome sets are assumed to satisfy latent unconfoundedness. When this condition may not hold in practice, sensitivity analyses for potential violations would be valuable extensions. *Third*, the cost functions are assumed known. Extending the framework to estimated costs would require accounting for cost estimation error in both design and treatment effect inference.

GENERATIVE ARTIFICIAL INTELLIGENCE (AI) USE STATEMENT

After completing the full manuscript draft and core research, the author(s) used generative AI tools (e.g., ChatGPT) for auxiliary tasks such as language refinement, formatting, and minor code editing. The authors reviewed, tested, and verified all AI-assisted outputs and take full responsibility for the content of the publication.

ETHICS STATEMENT

Our work adheres to the ICLR Code of Ethics. It is methodological and uses only synthetic simulation data. It does not involve human subjects or personally identifiable information.

REPRODUCIBILITY STATEMENT

We provide complete proofs and additional methodological details in the appendix. The simulation data generating processes, implementation details, and evaluation procedures are also documented there. Code for reproducing the numerical experiments will be included in the supplementary materials.

REFERENCES

Raghavendra Addanki, David Arbour, Tung Mai, Cameron Musco, and Anup Rao. Sample constrained treatment effect estimation. In *Advances in Neural Information Processing Systems*, volume 35, pp. 5417–5430, 2022.

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

Susan Athey, Raj Chetty, and Guido Imbens. The experimental selection correction estimator: Using experiments to remove biases in observational estimates. Technical report, National Bureau of Economic Research, 2025.

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

Yuehao Bai. Optimality of matched-pair designs in randomized controlled trials. *American Economic Review*, 112(12):3911–3940, 2022.

Alexandre Belloni, Victor Chernozhukov, and Lie Wang. Pivotal estimation via square-root Lasso in nonparametric regression. *The Annals of Statistics*, 42(2):757–788, 2014.

Barbara Biasi, Julien Lafortune, and David Schönholzer. What works and for whom? effectiveness and efficiency of school capital investments across the us. *The Quarterly Journal of Economics*, 140(3):2329–2379, 2025.

Aurélien Bibaut, Nathan Kallus, Simon Ejdemyr, and Michael Zhao. Long-term causal inference with imperfect surrogates using many weak experiments, proxies, and cross-fold moments. *arXiv preprint arXiv:2311.04657*, 2023.

Peter J. Bickel, Ya'acov Ritov, and Alexandre B. Tsybakov. Simultaneous analysis of Lasso and Dantzig selector. *The Annals of Statistics*, 37(4):1705–1732, 2009.

Stefano B. Blumberg, Paddy J. Slator, and Daniel C. Alexander. Experimental design for multi-channel imaging via task-driven feature selection. In *International Conference on Learning Representations*, 2024.

Yong Cai and Ahnaf Rafi. On the performance of the neyman allocation with small pilots. *Journal of Econometrics*, 242(1):105793, 2024.

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

Xiaohong Chen and Halbert White. Improved rates and asymptotic normality for nonparametric neural network estimators. *IEEE Transactions on Information Theory*, 45(2):682–691, 1999.

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

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

Thomas Cook, Alan Mishler, and Aaditya Ramdas. Semiparametric efficient inference in adaptive experiments. In *Proceedings of the Third Conference on Causal Learning and Reasoning*, volume 236 of *Proceedings of Machine Learning Research*, pp. 1033–1064. PMLR, 2024.

Ian Connick Covert, Wei Qiu, Mingyu Lu, Na Yoon Kim, Nathan J. White, and Su-In Lee. Learning to maximize mutual information for dynamic feature selection. In *Proceedings of the 40th International Conference on Machine Learning*, volume 202 of *Proceedings of Machine Learning Research*, pp. 6424–6447. PMLR, 2023.

Jessica Dai, Paula Gradu, and Christopher Harshaw. CLIP-OGD: An experimental design for adaptive neyman allocation in sequential experiments. In *Advances in Neural Information Processing Systems*, volume 36, pp. 32235–32269, 2023.

Aristotelis Epanomeritakis and Davide Viviano. Learning what to learn: Experimental design when combining experimental with observational evidence. *arXiv preprint arXiv:2510.23434*, 2025.

Yingying Fan, Zihan Wang, and Waverly Wei. SLOACI: Surrogate-leveraged online adaptive causal inference. *arXiv preprint arXiv:2512.06872*, 2025.

Max H. Farrell, Tengyuan Liang, and Sanjog Misra. Deep neural networks for estimation and inference. *Econometrica*, 89(1):181–213, 2021.

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

Kristina Gligorić, Tijana Zrnic, Cinoo Lee, Emmanuel J. Candès, and Dan Jurafsky. Can unconfident LLM annotations be used for confident conclusions? In *Proceedings of the 2025 Conference of the Nations of the Americas Chapter of the Association for Computational Linguistics: Human Language Technologies (Volume 1: Long Papers)*, pp. 3514–3533. Association for Computational Linguistics, 2025.

Jinyong Hahn. On the role of the propensity score in efficient semiparametric estimation of average treatment effects. *Econometrica*, 66(2):315–331, 1998.

Jinyong Hahn, Keisuke Hirano, and Dean Karlan. Adaptive experimental design using the propensity score. *Journal of Business & Economic Statistics*, 29(1):96–108, 2011.

Joseph W. Hogan, Jason Roy, and Christina Korkontzelou. Handling drop-out in longitudinal studies. *Statistics in Medicine*, 23(9):1455–1497, 2004.

Feifang Hu and William F Rosenberger. *The Theory of Response-Adaptive Randomization in Clinical Trials*. John Wiley & Sons, 2006.

Alan E. Hubbard, Sara Kherad-Pajouh, and Mark J. van der Laan. Statistical inference for data adaptive target parameters. *The International Journal of Biostatistics*, 12(1):3–19, 2016.

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

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

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

Alan B Krueger. Experimental estimates of education production functions. *The Quarterly Journal of Economics*, 114(2):497–532, 1999.

Yang Li and Junier Oliva. Active feature acquisition with generative surrogate models. In *Proceedings of the 38th International Conference on Machine Learning*, volume 139 of *Proceedings of Machine Learning Research*, pp. 6450–6459. PMLR, 2021.

Ye Luo and Martin Spindler. l_2-boosting for economic applications. *American Economic Review*, 107(5):270–273, 2017.

Chao Ma, Sebastian Tschiatschek, Konstantina Palla, José Miguel Hernández-Lobato, Sebastian Nowozin, and Cheng Zhang. EDDI: Efficient dynamic discovery of high-value information with partial VAE. In *Proceedings of the 36th International Conference on Machine Learning*, volume 97 of *Proceedings of Machine Learning Research*, pp. 4234–4243. PMLR, 2019.

Ojash Neopane, Aaditya Ramdas, and Aarti Singh. Logarithmic neyman regret for adaptive estimation of the average treatment effect. In *Proceedings of the 28th International Conference on Artificial Intelligence and Statistics*, volume 258 of *Proceedings of Machine Learning Research*, pp. 4303–4311. PMLR, 2025a.

Ojash Neopane, Aaditya Ramdas, and Aarti Singh. Optimistic algorithms for adaptive estimation of the average treatment effect. In *Proceedings of the 42nd International Conference on Machine Learning*, volume 267 of *Proceedings of Machine Learning Research*, pp. 45895–45910. PMLR, 2025b.

Whitney K. Newey and Daniel McFadden. Large sample estimation and hypothesis testing. In Robert F. Engle and Daniel L. McFadden (eds.), *Handbook of Econometrics*, volume 4, chapter 36, pp. 2111–2245. Elsevier, 1994.

Georgy Noarov, Riccardo Fogliato, Martin Andres Bertran, and Aaron Roth. Stronger neyman regret guarantees for adaptive experimental design. In *Proceedings of the 42nd International Conference on Machine Learning*, volume 267 of *Proceedings of Machine Learning Research*, pp. 46735–46761. PMLR, 2025.

Miruna Oprescu, Brian M. Cho, and Nathan Kallus. Efficient adaptive experimentation with noncompliance. In *Advances in Neural Information Processing Systems*, volume 38, pp. 47432–47468, 2025.

Donald B. Rubin. Comment: Randomization analysis of experimental data: The fisher randomization test. *Journal of the American Statistical Association*, 75(371):591–593, 1980.

Lei Shi, Waverly Wei, and Jingshen Wang. Using surrogates in covariate-adjusted response-adaptive randomized experiments with delayed outcomes. In *Advances in Neural Information Processing Systems*, volume 37, pp. 71085–71121, 2024.

Max Tabord-Meehan. Stratification trees for adaptive randomisation in randomised controlled trials. *The Review of Economic Studies*, 90(5):2646–2673, 2023.

Michael Valancius, Maxwell Lennon, and Junier Oliva. Acquisition conditioned oracle for nongreedy active feature acquisition. In *Proceedings of the 41st International Conference on Machine Learning*, volume 235 of *Proceedings of Machine Learning Research*, pp. 48957–48975. PMLR, 2024.

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

Aad W. van der Vaart and Jon A. Wellner. *Weak Convergence and Empirical Processes: With Applications to Statistics*. Springer Series in Statistics. Springer, New York, 1996.

Davide Viviano. Experimental design under network interference. *arXiv preprint arXiv:2003.08421*, 2020.

Stefan Wager and Guenther Walther. Adaptive concentration of regression trees, with application to random forests. *arXiv preprint arXiv:1503.06388*, 2015.

Yutaro Yamada, Ofir Lindenbaum, Sahand Negahban, and Yuval Kluger. Feature selection using stochastic gates. In *Proceedings of the 37th International Conference on Machine Learning*, volume 119 of *Proceedings of Machine Learning Research*, pp. 10648–10659. PMLR, 2020.

Jinsung Yoon, James Jordon, and Mihaela van der Schaar. INVASE: Instance-wise variable selection using neural networks. In *International Conference on Learning Representations*, 2019.

Zhenghao Zeng, David Arbour, Avi Feller, Raghavendra Addanki, Ryan Rossi, Ritwik Sinha, and Edward H. Kennedy. Continuous treatment effects with surrogate outcomes. In *Proceedings of the 41st International Conference on Machine Learning*, volume 235 of *Proceedings of Machine Learning Research*, pp. 58306–58328. PMLR, 2024.

Lixin Zhang, Feifang Hu, Siu Hung Cheung, and Wai Sum Chan. Asymptotic properties of covariate-adjusted response-adaptive designs. *The Annals of Statistics*, 35(3):1166–1182, 2007.

Wenxin Zhang, Aaron Hudson, Maya Petersen, and Mark van der Laan. An online meta-level adaptive design framework with targeted learning inference: Applications to evaluating and utilizing surrogate outcomes in adaptive designs. *Journal of the American Statistical Association*, 2026.

Tijana Zrnic and Emmanuel J. Candès. Active statistical inference. In *Proceedings of the 41st International Conference on Machine Learning*, volume 235 of *Proceedings of Machine Learning Research*, pp. 62993–63010, 2024.

APPENDIX

In the appendix, we establish the fixed-design asymptotic results in Appendix A, and provide the proofs of the main theoretical results in Appendix B, the additional methodological details in Appendix C, as well as the additional simulation results in Appendix D.

We first introduce some definitions and notation. For each candidate set $A \in \mathcal{A}$, define

$$\mathbf{Z}_{O,A,i} = (Y_i, W_i, \mathbf{S}_{i,A}^{\mathrm{T}}, \mathbf{X}_i^{\mathrm{T}})^{\mathrm{T}}, \qquad \mathbf{Z}_{E,A,j} = (W_j, \mathbf{S}_{j,A}^{\mathrm{T}}, \mathbf{X}_j^{\mathrm{T}})^{\mathrm{T}}.$$

For a measurable function f defined on the corresponding space, we use the same notation f for generic functions whose arguments are determined by the input under consideration. For index sets I_O and I_E, define the empirical measures:

$$\widehat{P}_{O,A,I_O} = \frac{1}{|I_O|} \sum_{i \in I_O} \delta_{\mathbf{Z}_{O,A,i}}, \qquad \widehat{P}_{E,A,I_E} = \frac{1}{|I_E|} \sum_{j \in I_E} \delta_{\mathbf{Z}_{O,A,j}},$$

where $\delta_{\mathbf{Z}_{O,A,i}}$ and $\delta_{\mathbf{Z}_{E,A,j}}$ denote the corresponding Dirac probability measures. The pilot observational and experimental samples are further divided into cross-fitting folds, whose index sets are defined in Appendix C.2. The observational and experimental samples in the main stage are also divided into cross-fitting folds, defined in Section 4.2. Then, for any measurable function f, let $\mathbb{E}_{\widehat{P}_{O,A}}(f) := n^{-1} \sum_{i \in \mathcal{I}_O} f(\mathbf{Z}_{O,A,i})$. Define the notation below:

$$\begin{aligned}
&\mathbb{E}_{\widehat{P}^{\mathrm{pilot}}_{E,A_P}}(f) = \frac{1}{n_P} \sum_{j \in I_P} f(\mathbf{Z}^{\mathrm{pilot}}_{E,A_P,j}), && \mathbb{E}_{\widehat{P}^{\mathrm{main}}_{E,A}}(f) = \frac{1}{n_M} \sum_{j \in I_M} f(\mathbf{Z}^{\mathrm{main}}_{E,A,j}),\\
&\mathbb{E}_{\widehat{P}^{\mathrm{pilot}}_{O,A_P,k}}(f) = \frac{1}{|I^P_{O,k}|} \sum_{i \in I^P_{O,k}} f(\mathbf{Z}_{O,A_P,i}), && \mathbb{E}_{\widehat{P}^{\mathrm{main}}_{O,A,k}}(f) = \frac{1}{|I_{O,k}|} \sum_{i \in I_{O,k}} f(\mathbf{Z}_{O,A,i}),\\
&\mathbb{E}_{\widehat{P}^{\mathrm{pilot}}_{E,A_P,k}}(f) = \frac{1}{|I^P_{E,k}|} \sum_{j \in I^P_{E,k}} f(\mathbf{Z}^{\mathrm{pilot}}_{E,A_P,j}), && \mathbb{E}_{\widehat{P}^{\mathrm{main}}_{E,A,k}}(f) = \frac{1}{|I_{E,k}|} \sum_{j \in I_{E,k}} f(\mathbf{Z}^{\mathrm{main}}_{E,A,j}).
\end{aligned}$$

For a sequence of random vectors X_n, we write $\mathcal{L}(X_n \mid \mathcal{F}) \rightsquigarrow_p \mathcal{L}(X)$ if, for every bounded continuous function $g : \mathbb{R}^d \to \mathbb{R}$, $\mathbb{E}\{g(X_n) \mid \mathcal{F}\} \to_p \mathbb{E}\{g(X)\}$.

A ADDITIONAL THEORETICAL RESULTS

This section establishes the fixed-design asymptotic normality of the infeasible ATE estimator $\widehat{\tau}^*(A, \varrho_A)$ and derives its asymptotic variance, which corresponds to the semiparametric efficiency bound used to define the design objective and oracle design in Section 3.4.

A.1 ASYMPTOTIC PROPERTIES OF THE ESTIMATOR UNDER FIXED DESIGN

For a fixed design (A, ϱ_A), let $n_M = n_M(A, \varrho_A)$ and consider the observational sample together with an independent main experimental sample $\mathcal{D}_M(A, \varrho_A; n_M)$. We first define the infeasible ATE estimator $\widehat{\tau}^*(A, \varrho_A)$, which uses the true nuisance functions $\boldsymbol{\eta}_A$, as the solution to

$$\begin{aligned}
0 = &\frac{1}{n} \sum_{i \in \mathcal{I}_O} \psi_{O,A}\big(Y_i, W_i, \mathbf{S}_{i,A}, \mathbf{X}_i; \widehat{\tau}^*(A, \varrho_A), \boldsymbol{\eta}_A\big)\\
&+ \frac{1}{n_M} \sum_{j \in \mathcal{I}_M} \psi_{E,A}\big(W_j, \mathbf{S}_{j,A}, \mathbf{X}_j; \boldsymbol{\eta}_A, \varrho_A\big).
\end{aligned} \tag{A.1}$$

To state the required rates, recall the nuisance vector $\boldsymbol{\eta}_A = (\boldsymbol{\mu}_A^{\mathrm{T}}, \boldsymbol{\kappa}_A^{\mathrm{T}})^{\mathrm{T}}$ from Section 3.2. Let

$$\widetilde{\boldsymbol{\eta}}_A = (\widetilde{\boldsymbol{\mu}}_A^{\mathrm{T}}, \widetilde{\boldsymbol{\kappa}}_A^{\mathrm{T}})^{\mathrm{T}} \tag{A.2}$$

denote a generic candidate nuisance vector, where

$$\widetilde{\boldsymbol{\mu}}_A = (\widetilde{\mu}_{1,A}, \widetilde{\mu}_{0,A}, \widetilde{\overline{\mu}}_{1,A}, \widetilde{\overline{\mu}}_{0,A})^{\mathrm{T}}, \qquad \widetilde{\boldsymbol{\kappa}}_A = (\widetilde{\rho}_{1,A}, \widetilde{\rho}_{0,A}, \widetilde{r})^{\mathrm{T}}.$$

Each component is an arbitrary measurable function with the same domain and codomain as its corresponding component in $\boldsymbol{\eta}_A$. For each scalar component f of $\boldsymbol{\eta}_A$, write $\Delta f = \widetilde{f} - f$.

For a fixed design (A, ϱ_A) and $q \geq 1$, define

$$\|\widetilde{\boldsymbol{\mu}}_A - \boldsymbol{\mu}_A\|_{\mu,q} := \max_{w\in\{0,1\}} \Big(\|\Delta\mu_{w,A}\|_{P_O,q} \vee \|\Delta\mu_{w,A}\|_{P_{E,A}^{\varrho_A},q} \vee \|\Delta\overline{\mu}_{w,A}\|_{P_O^X,q} \vee \|\Delta\overline{\mu}_{w,A}\|_{P_E^X,q} \Big), \tag{A.3}$$

$$\|\widetilde{\boldsymbol{\kappa}}_A - \boldsymbol{\kappa}_A\|_{\kappa,q} := \left(\max_{w\in\{0,1\}} \|\Delta\rho_{w,A}\|_{P_O,q} \right) \vee \|\Delta r\|_{P_E^X,q}, \tag{A.4}$$

$$\|\widetilde{\boldsymbol{\eta}}_A - \boldsymbol{\eta}_A\|_{\eta,q} := \|\widetilde{\boldsymbol{\mu}}_A - \boldsymbol{\mu}_A\|_{\mu,q} \vee \|\widetilde{\boldsymbol{\kappa}}_A - \boldsymbol{\kappa}_A\|_{\kappa,q}. \tag{A.5}$$

For a fixed candidate design (A, ϱ_A), we impose the following conditions.

Assumption A.1. *(i) For some constants $q > 2$ and $0 < c < C < \infty$, $\|\psi_{O,A}(\cdot\,; \tau, \boldsymbol{\eta}_A)\|_{P_O,q} + \|\psi_{E,A}(\cdot\,; \boldsymbol{\eta}_A, \varrho_A)\|_{P_{E,A}^{\varrho_A},q} \leq C$, and $V_O(A) \wedge V_E(A, \varrho_A) \geq c$.*
(ii) There exist constants $\epsilon \in (0, 1/2)$ and $0 < \underline{r} < \overline{r} < \infty$ such that, for $w \in \{0,1\}$, $\epsilon < \rho_{w,A}(\mathbf{S}_A, \mathbf{X}) < 1 - \epsilon$, P_*-a.s., and $\underline{r} < r(\mathbf{X}) < \overline{r}$, P_E^X-a.s. The main-stage estimates of $\rho_{w,A}$ and r are clipped to $[\epsilon, 1-\epsilon]$ and $[\underline{r}, \overline{r}]$, respectively.
(iii) The nuisance estimates satisfy $\max_{k\leq K_M} \|\widehat{\boldsymbol{\eta}}_A^{\mathrm{main},(-k)} - \boldsymbol{\eta}_A\|_{\eta,2} = o_p(1)$, and $\max_{k\leq K_M} \|\widehat{\boldsymbol{\mu}}_A^{\mathrm{main},(-k)} - \boldsymbol{\mu}_A\|_{\mu,2} \times \|\widehat{\boldsymbol{\kappa}}_A^{\mathrm{main},(-k)} - \boldsymbol{\kappa}_A\|_{\kappa,2} = o_p(n^{-1/2})$, where the norms are defined in equation A.5.

Assumption A.1 is the fixed-design counterpart of the corresponding conditions in Assumptions 2 and 3. Part (i) imposes similar moment and nondegeneracy conditions, part (ii) uses similar bounds and clipping, and part (iii) imposes similar main-stage nuisance estimation convergence conditions.

Let $\widehat{\tau}(A, \varrho_A)$ denote the cross-fitted estimator in equation 9 evaluated at the fixed design (A, ϱ_A), and let $\widehat{\sigma}_h^2(A, \varrho_A)$ denote the corresponding variance estimator similar to equation 10. We have the following results.

Lemma A.1 (Consistency of the estimated expected cost)**.** *For any fixed candidate design (A, ϱ_A), if $|\mathcal{I}_E| \to \infty$ as $n \to \infty$, then $\widehat{\overline{c}}_A(\varrho_A) \to_p \overline{c}_A(\varrho_A)$ as $n \to \infty$.*

Proof. Since $\{\mathbf{X}_j : j \in \mathcal{I}_E\}$ are i.i.d. draws from P_E^X, $|\mathcal{I}_E| \to \infty$, and the cost functions are bounded, the law of large numbers gives $\widehat{\overline{c}}_A(\varrho_A) \to_p \overline{c}_A(\varrho_A)$. □

Theorem A.1. *Suppose Assumptions 1 and A.1 hold, $\lambda_{P,b} \to 0$, and $\lambda_{P,b} b \to \infty$. Then, as $n \to \infty$,*

$$\sqrt{n}\{\widehat{\tau}(A, \varrho_A) - \tau\} \to_d \mathcal{N}\big(0, \sigma_h^2(A, \varrho_A)\big), \qquad \frac{\widehat{\sigma}_h^2(A, \varrho_A)}{\sigma_h^2(A, \varrho_A)} \to_p 1.$$

Corollary A.1. *Under the conditions of Theorem A.1, as $n \to \infty$,*

$$\sqrt{n}\{\widehat{\tau}^*(A, \varrho_A) - \tau\} \to_d \mathcal{N}\big(0, \sigma_h^2(A, \varrho_A)\big).$$

Corollary A.1 follows directly from Theorem A.1, and its proof is therefore omitted.

A.2 Proof of Theorem A.1

Fix a candidate design (A, ϱ_A), and write $n_M = n_M(A, \varrho_A)$, $P_E = P_{E,A}^{\varrho_A}$, $\varrho_{1,A}(\mathbf{X}) = \varrho_A(\mathbf{X})$,, and $\varrho_{0,A}(\mathbf{X}) = 1 - \varrho_A(\mathbf{X})$. By the uniform overlap condition imposed on $\mathcal{R}_A$, both assignment probabilities are bounded away from zero. By the definition of the implemented sample size, $n_M(A, \varrho_A) = \left\lfloor (1 - \lambda_{P,b}) b / \widehat{\overline{c}}_A(\varrho_A) \right\rfloor$.

The cost consistency condition in Assumption A.1 implies $\widehat{\overline{c}}_A(\varrho_A)/\overline{c}_A(\varrho_A) \to_p 1$. Since the integer rounding error is negligible,

$$\frac{n_M(A, \varrho_A)}{(1 - \lambda_{P,b}) b} \to_p \frac{1}{\overline{c}_A(\varrho_A)}.$$

Therefore,

$$\frac{n}{n_M(A, \varrho_A)} \to_p h\bar{c}_A(\varrho_A).$$

Define the sigma-field containing the information used for design selection and sample size determination:

$$\mathcal{F}_C = \sigma(\mathcal{D}_E^X, \mathcal{F}_{\text{pilot}}).$$

which contains the information used to determine the implemented design and sample size. Define

$$\phi_{O,A}(\mathbf{Z}_{O,A}) = \psi_{O,A}(\mathbf{Z}_{O,A}; \tau, \boldsymbol{\eta}_A), \qquad \phi_{E,A}(\mathbf{Z}_{E,A}^{\text{main}}) = \psi_{E,A}\left(\mathbf{Z}_{E,A}^{\text{main}}; \boldsymbol{\eta}_A, \varrho_A\right).$$

Then

$$\mathbb{E}_{P_O}(\phi_{O,A}) = 0, \qquad \mathbb{E}_{P_E}(\phi_{E,A}) = 0,$$

and

$$\mathbb{E}_{P_O}(\phi_{O,A}^2) = V_O(A), \qquad \mathbb{E}_{P_E}(\phi_{E,A}^2) = V_E(A, \varrho_A).$$

The following bias and continuity bounds control the effect of nuisance estimation.

Lemma A.2. *Let $\widetilde{\boldsymbol{\eta}}_A$ be a nuisance-function vector as defined in equation A.2. Suppose $\epsilon \leq \rho_{w,A}, \widetilde{\rho}_{w,A} \leq 1 - \epsilon$ and $\underline{r} \leq r, \widetilde{r} \leq \overline{r}$. Then*

$$\left|\mathbb{E}_{P_O}\{\psi_{O,A}(\tau, \widetilde{\boldsymbol{\eta}}_A)\} + \mathbb{E}_{P_E}\{\psi_{E,A}(\widetilde{\boldsymbol{\eta}}_A, \varrho_A)\}\right| \leq C \left\|\widetilde{\boldsymbol{\mu}}_A - \boldsymbol{\mu}_A\right\|_{\mu,2} \left\|\widetilde{\boldsymbol{\kappa}}_A - \boldsymbol{\kappa}_A\right\|_{\kappa,2}, \tag{A.6}$$

where $C < \infty$ may depend on A, ϱ_A, and the clipping bounds but not on $\widetilde{\boldsymbol{\eta}}_A$.

Proof. Define $\mathcal{M}_A(\widetilde{\boldsymbol{\mu}}_A, \widetilde{\boldsymbol{\kappa}}_A) = \mathbb{E}_{P_O}\{\psi_{O,A}(\tau, \widetilde{\boldsymbol{\eta}}_A)\} + \mathbb{E}_{P_E}\{\psi_{E,A}(\widetilde{\boldsymbol{\eta}}_A, \varrho_A)\}$. The ESC moment and its double-robust structure imply $\mathcal{M}_A(\boldsymbol{\mu}_A, \boldsymbol{\kappa}_A) = \mathcal{M}_A(\widetilde{\boldsymbol{\mu}}_A, \boldsymbol{\kappa}_A) = \mathcal{M}_A(\boldsymbol{\mu}_A, \widetilde{\boldsymbol{\kappa}}_A) = 0$.

Let $\Delta\mu_{w,A} = \widetilde{\mu}_{w,A} - \mu_{w,A}$ and $\Delta\overline{\mu}_{w,A} = \widetilde{\overline{\mu}}_{w,A} - \overline{\mu}_{w,A}$. Subtracting the preceding zero-valued moments gives

$$\begin{aligned}
&\mathcal{M}_A(\widetilde{\boldsymbol{\mu}}_A, \widetilde{\boldsymbol{\kappa}}_A) \\
&= \sum_{w=0}^{1}(2w-1)\mathbb{E}_{P_O}\left[\mathbb{1}\{W = w\}\Delta\mu_{w,A}\left(\frac{1}{\rho_{w,A}} - \frac{1}{\widetilde{\rho}_{w,A}}\right)\right] \\
&\quad + \sum_{w=0}^{1}(2w-1)\mathbb{E}_{P_E}\left[(\widetilde{r} - r)\frac{\mathbb{1}\{W = w\}}{\varrho_{w,A}(\mathbf{X})}\left(\Delta\mu_{w,A} - \Delta\overline{\mu}_{w,A}\right)\right].
\end{aligned}$$

The weighting bounds imply $\left|\rho_{w,A}^{-1} - \widetilde{\rho}_{w,A}^{-1}\right| \leq C\left|\widetilde{\rho}_{w,A} - \rho_{w,A}\right|$, while $\varrho_{w,A}^{-1}$ is bounded by the uniform overlap condition imposed on $\mathcal{R}_A$. The result follows from the Cauchy-Schwarz inequality and the norm definitions in equation A.5. □

Lemma A.3. *Under Assumption A.1(i), suppose a possibly random nuisance collection $\widetilde{\boldsymbol{\eta}}_A$ satisfies the weighting bounds in Lemma A.2 and $\left\|\widetilde{\boldsymbol{\eta}}_A - \boldsymbol{\eta}_A\right\|_{\eta,2} = o_p(1)$. Then*

$$\left\|\psi_{O,A}(\tau, \widetilde{\boldsymbol{\eta}}_A) - \psi_{O,A}(\tau, \boldsymbol{\eta}_A)\right\|_{P_O,2} + \left\|\psi_{E,A}(\widetilde{\boldsymbol{\eta}}_A, \varrho_A) - \psi_{E,A}(\boldsymbol{\eta}_A, \varrho_A)\right\|_{P_E,2} = o_p(1).$$

Proof. For $w \in \{0,1\}$, define $R_{O,w} = \mathbb{1}\{W = w\}\{Y - \mu_{w,A}(\mathbf{S}_A, \mathbf{X})\}/\rho_{w,A}(\mathbf{S}_A, \mathbf{X})$, and $B_A(\mathbf{X}) = \overline{\mu}_{1,A}(\mathbf{X}) - \overline{\mu}_{0,A}(\mathbf{X}) - \tau$. The observational oracle score satisfies $\phi_{O,A} = R_{O,1} - R_{O,0} + B_A$. Since the residual terms have conditional mean zero given $\mathbf{X}$, $B_A(\mathbf{X}) = \mathbb{E}_{P_O}(\phi_{O,A} \mid \mathbf{X})$. Conditional Jensen's inequality, mutual exclusivity of the treatment arms, and Assumption A.1(i) imply

$$\max_{w\in\{0,1\}} \|R_{O,w}\|_{P_O,q} \leq 2C. \tag{A.7}$$

Similarly, define $R_{E,w} = r(\mathbf{X})\mathbb{1}\{W = w\}\varrho_{w,A}(\mathbf{X})^{-1}\left\{\mu_{w,A}(\mathbf{S}_A, \mathbf{X}) - \overline{\mu}_{w,A}(\mathbf{X})\right\}$. Since $\phi_{E,A} = R_{E,1} - R_{E,0}$, mutual exclusivity and Assumption A.1(i) give

$$\max_{w\in\{0,1\}} \|R_{E,w}\|_{P_E,q} \leq C. \tag{A.8}$$

For the observational score, the armwise perturbation satisfies

$$\begin{aligned}
&\mathbb{1}\{W=w\}\left(\frac{Y-\widetilde{\mu}_{w,A}}{\widetilde{\rho}_{w,A}}-\frac{Y-\mu_{w,A}}{\rho_{w,A}}\right)\\
=&-\frac{\mathbb{1}\{W=w\}(\widetilde{\mu}_{w,A}-\mu_{w,A})}{\widetilde{\rho}_{w,A}}\\
&+\mathbb{1}\{W=w\}(Y-\mu_{w,A})\left(\frac{1}{\widetilde{\rho}_{w,A}}-\frac{1}{\rho_{w,A}}\right).
\end{aligned}\tag{A.9}$$

The first term is $o_p(1)$ in $L_2(P_O)$ by nuisance consistency and bounded inverse weights.

Let $a = 2q/(q-2)$, so that $1/2 = 1/q + 1/a$. Because the clipped weighting differences are uniformly bounded and converge to zero in L_2, interpolation gives convergence to zero in L_a. Hölder's inequality and equation A.7 then show that the second term in equation A.9 is $o_p(1)$ in $L_2(P_O)$. The perturbations involving $\overline{\mu}_{w,A}$ are controlled by equation A.5.

For the experimental score,

$$\begin{aligned}
&\widetilde{r}\frac{\mathbb{1}\{W=w\}}{\varrho_{w,A}}\left(\widetilde{\mu}_{w,A}-\widetilde{\overline{\mu}}_{w,A}\right)-r\frac{\mathbb{1}\{W=w\}}{\varrho_{w,A}}\left(\mu_{w,A}-\overline{\mu}_{w,A}\right)\\
=&\widetilde{r}\frac{\mathbb{1}\{W=w\}}{\varrho_{w,A}}\left\{(\widetilde{\mu}_{w,A}-\mu_{w,A})-(\widetilde{\overline{\mu}}_{w,A}-\overline{\mu}_{w,A})\right\}\\
&+(\widetilde{r}-r)\frac{\mathbb{1}\{W=w\}}{\varrho_{w,A}}\left(\mu_{w,A}-\overline{\mu}_{w,A}\right).
\end{aligned}\tag{A.10}$$

The first term is $o_p(1)$ in $L_2(P_E)$ by nuisance consistency and bounded weights. The lower bound on r, together with equation A.8, gives $\left\|\mathbb{1}\{W=w\}\varrho_{w,A}^{-1}\left(\mu_{w,A}-\overline{\mu}_{w,A}\right)\right\|_{P_E,q} \le C$. The same interpolation and Hölder argument therefore controls the second term in equation A.10. Summing over the two arms proves the claimed L_2-continuity. □

Proof of Theorem A.1. Let $I_{O,k}$ and $I_{E,k}$ denote the observational and experimental evaluation folds, and write $\widehat{\boldsymbol{\eta}}_{A,k} = \widehat{\boldsymbol{\eta}}_A^{\text{main},(-k)}$.

The weighting components of $\widehat{\boldsymbol{\eta}}_{A,k}$ satisfy the bounds in Lemma A.2 by equation C.37 and equation C.38. Assumption A.1(ii) ensures that clipping does not alter the population nuisance functions.

For clarity, suppose first that the folds have equal sizes. For approximately equal folds whose sizes differ by at most one, replacing fold-size-weighted averages by the K_M^{-1}-weighted averages in equation 9 contributes $O_p(n^{-1}+n_M^{-1}) = o_p(n^{-1/2})$, because n/n_M converges to a finite positive constant.

Since the observational score is affine in τ with derivative -1, the estimating equation gives

$$\widehat{\tau}(A,\varrho_A)-\tau=\frac{1}{K_M}\sum_{k=1}^{K_M}\left[\mathbb{E}_{\widehat{P}^{\text{main}}_{O,A,k}}\left\{\psi_{O,A}\left(\tau,\widehat{\boldsymbol{\eta}}_{A,k}\right)\right\}+\mathbb{E}_{\widehat{P}^{\text{main}}_{E,A,k}}\left\{\psi_{E,A}\left(\widehat{\boldsymbol{\eta}}_{A,k},\varrho_A\right)\right\}\right]+o_p(n^{-1/2}).$$

Define

$$\Delta_{O,k}=\psi_{O,A}\left(\tau,\widehat{\boldsymbol{\eta}}_{A,k}\right)-\phi_{O,A},\qquad \Delta_{E,k}=\psi_{E,A}\left(\widehat{\boldsymbol{\eta}}_{A,k},\varrho_A\right)-\phi_{E,A}.$$

Adding and subtracting their population expectations yields

$$\begin{aligned}
\widehat{\tau}(A,\varrho_A)-\tau=&\mathbb{E}_{\widehat{P}_{O,A}}(\phi_{O,A})+\mathbb{E}_{\widehat{P}^{\text{main}}_{E,A}}(\phi_{E,A})\\
&+\frac{1}{K_M}\sum_{k=1}^{K_M}\left\{\mathbb{E}_{\widehat{P}^{\text{main}}_{O,A,k}}(\Delta_{O,k})-\mathbb{E}_{P_O}(\Delta_{O,k})+\mathbb{E}_{\widehat{P}^{\text{main}}_{E,A,k}}(\Delta_{E,k})-\mathbb{E}_{P_E}(\Delta_{E,k})\right\}\\
&+\frac{1}{K_M}\sum_{k=1}^{K_M}\left\{\mathbb{E}_{P_O}(\Delta_{O,k})+\mathbb{E}_{P_E}(\Delta_{E,k})\right\}+o_p(n^{-1/2}).
\end{aligned}$$

Conditional on $\mathcal{F}_C$ and the training samples, cross-fitting makes the evaluation observations independent of the corresponding nuisance estimates, and the implemented sample size $n_M(A, \varrho_A)$ is fixed. Conditional Chebyshev inequalities and Lemma A.3 therefore give

$$\mathbb{E}_{\widehat{P}^{\text{main}}_{O,A,k}}(\Delta_{O,k}) - \mathbb{E}_{P_O}(\Delta_{O,k}) = o_p(n^{-1/2}), \qquad \mathbb{E}_{\widehat{P}^{\text{main}}_{E,A,k}}(\Delta_{E,k}) - \mathbb{E}_{P_E}(\Delta_{E,k}) = o_p(n_M^{-1/2}).$$

Since $n/n_M(A, \varrho_A) \to_p h\bar{c}_A(\varrho_A)$, we have $n_M \asymp n$ in probability. The entire empirical-process term is $o_p(n^{-1/2})$.

Lemma A.2 and Assumption A.1(iii) imply, uniformly over $k \leq K_M$, $|\mathbb{E}_{P_O}(\Delta_{O,k}) + \mathbb{E}_{P_E}(\Delta_{E,k})| = o_p(n^{-1/2})$. Consequently,

$$\widehat{\tau}(A, \varrho_A) - \tau = \frac{1}{n}\sum_{i\in\mathcal{I}_O} \phi_{O,A}(\mathbf{Z}_{O,A,i}) + \frac{1}{n_M}\sum_{j\in\mathcal{I}_M} \phi_{E,A}(\mathbf{Z}^{\text{main}}_{E,A,j}) + o_p(n^{-1/2}). \tag{A.11}$$

The two leading averages equal $\widehat{\tau}^*(A, \varrho_A) - \tau$. Thus, $\sqrt{n}\{\widehat{\tau}(A, \varrho_A) - \widehat{\tau}^*(A, \varrho_A)\} = o_p(1)$. Multiplying equation A.11 by $\sqrt{n}$, the leading term becomes $n^{-1/2}\sum_{i\in\mathcal{I}_O} \phi_{O,A}(\mathbf{Z}_{O,A,i}) + n_M^{-1}\sqrt{n}\sum_{j\in\mathcal{I}_M} \phi_{E,A}(\mathbf{Z}^{\text{main}}_{E,A,j})$. Conditional on the information used to determine the implemented design and sample size, the variance of the normalized estimator is

$$V_O(A) + \frac{n}{n_M(A, \varrho_A)} V_E(A, \varrho_A).$$

By the previous argument,

$$V_O(A) + \frac{n}{n_M(A, \varrho_A)} V_E(A, \varrho_A) \to_p \sigma_h^2(A, \varrho_A).$$

To verify the central limit theorem, define the two triangular-array summands $X_{n,i} = n^{-1/2}\phi_{O,A}(\mathbf{Z}_{O,A,i})$ and $Y_{n,j} = n_M^{-1}\sqrt{n}\phi_{E,A}(\mathbf{Z}^{\text{main}}_{E,A,j})$. By Assumption A.1(i),

$$\sum_{i\in\mathcal{I}_O} \mathbb{E}|X_{n,i}|^q + \sum_{j\in\mathcal{I}_M} \mathbb{E}|Y_{n,j}|^q \leq C\left(n^{1-q/2} + n^{q/2}n_M^{1-q}\right) = o(1),$$

where the last equality uses $n_M \asymp n$ and $q > 2$. Conditional on $\mathcal{F}_C$, the experimental observations remain i.i.d. from $P^{\varrho_A}_{E,A}$, and $n_M(A, \varrho_A)$ is treated as fixed. The moment assumptions in Assumption A.1(i) therefore imply the Lyapunov condition for the triangular array with sample size $n_M(A, \varrho_A)$. Conditional on $\mathcal{F}$, the conditional weak convergence gives

$$\mathcal{L}\left(\sqrt{n}(\widehat{\tau} - \tau) \mid \mathcal{F}\right) \rightsquigarrow_p N(0, \sigma_h^2(A, \varrho_A)).$$

and the unconditional result follows. Hence $\sqrt{n}\left\{\widehat{\tau}(A, \varrho_A) - \tau\right\} \to_d \mathcal{N}\big(0, \sigma_h^2(A, \varrho_A)\big)$.

It remains to establish consistency of the variance estimator. First, equation A.11 implies $\widehat{\tau}(A, \varrho_A) - \tau = o_p(1)$. Let $\widehat{\psi}^{\text{main}}_{O,i}$ and $\widehat{\psi}^{\text{main}}_{E,j}$ denote the cross-fitted estimates of the EIF components for this fixed pair, evaluated at $\widehat{\tau}(A, \varrho_A)$ and the corresponding nuisance fits, as in Appendix C.2.2.

For $i \in I_{O,k}$, define $\widetilde{\psi}^{\text{main}}_{O,i} = \psi_{O,A}\left(\mathbf{Z}_{O,A,i}; \tau, \widehat{\boldsymbol{\eta}}_{A,k}\right)$. Because the observational score is affine in τ, $\widehat{\psi}^{\text{main}}_{O,i} - \phi_{O,A}(\mathbf{Z}_{i,O,A}) = \widetilde{\psi}^{\text{main}}_{O,i} - \phi_{O,A}(\mathbf{Z}_{i,O,A}) - \{\widehat{\tau}(A, \varrho_A) - \tau\}$. Lemma A.3, cross-fitting, and conditional Markov inequalities give $n^{-1}\sum_{i\in\mathcal{I}_O}\left\{\widehat{\psi}^{\text{main}}_{O,i} - \phi_{O,A}(\mathbf{Z}_{i,O,A})\right\}^2 = o_p(1)$. The same argument gives $n_M^{-1}\sum_{j\in\mathcal{I}_M}\left\{\widehat{\psi}^{\text{main}}_{E,j} - \phi_{E,A}(\mathbf{Z}^{\text{main}}_{E,A,j})\right\}^2 = o_p(1)$.

The Cauchy-Schwarz inequality, the laws of large numbers for the oracle scores, and the preceding approximations together imply that

$$\widehat{V}^{\text{main}}_O \to_p V_O(A), \qquad \widehat{V}^{\text{main}}_E \to_p V_E(A, \varrho_A). \tag{A.12}$$

Using equation 10, equation A.12, and Lemma A.1, we have $n/n_M(A, \varrho_A) \to_p h\bar{c}_A(\varrho_A)$.

$$\begin{aligned}
&\widehat{\sigma}_h^2(A, \varrho_A) - \sigma_h^2(A, \varrho_A) \\
&= \{\widehat{V}_O^{\text{main}}(A) - V_O(A)\} \\
&\quad + \frac{n}{n_M(A, \varrho_A)}\{\widehat{V}_E^{\text{main}}(A, \varrho_A) - V_E(A, \varrho_A)\} \\
&\quad + \left\{\frac{n}{n_M(A, \varrho_A)} - h\bar{c}_A(\varrho_A)\right\} V_E(A, \varrho_A).
\end{aligned}$$

Since Assumption A.1(i) implies $\sigma_h^2(A, \varrho_A) \geq V_O(A) \geq c$, $\widehat{\sigma}_h^2(A, \varrho_A)/\sigma_h^2(A, \varrho_A) \to_p 1$. □

B Proofs of main theoretical results

B.1 Proof of Theorem 1

Under the conditions of Theorem 1, recall that $Q_A(\varrho_A) = \bar{c}_A(\varrho_A)V_E(A, \varrho_A)$ and $\widehat{Q}_A(\varrho_A) = \widehat{\bar{c}}_A(\varrho_A)\widehat{V}_E^{\text{pilot}}(A, \varrho_A)$. By equation 5 and equation 7, $V_n(A, \varrho_A) = V_O(A) + h_n Q_A(\varrho_A)$, and $\widehat{V}_n^{\text{pilot}}(A, \varrho_A) = \widehat{V}_O^{\text{pilot}}(A) + h_n\widehat{Q}_A(\varrho_A)$.

Lemma B.4 (Consistency of the estimated design criterion)**.** *Under the conditions of Theorem 1,*

$$\max_{A\in\mathcal{A}}\left|\widehat{V}_O^{\text{pilot}}(A) - V_O(A)\right| = o_p(1), \tag{B.13}$$

$$\max_{A\in\mathcal{A}}\sup_{\varrho_A\in\mathcal{R}_A}\left|\widehat{\bar{c}}_A(\varrho_A) - \bar{c}_A(\varrho_A)\right| = o_p(1), \tag{B.14}$$

and

$$\max_{A\in\mathcal{A}}\sup_{\varrho_A\in\mathcal{R}_A}\left|\widehat{V}_E^{\text{pilot}}(A, \varrho_A) - V_E(A, \varrho_A)\right| = o_p(1). \tag{B.15}$$

Consequently,

$$\max_{A\in\mathcal{A}}\sup_{\varrho_A\in\mathcal{R}_A}\left|\widehat{Q}_A(\varrho_A) - Q_A(\varrho_A)\right| = o_p(1). \tag{B.16}$$

Proof. Because $\mathcal{A}$ and K_P are finite, it suffices to prove the claims for a fixed $A \in \mathcal{A}$ and $k \leq K_P$.

For $w \in \{0, 1\}$, define $U_{w,A}(\mathbf{s}, \mathbf{x}) = r(\mathbf{x})\left\{\mu_{w,A}(\mathbf{s}, \mathbf{x}) - \overline{\mu}_{w,A}(\mathbf{x})\right\}$, and abbreviate $U_{w,A} = U_{w,A}(\mathbf{S}_A, \mathbf{X})$ when evaluated on a pilot unit.

Under a candidate treatment-assignment rule ϱ_A, treatment-arm mutual exclusivity and equation 2 together give that

$$\begin{aligned}
\mathbb{E}_{P_{E,A}^{\varrho_A}}\{|\psi_{E,A}(\,\cdot\,; \boldsymbol{\eta}_A, \varrho_A)|^q\} =& \mathbb{E}_{P_*}[\varrho_A(\mathbf{X})^{1-q}|U_{1,A}\big(\mathbf{S}_A(1), \mathbf{X}\big)|^q \\
&+ \{1 - \varrho_A(\mathbf{X})\}^{1-q}|U_{0,A}\big(\mathbf{S}_A(0), \mathbf{X}\big)|^q \mid G = E].
\end{aligned}$$

Under the balanced pilot assignment,

$$\mathbb{E}_{P_{E,A}^{\varrho_{\text{pilot}}}}[\mathbb{1}\{W = 1\}|U_{1,A}|^q] = \mathbb{E}_{P_*}\left\{\left|U_{1,A}\big(\mathbf{S}_A(1), \mathbf{X}\big)\right|^q \mid G = E\right\}/2,$$

and the analogous identity holds for the control arm. Assumption 2(ii) therefore implies

$$\max_{A\in\mathcal{A},\, w\in\{0,1\}} \|\mathbb{1}\{W = w\}U_{w,A}\|_{P_{E,A}^{\varrho_{\text{pilot}}},q} \leq C. \tag{B.17}$$

For an evaluation fold k, define $\widehat{U}_{w,A,k} = \widehat{r}^{\text{pilot},(-k)}(\mathbf{X})\left\{\widehat{\mu}_{w,A}^{\text{pilot},(-k)}(\mathbf{S}_A, \mathbf{X}) - \widehat{\overline{\mu}}_{w,A}^{\text{pilot},(-k)}(\mathbf{X})\right\}$. Assumption 2(i) and (v), the score perturbation identities in equation A.9 and equation A.10, and equation B.17 imply

$$\begin{aligned}
&\max_{A\in\mathcal{A},\, k\leq K_P}\left\|\psi_{O,A}\left(\,\cdot\,; \tau, \widehat{\boldsymbol{\eta}}_A^{\text{pilot},(-k)}\right) - \psi_{O,A}\left(\,\cdot\,; \tau, \boldsymbol{\eta}_A\right)\right\|_{P_O,2} = o_p(1), \\
&\max_{A\in\mathcal{A},\, k\leq K_P, w\in\{0,1\}}\left\|\mathbb{1}\{W = w\}\left(\widehat{U}_{w,A,k} - U_{w,A}\right)\right\|_{P_{E,A}^{\varrho_{\text{pilot}}},2} = o_p(1).
\end{aligned} \tag{B.18}$$

Indeed, if $a = 2q/(q-2)$, $1/2 = 1/q + 1/a$, then interpolation between L_2 and L_∞ gives convergence of the clipped weighting errors in L_a. Hölder's inequality and the L_q bounds above then establish equation B.18.

Use the pilot target estimate $\widehat{\tau}_A^{\text{pilot}}$ and the cross-fitted observational components $\widehat{\psi}_{O,A,i}^{\text{pilot}}$ from Appendix C.2. An overbar denotes their sample average. We first prove equation B.13. For $i \in I_{O,k}^P$, define $\widetilde{\psi}_{O,A,i}^{\text{pilot}} = \psi_{O,A}\left(Y_i, W_i, \mathbf{S}_{i,A}, \mathbf{X}_i; \tau, \widehat{\boldsymbol{\eta}}_A^{\text{pilot},(-k)}\right)$. Because the observational score is affine in its target argument, $\widehat{\psi}_{O,A,i}^{\text{pilot}} = \widetilde{\psi}_{O,A,i}^{\text{pilot}} - \left(\widehat{\tau}_A^{\text{pilot}} - \tau\right)$. The last term remains the same for all observations and disappears under empirical centering: $\widehat{\psi}_{O,A,i}^{\text{pilot}} - \overline{\widehat{\psi}}_{O,A}^{\text{pilot}} = \widetilde{\psi}_{O,A,i}^{\text{pilot}} - \overline{\widetilde{\psi}}_{O,A}^{\text{pilot}}$.

Conditional on the corresponding training observations, cross-fitting makes the evaluation observations independent of the fold-specific nuisance estimates. Therefore, equation B.18 and conditional Markov inequalities imply that replacing the oracle observational score by the cross-fitted score changes its empirical first and second moments by $o_p(1)$. The law of large numbers and the definition of $V_O(A)$ then give $\widehat{V}_O^{\text{pilot}}(A) = V_O(A) + o_p(1)$. Finiteness of $\mathcal{A}$ proves equation B.13.

The expected cost per unit conditional on $\mathbf{X}$ is $c_0(A, \mathbf{X}) + \{c_1(A, \mathbf{X}) - c_0(A, \mathbf{X})\}\varrho_A(\mathbf{X})$. For each $A \in \mathcal{A}$, the VC-subgraph condition on $\mathcal{R}_A$ and boundedness of the cost functions imply that the corresponding class is Glivenko–Cantelli under P_E^X. Since $\mathcal{A}$ is finite, $\{\mathbf{X}_j : j \in \mathcal{I}_E\}$ are i.i.d. from P_E^X, and $|\mathcal{I}_E| \to \infty$, the uniform law of large numbers gives equation B.14.

It remains to establish equation B.15. Under the balanced pilot randomization, define $H_{A,\varrho_A} = 2WU_{1,A}^2/\varrho_A(\mathbf{X}) + 2(1-W)U_{0,A}^2/\{1-\varrho_A(\mathbf{X})\}$. A change of assignment measure gives $\mathbb{E}_{P_{E,A}^{\varrho_{\text{pilot}}}}(H_{A,\varrho_A}) = \mathbb{E}_{P_{E,A}^{\varrho_A}}\left\{\psi_{E,A}^2(\,\cdot\,; \boldsymbol{\eta}_A, \varrho_A)\right\} = V_E(A, \varrho_A)$. By the uniform overlap condition imposed on $\mathcal{R}_A$, $\sup_{\varrho_A \in \mathcal{R}_A} |H_{A,\varrho_A}| \le 2\epsilon_\varrho^{-1}\left\{WU_{1,A}^2 + (1-W)U_{0,A}^2\right\}$. equation B.17 implies that this envelope belongs to $L_{q/2}$, where $q/2 > 1$. The inverse maps appearing in H_{A,ϱ_A} are Lipschitz on $[\epsilon_\varrho, 1-\epsilon_\varrho]$. Hence $\{H_{A,\varrho_A} : \varrho_A \in \mathcal{R}_A\}$ is a measurable Glivenko-Cantelli class, and $\sup_{\varrho_A \in \mathcal{R}_A}\left|\mathbb{E}_{\widehat{P}_{E,A_P}^{\text{pilot}}}(H_{A,\varrho_A}) - V_E(A, \varrho_A)\right| = o_p(1)$.

To bound the error from nuisance-function estimation, define $D_{A,k} = \epsilon_\varrho^{-1}\sum_{w=0}^{1}\mathbb{1}\{W = w\}\left|\widehat{U}_{w,A,k} - U_{w,A}\right|$. The second relation in equation B.18 gives

$$\max_{A \in \mathcal{A},\, k \le K_P} \|D_{A,k}\|_{P_{E,A}^{\varrho_{\text{pilot}}},2} = o_p(1).$$

For fold k, write $\widehat{\psi}_{E,A,k}^{\text{pilot}}(\,\cdot\,; \varrho_A) = \psi_{E,A}(\,\cdot\,; \widehat{\boldsymbol{\eta}}_A^{\text{pilot},(-k)}, \varrho_A)$. Uniformly over $\varrho_A \in \mathcal{R}_A$, we have $\left|\widehat{\psi}_{E,A,k}^{\text{pilot}}(\varrho_A) - \psi_{E,A}(\boldsymbol{\eta}_A, \varrho_A)\right| \le D_{A,k}$. Moreover, $\sup_{\varrho_A \in \mathcal{R}_A}|\psi_{E,A}(\boldsymbol{\eta}_A, \varrho_A)| \le \epsilon_\varrho^{-1}\{W|U_{1,A}| + (1-W)|U_{0,A}|\} =: G_A$, where $\|G_A\|_{P_{E,A}^{\varrho_{\text{pilot}}},2} \le C$.

For a generic pilot observation $(W, \mathbf{X}, \mathbf{S}_A)$, define $\omega_A^{\text{pilot}}(\varrho_A) = 2[W\varrho_A(\mathbf{X}) + (1-W)\{1-\varrho_A(\mathbf{X})\}]$. These importance weights are uniformly bounded over $\varrho_A \in \mathcal{R}_A$. Using $|a^2 - b^2| \le |a-b|(|a|+|b|)$, the Cauchy-Schwarz inequality, and conditional Markov inequalities, we obtain

$$\max_{A \in \mathcal{A},\, k \le K_P} \sup_{\varrho_A \in \mathcal{R}_A} \mathbb{E}_{\widehat{P}_{E,A_P,k}^{\text{pilot}}}\left\{\omega_A^{\text{pilot}}(\varrho_A)\left|\widehat{\psi}_{E,A,k}^{\text{pilot}}(\varrho_A)^2 - \psi_{E,A}(\boldsymbol{\eta}_A, \varrho_A)^2\right|\right\} = o_p(1).$$

Combining this relation with the preceding uniform law of large numbers proves equation B.15.

Finally, the cost and experimental variance components are uniformly bounded under Assumption 2(ii). Since

$$\widehat{Q}_A(\varrho_A) - Q_A(\varrho_A) = \left\{\widehat{\overline{c}}_A(\varrho_A) - \overline{c}_A(\varrho_A)\right\}\widehat{V}_E^{\text{pilot}}(A, \varrho_A) + \overline{c}_A(\varrho_A)\left\{\widehat{V}_E^{\text{pilot}}(A, \varrho_A) - V_E(A, \varrho_A)\right\},$$

equations equation B.14 and equation B.15 prove equation B.16. □

Lemma B.5. *Under the conditions of Theorem 1,*

$$\max_{A \in \mathcal{A}} \sup_{\varrho_A \in \mathcal{R}_A}\left|\widehat{V}_n^{\text{pilot}}(A, \varrho_A) - \sigma_h^2(A, \varrho_A)\right| = o_p(1).$$

Proof. By equation 7, equation 5, and equation 3,

$$\left|\widehat{V}_n^{\text{pilot}}(A,\varrho_A)-\sigma_h^2(A,\varrho_A)\right| \le \left|\widehat{V}_O^{\text{pilot}}(A)-V_O(A)\right|+h_n\left|\widehat{Q}_A(\varrho_A)-Q_A(\varrho_A)\right|+|h_n-h|\,Q_A(\varrho_A).$$

Assumption 2(ii) implies $\max_{A\in\mathcal{A}}\sup_{\varrho_A\in\mathcal{R}_A} Q_A(\varrho_A)<\infty$. Since $h_n\to h$, the sequence h_n is bounded. The result therefore follows from equation B.13, equation B.16, and equation 6. □

Lemma B.6 (Consistency of the treatment-assignment rules)**.** *Under the conditions of Theorem 1,*

$$\max_{A\in\mathcal{A}}\|\widehat{\varrho}_A-\varrho_A^*\|_{P_E^X,2}=o_p(1). \tag{B.19}$$

Proof. Fix $A\in\mathcal{A}$, and define $s_{A,b}=\sup_{\varrho_A\in\mathcal{R}_A}\left|\widehat{Q}_A(\varrho_A)-Q_A(\varrho_A)\right|$. By equation B.16, $\max_{A\in\mathcal{A}} s_{A,b}=o_p(1)$. Assumption 2(iii) implies $Q_A(\widehat{\varrho}_A)-Q_A(\varrho_A^*)\le 2s_{A,b}+\xi_{A,b}$.

For every $\varepsilon>0$, compactness of $\mathcal{R}_A$, continuity of Q_A, and uniqueness of ϱ_A^* imply

$$\Delta_A(\varepsilon):=\inf_{\varrho_A\in\mathcal{R}_A,\|\varrho_A-\varrho_A^*\|_{P_E^X,2}\ge\varepsilon}\{Q_A(\varrho_A)-Q_A(\varrho_A^*)\}>0.$$

Consequently, $\mathbb{P}\left(\|\widehat{\varrho}_A-\varrho_A^*\|_{P_E^X,2}\ge\varepsilon\right)\le\mathbb{P}\left(2s_{A,b}+\xi_{A,b}\ge\Delta_A(\varepsilon)\right)\to 0$. Because $\mathcal{A}$ is finite and $\max_A\xi_{A,b}=o_p(1)$, the convergence is uniform over $A\in\mathcal{A}$, proving equation B.19. □

Proof of Theorem 1. Let $e_b=\max_{A\in\mathcal{A}}\sup_{\varrho_A\in\mathcal{R}_A}\left|\widehat{V}_n^{\text{pilot}}(A,\varrho_A)-\sigma_h^2(A,\varrho_A)\right|$. Lemma B.5 gives $e_b=o_p(1)$.

By Lemma B.6, continuity of Q_A, and finiteness of $\mathcal{A}$, $\max_{A\in\mathcal{A}}|Q_A(\widehat{\varrho}_A)-Q_A(\varrho_A^*)|=o_p(1)$. Because $\sigma_h^2(A,\varrho_A)=V_O(A)+hQ_A(\varrho_A)$, it follows that $\max_{A\in\mathcal{A}}\left|\sigma_h^2(A,\widehat{\varrho}_A)-\sigma_h^2(A,\varrho_A^*)\right|=o_p(1)$. Therefore,

$$\max_{A\in\mathcal{A}}\left|\widehat{V}_n^{\text{pilot}}\left(A,\widehat{\varrho}_A\right)-\sigma_h^2\left(A,\varrho_A^*\right)\right|=o_p(1). \tag{B.20}$$

Because $\mathcal{A}$ is finite and A^* is the unique minimizer, Assumption 2(iv) implies $\min_{A\in\mathcal{A}\setminus\{A^*\}}\left\{\sigma_h^2(A,\varrho_A^*)-\sigma_h^2(A^*,\varrho_{A^*}^*)\right\}>0$. Together with equation B.20, this implies that, with probability approaching one, $\widehat{V}_n^{\text{pilot}}\left(A,\widehat{\varrho}_A\right)>\widehat{V}_n^{\text{pilot}}\left(A^*,\widehat{\varrho}_{A^*}\right)$, for every $A\neq A^*$. By the definition of $\widehat{A}$ in equation 8,

$$\mathbb{P}(\widehat{A}=A^*)\to 1.$$

Since $\widehat{\varrho}_{\widehat{A}}=\widehat{\varrho}_{A^*}$ on the event $\{\widehat{A}=A^*\}$, equation B.19 further implies

$$\|\widehat{\varrho}_{\widehat{A}}-\varrho_{A^*}^*\|_{P_E^X,2}=o_p(1).$$

This proves the first two conclusions of Theorem 1.

Finally, on $\{\widehat{A}=A^*\}$,

$$\left|\widehat{V}_n^{\text{pilot}}\left(\widehat{A},\widehat{\varrho}_{\widehat{A}}\right)-\sigma_h^2(A^*,\varrho_{A^*}^*)\right|\le e_b+\left|\sigma_h^2\left(A^*,\widehat{\varrho}_{A^*}\right)-\sigma_h^2\left(A^*,\varrho_{A^*}^*\right)\right|=o_p(1).$$

Since $\mathbb{P}(\widehat{A}=A^*)\to 1$, the same convergence holds unconditionally, completing the proof. □

B.2 Proof of Theorem 2

The main experiment depends on data from the pilot stage, including the observational sample reused in the main stage. The proof below controls the observational empirical remainder by stochastic equicontinuity and handles the main experimental sample conditional on the pilot-stage information. For the proof of Theorem 2, write $\varrho^*=\varrho_{A^*}^*$.

Lemma B.7. *Under the conditions of Theorem 2, there exists a deterministic sequence $a_b\downarrow 0$ such that the event*

$$\mathcal{E}_b=\left\{\widehat{A}=A^*,\quad\|\widehat{\varrho}_{A^*}-\varrho^*\|_{P_E^X,2}\le a_b\right\}$$

satisfies $\mathbb{P}(\mathcal{E}_b) \to 1$. *Moreover, on* $\mathcal{E}_b$,

$$\frac{n}{n_M} \to_p h\,\overline{c}_{A^*}(\varrho^*), \tag{B.21}$$

$V_E\left(A^*, \widehat{\varrho}_{A^*}\right) \to_p V_E(A^*, \varrho^*)$, *and*

$$\sigma_h^2(\widehat{A}, \widehat{\varrho}_{\widehat{A}}) \to_p \sigma_h^2(A^*, \varrho^*_{A^*}). \tag{B.22}$$

Proof. Theorem 1 gives $\mathbb{P}(\widehat{A} = A^*) \to 1$ and $\|\widehat{\varrho}_{A^*} - \varrho^*\|_{P_E^X,2} = o_p(1)$. For every random sequence $X_b = o_p(1)$, there exists a deterministic sequence $a_b \downarrow 0$ such that $\mathbb{P}(|X_b| > a_b) \to 0$. Applying this result to the preceding norm proves $\mathbb{P}(\mathcal{E}_b) \to 1$.

By equation B.14, $\sup_{\varrho \in \mathcal{R}_{A^*}} \left|\widehat{\overline{c}}_{A^*}(\varrho) - \overline{c}_{A^*}(\varrho)\right| = o_p(1)$. Moreover, boundedness of the cost functions gives $|\overline{c}_{A^*}(\varrho) - \overline{c}_{A^*}(\varrho^*)| \le \|c_1(A^*, \cdot) - c_0(A^*, \cdot)\|_{P_E^X,2} \|\varrho - \varrho^*\|_{P_E^X,2}$. Consequently, on $\mathcal{E}_b$, $\widehat{\overline{c}}_{A^*}\left(\widehat{\varrho}_{A^*}\right) - \overline{c}_{A^*}(\varrho^*) = o_p(1)$.

By the definition of n_M, we have $n_M = (1 - \lambda_{P,b})b/\widehat{\overline{c}}_{A^*}(\widehat{\varrho}_{A^*}) + O_p(1)$ on $\mathcal{E}_b$. Because the main budget diverges and the expected cost is bounded away from zero, $n_M\,\overline{c}_{A^*}(\widehat{\varrho}_{A^*})/\{(1-\lambda_{P,b})b\} \to_p 1$. It follows from equation 6 that $(n/n_M) = h_n\,\overline{c}_{A^*}\left(\widehat{\varrho}_{A^*}\right) + o_p(1) \to_p h\,\overline{c}_{A^*}(\varrho^*)$, which proves equation B.21. In particular, $n_M \to_p \infty$ and $n_M/n \to_p \{h\,\overline{c}_{A^*}(\varrho^*)\}^{-1}$.

The function $Q_{A^*}(\varrho) = \overline{c}_{A^*}(\varrho)V_E(A^*, \varrho)$ is continuous at ϱ^* by Assumption 2(ii). The expected cost is also continuous there and is bounded away from zero. Hence $V_E(A^*, \varrho) = Q_{A^*}(\varrho)/\overline{c}_{A^*}(\varrho)$ is continuous at ϱ^*. Therefore, $V_E\left(A^*, \widehat{\varrho}_{A^*}\right) \to_p V_E(A^*, \varrho^*)$.

Finally, on $\mathcal{E}_b$, $\sigma_h^2(\widehat{A}, \widehat{\varrho}_{\widehat{A}}) = V_O(A^*) + h\,Q_{A^*}\left(\widehat{\varrho}_{A^*}\right)$. Continuity of Q_{A^*} and convergence of $\widehat{\varrho}_{A^*}$ to ϱ^* in $\|\cdot\|_{P_E^X,2}$ imply that

$$\sigma_h^2(\widehat{A}, \widehat{\varrho}_{\widehat{A}}) \to_p V_O(A^*) + hQ_{A^*}(\varrho^*) = \sigma_h^2(A^*, \varrho^*_{A^*}),$$

proving equation B.22. □

Lemma B.8. *On* $\mathcal{E}_b$ *defined in Lemma B.7, define*

$$\Delta_{O,k} = \psi_{O,A^*}\left(\tau, \widehat{\boldsymbol{\eta}}_{A^*}^{\text{main},(-k)}\right) - \psi_{O,A^*}\left(\tau, \boldsymbol{\eta}_{A^*}\right),$$
$$\Delta_{E,k} = \psi_{E,A^*}\left(\widehat{\boldsymbol{\eta}}_{A^*}^{\text{main},(-k)}, \widehat{\varrho}_{A^*}\right) - \psi_{E,A^*}\left(\boldsymbol{\eta}_{A^*}, \widehat{\varrho}_{A^*}\right).$$

Under the conditions of Theorem 2,

$$\max_{k \le K_M} \|\Delta_{O,k}\|_{P_O,2} = o_p(1), \qquad \max_{k \le K_M} \|\Delta_{E,k}\|_{P^{\widehat{\varrho}_{A^*}}_{E,A^*},2} = o_p(1). \tag{B.23}$$

Proof. On $\mathcal{E}_b$, $\widehat{\varrho}_{A^*} \in \mathcal{R}_{A^*}$, so the uniform overlap bounds apply to the main experimental distribution. The estimated weighting functions satisfy the clipping bounds in equation C.37 and equation C.38.

Assumption 2(ii), conditional Jensen's inequality, and treatment-arm mutual exclusivity give the armwise q-moment bounds in equation A.7 and equation A.8, uniformly over the candidate treatment-assignment rule class.

Let $a = 2q/(q-2)$, $1/2 = 1/q + 1/a$. The clipped weighting errors are uniformly bounded and converge to zero in L_2 under Assumption 3(ii). Interpolation therefore gives convergence in L_a. Applying Hölder's inequality to equation A.9 and equation A.10, together with the regression-nuisance consistency in Assumption 3(ii), proves equation B.23. □

Define the oracle scores

$$\phi_{O,i} = \psi_{O,A^*}\left(Y_i, W_i, \mathbf{S}_{i,A^*}, \mathbf{X}_i; \tau, \boldsymbol{\eta}_{A^*}\right), \qquad \phi_{E,j}(\varrho) = \psi_{E,A^*}\left(W_j, \mathbf{S}_{j,A^*}, \mathbf{X}_j; \boldsymbol{\eta}_{A^*}, \varrho\right).$$

Lemma B.9. *Under the conditions of Theorem 2, on* $\mathcal{E}_b$ *defined in Lemma B.7,*

$$\widehat{\tau} - \tau = \frac{1}{n}\sum_{i \in \mathcal{I}_O} \phi_{O,i} + \frac{1}{n_M}\sum_{j \in \mathcal{I}_M} \phi_{E,j}\left(\widehat{\varrho}_{A^*}\right) + o_p(n^{-1/2}). \tag{B.24}$$

Proof. On $\mathcal{E}_b$, write $\widehat{\boldsymbol{\eta}}_k = \widehat{\boldsymbol{\eta}}_{A^*}^{\text{main},(-k)}$. Approximately equal fold sizes contribute $O_p(n^{-1} + n_M^{-1}) = o_p(n^{-1/2})$, because n/n_M converges in probability to a finite, positive constant by equation B.21.

The estimating equation and the affine dependence of the observational score on τ give

$$\widehat{\tau} - \tau = \frac{1}{K_M} \sum_{k=1}^{K_M} \left[\mathbb{E}_{\widehat{P}^{\text{main}}_{O,A^*,k}} \{\psi_{O,A^*}(\tau, \widehat{\boldsymbol{\eta}}_k)\} + \mathbb{E}_{\widehat{P}^{\text{main}}_{E,A^*,k}} \{\psi_{E,A^*}(\widehat{\boldsymbol{\eta}}_k, \widehat{\varrho}_{A^*})\} \right] + o_p(n^{-1/2}).$$

Adding and subtracting the oracle scores yields

$$\widehat{\tau} - \tau = \frac{1}{n} \sum_{i \in \mathcal{I}_O} \phi_{O,i} + \frac{1}{n_M} \sum_{j \in \mathcal{I}_M} \phi_{E,j}(\widehat{\varrho}_{A^*}) + R_{\text{bias}} + R_{O,\text{emp}} + R_{E,\text{emp}} + o_p(n^{-1/2}),$$

where

$$R_{\text{bias}} = \frac{1}{K_M} \sum_{k=1}^{K_M} \left\{ \mathbb{E}_{P_O}(\Delta_{O,k}) + \mathbb{E}_{P^{\widehat{\varrho}_{A^*}}_{E,A^*}}(\Delta_{E,k}) \right\},$$

$$R_{O,\text{emp}} = \frac{1}{K_M} \sum_{k=1}^{K_M} \left\{ \mathbb{E}_{\widehat{P}^{\text{main}}_{O,A^*,k}}(\Delta_{O,k}) - \mathbb{E}_{P_{O,A^*}}(\Delta_{O,k}) \right\},$$

$$R_{E,\text{emp}} = \frac{1}{K_M} \sum_{k=1}^{K_M} \left\{ eE_{\widehat{P}^{\text{main}}_{E,A^*,k}}(\Delta_{E,k}) - \mathbb{E}_{P^{\widehat{\varrho}_{A^*}}_{E,A^*}}(\Delta_{E,k}) \right\}.$$

Under Assumption 2(v), the second-order bound in equation A.6 is uniform over $\varrho \in \mathcal{R}_{A^*}$. Assumption 3(ii) therefore implies $R_{\text{bias}} = o_p(n^{-1/2})$.

For the observational empirical remainder, the score perturbations belong to differences of functions in the fixed Donsker class $\mathcal{F}_O$ with probability approaching one. The first relation in equation B.23 and stochastic equicontinuity imply $R_{O,\text{emp}} = o_p(n^{-1/2})$. This step does not require the score perturbation to be independent of the observational evaluation fold. Indeed, such independence is generally unavailable here because the learned design $(\widehat{A}, \widehat{\varrho}_{\widehat{A}})$ is learned using the full observational sample $\mathcal{D}_O$, including the k-th evaluation fold, and determines the main experimental data entering $\widehat{\boldsymbol{\eta}}_k$. The stochastic equicontinuity argument controls this remaining dependence directly.

For the experimental empirical remainder, condition on $\mathcal{F}_{\text{pilot}}$, the main-stage fold assignment, and the main experimental observations used to train $\widehat{\boldsymbol{\eta}}_k$. The observations in the experimental evaluation fold remain conditionally independent under $P^{\widehat{\varrho}_{A^*}}_{E,A^*}$. Conditional Chebyshev inequalities and the second relation in equation B.23 give $R_{E,\text{emp}} = o_p(n_M^{-1/2}) = o_p(n^{-1/2})$. Combining the three remainder bounds proves equation B.24. □

Lemma B.10. *Under the conditions of Theorem 2, on $\mathcal{E}_b$ defined in Lemma B.7, we have*

$$\mathcal{L}\left(\frac{n_M^{-1/2} \sum_{j \in I_M} \phi_{E,j}(\widehat{A}^*)}{V_E(A^*, \widehat{A}^*)^{1/2}} \middle| \mathcal{F}_{\text{pilot}} \right) \rightsquigarrow_p N(0,1).$$

Proof. Conditional on $\mathcal{F}_{\text{pilot}}$, $\widehat{\varrho}_{A^*}$ and n_M are fixed, and the main experimental observations are i.i.d. under $P^{\widehat{\varrho}_{A^*}}_{E,A^*}$. Their oracle scores have conditional mean zero and conditional variance $V_E(A^*, \widehat{\varrho}_{A^*})$.

The uniform q-moment bound in Assumption 2(ii) implies that the conditional Lyapunov ratio is bounded by $C n_M^{1-q/2} V_E(A^*, \widehat{\varrho}_{A^*})^{-q/2}$. By Assumption 3(i), continuity of V_E, and $n_M \to_p \infty$, this expression converges to zero in probability. The conditional Lyapunov central limit theorem gives the result. □

Proof of Theorem 2. Because $\mathbb{P}(\mathcal{E}_b) \to 1$, it suffices to establish the result on $\mathcal{E}_b$.

Multiplying equation B.24 by $\sqrt{n}$ gives

$$\sqrt{n}(\widehat{\tau} - \tau) = \frac{1}{\sqrt{n}} \sum_{i \in \mathcal{I}_O} \phi_{O,i} + \frac{\sqrt{n}}{n_M} \sum_{j \in \mathcal{I}_M} \phi_{E,j}(\widehat{\varrho}_{A^*}) + o_p(1).$$

The observational leading term satisfies $n^{-1/2}\sum_{i\in\mathcal{I}_O}\phi_{O,i} \to_d \mathcal{N}\big(0, V_O(A^*)\big)$ by the Lyapunov central limit theorem.

Conditional on $\mathcal{F}_{\text{pilot}}$, the variance of the experimental leading term is $(n/n_M)V_E\left(A^*, \widehat{\varrho}_{A^*}\right)$. By Lemma B.7, $(n/n_M)V_E\left(A^*, \widehat{\varrho}_{A^*}\right) \to_p h\,\bar{c}_{A^*}(\varrho^*)V_E(A^*, \varrho^*)$. Lemma B.10 therefore implies that, for every $t \in \mathbb{R}$,

$$\mathbb{E}\left[\exp\left\{it\frac{\sqrt{n}}{n_M}\sum_{j\in\mathcal{I}_M}\phi_{E,j}\left(\widehat{\varrho}_{A^*}\right)\right\}\middle|\,\mathcal{F}_{\text{pilot}}\right] \to_p \exp\left\{-\frac{t^2}{2}h\,\bar{c}_{A^*}(\varrho^*)V_E(A^*, \varrho^*)\right\}.$$

The conditional characteristic functions are bounded by one, so this convergence also holds in L_1.

The observational leading term is $\mathcal{F}_{\text{pilot}}$-measurable. Iterated expectations therefore give

$$\frac{1}{\sqrt{n}}\sum_{i\in\mathcal{I}_O}\phi_{O,i} + \frac{\sqrt{n}}{n_M}\sum_{j\in\mathcal{I}_M}\phi_{E,j}\left(\widehat{\varrho}_{A^*}\right) \to_d \mathcal{N}\big(0, \sigma_h^2(A^*, \varrho^*_{A^*})\big). \tag{B.25}$$

Combining equation B.24 and equation B.25 proves $\sqrt{n}(\widehat{\tau} - \tau) \to_d \mathcal{N}\big(0, \sigma_h^2(A^*, \varrho^*_{A^*})\big)$.

It remains to prove variance consistency. Empirical centering in $\widehat{V}_O^{\text{main}}$ removes the additive difference between evaluating the observational score at $\widehat{\tau}$ and at τ. Assumption 3(iii) implies that the relevant observational score class is Glivenko-Cantelli after squaring and centering. Together with the first relation in equation B.23, this yields $\widehat{V}_O^{\text{main}} - V_O(A^*) = o_p(1)$.

For the experimental component, condition on $\mathcal{F}_{\text{pilot}}$ and the fold-specific main-stage training observations. The uniform q-moment bound, the conditional law of large numbers, and the second relation in equation B.23 give $\widehat{V}_E^{\text{main}} - V_E\left(A^*, \widehat{\varrho}_{A^*}\right) = o_p(1)$.

Using equation 10,

$$\begin{aligned}
&\widehat{\sigma}_h^2(\widehat{A}, \widehat{\varrho}_{\widehat{A}}) - \sigma_h^2(A^*, \varrho^*_{A^*}) \\
=&\left\{\widehat{V}_O^{\text{main}} - V_O(A^*)\right\} \\
&+ \frac{n}{n_M}\left\{\widehat{V}_E^{\text{main}} - V_E\left(A^*, \widehat{\varrho}_{A^*}\right)\right\} \\
&+ \left\{\frac{n}{n_M}V_E\left(A^*, \widehat{\varrho}_{A^*}\right) - h\,\bar{c}_{A^*}(\varrho^*)V_E(A^*, \varrho^*)\right\} \\
=&o_p(1).
\end{aligned}$$

Assumption 3(i) implies $\sigma_h^2(A^*, \varrho^*_{A^*}) \geq V_O(A^*) \geq c$. Therefore, $\widehat{\sigma}_h^2(\widehat{A}, \widehat{\varrho}_{\widehat{A}})/\sigma_h^2(A^*, \varrho^*_{A^*}) \to_p 1$.

Combining variance consistency with the preceding central limit theorem, Slutsky's theorem gives asymptotic $(1-\alpha)$ coverage for the confidence interval in equation 11. □

C Additional Methodological Details

C.1 Pilot budget proportion

This section formalizes the pilot-budget tradeoff as discussed in Remark 1 by deriving conditions and a rate for $\lambda_{P,b}$ that balances pilot-stage design estimation against the budget remaining for the main experiment.

The pilot must supply enough observations to estimate the design criterion without using a nonvanishing share of the budget. Under the conditions for design criterion consistency, $\lambda_{P,b} \to 0$ and $\lambda_{P,b}b \to \infty$ suffice for recovery of the oracle design. A more specific allocation rate requires balancing the loss of budget for the main experiment against the error in estimating the design criterion. We derive this balance under an additional rate condition.

For (A, ϱ_A), write $Q_A(\varrho_A) = \bar{c}_A(\varrho_A)V_E(A, \varrho_A)$. For a generic pilot budget proportion $0 \leq \lambda < 1$, define

$$V_n(A, \varrho_A; \lambda) := V_O(A) + \frac{n}{(1-\lambda)b}Q_A(\varrho_A). \tag{C.26}$$

Here n and b are held fixed when different values of λ are compared. Their dependence is suppressed in the notation. The finite-sample design criterion in equation 5 is therefore $V_n(A, \varrho_A) = V_n(A, \varrho_A; \lambda_{P,b})$. The corresponding full budget benchmark is $V_n(A, \varrho_A; 0) = V_O(A) + (n/b)Q_A(\varrho_A)$.

For each $\lambda \in [0, 1)$, define

$$v^*(\lambda) := \inf_{(A,\varrho_A)\in\mathfrak{D}} V_n(A, \varrho_A; \lambda).$$

These optimized values compare alternative pilot budget proportions at a given n and b. They do not redefine the oracle design A^*, $\varrho^*_{A^*}$ in equation 4, which is determined by the fixed asymptotic design criterion $\sigma_h^2(A, \varrho_A)$.

Proposition C.1. *Suppose Assumptions 1 and 2 hold, and $\lambda_{P,b} \to 0$, $\lambda_{P,b}b \to \infty$, and the relative-budget condition in equation 6 hold. Then*

$$\max_{A\in\mathcal{A}} \sup_{\varrho_A\in\mathcal{R}_A} \left|\widehat{V}_n^{\text{pilot}}(A, \varrho_A) - \sigma_h^2(A, \varrho_A)\right| = o_p(1).$$

No additional rate restriction on $\lambda_{P,b}$ is required for this first-order conclusion.

Moreover, for every $0 \le \lambda < 1$,

$$v^*(0) \le v^*(\lambda) \le \frac{1}{1-\lambda} v^*(0). \tag{C.27}$$

If $v^*(0) > 0$, then

$$0 \le \frac{v^*(\lambda)}{v^*(0)} - 1 \le \frac{\lambda}{1-\lambda}. \tag{C.28}$$

Proof. By equation 7, $\widehat{V}_n^{\text{pilot}}(A, \varrho_A) = \widehat{V}_O^{\text{pilot}}(A) + h_n\widehat{Q}_A(\varrho_A)$, where $\widehat{Q}_A(\varrho_A) = \widehat{\overline{c}}_A(\varrho_A)\widehat{V}_E^{\text{pilot}}(A, \varrho_A)$. By equation 3, $\sigma_h^2(A, \varrho_A) = V_O(A) + hQ_A(\varrho_A)$. Therefore,

$$\left|\widehat{V}_n^{\text{pilot}}(A, \varrho_A) - \sigma_h^2(A, \varrho_A)\right| \le \left|\widehat{V}_O^{\text{pilot}}(A) - V_O(A)\right| + h_n \left|\widehat{Q}_A(\varrho_A) - Q_A(\varrho_A)\right| + |h_n - h| Q_A(\varrho_A).$$

Lemma B.4 gives uniform consistency of the estimated components. Moreover, $h_n \to h$, and Q_A is uniformly bounded under Assumption 2(ii). This proves the first conclusion.

By equation C.26, for every fixed candidate design (A, ϱ_A), nonnegativity of $V_O(A)$ and $Q_A(\varrho_A)$ gives $V_n(A, \varrho_A; 0) \le V_n(A, \varrho_A; \lambda) \le (1-\lambda)^{-1} V_n(A, \varrho_A; 0)$. Taking infima over the candidate designs proves equation C.27. If $v^*(0) > 0$, division by $v^*(0)$ gives equation C.28. □

For the remainder of this analysis of the pilot budget proportion, assume $v^*(0) > 0$. The positivity condition in equation C.28 is needed only for the relative comparison. A sufficient condition is $\min_{A\in\mathcal{A}} V_O(A) > 0$, because $v^*(0) \ge \min_{A\in\mathcal{A}} V_O(A)$. The sandwich in equation C.27 does not require this condition.

Define

$$c_P := \overline{c}_{A_P}(\varrho_{\text{pilot}}), \qquad \widehat{c}_P := \widehat{\overline{c}}_{A_P}(\varrho_{\text{pilot}}), \qquad N_b := \frac{b}{c_P}.$$

By equation B.14, $\widehat{c}_P / c_P \to_p 1$. Since

$$n_P = \left\lfloor \frac{\lambda_{P,b} b}{\widehat{c}_P} \right\rfloor = \lambda_{P,b} N_b \frac{c_P}{\widehat{c}_P} + O(1),$$

the condition $\lambda_{P,b}N_b \to \infty$ implies

$$\frac{n_P}{\lambda_{P,b}N_b} \to_p 1. \tag{C.29}$$

Moreover, $\lambda_{P,b}b = c_P\lambda_{P,b}N_b$. Hence the regime for the pilot budget proportion is equivalently

$$\lambda_{P,b} \to 0, \qquad \lambda_{P,b}N_b \to \infty. \tag{C.30}$$

This representation is invariant to a common rescaling of the units used to measure the budget and expected per-unit costs.

For example, if $\lambda_{P,b} = N_b^{-\beta}$, $0 < \beta < 1$, then $\lambda_{P,b} \to 0$ and $\lambda_{P,b} N_b = N_b^{1-\beta} \to \infty$. Thus, every $0 < \beta < 1$ satisfies equation C.30. First-order consistency alone therefore does not determine a unique pilot budget proportion.

All estimator-based statements below are understood along deterministic sequences $\lambda = \lambda_b$ satisfying $\lambda_b \to 0$ and $\lambda_b N_b \to \infty$. For $0 < \lambda < 1$, define the implemented pilot sample size by $n_P(\lambda) := \lfloor \lambda b / \widehat{c}_P \rfloor$. Then, along every such sequence,

$$\frac{n_P(\lambda_b)}{\lambda_b N_b} \to_p 1, \qquad n_P(\lambda_b) \to_p \infty.$$

Hence the pilot sample accommodates the fixed number K_P of folds with probability tending to one. Let

$$\widehat{V}_n^{\mathrm{pilot}}(A, \varrho_A; \lambda) = \widehat{V}_{O,\lambda}^{\mathrm{pilot}}(A) + \frac{n}{(1-\lambda)b} \widehat{Q}_{A,\lambda}(\varrho_A),$$

where

$$\widehat{Q}_{A,\lambda}(\varrho_A) = \widehat{\overline{c}}_{A,\lambda}(\varrho_A) \widehat{V}_{E,\lambda}^{\mathrm{pilot}}(A, \varrho_A).$$

The subscript λ indicates that all pilot-stage nuisance functions and criterion components are constructed using a pilot sample of size $n_P(\lambda)$.

Define

$$(\widehat{A}_\lambda, \widehat{\varrho}_{\widehat{A}_\lambda,\lambda}) = \arg \min_{(A,\varrho_A) \in = \mathfrak{D}} \widehat{V}_n^{\mathrm{pilot}}(A, \varrho_A; \lambda).$$

For this auxiliary calculation, assume that a measurable exact minimizer exists.

Define the relative total criterion gap by

$$\begin{aligned} \mathcal{G}_{P,b}(\lambda) &:= \frac{V_n(\widehat{A}_\lambda, \widehat{\varrho}_{\widehat{A}_\lambda,\lambda}; \lambda) - v^*(0)}{v^*(0)} \\ &= \underbrace{\frac{v^*(\lambda) - v^*(0)}{v^*(0)}}_{\mathcal{G}_{F,b}(\lambda)} + \underbrace{\frac{V_n(\widehat{A}_\lambda, \widehat{\varrho}_{\widehat{A}_\lambda,\lambda}; \lambda) - v^*(\lambda)}{v^*(0)}}_{\mathcal{G}_{\mathrm{main},b}(\lambda)}. \end{aligned} \tag{C.31}$$

The term $\mathcal{G}_{F,b}(\lambda)$ measures the increase in the design criterion caused by reserving a proportion λ of the budget for the pilot experiment. The term $\mathcal{G}_{\mathrm{main},b}(\lambda)$ measures the error from estimating the criterion.

By Proposition C.1,

$$0 \le \mathcal{G}_{F,b}(\lambda) \le \frac{\lambda}{1-\lambda} = \lambda + O(\lambda^2). \tag{C.32}$$

Define the uniform criterion estimation error by

$$\Delta_{P,b}(\lambda) := \frac{2}{v^*(0)} \sup_{(A,\varrho_A) \in \mathfrak{D}} \left| \widehat{V}_n^{\mathrm{pilot}}(A, \varrho_A; \lambda) - V_n(A, \varrho_A; \lambda) \right|.$$

This quantity includes every error arising from pilot-stage criterion estimation. In particular, it includes the effect of pilot-based nuisance estimation on both $\widehat{V}_{O,\lambda}^{\mathrm{pilot}}$ and $\widehat{Q}_{A,\lambda}$.

The empirical-minimization inequality gives

$$V_n(\widehat{A}_\lambda, \widehat{\varrho}_{\widehat{A}_\lambda,\lambda}; \lambda) - v^*(\lambda) \le 2 \sup_{(A,\varrho_A) \in \mathfrak{D}} \left| \widehat{V}_n^{\mathrm{pilot}}(A, \varrho_A; \lambda) - V_n(A, \varrho_A; \lambda) \right|.$$

Consequently,

$$0 \le \mathcal{G}_{\mathrm{main},b}(\lambda) \le \Delta_{P,b}(\lambda). \tag{C.33}$$

Combining equation C.31, equation C.32, and equation C.33 yields $0 \le \mathcal{G}_{P,b}(\lambda) \le \lambda/(1-\lambda) + \Delta_{P,b}(\lambda)$.

Suppose additionally that, for some $\gamma_P > 0$, every deterministic sequence $\lambda = \lambda_b$ satisfying $\lambda_b \to 0$ and $\lambda_b N_b \to \infty$ obeys

$$\Delta_{P,b}(\lambda_b) = O_p\left\{ (n_P(\lambda_b))^{-\gamma_P} \right\} = O_p\left\{ (\lambda_b N_b)^{-\gamma_P} \right\}. \tag{C.34}$$

In particular, this condition applies to the implemented sequence $\lambda_{P,b}$.

Under equation C.34, the comparison across pilot budget proportion is between $\lambda/(1-\lambda)$, which represents the reduction in the main experimental budget, and $(\lambda N_b)^{-\gamma_P}$, which represents the order of the pilot-stage criterion estimation error. A deterministic proxy for these two components is

$$\mathcal{U}_b(\lambda) = C_F \frac{\lambda}{1-\lambda} + C_P \left(\lambda N_b\right)^{-\gamma_P}, \qquad 0 < \lambda < 1, \tag{C.35}$$

where $C_F, C_P > 0$. The proxy is not asserted to equal $\mathcal{G}_{P,b}(\lambda)$. It is used only to determine the order of the pilot budget proportion.

The proxy is strictly convex because $\mathcal{U}_b''(\lambda) = 2C_F/(1-\lambda)^3 + \gamma_P(1+\gamma_P)C_P N_b^{-\gamma_P}\lambda^{-(2+\gamma_P)} > 0$. It also diverges as $\lambda \downarrow 0$ or $\lambda \uparrow 1$, and therefore has a unique interior minimizer $\lambda^*_{P,b}$. Its first-order condition is $C_F/(1-\lambda^*_{P,b})^2 = \gamma_P C_P N_b^{-\gamma_P}\left(\lambda^*_{P,b}\right)^{-(1+\gamma_P)}$. Equivalently, $(\lambda^*_{P,b})^{1+\gamma_P}/(1-\lambda^*_{P,b})^2 = (\gamma_P C_P/C_F)N_b^{-\gamma_P}$. The right-hand side converges to zero, so $\lambda^*_{P,b} \to 0$. It follows that

$$\lambda^*_{P,b} = \left(\frac{\gamma_P C_P}{C_F}\right)^{1/(1+\gamma_P)} N_b^{-\gamma_P/(1+\gamma_P)}\{1 + o(1)\}. \tag{C.36}$$

Therefore, $\lambda^*_{P,b} \asymp N_b^{-\gamma_P/(1+\gamma_P)}$. Moreover, $n_P(\lambda^*_{P,b})/\lambda^*_{P,b}N_b \to_p 1$, so the implemented pilot sample size has stochastic order $N_b^{1/(1+\gamma_P)}$. In particular, $\lambda^*_{P,b} \to 0$ and $\lambda^*_{P,b}N_b \to \infty$, so the rate-balancing allocation satisfies equation C.30.

For $\gamma_P = 1/2$, equation C.36 gives

$$\lambda^*_{P,b} = \left(\frac{C_P}{2C_F}\right)^{2/3} N_b^{-1/3}\{1 + o(1)\}, \qquad n_P(\lambda^*_{P,b}) \asymp N_b^{2/3}.$$

When the leading constants are unknown, this calculation motivates the rate class $\lambda_{P,b} \asymp N_b^{-1/3}$. In Section 6, we use $\lambda_{P,b} = 3N_b^{-1/3}$. The multiplier 3 changes the finite-sample pilot budget proportion but not the exponent $1/3$. Thus, the exponent follows from the rate proxy in equation C.35 under equation C.34, but it is not asserted to be a universal finite-sample optimum.

C.2 The construction and estimation of nuisance functions

This section provides the cross-fitted nuisance-function, score, and variance constructions used to estimate the pilot design criterion in Section 4.1 and to conduct final treatment effect inference in Section 4.2.

Partition the observational and pilot index sets into a fixed $K_P \geq 2$ approximately equal-sized folds, $\{I^P_{O,k}\}_{k=1}^{K_P}$ and $\{I^P_{E,k}\}_{k=1}^{K_P}$, respectively, using the same partitions for every candidate secondary-outcome set. Here $I^P_{O,-k} = \mathcal{I}_O \setminus I^P_{O,k}$ and $I^P_{E,-k} = \mathcal{I}_P \setminus I^P_{E,k}$. We provide the nuisance-function and score construction used in Section 4.1. For each candidate $A \in \mathcal{A}$, treatment arm $w \in \{0,1\}$, and pilot-stage fold k, nuisance functions used on the evaluation folds $I^P_{O,k}$ and $I^P_{E,k}$ are estimated using the samples $I^P_{O,-k}$ and $I^P_{E,-k}$. We suppress the superscript $(-k)$ on individual nuisance estimates below when no confusion arises.

First, we fit the primary outcome regression $\mu_{w,A}$ on the observational training sample, yielding $\widehat{\mu}_{w,A}(\mathbf{s}, \mathbf{x})$.

The latent propensity score $\rho_{w,A}$ admits an observed-data representation (Chen & Ritzwoller, 2023). Define

$$\begin{aligned} a_{w,A}(\mathbf{s}, \mathbf{x}) &= \mathbb{P}\left(G = O \mid \mathbf{S}_A = \mathbf{s},\, W = w,\, \mathbf{X} = \mathbf{x}\right), \\ b_w(\mathbf{x}) &= \mathbb{P}\left(G = O \mid W = w,\, \mathbf{X} = \mathbf{x}\right), \\ e_{O,w}(\mathbf{x}) &= \mathbb{P}\left(W = w \mid \mathbf{X} = \mathbf{x},\, G = O\right). \end{aligned}$$

Then

$$\rho_{w,A}(\mathbf{s}, \mathbf{x}) = \frac{a_{w,A}(\mathbf{s}, \mathbf{x})}{1 - a_{w,A}(\mathbf{s}, \mathbf{x})} \frac{1 - b_w(\mathbf{x})}{b_w(\mathbf{x})} e_{O,w}(\mathbf{x}).$$

For each $w \in \{0,1\}$, we estimate $a_{w,A}$ and b_w using the observations indexed by $\{i \in I^P_{O,-k} \cup I^P_{E,-k} : W_i = w\}$, and estimate $e_{O,w}$ using the observations indexed by $I^P_{O,-k}$. Let $\widetilde{a}_{w,A}$, $\widetilde{b}_w$, and $\widetilde{e}_{O,w}$ denote the corresponding raw estimates.

Let $\nu_b = n_P/n$. Since $\nu_b \to 0$, the observational sample accounts for an asymptotically dominant proportion of $I^P_{O,-k} \cup I^P_{E,-k}$. Consequently, $a_{w,A}(\mathbf{s},\mathbf{x})$ and $b_w(\mathbf{x})$ approach one. We therefore choose a common clipping level $\delta_b \in (0, 1/2)$ satisfying $\delta_b/\nu_b \to 0$. Under fixed overlap and density-ratio bounds for all candidate secondary-outcome sets, this restriction ensures that clipping leaves the true membership probabilities unchanged for sufficiently large b. Here, membership probabilities refer to the pooled training distribution, with mixing proportions determined by the training sample sizes. Let $\epsilon \in (0, 1/2)$ be the clipping level for the observational treatment probabilities.

For $z \in \mathbb{R}$ and $l < u$, define $\textsc{Clip}(z; l, u) = l \vee (z \wedge u)$. Define

$$\begin{aligned}
\widehat{a}_{w,A}(\mathbf{s},\mathbf{x}) &= \textsc{Clip}\left(\widetilde{a}_{w,A}(\mathbf{s},\mathbf{x}); \delta_b, 1-\delta_b\right), \\
\widehat{b}_w(\mathbf{x}) &= \textsc{Clip}\left(\widetilde{b}_w(\mathbf{x}); \delta_b, 1-\delta_b\right), \\
\widehat{e}_{O,w}(\mathbf{x}) &= \textsc{Clip}\left(\widetilde{e}_{O,w}(\mathbf{x}); \epsilon, 1-\epsilon\right).
\end{aligned}$$

Substituting these estimates into the observed-data representation gives

$$\widetilde{\rho}_{w,A}(\mathbf{s},\mathbf{x}) = \frac{\widehat{a}_{w,A}(\mathbf{s},\mathbf{x})}{1-\widehat{a}_{w,A}(\mathbf{s},\mathbf{x})} \frac{1-\widehat{b}_w(\mathbf{x})}{\widehat{b}_w(\mathbf{x})} \widehat{e}_{O,w}(\mathbf{x}),$$

which we stabilize by

$$\widehat{\rho}_{w,A}(\mathbf{s},\mathbf{x}) = \textsc{Clip}\left(\widetilde{\rho}_{w,A}(\mathbf{s},\mathbf{x}); \epsilon, 1-\epsilon\right). \tag{C.37}$$

For the covariate density ratio, under the pooled observational-pilot distribution define

$$\pi = \mathbb{P}(G = O), \qquad t(\mathbf{x}) = \mathbb{P}(G = O \mid \mathbf{X} = \mathbf{x}).$$

Then

$$r(\mathbf{x}) = \frac{1-\pi}{\pi} \frac{t(\mathbf{x})}{1-t(\mathbf{x})}.$$

Let $\widetilde{t}(\mathbf{x})$ denote the raw estimate of $t(\mathbf{x})$ from the pooled training sample and set

$$\widehat{\pi} = \frac{|I^P_{O,-k}|}{|I^P_{O,-k}| + |I^P_{E,-k}|}, \qquad \widehat{t}(\mathbf{x}) = \textsc{Clip}\left(\widetilde{t}(\mathbf{x}); \delta_b, 1-\delta_b\right).$$

Define

$$\widetilde{r}(\mathbf{x}) = \frac{1-\widehat{\pi}}{\widehat{\pi}} \frac{\widehat{t}(\mathbf{x})}{1-\widehat{t}(\mathbf{x})}.$$

For prespecified bounds $0 < \underline{r} < \overline{r} < \infty$, set

$$\widehat{r}(\mathbf{x}) = \textsc{Clip}\left(\widetilde{r}(\mathbf{x}); \underline{r}, \overline{r}\right). \tag{C.38}$$

The restriction on δ_b prevents clipping from altering the population membership probabilities. It does not by itself establish consistency of the fitted nuisance functions. The resulting estimates must still satisfy Assumption 2(i).

It is noteworthy that, in finite samples, the pooled observational-experimental classification problem can be highly imbalanced when n_P/n is small. Balanced subsampling, class-weighted fitting, or direct density-ratio estimation may therefore be useful computational safeguards. When the class proportions are altered during fitting, the resulting membership odds must be recalibrated to the original sampling proportion before being used in the formulas above.

Finally, we estimate the projected regression $\overline{\mu}_{w,A}$ using the pilot training sample. Within treatment arm w, we regress

$$\widehat{\mu}_{w,A}\left(\mathbf{S}_{j,A}, \mathbf{X}_j\right), \qquad j \in I^P_{E,-k}, \quad W_j = w,$$

on $\mathbf{X}_j$, yielding $\widehat{\overline{\mu}}_{w,A}(\mathbf{x})$. Together, these estimates form $\widehat{\boldsymbol{\eta}}^{\mathrm{pilot},(-k)}_A$.

For each candidate A, define the cross-fitted pilot target estimate

$$\widehat{\tau}_A^{\text{pilot}} = \frac{1}{n}\sum_{k=1}^{K_P}\sum_{i\in I_{O,k}^P}\left\{\widehat{\overline{\mu}}_{1,A}^{(-k)}(\mathbf{X}_i) - \widehat{\overline{\mu}}_{0,A}^{(-k)}(\mathbf{X}_i)\right\}.$$

For $i \in I_{O,k}^P$, define

$$\widehat{\psi}_{O,A,i}^{\text{pilot}} = \psi_{O,A}\left(Y_i, W_i, \mathbf{S}_{i,A}, \mathbf{X}_i; \widehat{\tau}_A^{\text{pilot}}, \widehat{\boldsymbol{\eta}}_A^{\text{pilot},(-k)}\right),$$

and for $j \in I_{E,k}^P$,

$$\widehat{\psi}_{E,A,j}^{\text{pilot}}(\varrho_A) = \psi_{E,A}\left(W_j, \mathbf{S}_{j,A}, \mathbf{X}_j; \widehat{\boldsymbol{\eta}}_A^{\text{pilot},(-k)}, \varrho_A\right).$$

These cross-fitted estimates of the influence-function components are used to construct $\widehat{V}_O^{\text{pilot}}(A)$ and $\widehat{V}_E^{\text{pilot}}(A, \varrho_A)$.

C.2.1 Estimation of the pilot criterion

The pilot uses balanced randomization, whereas each candidate rule induces a different treatment distribution. We account for this difference using the importance weights

$$\omega_{A,j}(\varrho_A) = 2\left[W_j\varrho_A(\mathbf{X}_j) + (1-W_j)\{1-\varrho_A(\mathbf{X}_j)\}\right],\ j \in \mathcal{I}_P.$$

These weights allow the same pilot sample to estimate the experimental variance under every candidate rule. We estimate $V_O(A)$ by the empirical variance of the observational components, $V_E(A, \varrho_A)$ by the weighted second moment of the experimental components, and $\overline{c}_A(\varrho_A)$ using $\widehat{\overline{c}}_A(\varrho_A)$, obtained by averaging the known conditional costs over the prescreening covariates $\mathcal{D}_E^X$, as defined in Section 3.3. Let

$$\widehat{\overline{\psi}}_{O,A}^{\text{pilot}} = \frac{1}{n}\sum_{i\in\mathcal{I}_O}\widehat{\psi}_{O,A,i}^{\text{pilot}}, \qquad \widehat{V}_O^{\text{pilot}}(A) = \frac{1}{n}\sum_{i\in\mathcal{I}_O}\left(\widehat{\psi}_{O,A,i}^{\text{pilot}} - \widehat{\overline{\psi}}_{O,A}^{\text{pilot}}\right)^2,$$
$$\widehat{V}_E^{\text{pilot}}(A, \varrho_A) = \frac{1}{n_P}\sum_{j\in\mathcal{I}_P}\omega_{A,j}(\varrho_A)\left\{\widehat{\psi}_{E,A,j}^{\text{pilot}}(\varrho_A)\right\}^2.$$

We use a weighted second moment because the true experimental component has mean zero under the candidate rule. Using $W_j(1-W_j) = 0$, the importance weight cancels one power of the assignment probability in the squared experimental score. Thus, conditional on the fitted nuisance functions, $\widehat{V}_E^{\text{pilot}}(A, \varrho_A)$ depends on the assignment rule through terms proportional to $1/\varrho_A(\mathbf{X}_j)$ and $1/\{1-\varrho_A(\mathbf{X}_j)\}$, and is convex in the pointwise assignment probabilities. When the expected per-unit cost is independent of the assignment rule, this also yields a convex assignment optimization problem. With arm-specific costs, however, $\widehat{\overline{c}}_A(\varrho_A)\widehat{V}_E^{\text{pilot}}(A, \varrho_A)$ needs not be convex, and our theory therefore does not require $\mathcal{R}_A$ to be convex and only imposes the approximate optimization condition in Assumption 2(iii).

C.2.2 Variance estimation for the main experiment

In the main stage, we refit all nuisance functions using the observational and main experimental training folds, applying the construction above with the corresponding sample sizes. Neither pilot observations nor pilot-stage fits are reused.

Let $\widehat{\psi}_{O,\widehat{A},i}^{\text{main}}$ and $\widehat{\psi}_{E,\widehat{A},j}^{\text{main}}$ denote estimates of the two influence-function components in equation 9, evaluated at $\widehat{\tau}$ and the nuisance estimates fitted without the corresponding evaluation fold. Write their sample means as $\widehat{\overline{\psi}}_{O,\widehat{A}}^{\text{main}}$ and $\widehat{\overline{\psi}}_{E,\widehat{A}}^{\text{main}}$. We estimate the variance components by

$$\widehat{V}_O^{\text{main}} = \frac{1}{n}\sum_{i\in\mathcal{I}_O}\left(\widehat{\psi}_{O,\widehat{A},i}^{\text{main}} - \widehat{\overline{\psi}}_{O,\widehat{A}}^{\text{main}}\right)^2, \quad \widehat{V}_E^{\text{main}} = \frac{1}{n_M}\sum_{j\in\mathcal{I}_M}\left(\widehat{\psi}_{E,\widehat{A},j}^{\text{main}} - \widehat{\overline{\psi}}_{E,\widehat{A}}^{\text{main}}\right)^2.$$

C.3 VARIANCE DECOMPOSITION OF THE TWO COMPONENTS

This section derives explicit expressions for the observational and experimental variance components $V_O(A)$ and $V_E(A, \varrho_A)$, which are introduced in Section 3.4 and form the basis of the design criterion.

For $w \in \{0, 1\}$, define the conditional outcome variance

$$\sigma^2_{w,A}(\mathbf{s}, \mathbf{x}) = \mathrm{Var}\left(Y_i \mid W_i = w,\, \mathbf{S}_{i,A} = \mathbf{s},\, \mathbf{X}_i = \mathbf{x},\, G_i = O\right).$$

Using the conditional mean-zero property of $Y_i - \mu_{w,A}(\mathbf{S}_{i,A}, \mathbf{X}_i)$, together with $W_i(1 - W_i) = 0$, the observational variance component satisfies

$$V_O(A) = \mathbb{E}_{P_O}\left[\frac{W_i \sigma^2_{1,A}(\mathbf{S}_{i,A}, \mathbf{X}_i)}{\rho^2_{1,A}(\mathbf{S}_{i,A}, \mathbf{X}_i)} + \frac{(1 - W_i)\sigma^2_{0,A}(\mathbf{S}_{i,A}, \mathbf{X}_i)}{\rho^2_{0,A}(\mathbf{S}_{i,A}, \mathbf{X}_i)} + \left\{\overline{\mu}_{1,A}(\mathbf{X}_i) - \overline{\mu}_{0,A}(\mathbf{X}_i) - \tau\right\}^2\right].$$

Since $W_i^2 = W_i$, $(1 - W_i)^2 = 1 - W_i$, and $W_i(1 - W_i) = 0$, the experimental variance component can be written as

$$V_E(A, \varrho_A) = \mathbb{E}_{P^{\varrho_A}_{E,A}}\left[r^2(\mathbf{X}_i)\left\{\frac{W_i}{\varrho^2_A(\mathbf{X}_i)}\left[\mu_{1,A}(\mathbf{S}_{i,A}, \mathbf{X}_i) - \overline{\mu}_{1,A}(\mathbf{X}_i)\right]^2 + \frac{1 - W_i}{\{1 - \varrho_A(\mathbf{X}_i)\}^2}\left[\mu_{0,A}(\mathbf{S}_{i,A}, \mathbf{X}_i) - \overline{\mu}_{0,A}(\mathbf{X}_i)\right]^2\right\}\right].$$

C.4 INTEGER ROUNDING UNDER THE EXPECTED-BUDGET CONSTRAINT

Because the implemented sample size must be integer-valued, we define the main experimental sample size as $n_M(A, \varrho_A) = \left\lfloor b_F / \widehat{\overline{c}}_A(\varrho_A) \right\rfloor$, where $b_F = (1 - \lambda_{P,b})b$ denotes the budget allocated to the main experiment. The floor operation introduces a rounding error in the relation between the implemented sample size and the estimated expected cost. We first characterize this rounding error, and then combine it with Lemma A.1 to relate the implemented sample size to the population-level variance criterion. For each candidate design (A, ϱ_A), define the rounding residual

$$\zeta_{A,b}(\varrho_A) = \frac{b_F}{\widehat{\overline{c}}_A(\varrho_A)} - n_M(A, \varrho_A).$$

By the definition of the floor operation, $0 \le \zeta_{A,b}(\varrho_A) < 1$. Hence,

$$n_M(A, \varrho_A) = \frac{b_F}{\widehat{\overline{c}}_A(\varrho_A)} - \zeta_{A,b}(\varrho_A).$$

Assume that the estimated costs are uniformly bounded away from zero and infinity with probability tending to one, i.e.,

$$\mathbb{P}\left(\underline{c} \le \min_{A \in \mathcal{A}} \inf_{\varrho_A \in \mathcal{R}_A} \widehat{\overline{c}}_A(\varrho_A) \le \max_{A \in \mathcal{A}} \sup_{\varrho_A \in \mathcal{R}_A} \widehat{\overline{c}}_A(\varrho_A) \le C\right) \to 1$$

for some constants $0 < \underline{c} < C < \infty$. Since $b_F \to \infty$, uniformly over the candidate design class, $b_F - \zeta_{A,b}(\varrho_A)\widehat{\overline{c}}_A(\varrho_A) \ge b_F/2$ with probability tending to one. Therefore,

$$\frac{1}{n_M(A, \varrho_A)} - \frac{\widehat{\overline{c}}_A(\varrho_A)}{b_F} = \frac{\zeta_{A,b}(\varrho_A)\widehat{\overline{c}}^2_A(\varrho_A)}{b_F\{b_F - \zeta_{A,b}(\varrho_A)\widehat{\overline{c}}_A(\varrho_A)\}}.$$

It follows that

$$\max_{A \in \mathcal{A}} \sup_{\varrho_A \in \mathcal{R}_A} \left| \frac{1}{n_M(A, \varrho_A)} - \frac{\widehat{\overline{c}}_A(\varrho_A)}{b_F} \right| = O_p(b_F^{-2}).$$

Consequently,

$$\max_{A\in\mathcal{A}}\sup_{\varrho_A\in\mathcal{R}_A}\left|V_O(A)+\frac{n}{n_M(A,\varrho_A)}V_E(A,\varrho_A)-\left\{V_O(A)+h_n\widehat{\bar{c}}_A(\varrho_A)V_E(A,\varrho_A)\right\}\right|=O_p(b_F^{-1}),$$

where $h_n = n/b_F$. Finally, by the uniform cost consistency result,

$$\max_{A\in\mathcal{A}}\sup_{\varrho_A\in\mathcal{R}_A}\left|\widehat{\bar{c}}_A(\varrho_A)-\bar{c}_A(\varrho_A)\right|=o_p(1),$$

we obtain

$$\max_{A\in\mathcal{A}}\sup_{\varrho_A\in\mathcal{R}_A}\left|V_O(A)+\frac{n}{n_M(A,\varrho_A)}V_E(A,\varrho_A)-V_n(A,\varrho_A)\right|=o_p(1).$$

Together with $h_n \to h$, this yields

$$\max_{A\in\mathcal{A}}\sup_{\varrho_A\in\mathcal{R}_A}\left|V_O(A)+\frac{n}{n_M(A,\varrho_A)}V_E(A,\varrho_A)-\sigma_h^2(A,\varrho_A)\right|=o_p(1).$$

D SIMULATION DETAILS

This section provides the simulation setup and implementation details, followed by additional results on empirical size and power of our testing procedure, oracle-design recovery, and sensitivity to the pilot budget proportion and observation-to-budget ratio (i.e., the ratio of the observational sample size to the budget). It also compares EFFICIENTESC with the balanced designs that use the full experimental budget without a pilot.

D.1 SIMULATION SETUP

D.1.1 DATA GENERATING PROCESS AND IDENTIFICATION

The candidate secondary-outcome collection is $\mathcal{A} = \{\{1\},\{1,2\},\{1,2,3\}\}$, and the pilot measures $A_P = \{1,2,3\}$, so every candidate set contains the first secondary outcome S_1. For each unit, let $X \sim \text{Bernoulli}(0.5)$, $U \sim \text{Unif}[-1,1]$, and $V, Z, \varepsilon_0, \varepsilon_1 \overset{\text{i.i.d.}}{\sim} \mathcal{N}(0,1)$, with these variables mutually independent.

For $w \in \{0,1\}$, define the potential secondary outcomes by

$$S_1(w) = U + 0.5w, \qquad S_2(w) = V + 0.5w, \qquad S_3(w) = V + 0.5w + Z,$$

and the potential primary outcome by

$$Y(w) = 0.9w + 0.5X + a_{w,X}S_1(w) + b_{w,X}S_2(w) + \varepsilon_w,$$

where

$$(a_{1,x}, a_{0,x}) = \begin{cases}(2, 0.5), & x = 0,\\ (0.5, 2), & x = 1,\end{cases} \qquad (b_{1,x}, b_{0,x}) = \begin{cases}(1.9, 0.9), & x = 0,\\ (0.9, 1.9), & x = 1.\end{cases}$$

The three data sources are generated as follows:

- *Observational sample.* Treatment is generated according to $W^O \mid X, U \sim \text{Bernoulli}(e_O(U))$, where $e_O(U) = \text{expit}(2U)$. The observed outcomes are $Y^O = Y(W^O)$ and $S_k^O = S_k(W^O)$, $k = 1,2,3$.
- *Pilot experiment.* The pilot measures $A_P = \{1,2,3\}$ under balanced randomization, $W \mid X \sim \text{Bernoulli}(1/2)$. Experimental primary outcomes are generated only for validating the data generating process (DGP) and are not supplied to any estimator.
- *Main experiment.* For each $A \in \mathcal{A}$, the assignment-rule class is

  $$\mathcal{R}_A = \left\{\varrho_A : \varrho_A(x) = p_{A,0}\mathbf{1}\{x=0\} + p_{A,1}\mathbf{1}\{x=1\},\ (p_{A,0}, p_{A,1}) \in [0.1, 0.9]^2\right\}.$$

  Conditional on the pilot-stage information $\mathcal{F}_{\text{pilot}}$, treatment in the main experiment is assigned according to $\widehat{\varrho}_{\widehat{A}}(X)$.

The observational sample and pilot experiment are generated independently. Conditional on $\mathcal{F}_{\text{pilot}}$, the main experimental observations are independent draws under $P^{\widehat{\varrho}_{\widehat{A}}}_{E,\widehat{A}}$. In the simulation, the per-unit experimental costs are constant with respect to X. Therefore, the expected cost coincides with the conditional cost, and the cost function is known without estimation.

All candidate secondary-outcome sets satisfy Assumption 1:

- *Experimental unconfoundedness.* Treatment is randomized conditional on X in the pilot and, conditional on $\mathcal{F}_{\text{pilot}}$, in the main experiment.
- *Conditional external validity.* The observational and experimental regimes share the same structural equations conditional on X, with $P^X_O = P^X_E$; hence $r(x) = 1$.
- *Latent unconfoundedness.* Since every $A \in \mathcal{A}$ contains S_1 and $S_1(w) = U + 0.5w$, conditioning on $S_1(w)$ determines the latent variable U. The remaining variation in $Y(w)$ is independent of observational treatment, yielding

  $$W^O \perp\!\!\!\perp Y(w) \mid \mathbf{S}_A(w), X, G = O$$

  for every $A \in \mathcal{A}$.
- *Overlap.* The observational latent propensity lies in

  $$[\text{expit}(-2), \text{expit}(2)] = [0.119, 0.881],$$

  while all candidate experimental assignment probabilities lie in $[0.1, 0.9]$.

The DGP nevertheless contains nontrivial observational confounding. The unadjusted observational contrast is

$$\tau_{\text{unadj}} = \mathbb{E}(Y^O \mid W^O = 1) - \mathbb{E}(Y^O \mid W^O = 0) = 2.928,$$

whereas the true ATE is

$$\tau = \mathbb{E}\{Y(1) - Y(0)\} = 2.225.$$

D.1.2 Oracle benchmarks

The DGP yields the following primary-outcome regressions:

$$\mu_{w,\{1\}}(s_1, x) = (0.9 + 0.5b_{w,x})w + 0.5x + a_{w,x}s_1,$$

$$\mu_{w,\{1,2\}}(s_1, s_2, x) = 0.9w + 0.5x + a_{w,x}s_1 + b_{w,x}s_2,$$

$$\mu_{w,\{1,2,3\}}(s_1, s_2, s_3, x) = \mu_{w,\{1,2\}}(s_1, s_2, x).$$

For every $A \in \mathcal{A}$,

$$\overline{\mu}_{w,A}(x) = 0.9w + 0.5x + \frac{w}{2}(a_{w,x} + b_{w,x}).$$

Writing s_1 for the first component of $\mathbf{s}_A$, the observational latent propensity scores are

$$\rho_{1,A}(\mathbf{s}_A, x) = \text{expit}\{2(s_1 - 0.5)\}, \qquad \rho_{0,A}(\mathbf{s}_A, x) = 1 - \text{expit}(2s_1),$$

which are the same for all three candidate secondary-outcome sets.

The simulation budget and measurement costs are specified as follows:

- *Measurement costs.* Costs are invariant across treatment arms and covariate strata:

  $$c_1(A, x) = c_0(A, x) = c(A) = 0.70 + 0.30\mathbf{1}\{1 \in A\} + 0.15\mathbf{1}\{2 \in A\} + 0.15\mathbf{1}\{3 \in A\}.$$

  Consequently,

  $$c(\{1\}) = 1.00, \qquad c(\{1, 2\}) = 1.15, \qquad c(\{1, 2, 3\}) = 1.30.$$

- *Budget and pilot allocation.* The observational sample-size grid is $n \in \{500, 1000, \ldots, 10000\}$, with $b = n/0.8$. Since the pilot measures $A_P = \{1, 2, 3\}$, $c_P = 1.30$ and $N_b = b/c_P$. The pilot allocation is

$$\lambda_{P,b} = 3N_b^{-1/3}, \qquad n_P = \lfloor \lambda_{P,b} N_b \rfloor = \lfloor 3N_b^{2/3} \rfloor.$$

This allocation satisfies $\lambda_{P,b} \to 0$, $\lambda_{P,b} b \to \infty$, and

$$h_n = \frac{n}{(1 - \lambda_{P,b})b} \to 0.8.$$

Main experimental size. For a fixed candidate design,

$$n_M(A, \varrho_A) = \left\lfloor \frac{(1 - \lambda_{P,b})b}{c(A)} \right\rfloor.$$

Monte Carlo replications. All reported results use 5,000 replications.

The resulting observational variance components are

$$V_O(\{1\}) = 18.453, \qquad V_O(\{1, 2\}) = V_O(\{1, 2, 3\}) = 6.017.$$

For $w, x \in \{0, 1\}$, define

$$M_{w,A}(x) = \mathbb{E}\left[\left\{ \mu_{w,A}\{\mathbf{S}_A(w), x\} - \overline{\mu}_{w,A}(x) \right\}^2 \mid X = x, G = E \right].$$

The structural equations imply

$$M_{w,\{1\}}(x) = \frac{a_{w,x}^2}{3},$$

and

$$M_{w,\{1,2\}}(x) = M_{w,\{1,2,3\}}(x) = \frac{a_{w,x}^2}{3} + b_{w,x}^2.$$

For an assignment rule with $\varrho_A(x) = p_{A,x}$, the experimental variance component is

$$V_E(A, \varrho_A) = \frac{1}{2} \sum_{x \in \{0,1\}} \left\{ \frac{M_{1,A}(x)}{p_{A,x}} + \frac{M_{0,A}(x)}{1 - p_{A,x}} \right\}.$$

Because measurement costs do not depend on treatment assignment, minimizing $\overline{c}_A(\varrho_A) V_E(A, \varrho_A)$ for a fixed A is equivalent to minimizing $V_E(A, \varrho_A)$. The optimization separates across the two covariate strata, with

$$p_{A,x}^* = \frac{\sqrt{M_{1,A}(x)}}{\sqrt{M_{1,A}(x)} + \sqrt{M_{0,A}(x)}}.$$

All resulting probabilities lie in $[0.1, 0.9]$, so they are also the minimizers over $\mathcal{R}_A$. The optimal assignment probabilities and corresponding experimental variance components are

A	$\varrho_A^*(0)$	$\varrho_A^*(1)$	$V_E(A, \varrho_A^*)$
$\{1\}$	0.800	0.200	2.083
$\{1, 2\}$	0.702	0.298	10.040
$\{1, 2, 3\}$	0.702	0.298	10.040

Since $h_n \to 0.8$, the asymptotic variance relevant for the oracle design is

$$\sigma_{0.8}^2(A, \varrho_A) = V_O(A) + 0.8 c(A) V_E(A, \varrho_A).$$

Under the optimal treatment-assignment rule for each candidate set, the resulting asymptotic variances are

A	$V_O(A)$	$c(A)$	$V_E(A, \varrho_A^*)$	$\sigma_{0.8}^2(A, \varrho_A^*)$
$\{1\}$	18.453	1.000	2.083	20.120
$\{1, 2\}$	6.017	1.150	10.040	15.254
$\{1, 2, 3\}$	6.017	1.300	10.040	16.459

Thus, the unique oracle design is

$$A^* = \{1, 2\}, \qquad (\varrho^*_{A^*}(0), \varrho^*_{A^*}(1)) = (0.702, 0.298).$$

The set $\{1\}$ omits useful prognostic information, whereas $\{1, 2, 3\}$ incurs the cost of measuring the redundant outcome S_3. For $A = \{1, 2\}$, balanced assignment gives $\sigma^2_{0.8}(\{1, 2\}, 1/2) = 16.757$, compared with 15.254 under the oracle assignment rule.

At a finite budget, the corresponding design criterion is

$$V_n(A, \varrho_A) = V_O(A) + h_n c(A) V_E(A, \varrho_A).$$

Direct evaluation over the full simulation grid confirms that $A^* = \{1, 2\}$ is also the unique minimizer of $V_n(A, \varrho^*_A)$ for every $n \in \{500, 1000, \ldots, 10000\}$.

The main simulation compares nine methods under a common pilot-budget convention:

- six fixed-set designs, obtained by combining each $A \in \mathcal{A}$ with either balanced assignment or a pilot-learned adaptive assignment rule;
- EFFICIENTESC, which learns both the secondary-outcome set and treatment-assignment rule;
- the regret-optimal method adapted from Epanomeritakis & Viviano (2025);
- the oracle-rule benchmark, which implements $(A^*, \varrho^*_{A^*})$.

The same pilot expenditure is deducted for all nine methods, and their main experimental sample sizes are determined from the remaining budget. For the oracle-rule benchmark, only the design is treated as known: the nuisance functions, treatment effect, and standard error are estimated from the realized observational and main experimental samples.

D.2 IMPLEMENTATION DETAILS

Pilot-stage estimation uses $K_P = 5$ folds, and final effect estimation uses $K_M = 2$ folds. Within each stage, nuisance functions are fitted on the training folds and evaluated on the corresponding held-out fold. For final estimation, all nuisance functions are refitted using the observational and main experimental samples. Pilot-stage fits are not reused.

For each treatment arm, $\mu_{w,A}$ is fitted by ordinary least squares using X, the outcomes in $\mathbf{S}_A$, and interactions between X and each outcome in $\mathbf{S}_A$. The regression for $\overline{\mu}_{w,A}$ regresses the fitted values of $\mu_{w,A}$ in the experimental sample on X. The binary nuisance models are fitted by logistic regression. For each arm, the sample-membership regression conditional on $(\mathbf{S}_A, X)$ includes X, the outcomes in $\mathbf{S}_A$, and their interactions with X. The corresponding regression conditional only on X, the observational treatment-probability regression, and the marginal sample-membership regression use X as the predictor. No regularization or hyperparameter tuning is used. The logistic regressions using $(\mathbf{S}_A, X)$ to distinguish observational from experimental observations are intentionally misspecified under this DGP, while the outcome regressions are correctly specified. This implementation provides a stress test of the finite-sample robustness of EFFICIENTESC to nuisance-model misspecification.

Sample-membership probabilities entering the ratio representations are clipped to $[0.01, 0.99]$. Estimated observational treatment probabilities and the resulting $\rho_{w,A}$ are clipped to $[0.05, 0.95]$, the estimated density ratio is restricted to $[0.1, 10]$, and candidate treatment-assignment probabilities are restricted to $[0.1, 0.9]$. The fixed clipping constants used in the simulations are finite-sample stabilization choices and are distinct from the asymptotic clipping sequence in Appendix C.2.

For each candidate set A, the two stratum-specific assignment probabilities are obtained by minimizing $\widehat{c}_A(\varrho_A)\widehat{V}^{\text{pilot}}_E(A, \varrho_A)$ over $[0.1, 0.9]^2$. We use L-BFGS-B, initialized at $(0.5, 0.5)$, with a maximum of 500 iterations and `factr`$= 10^7$. The three candidate secondary-outcome sets are enumerated directly and compared using $\widehat{V}^{\text{pilot}}_n(A, \widehat{\varrho}_A)$. Because X has only two strata in the simulation, this is a two-dimensional numerical optimization. We do not claim that the same optimization is computationally simple for arbitrary assignment-rule class or treatment-dependent costs.

We additionally implement a regret-optimal method following Epanomeritakis & Viviano (2025), adapted to the candidate secondary-outcome sets and treatment-assignment rules considered in our simulation. For each $A \in \mathcal{A}$, the observational sample and the balanced pilot are used to estimate the quantities entering the regret criterion, including the observational variance contribution, the experimental precision attainable under the corresponding assignment rule, and the sensitivity of the target estimand to secondary outcomes that are not measured experimentally. No main experimental outcomes are used in this design step. We restrict the shrinkage weights of Epanomeritakis & Viviano (2025) to one for experimentally measured secondary outcomes, so that the benchmark differs across candidate sets through which secondary outcomes are learned experimentally and through the resulting treatment allocation.

For each A, let $\widehat{\varrho}_A^R$ denote the sensitivity-weighted treatment-assignment rule described below, with the main experimental sample size determined by the same remaining budget and measurement cost $c(A)$ as for the other methods. Let $\widehat{\alpha}(A, \widehat{\varrho}_A^R)$ denote the estimated variance component and $\widehat{\beta}(A)$ the estimated worst-case squared-bias component. We first compute

$$\widehat{\alpha}^* = \min_{A \in \mathcal{A}} \widehat{\alpha}(A, \widehat{\varrho}_A^R), \qquad \widehat{\beta}^* = \min_{A \in \mathcal{A}} \widehat{\beta}(A),$$

and select

$$\widehat{A}_R \in \arg\min_{A \in \mathcal{A}} \max \left\{ \frac{\widehat{\alpha}(A, \widehat{\varrho}_A^R)}{\widehat{\alpha}^*}, \frac{\widehat{\beta}(A)}{\widehat{\beta}^*} \right\},$$

using the convention $0/0 = 1$. The resulting main experiment uses $(\widehat{A}_R, \widehat{\varrho}_{\widehat{A}_R}^R)$. Thus, the regret-optimal method balances normalized variance and worst-case bias sensitivity rather than directly minimizing the ESC variance criterion $\widehat{V}_n^{\text{pilot}}(A, \varrho_A)$.

All simulations use master seed 20260823, with separate deterministic random-number streams for observational sampling, the pilot, pilot-stage cross-fitting folds, main experimental potential outcomes and assignments, and final cross-fitting folds. Random-number generation uses the L'Ecuyer-CMRG generator.

D.3 Additional simulation results

D.3.1 Empirical size and power

We examine finite-sample inference for EfficientESC, the oracle-rule benchmark, regret-optimal, and the naive design used in the main comparison. The naive design fixes $A = \{1, 2, 3\}$ and $\varrho(X) = 1/2$, while retaining the same pilot-budget deduction as the other methods. It is therefore distinct from the no-pilot full-budget designs considered in Appendix D.3.5. For each method and replication, we use its estimated 95% Wald confidence interval. For EfficientESC, this is given by equation 11. A two-sided level-0.05 test of $H_0 : \tau = \tau_0$ rejects when τ_0 lies outside this interval.

Table D.1 reports rejection frequencies for $n = 1000, 2000, \ldots, 10000$, under the true null $\tau_0 = \tau = 2.225$ and the false null values $\tau_0 \in \{2.1, 2.0, 1.9\}$. The empirical sizes of all four methods remain close to the nominal 5% level. Power increases with n and as the false null moves farther from the true effect. The power of EfficientESC closely tracks that of the oracle-rule benchmark and is generally larger than that of both the naive design and the regret-optimal method. Across the non-saturated cases, the regret-optimal method generally achieves power between EfficientESC and the naive design.

D.3.2 Recovery of the oracle design

Figure D.1 examines recovery of the two components of the oracle design. Under the simulation DGP, the optimal secondary-outcome set is $A^* = \{1, 2\}$, with assignment probabilities

$$\big(\varrho_{A^*}^*(0), \varrho_{A^*}^*(1)\big) = (0.702, 0.298).$$

The left panel reports the empirical probability $\mathbb{P}(\widehat{A} = A^*)$. The right panel reports the root mean squared error of the treatment-assignment rule learned by EfficientESC,

$$\text{RMSE}(\widehat{\varrho}_{\widehat{A}}, \varrho_{A^*}^*) = \left[\mathbb{E} \left\{ \frac{1}{2} \sum_{x \in \{0,1\}} \big(\widehat{\varrho}_{\widehat{A}}(x) - \varrho_{A^*}^*(x) \big)^2 \right\} \right]^{1/2}.$$

Table D.1: Empirical sizes and powers of the two-sided Wald tests at $\alpha = 5\%$. The true effect is $\tau = 2.225$. Naive denotes the balanced design with $A = \{1, 2, 3\}$ and $\varrho(X) = 1/2$, with the common pilot budget proportion. Regret-optimal denotes the comparator based on the minimax proportional-regret design of Epanomeritakis & Viviano (2025).

	Size: $\tau_0 = 2.225$				Power: $\tau_0 = 2.1$			
n	EFFICIENTESC	Oracle	Naive	Regret-optimal	EFFICIENTESC	Oracle	Naive	Regret-optimal
1000	0.056	0.053	0.054	0.049	0.160	0.150	0.131	0.140
2000	0.056	0.058	0.051	0.047	0.258	0.254	0.224	0.239
3000	0.051	0.054	0.057	0.052	0.377	0.380	0.324	0.356
4000	0.051	0.055	0.051	0.049	0.472	0.471	0.404	0.438
5000	0.048	0.048	0.047	0.047	0.590	0.594	0.509	0.555
6000	0.056	0.056	0.056	0.050	0.654	0.652	0.569	0.614
7000	0.052	0.055	0.053	0.053	0.732	0.733	0.648	0.699
8000	0.050	0.050	0.049	0.047	0.779	0.779	0.704	0.745
9000	0.047	0.047	0.053	0.047	0.831	0.829	0.759	0.799
10000	0.052	0.053	0.057	0.050	0.863	0.861	0.797	0.832
	Power: $\tau_0 = 2.0$				Power: $\tau_0 = 1.9$			
n	EFFICIENTESC	Oracle	Naive	Regret-optimal	EFFICIENTESC	Oracle	Naive	Regret-optimal
1000	0.371	0.359	0.300	0.344	0.635	0.621	0.530	0.608
2000	0.643	0.643	0.561	0.620	0.920	0.913	0.861	0.908
3000	0.824	0.821	0.740	0.800	0.988	0.989	0.967	0.982
4000	0.928	0.925	0.874	0.905	0.999	0.999	0.992	0.998
5000	0.971	0.976	0.940	0.962	1.000	1.000	0.999	1.000
6000	0.986	0.987	0.970	0.984	1.000	1.000	1.000	1.000
7000	0.995	0.996	0.987	0.994	1.000	1.000	1.000	1.000
8000	0.999	0.999	0.995	0.998	1.000	1.000	1.000	1.000
9000	1.000	1.000	0.998	1.000	1.000	1.000	1.000	1.000
10000	1.000	1.000	0.999	1.000	1.000	1.000	1.000	1.000

The probability of selecting A^* rises from 0.466 at $n = 500$ to 0.998 at $n = 5{,}000$. The assignment-rule RMSE decreases from 0.093 at $n = 500$ to 0.015 at $n = 10{,}000$. Together, these patterns show that both the selected secondary-outcome set and the associated treatment-assignment rule approach the oracle design as the sample size increases, consistent with Theorem 1.

D.3.3 SENSITIVITY TO THE PILOT BUDGET PROPORTION

The main simulation uses $\lambda_{P,b} = 3N_b^{-1/3}$. To examine sensitivity to this choice, we repeat the complete EFFICIENTESC method under

$$\lambda_{P,b} = a_P N_b^{-1/3}, \qquad a_P \in \{1, 2, 3, 4, 5\}.$$

The total experimental budget b is held fixed at each value of n. Changing a_P therefore changes only how that budget is divided between the pilot and the main experiment: a larger a_P provides more data for experimental design but leaves less budget for final effect estimation.

For each pilot budget proportion, we report $n \operatorname{Var}(\widehat{\tau})$ over the same observational sample-size grid as in the main simulation. All other features of the DGP and estimation procedure are unchanged. This comparison isolates the finite-sample trade-off between learning the design more accurately and reserving experimental resources for the main experiment.

Figure D.2 shows that the normalized variance generally decreases as a_P becomes smaller for a given observational sample size. In the range considered, the additional main experimental budget

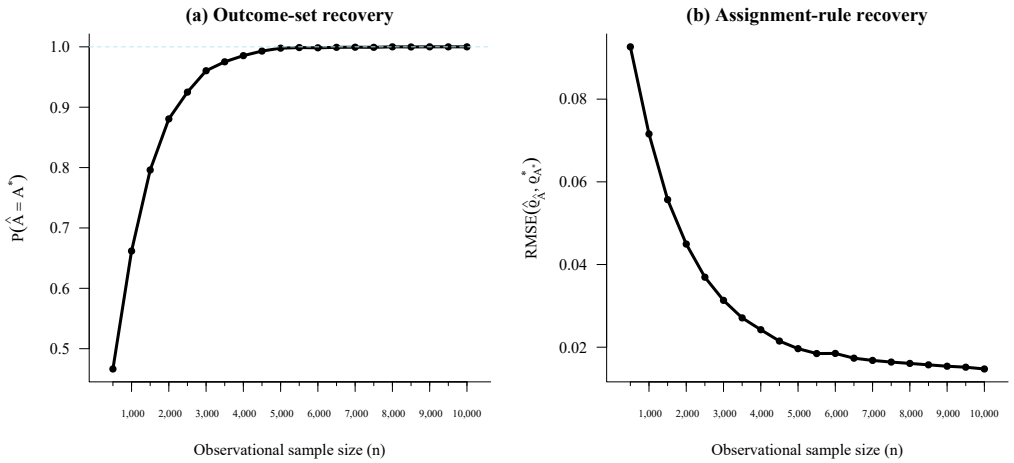

Figure D.1: Recovery of the oracle design over 5,000 independent Monte Carlo replications. Left: empirical probability of selecting $A^* = \{1, 2\}$. Right: root mean squared error of the learned treatment-assignment rule $\widehat{\varrho}_{\widehat{A}}$ relative to the oracle treatment-assignment rule $\varrho^*_{A^*}$, whose probabilities are $(0.702, 0.298)$.

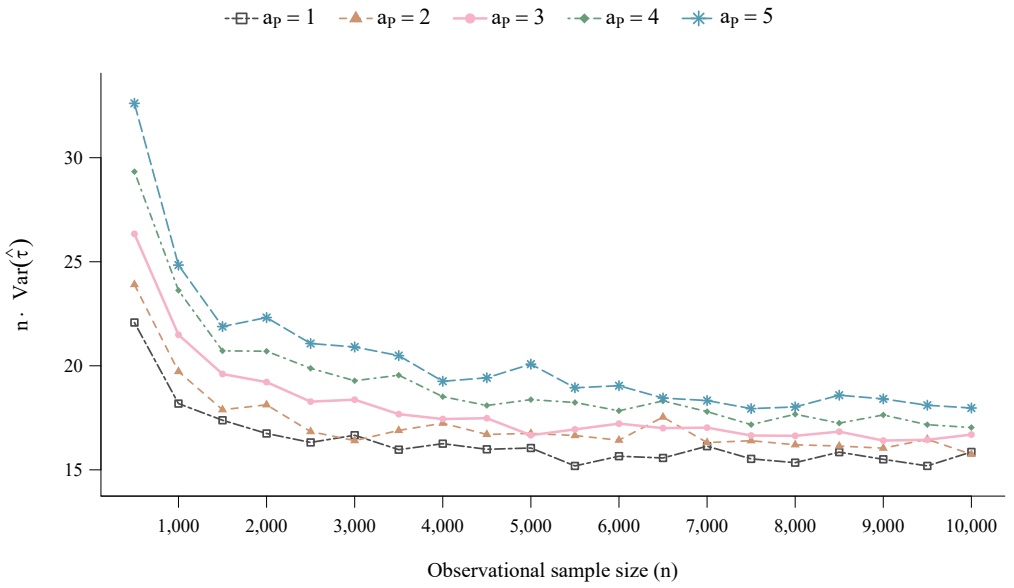

Figure D.2: Sensitivity of EFFICIENTESC to the pilot budget proportion $\lambda_{P,b} = a_P N_b^{-1/3}$ for $a_P \in \{1, 2, 3, 4, 5\}$. The total experimental budget is held fixed at each observational sample size.

obtained from using a smaller pilot outweighs the benefit of allocating more observations to pilot-stage design learning. The differences are most pronounced at smaller sample sizes and become less substantial as n increases. For all five choices of a_P, the normalized variance decreases and gradually stabilizes with increasing n, consistent with the asymptotic convergence of EFFICIENTESC. These results indicate that the method is not sensitive to a particular choice of the multiplicative constant in the $N_b^{-1/3}$ pilot budget proportion, while smaller constants can yield better finite-sample efficiency by reserving more of the experimental budget for the main experiment.

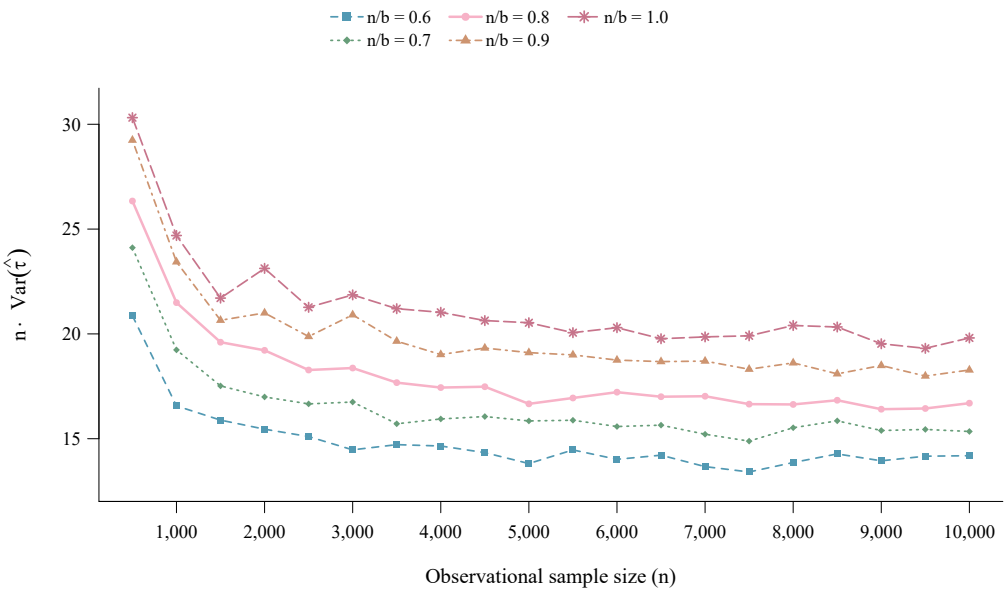

Figure D.3: Sensitivity of EFFICIENTESC to the observation-to-budget ratio. The figure plots $n\ \mathrm{Var}(\widehat{\tau})$ against the observational sample size n for $n/b \in \{0.6, 0.7, 0.8, 0.9, 1.0\}$. Smaller values of n/b correspond to larger experimental budgets at a fixed n.

D.3.4 SENSITIVITY TO THE OBSERVATION-TO-BUDGET RATIO

The main simulation fixes the observation-to-budget ratio at $n/b = 0.8$. We vary this ratio to examine how the performance of EFFICIENTESC changes when the amount of observational information and the experimental budget grow at different relative rates. For each prespecified value of n/b, we determine the experimental budget at each n and recompute N_b, the pilot size, and the remaining main experimental budget using the same allocation rule. The DGP, candidate secondary-outcome sets, assignment-rule class, and estimation procedure are otherwise unchanged.

For each value of n/b, we report $n\,\mathrm{Var}(\widehat{\tau})$ over the common observational sample-size grid. At a fixed n, different values of n/b imply different experimental budgets. The purpose of this analysis is to assess how the procedure behaves under different relative amounts of observational and experimental information.

Figure D.3 shows a clear ordering across the five observation-to-budget ratios. For a fixed observational sample size n, smaller values of n/b, which correspond to larger experimental budgets, lead to smaller normalized variances in the main experiment. Thus, holding the observational information fixed, a larger experimental budget improves the efficiency of the treatment effect estimator. At the same time, all curves stabilize as n increases, which is consistent with convergence to their respective asymptotic limits.

These results complement the main findings in two ways. First, they show that the relative amount of experimental budget has a direct effect on efficiency through the precision of the main experimental estimator. Second, together with the comparisons in the main simulation, they highlight that secondary outcomes are useful not only for debiasing under selection but also for improving estimation efficiency.

D.3.5 FULL-BUDGET COMPARISON FOR PRESPECIFIED DESIGNS

The main simulation uses a common pilot budget proportion for all methods so that their final experiments are conducted with the same post-pilot budget. This provides a comparison of design choices, but it is conservative for the balanced designs: once both the secondary-outcome set and the treatment-assignment rule are fixed in advance, these methods do not require an experimental

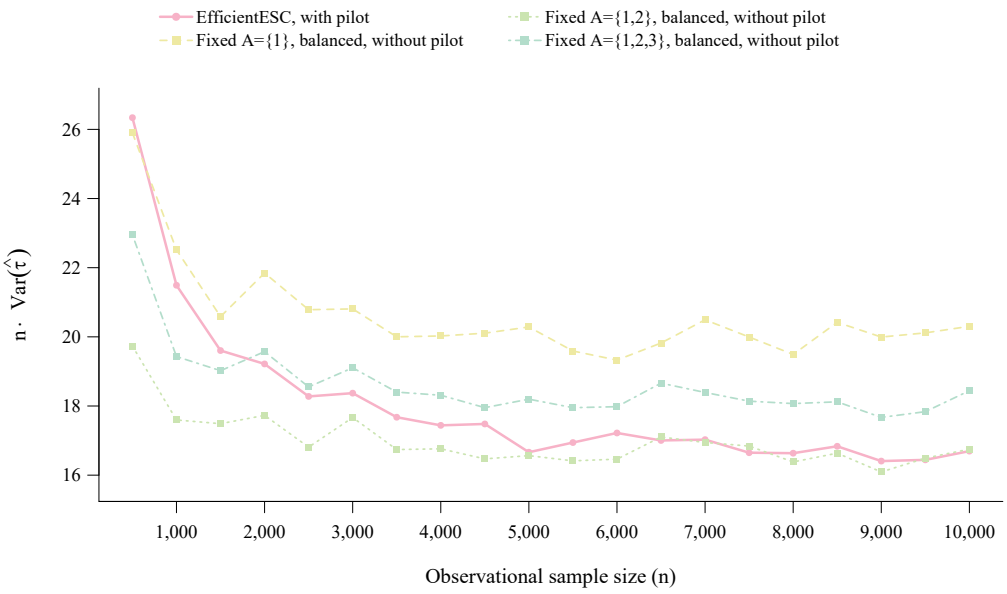

Figure D.4: Full-budget comparison of EFFICIENTESC and three prespecified balanced designs. At each n, all methods have the same total experimental budget $b = n/0.8$. The fixed designs measure the indicated secondary-outcome set, assign treatment with probability $1/2$, and use the entire budget without a pilot. EFFICIENTESC retains its pilot budget proportion and excludes pilot observations from final estimation. Points show $n \operatorname{Var}(\widehat{\tau})$.

pilot. We therefore conduct an additional comparison in which the three balanced designs,

$$A = \{1\}, \qquad A = \{1, 2\}, \qquad A = \{1, 2, 3\},$$

use $\varrho(X) = 1/2$, collect no pilot experiment, and devote the entire experimental budget b to the main experiment. The observational sample and final estimation procedure are otherwise the same as in the main simulation. EFFICIENTESC, by contrast, retains its original pilot-main allocation, uses the pilot to learn A and ϱ, and excludes the pilot observations from final effect estimation.

This comparison is deliberately different from the comparison in the main paper. It asks whether the variance reduction achieved by learning the secondary-outcome set and treatment-assignment rule is sufficient to offset the finite-sample experimental budget used for experimental design when the alternative is a fully prespecified design that can devote its entire budget to final estimation. We report $n \operatorname{Var}(\widehat{\tau})$ for all four methods over the same observational sample-size grid.

Figure D.4 shows that the comparison depends on sample size and the prespecified secondary-outcome set. At smaller sample sizes, the full-budget balanced design with $A = \{1, 2\}$ has smaller normalized variance than EFFICIENTESC. As n increases, the gap narrows, and the two methods have similar normalized variances at larger sample sizes. Over this larger-sample range, EFFICIENTESC also has smaller normalized variance than the full-budget balanced designs with $A = \{1\}$ and $A = \{1, 2, 3\}$.

These results illustrate the finite-sample cost of learning the design. When the informative secondary-outcome set is specified in advance, using the entire budget for final estimation can provide a substantial advantage at small sample sizes. Learning the design becomes more competitive as the pilot budget proportion decreases and the design estimates improve. Thus, the gains observed under a common pilot budget proportion should not be interpreted as uniform improvements over prespecified designs that require no pilot. This distinction is consistent with the discussion in Remark 2.