EXPLORE OVERLAP BOUNDARY: CAUSAL EFFECT BOUNDS VIA ASSIGNMENT REPLAY FROM STOCHASTIC ALGORITHM UNDER STRUCTURAL NON-OVERLAP

Anonymous authors
Paper under double-blind review

ABSTRACT

Stochastic algorithms assign personalized treatments such as videos or contexts whose features and outcomes vary across users. We study the conditional treatment-averaged causal effect of the treatment feature (C-TACE), with treatments generated from a stochastic algorithm. Under structural non-overlap, this estimand is unidentified as the non-realized treatment feature is concealed. We derive replay-aware causal bounds on C-TACE under non-overlap. When propensities are not logged, repeated replays are applied to estimate the propensity together with the non-realized feature. We show that the boundaries are in the form of function roots and that the bound is sharp under conditions, and we construct boundary estimators at finite replay counts. Rigorous theory has been developed on its asymptotic coverage. We also develop a replay-budget allocation method to improve bound efficiency. The method is comprehensively evaluated on semi-synthetic KuaiRec data, providing an applicable solution to non-overlap problem on causal effect under a stochastic algorithm.

1 INTRODUCTION

In many application scenarios of generative AI, we aim at evaluating the impact on users' feedback from contexts generated by large models with different tones, or evaluating the effect of AI-powered video advertisements in recommendation systems. The treatments, contexts or videos, are generated from a stochastic algorithm, and practitioners are interested in studying the causal effect of a treatment feature. For example, among users shown a video longer than the non-realized video available to them, what was the causal effect on user purchase? The two videos can differ in other attributes besides duration. However, the comparison between the duration of videos defines the population over which the treatment effect is of interest. Such a causal effect of the treatment feature is formalized as the conditional treatment-averaged causal effect (C-TACE) in Athey et al. (2026), adaptive to the setting in which treatments are generated by a stochastic algorithm.

In Athey et al. (2026), the C-TACE is identified and estimated based on available non-realized features, i.e., the features under the alternative treatment which is not realized in the observation must be logged from the stochastic algorithm, e.g., through replay. However, such availability essentially requires a positive assignment probability $e_i \neq 0, 1$ for each individual i, which is referred to as the overlap assumption in causal inference. Violation of the overlap assumption is common in applications, when only one assignment is available for some units. The non-overlap problem induces a critical challenge in effect identification from two aspects. First, similar to classical effects such as conditional average treatment effect (CATE), the non-overlap population with $e_i \in \{0, 1\}$ makes the inverse propensity ill-defined and therefore the full effect is unidentifiable. Second, under the mechanism where the non-realized feature can only be attained through replay, we can never log the non-realized feature in the non-overlap population, causing a challenge in population stratification. This challenge motivates us to develop new methods to evaluate C-TACE under structural non-overlap.

Constructing a partial identification bound on the targeted effect is an effective approach to evaluate causal effects under non-overlap in causal inference. Susmann et al. (2026) bound the full average treatment effect under non-overlap by combining identified and unsupported assignment regions. Lee & Weidner (2026) bound average effects using observations with shared covariates. However, these

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.

methods generally study causal bounds on ATE or CATE, and cannot be directly extended to C-TACE based on two different aspects under a stochastic algorithm. First, as the bound construction is different from existing partial identification bounds, this requires a new estimator and bound-validity theory. Second, as the propensity is not logged and needs to be estimated from the replay, we need to allocate the replay budget so as to stratify the overlap population by exploring the overlap boundary. This raises an extra problem in the construction of a causal bound under a stochastic algorithm. Further related works are discussed in Appendix A.

To deal with the challenges above, in this work, we develop a thorough method to evaluate C-TACE under non-overlap by constructing a causal bound. We start by re-visiting C-TACE identification on the overlap population, explain its implication that motivates us to construct a causal bound on C-TACE in the full population with non-overlap by bounding the effect outside an overlap population with propensity trimmed by threshold c. The boundaries can be expressed as unique roots of identifiable, decreasing functions. Under the stated feature-support condition, the bound converges to a sharp interval as c tends to zero. Then, we consider how the replay mechanism can construct a replay-aware C-TACE bound when the propensities need to be estimated from the replay, and develop the asymptotic properties of the estimator, which guarantee the asymptotic validity of the estimated causal bound. Adaptive to the replay mechanism, we further propose methods to choose the threshold and allocate the replay budget to improve the efficiency. The method comparison as well as ablation studies is conducted in the semi-synthetic KuaiRec benchmark. The framework of the paper is shown in Figure 1. The main contributions of this paper can be summarized as follows:

- We construct a causal bound on C-TACE under non-overlap which can be sharp under a choice of threshold, broadening its application in real-world scenarios.
- This method leverages the replay mechanism to discover the overlap boundary and considers the bound efficiency under a replay budget. This addresses the unique problems in causal effect under a stochastic algorithm, inspiring future work to study the replay mechanism in other scenarios.
- The bound is theoretically supported by the asymptotic properties discussed in this paper. The asymptotics of the boundaries as function roots provide a new theoretical tool for causal bounds in a similar form.

The paper is organized as follows. Section 2 defines the basic framework. Sections 3 and 4 develop identification and inference. Section 5 studies design choices, and Section 6 reports the experiments. The appendix reviews related work and gives proofs and additional results.

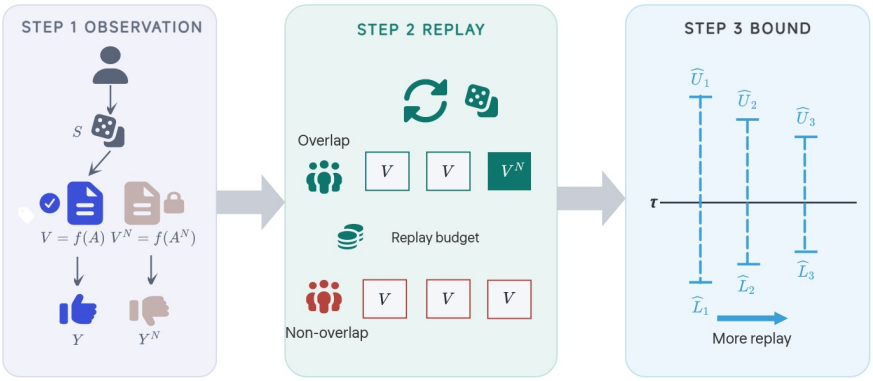


Figure 1: Framework on replay-aware C-TACE bound under non-overlap.

2 Causal Framework and Data Mechanism

2.1 Basic setting and scientific question

Consider an experiment with n units indexed by $i = 1, \ldots, n$. Unit i has covariates X_i and a pair of potential treatments $A_i(0), A_i(1) \in \mathcal{A}$, where $\mathcal{A}$ is a large treatment space, such as videos or responses. A binary *assignment* $W_i \in \{0, 1\}$ selects between them. The *realized treatment* is

$A_i = A_i(W_i)$, and the *non-realized treatment* is $A_i^N = A_i(1 - W_i)$. Let $Y_i(a)$ be the potential outcome under treatment a, and write $Y_i^w = Y_i(A_i(w))$ for $w \in \{0, 1\}$. The causal structure is summarized in Figure 2. We assume no interference, consistency, and bounded potential outcomes $Y_i^w \in [0, 1]$. Thus $Y_i = Y_i^{W_i}$ is the observed outcome, whereas $Y_i^N = Y_i^{1-W_i} = Y_i(A_i^N)$ is the counterfactual outcome under the non-realized treatment. The realized contrast $\tau_i = Y_i - Y_i^N$ equals $Y_i^1 - Y_i^0$ when $W_i = 1$ and $Y_i^0 - Y_i^1$ when $W_i = 0$.

Our scientific interest concerns a prespecified scalar feature of the treatments, such as video duration or the degree of guidance in a response. For a feature map $f : \mathcal{A} \to \mathbb{R}$, define the realized feature as $V_i = f(A_i)$ and the non-realized feature as $V_i^N = f(A_i^N)$. We restrict the study population to candidate pairs with $f(A_i(0)) \neq f(A_i(1))$ and assume $\mathbb{P}(V_i > V_i^N) > 0$. Using the treatment-characterizing function (TCF) $h(v, v') = \mathbf{1}(v > v')$, we define the population conditional treatment-averaged causal effect (C-TACE)

$$\tau = \mathbb{E}\{Y_i - Y_i^N \mid h(V_i, V_i^N) = 1\} = \mathbb{E}(Y_i - Y_i^N \mid V_i > V_i^N). \quad (1)$$

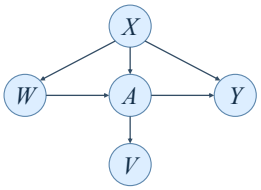


Figure 2: Causal structure.

Scientific question. Among users whose realized treatment has a larger feature value than their non-realized treatment, what is the average effect of receiving the realized treatment rather than the non-realized treatment?

For video duration, this C-TACE averages contrasts between complete videos among users whose realized video is longer than their non-realized video. Other video attributes may also differ within each pair. The feature comparison defines the target population, and the full video pair defines each causal contrast.

2.2 Stochastic assignment and replay

The stochastic algorithm assigns one of the two candidate treatments to each unit. Define the pre-assignment state $\mathcal{G}_i = (X_i, A_i(0), A_i(1), Y_i^0, Y_i^1)$ and propensity $e_i = \mathbb{P}(W_i = 1 \mid \mathcal{G}_i)$. States $\mathcal{G}_i$ are independent draws from a superpopulation. Conditional on $\mathcal{G}_i$, probabilities average over observation assignment and replay randomness. Unconditional expectations also average over the superpopulation. Thus Equation (1) conditions on the realized feature comparison in the joint population.

Assumption 1 (Within-pair unconfoundedness)**.**

$$W_i \perp\!\!\!\perp (Y_i^0, Y_i^1) \mid X_i, A_i(0), A_i(1), \qquad e_i = \mathbb{P}\{W_i = 1 \mid X_i, A_i(0), A_i(1)\}.$$

Assumption 1 states that, conditional on covariates and the candidate pair, assignment is independent of both potential outcomes. The displayed propensity equality then relates the assignment probability defined through $\mathcal{G}_i$ to the covariates and candidate pair. Outcomes are indexed by the received treatment, so the assignment label has no separate effect on Y_i.

Write $V_i^w = f(A_i(w))$. An offline replay reruns the fixed algorithm on the same inputs for unit i with a fresh random seed, and denotes $M_i \geq 1$ as the number of replay calls. Its assignment is $W_i^{(j)}$ on replay $j = 1, \ldots, M_i$, with $W_i^{(0)} = W_i$ denoting the observation assignment. We impose the following condition on the replay mechanism.

Assumption 2 (Replayable stochastic assignment)**.** Conditional on $\mathcal{G}_1, \ldots, \mathcal{G}_n$, the observed and replay assignments can be generated independently following the same mechanism as

$$W_i^{(j)} \mid \mathcal{G}_i \sim \text{Bernoulli}(e_i), \qquad i = 1, \ldots, n, \quad j = 0, \ldots, M_i.$$

Replay j returns the feature $V_i^{(j)} = f(A_i(W_i^{(j)}))$ of the treatment selected on that replay. It generates no outcome because no new user response is collected. The replay service retains the inputs and two candidate treatments needed to reproduce the fixed pair. The analyst's record is $\big(Y_i, V_i, V_i^{(1)}, \ldots, V_i^{(M_i)}\big)$, without the non-realized treatment A_i^N, its feature V_i^N, the assignment propensity, or the latent comparison Z_i. In Athey et al. (2026), the non-realized feature is logged.

Here a service can accept a saved user and pair identifier, rerun the selector with fresh seeds, and return the selected feature. If $e_i \in (0,1)$, the non-realized feature eventually appears almost surely as the replay count grows. At $e_i = 0$ or 1, replay never reveals it, so the comparison $V_i > V_i^N$ can remain unresolved.

Replay budgets. The replay count $M_i \geq 1$ is fixed before observation assignment by a rule based on pre-assignment information or an independent pilot. Conditional on $\mathcal{G}_i$, the count is independent of observation and replay assignments. The base method uses a common count $M_i = M$. A released pre-assignment stratum label may determine M_i for the allocation design.

3 Identification and Bounds under Non-overlap

3.1 Identification in the overlap population

Write $\mathbf{1}(\cdot)$ for an indicator. Let $J_i = \mathbf{1}(V_i^1 > V_i^0)$, $Z_i = \mathbf{1}(V_i > V_i^N) = \mathbf{1}(W_i = J_i)$ indicate whether the realized treatment has the higher feature value, and define its assignment probability as

$$Q_i = \mathbb{P}(Z_i = 1 \mid \mathcal{G}_i) = J_i e_i + (1 - J_i)(1 - e_i).$$

We first identify C-TACE in the *overlap population*, comprising units with $e_i \in (0,1)$ or, equivalently, $Q_i \in (0,1)$. With sufficiently many replays, V_i^N and hence J_i and Z_i are observed in this population. We also assume known propensities in this section, which determine Q_i. For a unit i in the overlap population, $\mathbb{E}\left[(1 - Z_i)\frac{Q_i Y_i}{1-Q_i} \,\middle|\, \mathcal{G}_i\right] = Q_i Y_i^{1-J_i}$. Let $I_i = \mathbf{1}(e_i \in (0,1))$ denote whether the individual belongs to the overlap population. C-TACE τ_{ov} in overlap population can be identified by

$$\tau_{\text{ov}} = \mathbb{E}(Y_i - Y_i^N \mid Z_i = 1, I_i = 1) = \frac{\mathbb{E}[I_i\{Z_i Y_i - (1 - Z_i)Q_i Y_i/(1 - Q_i)\}]}{\mathbb{E}[I_i Z_i]}. \tag{2}$$

The inverse-weight term in Equation (2) is evaluated only on $I_i = 1$ and defined as zero otherwise. At deterministic assignments, Q_i may be zero or one, and the non-realized feature and counterfactual outcome need not be observed. The next section bounds the full-population C-TACE.

3.2 Partial identification bound on C-TACE under non-overlap

To construct a partial identification bound on C-TACE, we first consider known assignment propensities and infinitely many replays. Therefore, for units with $I_i = 1$, replay eventually reveals the non-realized feature V_i^N, and for units with $I_i = 0$, assignment is deterministic and replay reveals only the realized feature.

For $0 \leq c < 1/2$, define the trimmed population by $O_i(c) = \mathbf{1}(e_i \in (c, 1-c))$ and set $B_i(c) = 1 - O_i(c)$. The threshold c excludes extreme propensities and trades a wider bound for more stable estimation near the overlap boundary (Susmann et al., 2026). Define $d(c) = \mathbb{E}[O_i(c)Z_i]$ and $a(c) = \mathbb{E}\left[O_i(c)\left\{Z_i Y_i - (1 - Z_i)\frac{Q_i Y_i}{1-Q_i}\right\}\right]$, where the integrands are defined as zero outside $O_i(c) = 1$. The full-population C-TACE satisfies

$$\tau = \frac{a(c) + \mathbb{E}[B_i(c)Z_i(Y_i - Y_i^N)]}{d(c) + \mathbb{E}[B_i(c)Z_i]}. \tag{3}$$

The terms $d(c)$ and $a(c)$ are identifiable since V_i^N and hence Z_i is known for individuals in the trimmed population with $O_i(c) = 1$. However, we need to bound the two remaining expectations since Z_i is not identifiable on the whole population. For a candidate effect t, bounded outcomes imply

$$\min\{0, Y_i - 1 - t\} \leq Z_i(Y_i - Y_i^N - t) \leq \max\{0, Y_i - t\}.$$

Therefore, the partial identification bound can be constructed as follows.

Theorem 1 (Full-population C-TACE bound)**.** *Define two identifiable, decreasing functions*

$$F_-(t; c) = a(c) - td(c) + \mathbb{E}[B_i(c)\min\{0, Y_i - 1 - t\}],$$
$$F_+(t; c) = a(c) - td(c) + \mathbb{E}[B_i(c)\max\{0, Y_i - t\}].$$

Under Assumptions 1–2, if $d(c) > 0$, each equation $F_-(L(c); c) = F_+(U(c); c) = 0$ has a unique root in $[-1, 1]$, and

$$\tau \in [L(c), U(c)].$$

Note that if $d(0) > 0$, then $L(c) \uparrow L(0)$ and $U(c) \downarrow U(0)$ as $c \downarrow 0$, and therefore the causal bound is tightened with a smaller choice of c. The interval $[L(0), U(0)]$ is sharp if, conditional on each observed deterministic record and any released design information, the non-realized feature may lie on either side of the realized feature and its counterfactual outcome may take any value in $[0, 1]$. Under these conditions, both endpoints can be attained by full-data distributions that generate the same observed data. If a stratum label or feature-support restriction rules out one ordering, the interval remains valid but may not be sharp.

4 Replay-Aware Estimation and Inference

4.1 Replay-aware bound when propensities are unlogged

Theorem 1 uses assignment propensities that may not be logged. We instead construct observable scores from replayed features. A distinct replay feature reveals the comparison on that record, while deterministic assignments remain unresolved.

For each unit, let $L_i = \sum_{j=1}^{M_i} \mathbf{1}(V_i^{(j)} < V_i)$ and $H_i = \sum_{j=1}^{M_i} \mathbf{1}(V_i^{(j)} > V_i)$. Only one count can be positive because the treatment pair is fixed. Put $N_i = M_i + 1$ and choose $k_i = \max\{1, \lfloor cN_i \rfloor\}$ for $0 \le c < 1/2$. We retain a record when each feature appears at least k_i times among the observation and replays. Define $R_i = \mathbf{1}\{k_i \le L_i + H_i \le N_i - k_i\}$ and $B_i = 1 - R_i$. Here $B_i = 1$ denotes replay non-retention, whereas $B_i(c)$ in Section 3.2 denotes exclusion by a known propensity threshold. The observable retained-target and outcome scores are

$$D_i = R_i \mathbf{1}(L_i > 0), \qquad T_i = Y_i \left(D_i - R_i \frac{H_i}{N_i - H_i} \right).$$

The denominator $N_i - H_i$ is at least k_i whenever $R_i = 1$.

Proposition 2 (Replay-aware C-TACE bound)**.** *Suppose M_i follows the pre-assignment rule in Section 2.2. For a fixed threshold c, define*

$$g_i = Q_i \mathbb{P}\{k_i \le \mathrm{Bin}(M_i, 1 - Q_i) \le N_i - k_i \mid \mathcal{G}_i, M_i\}.$$

Then

$$\mathbb{E}(D_i \mid \mathcal{G}_i, M_i) = g_i, \qquad \mathbb{E}(T_i \mid \mathcal{G}_i, M_i) = g_i(Y_i^{J_i} - Y_i^{1-J_i}). \tag{4}$$

With $a = \mathbb{E}T_i$, $d = \mathbb{E}D_i > 0$, the roots of

$$\tilde{F}_-(t) = a - td + \mathbb{E}[B_i \min\{0, Y_i - 1 - t\}], \qquad \tilde{F}_+(t) = a - td + \mathbb{E}[B_i \max\{0, Y_i - t\}]$$

satisfying $\tilde{F}_+(U) = \tilde{F}_-(L) = 0$ define a valid bound $[L, U]$ for τ.

The identity integrates over assignment without requiring e_i as an input. For $c = 0$, $k_i = 1$ and $g_i = Q_i - Q_i^{M_i+1}$. Increasing a common M contracts the population interval to $[L(0), U(0)]$ when $d(0) > 0$. For fixed $c > 0$, common $M \to \infty$, and $\mathbb{P}(e_i \in \{c, 1 - c\}) = 0$, the interval converges to $[L(c), U(c)]$ when $d(c) > 0$. This is an identification result for increasing replay counts. The inference result below holds for bounded replay counts.

The functions $\tilde{F}_-(t)$ and $\tilde{F}_+(t)$ need not be differentiable in t. We smooth their positive-part terms with $s_\epsilon(x) = \epsilon \log\{1 + \exp(x/\epsilon)\}$ for fixed $\epsilon > 0$. Define

$$\phi_{i,-}(t) = T_i - tD_i - B_i s_\epsilon(1 + t - Y_i), \qquad \phi_{i,+}(t) = T_i - tD_i + B_i s_\epsilon(Y_i - t). \tag{5}$$

The population roots solve $\mathbb{E}\phi_{i,-}(L_\epsilon) = \mathbb{E}\phi_{i,+}(U_\epsilon) = 0$. They provide a valid bound on τ with

$$L_\epsilon \le L \le \tau \le U \le U_\epsilon, \tag{6}$$

and $\max\{L - L_\epsilon, U_\epsilon - U\} \le \epsilon \log(2)\mathbb{E}(B_i)/d$. The estimators $\widehat{L}_\epsilon, \widehat{U}_\epsilon$ solve the corresponding sample equations on $\mathbb{R}$ when $\sum_i D_i > 0$. Each sample equation is then strictly decreasing and has a unique root. When $\sum_i D_i = 0$, Algorithm 1 reports $[-1, 1]$.

4.2 Asymptotic distribution and confidence coverage

We move on to discuss the bound estimation as well as the asymptotic properties of the estimated bound. Let $\ell(x) = (1 + e^{-x})^{-1}$ and define

$$\begin{aligned} A_- &= d + \mathbb{E}[B_i \ell\{(1 + L_\epsilon - Y_i)/\epsilon\}], \quad \psi_{i,-} = \phi_{i,-}(L_\epsilon)/A_-, \\ A_+ &= d + \mathbb{E}[B_i \ell\{(Y_i - U_\epsilon)/\epsilon\}], \qquad \psi_{i,+} = \phi_{i,+}(U_\epsilon)/A_+. \end{aligned} \tag{7}$$

For estimation, replace the expectations and population roots in $A_\pm$ by sample means and sample roots to obtain $\widehat{A}_\pm$. Set $\widehat{\psi}_{i,-} = \phi_{i,-}(\widehat{L}_\epsilon)/\widehat{A}_-$ and $\widehat{\psi}_{i,+} = \phi_{i,+}(\widehat{U}_\epsilon)/\widehat{A}_+$, and let $\widehat{\Sigma} = n^{-1}\sum_i (\widehat{\psi}_{i,-}, \widehat{\psi}_{i,+})^\mathsf{T}(\widehat{\psi}_{i,-}, \widehat{\psi}_{i,+})$. Fix c, ϵ and a replay rule $M_i = m(S_i)$ before collecting inference records, where S_i is pre-assignment information. The rule may be selected on an independent pilot, in which case all expectations and limits below are conditional on that pilot. Suppose the inference records are iid under this rule, $1 \le M_i \le M_{\max} < \infty$, $d > 0$, and both endpoint variances are positive. Then we have the following theorem.

Theorem 3 (Endpoint inference)**.** *Under these conditions, with* $\Sigma = \mathbb{E}[(\psi_{i,-}, \psi_{i,+})^\mathsf{T}(\psi_{i,-}, \psi_{i,+})]$,

$$\sqrt{n}\begin{pmatrix} \widehat{L}_\epsilon - L_\epsilon \\ \widehat{U}_\epsilon - U_\epsilon \end{pmatrix} \rightsquigarrow \mathcal{N}_2(0, \Sigma).$$

The empirical influence covariance $\widehat{\Sigma}$ *is consistent. Write* $\widehat{\sigma}_-^2 = \widehat{\Sigma}_{11}$, $\widehat{\sigma}_+^2 = \widehat{\Sigma}_{22}$ *and* $z_p = \Phi^{-1}(p)$, *where* Φ *is the standard normal distribution function. Then*

$$C_n = \left[\widehat{L}_\epsilon - z_{1-\alpha/2}\frac{\widehat{\sigma}_-}{\sqrt{n}}, \widehat{U}_\epsilon + z_{1-\alpha/2}\frac{\widehat{\sigma}_+}{\sqrt{n}}\right] \cap [-1, 1] \tag{8}$$

satisfies $\liminf_{n\to\infty} \mathbb{P}\{[L_\epsilon, U_\epsilon] \cap [-1, 1] \subseteq C_n\} \ge 1 - \alpha$ *and consequently* $\liminf_{n\to\infty} \mathbb{P}(\tau \in C_n) \ge 1 - \alpha$.

The estimators $\widehat{L}_\epsilon, \widehat{U}_\epsilon$ converge to the smoothed population endpoints, which enclose the replay-aware bound in Proposition 2. The resulting interval covers τ with asymptotic probability at least $1 - \alpha$. For a fixed collection of K candidate rules, use $z_{1-\alpha/(2K)}$ for each endpoint. Simultaneous coverage then permits selecting the shortest interval on the inference data. Algorithm 1 summarizes the procedure. Appendix D gives the proof and the stratified extension used for budget allocation.

Algorithm 1 Replay-aware C-TACE estimation

1: **Input:** observation $(Y_i, V_i)_{i=1}^n$, replay features, level α, and K rules (M_i, c, ϵ) fixed before observation assignment. Use $K = 1$ for one rule.
2: **for** each rule **do**
3: **Stage 1: Observable scores.** Count L_i, H_i and form R_i, D_i, B_i, T_i by definition.
4: **if** $\sum_i D_i = 0$ **then**
5: Set this rule's interval to $[-1, 1]$ and continue to the next rule.
6: **end if**
7: **Stage 2: Smooth endpoints.** Solve the two sample equations in (5) for $\widehat{L}_\epsilon, \widehat{U}_\epsilon$.
8: **Stage 3: Inference.** Estimate the influence variances from (7) at those roots.
9: Form (8) using $z_{1-\alpha/(2K)}$. Use $[-1, 1]$ if the clipped interval is empty.
10: **end for**
11: **Output:** The shortest candidate interval, with only one choice when $K = 1$.

5 Improve Efficiency under Replay Budget

5.1 Choosing the trimming threshold

In this section, we discuss how to improve the efficiency of the method by constructing more informative causal bounds through smart choices of the trimming threshold c and the replay budget allocation. A larger c limits inverse weights and retains fewer comparisons. We choose a threshold by minimizing an upper bound on population width plus normal-theory sampling uncertainty. Simultaneous intervals permit selection from a small prespecified grid.

Consider a common round of replay calls $M_i = M$, $N = M + 1 \geq 4$, and define effective threshold $\nu = k/N \in [1/N, 1/4]$. Suppose the retained target mass is at least $\delta > 0$ throughout this threshold range and, for some $C, \kappa > 0$,

$$\mathbb{P}\{0 < \min(e_i, 1 - e_i) \leq t\} \leq Ct^{\kappa}, \qquad 0 < t \leq 1/2. \tag{9}$$

Write $W_0 = U(0) - L(0)$ for the width of the bound with known-propensity and infinite replays, and let $z = z_{1-\alpha/2}$ for one confidence interval. The width from asymptotic normality $\mathcal{W}(\nu) = U_\epsilon - L_\epsilon + z(\sigma_- + \sigma_+)/\sqrt{n}$ obeys

$$\mathcal{W}(\nu) \leq W_0 + C_1\nu^{\kappa} + C_2N^{-\kappa} + \frac{2\epsilon \log 2}{\delta} + \frac{C_3}{\sqrt{n\nu}}, \tag{10}$$

where the constants depend on the stated tail and mass conditions, smoothing, and critical value. Appendix E derives the bound. Minimizing its threshold-dependent terms gives

$$\nu^* = \left[\left(\frac{C_3}{2\kappa C_1 \sqrt{n}} \right)^{1/(\kappa+1/2)} \right]_{[1/N,\, 1/4]}, \qquad k^* \approx N\nu^*, \tag{11}$$

and hence optimal $\nu^* \asymp n^{-1/(2\kappa+1)}$ is chosen at an inverse-polynomial rate of n. Here $[x]_{[a,b]} = \min\{b, \max(a, x)\}$. The oracle rule in Equation (11) depends on the unknown C_1, C_3, and κ. An integer k^* nearest to $N\nu^*$ gives $c^* = k^*/N$, with $1 \leq k^* \leq \lfloor N/4 \rfloor$. The choice $c = 0$ gives the same count threshold as $k = 1$. Away from the replay-resolution floor, $\nu^* \asymp n^{-1/(2\kappa+1)}$.

An independent tuning sample can choose a threshold without estimating C_1, C_3, or κ. Prespecify a finite grid $\mathcal{C} \subset [0, 1/4]$ with distinct effective thresholds $k(c) = \max\{1, \lfloor cN \rfloor\}$. For each $c \in \mathcal{C}$, compute the tuning-sample endpoints and influence standard deviations. With n_{inf} planned inference records, choose the c minimizing

$$\widehat{\mathcal{W}}_{\text{tun}}(c; n_{\text{inf}}) = \hat{U}_{\epsilon,\text{tun}}(c) - \hat{L}_{\epsilon,\text{tun}}(c) + \frac{z_{1-\alpha/2}}{\sqrt{n_{\text{inf}}}} \{\hat{\sigma}_{-,\text{tun}}(c) + \hat{\sigma}_{+,\text{tun}}(c)\}.$$

The independent choice permits ordinary endpoint calibration on inference data. The full-results comparison instead selects from a fixed grid on the inference sample using simultaneous intervals, as in Susmann et al. (2026). For $N < 4$, we use $k = 1$.

5.2 Allocating replay calls across units

For $c = 0$, conditional on $\mathcal{G}_i$, the probability that observation and all M_i replays show one feature is

$$r_i(M_i) = Q_i^{M_i+1} + (1 - Q_i)^{M_i+1} \leq 2\exp\{-\lambda_i(M_i + 1)\},$$

with $\lambda_i = -\log\{1 - \min(e_i, 1 - e_i)\}$. The non-discovery probability depends on the minority selection probability, regardless of which feature is larger. With a positive lower bound on retained target mass, excess effect width is bounded by a constant times unresolved interior mass. We minimize the sum of the exponential bounds subject to $\sum_i M_i = B$ and $m_{\min} \leq M_i \leq m_{\max}$. Its continuous optimum is

$$M_i^*(\eta) = \left[\frac{\log(2\lambda_i/\eta)}{\lambda_i} - 1 \right]_{[m_{\min}, m_{\max}]}, \tag{12}$$

satisfying $\sum_i M_i^*(\eta) = B$. For $\lambda_i = 0$, assign the minimum unless all positive-rate units reach the maximum. Find η by monotone search and round the continuous solution for an approximate integer allocation. An exact integer solution starts at $m_{\min}$ and assigns each remaining call to the largest marginal reduction $2(1 - e^{-\lambda_i})e^{-\lambda_i(M_i+1)}$; see Appendix E for details.

An independent pilot estimates stratum minority probabilities and fits the integer allocation. We compare that plan with uniform replay using a proxy for unresolved mass and score variability, then fix the selected plan before inference. Appendix E gives the pilot equations and selection rule.

6 Experiments

6.1 Experimental setup

Dataset. We evaluate on KuaiRec (Gao et al., 2022) using 1,687,549 user–video-pair records and 1,205 matched longer–shorter video pairs. Here X_i describes the user, $A_i(1)$ and $A_i(0)$ are the

candidate videos, and $W_i = 1$ means the logger selects the longer video with probability e_i. The realized and non-realized videos are $A_i = A_i(W_i)$ and $A_i^N = A_i(1 - W_i)$, with durations V_i and V_i^N. The outcome Y_i is completion of the shown video, and Y_i^N is completion under the unshown video. The question is the average effect of showing the longer rather than the shorter candidate among records where the logger selects the longer video. Both completion outcomes are available for evaluation. The simulated observational record releases Y_i and V_i, and each replay returns a fresh selected duration $V_i^{(j)}$.

Comparative methods. We compare replay with the observation-only set $[-1, 1]$, a pair-catalog bound, and three point estimators. The oracle overlap interval uses true propensities and targets τ_{ov}. The discovered-comparison estimator uses records with resolved feature order, and catalog inverse sampling applies Athey et al. (2026) using the recorded treatment pair. The replay variants select a threshold, allocate calls, or combine both rules.

Metrics and settings. We report full-target coverage, mean width, and correct-sign exclusion of zero, along with oracle overlap coverage of τ_{ov}. Across 500 replications, we vary deterministic assignment mass $\rho \in \{0, .1, .2, .4\}$, sample size $n \in \{2{,}000, 10{,}000, 50{,}000\}$, and replay count $M \in \{1, 2, 5, 10, 20, 50\}$. Basic replay uses $c = 0$, $\epsilon = .01$, and 95% calibration. Appendix F gives sampling and comparator details.

6.2 Main results

Table 1: Main comparison at $n = 50{,}000$ and $M = 20$ over 500 replications. Excluded is the percentage of full-target members outside overlap. Oracle coverage is evaluated against the full and overlap targets. Sharp is the known-propensity, infinite-replay identified width. Excess is replay confidence width minus sharp width, and sign is correct-sign exclusion of zero.

	Excluded	Oracle coverage (%)		Replay	Sharp	Replay	Excess	Sign
ρ	target (%)	Full	Own	coverage (%)	width	width	width	(%)
0.0	0.0	94.4	94.4	100.0	0.000	0.268	0.268	100.0
0.1	8.4	95.0	94.6	100.0	0.184	0.325	0.141	100.0
0.2	16.6	90.4	96.2	100.0	0.345	0.413	0.068	100.0
0.4	33.3	63.4	94.8	100.0	0.623	0.658	0.036	87.0

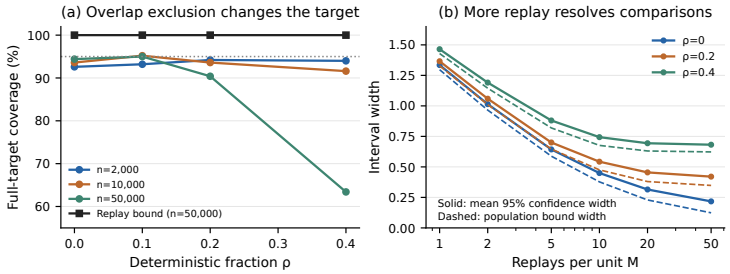


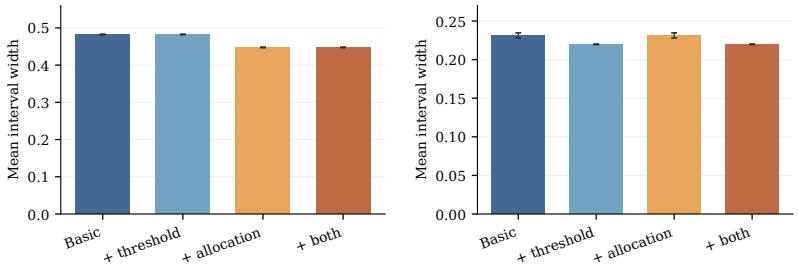
Figure 3: Main comparison over 500 replications. (a) The oracle overlap interval loses full-target coverage as non-overlap increases. (b) At $n = 10{,}000$, replay reduces interval width, while structural non-overlap leaves a positive limiting width.

The main results are shown in Table 1 and Figure 3, in which we address three key observations. (1) At $\rho = .4$, 33.3% of the full target lies outside overlap. The oracle overlap interval covers its own target in 94.8% of replications and the full C-TACE in 63.4%. Replay covers the full target in all 500 replications and excludes zero with the correct sign in 87.0%. The coverage gap tracks the target members excluded by overlap restriction. (2) Replays narrow the full-target interval. At $\rho = .4$, mean width is 0.658, compared with 2 for observation only and 1.008 for the pair-catalog bound. The sharp width is 0.623, leaving 0.036 of excess confidence width at $n = 50{,}000$ and $M = 20$. Figure 3(b) shows further narrowing as replay count rises. Deterministic assignments account for the positive

limiting width. (3) Point estimators lose full-target coverage as the extent of non-overlap increases. At $\rho = .4$, the discovered-comparison interval covers the full C-TACE in 71.0% of replications. Catalog inverse sampling has limiting bias $+.176$ and zero empirical full-target coverage in this setting (Appendix F). The replay bound keeps the uncertainty from unresolved feature order and counterfactual outcomes in its endpoints.

6.3 Ablation Study

In ablation study, we compare the C-TACE bound under different threshold selection and treatment-budget allocation strategies. While the basic replay uses a uniform replay allocation with threshold $c = 0$, we change its threshold selection and budget allocation to the method in Section 5 respectively and also combine both modifications together. We compare the variants under two different settings, the main experimental setting with non-overlap magnitude $\rho = .1$, and a sparse-overlap setting. Appendix F gives its construction and the full results.



(a) Main setting, $\rho = .1$ (b) Sparse-overlap

Figure 4: Mean interval width. Panel (a) uses 500 replications, 100,000 inference calls, and a 20,000-call tuning pilot. Panel (b) uses 5,000 replications, 200,000 inference calls, and a 40,000-call allocation pilot. Threshold is set as $c = .05$. Error bars are 95% confidence intervals.

Figure 4(a) shows that allocation reduces mean width from 0.482 to 0.448. Its reduction across the four ρ values is 4.4–7.2%. The pilot keeps the allocated plan and selects $c = 0$ in all replications. Thus threshold selection matches basic replay, the combined design matches allocation, and all four attain 100% empirical full-target coverage. With 10 replays, $c = .2$ widens the population bound by requiring two contrasting draws for retention.

In Figure 4(b), the higher-feature treatment has selection probability .99 for the top 5% of prediction gaps and .5 elsewhere. At $n = 2{,}000$ and $M = 100$, choosing threshold as $c = .05$ reduces mean width from 0.231 to 0.220 over 5,000 replications, a paired reduction of 0.012 ± 0.003 at 95% Monte Carlo confidence. The allocation chooses uniform replay in every replication in this setting. The combined design matches the threshold width with the same pilot cost. Coverage is 95.4% for basic and allocation and 100% for threshold and combined. Appendix F reports the wider population bound at $c = .05$.

7 Discussion

This paper studies C-TACE under non-overlap when deterministic assignment hides both the non-realized feature and the counterfactual outcome. We derive full-population bounds by varying these unknown quantities jointly, use replay counts to construct observable endpoint estimations, and obtain confidence intervals from smoothed root estimators. The KuaiRec experiments show that replay narrows the bounds while retaining full-target coverage across the evaluated overlap settings.

Our central contribution is to bound the original C-TACE when target membership itself is unknown. The limiting interval is sharp under the stated support conditions, and threshold and allocation rules address finite-replay width. Future work could consider more advanced replay allocation to reduce the total replay cost, for instance, under cost heterogeneity. Other directions include developing threshold selection by leveraging rare informative replays, extending the framework to multiple treatments, or consider other non-identifiable cases such as the existence of unmeasured confoundings.

Ethics Statement

We have read and adhere to the ICLR Code of Ethics throughout this research. This work does not involve new human-subject recruitment, data collection, or intervention. The semi-synthetic datasets are used only for method evaluation. The proposed framework is intended for methodological research, and any real-world deployment would require application-specific consideration of ethical, legal, and fairness requirements.

Reproducibility Statement

The implementation details on experimental configurations have been provided in the appendix. The source code and scripts for reproducing the reported results are provided at the following anonymous repository: `https://anonymous.4open.science/r/C-TACE-N-1847/`.

AI use statement

Generative AI assisted with manuscript editing, experiment-code changes, figure preparation, and drafting the interpretation of results. The numerical results were generated by the reported code and checked against saved replications. The authors retain responsibility for reviewing and verifying the final manuscript.

References

Susan Athey, Guido Imbens, and Zoe Ji. Estimating causal effects from data generated by stochastic algorithms. *arXiv preprint arXiv:2607.05792*, 2026.

Hui Bai, Jan G. Voelkel, Shane Muldowney, Johannes C. Eichstaedt, and Robb Willer. LLM-generated messages can persuade humans on policy issues. *Nature Communications*, 16:6037, 2025.

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

Chongming Gao, Shijun Li, Wenqiang Lei, Jiawei Chen, Biao Li, Peng Jiang, Xiangnan He, Jiaxin Mao, and Tat-Seng Chua. KuaiRec: A fully-observed dataset and insights for evaluating recommender systems. In *Proceedings of the 31st ACM International Conference on Information & Knowledge Management*, pp. 540–550, 2022.

Masahiro Kato, Akihiro Oga, Wataru Komatsubara, and Ryo Inokuchi. Active adaptive experimental design for treatment effect estimation with covariate choice. In *Proceedings of the 41st International Conference on Machine Learning*, volume 235 of *Proceedings of Machine Learning Research*, pp. 23291–23323, 2024.

Edward H. Kennedy. Nonparametric causal effects based on incremental propensity score interventions. *Journal of the American Statistical Association*, 114(526):645–656, 2019.

Sokbae Lee and Martin Weidner. Bounding treatment effects by pooling limited information across observations. *Journal of Econometrics*, 256:106254, 2026.

Victoria Lin, Louis-Philippe Morency, and Eli Ben-Michael. Isolated causal effects of natural language. In *Proceedings of the 42nd International Conference on Machine Learning*, volume 267 of *Proceedings of Machine Learning Research*, pp. 37919–37941, 2025.

Francesco Salvi, Manoel Horta Ribeiro, Riccardo Gallotti, and Robert West. On the conversational persuasiveness of GPT-4. *Nature Human Behaviour*, 9:1645–1653, 2025.

Herbert P. Susmann, Alec McClean, and Iván Díaz. Non-overlap average treatment effect bounds. *arXiv preprint arXiv:2509.20206*, 2026. Version 2, revised May 2026.

Zhiheng Zhang, Haoxiang Wang, Haoxuan Li, and Zhouchen Lin. Active treatment effect estimation via limited samples. In *Proceedings of the 42nd International Conference on Machine Learning*, volume 267 of *Proceedings of Machine Learning Research*, pp. 76225–76245, 2025.

A Related Work

LLM interventions and effects of content features. Randomized studies show that generated language can affect human responses. Bai et al. (2025) compare LLM-generated policy messages with human-written and neutral messages. Salvi et al. (2025) study persuasion in conversations with GPT-4 and human opponents. Their interventions can change several message attributes at once. Lin et al. (2025) study a focal linguistic intervention while accounting for non-focal language. For stochastic content selection, Athey et al. (2026) define C-TACE to average contrasts between complete realized and non-realized treatments within a population selected by an interpretable feature comparison. Their identification results use the logged non-realized content and relative assignment probabilities, or additional exposures and replay. In our observation model, the service releases the feature selected on each replay. At deterministic assignments, the non-realized feature remains unresolved.

Non-overlap and violations of positivity. Weak or absent overlap has motivated estimands and bounds with different targets. Crump et al. (2009) select an overlap subpopulation to estimate its average treatment effect more precisely. Kennedy (2019) define incremental propensity-score interventions whose effects are identified without conventional positivity. For bounded outcomes, Susmann et al. (2026) combine identified overlap contributions with worst-case contributions from unsupported assignments to bound the full-population average treatment effect and develop smooth inference.

Lee & Weidner (2026) bound average effects, including effects on the treated, by pooling observations with shared covariates. These methods address targets defined by treatment labels and observed covariates. For received-treatment C-TACE, deterministic assignment can also hide the feature comparison that defines target membership. The root equations therefore bound both the effect numerator and the target-population denominator.

Causal inference with limited resources. Experimental design determines how to spend finite data-collection resources. Kato et al. (2024) adaptively select units and treatment probabilities to reduce asymptotic variance for an average treatment effect. Zhang et al. (2025) study active sampling under a sample-size constraint. In the replay setting, Athey et al. (2026) use inverse sampling until a specified number of feature matches occurs. We start with a fixed set of inference units and a fixed total number of replay calls, allocating calls to reduce a bound on the probability that the non-realized feature remains undiscovered. Section 6 compares the resulting confidence width with uniform designs at equal total cost.

B Proofs of Identification and Sharpness

Overlap identification. Conditional on $\mathcal{G}_i$, $Z_i \sim \text{Bernoulli}(Q_i)$, $Y_i = Y_i^{J_i}$ when $Z_i = 1$, and $Y_i = Y_i^{1-J_i}$ otherwise. Hence

$$\mathbb{E}(Z_i Y_i \mid \mathcal{G}_i) = Q_i Y_i^{J_i}, \qquad \mathbb{E}\left[(1 - Z_i)\frac{Q_i Y_i}{1 - Q_i} \,\middle|\, \mathcal{G}_i\right] = Q_i Y_i^{1-J_i}$$

on overlap. These identities and iterated expectation prove Equation (2) and identify $a(c) = \mathbb{E}[O_i(c)Q_i(Y_i^{J_i} - Y_i^{1-J_i})]$. All expectations are integrable. The conditional expected absolute inverse-weight contribution is at most $Q_i \leq 1$, even if individual weights are unbounded at $c = 0$.

Proof of Theorem 1. The target probability is $p = \mathbb{E}Z_i > 0$. Splitting $\mathbb{E}[Z_i(Y_i - Y_i^N - \tau)] = 0$ by $O_i(c)$ and applying the pointwise inequality in Section 3.2 gives $F_-(\tau; c) \leq 0 \leq F_+(\tau; c)$. Each function is continuous and decreases by at least $d(c)(t_2 - t_1)$ between $t_1 < t_2$. Since $|a(c)| \leq d(c)$ and $Y_i \in [0, 1]$, both functions are nonnegative at $t = -1$ and nonpositive at $t = 1$. Each therefore has a unique root in $[-1, 1]$. Their order relative to τ proves validity.

For $c_1 < c_2$, a state moved into support has an exact conditional contribution $Q_i(Y_i^{J_i} - Y_i^{1-J_i} - t)$. This lies between the pointwise lower and upper contributions in Section 3.2. Therefore $F_-(t; c_1) \geq F_-(t; c_2)$ and $F_+(t; c_1) \leq F_+(t; c_2)$, proving nested intervals. Dominated convergence, applied to

the bounded conditional expressions, gives convergence of the functions as $c \downarrow 0$. Strict decrease with $d(0) > 0$ implies convergence of their roots.

At $c = 0$ all interior contributions are fixed by the observed law. For a deterministic record with observed outcome y, attain the lower root by setting its counterfactual outcome to 1 and selecting target membership $z = 1$ exactly when $y - 1 - L(0) < 0$. Set $z = 0$ when the expression is positive. Either value is allowed at equality. Choose a non-realized feature below or above the realized feature to implement that membership. These changes preserve the observed deterministic assignment and outcome. The lower score is then identically its pointwise minimum, giving effect $L(0)$. Similarly, set the counterfactual outcome to zero and $z = 1$ when $y - U(0) > 0$ to attain $U(0)$. The stipulated feature support makes both constructions admissible. Their target denominators are positive because they contain $d(0)$. Mixing the two observationally equivalent completions preserves the observed law, and the effect varies continuously between the endpoints. This proves sharpness of the full interval. Feature-support restrictions that rule out either ordering give a narrower identified set. □

Observational equivalence at $Q_i = 0$. A state with $Q_i = 0$ has no target members in the full-data law, but its observed deterministic record may coincide with that of a state with $Q_i = 1$. Their non-realized features differ, but these features are unobserved. The bound therefore retains both possible values of Z_i at deterministic records.

Observation-only information. Without replay or a logged pair catalog, the observed law does not reveal Z_i or a positive lower bound on its probability. In a binary-outcome population with both outcomes present, set $Z_i = 1$ only for observed $Y_i = 0$ records, their counterfactual outcomes to 1, and the others to $Z_i = 0$. A deterministic selector and a suitable unseen ordering of the non-realized feature yield $\tau = -1$. Selecting observed $Y_i = 1$ records with counterfactual outcomes zero yields $\tau = 1$. Both completions respect the same observation law, and mixtures fill the interval. Thus $[-1, 1]$ is sharp under this unrestricted observation-only model. With a complete pair catalog and no propensity information, membership is known and the sharp set has width one, as stated in Section 6.

C Replay Identities and Information Limits

Proof of Proposition 2. Fix $\mathcal{G}_i, M_i = m$ and write $q = Q_i$, $N = m + 1$, $k = k_i$ in this proof only. If observation selects the larger feature and l replays select the smaller feature, the joint probability is

$$\binom{m}{l} q^{N-l}(1-q)^l.$$

If observation selects the smaller feature and $s = N - l$ replays select the larger feature, the joint probability is

$$\binom{m}{l-1} q^{N-l}(1-q)^l.$$

The identity $\binom{m}{l-1}(N-l)/l = \binom{m}{l}$ pairs these configurations for every retained l with $k \le l \le N-k$. On the first configuration $D_i = 1$ and $H_i = 0$. On the second $D_i = 0$ and $H_i = N - l$, so its correction weight is $(N - l)/l$. Thus $\mathbb{E}[R_i H_i/(N_i - H_i) \mid \mathcal{G}_i, M_i] = \mathbb{E}[D_i \mid \mathcal{G}_i, M_i] = g_i$. The outcomes on the paired configurations are $Y_i^{J_i}$ and $Y_i^{1-J_i}$, giving Equation (4). At $q = 0$ or $q = 1$, retention has probability zero and both sides vanish.

Because $D_i = R_i Z_i$ pathwise and observation outcomes are fixed conditional on assignment and $\mathcal{G}_i$,

$$\mathbb{E}T_i = \mathbb{E}[R_i Z_i (Y_i - Y_i^N)], \qquad \mathbb{E}Z_i = \mathbb{E}D_i + \mathbb{E}(B_i Z_i).$$

The decomposition in Equation (3) therefore holds with (a, d, B_i) defined from replay. The pointwise inequality in Section 3.2 proves the bound using entirely observable scores. □

Untrimmed replay limit and nesting. For $k = 1$ and common M, write $v = Y_i^{J_i}$, $u = Y_i^{1-J_i}$, $q = Q_i$. Then

$$g_M = q - q^{M+1}, \qquad \mathbb{P}(B_i = 1 \mid \mathcal{G}_i) = q^{M+1} + (1-q)^{M+1}.$$

The second term accounts for unknown membership among units that repeatedly receive the smaller feature. For $\ell_t(y) = \min(0, y - 1 - t)$, the conditional lower score is

$$g_M(v - u - t) + q^{M+1}\ell_t(v) + (1 - q)^{M+1}\ell_t(u).$$

Its increase from M to $M + 1$ equals

$$q^{M+1}(1 - q)\{v - u - t - \ell_t(v)\} - (1 - q)^{M+1}q\ell_t(u) \geq 0.$$

The first bracket is nonnegative because $u \leq 1$. The second term is also nonnegative. Replacing ℓ_t by $u_t(y) = \max(0, y - t)$ reverses both inequalities because $u \geq 0$, so the upper score decreases. The roots are consequently nested. As $M \to \infty$, $g_M \to q\mathbf{1}(0 < q < 1)$ and the unresolved law retains exactly the two deterministic cases. Dominated convergence proves the limit in Section 4.1. Unit-specific budgets that increase pointwise obey the same conclusion when the limiting retained mass is positive.

Positive threshold limit. For $k = \max(1, \lfloor cN \rfloor)$ and $c > 0$, the law of large numbers for replay counts gives $R_i \to \mathbf{1}(c < e_i < 1 - c)$ almost surely except at threshold atoms. Therefore $g_i \to Q_i O_i(c)$ and the unresolved outcome contribution converges to that in the known-propensity root functions of Theorem 1. Bounded conditional moments justify dominated convergence. At positive thresholds, discretized intervals need not be nested as M grows. The no-atom condition determines the strict-trimming limit at threshold values.

Information in replay counts. If both features are discovered, their empirical frequencies estimate Q_i and $1 - Q_i$. Assignment labels also permit estimation of e_i. Without labels, the propensity is learned up to a swap, which suffices for symmetric trimming and minority-probability allocation. Discovery of a distinct feature supplies the sign of $V_i - V_i^N$. With fixed M_i, individual frequency estimates remain noisy as $n \to \infty$. The replay identity gives the exact retained-target contribution under this design.

D Smooth Estimation and Coverage

Smoothing direction and error. For all x, $x_+ \leq s_\epsilon(x) \leq x_+ + \epsilon \log 2$. Hence the smoothed lower score is no larger than the exact lower score, and the smoothed upper score is no smaller than the exact upper score. Each root shifts outward by at most $\epsilon \log(2)\mathbb{E}(B_i)/d$, since the magnitude of each score slope is at least d. This proves Equation (6). Roots exist on $\mathbb{R}$ because the slope is bounded away from zero. Clipping is applied only after confidence expansion.

Proof of Theorem 3. Conditional on any independent pilot that fixes the replay rule, the inference records are iid. The scores T_i, D_i, B_i are bounded, with $|T_i| \leq \max(1, M_{\max})$. On any compact neighborhood of either population root, the smooth scores and their first derivatives have bounded envelopes. Their empirical means converge uniformly to their population counterparts. Strict monotonicity, with slope magnitude at least $d > 0$, implies consistent empirical roots and eventually positive empirical retained mass.

A mean-value expansion of the lower estimating equation gives

$$\sqrt{n}(\widehat{L}_\epsilon - L_\epsilon) = \frac{1}{A_-}\frac{1}{\sqrt{n}}\sum_{i=1}^{n}\phi_{i,-}(L_\epsilon) + o_p(1).$$

The upper expansion has the same positive sign because its population derivative is $-A_+$. The multivariate central limit theorem yields the joint normal limit. Uniform convergence of the derivatives and consistency of the roots also yield $\widehat{A}_\pm \to_p A_\pm$ and consistent empirical second moments of the influence functions. Positive marginal variances justify studentization. The marginal bounds also apply when the joint covariance is singular.

Each one-sided endpoint error has limiting probability $\alpha/2$. Their union has limiting probability at most α. On the complement, the reported interval contains the smoothed population interval intersected with the logical range, and therefore contains τ. The zero-mass and empty-interval fallbacks cannot reduce coverage. For a fixed K, apply the same argument jointly to the $2K$ one-sided errors at level $\alpha/(2K)$. This proves simultaneous coverage and permits arbitrary selection among those intervals. No independence between endpoints or candidate rules is required. □

Independent design selection. Condition on an independent pilot sample that fixes the budget map, thresholds, and smoothing. The theorem applies to iid inference records under that fixed map, assuming its nondegeneracy conditions hold. Integrating conditional limiting coverage gives the unconditional guarantee. Growing design classes or tuning outside a fixed simultaneous grid require uniform analysis beyond this theorem.

Stratified extension for a literal budget. Let $g = 1, \dots, G$ index a fixed number of input strata with known population proportions $\omega_g > 0$. Independently sample n_g records per stratum, with $n_g/n \to \pi_g > 0$. Pilot data fix M_g before these samples are drawn, so $\sum_g n_g M_g = B$ is literal. Replace every population expectation estimate by $\sum_g \omega_g \overline{(\cdot)}_g$. If $\widehat{\phi}_{i,\pm}$ are the fitted root scores and $\widehat{A}_\pm$ the corresponding weighted derivatives, the variance of the endpoint estimate is consistently estimated by

$$\widehat{\mathrm{se}}_\pm^2 = \frac{1}{\widehat{A}_\pm^2} \sum_{g=1}^{G} \frac{\omega_g^2}{n_g} \widehat{\mathrm{Var}}_g(\widehat{\phi}_{i,\pm}). \tag{13}$$

Within-stratum iid central limit theorems and a finite sum give the same root expansion and Bonferroni coverage. This avoids treating an allocation chosen from the inference observations as fixed. All observed outcomes remain excluded from design selection.

E Derivation of the Efficiency Rules

From unresolved mass to width. Write $p_i = \min(e_i, 1 - e_i)$ within this appendix and let $r_{\mathrm{int}} = \mathbb{E}[B_i \mathbf{1}(p_i > 0)]$. Compare a replay score with its infinite-replay score at $c = 0$. Removed target numerator and mass have absolute magnitudes at most r_{int} each. The added tail term has absolute magnitude at most $2r_{\mathrm{int}}$ for $t \in [-1, 1]$. Thus each score changes by at most $4r_{\mathrm{int}}$. If its retained mass is at least δ, either endpoint moves by at most $4r_{\mathrm{int}}/\delta$, giving

$$U - L \le W_0 + 8r_{\mathrm{int}}/\delta.$$

Equation (4) supplies the retained target contribution despite dependence of retention on the observation draw. Smoothing adds at most $2\epsilon \log 2/\delta$ to width.

Discovery error under a propensity tail condition. The count of one feature among all $N = M+1$ assignments has the binomial law. With $k = N\nu$, retention fails if either feature count is below k. For $p_i > 2\nu$ and $\nu \le 1/4$, multiplicative binomial tail inequalities give

$$\mathbb{P}(B_i = 1 \mid \mathcal{G}_i) \le 2e^{-Np_i/8}.$$

On $0 < p_i \le 2\nu$, use the bound of one. Under Equation (9), integration of $e^{-Np/8}$ against the distribution of positive p yields a bound of order $N^{-\kappa}$, since

$$\int_{(0,1/2]} e^{-Np/8}\, dF(p) \le e^{-N/16} + C\Gamma(\kappa+1)(8/N)^\kappa.$$

Consequently $r_{\mathrm{int}} \le C_1'\nu^\kappa + C_2' N^{-\kappa}$. The deterministic mass $p_i = 0$ is excluded from this calculation and remains in W_0.

Variance bound. Let $C_i = R_i H_i/(N - H_i)$ in this proof. The replay identity gives $\mathbb{E}C_i = \mathbb{E}D_i = d$, and $D_i C_i = 0$. Since $C_i \le (1-\nu)/\nu$,

$$\mathbb{E}T_i^2 \le \mathbb{E}D_i + \mathbb{E}C_i^2 \le d/\nu.$$

Population smoothed roots lie within $[-1 - \epsilon \log 2/\delta, 1 + \epsilon \log 2/\delta]$. The remaining score terms therefore have bounded second moments depending only on δ, ϵ. Dividing by $A_\pm \ge \delta$ shows $\sigma_- + \sigma_+ \le C'/\sqrt{\nu}$. Combining these bounds proves Equation (10). Differentiating $C_1\nu^\kappa + C_3 n^{-1/2}\nu^{-1/2}$ gives Equation (11). The bound combines population width with normal-theory uncertainty. Observed confidence width remains sample-dependent.

Budget allocation. For known minority probabilities, the surrogate objective $\sum_i 2e^{-\lambda_i(M_i+1)}$ is convex in continuous budgets. At an interior optimum its Lagrange equation is $2\lambda_i e^{-\lambda_i(M_i+1)} = \eta$, which gives Equation (12). Clipping imposes the box constraints. The scalar constraint selects η, except when every informative coordinate is saturated. For integer budgets, successive marginal gains are $2(1-e^{-\lambda_i})e^{-\lambda_i(M_i+1)}$ and decrease with M_i. Starting at the lower bounds and selecting the largest remaining gains is optimal by an exchange argument. Replacing any selected smaller gain with an available larger gain decreases the objective. Decreasing within-unit gains ensure that all prerequisite calls are selected. Zero-rate coordinates have zero gain and receive surplus calls only after informative coordinates saturate. Pilot estimates define a fitted version of this surrogate.

Pilot allocation from observed replays. Let $g = 1, \ldots, G$ index pre-assignment strata, n_g^{pil} the pilot records in stratum g, and b_{pil} the replay outputs per pilot record. If C_{gr}^{pil} counts outputs matching the first output for pilot record r, estimate minority probability by

$$\hat{p}_g = \frac{1}{n_g^{\mathrm{pil}}} \sum_{r=1}^{n_g^{\mathrm{pil}}} \frac{\min\{C_{gr}^{\mathrm{pil}},\, b_{\mathrm{pil}} - C_{gr}^{\mathrm{pil}}\}}{b_{\mathrm{pil}}}, \qquad \hat{\lambda}_g = -\log(1-\hat{p}_g). \tag{14}$$

With n_g^{inf} planned inference records in stratum g, the fitted integer plan minimizes

$$\sum_{g=1}^{G} 2n_g^{\mathrm{inf}}(1-\hat{p}_g)^{m_g+1} \quad \text{subject to} \quad \sum_{g=1}^{G} n_g^{\mathrm{inf}} m_g \le B, \quad m_g \in \{m_{\min}, \ldots, m_{\max}\}. \tag{15}$$

Here B is the inference budget after pilot calls are charged. Let w_g be the population stratum weight. We compare the fitted plan $\hat{m}$ with uniform replay m^{unif} using

$$\widehat{R}(m) = \sum_{g=1}^{G} 2w_g(1-\hat{p}_g)^{m_g+1} + 2z_{1-\alpha/2}\left\{\sum_{g=1}^{G} \frac{w_g^2 \hat{v}_g(m_g)}{n_g^{\mathrm{inf}}}\right\}^{1/2}. \tag{16}$$

The second-moment proxy $\hat{v}_g$ is defined below. We choose the fitted plan only if its proxy is smaller; otherwise each stratum receives the uniform budget. This includes uniform replay in the pilot search and fixes the plan before inference. The pilot proxy does not guarantee the ordering of realized interval widths.

For the pilot comparison, let $p = \min(e, 1-e)$ and $H \sim \mathrm{Bin}(m, 1-p)$. The score coefficient with the higher feature selected more often has second moment

$$v(p,m) = (1-p)\{1-(1-p)^m\} + p\mathbb{E}\left[\left\{\frac{H}{m-H+1}\right\}^2 \mathbf{1}(H \ge 1)\right].$$

We insert $\hat{p}_g$ into $v(p,m)$ to obtain $\hat{v}_g(m)$ in Equation (16). The first term of that equation measures unresolved comparisons; the second penalizes large replay-score moments. The proxy is used only to choose between the fitted surrogate allocation and uniform replay on the independent pilot.

F Additional Experimental Results and Reproducibility

Dataset construction and sampling. We match longer and shorter public videos within content category on popularity, upload date, and aspect ratio. A retained user–pair record contains a recorded watch response to both videos. Completion is the indicator that play duration reaches video duration. The table has 1,687,549 retained records and 1,205 matched pairs. A rank-eight matrix-factorization logger trained on 250,000 earlier interactions scores both videos. Their score difference is the prediction gap. For interior records, the probability $q_i = Q_i$ of selecting the longer video is the logistic transform of four times the standardized gap, clipped to $[.01, .99]$. We set $q_i = 0$ and $q_i = 1$ for the bottom and top $\rho/2$ fractions of the gap distribution. Observation and replay assignments are simulated from this logger, and the full C-TACE is recomputed for each ρ. Each replication draws n records with replacement from the fixed table. Thus the sampling law concerns records rather than distinct users.

Comparator construction. The observation-only set is $[-1, 1]$ under unrestricted feature ordering and unknown target size. The pair-catalog baseline observes each non-realized feature and has set $[\mathbb{E}(Y_i \mid Z_i = 1) - 1, \mathbb{E}(Y_i \mid Z_i = 1)]$. The discovered-comparison point estimator uses $\overline{T}/\overline{D}$ at $M = 20, c = 0$. The oracle overlap point estimator uses true propensities and feature comparisons on units with overlap. Both point estimators use normal approximations to their ratio estimators for 95% intervals. For the catalog inverse-sampling estimator, units with $Z_i = 0$ replay until ten draws match the observed feature. The number of other-feature draws divided by ten estimates inverse odds. The replay confidence intervals use $\epsilon = .01$ and endpoint influence standard errors. The tuned $M = 20$ comparison selects among three thresholds using the simultaneous critical value $z_{1-\alpha/6}$. The ablation fixes its threshold from an independent pilot and uses $z_{1-\alpha/2}$ for inference.

Two distinct point-estimator limits. Let $p_{\text{ov}} = \mathbb{E}[Q_i \mathbf{1}(0 < Q_i < 1)]$ and $p_1 = \mathbb{P}(Q_i = 1)$. If $p_1 > 0$, write $\tau_1 = \mathbb{E}[Y_i^{J_i} - Y_i^{1-J_i} \mid Q_i = 1]$. Then

$$\tau_{\text{ov}} - \tau = \frac{p_1}{p_{\text{ov}} + p_1}(\tau_{\text{ov}} - \tau_1).$$

This bias persists as sample size and replay budget grow. The catalog inverse-sampling extension has numerator $Z_i Y_i - (1 - Z_i) Y_i F_i / 10$, where F_i counts nonmatching draws before ten matches on a $Z_i = 0$ unit. For $Q_i < 1$, $\mathbb{E}(F_i \mid \mathcal{G}_i, Z_i = 0) = 10 Q_i/(1 - Q_i)$. Its probability limit is

$$\frac{\mathbb{E}[Q_i Y_i^{J_i} - Q_i Y_i^{1-J_i} \mathbf{1}(Q_i < 1)]}{\mathbb{E}Q_i} = \tau + \frac{\mathbb{E}[\mathbf{1}(Q_i = 1) Y_i^{1-J_i}]}{\mathbb{E}Q_i}.$$

At $Q_i = 1$, no observation record has $Z_i = 0$, so inverse sampling does not recover the counterfactual-outcome contribution. The catalog benchmarks use the recorded treatment pair. The replay bound uses observed outcomes and replay features.

Table 2: Coverage (C, percent) and mean confidence width (W) over 500 replications. The oracle overlap point estimator uses true propensities and feature comparisons. Tuned replay intervals account for selection over three thresholds.

		Oracle point		Replay $M = 20$		Tuned $M = 20$		Replay $M = 50$	
ρ	n	C	W	C	W	C	W	C	W
0.0	2,000	92.6	0.223	99.8	0.420	99.8	0.462	98.4	0.331
0.0	10,000	93.6	0.102	100.0	0.315	100.0	0.334	100.0	0.217
0.0	50,000	94.4	0.047	100.0	0.268	100.0	0.276	100.0	0.166
0.1	2,000	93.2	0.194	100.0	0.469	100.0	0.509	100.0	0.404
0.1	10,000	95.2	0.087	100.0	0.370	100.0	0.388	100.0	0.306
0.1	50,000	95.0	0.039	100.0	0.325	100.0	0.333	100.0	0.261
0.2	2,000	94.2	0.194	100.0	0.546	100.0	0.583	100.0	0.509
0.2	10,000	93.6	0.087	100.0	0.455	100.0	0.472	100.0	0.420
0.2	50,000	90.4	0.039	100.0	0.413	100.0	0.421	100.0	0.380
0.4	2,000	94.0	0.204	100.0	0.776	100.0	0.808	100.0	0.759
0.4	10,000	91.6	0.091	100.0	0.694	100.0	0.708	100.0	0.682
0.4	50,000	63.4	0.041	100.0	0.658	100.0	0.664	100.0	0.649

Table 3: Mean confidence width in the four-way replay-design ablation at $n = 10{,}000$ over 500 replications. Every design uses 100,000 inference calls. The tuned designs share an additional 20,000-call pilot. Every design has 100% empirical full-target coverage.

ρ	Basic	Threshold	Allocation	Both
0.0	0.449	0.449	0.422	0.422
0.1	0.482	0.482	0.448	0.448
0.2	0.542	0.542	0.506	0.506
0.4	0.742	0.742	0.709	0.709

Ground-truth evaluation uses the full outcome table and true propensity. Observation and replay streams are coupled across budgets within a replication. The replay estimator receives relative replay counts. The evaluation includes propensity-zero and propensity-one strata.

Table 4: Sparse-overlap stress ablation over 5,000 fresh replications with $n = 2{,}000$ and 100 uniform replay calls per inference record. The threshold $c = .05$ was fixed before validation. Allocation uses an independent pilot of 20 records in each of 20 prediction-gap strata, with 100 calls per pilot record. The pilot width screen selects uniform replay in every replication.

Design	Mean width	Coverage (%)	Mean total calls
Basic	0.2313	95.4	200,000
Positive threshold	0.2198	100.0	200,000
Pilot allocation	0.2313	95.4	240,000
Both	0.2198	100.0	240,000

Table 5: Population benchmarks. The known-propensity, infinite-replay interval accounts for unknown boundary membership. Finite-replay width uses $M = 20, c = 0$ without smoothing or sampling correction. Position is the mean relative location of the true effect within the $n = 50{,}000$ replay confidence interval, calculated as $(\tau - C_{n,L})/(C_{n,U} - C_{n,L})$.

ρ	C-TACE	Sharp lower	Sharp upper	Replay width	Position
0.0	-0.4144	-0.4144	-0.4144	0.2291	0.313
0.1	-0.4141	-0.4738	-0.2898	0.2884	0.337
0.2	-0.4136	-0.5337	-0.1886	0.3798	0.365
0.4	-0.4110	-0.6457	-0.0230	0.6304	0.386

For the four-way ablation, an independent pilot samples 100 records from each of 20 strata defined by the absolute prediction gap and makes 10 replay calls per record. The pilot uses the smaller empirical feature frequency to estimate minority probability in each stratum. The same pilot chooses between $c = 0$ and $c = .2$ by projected inference width. Inference draws 500 fresh records per stratum and weights them by their population proportions using Equation (13). All four designs share inference records and nested replay streams within a replication. Each makes 100,000 inference calls, and the three tuned designs use 20,000 additional pilot calls. The replay service supplies the pre-assignment stratum label for allocation. Interval endpoints use observed outcomes, replay counts, and stratum weights.

The sparse-overlap stress logger uses the same KuaiRec outcome catalog. It assigns the higher-feature treatment with probability .99 in the top 5% of prediction gaps and probability .5 elsewhere. Each replication draws 2,000 inference records independently from the catalog. Basic and threshold designs use 100 replays per record. The allocation pilot samples 20 records from each of 20 prediction-gap strata and uses 100 replays per pilot record. Its discovery-surrogate plan sends about 200 calls to the sparse stratum and fewer than 100 to most other strata. The pilot screen in Equation (16) chooses uniform replay in all 5,000 replications, so allocation and basic use 200,000 inference calls each. The allocation design also uses 40,000 pilot calls. The fixed threshold $c = .05$ requires five contrasting replays for retention. At this threshold the finite-replay population bound has width 0.102, compared with 0.037 at $c = 0$. The confidence interval is narrower because the threshold reduces finite-sample variability. A separate trial of the pilot threshold criterion with 20 pilot records per stratum selected $c = 0$ in 464 of 500 replications. The positive-threshold result in Table 4 therefore concerns the prespecified threshold, not successful pilot selection.

Table 6 gives a separate equal-total-cost comparison using a larger pilot with 40 calls on each of 100 records per stratum. Its uniform 10-call and allocated designs each use 80,000 pilot calls and 100,000 inference calls. Its uniform 18-call design uses 180,000 inference calls. All three designs share inference records and nested replay streams within a replication.

The main experiment and the ρ ablation report empirical coverage over 500 replications, with Monte Carlo standard error $\sqrt{\widehat{p}(1-\widehat{p})/500}$. The sparse-overlap stress test uses 5,000 replications. The variance formulas correspond to independent draws from the fixed empirical table.

Table 6: Replay allocation for 10,000 sampled user–pair records under 180,000 total calls. U10 and allocated each use 80,000 pilot plus 100,000 inference calls. U18 uses all calls for inference and no pilot. Reductions compare allocated width with each uniform width. Negative values favor uniform. All three use stratified inference.

	Mean width			Reduction (%)		Coverage (%)		
ρ	U10	Alloc.	U18	vs U10	vs U18	U10	Alloc.	U18
0.0	0.449	0.421	0.331	+6.3	-27.1	100.0	100.0	100.0
0.1	0.482	0.446	0.382	+7.6	-16.7	100.0	100.0	100.0
0.2	0.542	0.504	0.463	+6.9	-8.9	100.0	100.0	100.0
0.4	0.742	0.710	0.696	+4.3	-2.0	100.0	100.0	100.0

Table 7: Catalog inverse-sampling estimator at $n = 50{,}000$. Bias is its exact limiting difference from the full C-TACE. Replay cost is the mean number of calls per observation unit, including zero calls for known target members. The estimator uses the recorded treatment pair.

ρ	Limiting bias	Coverage (%)	Mean width	Mean calls
0.0	0.0000	94.0	0.049	10.00
0.1	0.0453	1.0	0.037	9.50
0.2	0.0881	0.0	0.034	9.00
0.4	0.1761	0.0	0.028	8.00