EXACT POLICY CERTIFICATE LOSS UNDER NORMALIZATION COARSENING

Anonymous authors
Paper under double-blind review

ABSTRACT

Robust policy certificates lower-bound policy improvement under a sensitivity model for unobserved confounding. Group-level inverse-weight normalization can weaken these certificates, even when the policy and full-history propensity reference remain fixed. In a one-step binary-action model, we characterize this loss for arbitrary measurable normalization interfaces: it equals the excess weighted pinball risk incurred by requiring histories in the same group to share a threshold. For sensitivity radii greater than one, grouping is lossless if and only if a measurable group-level threshold belongs to the conditional treated-reward quantile set at almost every history, including with discrete rewards and nonunique quantiles. To certify fitted thresholds with estimated nuisance quantities, we develop rectangular augmented certification (RAC), combining nuisance correction, confidence-aware fitting and independent betting. For binary rewards and an always-treat target, simultaneous full-history nuisance envelopes yield a finite-sample-valid lower confidence bound on policy improvement after fitting. Experiments show that augmentation strengthens the matched protected weak dual, while confidence-aware fitting improves average certificate strength over empirical-mean fitting across tested positive-margin conditions; fixed thresholds remain competitive. Grouping trades structural loss against statistical uncertainty. Under correctly specified low-dimensional nuisance models, structured calibration reduces protection costs at larger history supports.

1 INTRODUCTION

Logged actions and rewards allow us to evaluate a candidate policy before replacing an existing decision rule. When unobserved factors affect both action choices and rewards, the policy's value is generally not identified. Sensitivity analysis considers causal laws compatible with the observations and a specified bound on hidden confounding (Tan, 2006). A robust policy certificate lower-bounds improvement over the observed assignment system across these laws; a positive bound supports improvement throughout the declared sensitivity class (Kallus & Zhou, 2018).

Large history spaces motivate summaries through learned features or fixed groups (Mandel et al., 2014; Chang et al., 2022). In an inverse-weight certificate, normalization imposes a separate balancing constraint at each history. Enforcing it only within groups permits weight redistribution across histories, enlarging the adversary's feasible set and weakening the population bound. In the dual, grouping requires histories to share a threshold. This reduces fitting flexibility and may help finite-sample evaluation, but can incur structural loss. We ask when grouping preserves the certificate, how much it loses otherwise, and how the loss interacts with statistical uncertainty.

We isolate normalization by keeping the policy, objective, full-history propensity reference, and sensitivity boxes fixed, while conditioning normalization on either the full history H or a summary $Z = \phi(H)$, where ϕ is the summary map. We call this relaxation *reference transport*. Replacing the reference itself by the propensity conditional on the summary changes the sensitivity class, so equal numerical radii need not describe comparable restrictions (Melnychuk et al., 2024). Weighted-quantile duality characterizes the resulting loss. For finite groups, the identity follows from Tan (2024, Propositions 1–2). Under overlap and integrable rewards, we extend it to arbitrary standard-Borel interfaces and establish attainment. The exact loss is the excess weighted pinball risk from restricting thresholds to the summary. At sensitivity radii greater than one, grouping is lossless

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.

precisely when a measurable summary-level threshold belongs to the treated-reward quantile set at almost every history. Complete quantile sets matter for discrete rewards: different selected quantiles may still admit a shared optimal threshold.

Fitted certificate strength also depends on score variance and nuisance correction, which can differ across population-equivalent thresholds. For binary rewards and an always-treat target, rectangular augmented certification (RAC) corrects augmented sensitivity scores (Dorn et al., 2025; Tan, 2024) using simultaneous full-history nuisance envelopes and fits thresholds with a variance-aware objective. Independent betting evaluation (Waudby-Smith & Ramdas, 2024) gives finite-sample coverage under overlap and the stated envelope conditions. Thresholds can be shared while nuisance estimates retain full-history resolution.

Experiments distinguish structural loss from fitting, nuisance, and evaluation uncertainty. Augmentation strengthens the matched protected weak dual. Across tested positive-margin conditions, confidence-aware fitting improves average certificate strength over empirical-mean fitting, while fixed thresholds remain competitive. Grouping gives modest gains when structural loss is small and weaker certificates when it dominates. Correctly specified structured nuisance models reduce protection costs at large supports, while cellwise calibration is stronger at small supports.

Our contributions are:

- **Exact population loss.** An attained weighted-pinball characterization and measurable shared-quantile criterion for arbitrary standard-Borel normalization interfaces (Theorem 1).
- **Finite-sample certification.** Rectangular nuisance correction and independent betting evaluation protect fitted RAC certificates for binary rewards and an always-treat target (Theorem 2).
- **Empirical trade-offs.** We quantify how structural loss, augmentation, threshold fitting, and nuisance calibration jointly affect certificate strength.

2 Related Work

Sensitivity analysis bounds causal effects when unobserved confounding cannot be ruled out. Under the marginal sensitivity model (Tan, 2006), quantile balancing yields sharp bounds (Dorn & Guo, 2023), and weighted quantile regression characterizes both sharp and relaxed bounds (Tan, 2024). Kernel conditional moment constraints provide a broader framework for approximating normalization restrictions (Ishikawa & He, 2023; Ishikawa et al., 2026). Building on this duality, we characterize the exact loss from a measurable normalization interface and give a necessary and sufficient condition for preserving the sharp endpoint.

Representation learning can improve causal estimation (Johansson et al., 2016; Shalit et al., 2017), but discarding confounders can introduce bias (Melnychuk et al., 2024). Coarsening can also improve inverse-weighted inference under unconfoundedness (Kalavasis et al., 2024). Our comparison keeps the full-history propensity reference fixed, isolating the effect of relaxing normalization. For finite-data certification, RAC builds on augmented sensitivity scores (Dorn et al., 2025; Tan, 2024), variance-aware robust fitting (Duchi & Namkoong, 2019), and betting inference (Waudby-Smith & Ramdas, 2024). Its contribution is finite-sample protection of the fitted certificate through simultaneous nuisance envelopes and independent evaluation. Appendix A gives further context.

3 Model and Normalization Interfaces

We evaluate always treating relative to the observed assignment system. The comparison fixes the full-history propensity reference and varies the information used by normalization under bounded unobserved confounding. Let H be the finest declared observed pre-action history in a standard-Borel space $\mathcal{H}$, U an unobserved pre-action state, $A \in \{0, 1\}$ the logged action, and $R = R(A)$ the observed reward, where $R(a)$ denotes the potential reward under action a. The observed law of (H, A, R) is fixed, with propensity $e(h) = e_{\text{full}}(h) = P(A = 1 \mid H = h)$ and uniform overlap $\eta \le e(H) \le 1 - \eta$ for $0 < \eta < 1/2$. Here P and $\mathbb{E}$ denote probability and expectation under the observed law unless subscripted; $\mathbb{E}_H$ averages over its history marginal. We assume $\mathbb{E}|R| < \infty$, and "a.s." means almost surely.

Candidate causal laws reproduce the observed law and satisfy consistency, latent exchangeability $A \perp \{R(0), R(1)\} \mid (H, U)$, and positivity. The sensitivity parameter $1 \leq \lambda < \infty$ limits how much the latent state can change treatment odds relative to $e(H)$. For the complete propensity $q(h, u) = P(A = 1 \mid H = h, U = u)$ and $\operatorname{odds}(x) = x/(1-x)$, the marginal sensitivity class is

$$\mathfrak{P}_{\text{full}}(\lambda) = \left\{ P : \lambda^{-1} \leq \frac{\operatorname{odds} q(H, U)}{\operatorname{odds} e(H)} \leq \lambda \quad P\text{-a.s.} \right\}, \qquad 1 \leq \lambda < \infty, \tag{1}$$

including the preceding compatibility and causal conditions. Write $V_P = \mathbb{E}_P R(1)$, $B = \mathbb{E}R$ for the factual value of the observed assignment system, and $\Delta_P = V_P - B$ for the improvement of always treating under P. Inverse weighting expresses the policy value using observed treated rewards. The odds restriction gives lower and upper full-reference inverse-weight endpoints $\ell(h)$ and $u(h)$, respectively:

$$\ell(h) = 1 + \frac{1 - e(h)}{\lambda e(h)}, \qquad u(h) = 1 + \frac{\lambda\{1 - e(h)\}}{e(h)}. \tag{2}$$

The normalization interface $Z = \phi(H)$ takes values in a standard-Borel space $\mathcal{Z}$, with measurable $\phi : \mathcal{H} \to \mathcal{Z}$. For normalization choice $X \in \{H, Z\}$, optimize over measurable inverse weights $w(H, R)$ to define

$$V_X(\lambda) = \inf_w \mathbb{E}\{ARw(H, R)\}, \quad \ell(H) \leq w(H, R) \leq u(H), \quad \mathbb{E}(Aw \mid X) = 1. \tag{3}$$

The constraint $\mathbb{E}(Aw \mid X) = 1$ restores unit weighted mass within each normalization stratum. For $X = H$, it must hold separately at almost every history; for $X = Z$, histories with the same summary can compensate for one another. The weights still depend on (H, R) and retain the same history-specific boxes. Every compatible causal law induces a feasible full-history weight, and full-history normalization implies grouped normalization (Appendix B.1.1). Thus $V_Z \leq V_H \leq V_P$ for every compatible P. Theorem 1 also establishes that the full-history bound is attained by a compatible causal law. Define the population improvement endpoints $M_X = V_X - B$, so $M_H = \inf_{P \in \mathfrak{P}_{\text{full}}(\lambda)} \Delta_P$. A positive M_X therefore certifies improvement throughout the full-reference sensitivity class.

4 Exact Normalization Loss

4.1 Exact loss and consequences

Full-history normalization permits history-specific dual thresholds; normalization through Z forces histories with the same summary to share one. The endpoint loss is exactly the increase in optimal weighted quantile risk. For $1 < \lambda < \infty$, define quantile level $\tau = (1 + \lambda)^{-1}$ and history weight $c(h) = \{1 - e_{\text{full}}(h)\}(\lambda - \lambda^{-1})$. Let F_h be the conditional treated reward law, with cumulative distribution function (CDF) $F_h(t)$ and left limit $F_h(t-) = P(R < t \mid H = h, A = 1)$. We view a threshold a measurable through the interface as a function of H by composition, so a Z-measurable threshold has $a(H) = a\{\phi(H)\}$. Let ρ_τ be the pinball loss and $\mathbf{1}$ an indicator; define the weighted pinball risk (Koenker & Bassett, 1978; Tan, 2024) by

$$\mathcal{R}_\lambda(a) = \mathbb{E}_H \left[c(H) \int \rho_\tau\{r - a(H)\} \, F_H(dr) \right], \qquad \rho_\tau(v) = v\{\tau - \mathbf{1}(v < 0)\}.$$

Pinball risk is minimized by a τ-quantile; the weight $c(h)$ determines how strongly a history contributes to the sensitivity endpoint. In the result below, L_ϕ denotes normalization loss, while a and b range over summary- and history-measurable thresholds, respectively.

Theorem 1 (Exact loss for general normalization interfaces)**.** *For $1 < \lambda < \infty$, under the assumptions of Section 3, the primal infima are attained and*

$$L_\phi(\lambda) := V_H(\lambda) - V_Z(\lambda) = \min_{a:\, Z\text{-measurable}} \mathcal{R}_\lambda(a) - \min_{b:\, H\text{-measurable}} \mathcal{R}_\lambda(b). \tag{4}$$

Both risk minima have integrable measurable minimizers. For the closed quantile set $Q_\tau(h) = \{t : F_h(t-) \leq \tau \leq F_h(t)\}$,

$$L_\phi(\lambda) = 0 \quad \Longleftrightarrow \quad \exists \text{ measurable } a : \mathcal{Z} \to \mathbb{R} : \; a\{\phi(H)\} \in Q_\tau(H) \text{ a.s.}$$

At $\lambda = 1$, the boxes are singletons and $V_H = V_Z$ for every interface, without a shared-quantile requirement. Moreover, V_H is the attained sharp lower causal value under the unrestricted one-step marginal sensitivity model.

For $\lambda > 1$, an interface preserves the endpoint exactly when it permits a shared optimal treated-outcome threshold almost surely. Treated reward distributions can differ without incurring loss if their quantile sets admit such a choice.

Dual interpretation. Let $\mu = \mathbb{E}_H \int r F_H(dr)$ be the standardized treated mean. For an integrable threshold measurable through $X \in \{H, Z\}$, the quantile-balancing dual value $D_X(a)$ (Dorn & Guo, 2023; Tan, 2024) is

$$D_X(a) = \mathbb{E}\{a(X)\} + \mathbb{E}\left[A \min_{v \in [\ell(H), u(H)]} v\{R - a(X)\}\right] = \mu - \mathcal{R}_\lambda(a). \tag{5}$$

Maximizing this dual minimizes weighted pinball risk, and its optimum is V_X. Restricting thresholds to the interface therefore gives the difference of optimal risks in Theorem 1. For finite interfaces, subtracting Tan (2024, Propositions 1–2) gives the identity; nonnegative conditional excess risk gives the shared-quantile condition. Appendix B.1.2 proves the arbitrary-interface result, integrability, attainment and causal sharpness.

The complete quantile sets are essential for discrete rewards. For example, at $\lambda = 2$, two histories h_1, h_2 with treated success probabilities $m_{h_1} = 2/3$ and $m_{h_2} = 0.8$ give $Q_\tau(h_1) = [0, 1]$ and $Q_\tau(h_2) = \{1\}$. Their conventional lower quantiles differ, yet the shared threshold $a = 1$ makes merging these histories lossless. Appendix B.1.2 gives the finite-cell expansion and extension to fixed deterministic binary policies (Corollary 1). Further coarsening weakens the population endpoint when the full-history reference boxes are retained (Appendix B.1.3).

4.2 Full-reference comparisons

Propensity envelopes conservatively contain the original boxes using separate upward and downward odds factors (directional), their cellwise maximum (symmetric), or the maximum across cells (global). Retaining history-specific boxes at the same Z normalization therefore gives a bound at least as strong as these envelopes. Adding propensity metadata $S = e(H)$ changes normalization from Z to (Z, S) and can tighten it further without changing the boxes. Appendix B.1.4 gives the complete ordering, and Appendix B.1.3 explains the metadata refinement. Recentering boxes at $P(A = 1 \mid Z)$ changes the sensitivity class and can falsely certify improvement under the original class (Appendix B.1.5).

5 Rectangular Augmented Certification

RAC turns a fitted threshold into an improvement lower confidence bound (LCB) $\widehat{L}$ for M_H. It corrects an augmented score using simultaneous nuisance envelopes, fits a variance-aware threshold, and evaluates the frozen score on independent data. Algorithm 1 summarizes the procedure.

5.1 Nuisance envelopes and the corrected score

Consider binary rewards, the always-treat target, fixed $1 \leq \lambda < \infty$, and histories in a declared standard-Borel domain $\mathcal{H}$ of full probability. Assume overlap $e_h = P(A = 1 \mid H = h) \in [\eta, 1-\eta]$, $0 < \eta < 1/2$. The construction sigma-algebra $\mathcal{F}$ records calibration $\mathcal{D}_e$, fitting $\mathcal{D}_f$, and interface training. An evaluation sample $\mathcal{D}_s$, independent of $\mathcal{F}$, contains a fixed number n_s of independent and identically distributed (i.i.d.) records $O_i = (H_i, A_i, R_i)$. This independence allows the fitted score to be treated as fixed when evaluating its conditional mean.

The nuisance functions are the propensity e_h and the treated success probability $\mu_h = P(R = 1 \mid H = h, A = 1)$. Their uncertainty must be controlled simultaneously over the declared history domain because the threshold is chosen using construction data. Let α_e, α_μ be the propensity and outcome envelope error budgets, and $\mathcal{B}(\mathcal{H})$ the Borel sigma-algebra on histories. Input intervals $[L_e(h), U_e(h)]$ and $[L_\mu(h), U_\mu(h)]$ have lower and upper endpoints jointly measurable with respect to $\mathcal{F} \otimes \mathcal{B}(\mathcal{H})$. On a simultaneous coverage event $\mathcal{E} \in \mathcal{F}$ with $P(\mathcal{E}) \geq 1 - \alpha_e - \alpha_\mu$, they contain both nuisance functions throughout $\mathcal{H}$. Intersect with $[\eta, 1 - \eta]$ and $[0, 1]$, respectively, obtaining $L_{eh}, U_{eh}, L_{\mu h}, U_{\mu h}$. A measurable rule abstains with $-\infty$ unless all intersections are nonempty. Project jointly measurable nuisance fits into these intervals, giving propensity and outcome estimates p_h and $\widehat{\mu}_h$, respectively. Admissible thresholds are jointly measurable $[0, 1]$-valued functions

through $X = Z$ or H, constructed using $\mathcal{F}$. Population quantities are $\mathcal{F}$-measurable for the frozen interface. The interface Z restricts normalization and thresholds, while nuisance functions and their envelopes retain full-history resolution.

For a chosen threshold, the transformed reward Q_a represents the lower sensitivity dual, and $\widehat{m}_h(a)$ estimates its treated conditional mean. Augmentation combines this estimate with an inverse-weighted residual. Write $X(h) = \phi(h)$ for $X = Z$ and $X(h) = h$ for $X = H$, and set $a_h = a\{X(h)\} \in [0,1]$; subscripts H evaluate history-indexed functions at the random history. With $(x)_+ = \max\{x, 0\}$, define the reward slope κ and augmented improvement score φ_a by

$$\begin{aligned} Q_a(r) &= a + (r-a)/\lambda - (\lambda - \lambda^{-1})(a-r)_+, \\ \kappa(a) &= Q_a(1) - Q_a(0) = \lambda^{-1} + (\lambda - \lambda^{-1})a, \qquad \widehat{m}_h(a) = Q_{a_h}(0) + \kappa(a_h)\widehat{\mu}_h, \\ \varphi_a(H, A, R) &= (1-A)\{\widehat{m}_H(a) - R\} + A\frac{1-p_H}{p_H}\{Q_{a_H}(R) - \widehat{m}_H(a)\}. \end{aligned} \tag{6}$$

This is the augmented lower sensitivity score of Dorn et al. (2025), also expressed in Tan (2024, Equation (20)), minus factual reward. For frozen nuisances, its signed conditional bias relative to the known-reference dual is $r_h(a) = (e_h - p_h)\kappa(a_h)(\mu_h - \widehat{\mu}_h)/p_h$. The bias is a product of propensity and outcome-estimation errors. On $\mathcal{E}$, both errors lie in the declared rectangle, so subtracting their largest possible product protects the score for every admissible threshold. Define the rectangle bias bound β_h and corrected score g_a by

$$\beta_h = \frac{1}{p_h} \max_{\substack{s \in \{L_{eh}, U_{eh}\} \\ t \in \{L_{\mu h}, U_{\mu h}\}}} (s - p_h)(t - \widehat{\mu}_h), \qquad g_a = \varphi_a - \kappa(a_H)\beta_H. \tag{7}$$

The bilinear maximum occurs at a rectangle corner. Let $\Delta_{e,h} = U_{eh} - L_{eh}$ and $\Delta_{\mu,h} = U_{\mu h} - L_{\mu h}$ be the band widths. Projection, $p_h \geq \eta$, and $\kappa(a_h) \leq \lambda$ give $0 \leq \kappa(a_h)\beta_h \leq (\lambda/\eta)\Delta_{e,h}\Delta_{\mu,h}$. The applied correction therefore scales with the product of the nuisance-band widths. On $\mathcal{E}$, every nonabstaining admissible fit satisfies

$$\mathbb{E}(g_a \mid \mathcal{F}) \leq D_X(a) - B \leq M_X \leq M_H. \tag{8}$$

Thus a lower confidence bound for the corrected score's conditional mean also bounds the full-history improvement endpoint on nuisance coverage.

5.2 Threshold fitting and independent certification

Population-equivalent thresholds can differ in variance and nuisance correction. RAC uses a variance-aware fitting surrogate to select thresholds for independent certification. With $n_f = |\mathcal{D}_f|$ and $\mathcal{A}_X$ the convex class of $[0,1]$ interface thresholds, RAC fits $\widehat{a}$ by applying Duchi & Namkoong (2019)'s chi-square robust fitting principle to g_a. Here w_i are fitting-record probability weights and $\mathcal{W}_\rho$ their chi-square neighborhood of uniform weights, with radius ρ:

$$\begin{aligned} &\max_{a \in \mathcal{A}_X} \min_{w \in \mathcal{W}_\rho} \sum_{i=1}^{n_f} w_i g_a(O_i), \qquad \rho = \frac{2\log 40}{n_s}, \\ &\mathcal{W}_\rho = \left\{ w_i \geq 0 : \sum_i w_i = 1, \quad \frac{1}{n_f}\sum_i (n_f w_i - 1)^2 \leq \rho \right\}. \end{aligned} \tag{9}$$

The fitting sample estimates score quality; the planned evaluation size n_s sets the variance penalty, with $40 = 1/\alpha_s$ for nominal betting error budget $\alpha_s = 0.025$. When the unconstrained weights are nonnegative, the inner value is the empirical mean minus a standard-deviation penalty; otherwise the full robust objective governs the fit (Appendix B.2.3). RAC retains its best feasible numerical fit. Coverage holds for any admissible threshold, including empirical-mean and fixed-threshold ablations (Appendix D.3).

Before observing evaluation data, freeze the corrected score and a finite $\mathcal{F}$-measurable enclosure $[b_-, b_+]$ of its support throughout $\mathcal{H} \times \{0,1\}^2$ on every nonabstaining construction, including combinations absent from construction data. Normalize to $[0,1]$ and apply predictable positive-capital betting (Waudby-Smith & Ramdas, 2024, Theorem 3 and Remark 3) to lower-bound its conditional mean. The resulting improvement bound is $\widehat{L}$ ($-\infty$ on abstention); certify when $\widehat{L} > 0$. Appendix B.2.4 gives conservative inversion and constant-support conventions.

Theorem 2 (Pointwise complete-procedure protection)**.** *Under this section's assumptions, for* $\alpha_e + \alpha_\mu < 1$ *and* $0 < \alpha_s < 1$*,* RAC *satisfies*

$$P(\widehat{L} \le M_H) \ge (1 - \alpha_e - \alpha_\mu)(1 - \alpha_s). \tag{10}$$

The guarantee combines two events: simultaneous nuisance coverage has probability at least $1 - \alpha_e - \alpha_\mu$, and independent betting covers the frozen score mean conditionally with probability at least $1 - \alpha_s$. Integrating the conditional guarantee gives (10). The nuisance events need not be independent, and abstention always covers (proof: Appendix B.2). Allocations $\alpha_e = \alpha_\mu = 0.0125$, $\alpha_s = 0.025$ give 0.950625 for each specified procedure; Appendix B.2.6 extends protection to simultaneous selection.

Algorithm 1 Rectangular augmented certification (RAC)

Require: Binary rewards, always-treat target, fixed λ, declared overlap, interface X, construction information $\mathcal{F}$, independent i.i.d. evaluation sample $\mathcal{D}_s$.

1: Construct simultaneous full-history nuisance envelopes. Intersect the declared ranges; abstain with $\widehat{L} = -\infty$ if infeasible.
2: Project nuisance fits and form the corrected score in (7).
3: Fit an admissible $a \in \mathcal{A}_X$ by (9).
4: Freeze the score and a finite valid support enclosure $[b_-, b_+]$.
5: Evaluate predictable betting capital through the fixed horizon and conservatively invert its running acceptance set.
6: **return** Improvement LCB $\widehat{L}$; certify when $\widehat{L} > 0$.

5.3 Envelope constructions and certificate shortfall

For K histories, cellwise calibration uses independent samples and Bonferroni Clopper–Pearson intervals (Clopper & Pearson, 1934) with coordinate errors α_e/K and α_μ/K (Appendix B.2.1: counts, empty cells and projection). All $4K$ declared history–action–reward combinations determine support: combinations absent from construction data can still occur during evaluation. Alternatively, correctly specified linear nuisance models with bounded feature maps and coefficient-norm bounds fixed before calibration and fitting yield simultaneous ridge envelopes (Appendix B.2.2, Corollary 2). Features may vary within normalization cells; the outcome design uses treated fitting records. Dimension and treated-design leverage control widths.

From normalization loss to the final certificate. For interface X, write $\widehat{L}_X$ for its bound. Fitting regret $J_X(a) = M_X - \{D_X(a) - B\}$ measures the chosen threshold's shortfall from its interface's optimal dual value. Set $J_X = J_X(\widehat{a})$. On nonabstention,

$$\begin{aligned} M_H - \widehat{L}_X = (M_H - M_X) + J_X + \big[D_X(\widehat{a}) - B - \mathbb{E}(g_{\widehat{a}} \mid \mathcal{F})\big] \\ + \big[\mathbb{E}(g_{\widehat{a}} \mid \mathcal{F}) - \widehat{L}_X\big]. \end{aligned} \tag{11}$$

The terms are normalization loss (L_ϕ for Z, zero for H), fitting regret, net nuisance deficit (correction minus signed bias; (24)), and evaluation gap (score mean minus betting bound). The last two are nonnegative on nuisance and score-mean coverage events, respectively.

6 Experiments

Population studies measure normalization loss; finite-data studies assess certificate strength and empirical coverage under Theorem 2. Known-law synthetic models and HC–MNIST (real images with simulated assignment, confounding and rewards; Jesson et al., 2021; Melnychuk et al., 2024) supply always-treat population evidence (Appendices C.2–C.3). Binary-reward studies compare RAC with the matched protected weak dual, empirical-mean fitting ($\rho = 0$) and fixed thresholds. Finite-sample comparisons precede component ablations and calibration studies.

6.1 POPULATION CHARACTERIZATION

We isolate outcome alignment in a balanced four-history law with propensities $(0.2, 0.8, 0.2, 0.8)$ and treated Bernoulli means $(\omega, \omega, 1-\omega, 1-\omega)$, where ω is reward noise. Outcome-pair and crossed-pair interfaces match two equal-mass cells, coarse propensity $1/2$ and directional odds distortion 4. At $\lambda = 2, \omega = 0.1$, outcome pairs preserve $V_H = 0.4625$; crossed pairs give $V_Z = 0.4025$, losing 0.0600. Crossed quantile intervals first intersect at $\omega = 1/3$, where loss vanishes (Figure 1). The equal-spread groupings in Figure 1 have different normalization losses.

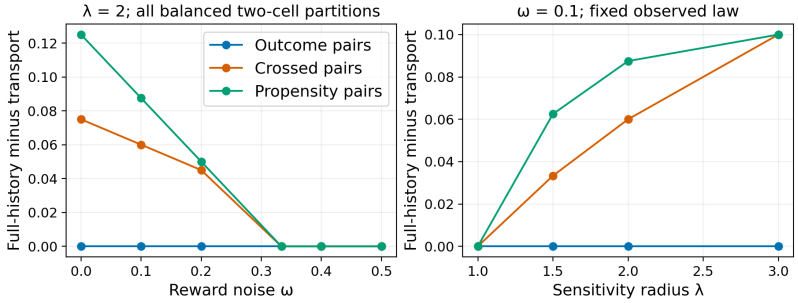

Figure 1: Normalization loss $V_H - V_Z$ across three groupings. $\lambda = 2$ (left), $\omega = 0.1$ (right). Outcome pairs are lossless; crossed and propensity pairs become lossless at $\omega = 1/3$ for $\lambda = 2$.

Across 24 positive-effect synthetic laws and three interfaces, coarsening retains 49/54 positive full-history certificates. On HC–MNIST, 2,700 learned interfaces at four positive effects retain all 3,600 positive certificates among 10,800 evaluations (Table 1). Mean/maximum loss over the complete image grid is $0.000603/0.001924$, despite matching decisions (Appendix D.1).

Table 1: Certificate retention under normalization coarsening. Counts of positive known-law certificates among N_{eval} policy–interface evaluations; $S = e(H)$ refines normalization. Grouping largely preserves positive decisions while weakening endpoints.

Grid	N_{eval} evaluations	Full-history reference V_H	Refined reference $V_{(Z,S)}$	**Coarse RAC target** (V_Z)	Directional envelope	Symmetric envelope	Global envelope
Synthetic	72	54	50	49	20	18	18
HC–MNIST	10,800	3,600	3,600	3,600	60	0	0

6.2 CERTIFICATE STRENGTH AND COVERAGE

We compare fitted certificate strength and empirical coverage. The four-history study crosses $\lambda = 2$, cutoff separation $\epsilon \in \{0, 0.01, 0.10\}$, margin $M_H \in \{-0.02, 0, 0.05\}$ and total budget $N \in \{10{,}000, 40{,}000\}$: 18 conditions, 1,000 datasets each. Treated means $2/3 \pm \epsilon$ yield nonunique thresholds at $\epsilon = 0$; $M_H = 0$ is the robust decision boundary. Coarse/full-history thresholds use two/four cells; nuisances retain four histories. Protected procedures allocate $N/2$ to calibration, $N/4$ to fitting and $N/4$ to evaluation (Appendix C.4).

Coverage is $\widehat{L} \leq M_H$; certification is $\widehat{L} > 0$. Doubly-valid/doubly-sharp (DVDS; Dorn et al., 2025) uses Wald inference; quantile balancing (QB; Dorn & Guo, 2023) uses five-fold cross-fitting and percentile bootstrap; kernel conditional moment constraints (KCMC; Ishikawa & He, 2023) uses a bootstrap adaptation. All use $N/2$ reward records; protected procedures reserve half for independent evaluation.

DVDS, QB and KCMC yield stronger LCBs and higher robust-null false certification than RAC (Table 2); DVDS's 5.8% rate has Wilson 95% interval [4.5,7.4]%. This discrete, binary-reward law falls outside the density assumptions for the DVDS preprint (v2) and QB. Appendices C.6 and D.5 give inference assumptions, sample use and all 18 conditions on the same 18,000 datasets.

Table 2: Certificate strength and robust-null false certification. Four histories, $\epsilon = 0, N = 40{,}000$; 1,000 datasets per condition, 95% nominal coverage. Stronger LCBs accompany higher null false-certification rates.

Procedure	Mean LCB ($M_H = 0.05$)	False certification ($M_H = 0$)
RAC (ours; cellwise, coarse)	0.039834	1.0%
DVDS	0.043503	5.8%
QB (five-fold cross-fitted)	0.045217	9.8%
KCMC (bootstrap adaptation)	0.045064	13.9%

6.3 Augmentation and threshold fitting

The direct protected weak dual matches fitting, independent betting and complete protection, with propensity and betting errors both 0.025. At $(\epsilon, M_H) = (0, 0.05)$, coarse RAC certifies 1,000/1,000 datasets at $N = 40{,}000$ and 997 versus direct's 702 at $N = 10{,}000$ (Table 3).

Table 3: Do augmentation and confidence-aware fitting strengthen certificates? Coarse thresholds, $\epsilon = 0$, $M_H = 0.05$; 1,000 paired datasets per budget, 95% nominal coverage. Brackets: 95% Wilson rate and Monte Carlo mean-LCB intervals.

Total budget N	Procedure	Coverage (%)	Certification (%)	Mean LCB
10,000	**RAC (ours; coarse)**	99.4 [98.7, 99.7]	99.7 [99.1, 99.9]	0.02702 [0.02647, 0.02757]
10,000	RAC ablation ($\rho = 0$)	99.7 [99.1, 99.9]	89.4 [87.3, 91.2]	0.02132 [0.02035, 0.02229]
10,000	RAC ablation (fixed $a = 0$)	99.4 [98.7, 99.7]	100.0 [99.6, 100.0]	0.02718 [0.02664, 0.02771]
10,000	Direct protected weak dual	100.0 [99.6, 100.0]	70.2 [67.3, 73.0]	0.00514 [0.00444, 0.00583]
40,000	**RAC (ours; coarse)**	99.6 [99.0, 99.8]	100.0 [99.6, 100.0]	0.03983 [0.03958, 0.04009]
40,000	RAC ablation ($\rho = 0$)	99.4 [98.7, 99.7]	99.9 [99.4, 100.0]	0.03607 [0.03559, 0.03655]
40,000	RAC ablation (fixed $a = 0$)	99.7 [99.1, 99.9]	100.0 [99.6, 100.0]	0.03995 [0.03970, 0.04020]
40,000	Direct protected weak dual	100.0 [99.6, 100.0]	100.0 [99.6, 100.0]	0.02731 [0.02697, 0.02765]

Augmentation strengthens bounds and certification through lower nuisance-protection costs and score variance (Appendix D.2). The direct comparator certifies 0/1,000 datasets in every robust-null condition.

Across six positive-margin coarse conditions, confidence-aware fitting raises mean LCB by 0.003152 and adds 238 certificates/6,000 over empirical-mean fitting. Most of this gain comes from smaller evaluation gaps and net nuisance deficits; at the exact cutoff, thresholds with essentially equal population value can yield different certificate strength. Fixed thresholds remain competitive (Appendices D.3–D.4: uncertainty, variance and orientation ablations).

6.4 Restriction trade-off

We test whether sharing thresholds offsets structural loss through lower statistical uncertainty. In the cutoff family, $L_\phi = 0.15\epsilon$. Coarse thresholds slightly strengthen certificates when this loss is zero; at $\epsilon = 0.01$, their advantage is unresolved. At $\epsilon = 0.10$, structural loss dominates and full-history RAC is stronger (Table 4). Appendix D.2.1 decomposes these comparisons into threshold regret, nuisance deficit and evaluation gap.

6.5 Structured calibration at finer history resolution

We test shared nuisance structure at $K \in \{4, 16, 64, 256\}$ histories. All procedures use two threshold cells, matched splits and protection. Structured RAC uses correctly specified four-dimensional propensity and treated-outcome models; the direct comparator uses propensity only. Cellwise RAC calibrates probabilities by history (Appendix C.7).

Table 4: When does coarsening strengthen a certificate? RAC at $N = 40{,}000, M_H = 0.05$; 1,000 paired datasets per row. Small loss permits modest gains; larger loss favors full history. Brackets: 95% paired Monte Carlo intervals.

Cutoff separation ϵ	Exact loss L_ϕ	Coarse RAC mean LCB	Full-history RAC mean LCB	Paired LCB difference (coarse minus full history)
0.00	0.0000	0.039834	0.039596	+0.000237 [+0.000182, +0.000293]
0.01	0.0015	0.038159	0.038085	+0.000074 [-0.000002, +0.000149]
0.10	0.0150	0.024667	0.037795	−0.013128 [-0.013250, -0.013006]

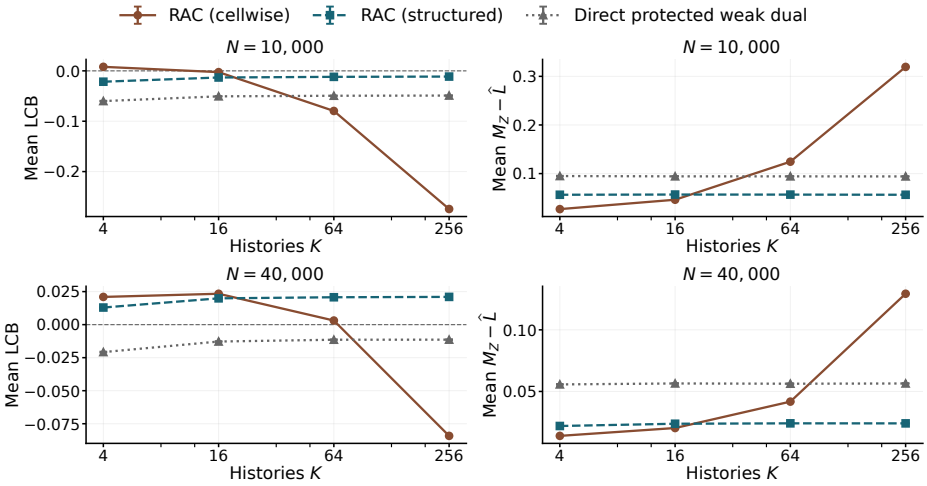

Figure 2: How does nuisance calibration scale with history support? Higher mean LCBs and lower remaining shortfall $M_Z - \widehat{L}$ are better. $M_H = 0.05$; 1,000 paired datasets per point, 95% Monte Carlo intervals. Structured mean bounds remain negative at $N = 10{,}000$.

Figure 2 favors cellwise RAC at $K = 4, 16$ and structured RAC at $K = 64, 256$ under both budgets. At $K = 256, N = 40{,}000$, structured RAC certifies 1,000/1,000 datasets, versus 0 for cellwise RAC and 17 for the direct comparator; all three cover 1,000/1,000. Smaller net nuisance deficits explain this gain. Appendix B.2.2 bounds the structured nuisance deficit using weighted products of prediction widths. Across supports, raw LCBs rise partly because structural loss falls, while $M_Z - \widehat{L}$ rises slightly. Appendix D.6 gives certification rates (Figure 3) and decompositions.

7 Conclusion

Our attained weighted-pinball characterization and shared-quantile criterion show exactly when a measurable normalization interface preserves the population certificate, including with atomic rewards. RAC provides finite-sample protection for fitted binary-reward, always-treat certificates through nuisance envelopes and independent betting. Experiments distinguish population preservation from statistical certificate strength. Augmentation strengthens the matched protected weak dual. Across tested positive-margin conditions, confidence-aware fitting improves average certificate strength over empirical-mean fitting, while fixed thresholds remain competitive. Coarsening trades structural loss against statistical uncertainty. Structured calibration reduces protection costs at large supports under correct specification.

The characterization separates what a normalization interface preserves from how accurately its certificate can be estimated. This distinction motivates choosing normalization resolution and nuisance-model resolution separately. Future work can develop interface-selection rules that balance structural loss against statistical certification costs, and extend calibration to learned nuisance features and misspecified models.

Reproducibility Statement

Sections 3–5 specify the estimands, assumptions, fitting objective and coverage allocation. Appendix B provides the proofs. Appendix C describes data generation, representation training, sample splits, inference settings and comparator implementations. Appendix D reports condition-level results, ablations and uncertainty estimates. We will release the code and experiment configurations upon acceptance.

AI Use Statement

In this work, we used generative AI tools to polish writing, assist with literature retrieval and discovery, and refinement of mathematical statements and proofs. All AI-assisted content was reviewed and verified by the authors. The authors take full responsibility for the final content of the submission.

References

Yasin Abbasi-Yadkori, Dávid Pál, and Csaba Szepesvári. Improved algorithms for linear stochastic bandits. In *Advances in Neural Information Processing Systems*, volume 24, 2011. URL https://proceedings.neurips.cc/paper_files/paper/2011/hash/e1d5be1c7f2f456670de3d53c7b54f4a-Abstract.html.

David Arthur and Sergei Vassilvitskii. k-means++: The advantages of careful seeding. In *Proceedings of the 18th Annual ACM-SIAM Symposium on Discrete Algorithms*, pp. 1027–1035. Society for Industrial and Applied Mathematics, 2007. URL https://dl.acm.org/doi/10.5555/1283383.1283494.

Andrew Bennett, Nathan Kallus, Miruna Oprescu, Wen Sun, and Kaiwen Wang. Efficient and sharp off-policy evaluation in robust Markov decision processes. In *Advances in Neural Information Processing Systems*, volume 37, pp. 112962–113000, 2024. doi: 10.52202/079017-3590. URL https://proceedings.neurips.cc/paper_files/paper/2024/hash/cd1da8043ba5c1c144ab4e10a8de6e53-Abstract-Conference.html.

Jonathan D. Chang, Kaiwen Wang, Nathan Kallus, and Wen Sun. Learning Bellman complete representations for offline policy evaluation. In *Proceedings of the 39th International Conference on Machine Learning*, volume 162 of *Proceedings of Machine Learning Research*, pp. 2938–2971. PMLR, 2022. URL https://proceedings.mlr.press/v162/chang22b.html.

C. J. Clopper and E. S. Pearson. The use of confidence or fiducial limits illustrated in the case of the binomial. *Biometrika*, 26(4):404–413, 1934. doi: 10.1093/biomet/26.4.404.

Marco Cuturi. Sinkhorn distances: Lightspeed computation of optimal transport. In *Advances in Neural Information Processing Systems*, volume 26, 2013. URL https://proceedings.neurips.cc/paper/2013/hash/af21d0c97db2e27e13572cbf59eb343d-Abstract.html.

Jacob Dorn and Kevin Guo. Sharp sensitivity analysis for inverse propensity weighting via quantile balancing. *Journal of the American Statistical Association*, 118(544):2645–2657, 2023. doi: 10.1080/01621459.2022.2069572.

Jacob Dorn, Kevin Guo, and Nathan Kallus. Doubly-valid/doubly-sharp sensitivity analysis for causal inference with unmeasured confounding. *Journal of the American Statistical Association*, 120(549):331–342, 2025. doi: 10.1080/01621459.2024.2335588.

John Duchi and Hongseok Namkoong. Variance-based regularization with convex objectives. *Journal of Machine Learning Research*, 20(68):1–55, 2019. URL https://jmlr.org/papers/v20/17-750.html.

Dennis Frauen, Fergus Imrie, Alicia Curth, Valentyn Melnychuk, Stefan Feuerriegel, and Mihaela van der Schaar. A neural framework for generalized causal sensitivity analysis. In *International Conference on Learning Representations*, 2024.

URL https://proceedings.iclr.cc/paper_files/paper/2024/hash/7e83fd2ff7ae58485d418685521c9608-Abstract-Conference.html.

Arthur Gretton, Karsten M. Borgwardt, Malte J. Rasch, Bernhard Schölkopf, and Alexander Smola. A kernel two-sample test. *Journal of Machine Learning Research*, 13(25):723–773, 2012. URL https://jmlr.org/papers/v13/gretton12a.html.

Konstantin Hess, Dennis Frauen, Valentyn Melnychuk, and Stefan Feuerriegel. Efficient and sharp off-policy learning under unobserved confounding. In *International Conference on Learning Representations*, 2026. URL https://proceedings.iclr.cc/paper_files/paper/2026/hash/5e972fe492946fc7a02aef64d0f4fd7e-Abstract-Conference.html.

Kei Ishikawa and Niao He. Kernel Conditional Moment Constraints for confounding robust inference. In *Proceedings of the 26th International Conference on Artificial Intelligence and Statistics*, volume 206 of *Proceedings of Machine Learning Research*, pp. 650–674. PMLR, 2023. URL https://proceedings.mlr.press/v206/ishikawa23a.html.

Kei Ishikawa, Niao He, and Takafumi Kanamori. A convex framework for confounding robust inference. *Journal of Machine Learning Research*, 27(62):1–53, 2026. URL https://jmlr.org/papers/v27/23-1434.html.

Andrew Jesson, Sören Mindermann, Yarin Gal, and Uri Shalit. Quantifying ignorance in individual-level causal-effect estimates under hidden confounding. In *Proceedings of the 38th International Conference on Machine Learning*, volume 139 of *Proceedings of Machine Learning Research*, pp. 4829–4838. PMLR, 2021. URL https://proceedings.mlr.press/v139/jesson21a.html.

Fredrik Johansson, Uri Shalit, and David Sontag. Learning representations for counterfactual inference. In *Proceedings of the 33rd International Conference on Machine Learning*, volume 48 of *Proceedings of Machine Learning Research*, pp. 3020–3029. PMLR, 2016. URL https://proceedings.mlr.press/v48/johansson16.html.

Alkis Kalavasis, Anay Mehrotra, and Manolis Zampetakis. Smaller confidence intervals from IPW estimators via data-dependent coarsening (extended abstract). In *Proceedings of the 37th Conference on Learning Theory*, volume 247 of *Proceedings of Machine Learning Research*, pp. 2767. PMLR, 2024. URL https://proceedings.mlr.press/v247/kalavasis24a.html.

Nathan Kallus and Angela Zhou. Confounding-robust policy improvement. In *Advances in Neural Information Processing Systems*, volume 31, 2018. URL https://proceedings.neurips.cc/paper_files/paper/2018/hash/3a09a524440d44d7f19870070a5ad42f-Abstract.html.

Nikos Karampatziakis, Paul Mineiro, and Aaditya Ramdas. Off-policy confidence sequences. In *Proceedings of the 38th International Conference on Machine Learning*, volume 139 of *Proceedings of Machine Learning Research*, pp. 5301–5310. PMLR, 2021. URL https://proceedings.mlr.press/v139/karampatziakis21a.html.

Diederik P. Kingma and Jimmy Lei Ba. Adam: A method for stochastic optimization. In *International Conference on Learning Representations*, 2015. URL https://arxiv.org/abs/1412.6980.

Roger Koenker and Gilbert Bassett, Jr. Regression quantiles. *Econometrica*, 46(1):33–50, 1978. doi: 10.2307/1913643.

Yann LeCun, Léon Bottou, Yoshua Bengio, and Patrick Haffner. Gradient-based learning applied to document recognition. *Proceedings of the IEEE*, 86(11):2278–2324, 1998. doi: 10.1109/5.726791. URL https://ieeexplore.ieee.org/document/726791.

Travis Mandel, Yun-En Liu, Sergey Levine, Emma Brunskill, and Zoran Popović. Offline policy evaluation across representations with applications to educational games. In *Proceedings of*

the 13th International Conference on Autonomous Agents and Multiagent Systems, pp. 1077–1084, 2014. URL https://www.ifaamas.org/Proceedings/aamas2014/aamas/p1077.pdf.

Valentyn Melnychuk, Dennis Frauen, and Stefan Feuerriegel. Bounds on representation-induced confounding bias for treatment effect estimation. In *International Conference on Learning Representations*, 2024. URL https://proceedings.iclr.cc/paper_files/paper/2024/hash/abe31a12e83111fdf2cfd54deed5a2ce-Abstract-Conference.html.

Valentyn Melnychuk, Dennis Frauen, Jonas Schweisthal, and Stefan Feuerriegel. Orthogonal representation learning for estimating causal quantities. In *Proceedings of the 29th International Conference on Artificial Intelligence and Statistics*, volume 300 of *Proceedings of Machine Learning Research*, pp. 5086–5094. PMLR, 2026. URL https://proceedings.mlr.press/v300/melnychuk26a.html.

Uri Shalit, Fredrik D. Johansson, and David Sontag. Estimating individual treatment effect: Generalization bounds and algorithms. In *Proceedings of the 34th International Conference on Machine Learning*, volume 70 of *Proceedings of Machine Learning Research*, pp. 3076–3085. PMLR, 2017. URL https://proceedings.mlr.press/v70/shalit17a.html.

Adith Swaminathan and Thorsten Joachims. Counterfactual risk minimization: Learning from logged bandit feedback. In *Proceedings of the 32nd International Conference on Machine Learning*, volume 37 of *Proceedings of Machine Learning Research*, pp. 814–823. PMLR, 2015. URL https://proceedings.mlr.press/v37/swaminathan15.html.

Zhiqiang Tan. A distributional approach for causal inference using propensity scores. *Journal of the American Statistical Association*, 101(476):1619–1637, 2006. doi: 10.1198/016214506000000023.

Zhiqiang Tan. Model-assisted sensitivity analysis for treatment effects under unmeasured confounding via regularized calibrated estimation. *Journal of the Royal Statistical Society Series B: Statistical Methodology*, 86(5):1339–1363, 2024. doi: 10.1093/jrsssb/qkae034. URL https://academic.oup.com/jrsssb/article/86/5/1339/7664041.

Ian Waudby-Smith and Aaditya Ramdas. Estimating means of bounded random variables by betting. *Journal of the Royal Statistical Society Series B: Statistical Methodology*, 86(1):1–27, 2024. doi: 10.1093/jrsssb/qkad009.

Ian Waudby-Smith, Lili Wu, Aaditya Ramdas, Nikos Karampatziakis, and Paul Mineiro. Anytime-valid off-policy inference for contextual bandits. *ACM / IMS Journal of Data Science*, 1(3):1–42, 2024. doi: 10.1145/3643693. URL https://dl.acm.org/doi/10.1145/3643693.

Edwin B. Wilson. Probable inference, the law of succession, and statistical inference. *Journal of the American Statistical Association*, 22(158):209–212, 1927. doi: 10.1080/01621459.1927.10502953.

Yi Zhang, Wenfu Xu, and Zhiqiang Tan. Distributionally enhanced marginal sensitivity model and bounds. arXiv preprint arXiv:2505.13868, 2025. URL https://arxiv.org/abs/2505.13868.

Yi Zhang, Wenfu Xu, and Zhiqiang Tan. Enhanced marginal sensitivity model and bounds. *Journal of the American Statistical Association*, pp. 1–13, 2026. doi: 10.1080/01621459.2026.2699465. Advance online publication.

Qingyuan Zhao, Dylan S. Small, and Bhaswar B. Bhattacharya. Sensitivity analysis for inverse probability weighting estimators via the percentile bootstrap. *Journal of the Royal Statistical Society Series B: Statistical Methodology*, 81(4):735–761, 2019. doi: 10.1111/rssb.12327. URL https://academic.oup.com/jrsssb/article/81/4/735/7048357.

A Extended Related Work

A.1 Sensitivity bounds and normalization restrictions

Sensitivity analysis replaces an untestable absence of confounding with an explicit restriction on its strength. The marginal sensitivity model of Tan (2006) controls departures from the observed propensity score through treatment odds. This framework supports both causal-effect bounds and robust decisions. Kallus & Zhou (2018) optimize worst-case policy improvement relative to a baseline, while Hess et al. (2026) derive sharp policy-value bounds and semiparametrically efficient estimators for policy learning and improvement. Bennett et al. (2024) extend robust policy evaluation to infinite-horizon Markov decision processes under bounded transition perturbations. These works motivate certification under uncertainty; our population question concerns how much of a fixed policy's lower endpoint is lost when its normalization restrictions use less information. NeuralCSA extends sensitivity analysis to multiple sensitivity models, treatment types and causal queries (Frauen et al., 2024). Related extensions constrain the outcome response to unmeasured confounding through conditional means (Zhang et al., 2026) or density ratios (Zhang et al., 2025).

Sharpness depends on respecting the constraints implied by the observed law, as well as the sensitivity boxes. Dorn & Guo (2023) use quantile balancing to obtain sharp bounds for inverse propensity weighting. Tan (2024) express sharp and finite-feature relaxed bounds through weighted quantile regression. For finite interfaces, the excess-risk identity in this paper follows from those representations. Our result establishes attainment and the exact loss for arbitrary measurable interfaces on standard Borel spaces, and characterizes losslessness by a measurable threshold shared by the full conditional quantile sets. Using the complete sets matters for discrete outcomes: two histories can have different conventional quantiles and still admit a common optimal threshold.

Conditional-moment approximations provide another route to tractable sensitivity bounds. Ishikawa & He (2023) introduce kernel conditional moment constraints, and Ishikawa et al. (2026) develop a broader convex framework with specification-error bounds and conditions for zero error. Their restricted dual programs explain why approximating normalization can make a lower bound conservative. We study the particular restriction imposed by a coarsened information interface under fixed marginal-sensitivity boxes. In this setting, the restriction cost has an exact weighted pinball-risk expression and a necessary and sufficient condition for vanishing loss. This identifies which distinctions between histories the certificate needs to retain, alongside the broader approximation guarantees in prior work.

A.2 Representations and coarsening for causal evaluation

Representation learning seeks to make counterfactual estimation more accurate with limited observational data. Johansson et al. (2016) connect counterfactual inference to domain adaptation, and Shalit et al. (2017) bound treatment-effect estimation error using prediction error and imbalance between treated and control representations under strong ignorability. These results motivate balancing and dimension reduction. Certificate preservation also depends on the outcome thresholds required by the robust endpoint. Melnychuk et al. (2026) connect learned representations to Neyman-orthogonal estimation and characterize improvements in estimation error under a low-dimensional manifold assumption.

The closest analysis of information loss is Melnychuk et al. (2024), who show that a representation can discard observed confounders and bound the resulting confounding bias in treatment-effect estimation. Their framework compares full-covariate and representation propensities, using observed odds distortion to calibrate bounds. Our normalization comparison retains the full-history propensity reference and its sensitivity boxes, while reducing the information available to the normalization equations. Replacing that reference by a representation propensity also changes the uncertainty class. Distinguishing these operations allows us to isolate normalization loss and compare it with a conservative representation-based envelope at the same full-history reference.

Coarsening can also improve statistical precision. Kalavasis et al. (2024) study data-dependent coarsening for inverse propensity weighting under unconfoundedness, particularly when propensities are extreme or inaccurately estimated. Their objective concerns the accuracy of a point-identified causal estimator. Here, coarsening changes a robust population endpoint and can also affect the strength of a fitted lower confidence bound. We therefore distinguish the exact population loss from

threshold-fitting regret, nuisance uncertainty, and evaluation uncertainty. The normalization interface controls threshold sharing, while nuisance models can retain full-history features.

A.3 Augmented estimation and finite-sample certification

For the marginal sensitivity model, Zhao et al. (2019) develop asymptotically valid percentile-bootstrap intervals for inverse-weighted estimators under a correctly specified propensity model. Augmentation reduces sensitivity to nuisance estimation, but sharpness and validity are distinct requirements. The doubly-valid/doubly-sharp estimators of Dorn et al. (2025) combine propensity, quantile, and transformed-outcome regressions. Under their stated conditions, misspecified quantile models can still yield valid conservative bounds, and suitably estimated nuisances give semiparametric efficiency. Tan (2024) develops related augmented estimating functions and calibrated estimation for relaxed bounds. RAC builds on these population targets. An imperfect threshold can still define a conservative dual value; our finite-sample construction additionally corrects for nuisance uncertainty using simultaneous rectangular envelopes.

The inference guarantee depends on how that uncertainty is controlled. Hess et al. (2026) derive efficient estimation under local continuous-density assumptions, while Bennett et al. (2024) obtain asymptotic Wald inference in robust Markov decision processes. RAC addresses binary rewards with a finite-sample guarantee conditional on valid simultaneous nuisance envelopes, declared overlap, and independent evaluation. Cellwise calibration and correctly specified structured nuisance models give two ways to construct those envelopes. Finite-sample coverage requires nuisance envelopes as well as independent evaluation.

Certificate fitting also needs to account for variability. Swaminathan & Joachims (2015) incorporate the variance of propensity-weighted empirical risk into learning from logged bandit feedback. Duchi & Namkoong (2019) connect chi-square distributionally robust optimization with variance regularization. RAC applies this established principle to its corrected score because thresholds with the same population value can have different variances and nuisance corrections. The resulting fit is frozen before independent evaluation. With observed logging probabilities, Karampatziakis et al. (2021) construct finite-sample confidence sequences for contextual-bandit policy evaluation; Waudby-Smith et al. (2024) extend this approach to adaptive data collection and doubly robust scores. Waudby-Smith & Ramdas (2024) provide the bounded-mean betting machinery used to protect the frozen score's conditional mean. Fitting aims to strengthen the certificate, while the envelope correction and independent betting evaluation provide its coverage guarantee for any admissible fitted threshold.

B Proofs and Additional Theoretical Results

B.1 Normalization loss and sensitivity bounds

B.1.1 Feasible weights and endpoint ordering

The boxes in (2) equivalently satisfy $\ell = e_{\mathrm{hi}}(e,\lambda)^{-1}$ and $u = e_{\mathrm{lo}}(e,\lambda)^{-1}$, where $e_{\mathrm{hi}}(p,\lambda) = \lambda p/[1+(\lambda-1)p]$ and $e_{\mathrm{lo}}(p,\lambda) = p/[\lambda+(1-\lambda)p]$.

Every compatible causal law induces

$$w_P(H,R) = \mathbb{E}_P\{q(H,U)^{-1} \mid H, A=1, R\}, \quad \mathbb{E}(Aw_P \mid H) = 1, \quad \mathbb{E}(ARw_P) = V_P. \tag{12}$$

Conditioning a full-history normalization equation on Z gives $\mathbb{E}(Aw \mid Z) = 1$, retaining the same boxes and objective. For a compatible causal law, the induced weight in (12) lies in these boxes. With $\mathbb{E}_{\mathrm{obs}}$ denoting observed-law expectation,

$$\mathbb{E}_{\mathrm{obs}}(Aw_P \mid H) = \mathbb{E}_P\{A/q(H,U) \mid H\} = 1, \qquad \mathbb{E}_{\mathrm{obs}}(ARw_P) = \mathbb{E}_P\{AR/q(H,U)\} = V_P,$$

by consistency and latent exchangeability. This proves the endpoint ordering in Section 3.

For a finite observed law, let i index atoms (H_i, A_i, R_i) with mass p_i, let $d(i)$ denote the history cell, and set $B_d^{\mathrm{full}}(\lambda) = [\ell(d), u(d)]$. Writing $\underline{V}_\phi^{\mathrm{tr}} = V_Z$ for the transported endpoint, $\underline{V}_{\mathrm{full}} = V_H$

for the full-history endpoint, and c for a summary cell, the grouped program is

$$\begin{aligned} \underline{V}_{\phi}^{\mathrm{tr}}(\lambda) &= \min_{w} \sum_{i:A_i=1} p_i R_i w_i, \\ \text{s.t.} \quad & w_i \in B_{d(i)}^{\mathrm{full}}(\lambda), \qquad \sum_{i:\phi\{d(i)\}=c, A_i=1} p_i w_i = P(Z=c) \quad \forall c. \end{aligned} \tag{13}$$

B.1.2 Dual and quantile characterization of transport loss

Proof of Theorem 1. Standard-Borel regularity gives a conditional reward kernel $F_h = \mathrm{Law}(R \mid H = h, A = 1)$. Uniform overlap and $\mathbb{E}|R| < \infty$ imply

$$\mathbb{E}_H \int |r| F_H(dr) \le \eta^{-1} \mathbb{E}(A|R|) < \infty.$$

Modify the kernel on a null set so its first moment is finite. For $X \in \{H, Z\}$, let $K_x(dh)$ be a regular conditional kernel for H given $X = x$; for $X = H$, take $K_h = \delta_h$, the point mass at h. For a Borel reward set $D \subseteq \mathbb{R}$, define the pooled weight mass $C_X(x)$ and weighted reward kernel G_x by

$$C_X(x) = \int c(h) K_x(dh), \qquad G_x(D) = \frac{\int c(h) F_h(D) K_x(dh)}{C_X(x)} \tag{14}$$

Because overlap bounds c above and away from zero, $C_X > 0$ almost surely; define G arbitrarily on the exceptional set. Equation 14 defines one probability kernel, avoiding incompatible versions of conditional CDFs.

The lower quantile, with $\mathbb{Q}$ denoting the rational numbers,

$$a_X(x) = \inf\{t \in \mathbb{Q} : G_x((-\infty, t]) \ge \tau\}$$

is measurable and satisfies $G_x(a_X-) \le \tau \le G_x(a_X)$. A τ-quantile a of any integrable law G obeys

$$|a| \le \frac{\int |r| G(dr)}{\min(\tau, 1-\tau)}. \tag{15}$$

Together with the upper and lower bounds on c, this makes $a_X(X)$ integrable. Conversely, pinball risk satisfies $\mathbb{E}\rho_\tau(R-a) \ge \min(\tau, 1-\tau)\{|a| - \mathbb{E}|R|\}$, so restricting to integrable thresholds does not change the finite optimum. The subgradient of the conditional pinball risk at a is $[F_h(a-) - \tau, F_h(a) - \tau]$; hence its minimizers are exactly $Q_\tau(h)$, including when F_h has an atom.

For primal attainment, write $e = e_{\mathrm{full}}$ and note

$$e(u - \ell) = c, \qquad 1 - e\ell = \tau c. \tag{16}$$

Let $p_-(x) = G_x((-\infty, a_X(x)))$ and $p_0(x) = G_x(\{a_X(x)\})$ be the pooled masses below and at the threshold. The fraction of threshold-atom mass assigned the upper weight is

$$\theta(x) = \begin{cases} (\tau - p_-(x))/p_0(x), & p_0(x) > 0, \\ 0, & p_0(x) = 0. \end{cases}$$

Kernel measurability and the measurable threshold make p_-, p_0, and θ measurable. The quantile inequalities give $0 \le \theta \le 1$ and $p_- + \theta p_0 = \tau$. Define the attaining inverse weight w_X by

$$w_X(h, r) = \ell(h) + \{u(h) - \ell(h)\}\,[\mathbf{1}\{r < a_X(X(h))\} + \theta(X(h))\mathbf{1}\{r = a_X(X(h))\}] \tag{17}$$

Here $X(h) = h$ for $X = H$ and $X(h) = \phi(h)$ for $X = Z$. The weight is measurable, lies in its box, and, by (14)–(16), satisfies $\mathbb{E}(A w_X \mid X) = 1$. The fractional value at a threshold atom is what ensures exact normalization.

For every feasible w and integrable X-measurable a, weak duality and pointwise minimization give

$$\mathbb{E}(ARw) \ge \mathbb{E}\{a(X)\} + \mathbb{E}\left[A \min_{v \in [\ell(H), u(H)]} v\{R - a(X)\}\right] = \mu - \mathcal{R}_\lambda(a), \tag{18}$$

where $\mu = \mathbb{E}_H \int r F_H(dr)$. The weight in (17) attains every pointwise minimum in (18), so

$$V_X(\lambda) = \mu - \min_{a:\, X\text{-measurable}} \mathcal{R}_\lambda(a).$$

Subtracting the cases $X = H$ and $X = Z$ proves (4). At an attained pooled threshold, the loss is the expectation of the nonnegative conditional excess risk. It is zero exactly when the pooled threshold lies in $Q_\tau(H)$ almost surely. Conversely, any measurable shared selection attains the full risk and is integrable by (15). For finite histories, this condition is the cellwise intersection criterion below, and expansion by interface cells yields (19). At $\lambda = 1$, $\ell = u = e^{-1}$, giving the same unique weight for both normalizations.

For causal sharpness, let $F_h^0 = \mathrm{Law}(R \mid H = h, A = 0)$. Every compatible causal law induces $w_P(h, r) = \mathbb{E}\{q(H, U)^{-1} \mid H = h, A = 1, R = r\}$, which is full-history feasible and satisfies $\mathbb{E}(ARw_P) = \mathbb{E}\{R(1)\}$. Thus every compatible value is at least V_H. Conversely, for any feasible full-history w, define the probability kernel

$$T_h(dr) = e(h)w(h, r)F_h(dr).$$

Given $H = h$, draw $Y_1 \sim T_h$ and $Y_0 \sim F_h^0$ independently, put $U = (Y_1, Y_0)$, draw A with probability $q(h, U) = 1/w(h, Y_1)$, and set $R(a) = Y_a$. Then

$$P(A = 1, Y_1 \in D \mid H = h) = e(h)F_h(D), \qquad P(A = 0, Y_0 \in D \mid H = h) = (1 - e(h))F_h^0(D).$$

Hence both observed arms are reproduced. Potential outcomes are deterministic given (H, U), the box for w is equivalent to the marginal-sensitivity odds band, and the causal value is $\mathbb{E}(Y_1) = \mathbb{E}(ARw)$. Applying this construction to the attaining weight proves attained sharpness. □

Finite cells and discrete outcomes. For finite H and $\lambda > 1$, let $P_d = P(H = d)$, $e_d = e_{\mathrm{full}}(d)$, and $C_d = P_d(1 - e_d)(\lambda - \lambda^{-1})$. Writing $d \subset c$ for $\phi(d) = c$, expansion by positive-mass interface cells gives

$$L_\phi(\lambda) = \sum_c \left[\min_a \sum_{d \subset c} C_d \mathbb{E}_{F_d} \rho_\tau(R - a) - \sum_{d \subset c} C_d \min_a \mathbb{E}_{F_d} \rho_\tau(R - a) \right]. \tag{19}$$

Each nonnegative bracket vanishes exactly when its histories share a minimizer: $\bigcap_{d:\phi(d)=c,\, P_d>0} Q_\tau(d) \neq \varnothing$ for every positive-mass cell c. In the Bernoulli experiments, if m_d is the treated success probability, the closed quantile set is

$$Q_\tau(d) = \begin{cases} \{0\}, & \tau < 1 - m_d, \\ [0, 1], & \tau = 1 - m_d, \\ \{1\}, & \tau > 1 - m_d. \end{cases}$$

Corollary 1 (Fixed deterministic binary policy)**.** *Let $\pi(H) = \widetilde{\pi}\{\phi(H)\}$ be fixed and measurable, put $M^\pi = \mathbf{1}\{A = \pi(H)\}$ and $s(h) = P\{M^\pi = 1 \mid H = h\}$, and assume uniform overlap for s. For $1 < \lambda < \infty$, replacing $(A, e_{\mathrm{full}}, F_h)$ in Theorem 1 by $(M^\pi, s, \mathrm{Law}\{R \mid H = h, M^\pi = 1\})$ gives the same loss and losslessness characterization, with $c^\pi(h) = \{1 - s(h)\}(\lambda - \lambda^{-1})$. The common factual baseline cancels from the two margins.*

Proof of Corollary 1. Let $F_h^\pi = \mathrm{Law}\{R \mid H = h, M^\pi = 1\}$ and $q_\pi(H, U) = P\{A = \pi(H) \mid H, U\}$. On $\{\pi(H) = 0\}$, the match-action odds ratio is the reciprocal of the action-one odds ratio, so it remains in $[\lambda^{-1}, \lambda]$; on $\{\pi(H) = 1\}$ it is unchanged. Thus the match propensity has nominal value $s(H)$ and box

$$\left[1 + \frac{1 - s(H)}{\lambda s(H)},\, 1 + \frac{\lambda(1 - s(H))}{s(H)} \right].$$

Because $\pi = \widetilde{\pi} \circ \phi$, the policy action is retained by the normalization interface. Conditional on H, M^π relabels the binary action, and consistency and exchangeability give $\mathbb{E}\{M^\pi / q_\pi(H, U) \mid H\} = 1$ and $\mathbb{E}\{M^\pi R / q_\pi(H, U)\} = \mathbb{E}[R\{\pi(H)\}]$. Theorem 1 now applies with (A, e, F_h) replaced by (M^π, s, F_h^π). The common factual baseline cancels from the two margins. □

B.1.3 Reference metadata and normalization-interface structure

Proposition 1 (Reference-metadata sufficiency and separation)**.** *For finite support, the law of (Z, S, A, R) with $S = e(H)$ determines V_Z, but need not determine V_H.*

For the always-treat target, set $S = e_{\text{full}}(H)$ and let $Z = \phi(H)$ be a deterministic normalization interface. Let P^H_{obs} denote the observed history–action–reward law. Aggregate its finite atoms by (Z, S, A, R) and define

$$p_{zsar} := P^H_{\text{obs}}(Z = z, S = s, A = a, R = r).$$

We write $\underline{V}^{\text{tr}}_Z$ for $\underline{V}^{\text{tr}}_\phi$ when indexing the endpoint by its normalization interface.

For the always-treat specialization, the exact-population reduced transported program uses one inverse weight v_{zsr} per treated (z, s, r) group:

$$\begin{aligned} \underline{V}^{\text{tr,meta}}_Z(\lambda) = \min_v & \sum_{z,s,r} p_{zs1r} r\, v_{zsr}, \\ \text{s.t.} \quad & v_{zsr} \in \left[\frac{1}{e_{\text{hi}}(s,\lambda)}, \frac{1}{e_{\text{lo}}(s,\lambda)}\right] \quad \forall z, s, r, \\ & \sum_{s,r} p_{zs1r} v_{zsr} = P(Z = z) \qquad \forall z. \end{aligned} \tag{20}$$

Proof of Proposition 1. For sufficiency, start from the H-indexed transported program (13). Within a fixed normalization cell z, every fine treated atom with common $(S = s, R = r)$ has the same box and the same reward coefficient. Replacing its individual weighted masses by their sum $p_{zs1r}v_{zsr}$ preserves both the objective and the z-normalization equality, which gives (20). Conversely, any solution of (20) can be assigned to each fine atom in its (z, s, r) group. It satisfies the same boxes and the same summed normalization equation. The two LPs therefore have equal optima, and the reduced LP uses only $\text{Law}(Z, S, A, R)$.

For separation, take two histories h_1, h_2, each with mass $1/2$, let $Z \equiv z$, and set $e_{\text{full}}(h_1) = e_{\text{full}}(h_2) = 1/2$. In P_1, treated h_1 has $R = 0$ and treated h_2 has $R = 1$. In P_2, each treated history has $R = 0, 1$ with equal probability. Controls have $R = 0$ in both laws. Thus, under both laws,

$$P(Z = z, S = 1/2, A = 1, R = 0) = P(Z = z, S = 1/2, A = 1, R = 1) = \tfrac{1}{4},$$

and $P(Z = z, S = 1/2, A = 0, R = 0) = 1/2$. The reduced laws are identical. In P_1, the separate full-history equality on h_2 fixes the inverse-propensity weight at 2, so $\underline{V}^{P_1}_{\text{full}}(\lambda) = 1/2$. In P_2, the adversary gives its $R = 1$ atoms the lower box endpoint $1 + 1/\lambda$, which is feasible after compensating on the $R = 0$ atoms. Hence

$$\underline{V}^{P_2}_{\text{full}}(\lambda) = \frac{1}{4}\left(1 + \frac{1}{\lambda}\right) < \frac{1}{2} \qquad (\lambda > 1).$$

The common reduced program has the latter value for both laws. □

The variable S supplies the full-history reference boxes, whereas Z indexes the normalization equations. If $h \mapsto (\phi(h), e_{\text{full}}(h))$ is injective on positive-mass history cells, (Z, S) retains history-cell identity and exact (Z, S) normalization recovers the full-history equations. Without injectivity, distinct histories can share (Z, S), so their separate normalization equations are generally not determined by the reduced law. For a fixed deterministic binary target, the same reduction uses (Z, S^π, M^π, R), where

$$S^\pi(H) = P^H_{\text{obs}}\{A = \pi(H) \mid H\},$$

which equals $e_{\text{full}}(H)$ when $\pi(H) = 1$ and $1 - e_{\text{full}}(H)$ otherwise.

Further coarsening. For a nested chain $Z_{k+1} = \rho_k(Z_k)$ retaining the full-history boxes, conditioning $\mathbb{E}(Aw \mid Z_k) = 1$ on Z_{k+1} enlarges the feasible set. Thus $V_{Z_{k+1}} \le V_{Z_k}$, and exact normalization loss is nondecreasing along the chain.

B.1.4 Envelope containment and reference changes

For each positive-mass fine history d in cell c, write $e_d = e(d)$ and $e_c = P(A = 1 \mid Z = c)$. Define the directional factors and odds envelope by

$$K^+_c = \max_{d \subset c} \frac{\text{odds}(e_d)}{\text{odds}(e_c)}, \quad K^-_c = \max_{d \subset c} \frac{\text{odds}(e_c)}{\text{odds}(e_d)}, \qquad \frac{\text{odds}(e_c)}{\lambda K^-_c} \le \text{odds}(q) \le \lambda K^+_c \,\text{odds}(e_c). \tag{21}$$

The symmetric factor is $K_c = \max\{K_c^+, K_c^-\}$ and the global factor is $K_\phi = \max_c K_c$. Let $\underline{V}_\phi^{\text{directional}}$, $\underline{V}_\phi^{\text{symmetric}}$, and $\underline{V}_\phi^{\text{global}}$ be their lower endpoints under grouped normalization. For $S = e(H)$, write $\underline{V}_{(Z,S)}^{\text{tr}} = V_{(Z,S)}$ for the original full-history boxes with normalization through (Z, S). The complete ordering is

$$\underline{V}_\phi^{\text{global}} \le \underline{V}_\phi^{\text{symmetric}} \le \underline{V}_\phi^{\text{directional}} \le V_Z \le \underline{V}_{(Z,S)}^{\text{tr}} \le V_H. \tag{22}$$

For a treated fine atom i, aggregation to (z, r) takes the mass-weighted average of its sensitivity weights. It preserves the reward objective and converts grouped normalization into $\mathbb{E}(Aw \mid Z) = 1$. The logit decomposition in (21) places every fine box inside the directional inverse-propensity envelope

$$[e_{\text{hi}}(e_z, \lambda K_z^+)^{-1}, e_{\text{lo}}(e_z, \lambda K_z^-)^{-1}].$$

A weighted average stays inside this box. Expanding directional factors to $K_z = \max(K_z^+, K_z^-)$ and then $K_\phi = \max_z K_z$ proves the first three inequalities of (22). Conditioning (Z, S) normalization on Z proves $V_Z \le V_{(Z,S)}$; full-history normalization implies (Z, S) normalization and gives the last inequality. The additional loss from an envelope is distinct from exact normalization loss.

B.1.5 A RECENTERING WITNESS

Write $\text{expit}(t) = 1/(1+\exp(-t))$ for the logistic function. Let independent $z, u \sim \text{Bernoulli}(1/2)$ and a binary history feature x_c satisfy $P(x_c = 1 \mid u) = 0.5 + 0.40(2u - 1)$. The observed history is $H = (z, x_c)$, the interface keeps z, and the complete propensity and potential rewards are

$$q(z, x_c, u) = \text{expit}\{-0.2 + 1.4(x_c - 0.5) + 0.40(2u - 1)\},$$
$$R(1) = 0.45 + 0.50x_c + 0.20(2u - 1), \qquad R(0) = 0.55 - 0.10x_c.$$

Consistency, exchangeability given (H, u), and positivity hold. The target value is 0.700000 and $B = 0.704799$. The generating law's full-history radius is 2.036312; its radius after centering at $P(A = 1 \mid z)$ is 3.143708. At the full-history radius, $V_H = 0.699930$ and $V_Z = 0.636684$. Centering new boxes at the coarse propensity but reusing radius 2.036312 instead gives lower value 0.789785, above both the target and baseline. The properly calibrated coarse radius gives 0.692954. These calculations illustrate the effect of changing the sensitivity reference. For the recentered model, applying the sharp construction in Appendix B.1.2 to history Z gives its own sharp endpoint; that endpoint pertains to a different class.

B.2 FINITE-SAMPLE CERTIFICATION

B.2.1 SIMULTANEOUS ENVELOPES AND THE CORRECTED EXPECTATION

Fix versions of the conditional nuisance functions on the declared standard-Borel history domain. The envelopes, projected points, interface, and fitted threshold are jointly measurable in construction information and history. Their simultaneous coverage event $\mathcal{E}$ belongs to $\mathcal{F}$. Let $\mathcal{G} \in \mathcal{F}$ be the nonabstention event of the measurable feasibility rule. On $\mathcal{G}$, all intersected bands are nonempty, $p_h \in [L_{eh}, U_{eh}] \subset [\eta, 1 - \eta]$, and $\widehat{\mu}_h \in [L_{\mu h}, U_{\mu h}] \subset [0, 1]$ throughout $\mathcal{H}$. An empty feasible envelope forces abstention. The rule may also abstain conservatively; its assigned bound $-\infty$ always covers.

Corrected expectation in (8). On $\mathcal{G}$, put $\mu_{0h} = \mathbb{E}(R \mid H = h, A = 0)$ and $m_h(a) = Q_{a_h}(0) + \kappa(a_h)\mu_h$, and condition on $\mathcal{F}$. Independence of a fresh observation from $\mathcal{F}$ leaves its conditional action and reward laws unchanged, even when the fit and envelopes share data. Direct expansion of (6) gives

$$\mathbb{E}(\varphi_a \mid H = h, \mathcal{F}) = (1 - e_h)(m_h(a) - \mu_{0h}) + \frac{e_h - p_h}{p_h}\kappa(a_h)(\mu_h - \widehat{\mu}_h). \tag{23}$$

The first term is the conditional known-reference dual improvement. For $A = 0$, the binary direct-dual score is $a_h - R$. For $A = 1$, its values are $-(u_h - 1)a_h$ at $R = 0$ and $(\ell_h - 1)(1 - a_h)$ at $R = 1$. Averaging them yields $(1 - e_h)[(1 - \lambda)a_h + \{\lambda^{-1} + (\lambda - \lambda^{-1})a_h\}\mu_h - \mu_{0h}]$, the first term in (23).

The remainder is bilinear in the two true nuisance probabilities with the fit fixed. Its maximum over their rectangle is $\kappa(a_h)\beta_h$. On $\mathcal{E} \cap \mathcal{G}$ it upper-bounds the actual remainder for every admissible threshold, even if fitted from the same data as the outcome bands. Subtracting and averaging over H gives $\mathbb{E}(g_a \mid \mathcal{F}) = D_X(a) - B + \mathbb{E}_H[r_H(a) - \kappa(a_H)\beta_H]$. At the fitted threshold,

$$D_X(\widehat{a}) - B - \mathbb{E}(g_{\widehat{a}} \mid \mathcal{F}) = \mathbb{E}_H\big[\kappa(\widehat{a}_H)\beta_H - r_H(\widehat{a})\big]. \tag{24}$$

This proves the first inequality in (8). Weak duality and normalization coarsening give the remaining inequalities. Projection places a zero product inside the rectangle, hence its maximal product is nonnegative and the implemented nonnegative truncation leaves the exact maximum unchanged. Joint measurability and the bounded score enclosure below justify the conditional expectations and integration for general histories. □

Cellwise nuisance envelopes. For a finite declared support of size K, let $n_{e,h}$ count calibration records with history h and $k_{e,h}$ those with $A = 1$. Let $n_{\mu,h}$ count fitting records with $H = h, A = 1$ and $k_{\mu,h}$ those with $R = 1$. Conditional on the corresponding totals, $k_{e,h}$ and $k_{\mu,h}$ are binomial with parameters e_h and μ_h. Two-sided Clopper–Pearson intervals with coordinate errors α_e/K and α_μ/K cover each family by a union bound, conditionally on its totals and unconditionally. The intersection of these two family events is $\mathcal{E}$, with $P(\mathcal{E}) \geq 1 - \alpha_e - \alpha_\mu$; their independence is unnecessary. Intersect propensity intervals with $[\eta, 1-\eta]$ and outcome intervals with $[0, 1]$. For k successes in n trials, project $(k+1/2)/(n+1)$ into the intersected interval. Empty cells receive $[0, 1]$ before intersection. If any overlap intersection is empty, return $-\infty$; on $\mathcal{E}$, valid declared overlap makes all intersections nonempty. This finite check is measurable and gives the projected p_h and $\widehat{\mu}_h$ above. All nuisance estimates are computed from observed counts. The studies in Appendix D.2 use this construction and enumerate all $4K$ declared history/action/reward categories for support, including categories absent from calibration or fitting.

B.2.2 Structured-model nuisance envelopes

Model and design assumptions. For nuisance index $j \in \{e, \mu\}$, the d_j-dimensional feature maps $x_j : \mathcal{H} \to \mathbb{R}^{d_j}$ are fixed, measurable, and uniformly bounded in Euclidean norm by known L_j^x. With true coefficient vectors $\theta_j^\star$, the observed conditional means are correctly specified as $e_h = x_e(h)^\top \theta_e^\star$ and $\mu_h = x_\mu(h)^\top \theta_\mu^\star$, with known bounds $\|\theta_j^\star\|_2 \leq S_j$ and values in the declared ranges $C_e = [\eta, 1-\eta]$, $C_\mu = [0, 1]$ throughout $\mathcal{H}$. The features and norm bounds are specified before these samples are observed. Calibration and fitting records are i.i.d. from the observed law and are independent of final evaluation. The threshold interface Z has a separate role: it restricts normalization and thresholds, while x_j describes full-history nuisance variation, including variation within a Z cell. The regularization $\gamma > 0$ is distinct from sensitivity λ.

For ridge regularization $\gamma > 0$, let I denote the identity matrix and define design matrices V_j, coefficient radii B_j, and prediction half-widths $w_j(h)$ by

$$\begin{aligned} &V_e = \gamma I + \sum_{i \in \mathcal{D}_e} x_e(H_i) x_e(H_i)^\top, \quad V_\mu = \gamma I + \sum_{i \in \mathcal{D}_f} A_i x_\mu(H_i) x_\mu(H_i)^\top, \\ &B_j = \tfrac{1}{2}\sqrt{\log\det V_j - d_j \log\gamma + 2\log(1/\alpha_j)} + \sqrt{\gamma} S_j, \quad w_j(h) = B_j \sqrt{x_j(h)^\top V_j^{-1} x_j(h)}. \end{aligned} \tag{25}$$

With $\widehat{\theta}_j$ the ridge coefficient estimate, the predictions $q_j(h) = x_j(h)^\top \widehat{\theta}_j$ have envelopes $[q_j(h) - w_j(h), q_j(h) + w_j(h)]$ intersected with their probability ranges; project $q_j(h)$ into these intervals. The radii specialize Abbasi-Yadkori et al. (2011, Theorems 1–2) to Bernoulli noise with sub-Gaussian parameter $1/2$; the outcome design uses treated fitting records.

Corollary 2 (Structured nuisance protection)**.** *Under these fixed-feature model assumptions and Theorem 2's construction conditions, the structured envelopes give simultaneous coverage at least* $1 - \alpha_e - \alpha_\mu$ *and RAC satisfies* (10)*.*

The designs in (25) have right-hand sides and exact ridge fits

$$b_e = \sum_{i \in \mathcal{D}_e} x_e(H_i) A_i, \qquad b_\mu = \sum_{i \in \mathcal{D}_f} A_i x_\mu(H_i) R_i, \qquad \widehat{\theta}_j = V_j^{-1} b_j. \tag{26}$$

For finite histories the designs can be aggregated by cell using calibration totals and treatment successes for e, and treated fitting totals and reward successes for μ. Only treated fitting records enter the outcome design, which estimates the observed treated mean $\mu(h)$.

Coefficient confidence and the treated design. For propensity, condition on all calibration histories. Its independent Bernoulli residuals $A_i - e(H_i)$ have conditional moment generating function at most $\exp(t^2/8)$. For outcome, condition on all fitting histories and actions, reveal rewards in record order, and use predictor $A_i x_\mu(H_i)$ and response $A_i R_i$. The residual is $A_i\{R_i - \mu(H_i)\}$: it is zero when $A_i = 0$ and centered Bernoulli noise when $A_i = 1$, with the same bound $\exp(t^2/8)$. These conditionings produce predictable fixed designs and conditional sub-Gaussian parameter $1/2$. They are valid under i.i.d. sampling even if action assignment is confounded, because μ is the observed conditional mean given $H, A = 1$.

Let $\xi_{j,i}$ denote these predictors and $\epsilon_{j,i}$ their centered responses, and write $s_j = \sum_i \xi_{j,i}\epsilon_{j,i}$. For a positive-definite matrix M, write $\|v\|_M = (v^\top M v)^{1/2}$; $M \succeq N$ means $M - N$ is positive semidefinite. The self-normalized inequality of Abbasi-Yadkori et al. (2011, Theorem 1), with initial matrix γI, yields, with probability at least $1 - \alpha_j$ conditional on the corresponding design,

$$\|s_j\|_{V_j^{-1}} \le \tfrac{1}{2}\sqrt{\log\det V_j - d_j \log\gamma + 2\log(1/\alpha_j)}. \tag{27}$$

The identity $\widehat{\theta}_j - \theta_j^\star = V_j^{-1}(s_j - \gamma\theta_j^\star)$ and $V_j \succeq \gamma I$ give $\|\gamma\theta_j^\star\|_{V_j^{-1}} \le \sqrt{\gamma} S_j$. Consequently the coefficient event

$$\mathcal{E}_j = \{\|\widehat{\theta}_j - \theta_j^\star\|_{V_j} \le B_j\} \quad \text{satisfies} \quad P(\mathcal{E}_j) \ge 1 - \alpha_j, \tag{28}$$

with exactly B_j in (25), as in Abbasi-Yadkori et al. (2011, Theorem 2). Here $N_T = \sum_{i\in\mathcal{D}_f} A_i$ is random, and V_μ contains only its N_T treated contributions. Conditioning on the full fitting design and then integrating over its distribution establishes (28). An empty treated design has $V_\mu = \gamma I$, $b_\mu = 0$ and a valid broad band determined by the norm bound.

Simultaneous prediction coverage and physical intersection. On $\mathcal{E}_j$, Cauchy–Schwarz gives $|x_j(h)^\top(\widehat{\theta}_j - \theta_j^\star)| \le w_j(h)$ for *every* $h \in \mathcal{H}$. Thus one coefficient event supplies simultaneous full-history coverage, irrespective of the number of prediction locations. For $q_j(h) = x_j(h)^\top\widehat{\theta}_j$, form

$$I_j(h) = [q_j(h) - w_j(h), q_j(h) + w_j(h)] \cap C_j, \qquad t_j(h) = \mathrm{proj}_{I_j(h)}\, q_j(h),$$

with $t_e = p$ and $t_\mu = \widehat{\mu}$. Physical intersection preserves truth on $\mathcal{E}_j$. These scalar intervals conservatively enclose predictions from the coefficient ellipsoid intersected with all-history physical constraints. Projection can produce nonlinear fits; the RAC bias identity applies to these measurable projected functions. Empty scalar intersections, failed numerical construction, or an unavailable valid support enclosure force abstention. For general histories the feasibility rule is measurable as required by Theorem 2; finite declared supports permit direct checking. Bounded features and $V_j \succeq \gamma I$ give finite measurable prediction widths. The universal enclosure in Appendix B.2.4 supplies finite score support on any nonabstaining construction.

Proof of Corollary 2. The coefficient events belong to $\mathcal{F}$. Their intersection $\mathcal{E} = \mathcal{E}_e \cap \mathcal{E}_\mu$ has probability at least $1 - \alpha_e - \alpha_\mu$ by a union bound; their independence is unnecessary. On $\mathcal{E}$ the measurable envelopes contain both true nuisance functions throughout the history domain. Projection, admissible fitting and the frozen support satisfy Theorem 2's remaining assumptions. The fit may use the same observations as the outcome envelope. The corrected expectation still satisfies (8), and independent betting has conditional coverage $1 - \alpha_s$. Integrating gives (10). Numerical envelopes obey the same argument when their outward bounds enclose these analytic prediction intervals, as specified below. □

Projection distances, correction and net deficit. For any centered interval $[q - W, q + W]$ intersected with a scalar range C, let $[l, u]$ be its nonempty intersection and $t = \mathrm{proj}_{[l,u]}\, q$. Then

$$\delta = \max\{t - l, u - t\} \le W. \tag{29}$$

If $q \in C$, $t = q$. If q lies below C, $t = l \geq q$ and $u - t \leq q + W - t \leq W$; the case above C is symmetric. Let $\delta_e(h)$ and $\delta_\mu(h)$ be these endpoint distances. On nuisance coverage and nonabstention,

$$|r_h(a)| \leq \frac{\kappa(a_h)}{p_h}\delta_e(h)\delta_\mu(h) \leq \frac{\kappa(a_h)}{p_h}w_e(h)w_\mu(h),$$
$$0 \leq \kappa(a_h)\beta_h \leq \frac{\kappa(a_h)}{p_h}\delta_e(h)\delta_\mu(h) \leq \frac{\kappa(a_h)}{p_h}w_e(h)w_\mu(h). \tag{30}$$

The correction upper bound holds on any nonabstaining construction; the bias bound uses coverage. Rectangle maximization also gives $r_h(a) \leq \kappa(a_h)\beta_h$ on coverage. Hence

$$0 \leq \kappa(a_h)\beta_h - r_h(a) \leq 2\frac{\kappa(a_h)}{p_h}w_e(h)w_\mu(h).$$

For centered structured envelopes, integration yields, on coverage and nonabstention,

$$0 \leq D_X(a) - B - \mathbb{E}(g_a \mid \mathcal{F}) \leq 2\mathbb{E}_H\left[\frac{\kappa(a_H)}{p_H}w_e(H)w_\mu(H)\right],$$
$$M_H - \mathbb{E}(g_a \mid \mathcal{F}) \leq L_\phi + J_Z(a) + 2\mathbb{E}_H\left[\frac{\kappa(a_H)}{p_H}w_e(H)w_\mu(H)\right], \qquad X = Z. \tag{31}$$

Integrating the exact conditional identity proves the first bound in (31). For $X = Z$, Theorem 1 and weak duality give $M_H - \{D_Z(a) - B\} = L_\phi + J_Z(a)$; adding the net deficit proves its second bound. At $\lambda = 1$ the structural loss is zero, and the same identity applies. The net nuisance term is the applied correction minus signed bias. Both statements bound the protected score's population deficit, before the evaluation gap in (11). The latter gap is nonnegative on betting's conditional score-mean coverage event.

Numerical outer envelopes. The analytic radius above is centered at the exact ridge solve. Intervals centered at computed predictions include outward allowances for numerical error. Let W_j be a computed symmetric inverse and $\theta_{c,j}$ a computed center. Suppose $\tau_j \geq \|I - V_jW_j\|_2$, $\varrho_j \geq \|b_j - V_j\theta_{c,j}\|_2$, $\ell_j^+ \geq \log\det V_j$, and $\epsilon_{q,j}(h)$ bounds the error in computing $x_j(h)^\top\theta_{c,j}$; $\epsilon_{\mathrm{pad},j}(h) \geq 0$ is additional outward padding. In exact arithmetic,

$$\|W_j - V_j^{-1}\|_2 \leq \tau_j/\gamma, \qquad \|\theta_{c,j} - \widehat{\theta}_j\|_{V_j} \leq \varrho_j/\sqrt{\gamma},$$
$$v_j^+(h) = \max\{x_j(h)^\top W_j x_j(h) + (\tau_j/\gamma)\|x_j(h)\|_2^2, 0\},$$
$$\widetilde{B}_j = \tfrac{1}{2}\sqrt{\ell_j^+ - d_j\log\gamma + 2\log(1/\alpha_j)} + \sqrt{\gamma}S_j + \varrho_j/\sqrt{\gamma},$$
$$\widetilde{w}_j(h) = \widetilde{B}_j\sqrt{v_j^+(h)} + \epsilon_{q,j}(h) + \epsilon_{\mathrm{pad},j}(h), \tag{32}$$

A leverage computation also requires an outward bound on its dot-product error. With these allowances, intervals centered at the computed prediction with half-width $\widetilde{w}_j$ enclose the analytic prediction intervals. Intersect with C_j and project that computed prediction. The same projection-distance argument proves (30) and (31) with $\widetilde{w}_j$ throughout.

The numerical-enclosure statement assumes valid outward bounds for all computed residuals, leverage and determinant quantities, including error in assembling the sufficient statistics.

Dimension, regularization and treated-sample dependence. Set $n_e = |\mathcal{D}_e|$ and $n_\mu = N_T$. The trace bound and arithmetic–geometric mean inequality give

$$B_j \leq \overline{B}_j := \tfrac{1}{2}\sqrt{d_j\log\left(1 + \frac{n_j(L_j^x)^2}{\gamma d_j}\right) + 2\log(1/\alpha_j)} + \sqrt{\gamma}S_j. \tag{33}$$

The leverage $x_j(h)^\top V_j^{-1}x_j(h)$ remains essential. For example, under the explicit design condition $G_j := V_j - \gamma I \succeq (n_j\nu_j/d_j)I$ with stated $\nu_j > 0$ and $n_j > 0$,

$$w_j(h) \leq \frac{\overline{B}_j L_j^x}{\sqrt{\gamma + n_j\nu_j/d_j}}, \qquad D_X(a) - B - \mathbb{E}(g_a \mid \mathcal{F}) \leq \frac{2\lambda}{\eta}\prod_{j\in\{e,\mu\}}\frac{\overline{B}_j L_j^x}{\sqrt{\gamma + n_j\nu_j/d_j}}. \tag{34}$$

The deficit bound holds on nuisance coverage and nonabstention. Its design constants ν_j describe the actual matrices and may depend on dimension and sampling probabilities. For an exactly balanced subset-Walsh design, $\|x_j(h)\|_2^2 = d_j$ and $G_j = n_j I$, so $w_j(h) \le \overline{B}_j\sqrt{d_j/(\gamma + n_j)}$; both the radius's d_j dependence and the leverage's d_j dependence remain. An unobserved feature direction can instead retain substantial width despite many records. Shared structure allows histories to contribute to common coefficients and can accommodate finer resolution when dimensions and leverage stay controlled. It permits unequal nuisances within normalization cells while using shared confidence regions. Coverage requires correct fixed-feature models and the actual treated design. Learned or misspecified features require additional calibration.

B.2.3 Concavity of the executed fitting criterion

Variance-aware fitting. At $\lambda = 2$, $\mu_h = 2/3$, thresholds 0 and 1 are optimal, but the treated-residual variance component at exact nuisances is 16 times larger at 1. With $\bar{g}_f(a) = n_f^{-1}\sum_i g_a(O_i)$ and $\widehat{\mathrm{Var}}_f(g_a) = n_f^{-1}\sum_i\{g_a(O_i) - \bar{g}_f(a)\}^2$, the inner value, when its unconstrained weights are nonnegative, is

$$\bar{g}_f(a) - \sqrt{\rho\,\widehat{\mathrm{Var}}_f(g_a)} = \bar{g}_f(a) - \sqrt{\frac{2\log(1/\alpha_s)}{n_s}\widehat{\mathrm{Var}}_f(g_a)}, \qquad \alpha_s = 0.025.$$

Concavity. For frozen nuisances, $Q_a(0) = (1-\lambda)a$, $Q_a(1) = \lambda^{-1} + (1-\lambda^{-1})a$, and $\widehat{m}_h(a) = (1-\lambda)a_h + \kappa(a_h)\widehat{\mu}_h$ are affine in a_h. The coefficient β_h in (7) depends only on the frozen nuisance rectangle, so the corrected score $g_a(h, A, R)$ is affine in the threshold function. Therefore each weighted fitting mean is affine; its pointwise infimum over the fixed convex set $\mathcal{W}_\rho$ is concave on a convex admissible threshold class. The restriction $a_h = a\{\phi(h)\}$ is linear. This verifies concavity for the corrected score itself.

In the finite-support implementation, repeated fitting categories can be compressed without changing the inner problem. If present category k has count n_k, frequency $f_k = n_k/n_f > 0$, and score $s_k(a)$, equal record weights within a category minimize chi-square distance for any fixed category mass. Thus the inner problem, with q_k denoting category probability mass, is exactly

$$\min_{q_k \ge 0}\sum_k q_k s_k(a), \qquad \sum_k q_k = 1, \quad \sum_k \frac{(q_k - f_k)^2}{f_k} \le \rho. \tag{35}$$

Absent categories receive zero fitting weight; they remain part of the frozen evaluation support. With $\bar{s} = \sum_k f_k s_k$ and $v = \sum_k f_k(s_k - \bar{s})^2 > 0$, the unconstrained density is $q_k/f_k = 1 - \sqrt{\rho}(s_k - \bar{s})/\sqrt{v}$. Only if these values are all nonnegative does the inner optimum equal $\bar{s} - \sqrt{\rho v}$. Zero radius or zero variance returns $\bar{s}$.

When nonnegativity binds, an active set I has mass $F_I = \sum_{k\in I} f_k$, mean $\bar{s}_I = \sum_{k\in I} f_k s_k/F_I$, and variation $V_I = \sum_{k\in I} f_k(s_k - \bar{s}_I)^2$. The KKT density on I is

$$\frac{q_k}{f_k} = \frac{1}{F_I} - \sqrt{\frac{\rho - (F_I^{-1} - 1)}{V_I}}(s_k - \bar{s}_I), \qquad k \in I,$$

and zero off I. The implementation removes the largest-score category until all active densities are nonnegative. If $\rho \ge f_{\min}^{-1} - 1$, where $f_{\min}$ is total mass of the minimum score, all mass can be placed on its minimizers. This active-set solution enforces nonnegativity, total mass, and radius feasibility.

For two or four threshold cells, the fitter uses multistart SLSQP and retains the highest objective among feasible outputs and predefined cube corners and the midpoint. Ties favor lexicographically smaller cell thresholds.

B.2.4 Complete support, predictable bets, and inversion

The theorem requires finite $\mathcal{F}$-measurable $b_- \le b_+$ enclosing every declared score on $\mathcal{G}$, whether or not nuisance coverage holds. It does not require attained extrema. Such an enclosure exists under the stated binary-reward and overlap assumptions: $Q_a(0) \in [1-\lambda, 0]$, $Q_a(1) \in [\lambda^{-1}, 1]$, and $\widehat{m}_h(a)$

is their convex combination. Consequently $|\varphi_a| \le \lambda/\eta$. The projected band widths are at most one, so $0 \le \kappa(a_h)\beta_h \le \lambda/\eta$ and

$$-2\lambda/\eta \le g_a(h, A, R) \le \lambda/\eta.$$

This universal enclosure permits general histories. The experiments instead use the tighter minimum and maximum over all $4K$ declared categories, not just observed categories. These catalogue extrema need not equal the essential extrema under the evaluation law. Let $d = b_+ - b_-$ be the support width. For $d > 0$ put $x_i = (g_{\widehat{a}}(O_i) - b_-)/d$, the normalized evaluation score. For candidate normalized mean m, let $K_t(m)$ be capital after t observations, $\zeta_i(m)$ the predictable bet, and $\widehat{v}_{i-1}$ the running variance estimate. Then

$$K_t(m) = \prod_{i=1}^{t}[1+\zeta_i(m)(x_i-m)], \qquad \zeta_i(m) = \min\left\{\sqrt{\frac{2\log(1/\alpha_s)}{n_s\widehat{v}_{i-1}}}, \frac{1}{2m}\right\}, \quad m \in [0,1], \tag{36}$$

with $1/(2m) = \infty$ at $m = 0$. With n_s fixed in advance, the executed recursion uses running mean estimate $\widehat{q}_t$ and uncapped bet $\widetilde{\zeta}_i$:

$$\widehat{q}_t = \frac{1/2 + \sum_{i=1}^{t} x_i}{t+1}, \qquad \widehat{v}_t = \frac{1/4 + \sum_{i=1}^{t}(x_i - \widehat{q}_i)^2}{t+1}, \qquad \widehat{v}_0 = 1/4,$$

$$\widetilde{\zeta}_i = \sqrt{\frac{2\log(1/\alpha_s)}{n_s\widehat{v}_{i-1}}}, \qquad \zeta_i(m) = \min\{\widetilde{\zeta}_i, 1/(2m)\}. \tag{37}$$

Although $\widehat{q}_i$ uses the current observation in the variance update, the bet uses $\widehat{v}_{i-1}$ and is predictable. Every capital factor is at least $1/2$ for $m > 0$; at zero the factors are at least one. Capital is computed by cumulative $\log 1p$ values, and the acceptance set is

$$C_n = \{m \in [0,1] : \max_{1 \le t \le n_s} \log K_t(m) < \log(1/\alpha_s)\}. \tag{38}$$

For each observation, a factor is $1 + \widetilde{\zeta}_i(x_i - m)$ in the uncapped region and $1/2 + x_i/(2m)$ in the capped region. Both decrease continuously in m and join at the cap boundary; products and their running maximum are nonincreasing. Thus C_n is an upper interval and $m = 1$ is accepted. If zero is accepted, return normalized bound zero. Otherwise bisect from rejected zero to accepted one, retaining the rejected endpoint. Returning the rejected endpoint is conservative even when the exact infimum is excluded from the open acceptance set. A failed inversion forces abstention. For $d = 0$, return the constant score.

Proof of Theorem 2. On $\mathcal{G}$, condition on $\mathcal{F}$. All score values and valid support endpoints are frozen independently of the final i.i.d. sample. For the true normalized score mean $m_0 = \mathbb{E}(x_i \mid \mathcal{F})$, $\mathbb{E}[x_i - m_0 \mid \mathcal{F}, x_{1:i-1}] = 0$. Hence $K_t(m_0)$ is a nonnegative martingale with initial capital one. Ville's inequality gives $P(m_0 \in C_n \mid \mathcal{F}) \ge 1 - \alpha_s$. The conservative inversion therefore supplies $\widehat{L} \le \mathbb{E}(g_{\widehat{a}} \mid \mathcal{F})$ with at least that conditional probability. This is the one-sided positive-capital construction of Waudby-Smith & Ramdas (2024). If $d = 0$, the score is constant and its returned value equals its conditional mean.

Define $\mathcal{E}_s$ to hold on abstention and, on $\mathcal{G}$, whenever $\widehat{L} \le \mathbb{E}(g_{\widehat{a}} \mid \mathcal{F})$. Then $P(\mathcal{E}_s \mid \mathcal{F}) \ge 1 - \alpha_s$ almost surely. The corrected expectation proof gives (8) on $\mathcal{E} \cap \mathcal{G}$, at the data-dependent but frozen threshold. On $\mathcal{G}^c$, $\widehat{L} = -\infty$ always covers. Thus $\mathcal{E} \cap \mathcal{E}_s \subseteq \{\widehat{L} \le M_H\}$ and

$$P(\mathcal{E} \cap \mathcal{E}_s) = \mathbb{E}[\mathbf{1}_{\mathcal{E}} P(\mathcal{E}_s \mid \mathcal{F})] \ge (1 - \alpha_s)P(\mathcal{E}) \ge (1 - \alpha_s)(1 - \alpha_e - \alpha_\mu).$$

On this intersection, a nonabstaining LCB is below the fitted dual, the interface endpoint, and M_H. The nuisance family events need not be independent, and thresholds may depend on all fitting and calibration data in $\mathcal{F}$. Independence is required between that entire frozen construction and evaluation. The argument allows nonunique quantiles and general standard-Borel histories. □

When the simulator stores multinomial category counts, an independent uniform shuffle orders the complete evaluation multiset and recovers the i.i.d. sequence law. The guarantee conditions on calibration and fitting and applies to the resulting i.i.d. evaluation sequence.

B.2.5 DIRECT PROTECTED WEAK DUAL

The comparator uses simultaneous propensity envelopes with family error 0.025. In the four-history cellwise study, each interval has coordinate error $0.025/4$. It intersects overlap and expands inverse-weight boxes to $[\widetilde{\ell}_h, \widetilde{u}_h] = [e_{\mathrm{hi}}(U_{eh}, \lambda)^{-1}, e_{\mathrm{lo}}(L_{eh}, \lambda)^{-1}]$. Its score is

$$g_a^{\mathrm{dir}} = a_H + A \min_{v \in [\widetilde{\ell}_H, \widetilde{u}_H]} v(R - a_H) - R. \tag{39}$$

On reference coverage, monotonicity gives $[\ell_h, u_h] \subset [\widetilde{\ell}_h, \widetilde{u}_h]$ for all h. For binary rewards its score is affine:

$$g_a^{\mathrm{dir}} = A\widetilde{\ell}_H R - R + \{1 - A[\widetilde{\ell}_H R + \widetilde{u}_H(1 - R)]\}a_H.$$

It independently fits the chi-square-weight objective using its own score, the same ρ, and the prescribed interface, then freezes all $4K$ category scores. For any true-reference feasible sensitivity weight, the expanded box minimum is no larger than its dual contribution. Taking expectations, the normalization equation cancels $a_H(1 - Aw)$, so $\mathbb{E}(g_a^{\mathrm{dir}} \mid \mathcal{F}) \le M_X \le M_H$. The same betting proof gives complete pointwise coverage $(1 - 0.025)^2 = 0.950625$. The comparator uses the same data partitions as RAC; its fitting sample estimates thresholds and no treated-outcome nuisance.

B.2.6 SIMULTANEOUS-SELECTION EXTENSION

For a finite family of scores frozen independently of evaluation, sharing a simultaneous nuisance event, assign score errors δ_j with $\sum_j \delta_j \le \alpha_s$ and use each δ_j in its capital process. A conditional union bound and the same nuisance-event integration give

$$P\{\forall j : \widehat{L}_j \le M_{X_j} \le M_H\} \ge (1 - \alpha_e - \alpha_\mu)(1 - \alpha_s).$$

Any selected family member is protected on this event. The thresholds may be fitted separately, but no score range or candidate can be added using evaluation outcomes. The executed comparisons instead use the full pointwise score allocation for each specified procedure and report their results separately; their evaluation-selected maximum has no asserted guarantee.

C EXPERIMENTAL SETTINGS

C.1 EVALUATION PROTOCOL AND UNCERTAINTY REPORTING

The known-law studies evaluate population endpoints by exact enumeration; the finite-data studies evaluate fitted lower confidence bounds. Coverage is $\widehat{L} \le M_H$, and certification is $\widehat{L} > 0$. A positive bound at the robust null $M_H = 0$ is a false certificate. Matched protected procedures use independent calibration, fitting and evaluation samples with total budget N split as $N/2$, $N/4$ and $N/4$. RAC allocates $\alpha_e = \alpha_\mu = 0.0125$ and $\alpha_s = 0.025$; the direct comparator allocates $\alpha_e = \alpha_s = 0.025$. Both have the same complete-procedure coverage lower bound 0.950625. Published comparators use their stated inference constructions and all $N/2$ observed reward records. Across the reported finite-data studies, no procedure abstained or had a numerical failure. The changing-threshold study also recorded no optimizer fallback.

Datasets are the independent units within finite-data conditions; procedures and interfaces are paired measurements on those datasets. Rates use 95% Wilson intervals. Mean lower bounds and within-dataset differences have 95% Monte Carlo intervals computed over independent datasets. Aggregates equally weight condition means and retain the pairing within each condition. Intervals are descriptive and are not multiplicity-adjusted, except for the explicitly adjusted published- comparator undercoverage tests.

For the six-condition threshold ablation, with $n_c = 1000$ in each condition, let c index conditions and $\bar{d}_c, s_c^2$ be the mean and sample variance of paired LCB differences. The equal-weight aggregate is $\bar{d} = \sum_c \bar{d}_c/6$, with variance estimate $v = \sum_c v_c$, $v_c = s_c^2/(36n_c)$, and 95% interval $\bar{d} \pm t_{0.975,\nu}\sqrt{v}$, with $t_{0.975,\nu}$ the 0.975 quantile of Student's t distribution and Welch–Satterthwaite degrees of freedom

$$\nu = \frac{v^2}{\sum_c \dfrac{v_c^2}{n_c - 1}}.$$

The threshold-orientation aggregate uses the same calculation with eight strata. Representation replicates are the independent training units in the HC–MNIST study; interface/effect evaluations within a replicate are correlated.

C.2 Synthetic population laws

C.2.1 Matched reward-alignment illustration

The four histories have equal mass, propensities $(0.2, 0.8, 0.2, 0.8)$, treated Bernoulli means $(\omega, \omega, 1-\omega, 1-\omega)$, and zero control rewards. Here ω is reward noise. Action and potential rewards are independent given history. Outcome pairs merge $(0,1)$ and $(2,3)$; crossed pairs merge $(0,3)$ and $(1,2)$; propensity pairs merge $(0,2)$ and $(1,3)$. The grid $\omega \in \{0, 0.1, 0.2, 1/3, 0.4, 0.5\}$ and $\lambda \in \{1, 1.5, 2, 3\}$ gives 72 deterministic comparisons.

C.2.2 Synthetic hidden-confounding utility laws

The synthetic utility grid uses $U \sim \mathrm{Unif}\{0,1,2\}$, $H = (X, Y)$ with binary coordinates X, Y, and $Y \sim \mathrm{Bernoulli}(1/2)$ independently of X given U. With $\mathrm{logit}(p) = \log\{p/(1-p)\}$, $P(X = 1 \mid U) = \mathrm{expit}\{\mathrm{logit}(0.95)(U-1)\}$. For confounding strength $\alpha_U \in \{0, 0.15, 0.30, 0.60\}$, the complete propensity and treated reward mean are

$$q(X,U) = \mathrm{expit}\{-0.20 + 2(2X-1) + \alpha_U(U-1)\}, \qquad m_1(Y,U) = 0.42 + 0.40Y + 0.16(U-1).$$

Controls have mean $m_0 = [\mathbb{E}\{(1-q)m_1\} - \delta]/\mathbb{E}(1-q)$ for $\delta \in \{-0.02, 0, 0.02, 0.05, 0.10, 0.15, 0.20, 0.25\}$, giving true improvement δ. Every mean lies in $[0.0545, 0.98]$ without clipping. The common evaluated radius is the smallest full-history odds factor of each law, approximately $1.000, 1.268, 1.599, 2.496$ across the four confounding levels.

C.3 HC–MNIST representation study

HC–MNIST was introduced by Jesson et al. (2021). Our experiment uses the canonical 60,000/10,000 MNIST train/test split (LeCun et al., 1998) and the assignment construction used by Melnychuk et al. (2024). The 10,000 test images form the fixed exact-population evaluation. The complete image and deterministic digit metadata define H; the synthetic $U \sim \mathrm{Bernoulli}(1/2)$ remains hidden. The image score $\zeta(h)$ is constructed within each digit $y \in \{0, \ldots, 9\}$: mean standardized pixel intensity is standardized by the digit-y training mean and sample standard deviation. The resulting score is clipped to $[-1.4, 1.4]$ and affinely mapped from its split-specific observed range to $[-2 + 0.4y, -2 + 0.4(y+1)]$. For assignment parameter $G \in \{\exp(0.25), \exp(0.50), \exp(1.00)\}$, set $p(h) = \mathrm{expit}\{0.75\zeta(h) + 0.5\}$. The complete propensity is

$$q_G(h,0) = \frac{p(h)}{G + (1-G)p(h)}, \qquad q_G(h,1) = \frac{Gp(h)}{1 + (G-1)p(h)}.$$

Its full-history reference and smallest covering radius are

$$e_{\mathrm{full}}(h) = \tfrac{1}{2}q_G(h,0) + \tfrac{1}{2}q_G(h,1), \qquad \lambda_{\mathrm{true}}(G) = \max_{h,u} \max\left\{ \frac{\mathrm{odds}\, q_G(h,u)}{\mathrm{odds}\, e_{\mathrm{full}}(h)}, \frac{\mathrm{odds}\, e_{\mathrm{full}}(h)}{\mathrm{odds}\, q_G(h,u)} \right\}.$$

The always-treat target is measurable under every representation. The baseline treated mean is

$$m_1(h,u) = 0.10 + 0.80\ \mathrm{expit}\left(\frac{\zeta(h) + 1 - 2\sin\{2\zeta(h)\} - (4u-2)\{1 + 0.5\zeta(h)\}}{4} \right).$$

For each requested improvement Δ_P, the constant control mean is

$$m_0(\Delta_P) = \frac{\mathbb{E}_{H,U}[\{1 - q_G(H,U)\}m_1(H,U)] - \Delta_P}{\mathbb{E}_H\{1 - e_{\mathrm{full}}(H)\}}.$$

Bernoulli potential rewards use means $m_1(H,U)$ and $m_0(\Delta_P)$. This baseline has hidden confounding because U affects both assignment and treated reward. It is the outcome law used for representation training and the population certificate-preservation grid, whose effects are $\Delta_P \in \{-0.02, 0, 0.03, 0.06, 0.10, 0.15\}$.

Each representation replicate supplies one shared logged training sample. Five fixed convolutional objectives are crossed with bottleneck $d \in \{8, 32, 128\}$ and requested k-means normalization-interface count $C \in \{8, 32, 128\}$. For objective m, $f_{m,d}$ is the fitted d-dimensional encoder and k_C its clustering map; the evaluated representation is $\phi_{m,d,C} = k_C \circ f_{m,d}$. The study uses 20 independent representation replicates with the same architecture, objectives and evaluation grids.

Representation training and clustering. Table 5 specifies the architecture, optimization, and clustering shared by the five objectives. Here θ denotes network parameters and $\mathrm{ReLU}(x) = \max(0, x)$.

Table 5: HC–MNIST representation-training specification.

Component	Implemented specification
Images	Canonical MNIST 28×28 grayscale pixels scaled to $[0, 1]$, then standardized by mean 0.1307 and s.d. 0.3081.
Split	For each assignment mechanism and representation replicate, a random split of the 60,000 training images gives 48,000 training and 12,000 validation images.
Encoder f_θ	5×5 convolution $1 \to 32$, ReLU, 2×2 max pooling; 5×5 convolution $32 \to 64$, ReLU, 2×2 max pooling; flatten; linear $1024 \to 256$, ReLU; linear $256 \to d$ for $d \in \{8, 32, 128\}$.
Heads	Two action-specific reward heads, each $d \to 128 \to 1$ with an intermediate ReLU. InvTARNet adds a decoder $d \to 256 \to 784$ with ReLU; propensity-aware training adds an action head $d \to 128 \to 1$ with ReLU.
Optimization	Adam (Kingma & Ba, 2015), learning rate 10^{-3}, weight decay 10^{-4}, batch size 256, 50 epochs.
Checkpoint	The terminal 50-epoch model is retained.
Grouping	K-means++ (Arthur & Vassilvitskii, 2007) with 20 initializations, fit to the 12,000 validation embeddings before assigning the 10,000 test embeddings.

Let $\mathcal{L}_{\mathrm{F}}$ be binary cross-entropy for the factual outcome logit from the appropriate reward head. Let MMD denote maximum mean discrepancy (Gretton et al., 2012), $\widehat{W}$ the entropic optimal-transport (OT) loss (Cuturi, 2013), $\mathcal{L}_{\mathrm{rec}}$ the reconstruction loss, and $\mathcal{L}_A$ the action-prediction loss. TARNet and CFR adapt Shalit et al. (2017) to the binary reward. The five implemented objectives are TARNet: $\mathcal{L}_{\mathrm{F}}$; CFR–MMD: $\mathcal{L}_{\mathrm{F}} + 0.1\ \mathrm{MMD}^2$; CFR–Wasserstein: $\mathcal{L}_{\mathrm{F}} + \widehat{W}$; InvTARNet: $\mathcal{L}_{\mathrm{F}} + \mathcal{L}_{\mathrm{rec}}$; and propensity-aware: $\mathcal{L}_{\mathrm{F}} + \mathcal{L}_A$. The MMD averages Gaussian kernels with bandwidths $0.5, 1, 2$. The OT loss uses squared Euclidean costs divided by d and entropic regularization $\sigma_{\mathrm{OT}} = 0.1$. The reconstruction term $\mathcal{L}_{\mathrm{rec}}$ is pixelwise binary cross-entropy between the decoder logits and the original unstandardized $[0, 1]$ image; the action term $\mathcal{L}_A$ is binary cross-entropy between the propensity-head logit and observed A. All auxiliary coefficients not shown explicitly equal one. Training actions and Bernoulli rewards use the corresponding HC–MNIST assignment mechanism and a fixed training improvement 0.10.

C.4 Four-history finite-data study

Let $H \in \{0, 1, 2, 3\}$ and define the two-cell coarse interface by $\phi(0) = \phi(1) = 0$ and $\phi(2) = \phi(3) = 1$. For history masses $P_h = 1/4$, propensities $e = (0.3, 0.6, 0.3, 0.6)$, and treated means $\mu = (2/3 - \epsilon, 2/3 + \epsilon, 0.4, 0.5)$, let $U \in \{-1, 1\}$ have equal probabilities. Conditional on $(H = h, U = u)$, action probability is $e_h + 0.08u$, treated reward mean is $\mu_h - 0.008/e_h + 0.1u$, and controls have constant reward mean μ_0, using the binary lower sensitivity transform g_λ:

$$\mu_0 = \frac{\sum_h P_h(1 - e_h) g_2(\mu_h) - M_H}{\sum_h P_h(1 - e_h)}, \qquad g_\lambda(t) = \max\{t/\lambda, 1 - \lambda + \lambda t\}.$$

Potential Bernoulli rewards are independent of assignment given (H, U). Enumerating U yields exactly the declared observed e and treated μ, positive assignment–reward covariance 0.008, and latent odds factor at most 1.519481, below 2. All probabilities are feasible without clipping; the smallest declared (H, A, R) probability is 0.018485.

Write $v_h = P_h(1 - e_h)$. The known-law robust margin is $\sum_h v_h\{g_2(\mu_h) - \mu_0\} = M_H$. With success-probability cutoff t_λ, the binary threshold regret is

$$T(a) = \sum_h v_h(\lambda - \lambda^{-1}) \left[(\mu_h - t_\lambda)_+ - a_h(\mu_h - t_\lambda)\right], \qquad t_\lambda = \frac{\lambda}{1+\lambda}. \tag{40}$$

Only histories 0 and 1 straddle the cutoff, and their optimal coarse threshold is zero when $\epsilon > 0$. Thus $L_\phi = (2 - 1/2)\epsilon \min(v_0, v_1) = 0.15\epsilon$. At zero, the entire interval $[0, 1]$ is optimal in this cell, explaining zero regret for its nonunique fitted thresholds. At exact nuisances, the treated-residual variance in (6) is $(1-e_h)^2\kappa(a_h)^2\mu_h(1-\mu_h)/e_h$ conditional on $H = h$. Since $\kappa(1)/\kappa(0) = \lambda^2$, this component is $\lambda^4 = 16$ times larger at $a = 1$ than at $a = 0$ at the cutoff, although both are optimal.

The cutoff-family RAC results use the cellwise calibration in Appendix B.2.1. Both interfaces declare overlap $\eta = 0.01$. The study crosses $\epsilon \in \{0, 0.01, 0.10\}$, $M_H \in \{-0.02, 0, 0.05\}$ and both budgets at $\lambda = 2$, with 1,000 datasets per condition. It uses independent randomization for calibration, fitting, evaluation, evaluation order, and bootstrap. The budget configurations are $(n_e, n_f, n_s) = (5000, 2500, 2500)$ and $(20000, 10000, 10000)$. Each RAC setting fits independently using the rectangular score and $\rho = 2\log 40/n_s$. Every score uses all 16 declared categories for its support. Each procedure is evaluated pointwise under the coverage allocation in Appendix C.1.

C.5 Threshold-choice studies

C.5.1 Threshold-choice ablations in the cutoff family

The fitting ablation changes only the radius in (9), from $2\log 40/n_s$ to $\rho = 0$ (empirical-mean fitting). It uses the same 18,000 datasets as the four-history study, with paired calibration, fitting and evaluation samples and the same evaluation order. Both interfaces retain the nuisance construction, threshold domain, tie convention and error allocations of Appendix C.4. Each rule recomputes its threshold-dependent correction and complete 16-category support before betting evaluation. Empirical-mean fits are checked against the exact linear box-program solutions. A matched fixed-zero ablation evaluates $a_h = 0$ on the same datasets, recomputing its correction and support. Theorem 2 protects each specified threshold rule under the same allocation.

C.5.2 Changing population-optimal thresholds

This component study retains the four-history latent construction of Appendix C.4, propensities, coarse map and splits. The original orientation has treated means $(2/3 - \epsilon, 2/3 + \epsilon, 0.4, 0.5)$; the reversed orientation swaps the first two means. Control means are recalibrated to $M_H = 0$ or 0.05 under the same latent assignment and reward shifts. The first coarse cell's population-optimal threshold changes from zero to one, while the second remains zero. Writing $a = (a_1, a_2)$ for its two cell thresholds,

$$D_Z(1,0) - D_Z(0,0) = \begin{cases} -0.1125\epsilon, & \text{original orientation,} \\ +0.1125\epsilon, & \text{reversed orientation.} \end{cases}$$

Both orientations have exact normalization loss $L_\phi = 0.15\epsilon$.

The configuration crosses two orientations, $\epsilon \in \{0.01, 0.10\}$, $M_H \in \{0, 0.05\}$ and both budgets: 16 conditions with 1,000 independent datasets each. The four separately specified rules are confidence-aware fitting, empirical-mean fitting, fixed $(0, 0)$ and fixed $(1, 0)$. Each uses the same projected cellwise nuisances and evaluation order, with its own recomputed rectangular correction and complete 16-category support. Theorem 2 applies pointwise to each rule with the same allocation; evaluation does not select a maximum over them.

C.6 Published comparator implementations

DVDS (Dorn et al., 2025) uses the lower recentered influence-function score in Section 4 of its July 2022 preprint (v2),[1] with the binary nuisance specialization in Equation (10).

[1] Version 2, 22 July 2022: `https://arxiv.org/abs/2112.11449v2`.

Five treatment-balanced folds partition all $N/2$ observed reward records. For each held-out fold, treated means are fit on the other four folds; propensities combine the independent calibration records with those training folds' action counts. Success proportions use the same $1/2$ smoothing and propensities are clipped to [0.01,0.99]. Factual reward is subtracted from each cross-fitted counterfactual score before computing the variance, retaining its covariance with the bound estimator. The LCB is the mean minus the one-sided 95% normal critical value times its standard error. This binary specialization uses nominal Wald inference.

Version-specific DVDS inference assumptions. The assumption comparison uses this v2; condition and theorem numbers below refer to that version. Conditions 1–2 impose outcome moments, overlap and moment bounds for fitted nuisances in each cross-fitting fold. Its binary Condition 3 requires a bounded density for the random regression value $\mu(X, Z)$, with X and Z denoting that paper's covariates and treatment. Writing n for sample size, $\|\cdot\|_\infty$ for the uniform norm, and o_P for a term negligible in probability at the specified rate, Theorems 2–3 add $\|\widehat{\mu}-\mu\|_\infty = o_P(n^{-1/4})$ and the propensity–outcome mean-square (L^2) product rate $o_P(n^{-1/2})$ for sharp Wald inference. Theorem 4 retains Conditions 1–3 and consistent nuisance fits with that product rate, while dropping the uniform outcome-regression rate. Its binary proof is in Appendix C.10.5. A finite-history regression distribution has atoms and fails this version's density condition. Our rates report empirical behavior of the implemented Wald procedure under the stated discrete law.

KCMC implements the native full-history estimating program with saturated history indicators and point propensities (Ishikawa & He, 2023). Its point propensities are calibration-only empirical fractions clipped to $[0.01, 0.99]$, with empty calibration cells assigned 0.5. Inference uses a percentile-bootstrap adaptation. Its 999 resamples independently resample calibration and observed records, and refit propensities within every resample. The lower percentile at 0.05 is the one-sided LCB. All three published comparators have total budget N and use all $N/2$ reward observations; RAC reserves half those reward records for independent evaluation.

Cross-fitted QB. We specialize Dorn & Guo (2023, Sections 4.1–4.3 and Appendix A.1) using five folds, following their sample-splitting recommendation and five-fold empirical examples. Use five balanced, unstratified folds. For each record i, fit the fine-history empirical treated lower quantile Q_i at level $1/(1+\lambda)$ outside its fold. At a binary tie choose zero; empty treated training cells also predict zero. Preserve these record-specific out-of-fold predictions, including distinct predictions within one history. Let $\widetilde{e}_h$ be the calibration-only empirical propensity, clipped to $[0.01, 0.99]$ (empty calibration cell: 0.5), and put $C_Q = \sum_{i:A_i=1} \widetilde{e}_{H_i}^{-1}$. One pooled stabilized lower program estimates the counterfactual endpoint:

$$\begin{aligned} \widehat{V}_{\mathrm{QB}} &= \min_w C_Q^{-1} \sum_{i:A_i=1} R_i w_i, \\ \sum_{i:A_i=1} w_i = C_Q, \qquad & \sum_{i:A_i=1} Q_i w_i = \sum_{i:A_i=1} Q_i/\widetilde{e}_{H_i}, \\ 1 + \frac{1-\widetilde{e}_{H_i}}{\lambda \widetilde{e}_{H_i}} \le w_i \le 1 + \frac{\lambda(1-\widetilde{e}_{H_i})}{\widetilde{e}_{H_i}}&. \end{aligned} \tag{41}$$

Subtract the observed factual mean from this pooled endpoint to estimate improvement. In 999 independent calibration/observed record-bootstrap resamples, refit calibration-only propensities, preserve attached Q_i, and subtract the same observed draw's factual mean. The 0.05 percentile gives the nominal one-sided 95% LCB.

QB's conditional-density guarantees require continuous outcomes with a bounded density positive at the relevant quantiles; its partial-identification results include atoms. The binary family here is therefore outside those estimation assumptions. The QB comparison uses the same 18,000 datasets as the other procedures.

Undercoverage diagnostics. For a fixed condition and published method, let k of 1,000 bounds cover M_H. The one-sided exact binomial undercoverage probability is $P\{\mathrm{Binomial}(1000, 0.95) \le k\}$; Wilson intervals instead summarize coverage and certification uncertainty. At the robust boundary $M_H = 0$, the two counts are complements. The targeted comparator analysis uses one QB

implementation throughout and Bonferroni adjustment over $18 \times 3 = 54$ tests, with familywise level 0.05.

C.7 Structured nuisance scaling

The scaling study instantiates Corollary 2 with shared nuisance coefficients. It uses $K \in \{4, 16, 64, 256\}$ histories, total budgets $N \in \{10{,}000, 40{,}000\}$, sharp margins $M_H \in \{0, 0.05\}$, and 1,000 paired datasets per condition: 16,000 datasets in 16 conditions. Each dataset supplies paired, separately specified evaluations of the two RAC envelope constructions and the direct protected weak dual.

Write the observed history as $H = (Z, W)$, with $Z \in \{-1, +1\}$ and W on $J = K/2$ equally spaced points in $[-1, 1]$, including endpoints. The uniform history law is used only by the simulator. Fixed features and observed means are

$$\begin{aligned} x_e(H) = x_\mu(H) &= (1, Z, W, ZW)^\top, \\ e(H) &= 0.45 + 0.05Z + 0.06W + 0.02ZW, \\ \mu(H) &= 2/3 + 0.02Z + 0.055W + 0.015ZW. \end{aligned} \tag{42}$$

Every grid point has a distinct conditional law; both nuisance functions vary within the two fixed Z threshold cells. Coefficients and population history probabilities are not supplied to estimators. The fixed feature class correctly specifies both nuisance functions.

For the four corner histories $(Z, W) \in \{-1, +1\}^2$, let Q be the matrix of feature rows. Since $Q^\top Q = 4I$, any coefficient vector with corner predictions $f \in [0, 1]^4$ satisfies $\|\theta\|_2^2 = \|f\|_2^2/4 \le 1$. Thus declared physical constraints justify $S_e = S_\mu = 1$ without knowing sampling probabilities or true coefficients. This corner argument applies to the refined grids. In the uniform simulator law the feature Gram is $\mathrm{diag}(1, 1, v, v)$, $v = (J+1)/\{3(J-1)\}$, giving population condition numbers $1, 2.3333, 2.8182, 2.9535$ as K increases. The bound $\|x(H)\|_2 \le 2$ holds on every grid.

Hidden-confounding law and exact targets. An independent equiprobable latent $U \in \{-1, +1\}$ changes treatment probability to $e(H) + 0.06U$ and treated potential success probability to $\mu(H) - 0.0048/e(H) + 0.08U$. Assignment and potential rewards are independent given (H, U). Marginalizing gives the observed e, μ in (42), with assignment–treated-potential covariance 0.0048. Observed propensities lie in $[0.36, 0.58]$; the largest latent-to-observed odds factor is 1.3125, below sensitivity $\lambda = 2$. Control potential success is constant

$$q_0 = \frac{\mathbb{E}_H[(1 - e(H)) \max\{\mu(H)/2, -1 + 2\mu(H)\}] - M_H}{\mathbb{E}_H[1 - e(H)]},$$

computed without clipping. Exact latent enumeration verifies every target.

Equal protection and unchanged fitting/evaluation. Every total budget is split into $N/2$ calibration records and $N/4$ each for fitting and final evaluation. Cellwise RAC uses coordinate errors $0.0125/K$ in each nuisance family. Structured RAC uses $d_e = d_\mu = 4$, $\gamma = 1$, $S_e = S_\mu = 1$, $\alpha_e = \alpha_\mu = 0.0125$ and the prediction envelopes in Appendix B.2.2. Its propensity design uses calibration records; its outcome design uses the actual treated fitting records. Declared overlap is $\eta = 0.01$ in all procedures. The structured direct protected weak dual independently recomputes its propensity envelope with $\alpha_e = 0.025$ and no outcome event. It fits its own unaugmented weak-dual objective. With $\alpha_s = 0.025$, every procedure has the same complete pointwise lower coverage $0.975^2 = 0.950625$ under its stated assumptions. The allocations match complete protection while permitting procedure-specific propensity envelopes.

The corrected RAC score, rectangle-corner correction, fitting radius $\rho = 2 \log 40/n_s$, admissible soft thresholds and fixed-horizon betting are unchanged. All procedures share each dataset and evaluation order. Complete support enumerates all $4K$ declared categories, including absent ones. Noninteger-feature designs include sufficient-statistic assembly error in the outward bounds of Appendix B.2.2; empty envelopes or numerical/inference failures return $-\infty$, while a failed threshold optimization retains a valid predefined fit.

D Additional Experimental Results

D.1 Population normalization comparisons

The synthetic laws are specified in Appendix C.2; the representation study is specified in Appendix C.3.

D.1.1 Matched reward-alignment illustration

At $\lambda = 2$, the crossed and propensity partitions become lossless at $\omega = 1/3$, when their closed quantile sets first meet. The outcome and crossed partitions have identical cell masses, coarse propensities, and directional factors 4; the propensity partition has directional factors one.

Table 6: Controlled population comparison at $\omega = 0.1$ and $\lambda = 2$. The first two partitions match cell count, cell masses, coarse propensities, and propensity distortion. The third has different distortion; the final column is certificate-value loss.

partition	full	(Z, S)	transport	directional	symmetric/global	loss
outcome pairs	0.4625	0.4625	0.4625	0.303125	0.303125	0.0000
crossed pairs	0.4625	0.4625	0.4025	0.281250	0.281250	0.0600
propensity pairs	0.4625	0.3750	0.3750	0.375000	0.375000	0.0875

The (Z, S) index identifies all four histories in the outcome and crossed partitions because their propensities differ within each cell. In the propensity partition it adds no cells, so $V_{(Z,S)} = V_Z$.

D.1.2 Synthetic hidden-confounding utility laws

Table 1 uses three coarse interfaces (constant, retain X, retain Y) on the 24 positive-effect laws, giving 72 rows. Full history certifies 54 of them; V_Z certifies 49, with finer (Z, S) giving 50. The conditional quantile and saturated conditional-moment population programs of QB and KCMC both reproduce V_H on these 32 laws and the 18 image laws (Dorn & Guo, 2023; Ishikawa & He, 2023). These are known-law limits, distinct from the finite-data implementations in Section 6.2.

D.1.3 Known-law certificate retention and endpoint loss

The complete grid contains 16,200 paired interface/effect evaluations, including 10,800 positive-effect and 5,400 nonpositive-effect rows. All valid procedures give zero certificates on the nonpositive effects. On the positive rows, H and Z both certify 3,600; directional envelopes certify 60 and symmetric/global envelopes certify zero. The mean transport advantage over the directional envelope is 0.08182 with range [0.03001,0.12556]. Mean and maximum $V_H - V_Z$ are 0.000603 and 0.001924 over the complete grid. The 900 trained models provide three mappings each; these mappings and effects are correlated paired evaluations. The independent training unit is one of 20 representation replicates within a fixed assignment/learner/bottleneck/cell-count condition. Counts over interfaces are descriptive population summaries.

Table 7: Complete learner-level HC–MNIST aggregates. Each learner contributes 180 models, 2,160 positive-effect configurations, and 1,080 null or harmful configurations. Lower-endpoint loss is relative to sharp full history, averaged over all 3,240 configurations. Occupancy counts concern 540 trained interfaces per learner.

Learner	Transport cert.	Mean K_ϕ	Lower loss	Fewer occupied than requested cells
TARNet	720/2160	6.247	0.000654	112/540
CFR–MMD	720/2160	5.586	0.000654	130/540
CFR–Wasserstein	720/2160	5.318	0.000650	266/540
InvTARNet	720/2160	11.390	0.000404	0/540
propensity-aware	720/2160	9.416	0.000653	21/540

Table 8: HC–MNIST normalization-resolution summary over positive full-history certificate decisions, grouped by requested k-means count C. Loss is $\underline{V}_{\text{full}} - \underline{V}_{\phi}^{\text{tr}}$.

interface	retention	mean loss	median loss	max. loss
$C = 8$	1200/1200	0.001429	0.001924	0.001924
$C = 32$	1200/1200	0.001313	0.001900	0.001924
$C = 128$	1200/1200	0.001252	0.001862	0.001924
full history (sharp reference)	1200/1200	0	0	0

The learner summary shows that InvTARNet has smaller mean normalization loss than CFR–Wasserstein despite larger propensity distortion. Threshold-relevant outcome geometry explains why distortion alone cannot order these losses.

D.2 Protected certification and normalization trade-offs

The experimental construction and sample splits are specified in Appendix C.4.

D.2.1 Additional fitting regret and total regret within RAC

Table 9: **RAC (ours): coarse versus full-history thresholds.** $M_H = 0.05$; 1,000 paired datasets per row. Brackets: 95% paired Monte Carlo intervals; positive favors coarse. Regret: Appendix D.2.1.

Cutoff separation ϵ	Total budget N	Exact loss L_ϕ	Coarse RAC mean LCB	Full-history RAC mean LCB	Paired LCB difference (coarse minus full history)
0.00	10,000	0.0000	0.027022	0.026594	+0.000428 [+0.000316, +0.000539]
0.00	40,000	0.0000	0.039834	0.039596	+0.000237 [+0.000182, +0.000293]
0.01	10,000	0.0015	0.025641	0.025134	+0.000507 [+0.000409, +0.000606]
0.01	40,000	0.0015	0.038159	0.038085	+0.000074 [-0.000002, +0.000149]
0.10	10,000	0.0150	0.011300	0.021136	−0.009836 [-0.010091, -0.009581]
0.10	40,000	0.0150	0.024667	0.037795	−0.013128 [-0.013250, -0.013006]

Table 10 separates the structural cost of coarsening from additional fitting regret at $M_H = 0.05$. Recall that $J_X = M_X - \{D_X(\widehat{a}) - B\}$ measures regret within the chosen interface, whereas $T_X = M_H - \{D_X(\widehat{a}) - B\}$ measures total regret relative to full history. Thus $T_Z = L_\phi + J_Z$ and $T_H = J_H$. These known-law diagnostics exclude nuisance correction and betting gaps. At $\epsilon = 0.01$, the coarse fit has less additional fitting regret but greater total regret; its small LCB advantage in Table 4 comes from lower nuisance-protection and evaluation costs. At $\epsilon = 0.01$, $N = 10{,}000$, coarse RAC has 0.000204 more mean total threshold regret, but its remaining deficit $M_H - T_X - \widehat{L}$ is 0.000711 smaller, yielding a 0.000507 larger mean LCB. This remaining deficit combines the net nuisance deficit and evaluation gap in (11); their respective reductions are 0.000271 and 0.000440. At $\epsilon = 0$, both settings have essentially zero total threshold regret, so their LCB difference reflects these costs. At $\epsilon = 0.10$, full-history fitting recovers the material structural loss.

D.2.2 Condition-level coverage and certification

The tables below retain all 1,000 repetitions in each condition for the four protected procedure/interface combinations. Percentages in brackets are 95% Wilson intervals (Wilson, 1927). Mean LCB intervals use independent repetitions, and contrasts are paired. The data budget N includes calibration; every condition receives $N/2$ reward observations. Both RAC settings and the direct comparator covered their indexed coarsened endpoints in at least 98.9% of repetitions in every condition. Observed coverage is consistent with the finite-sample guarantee. At $\epsilon = 0$, $M_H = 0.05$, the paired coarse RAC-minus-direct mean LCB gain is 0.021885 [0.021336,0.022435] at $N = 10{,}000$, adding 295 certificates, and 0.012521 [0.012253,0.012789] at $N = 40{,}000$, with both procedures certifying all 1,000 datasets. At the latter condition, RAC's mean applied correction is 0.001202 and the direct comparator's nuisance-protection cost is 0.012140. Their evaluation gaps

Table 10: **RAC (ours)** population regret decomposition at $M_H = 0.05$. Each row averages 1,000 datasets; regrets refer to the fitted known-law dual. J_X is additional fitting regret within an interface; T_X is total regret relative to full history. Rounded zeros include values below 5×10^{-7}.

ϵ	Total budget N	Exact loss L_ϕ	Coarse: additional fitting regret J_Z	Coarse: total regret T_Z	Full history: total regret $T_H = J_H$
0.00	10,000	0.0000	0.000000	0.000000	0.000000
0.00	40,000	0.0000	0.000000	0.000000	0.000000
0.01	10,000	0.0015	0.000006	0.001506	0.001302
0.01	40,000	0.0015	0.000013	0.001513	0.001012
0.10	10,000	0.0150	0.000003	0.015003	0.000064
0.10	40,000	0.0150	0.000000	0.015000	0.000000

are 0.008962/0.010547 and conditional score variances 0.1458/0.2121. These diagnostics isolate augmentation with the interface and coverage requirement held fixed.

Table 11: RAC with coarse thresholds: complete 18-condition protected results. Each row has 1,000 datasets. Coverage targets M_H; certification is a positive LCB. Brackets give 95% Wilson or mean Monte Carlo intervals.

ϵ	M_H	N	Coverage: count; %	Certificates: count; %	Mean LCB
0.00	-0.02	10,000	995/1000; 99.5 [98.8, 99.8]	0/1000; 0.0 [0.0, 0.4]	-0.04477 [-0.04536, -0.04418]
0.00	-0.02	40,000	990/1000; 99.0 [98.2, 99.5]	0/1000; 0.0 [0.0, 0.4]	-0.03113 [-0.03142, -0.03084]
0.00	+0.00	10,000	996/1000; 99.6 [99.0, 99.8]	4/1000; 0.4 [0.2, 1.0]	-0.02451 [-0.02509, -0.02394]
0.00	+0.00	40,000	990/1000; 99.0 [98.2, 99.5]	10/1000; 1.0 [0.5, 1.8]	-0.01085 [-0.01114, -0.01057]
0.00	+0.05	10,000	994/1000; 99.4 [98.7, 99.7]	997/1000; 99.7 [99.1, 99.9]	0.02702 [0.02647, 0.02757]
0.00	+0.05	40,000	996/1000; 99.6 [99.0, 99.8]	1000/1000; 100.0 [99.6, 100.0]	0.03983 [0.03958, 0.04009]
0.01	-0.02	10,000	999/1000; 99.9 [99.4, 100.0]	0/1000; 0.0 [0.0, 0.4]	-0.04667 [-0.04726, -0.04608]
0.01	-0.02	40,000	998/1000; 99.8 [99.3, 99.9]	0/1000; 0.0 [0.0, 0.4]	-0.03269 [-0.03296, -0.03242]
0.01	+0.00	10,000	998/1000; 99.8 [99.3, 99.9]	2/1000; 0.2 [0.1, 0.7]	-0.02610 [-0.02666, -0.02553]
0.01	+0.00	40,000	995/1000; 99.5 [98.8, 99.8]	5/1000; 0.5 [0.2, 1.2]	-0.01226 [-0.01254, -0.01197]
0.01	+0.05	10,000	999/1000; 99.9 [99.4, 100.0]	1000/1000; 100.0 [99.6, 100.0]	0.02564 [0.02515, 0.02613]
0.01	+0.05	40,000	998/1000; 99.8 [99.3, 99.9]	1000/1000; 100.0 [99.6, 100.0]	0.03816 [0.03789, 0.03842]
0.10	-0.02	10,000	1000/1000; 100.0 [99.6, 100.0]	0/1000; 0.0 [0.0, 0.4]	-0.06022 [-0.06079, -0.05964]
0.10	-0.02	40,000	1000/1000; 100.0 [99.6, 100.0]	0/1000; 0.0 [0.0, 0.4]	-0.04632 [-0.04661, -0.04603]
0.10	+0.00	10,000	1000/1000; 100.0 [99.6, 100.0]	0/1000; 0.0 [0.0, 0.4]	-0.03977 [-0.04035, -0.03919]
0.10	+0.00	40,000	1000/1000; 100.0 [99.6, 100.0]	0/1000; 0.0 [0.0, 0.4]	-0.02637 [-0.02664, -0.02610]
0.10	+0.05	10,000	1000/1000; 100.0 [99.6, 100.0]	901/1000; 90.1 [88.1, 91.8]	0.01130 [0.01076, 0.01184]
0.10	+0.05	40,000	1000/1000; 100.0 [99.6, 100.0]	1000/1000; 100.0 [99.6, 100.0]	0.02467 [0.02441, 0.02492]

Table 12: RAC with full history thresholds: complete 18-condition protected results. Each row has 1,000 datasets. Coverage targets M_H; certification is a positive LCB. Brackets give 95% Wilson or mean Monte Carlo intervals.

ϵ	M_H	N	Coverage: count; %	Certificates: count; %	Mean LCB
0.00	-0.02	10,000	995/1000; 99.5 [98.8, 99.8]	0/1000; 0.0 [0.0, 0.4]	-0.04525 [-0.04586, -0.04465]
0.00	-0.02	40,000	991/1000; 99.1 [98.3, 99.5]	0/1000; 0.0 [0.0, 0.4]	-0.03142 [-0.03172, -0.03113]
0.00	+0.00	10,000	996/1000; 99.6 [99.0, 99.8]	4/1000; 0.4 [0.2, 1.0]	-0.02503 [-0.02563, -0.02444]
0.00	+0.00	40,000	990/1000; 99.0 [98.2, 99.5]	10/1000; 1.0 [0.5, 1.8]	-0.01106 [-0.01135, -0.01077]
0.00	+0.05	10,000	994/1000; 99.4 [98.7, 99.7]	996/1000; 99.6 [99.0, 99.8]	0.02659 [0.02603, 0.02716]
0.00	+0.05	40,000	995/1000; 99.5 [98.8, 99.8]	1000/1000; 100.0 [99.6, 100.0]	0.03960 [0.03933, 0.03986]
0.01	-0.02	10,000	999/1000; 99.9 [99.4, 100.0]	0/1000; 0.0 [0.0, 0.4]	-0.04715 [-0.04776, -0.04653]
0.01	-0.02	40,000	998/1000; 99.8 [99.3, 99.9]	0/1000; 0.0 [0.0, 0.4]	-0.03272 [-0.03300, -0.03244]
0.01	+0.00	10,000	998/1000; 99.8 [99.3, 99.9]	2/1000; 0.2 [0.1, 0.7]	-0.02661 [-0.02719, -0.02603]
0.01	+0.00	40,000	994/1000; 99.4 [98.7, 99.7]	6/1000; 0.6 [0.3, 1.3]	-0.01233 [-0.01262, -0.01204]
0.01	+0.05	10,000	999/1000; 99.9 [99.4, 100.0]	999/1000; 99.9 [99.4, 100.0]	0.02513 [0.02463, 0.02564]
0.01	+0.05	40,000	995/1000; 99.5 [98.8, 99.8]	1000/1000; 100.0 [99.6, 100.0]	0.03808 [0.03781, 0.03836]
0.10	-0.02	10,000	999/1000; 99.9 [99.4, 100.0]	0/1000; 0.0 [0.0, 0.4]	-0.04985 [-0.05049, -0.04921]
0.10	-0.02	40,000	994/1000; 99.4 [98.7, 99.7]	0/1000; 0.0 [0.0, 0.4]	-0.03304 [-0.03336, -0.03271]
0.10	+0.00	10,000	995/1000; 99.5 [98.8, 99.8]	5/1000; 0.5 [0.2, 1.2]	-0.02913 [-0.02979, -0.02847]
0.10	+0.00	40,000	992/1000; 99.2 [98.4, 99.6]	8/1000; 0.8 [0.4, 1.6]	-0.01324 [-0.01354, -0.01293]
0.10	+0.05	10,000	996/1000; 99.6 [99.0, 99.8]	984/1000; 98.4 [97.4, 99.0]	0.02114 [0.02052, 0.02176]
0.10	+0.05	40,000	997/1000; 99.7 [99.1, 99.9]	1000/1000; 100.0 [99.6, 100.0]	0.03779 [0.03750, 0.03809]

Table 13: Direct protected weak dual with coarse thresholds: complete 18-condition protected results. Each row has 1,000 datasets. Coverage targets M_H; certification is a positive LCB. Brackets give 95% Wilson or mean Monte Carlo intervals.

ϵ	M_H	N	Coverage: count; %	Certificates: count; %	Mean LCB
0.00	-0.02	10,000	1000/1000; 100.0 [99.6, 100.0]	0/1000; 0.0 [0.0, 0.4]	-0.06807 [-0.06885, -0.06730]
0.00	-0.02	40,000	1000/1000; 100.0 [99.6, 100.0]	0/1000; 0.0 [0.0, 0.4]	-0.04462 [-0.04500, -0.04424]
0.00	+0.00	10,000	1000/1000; 100.0 [99.6, 100.0]	0/1000; 0.0 [0.0, 0.4]	-0.04750 [-0.04823, -0.04676]
0.00	+0.00	40,000	1000/1000; 100.0 [99.6, 100.0]	0/1000; 0.0 [0.0, 0.4]	-0.02430 [-0.02467, -0.02393]
0.00	+0.05	10,000	1000/1000; 100.0 [99.6, 100.0]	702/1000; 70.2 [67.3, 73.0]	0.00514 [0.00444, 0.00583]
0.00	+0.05	40,000	1000/1000; 100.0 [99.6, 100.0]	1000/1000; 100.0 [99.6, 100.0]	0.02731 [0.02697, 0.02765]
0.01	-0.02	10,000	1000/1000; 100.0 [99.6, 100.0]	0/1000; 0.0 [0.0, 0.4]	-0.07001 [-0.07079, -0.06922]
0.01	-0.02	40,000	1000/1000; 100.0 [99.6, 100.0]	0/1000; 0.0 [0.0, 0.4]	-0.04632 [-0.04668, -0.04596]
0.01	+0.00	10,000	1000/1000; 100.0 [99.6, 100.0]	0/1000; 0.0 [0.0, 0.4]	-0.04920 [-0.04993, -0.04846]
0.01	+0.00	40,000	1000/1000; 100.0 [99.6, 100.0]	0/1000; 0.0 [0.0, 0.4]	-0.02586 [-0.02624, -0.02549]
0.01	+0.05	10,000	1000/1000; 100.0 [99.6, 100.0]	661/1000; 66.1 [63.1, 69.0]	0.00374 [0.00308, 0.00439]
0.01	+0.05	40,000	1000/1000; 100.0 [99.6, 100.0]	999/1000; 99.9 [99.4, 100.0]	0.02523 [0.02489, 0.02558]
0.10	-0.02	10,000	1000/1000; 100.0 [99.6, 100.0]	0/1000; 0.0 [0.0, 0.4]	-0.08270 [-0.08344, -0.08197]
0.10	-0.02	40,000	1000/1000; 100.0 [99.6, 100.0]	0/1000; 0.0 [0.0, 0.4]	-0.05944 [-0.05982, -0.05907]
0.10	+0.00	10,000	1000/1000; 100.0 [99.6, 100.0]	0/1000; 0.0 [0.0, 0.4]	-0.06174 [-0.06245, -0.06103]
0.10	+0.00	40,000	1000/1000; 100.0 [99.6, 100.0]	0/1000; 0.0 [0.0, 0.4]	-0.03909 [-0.03945, -0.03872]
0.10	+0.05	10,000	1000/1000; 100.0 [99.6, 100.0]	182/1000; 18.2 [15.9, 20.7]	-0.00965 [-0.01033, -0.00898]
0.10	+0.05	40,000	1000/1000; 100.0 [99.6, 100.0]	989/1000; 98.9 [98.0, 99.4]	0.01221 [0.01189, 0.01254]

Table 14: Direct protected weak dual with full history thresholds: complete 18-condition protected results. Each row has 1,000 datasets. Coverage targets M_H; certification is a positive LCB. Brackets give 95% Wilson or mean Monte Carlo intervals.

ϵ	M_H	N	Coverage: count; %	Certificates: count; %	Mean LCB
0.00	-0.02	10,000	1000/1000; 100.0 [99.6, 100.0]	0/1000; 0.0 [0.0, 0.4]	-0.06861 [-0.06941, -0.06781]
0.00	-0.02	40,000	1000/1000; 100.0 [99.6, 100.0]	0/1000; 0.0 [0.0, 0.4]	-0.04497 [-0.04537, -0.04457]
0.00	+0.00	10,000	1000/1000; 100.0 [99.6, 100.0]	0/1000; 0.0 [0.0, 0.4]	-0.04865 [-0.04944, -0.04786]
0.00	+0.00	40,000	1000/1000; 100.0 [99.6, 100.0]	0/1000; 0.0 [0.0, 0.4]	-0.02464 [-0.02502, -0.02426]
0.00	+0.05	10,000	1000/1000; 100.0 [99.6, 100.0]	678/1000; 67.8 [64.8, 70.6]	0.00433 [0.00358, 0.00508]
0.00	+0.05	40,000	1000/1000; 100.0 [99.6, 100.0]	1000/1000; 100.0 [99.6, 100.0]	0.02693 [0.02657, 0.02728]
0.01	-0.02	10,000	1000/1000; 100.0 [99.6, 100.0]	0/1000; 0.0 [0.0, 0.4]	-0.07068 [-0.07151, -0.06986]
0.01	-0.02	40,000	1000/1000; 100.0 [99.6, 100.0]	0/1000; 0.0 [0.0, 0.4]	-0.04654 [-0.04691, -0.04617]
0.01	+0.00	10,000	1000/1000; 100.0 [99.6, 100.0]	0/1000; 0.0 [0.0, 0.4]	-0.04983 [-0.05060, -0.04907]
0.01	+0.00	40,000	1000/1000; 100.0 [99.6, 100.0]	0/1000; 0.0 [0.0, 0.4]	-0.02592 [-0.02631, -0.02553]
0.01	+0.05	10,000	1000/1000; 100.0 [99.6, 100.0]	634/1000; 63.4 [60.4, 66.3]	0.00307 [0.00239, 0.00375]
0.01	+0.05	40,000	1000/1000; 100.0 [99.6, 100.0]	999/1000; 99.9 [99.4, 100.0]	0.02499 [0.02463, 0.02536]
0.10	-0.02	10,000	1000/1000; 100.0 [99.6, 100.0]	0/1000; 0.0 [0.0, 0.4]	-0.07477 [-0.07560, -0.07395]
0.10	-0.02	40,000	1000/1000; 100.0 [99.6, 100.0]	0/1000; 0.0 [0.0, 0.4]	-0.04737 [-0.04781, -0.04693]
0.10	+0.00	10,000	1000/1000; 100.0 [99.6, 100.0]	0/1000; 0.0 [0.0, 0.4]	-0.05418 [-0.05500, -0.05335]
0.10	+0.00	40,000	1000/1000; 100.0 [99.6, 100.0]	0/1000; 0.0 [0.0, 0.4]	-0.02732 [-0.02776, -0.02689]
0.10	+0.05	10,000	1000/1000; 100.0 [99.6, 100.0]	377/1000; 37.7 [34.7, 40.7]	-0.00372 [-0.00447, -0.00298]
0.10	+0.05	40,000	1000/1000; 100.0 [99.6, 100.0]	1000/1000; 100.0 [99.6, 100.0]	0.02346 [0.02306, 0.02386]

At $M_H = 0$, false certification equals sharp-endpoint noncoverage. For coarse RAC, boundary false certificates are 4, 10, 2, 5, 0, and 0 out of 1,000, ordered by $\epsilon = 0, 0.01, 0.10$ and the two increasing budgets. Full-history RAC has 4, 10, 2, 6, 5, and 8; both direct settings have zero in every boundary condition. Harmful-margin controls have zero certifications for every protected row. Pooling those easy controls would obscure the more informative boundary assessment.

D.3 Threshold fitting and fixed-threshold ablations

These paired ablations use the protocol in Appendix C.5.1.

Positive-margin comparisons and uncertainty. Table 15 gives all six $M_H = 0.05$ conditions separately for coarse and full-history thresholds. The coarse aggregate equally weights the six condition means: mean LCBs are 0.027770 for RAC and 0.024618 for the ablation, with paired gain 0.003152 [0.002946,0.003359]. Certification increases from 5,660/6,000 to 5,898/6,000: 238 gained certificates and none lost, a paired rate change of 3.97 percentage points [3.48,4.45]. The strongest frequency gains occur at $N = 10{,}000$ and $\epsilon = 0, 0.01$, with 103 and 95 added certificates. At the primary condition $\epsilon = 0, N = 40{,}000$, the mean LCB gain is 0.003761 [0.003395,0.004127], but certification is already nearly saturated (1,000 versus 999).

The full-history aggregate has mean LCB gain 0.003059 [0.002887,0.003232], with 5,979 versus 5,780 certificates among the same 6,000 datasets. At $\epsilon = 0.10, N = 10{,}000$, confidence-aware fitting has a small regression, -0.000037 [$-0.000070, -0.000004$], with 984 certificates under both criteria. At $N = 40{,}000$ the full-history difference is numerically negligible. These intervals describe conditionwise Monte Carlo uncertainty and are not multiplicity-adjusted.

Appendix C.1 specifies the paired and stratified uncertainty calculations.

Table 15: **RAC (ours): fitting ablation** at $M_H = 0.05$. Each row has 1,000 paired datasets. The ablation sets $\rho = 0$; all other rules are matched. Differences are RAC minus empirical-mean fitting, with descriptive 95% paired Monte Carlo intervals. Certificate counts are out of 1,000 under each criterion. Both criteria have the same formal protection; Table 19 gives empirical coverage. Rounded zeros include differences below 10^{-12}.

ϵ	Total budget N	RAC mean LCB	Mean-fit ablation mean LCB	Paired LCB difference (RAC minus mean fit)	Certificates RAC / mean fit
(a) Coarse thresholds					
0.00	10,000	0.027022	0.021319	+0.005703 [+0.004964, +0.006441]	997 / 894
0.00	40,000	0.039834	0.036072	+0.003761 [+0.003395, +0.004127]	1000 / 999
0.01	10,000	0.025641	0.020802	+0.004839 [+0.004144, +0.005533]	1000 / 905
0.01	40,000	0.038159	0.035257	+0.002901 [+0.002554, +0.003248]	1000 / 1000
0.10	10,000	0.011300	0.009693	+0.001608 [+0.001109, +0.002106]	901 / 865
0.10	40,000	0.024667	0.024564	+0.000103 [+0.000012, +0.000194]	1000 / 997
(b) Full-history thresholds					
0.00	10,000	0.026594	0.019486	+0.007108 [+0.006411, +0.007805]	996 / 887
0.00	40,000	0.039596	0.035684	+0.003913 [+0.003588, +0.004237]	1000 / 1000
0.01	10,000	0.025134	0.020098	+0.005036 [+0.004410, +0.005663]	999 / 909
0.01	40,000	0.038085	0.035748	+0.002337 [+0.002036, +0.002637]	1000 / 1000
0.10	10,000	0.021136	0.021173	−0.000037 [-0.000070, -0.000004]	984 / 984
0.10	40,000	0.037795	0.037795	0.000000 [0.000000, 0.000000]	1000 / 1000

Population-optimal fixed threshold. For this coarse family, $a = (0, 0)$ is population-optimal at every separation. Its mean LCB is 0.027851, versus 0.027770 for confidence-aware fitting and 0.024618 for empirical-mean fitting across the six positive conditions. Fixed zero minus confidence-aware fitting is $+0.000081$ [+0.000059,+0.000103], with three certificate gains and none lost: 5,901 versus 5,898/6,000. Relative to empirical-mean fitting, its gain is $+0.003233$ [+0.003023,+0.003443], with 241 gains and no losses. Table 16 retains the condition-level compar-

ison. Confidence-aware fitting's benefit in this family is therefore relative to empirical-mean fitting; a simple population-informed fixed rule is slightly stronger.

For full-history thresholds, fixed zero is generally suboptimal. Its secondary aggregate LCB difference is -0.003539 [-0.003598,-0.003479] versus confidence-aware fitting, and -0.000479 [-0.000666,-0.000293] versus empirical-mean fitting. The coarse and full-history analyses use the same datasets and retain signed nuisance and evaluation terms in the decomposition.

Table 16: **RAC (ours; coarse): fixed-zero ablation** in the original cutoff family at $M_H = 0.05$. Each row has 1,000 paired datasets. Brackets are descriptive 95% Monte Carlo intervals (paired for differences). Counts are certificates/1,000; confidence-aware and empirical-mean results appear in Table 15.

ϵ	N	Fixed-zero mean LCB	Fixed zero minus confidence-aware	Certificates
0.00	10,000	0.027176 [0.026641, 0.027711]	$+0.000154$ [+0.000062, +0.000246]	1000
0.00	40,000	0.039952 [0.039703, 0.040201]	$+0.000118$ [+0.000061, +0.000175]	1000
0.01	10,000	0.025721 [0.025235, 0.026207]	$+0.000080$ [+0.000035, +0.000125]	1000
0.01	40,000	0.038284 [0.038030, 0.038538]	$+0.000125$ [+0.000064, +0.000186]	1000
0.10	10,000	0.011307 [0.010771, 0.011843]	$+0.000007$ [-0.000002, +0.000016]	901
0.10	40,000	0.024667 [0.024411, 0.024923]	$+0.000000$ [+0.000000, +0.000000]	1000

At $M_H = 0$, fixed zero gives 21/6,000 coarse false certificates, the same as confidence-aware fitting, versus 20/6,000 for empirical-mean fitting. All 6,000 negative-margin datasets give zero certificates under each rule. Condition-level robust-null rates and positive-margin coverage for fixed zero are given in Table 17; full-history fixed-zero null counts coincide with the coarse counts because their scores are identical.

Table 17: **RAC (ours): fixed-zero coverage and robust-null controls.** Every cell has 1,000 datasets. Entries are counts/1,000; percentage [95% Wilson interval]. Positive-margin coverage means $\widehat{L} \le M_H$ at $M_H = 0.05$; robust-null false certification means $\widehat{L} > 0$ at $M_H = 0$. Fixed-zero scores are identical under both threshold interfaces.

ϵ	N	Positive-margin coverage	Robust-null false certification
0.00	10,000	994/1000; 99.4 [98.7, 99.7]	4/1000; 0.4 [0.2, 1.0]
0.00	40,000	997/1000; 99.7 [99.1, 99.9]	10/1000; 1.0 [0.5, 1.8]
0.01	10,000	999/1000; 99.9 [99.4, 100.0]	2/1000; 0.2 [0.1, 0.7]
0.01	40,000	998/1000; 99.8 [99.3, 99.9]	5/1000; 0.5 [0.2, 1.2]
0.10	10,000	1000/1000; 100.0 [99.6, 100.0]	0/1000; 0.0 [0.0, 0.4]
0.10	40,000	1000/1000; 100.0 [99.6, 100.0]	0/1000; 0.0 [0.0, 0.4]

Population-equivalent thresholds and statistical uncertainty. Subtracting (11) between fitting criteria cancels structural loss; Table 18 separates the remaining contributions. For coarse thresholds, a smaller evaluation gap contributes about 72.0% of the mean LCB gain, a smaller net nuisance deficit 22.5%, and lower additional fitting regret 5.5%. The inference algorithm and structural normalization loss are unchanged. The net nuisance term is applied correction minus signed bias, as in (24). Mean signed biases are of order 10^{-6} here, so interval uncertainty accounts for most of the larger protection term.

At $\epsilon = 0, M_H = 0.05, N = 40{,}000$ with coarse thresholds, both criteria have negligible population threshold regret. Yet mean conditional score variance is 0.1458 under RAC versus 0.2968 under empirical-mean fitting, and mean declared score range is 1.5416 versus 2.7280. Their mean first-cell thresholds are 0.0136 and 0.3860, respectively; every threshold in that cell is population-optimal. Mean net nuisance deficits are 0.001204/0.001850, whereas applied corrections are 0.001202/0.001840. Evaluation gaps are 0.008962/0.012078. Thus population-equivalent thresholds can yield different statistical uncertainty under the same protected evaluation.

Table 18: **RAC (ours): signed fitting-gain decomposition.** Each interface separately aggregates six positive-margin conditions, with 1,000 pairs per condition. Contributions are ablation minus RAC deficits, so their sum equals the RAC-minus-ablation LCB gain. Structural loss cancels; the nuisance column is the net deficit including signed bias. Last two columns give RAC / ablation score diagnostics.

Interface	Fitting-regret contribution	Net nuisance contribution	Evaluation-gap contribution	LCB gain	Conditional variance	Declared range
Coarse	$+0.000172$	$+0.000709$	$+0.002271$	$+0.003152$	0.1470 / 0.2213	1.5232 / 2.0975
Full history	-0.000025	$+0.000724$	$+0.002360$	$+0.003059$	0.1701 / 0.2425	1.6139 / 2.1759

Coverage and robust-null controls. Table 19 reports positive-condition coverage alongside the certification counts in Table 15. For $M_H = 0$, Table 20 gives conditionwise false certification; a positive LCB is exactly sharp-endpoint noncoverage. All negative-margin controls ($M_H = -0.02$) give zero certificates among 6,000 datasets per interface and criterion. They are kept separate from the more informative robust-null assessment. Across all 18 conditions, minimum score-mean coverage is 97.6%/97.8% for coarse thresholds and 97.8%/98.1% for full history, against the 97.5% conditional target. Both criteria retain the finite-sample guarantee of Theorem 2.

Table 19: **RAC (ours): coverage in the fitting ablation**, $M_H = 0.05$. Each row has 1,000 datasets per interface; coverage means $\widehat{L} \leq M_H$ and the nominal target is 95%. Each cell lists confidence-aware RAC first and the empirical-mean ablation second: count/1,000; percentage [95% Wilson interval]. Certification is reported in Table 15.

ϵ	Total budget N	Coarse threshold coverage	Full-history threshold coverage
0.00	10,000	994/1000; 99.4 [98.7, 99.7] 997/1000; 99.7 [99.1, 99.9]	994/1000; 99.4 [98.7, 99.7] 997/1000; 99.7 [99.1, 99.9]
0.00	40,000	996/1000; 99.6 [99.0, 99.8] 994/1000; 99.4 [98.7, 99.7]	995/1000; 99.5 [98.8, 99.8] 993/1000; 99.3 [98.6, 99.7]
0.01	10,000	999/1000; 99.9 [99.4, 100.0] 999/1000; 99.9 [99.4, 100.0]	999/1000; 99.9 [99.4, 100.0] 1000/1000; 100.0 [99.6, 100.0]
0.01	40,000	998/1000; 99.8 [99.3, 99.9] 997/1000; 99.7 [99.1, 99.9]	995/1000; 99.5 [98.8, 99.8] 994/1000; 99.4 [98.7, 99.7]
0.10	10,000	1000/1000; 100.0 [99.6, 100.0] 1000/1000; 100.0 [99.6, 100.0]	996/1000; 99.6 [99.0, 99.8] 996/1000; 99.6 [99.0, 99.8]
0.10	40,000	1000/1000; 100.0 [99.6, 100.0] 1000/1000; 100.0 [99.6, 100.0]	997/1000; 99.7 [99.1, 99.9] 997/1000; 99.7 [99.1, 99.9]

Table 20: **RAC (ours): robust-null fitting controls**, $M_H = 0$. Each row has 1,000 paired datasets. Entries are false certificates/1,000; percentage [95% Wilson interval]. Endpoint coverage is one minus the false-certification rate, so its interval is obtained by complementing the endpoints. Negative-margin controls are reported separately in the text.

ϵ	Total budget N	Interface	Confidence-aware RAC	Mean-fit ablation ($\rho = 0$)
0.00	10,000	Coarse	4/1000; 0.4 [0.2, 1.0]	4/1000; 0.4 [0.2, 1.0]
0.00	40,000	Coarse	10/1000; 1.0 [0.5, 1.8]	10/1000; 1.0 [0.5, 1.8]
0.01	10,000	Coarse	2/1000; 0.2 [0.1, 0.7]	2/1000; 0.2 [0.1, 0.7]
0.01	40,000	Coarse	5/1000; 0.5 [0.2, 1.2]	4/1000; 0.4 [0.2, 1.0]
0.10	10,000	Coarse	0/1000; 0.0 [0.0, 0.4]	0/1000; 0.0 [0.0, 0.4]
0.10	40,000	Coarse	0/1000; 0.0 [0.0, 0.4]	0/1000; 0.0 [0.0, 0.4]
0.00	10,000	Full history	4/1000; 0.4 [0.2, 1.0]	4/1000; 0.4 [0.2, 1.0]
0.00	40,000	Full history	10/1000; 1.0 [0.5, 1.8]	8/1000; 0.8 [0.4, 1.6]
0.01	10,000	Full history	2/1000; 0.2 [0.1, 0.7]	2/1000; 0.2 [0.1, 0.7]
0.01	40,000	Full history	6/1000; 0.6 [0.3, 1.3]	9/1000; 0.9 [0.5, 1.7]
0.10	10,000	Full history	5/1000; 0.5 [0.2, 1.2]	5/1000; 0.5 [0.2, 1.2]
0.10	40,000	Full history	8/1000; 0.8 [0.4, 1.6]	8/1000; 0.8 [0.4, 1.6]

D.4 Changing population-optimal thresholds

The original and reversed orientations are specified in Appendix C.5.2. Each condition evaluates all four threshold rules on 1,000 paired datasets.

Certificate strength and adaptation. Across eight positive-margin conditions, confidence-aware fitting exceeds empirical-mean fitting by 0.003337 [0.003156,0.003518] in mean LCB, with 455 certificate gains and 11 losses, or 444 net gains. Table 21 compares all three ablations; Table 22 keeps every positive condition. The mean LCB gain over empirical-mean fitting decomposes as -0.001329 from additional fitting regret, $+0.001162$ from net nuisance deficit and $+0.003505$ from evaluation gap. Thus reduced statistical deficits can outweigh greater population fitting regret. Net deficit is correction minus signed bias, and the betting algorithm is unchanged.

At reversed $\epsilon = 0.10, N = 40{,}000$, the mean fitted first threshold is 0.578392, versus approximately zero in the original orientation. Confidence-aware fitting has mean LCB 0.014660 and 990/1,000 certificates; empirical-mean fitting gives 0.015316 and 980, fixed zero 0.013293 and 998, and fixed $(1, 0)$ 0.015421 and 980. The fitted threshold responds to the changed population optimum, but mean certificate strength and certification frequency rank the rules differently. At $N = 10{,}000$ in this reversed condition, fixed zero is stronger on both measures (0.000288 and 519 certificates, versus -0.000913 and 481 for confidence-aware fitting). Fixed zero remains slightly stronger in the aggregate. These comparisons retain population regret, mean LCB and certification frequency as separate quantities.

Table 21: **RAC (ours; coarse): threshold-choice aggregate** in the changing-threshold study. Eight equally weighted positive conditions, 1,000 paired datasets each. Differences are confidence-aware minus the specified ablation, with descriptive 95% stratified paired Monte Carlo intervals. Counts list certificate gains/losses. All rules share the same validity guarantee.

Ablation	Paired mean LCB difference	Certificate gains / losses
Empirical mean ($\rho = 0$)	$+0.003337$ [+0.003156, +0.003518]	455 / 11
Fixed $(0, 0)$	-0.000062 [-0.000112, -0.000011]	2 / 53
Fixed $(1, 0)$	$+0.017264$ [+0.017071, +0.017457]	2,475 / 13

Table 22: **RAC (ours; coarse): changing-threshold positive conditions.** $M_H = 0.05$; 1,000 paired datasets per row. Entries give mean LCB; certificates/1,000. All four rules are specified before evaluation. Paired aggregate intervals are in Table 21.

ϵ	N	**RAC (ours)** confidence-aware	Empirical-mean ablation	Fixed $(0, 0)$	Fixed $(1, 0)$
Original orientation					
0.01	10,000	0.025705; 997	0.020125; 884	0.025880; 999	-0.000385; 489
0.01	40,000	0.038270; 1000	0.035088; 1000	0.038337; 1000	0.026989; 1000
0.10	10,000	0.011102; 909	0.009617; 877	0.011110; 910	-0.025844; 59
0.10	40,000	0.024669; 1000	0.024608; 1000	0.024669; 1000	0.002929; 643
Reversed orientation					
0.01	10,000	0.024134; 995	0.017467; 859	0.024348; 997	0.000703; 502
0.01	40,000	0.036890; 1000	0.033369; 1000	0.037085; 1000	0.027893; 1000
0.10	10,000	-0.000913; 481	-0.007767; 328	0.000288; 519	-0.011304; 237
0.10	40,000	0.014660; 990	0.015316; 980	0.013293; 998	0.015421; 980

For the highlighted reversed condition ($\epsilon = 0.10, N = 40{,}000, M_H = 0.05$), confidence-aware minus fixed-zero mean LCB is $+0.001368$ [+0.001133,+0.001602]. The certification-rate difference is -0.80 $[-1.35, -0.25]$ percentage points, with zero certification gains and eight losses among 1,000 paired datasets. Both are approximate paired Student-t 95% Monte Carlo intervals from unrounded within-dataset differences; the rate interval uses differences of certification indicators. Thus stronger average LCBs coexist with fewer positive certificates.

Coverage and robust-null controls. Across positive conditions, confidence-aware, empirical-mean, fixed-zero and fixed-$(1, 0)$ rules certify 7,372, 6,928, 7,423 and 4,910/8,000, respectively. At the robust null, false certificates total 12, 14, 12 and 19/8,000. Table 23 gives condition-level counts and uncertainty. At $M_H = 0$, coverage equals one minus false certification; the displayed intervals complement to coverage intervals. Equal-stratum aggregate intervals use the construction in Appendix C.1, with eight strata in place of six. Each rule has the pointwise guarantee of Theorem 2; the intervals describe simulation uncertainty.

Table 23: **RAC (ours; coarse): changing-threshold controls.** Each cell has 1,000 datasets. Entries are robust-null false-certificate counts with 95% Wilson rate intervals in percent; nominal endpoint coverage is 95%. All eight null conditions are retained.

ϵ	N	**RAC (ours)** confidence-aware	Empirical-mean ablation	Fixed $(0, 0)$	Fixed $(1, 0)$
Original orientation					
0.01	10,000	4 [0.2, 1.0]	4 [0.2, 1.0]	4 [0.2, 1.0]	3 [0.1, 0.9]
0.01	40,000	4 [0.2, 1.0]	3 [0.1, 0.9]	4 [0.2, 1.0]	4 [0.2, 1.0]
0.10	10,000	0 [0.0, 0.4]	0 [0.0, 0.4]	0 [0.0, 0.4]	0 [0.0, 0.4]
0.10	40,000	0 [0.0, 0.4]	0 [0.0, 0.4]	0 [0.0, 0.4]	0 [0.0, 0.4]
Reversed orientation					
0.01	10,000	3 [0.1, 0.9]	3 [0.1, 0.9]	3 [0.1, 0.9]	4 [0.2, 1.0]
0.01	40,000	1 [0.0, 0.6]	3 [0.1, 0.9]	1 [0.0, 0.6]	7 [0.3, 1.4]
0.10	10,000	0 [0.0, 0.4]	1 [0.0, 0.6]	0 [0.0, 0.4]	1 [0.0, 0.6]
0.10	40,000	0 [0.0, 0.4]	0 [0.0, 0.4]	0 [0.0, 0.4]	0 [0.0, 0.4]

D.5 PUBLISHED-COMPARATOR RESULTS

Appendix C.6 specifies the implementations, sample use and undercoverage diagnostics.

Table 24: **RAC (ours; coarse)** versus published procedures: $\epsilon = 0$, total record budget $N = 40{,}000$, 1,000 datasets per procedure per panel; nominal coverage 95%. Brackets: 95% Wilson (rates) and Monte Carlo (mean LCBs) intervals. All panel (a) certification: 100.0% [99.6, 100.0].

(a) Positive margin $M_H = 0.05$: all certify 1,000/1,000

Procedure	Coverage (%)	Mean LCB
RAC (ours; coarse)	99.6 [99.0, 99.8]	0.039834 [0.039577, 0.040091]
DVDS	94.3 [92.7, 95.6]	0.043503 [0.043256, 0.043751]
QB (five-fold cross-fitted)	90.6 [88.6, 92.3]	0.045217 [0.044984, 0.045450]
KCMC (bootstrap adaptation)	84.8 [82.4, 86.9]	0.045064 [0.044774, 0.045354]

(b) Robust null $M_H = 0$: a positive LCB is a false robust certificate

Procedure	False certificates	False certification (%)
RAC (ours; coarse)	10/1,000	1.0 [0.5, 1.8]
DVDS	58/1,000	5.8 [4.5, 7.4]
QB (five-fold cross-fitted)	98/1,000	9.8 [8.1, 11.8]
KCMC (bootstrap adaptation)	139/1,000	13.9 [11.9, 16.2]

There are no DVDS flags, six QB flags (the six $\epsilon = 0$ conditions), and nine KCMC flags (all six $\epsilon = 0$ conditions and the three $\epsilon = 0.01, N = 10{,}000$ conditions). There are no QB flags for either positive ϵ value. Table 28 isolates boundary false certification with Wilson intervals.

D.6 Structured nuisance calibration

The nuisance models, confidence allocations and fixed interfaces are specified in Appendix C.7.

D.6.1 Complete condition-level certification and coverage

Table 29 reports all conditions. At $K = 256, N = 40{,}000, M_H = 0.05$, structured RAC certifies 1,000/1,000, versus 0/1,000 for cellwise RAC and 17/1,000 for the equally protected structured direct comparator. Its paired LCB gains are $0.105190\,[0.105107, 0.105273]$ over cellwise RAC and $0.032369\,[0.032112, 0.032626]$ over the direct comparator (95% paired Monte Carlo intervals), with 1,000 and 983 gained certificates, respectively, and zero lost certificates. At $N = 10{,}000$, structured RAC retains negative mean LCBs at every K and certifies only 81/1,000 at $K = 256$. Cellwise RAC has stronger mean LCBs at $K = 4, 16$ under both budgets.

Empirical coverage is 1,000/1,000 in each displayed procedure–condition cell (95% Wilson interval $[99.6, 100.0]\%$). Each robust-null cell has 0/1,000 false certificates (Wilson interval $[0.0, 0.4]\%$). These condition-level observations use paired datasets. Nuisance calibration, score-mean coverage and sharp-endpoint coverage are distinct diagnostics. Structured validity follows under the model and construction assumptions in Appendix B.2.2. Certificate strength additionally depends on sufficiently observed feature directions. Uncertainty and pairing follow Appendix C.1.

Table 25: DVDS on all 18 conditions; 1,000 datasets per row. Coverage targets the sharp M_H endpoint. Percentages have 95% Wilson intervals and LCB means have 95% Monte Carlo intervals.

ϵ	M_H	N	Coverage: count; %	Certificates: count; %	Mean LCB
0.00	-0.02	10,000	938/1000; 93.8 [92.1, 95.1]	0/1000; 0.0 [0.0, 0.4]	-0.03390 [-0.03442, -0.03339]
0.00	-0.02	40,000	949/1000; 94.9 [93.4, 96.1]	0/1000; 0.0 [0.0, 0.4]	-0.02692 [-0.02718, -0.02666]
0.00	+0.00	10,000	940/1000; 94.0 [92.4, 95.3]	60/1000; 6.0 [4.7, 7.6]	-0.01395 [-0.01447, -0.01344]
0.00	+0.00	40,000	942/1000; 94.2 [92.6, 95.5]	58/1000; 5.8 [4.5, 7.4]	-0.00671 [-0.00697, -0.00645]
0.00	+0.05	10,000	929/1000; 92.9 [91.1, 94.3]	1000/1000; 100.0 [99.6, 100.0]	0.03697 [0.03646, 0.03748]
0.00	+0.05	40,000	943/1000; 94.3 [92.7, 95.6]	1000/1000; 100.0 [99.6, 100.0]	0.04350 [0.04326, 0.04375]
0.01	-0.02	10,000	961/1000; 96.1 [94.7, 97.1]	0/1000; 0.0 [0.0, 0.4]	-0.03470 [-0.03519, -0.03420]
0.01	-0.02	40,000	974/1000; 97.4 [96.2, 98.2]	0/1000; 0.0 [0.0, 0.4]	-0.02723 [-0.02746, -0.02700]
0.01	+0.00	10,000	967/1000; 96.7 [95.4, 97.6]	33/1000; 3.3 [2.4, 4.6]	-0.01428 [-0.01475, -0.01380]
0.01	+0.00	40,000	970/1000; 97.0 [95.7, 97.9]	30/1000; 3.0 [2.1, 4.3]	-0.00688 [-0.00711, -0.00666]
0.01	+0.05	10,000	970/1000; 97.0 [95.7, 97.9]	1000/1000; 100.0 [99.6, 100.0]	0.03613 [0.03568, 0.03658]
0.01	+0.05	40,000	975/1000; 97.5 [96.3, 98.3]	1000/1000; 100.0 [99.6, 100.0]	0.04315 [0.04293, 0.04337]
0.10	-0.02	10,000	953/1000; 95.3 [93.8, 96.4]	0/1000; 0.0 [0.0, 0.4]	-0.03121 [-0.03162, -0.03079]
0.10	-0.02	40,000	948/1000; 94.8 [93.2, 96.0]	0/1000; 0.0 [0.0, 0.4]	-0.02563 [-0.02584, -0.02542]
0.10	+0.00	10,000	948/1000; 94.8 [93.2, 96.0]	52/1000; 5.2 [4.0, 6.8]	-0.01104 [-0.01147, -0.01060]
0.10	+0.00	40,000	952/1000; 95.2 [93.7, 96.4]	48/1000; 4.8 [3.6, 6.3]	-0.00569 [-0.00590, -0.00549]
0.10	+0.05	10,000	946/1000; 94.6 [93.0, 95.8]	1000/1000; 100.0 [99.6, 100.0]	0.03923 [0.03882, 0.03964]
0.10	+0.05	40,000	944/1000; 94.4 [92.8, 95.7]	1000/1000; 100.0 [99.6, 100.0]	0.04495 [0.04476, 0.04514]

Table 26: Five-fold cross-fitted QB on all 18 conditions; 1,000 datasets per row. Coverage targets the sharp M_H endpoint. Percentages have 95% Wilson intervals and LCB means have 95% Monte Carlo intervals.

ϵ	M_H	N	Coverage: count; %	Certificates: count; %	Mean LCB
0.00	-0.02	10,000	908/1000; 90.8 [88.8, 92.4]	0/1000; 0.0 [0.0, 0.4]	-0.02995 [-0.03042, -0.02948]
0.00	-0.02	40,000	911/1000; 91.1 [89.2, 92.7]	0/1000; 0.0 [0.0, 0.4]	-0.02502 [-0.02525, -0.02478]
0.00	+0.00	10,000	912/1000; 91.2 [89.3, 92.8]	88/1000; 8.8 [7.2, 10.7]	-0.01004 [-0.01050, -0.00958]
0.00	+0.00	40,000	902/1000; 90.2 [88.2, 91.9]	98/1000; 9.8 [8.1, 11.8]	-0.00488 [-0.00511, -0.00466]
0.00	+0.05	10,000	897/1000; 89.7 [87.7, 91.4]	1000/1000; 100.0 [99.6, 100.0]	0.04060 [0.04015, 0.04106]
0.00	+0.05	40,000	906/1000; 90.6 [88.6, 92.3]	1000/1000; 100.0 [99.6, 100.0]	0.04522 [0.04498, 0.04545]
0.01	-0.02	10,000	939/1000; 93.9 [92.2, 95.2]	0/1000; 0.0 [0.0, 0.4]	-0.03171 [-0.03217, -0.03124]
0.01	-0.02	40,000	956/1000; 95.6 [94.1, 96.7]	0/1000; 0.0 [0.0, 0.4]	-0.02627 [-0.02649, -0.02605]
0.01	+0.00	10,000	943/1000; 94.3 [92.7, 95.6]	57/1000; 5.7 [4.4, 7.3]	-0.01127 [-0.01172, -0.01082]
0.01	+0.00	40,000	946/1000; 94.6 [93.0, 95.8]	54/1000; 5.4 [4.2, 7.0]	-0.00595 [-0.00617, -0.00572]
0.01	+0.05	10,000	946/1000; 94.6 [93.0, 95.8]	1000/1000; 100.0 [99.6, 100.0]	0.03915 [0.03874, 0.03957]
0.01	+0.05	40,000	959/1000; 95.9 [94.5, 97.0]	1000/1000; 100.0 [99.6, 100.0]	0.04386 [0.04364, 0.04408]
0.10	-0.02	10,000	952/1000; 95.2 [93.7, 96.4]	0/1000; 0.0 [0.0, 0.4]	-0.03314 [-0.03361, -0.03266]
0.10	-0.02	40,000	951/1000; 95.1 [93.6, 96.3]	0/1000; 0.0 [0.0, 0.4]	-0.02656 [-0.02680, -0.02631]
0.10	+0.00	10,000	956/1000; 95.6 [94.1, 96.7]	44/1000; 4.4 [3.3, 5.9]	-0.01312 [-0.01361, -0.01263]
0.10	+0.00	40,000	960/1000; 96.0 [94.6, 97.0]	40/1000; 4.0 [3.0, 5.4]	-0.00676 [-0.00700, -0.00651]
0.10	+0.05	10,000	950/1000; 95.0 [93.5, 96.2]	1000/1000; 100.0 [99.6, 100.0]	0.03697 [0.03648, 0.03746]
0.10	+0.05	40,000	946/1000; 94.6 [93.0, 95.8]	1000/1000; 100.0 [99.6, 100.0]	0.04396 [0.04372, 0.04420]

Table 27: KCMC on all 18 conditions; 1,000 datasets per row. Coverage targets the sharp M_H endpoint. Percentages have 95% Wilson intervals and LCB means have 95% Monte Carlo intervals.

ϵ	M_H	N	Coverage: count; %	Certificates: count; %	Mean LCB
0.00	-0.02	10,000	884/1000; 88.4 [86.3, 90.2]	1/1000; 0.1 [0.0, 0.6]	-0.03142 [-0.03201, -0.03082]
0.00	-0.02	40,000	881/1000; 88.1 [85.9, 90.0]	0/1000; 0.0 [0.0, 0.4]	-0.02585 [-0.02615, -0.02555]
0.00	+0.00	10,000	884/1000; 88.4 [86.3, 90.2]	116/1000; 11.6 [9.8, 13.7]	-0.01164 [-0.01224, -0.01103]
0.00	+0.00	40,000	861/1000; 86.1 [83.8, 88.1]	139/1000; 13.9 [11.9, 16.2]	-0.00560 [-0.00590, -0.00530]
0.00	+0.05	10,000	872/1000; 87.2 [85.0, 89.1]	1000/1000; 100.0 [99.6, 100.0]	0.03986 [0.03930, 0.04043]
0.00	+0.05	40,000	848/1000; 84.8 [82.4, 86.9]	1000/1000; 100.0 [99.6, 100.0]	0.04506 [0.04477, 0.04535]
0.01	-0.02	10,000	919/1000; 91.9 [90.0, 93.4]	2/1000; 0.2 [0.1, 0.7]	-0.03296 [-0.03355, -0.03237]
0.01	-0.02	40,000	940/1000; 94.0 [92.4, 95.3]	0/1000; 0.0 [0.0, 0.4]	-0.02721 [-0.02750, -0.02693]
0.01	+0.00	10,000	919/1000; 91.9 [90.0, 93.4]	81/1000; 8.1 [6.6, 10.0]	-0.01300 [-0.01358, -0.01242]
0.01	+0.00	40,000	934/1000; 93.4 [91.7, 94.8]	66/1000; 6.6 [5.2, 8.3]	-0.00696 [-0.00725, -0.00666]
0.01	+0.05	10,000	915/1000; 91.5 [89.6, 93.1]	1000/1000; 100.0 [99.6, 100.0]	0.03802 [0.03748, 0.03855]
0.01	+0.05	40,000	931/1000; 93.1 [91.4, 94.5]	1000/1000; 100.0 [99.6, 100.0]	0.04329 [0.04302, 0.04357]
0.10	-0.02	10,000	943/1000; 94.3 [92.7, 95.6]	0/1000; 0.0 [0.0, 0.4]	-0.03554 [-0.03614, -0.03494]
0.10	-0.02	40,000	960/1000; 96.0 [94.6, 97.0]	0/1000; 0.0 [0.0, 0.4]	-0.02800 [-0.02828, -0.02771]
0.10	+0.00	10,000	948/1000; 94.8 [93.2, 96.0]	52/1000; 5.2 [4.0, 6.8]	-0.01513 [-0.01570, -0.01456]
0.10	+0.00	40,000	944/1000; 94.4 [92.8, 95.7]	56/1000; 5.6 [4.3, 7.2]	-0.00774 [-0.00804, -0.00744]
0.10	+0.05	10,000	958/1000; 95.8 [94.4, 96.9]	1000/1000; 100.0 [99.6, 100.0]	0.03544 [0.03490, 0.03599]
0.10	+0.05	40,000	959/1000; 95.9 [94.5, 97.0]	1000/1000; 100.0 [99.6, 100.0]	0.04284 [0.04258, 0.04311]

Table 28: False certification at the robust boundary $M_H = 0$. Each entry has denominator 1,000 and a 95% Wilson interval in percent. The direct protected weak dual makes 0/1000 ([0.0,0.4]%) in all six conditions for both interfaces. QB uses five-fold cross-fitting; KCMC uses the bootstrap adaptation.

ϵ	N	**RAC (ours; coarse)**	DVDS	QB (five-fold cross-fitted)	KCMC (bootstrap adaptation)
0.00	10,000	4/1000 0.4 [0.2, 1.0]	60/1000 6.0 [4.7, 7.6]	88/1000 8.8 [7.2, 10.7]	116/1000 11.6 [9.8, 13.7]
0.00	40,000	10/1000 1.0 [0.5, 1.8]	58/1000 5.8 [4.5, 7.4]	98/1000 9.8 [8.1, 11.8]	139/1000 13.9 [11.9, 16.2]
0.01	10,000	2/1000 0.2 [0.1, 0.7]	33/1000 3.3 [2.4, 4.6]	57/1000 5.7 [4.4, 7.3]	81/1000 8.1 [6.6, 10.0]
0.01	40,000	5/1000 0.5 [0.2, 1.2]	30/1000 3.0 [2.1, 4.3]	54/1000 5.4 [4.2, 7.0]	66/1000 6.6 [5.2, 8.3]
0.10	10,000	0/1000 0.0 [0.0, 0.4]	52/1000 5.2 [4.0, 6.8]	44/1000 4.4 [3.3, 5.9]	52/1000 5.2 [4.0, 6.8]
0.10	40,000	0/1000 0.0 [0.0, 0.4]	48/1000 4.8 [3.6, 6.3]	40/1000 4.0 [3.0, 5.4]	56/1000 5.6 [4.3, 7.2]

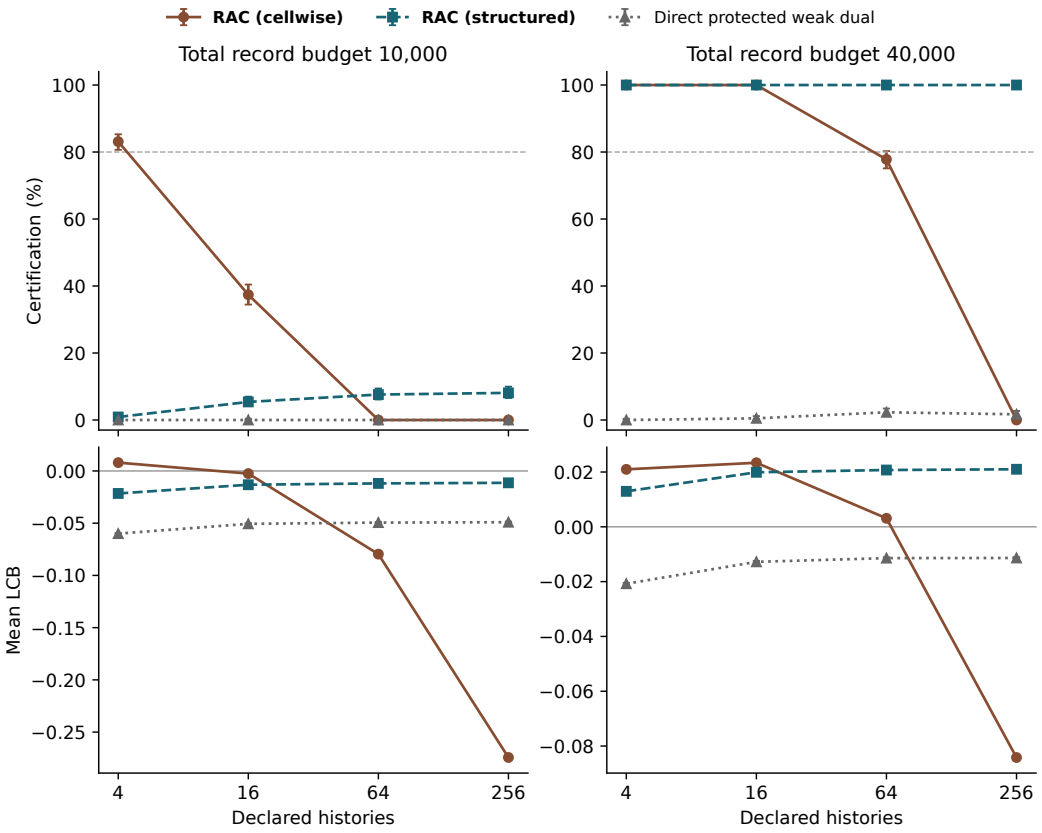

Figure 3: Structured versus cellwise nuisance calibration at $M_H = 0.05$. Columns: $N = 10{,}000$ (left), $N = 40{,}000$ (right). Rows: certification rate (top) and mean LCB (bottom); higher is better. Each point uses 1,000 paired datasets; 95% Wilson rate and Monte Carlo mean intervals may be hidden by markers. Both RAC variants use the same two-cell interface; the direct comparator has equal complete protection. Structured RAC sustains certification at large supports with the larger budget.

Table 29: **RAC (ours): complete structured scaling results.** Each procedure–condition cell has 1,000 datasets and coverage 1,000/1,000 (Wilson 95% $[99.6, 100.0]\%$); nominal target 95%. Mean LCB brackets are 95% Monte Carlo intervals. Certificate brackets are Wilson intervals in percent, with numerator out of 1,000. At $M_H = 0$, these are false certificates. The direct comparator uses structured propensity calibration and the equal-protection allocation in Appendix C.7.

K	Procedure	Certificates /1,000 (Wilson, %)	Mean LCB [95% MC]
$M_H = 0.05$, total budget $N = 10,000$			
4	**RAC (cellwise)**	831 [80.7, 85.3]	0.007994 [0.007437, 0.008552]
	RAC (structured)	9 [0.5, 1.7]	-0.021538 [-0.022062, -0.021014]
	Direct protected weak dual	0 [0.0, 0.4]	-0.060002 [-0.060627, -0.059376]
16	**RAC (cellwise)**	374 [34.5, 40.4]	-0.002506 [-0.003020, -0.001992]
	RAC (structured)	54 [4.2, 7.0]	-0.013183 [-0.013696, -0.012671]
	Direct protected weak dual	0 [0.0, 0.4]	-0.050611 [-0.051222, -0.050000]
64	**RAC (cellwise)**	0 [0.0, 0.4]	-0.079608 [-0.080159, -0.079057]
	RAC (structured)	76 [6.1, 9.4]	-0.011900 [-0.012423, -0.011377]
	Direct protected weak dual	0 [0.0, 0.4]	-0.049442 [-0.050081, -0.048803]
256	**RAC (cellwise)**	0 [0.0, 0.4]	-0.274306 [-0.275053, -0.273559]
	RAC (structured)	81 [6.6, 10.0]	-0.011346 [-0.011858, -0.010835]
	Direct protected weak dual	0 [0.0, 0.4]	-0.049055 [-0.049674, -0.048436]
$M_H = 0.05$, total budget $N = 40,000$			
4	**RAC (cellwise)**	1000 [99.6, 100.0]	0.020972 [0.020681, 0.021264]
	RAC (structured)	1000 [99.6, 100.0]	0.012947 [0.012687, 0.013207]
	Direct protected weak dual	0 [0.0, 0.4]	-0.020783 [-0.021127, -0.020438]
16	**RAC (cellwise)**	1000 [99.6, 100.0]	0.023375 [0.023110, 0.023641]
	RAC (structured)	1000 [99.6, 100.0]	0.019903 [0.019644, 0.020162]
	Direct protected weak dual	5 [0.2, 1.2]	-0.012784 [-0.013110, -0.012457]
64	**RAC (cellwise)**	778 [75.1, 80.3]	0.003104 [0.002840, 0.003368]
	RAC (structured)	1000 [99.6, 100.0]	0.020727 [0.020466, 0.020989]
	Direct protected weak dual	23 [1.5, 3.4]	-0.011426 [-0.011773, -0.011079]
256	**RAC (cellwise)**	0 [0.0, 0.4]	-0.084191 [-0.084464, -0.083918]
	RAC (structured)	1000 [99.6, 100.0]	0.020999 [0.020744, 0.021255]
	Direct protected weak dual	17 [1.1, 2.7]	-0.011370 [-0.011704, -0.011036]
$M_H = 0.00$, total budget $N = 10,000$			
4	**RAC (cellwise)**	0 [0.0, 0.4]	-0.042896 [-0.043510, -0.042282]
	RAC (structured)	0 [0.0, 0.4]	-0.072298 [-0.072877, -0.071720]
	Direct protected weak dual	0 [0.0, 0.4]	-0.112049 [-0.112761, -0.111337]

Continued on next page

Table 29 continued

K	Procedure	Certificates /1,000 (Wilson, %)	Mean LCB [95% MC]
16	**RAC (cellwise)**	0 [0.0, 0.4]	-0.053059 [-0.053611, -0.052506]
	RAC (structured)	0 [0.0, 0.4]	-0.063779 [-0.064330, -0.063228]
	Direct protected weak dual	0 [0.0, 0.4]	-0.102365 [-0.103075, -0.101655]
64	**RAC (cellwise)**	0 [0.0, 0.4]	-0.130337 [-0.130900, -0.129774]
	RAC (structured)	0 [0.0, 0.4]	-0.062636 [-0.063177, -0.062094]
	Direct protected weak dual	0 [0.0, 0.4]	-0.101446 [-0.102110, -0.100783]
256	**RAC (cellwise)**	0 [0.0, 0.4]	-0.325785 [-0.326589, -0.324980]
	RAC (structured)	0 [0.0, 0.4]	-0.062477 [-0.063041, -0.061913]
	Direct protected weak dual	0 [0.0, 0.4]	-0.101059 [-0.101739, -0.100379]
$M_H = 0.00$, total budget $N = 40,000$			
4	**RAC (cellwise)**	0 [0.0, 0.4]	-0.029630 [-0.029944, -0.029317]
	RAC (structured)	0 [0.0, 0.4]	-0.037605 [-0.037885, -0.037326]
	Direct protected weak dual	0 [0.0, 0.4]	-0.071844 [-0.072197, -0.071490]
16	**RAC (cellwise)**	0 [0.0, 0.4]	-0.027285 [-0.027560, -0.027010]
	RAC (structured)	0 [0.0, 0.4]	-0.030767 [-0.031032, -0.030502]
	Direct protected weak dual	0 [0.0, 0.4]	-0.063542 [-0.063903, -0.063182]
64	**RAC (cellwise)**	0 [0.0, 0.4]	-0.047301 [-0.047578, -0.047025]
	RAC (structured)	0 [0.0, 0.4]	-0.029667 [-0.029945, -0.029389]
	Direct protected weak dual	0 [0.0, 0.4]	-0.062654 [-0.063018, -0.062290]
256	**RAC (cellwise)**	0 [0.0, 0.4]	-0.134711 [-0.134999, -0.134424]
	RAC (structured)	0 [0.0, 0.4]	-0.029526 [-0.029798, -0.029254]
	Direct protected weak dual	0 [0.0, 0.4]	-0.062500 [-0.062870, -0.062131]

D.6.2 NUISANCE UNCERTAINTY, SUPPORT AND SIGNED DEFICITS

Table 30 separates prediction widths from score variance and complete support range. The reported widths are full prediction-interval widths after physical intersection. Structured widths remain similar as these balanced grids are refined; cellwise widths grow with sparse fine cells and their K-fold allocation. The comparison evaluates jointly fitted points and envelopes under the exact-model assumption; unrestricted cellwise calibration uses a larger model class.

Table 30: **RAC (ours): structured scaling diagnostics at $M_H = 0.05$.** Means of 1,000 datasets per procedure–condition cell. Widths are full prediction-interval widths; range encloses all declared score categories. The direct comparator uses propensity envelopes only.

K	Procedure	Propensity full width	Outcome full width	Score variance	Score range
Total budget $N = 10,000$					
4	**RAC (cellwise)**	0.0827	0.1692	0.1607	1.1400
	RAC (structured)	0.2417	0.4868	0.1529	1.1042
	Direct protected weak dual	0.2387	—	0.1930	1.5454
16	**RAC (cellwise)**	0.1887	0.3754	0.1501	1.1546
	RAC (structured)	0.2325	0.4618	0.1489	1.1208
	Direct protected weak dual	0.2295	—	0.1901	1.5453
64	**RAC (cellwise)**	0.4076	0.7187	0.1566	1.5320
	RAC (structured)	0.2315	0.4597	0.1485	1.1278
	Direct protected weak dual	0.2284	—	0.1897	1.5457
256	**RAC (cellwise)**	0.7576	0.9625	0.2002	4.3442
	RAC (structured)	0.2312	0.4600	0.1483	1.1303
	Direct protected weak dual	0.2282	—	0.1896	1.5452
Total budget $N = 40,000$					
4	**RAC (cellwise)**	0.0412	0.0842	0.1826	1.2901
	RAC (structured)	0.1267	0.2565	0.1521	1.0835
	Direct protected weak dual	0.1252	—	0.2202	1.6827
16	**RAC (cellwise)**	0.0943	0.1911	0.1553	1.1467
	RAC (structured)	0.1220	0.2458	0.1483	1.0874
	Direct protected weak dual	0.1206	—	0.2158	1.6520
64	**RAC (cellwise)**	0.2086	0.4115	0.1497	1.2070
	RAC (structured)	0.1215	0.2447	0.1481	1.0905
	Direct protected weak dual	0.1201	—	0.2152	1.6500
256	**RAC (cellwise)**	0.4397	0.7565	0.1566	1.7806
	RAC (structured)	0.1214	0.2444	0.1478	1.0890
	Direct protected weak dual	0.1200	—	0.2150	1.6472

Table 31 reports the actual treated fitting counts and empirical feature leverage. Observations at other histories inform shared coefficients, allowing structured prediction bands at histories with no treated fitting records. The empirical outcome design still depends on which fitting observations are treated, as required in (25).

At $K = 256, N = 40{,}000, M_H = 0.05$, structural loss is the same across procedures, while the mean net nuisance deficit is 0.113862 for cellwise RAC and 0.008937 for structured RAC (Table 32). Their paired LCB difference decomposes into $+0.000009$ from additional fitting regret, $+0.104924$ from net nuisance deficit and $+0.000257$ from evaluation gap. Thus the gain primarily comes from cheaper nuisance protection under the model assumption.

Table 31: **RAC (ours; structured): treated information and leverage** at $M_H = 0.05$; 1,000 datasets per row. Counts and maximum leverages are means across datasets. Empty-cell counts are datasets with at least one empty treated fitting history.

K	Total budget N	Mean N_T	Empty treated-cell datasets /1,000	Max propensity leverage	Max outcome leverage
4	10,000	1124.2	0	0.000824	0.004470
4	40,000	4503.6	0	0.000203	0.001112
16	10,000	1126.4	0	0.001375	0.007128
16	40,000	4503.6	0	0.000338	0.001773
64	10,000	1125.9	0	0.001571	0.008103
64	40,000	4501.9	0	0.000387	0.002016
256	10,000	1123.9	975	0.001626	0.008429
256	40,000	4500.7	0	0.000401	0.002086

Table 32: **RAC (ours): primary structured-scaling shortfall.** $K = 256, N = 40{,}000, M_H = 0.05$; means of 1,000 datasets per procedure. Terms follow (11); J_Z is within-interface fitting regret. RAC's net nuisance deficit is correction minus signed bias. For the direct protected weak dual, the net deficit comes from expanding inverse-weight boxes; dashes mark RAC-specific correction and bias terms.

Procedure	L_ϕ	J_Z	Applied correction	Signed bias	Net nuisance deficit	Evaluation gap
RAC (cellwise)	0.004900	0.006078	0.113922	6.0e-05	0.113862	0.009352
RAC (structured)	0.004900	0.006069	0.008937	-5.1e-08	0.008937	0.009094
Direct protected weak dual	0.004900	0.006037	—	—	0.039298	0.011135

Remaining shortfall at a fixed interface. The exact Z normalization losses are $0.015075000, 0.006295791, 0.005145304, 0.004899641$ for $K = 4, 16, 64, 256$. For these binary laws at $\lambda = 2$, with $v_h = P(H = h)(1 - e_h)(\mu_h - 2/3)$, the population formula is

$$L_\phi = (\lambda - \lambda^{-1}) \left[\sum_h (v_h)_+ - \sum_z \left(\sum_{h:\phi(h)=z} v_h \right)_+ \right].$$

On nonabstention, subtracting structural loss from (11) gives

$$M_Z - \widehat{L} = M_H - \widehat{L} - L_\phi = J_Z + \left[D_Z(\widehat{a}) - B - \mathbb{E}(g_{\widehat{a}} \mid \mathcal{F}) \right] + \left[\mathbb{E}(g_{\widehat{a}} \mid \mathcal{F}) - \widehat{L} \right]. \tag{43}$$

For the direct comparator, its protected direct-score mean replaces the augmented mean. The signed nuisance and evaluation terms are retained on every dataset; their nonnegativity holds on their respective coverage events. Table 33 reports remaining shortfall for both budgets and all three procedures on the same paired datasets.

At $N = 40{,}000$, from $K = 4$ to 256, structured raw LCB rises from 0.012947029 to 0.020999356, structural loss falls from 0.015075000 to 0.004899641, and remaining shortfall rises from 0.021977971 to 0.024101003. Hence

$$0.008052327 = 0.010175359 - 0.002123032.$$

Support grids change the population law and design. Same-support paired gains quantify reduced nuisance-protection cost under correct models; across supports, structural loss also changes.

Table 33: **RAC (ours): remaining scaling shortfall** $M_Z - \widehat{L}$, $M_H = 0.05$; means [95% Monte Carlo intervals] from 1,000 datasets per entry. Structural loss is removed; signed nuisance and evaluation terms remain.

K	N	**RAC (cellwise)**	**RAC (structured)**	Direct protected weak dual
4	10,000	0.02693 [0.02637, 0.02749]	0.05646 [0.05594, 0.05699]	0.09493 [0.09430, 0.09555]
16	10,000	0.04621 [0.04570, 0.04672]	0.05689 [0.05638, 0.05740]	0.09431 [0.09370, 0.09493]
64	10,000	0.12446 [0.12391, 0.12501]	0.05675 [0.05623, 0.05728]	0.09430 [0.09366, 0.09494]
256	10,000	0.31941 [0.31866, 0.32015]	0.05645 [0.05593, 0.05696]	0.09416 [0.09354, 0.09477]
4	40,000	0.01395 [0.01366, 0.01424]	0.02198 [0.02172, 0.02224]	0.05571 [0.05536, 0.05605]
16	40,000	0.02033 [0.02006, 0.02059]	0.02380 [0.02354, 0.02406]	0.05649 [0.05616, 0.05681]
64	40,000	0.04175 [0.04149, 0.04201]	0.02413 [0.02387, 0.02439]	0.05628 [0.05593, 0.05663]
256	40,000	0.12929 [0.12902, 0.12956]	0.02410 [0.02385, 0.02436]	0.05647 [0.05614, 0.05680]

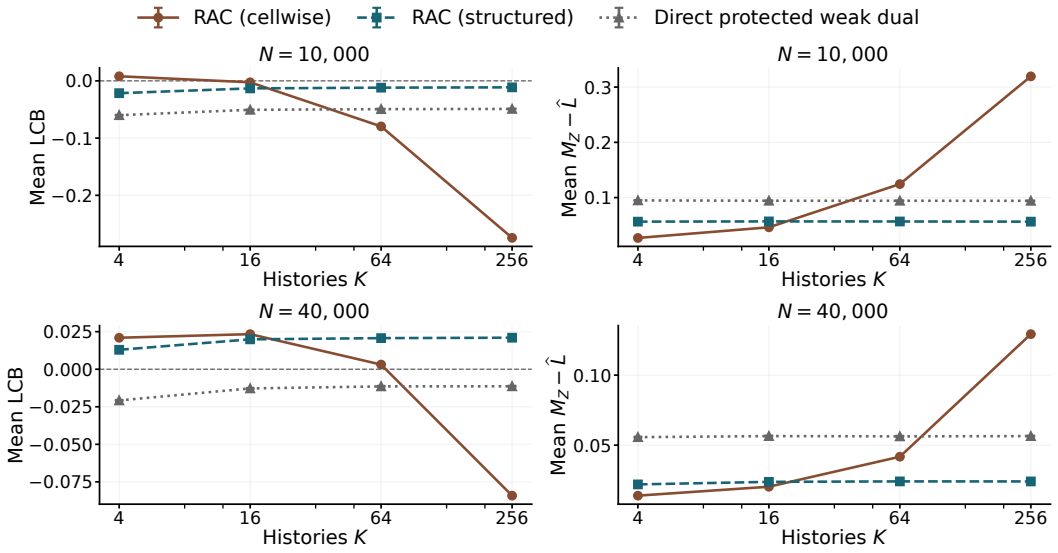

Figure 4: How does nuisance calibration affect certificate strength? Enlarged Figure 2: one budget per row and metric per column. Higher mean LCBs and lower remaining shortfall are better. $M_H = 0.05$; 1,000 paired datasets per point, 95% Monte Carlo intervals. Across supports, the population law and structural loss also change.