BEST-OF-N CAN RECOVER AFTER DECLINE: SHARP REVERSALS AND AUDIT COMPLEXITY

Anonymous authors
Paper under double-blind review

ABSTRACT

A decline in Best-of-N utility at small N does not rule out recovery at larger N, even when calibration is smooth and the early utilities are known exactly. For every finite horizon $H \geq 64$, we construct a fixed generator–verifier law whose utility first peaks at $N = 2$, decreases at $N = 3$, and then exceeds that first peak by at least 0.15 at $N = H$. A task-conditional rank transform explains the phenomenon: for each N, expected utility is the average of one calibration profile under a Beta$(N, 1)$ density. The ordered sign changes of the profile derivative then give a sharp bound on the number of utility reversals, whereas smoothness alone gives no finite bound. We next study how much evaluation is needed to tell whether utility recovers. For an explicit smooth recovery/no-recovery pair, passive evaluation requires on the order of H utility labels at fixed confidence, while direct winner evaluation at the two target breadths needs a number of labels independent of H for a constant utility gap. More generally, we derive minimax bounds that quantify the roles of distribution overlap, generation cost, and calibration smoothness. Thus an early decline, by itself, is not a reliable stopping criterion for Best-of-N scaling.

1 INTRODUCTION

Best-of-N samples N candidates and returns the one with the largest verifier score, making it a basic form of test-time scaling (Snell et al., 2024). Under nested sampling, the maximum verifier score can only improve as N grows, but the true utility of the selected answer need not. This raises a practical question: if utility falls once, is it safe to stop increasing N? Without further assumptions, it is not. We study which structural properties limit such reversals and what evaluation data are needed to detect a later recovery.

We study the population curve

$$J_N = \mathbb{E}[U_{\hat{i}_N}], \qquad \hat{i}_N \in \arg\max_{i \leq N} S_i,$$

for a fixed generator, verifier, bounded utility law, and task distribution. Our first result is a finite-horizon counterexample to stopping after the first decline. For every $H \geq 64$, there is a smooth calibration profile whose utility first peaks at $N = 2$, falls by at least 0.01 at $N = 3$, and then rises to at least 0.15 above that first peak by $N = H$ (Theorem 3.1). Thus a first-decline stopping rule can miss a constant improvement even when the early population utilities are known exactly.

The key coordinate is score rank *within the task*. For continuous conditional score CDFs F_z, write $X = F_Z(S)$ and $\psi_z(x) = \mathbb{E}[U \mid X = x, Z = z]$. The winning rank has density $p_N(x) = Nx^{N-1}$, yielding

$$J_N = N \int_0^1 \bar{\psi}(x) x^{N-1}\, dx, \qquad \bar{\psi}(x) = \mathbb{E}_Z[\psi_Z(x)]. \tag{1}$$

For each N, (1) averages the same calibration profile against a Beta$(N, 1)$ kernel, which concentrates increasingly near rank one as N grows. If verification quality varies across high-rank regions, the curve can therefore decline and later recover. Smoothness alone does not limit the number of such reversals; what matters is the ordered sign structure of the profile derivative.

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.

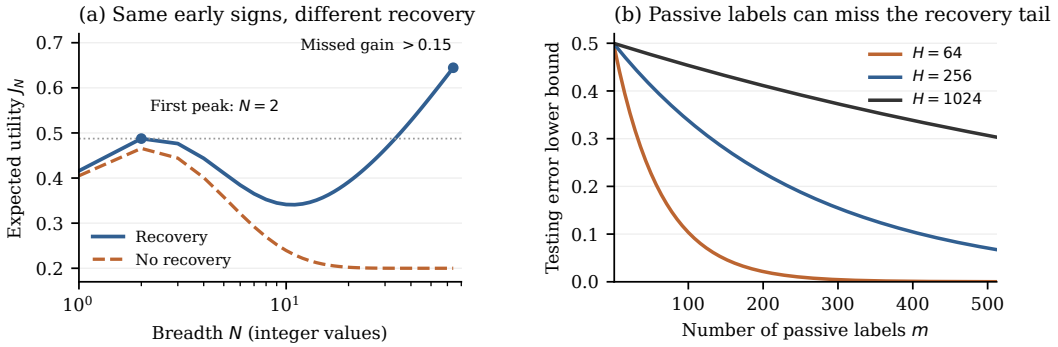

Figure 1: **A first decline does not rule out recovery at larger N.** Left: exact utility curves for the step-profile pair used in Theorems 3.1 and 5.1, with $H = 64$. Both curves first peak at $N = 2$ and then decline, but only the profile with a reliable extreme tail later recovers. The smooth constructions preserve the stated gaps. Right: the passive-evaluation lower bound $\frac{1}{2}(1 - 1/H)^m$ at three horizons. The right panel plots a lower bound on testing error, not the error of a particular test.

Contributions.

1. **Finite-horizon recovery.** We give a smooth bounded construction with an explicit first decline and a missed-utility gap of at least 0.15, uniformly over the horizon H (Theorem 3.1).
2. **Sharp reversal bounds.** Both discrete utility gains and derivatives of the continuous-breadth curve can be written as transforms of the same Stieltjes measure. Its ordered sign complexity sharply bounds the number of reversals, and every finite bound is attained by a smooth profile. We also give a positive counterexample showing why strict TP_2 alone does not imply a one-extremum shape (Theorem 4.1 and Proposition 4.3).
3. **Statistical evidence for recovery.** For an explicit pair of smooth profiles, passive evaluation needs on the order of H labels to distinguish recovery from no recovery, while direct winner evaluation removes this dependence from the label count. We then derive general minimax bounds for design mismatch, generation–label tradeoffs, and smoothness-dependent resolution (Theorems 5.1, 6.1, 6.4 and 6.5).

Our results complement Dorner et al. (2026), who show that early Best-of-N behavior need not determine limiting accuracy. Here the emphasis is finite-horizon behavior: we quantify the loss from stopping after the first decline, give sharp reversal bounds, and study the sample complexity of testing for recovery. The analysis combines order statistics, moment theory, total positivity, and importance sampling. Section B also gives a complete characterization of realizable utility sequences.

2 Model and Conditional Quantile Geometry

Let Z take values in a standard Borel space with a fixed probability law. Conditional on $Z = z$, let $(U_i, S_i)_{i\geq 1}$ be one infinite i.i.d. sequence from a fixed measurable probability kernel, with $U_i \in [0, 1]$ and scalar $S_i \in \mathbb{R}$. Breadth $N \in \mathbb{N} = \{1, 2, \ldots\}$ uses its first N candidates; none of these laws changes with N. Assume first that F_z is continuous for almost every task and that selection compares candidates only within that task. Put $X_i = F_z(S_i)$ and take a jointly measurable version of $\psi_z(x) = \mathbb{E}[U \mid X = x, Z = z]$. Since $X \mid Z$ is uniform, X is independent of Z, and $\bar{\psi}$ is defined Lebesgue-a.e.

Lemma 2.1 (Task-conditional rank transform)**.** *For every $N \geq 1$,*

$$J_N = N \int_0^1 \bar{\psi}(x) x^{N-1}\, dx. \tag{2}$$

Conversely, every measurable $\psi : [0, 1] \to [0, 1]$ is realized by $X \sim \mathrm{Unif}(0, 1)$ and $U \mid X = x \sim \mathrm{Bernoulli}(\psi(x))$.

The reduction is unchanged by any jointly measurable, task-dependent strictly increasing transformation of the score. It identifies only the average profile, not the distribution of taskwise risks (Theorem A.1). This task conditioning matters: pooling heterogeneous tasks through one marginal CDF and then applying the homogeneous Beta formula can reverse the inferred direction even when the verifier ranks perfectly within each task (Proposition A.3). From here on we write $\psi = \bar{\psi}$.

Remark 2.2 (Ties). For discrete scores, the distributional transform $X = F_Z(S^-) + V\{F_Z(S) - F_Z(S^-)\}$ with an independent uniform V is conditionally uniform and implements independent randomized tie breaking (Brockwell, 2007). All results extend under that rule.

Let $(\mathsf{D}a)_n = a_{n+1} - a_n$. A sequence is completely monotone when $(-1)^k(\mathsf{D}^k a)_n \geq 0$ for all $n, k \geq 0$.

3 Finite-Horizon Recovery After a First Decline

For a finite horizon H, define the first-decline choice

$$\tau_H = \min\{n \in \{1, \ldots, H-1\} : J_{n+1} < J_n\},$$

with $\tau_H = H$ if the set is empty. This idealized stopping rule observes the exact population utilities in sequence and keeps the breadth immediately before the first decline. We measure its loss by $\max_{1 \leq n \leq H} J_n - J_{\tau_H}$, without subtracting compute cost from utility.

Theorem 3.1 (A constant improvement missed within a finite horizon)**.** *For every integer $H \geq 64$, there is a $C^\infty[0,1]$ profile ψ_H taking values in $[0.1, 0.9]$, realizable by a fixed single-task law with Bernoulli utility, such that*

$$J_2 - J_1 \geq 0.06, \qquad J_2 - J_3 \geq 0.01, \qquad J_H - J_2 \geq 0.15. \tag{3}$$

Consequently $\tau_H = 2$ and $\max_{n \leq H} J_n - J_{\tau_H} \geq 0.15$. The profile has exactly two ordered derivative sign changes, and its complete discrete utility curve has exactly two reversals.

Construction and mechanism. Put $b_H = 1 - 1/H$. Start with utility levels $0.1, 0.9, 0.2, 0.9$ on the successive intervals cut at $0.4, 0.75, b_H$. Direct integration gives

$$J_N^+ = 0.9 - 0.8(0.4)^N + 0.7(0.75)^N - 0.7 b_H^N. \tag{4}$$

The final interval has passive rank mass only $1/H$, yet the winner of an H-candidate pool lands there with probability $1 - b_H^H \geq 1 - e^{-1}$. For example, at $H = 64$, $J_2^+ \approx 0.4875$ while $J_{64}^+ \approx 0.6445$. We smooth each jump over a monotone transition of width $1/(10^4 H)$; this changes every J_N with $N \leq H$ by at most 0.00022, so the margins above remain intact (Section E). The generator–verifier law is fixed as N varies. The construction itself may depend on H, however, and its derivative norms are not uniform in H; therefore it does not imply the same constant-gap result over a fixed Hölder ball.

4 Sharp Reversal Bounds and the Limit of TP_2

Define the continuous-breadth extension $J(t) = t \int_0^1 \psi(x) x^{t-1} dx$ for real $t > 0$, so $J(N) = J_N$. For a finite or infinite sequence, let S^- count sign changes after zeros are removed, with $\mathsf{S}^-(0) = 0$. For a function, take the supremum of this count over finite increasing samples. A *discrete reversal* is a sign change of $G_N = J_{N+1} - J_N$; a *continuous reversal* is a sign change of $J'(t)$. Their counts need not coincide, although $G_N = \int_N^{N+1} J'(t) dt$. Flat increments and stationary points without a sign change do not count as reversals.

For a finite signed Borel measure ν on $[0,1]$, define its *ordered sign complexity* $\mathrm{osc}(\nu)$ as the least $k \geq 0$ for which

$$\nu = \sigma \sum_{j=0}^{k} (-1)^j \mu_j, \qquad \sigma \in \{-1, 1\}, \tag{5}$$

where the μ_j are nonzero finite positive measures carried by ordered compact intervals $I_0 \leq I_1 \leq \cdots \leq I_k$ with $\sup I_j \leq \inf I_{j+1}$. Set $\mathrm{osc}(0) = 0$ and $\mathrm{osc}(\nu) = \infty$ when no finite decomposition

exists. Equal boundary atoms are combined before counting. For a finite atomic measure this is the sign variation of its ordered nonzero weights; for $d\nu = f(x)dx$ with finitely many essential sign intervals, it is the usual number of sign transitions. For a BV profile we use its right-continuous representative on $[0, 1)$ and set $\psi(1) = \psi(1-)$. Its Stieltjes measure on $(0, 1)$ is extended by zero mass at both endpoints. Thus endpoint point values, which do not affect J, do not enter $\mathrm{osc}(d\psi)$. Interior jumps and singular components are allowed.

Theorem 4.1 (Sharp reversal bound for bounded-variation profiles)**.** *Let $\psi : [0, 1] \to \mathbb{R}$ have bounded variation, with the representative and Stieltjes convention just specified. For every integer $N \geq 1$ and every real $t > 0$,*

$$G_N := J_{N+1} - J_N = \int_{[0,1]} x^N (1 - x)\, d\psi(x), \tag{6}$$

$$J'(t) = \int_{[0,1]} x^t (-\log x)\, d\psi(x), \qquad t > 0, \tag{7}$$

where the endpoint weights are interpreted by continuity. Consequently,

$$\mathsf{S}^-((G_N)_{N\geq 1}) \leq \mathrm{osc}(d\psi), \qquad \mathsf{S}^-(J') \leq \mathrm{osc}(d\psi). \tag{8}$$

Each bound is sharp over bounded-utility profiles: for every integer $k \geq 0$ there is a C^∞ profile with $0 < \psi < 1$ and $\mathrm{osc}(d\psi) = k$ whose complete gain sequence has exactly k sign changes. A possibly different smooth profile with $\mathrm{osc}(d\psi) = k$ makes J' have exactly k sign changes.

The quantity controlling reversals is ordered sign complexity, not smoothness: even a C^∞ calibration profile can have arbitrarily many reversals. The key step is that (6)–(7) express both objects as transforms of one signed measure that does not depend on N or t. We can then apply total positivity of all orders, rather than TP_2 alone. Approximating the positive components in (5) by ordered atoms reduces the argument to variation diminution for a totally nonnegative generalized-Vandermonde matrix; weak convergence yields (8). The formulation also covers jumps and singular components of the calibration profile. Sharpness is obtained by prescribing alternating signs at finitely many points and smoothing the resulting atomic construction without changing those signs.

Corollary 4.2 (A structural one-reversal condition)**.** *If $\mathrm{osc}(d\psi) \leq 1$, then the integer curve changes monotonicity at most once and the continuous breadth curve has at most one strict interior extremum. If $d\psi \geq 0$, then all gains are nonnegative and completely monotone; in particular they are nonincreasing and log-convex.*

4.1 A TP_2 COUNTEREXAMPLE AND ITS SCOPE

Theorem 1 in the NeurIPS proceedings version and arXiv v2 of Khalaf et al. (2025) states that, on common support, strict TP_2 and a continuous score increasing in the outcome imply that the expectation of every bounded nonnegative reward is monotone or has a unique interior extremum. The following reward satisfies those stated hypotheses but not the stated shape conclusion.

Proposition 4.3 (Positive reward with infinitely many reversals)**.** *Fix $0 < \varepsilon < 1$, $\rho > 0$, and $\omega > 0$, and let*

$$r(x) = \frac{1}{2} + \frac{\varepsilon}{2}(1 - x)^\rho \cos\big(\omega \log(1 - x)\big), \qquad r(1) := \frac{1}{2}. \tag{9}$$

For $p_t(x) = tx^{t-1}$ and $J(t) = \int_0^1 p_t(x) r(x) dx$, the family p_t has common support $(0, 1)$, is strictly TP_2, and has score $t^{-1} + \log x$ continuous and strictly increasing in x. Nevertheless $J(t)$ has infinitely many local extrema and $(J_{N+1} - J_N)$ changes sign infinitely often. If $\rho > 1$, then $r \in C^1[0, 1]$, $r'(0^+) = -\varepsilon\rho/2 < 0$, and $r'(1^-) = 0$. Thus it satisfies the reward-derivative sign inequalities displayed in the "immediate decline" row of the cited Corollary 3, while still exhibiting infinitely many reversals.

Indeed,

$$J(t) = \frac{1}{2} + \frac{\varepsilon}{2} \mathrm{Re}\left[\Gamma(1 + \rho + i\omega) \frac{\Gamma(t + 1)}{\Gamma(t + 1 + \rho + i\omega)}\right], \tag{10}$$

so the leading correction is a decaying sinusoid in $\log t$. The issue is not the score parameterization. Differentiation gives $J'(t) = \int p_t(x) h_t(x) dx$ with $h_t = r\, \partial_t \log p_t$, but now the signed input h_t

varies with the same parameter t as the kernel. Classical variation diminution does not give a uniform reversal bound in that form because it requires a fixed signed input. For the Beta family, Theorem 4.1 restores a fixed-measure representation; a one-reversal conclusion then requires the additional condition $\mathrm{osc}(d\psi) \le 1$. This does not contradict Theorem 4.1, since the derivative dr in the present example has infinitely many ordered sign changes and hence $\mathrm{osc}(dr) = \infty$.

The cited work's toy reward $r(x) \propto x^p(1-x)$, $p > 0$, has a derivative with exactly one ordered sign change, so Corollary 4.2 recovers the intended one-extremum conclusion for that example. Appendix F.1 verifies the assumptions and the continuous, integer, and endpoint-sign conclusions.

5 Statistical Evidence for Recovery

The region responsible for recovery can carry very little probability under passive sampling. To quantify the resulting difficulty, we cast recovery detection as a statistical decision problem. For a profile ψ, let $P_{\psi,K}$ denote the law of one observed pair (X, Y), where $X \sim p_K$ and $Y \mid X \sim \mathrm{Bernoulli}(\psi(X))$. Repeating this experiment corresponds to labeling the winner of an independent K-candidate pool while observing its rank; $K = 1$ reduces to passive sampling. Only the selected candidate receives a utility label.

Theorem 5.1 (Testing recovery: passive versus targeted evaluation)**.** *For every $H \ge 64$, there are smooth profiles $\psi_{0,H}, \psi_{1,H} \in [0.1, 0.9]$ with the first two gain margins in* (3)*, but opposite signs of $\Delta_i = J_H(\psi_{i,H}) - J_2(\psi_{i,H})$:*

$$\Delta_0 \le -0.15, \qquad \Delta_1 \ge 0.15.$$

They agree on $[0, 1 - 1/H]$. For every $m \ge 1$ and $1 \le K \le H$, the minimax error over this known pair, allowing randomized tests based on m i.i.d. observations, satisfies

$$\tfrac{1}{2}(1 - 1/H)^{Km} \le \inf_T \max_{i \in \{0,1\}} P^{\otimes m}_{\psi_{i,H},K}(T \ne i) \le e^{-mK/(24H)}. \tag{11}$$

In particular, passive labels give error at least $3/8$ when $m \le H/4$. Error at most $\delta < 1/2$ requires

$$m \ge \frac{H-1}{K} \log \frac{1}{2\delta}. \tag{12}$$

The sufficient count $m \ge 24(H/K)\log(1/\delta)$ matches this order for $\delta \le 1/4$. This pair-specific upper test uses known ranks and the specified profiles. Separately, for any bounded-utility law with $|J_H - J_2| \ge d > 0$, the sign of the difference of independent winner-label means, using r labels at each breadth, has error at most $\exp(-rd^2/2)$. Thus $r \ge 2d^{-2}\log(1/\delta)$ suffices with no rank knowledge. For the constructed pair, $d = 0.15$ is uniform in H.

Why targeted labels help. The lower bound arises because the two profiles induce exactly the same data below rank $1 - 1/H$. A matching test uses only labeled ranks above the smooth transition, where the Bernoulli means are 0.2 and 0.9. If ranks are unavailable, one can instead compare the utilities of independently sampled winners at the two breadths. This reduces the number of labels, but not the generation cost: the rank-free comparison uses $r(H+2)$ generated candidates. The next section extends this design-mismatch calculation beyond the two-profile construction.

6 Minimax Rates for Auditing

Fix a target breadth $N \ge 1$ and a known design density q on $[0, 1]$, independent of the unknown profile, with $p_N(x)dx \ll q(x)dx$, where $p_N(x) = Nx^{N-1}$. Observe $m \ge 1$ *i.i.d. pairs* $(X_j, Y_j) \sim P$, with $X_j \sim q$, $Y_j \in [0, 1]$, and $\mathbb{E}_P[Y_j \mid X_j = x] = \psi(x)$. Let $\mathfrak{P}(\psi, q)$ contain all such one-pair laws; Bernoulli labels are a subclass. The parameter class and squared-error minimax risk are

$$\Psi_{\mathrm{all}} = \{\psi : [0,1] \to [0,1] \text{ measurable}\},$$

$$\mathcal{R}_{m,N}(\Psi; q) = \inf_{\widehat{J}} \sup_{\psi \in \Psi} \sup_{P \in \mathfrak{P}(\psi,q)} \mathbb{E}_{P^{\otimes m}}(\widehat{J} - J_N(\psi))^2.$$

The estimator ranges over measurable functions of the observed pairs and the known N, q. Write $w = p_N/q$ q-a.e. and set it to zero on null sets where $q = 0$. Define

$$A_N(q) = \mathbb{E}_q w^2 = 1 + \chi^2(p_N \| q), \qquad B_N(q) = \|w\|_{L^\infty(q)}. \tag{13}$$

Both may be infinite. For heterogeneous tasks, this experiment assumes the design preserves the target task distribution and applies the same rank density q within each task. It assumes observed ranks; raw-score implementations with task-matched reference data are treated separately in Section H. Throughout, $a \wedge b = \min\{a, b\}$ and $\asymp$ denotes bounds up to positive constants, not equality of minimax constants.

Theorem 6.1 (General nonasymptotic minimax bounds)**.** *For all integers $m, N \geq 1$, if $B = B_N(q) < \infty$, then $A = A_N(q) \leq B < \infty$ and universal constants $0 < c_* < C_* < \infty$ satisfy*

$$c_* \left(\frac{A}{m} \wedge \frac{A^2}{B^2} \right) \leq \mathcal{R}_{m,N}(\Psi_{\mathrm{all}}; q) \leq C_* \left(1 \wedge \frac{A}{m} \right). \tag{14}$$

One may take $c_* = 1/256$ and $C_* = 1$; no optimality of these constants is asserted. The upper bound only requires $A < \infty$ and is attained by the clipped importance estimator

$$\widehat{J}^{\mathrm{IW}}_{N,q} = \Pi_{[0,1]} \left\{ \frac{1}{m} \sum_{j=1}^{m} \frac{p_N(X_j)}{q(X_j)} Y_j \right\}. \tag{15}$$

The upper bound follows from the variance of the importance estimator with weight $w = p_N/q$. For the lower bound, use the interior perturbation $h = w/B$: its effect on the target is A/B, while its squared size under q is A/B^2. Scaling this perturbation until the m-sample Bernoulli KL remains constant yields the two terms in (14). Here A_N measures average overlap between the target and design distributions, while B_N controls the largest importance weight. The upper and lower bounds match under the following leverage condition.

Corollary 6.2 (Matching minimax rate under bounded leverage)**.** *If $A_N(q) < \infty$ and $B_N(q) \leq \kappa A_N(q)$ for a fixed $\kappa \geq 1$, then, uniformly over integers $m, N \geq 1$ and designs satisfying this condition,*

$$\mathcal{R}_{m,N}(\Psi_{\mathrm{all}}; q) \asymp_{\kappa} 1 \wedge \frac{A_N(q)}{m}. \tag{16}$$

The condition is sufficient for matching rates. For infinite overlap, Proposition H.3 gives a truncation bound using the full weight tail.

Corollary 6.3 (Optimal targeting and passive impossibility)**.** *For every admissible density q, $A_N(q) \geq 1$, with equality exactly when $q = p_N$ a.e. For fixed target breadth N, labeling actual N-way winners has risk $\asymp 1/m$, and no design density improves this order uniformly over the profile class. Thus winner targeting is label-minimax rate-optimal; uniqueness is asserted only for minimizing A_N. Passive auditing has $q \equiv 1$, $A_N = N^2/(2N-1)$ and $B_N = N$, so*

$$\mathcal{R}_{m,N}(\Psi_{\mathrm{all}}; 1) \asymp 1 \wedge \frac{N}{m}. \tag{17}$$

The passive order remains unchanged with arbitrarily many unlabeled scores. Winner auditing is implemented by the mean of m independently sampled winners' utility labels; it requires neither their ranks nor knowledge of the task CDFs. Its advantage is a factor of order N in label complexity, with generation accounted for separately below.

Theorem 6.4 (Rate-optimal partial targeting)**.** *A K-pool audit labels the winner among $K \leq N$ conditionally i.i.d. candidates, independently for each labeled observation. In the known-rank experiment, its density is $q = p_K$. Let generation cost $c_g > 0$, label cost $c_y \geq 0$, total budget C, and assume the feasible set below is nonempty:*

$$\mathcal{K}_C = \{K \in \{1, \ldots, N\} : c_y + K c_g \leq C\}, \qquad m_K = \left\lfloor \frac{C}{c_y + K c_g} \right\rfloor.$$

For every $K \in \mathcal{K}_C$,

$$\mathcal{R}^{\mathrm{cost}}_{C,N}(K) := \mathcal{R}_{m_K,N}(\Psi_{\mathrm{all}}; p_K) \asymp 1 \wedge \frac{N^2}{m_K K(2N-K)} \asymp 1 \wedge \frac{N^2(c_y + K c_g)}{CK(2N-K)}. \tag{18}$$

Let $\lambda = c_y/c_g$, $K_{\max} = \max \mathcal{K}_C$, and

$$K^{\dagger}_{\mathrm{cont}} = \Pi_{[1,K_{\max}]} \left\{ \sqrt{\lambda^2 + 2\lambda N} - \lambda \right\}. \tag{19}$$

Let $\bar{K}$ be the better, under the last proxy in (18)*, of the at most two feasible integers adjacent to $K^{\dagger}_{\mathrm{cont}}$. Then a universal constant C_0 satisfies*

$$\mathcal{R}^{\mathrm{cost}}_{C,N}(\bar{K}) \le C_0 \inf_{K \in \mathcal{K}_C} \mathcal{R}^{\mathrm{cost}}_{C,N}(K). \tag{20}$$

The guarantee is only for the stated fixed-K designs with $K \le N$; it does not claim optimality over arbitrary adaptive sampling policies. Rank/CDF acquisition costs are also excluded. Because m_K is integer-valued, (19) *should be read as a continuous proxy rather than an exact finite-budget optimizer; see Remark I.1.*

For $s > 0$, put $r = \lceil s \rceil - 1$ and $\alpha = s - r \in (0, 1]$. The Hölder ball $\mathcal{H}^s(L)$ consists of functions with continuous derivatives through order r, each bounded in absolute value by L (including order zero), and $|f^{(r)}(x) - f^{(r)}(y)| \le L|x - y|^{\alpha}$. Endpoint derivatives are one-sided. Set $\Psi_s(L) = \mathcal{H}^s(L) \cap \Psi_{\mathrm{all}}$. Thus smoothness is imposed globally on the average calibration profile in rank space.

Theorem 6.5 (Sharp smoothness transition)**.** *For every fixed $s > 0$ and $L > 0$, uniformly over all integers $m, N \ge 1$,*

$$\mathcal{R}_{m,N}(\Psi_s(L); 1) \asymp_{s,L} \frac{N}{m} \wedge m^{-2s/(2s+1)}. \tag{21}$$

Constants depend only on s, L. The phase boundary is $N \asymp m^{1/(2s+1)}$: below it direct importance weighting is rate-optimal; above it a clipped boundary local-polynomial estimator attains the pointwise nonparametric rate. The passive estimator uses known ranks and known s to choose the regime and bandwidth. For $s > 1$ it retains the lower-order Taylor terms; smoothness alone does not make $J_N - \psi(1)$ of order N^{-s}.

Why the transition occurs. Direct importance weighting has variance of order N/m. A local polynomial fit over an endpoint interval of width $h \ge 1/N$ instead has squared bias $O(h^{2s})$ and variance $O((mh)^{-1})$. Balancing these terms with $h = m^{-1/(2s+1)}$ gives the second rate. The matching lower bound uses a Bernoulli pair that differs by a smooth bump of width $\max\{1/N, m^{-1/(2s+1)}\}$ (Section I). The phase transition occurs when the effective width $1/N$ of the target becomes smaller than the nonparametric resolution scale.

Audit implications. For a fixed sufficiently small target error, passive auditing needs order N more labels than labeling actual winners. Equivalently, in the nonsaturated regime $m \gtrsim N$, the two MSEs differ by a factor of order N. Partial targeting lets one trade this label inefficiency against generation cost. Existing passive labels can still be reused through the raw-score procedures in Section H; Theorem 6.5 does not claim the same smoothness rate when ranks are unknown.

Practical evaluation. For a prespecified breadth N, a direct estimator is simple: sample tasks from the deployment distribution, generate an independent group of N candidates for each task, label the selected winner, and average the labels. This estimates J_N without estimating task CDFs. If only passive labels are available, Theorem 6.1 quantifies the loss from design mismatch; the smoothness result additionally requires the stated rank-space model. Comparisons across several breadths require simultaneous uncertainty bounds (Corollary H.4). An observed local peak alone does not justify stopping globally. Throughout, selection uses the same fixed maximum-score rule.

7 Related Work

Classical ingredients. The Beta density of the winning rank is a standard order-statistic identity, and the selected utility is a concomitant of the maximizing score (David and Galambos, 1974; Wang et al., 2006). Expected-maxima theory asks a different question: it studies the expectation of the selected variable itself (Papadatos, 2017). Our realizability criterion is a specialization of the classical bounded-density Hausdorff moment problem, with Bernstein–Widder giving the continuous-breadth analogue (Hausdorff, 1921; Widder, 1941). Corollary D.4 uses the resulting inequalities only as a finite-sample rejection certificate.

We also use classical total positivity and variation-diminishing theory (Karlin, 1968; Pinkus, 2010). Our use of these tools starts from the Stieltjes representation in Theorem 4.1, which gives a single

fixed signed measure for both discrete gains and continuous-breadth derivatives, including profiles with jumps or singular components. Regular variation is used for the endpoint asymptotics in Section G (Bingham et al., 1987).

Audit theory and off-policy evaluation. The dependence of off-policy evaluation on distribution overlap is well established. Wang et al. (2017, Theorem 1) prove an agnostic contextual-bandit lower bound under moment and sample-size conditions, while Ma et al. (2022, Theorem 1) give an all-sample-size minimax characterization for finite-action bandits using switching. Our setting is a continuous rank-space problem in which the target density changes with N. Specializing the overlap calculation to this structure yields the winner-versus-passive label gap, the fixed-K generation–label tradeoff, and the Hölder phase transition. The smooth case uses standard local-polynomial and testing arguments (Gao, 2018; Stone, 1982; Tsybakov, 2009); Duan et al. (2020) study a different design-dependent setting based on linear function approximation. Ranked-set sampling provides a related cost-design perspective (Nahhas et al., 2002). Appendix H.1 lists the observation-model differences explicitly.

Best-of-N and imperfect verification. Existing work analyzes policy KL and win rate (Beirami et al., 2025), coverage and policy optimality (Huang et al., 2025; Sriraman and Block, 2026), and inference-aware training (Balashankar et al., 2025). Generation complexity for producing correct answers (Huang et al., 2026) is distinct from the true-utility label complexity for auditing a fixed selected distribution studied here. Reward overoptimization and pessimistic mitigation motivate the same problem (Gao et al., 2023; Skalse et al., 2022; Yu et al., 2026).

Dorner et al. (2026) analyze Best-of-N accuracy through ROC geometry, including monotonicity and high-compute limits. Their ground-truth utility is binary, although the verifier score need not be. The calibration-profile view used here describes the same score–utility relationship in a different coordinate system. Because every bounded profile is realizable with Bernoulli utility, allowing general bounded utility does not enlarge the set of attainable curves. Dorner et al.'s Proposition 7 already shows that two systems can agree approximately on any finite prefix yet have incompatible limiting accuracies. Our finite-horizon result instead puts an explicit constant loss on stopping after the first decline; the reversal theorem and recovery-testing result address the shape and statistical detectability of that behavior. Section B separately characterizes the full set of realizable utility sequences. Conditional-calibration work also warns against marginal pooling (Detommaso et al., 2024); Proposition A.3 gives a concrete example in which pooling reverses the conclusion.

Our TP_2 comparison is limited to the shape theorem as stated by Khalaf et al. (2025). The counterexample isolates the difficulty: after differentiation, the signed input depends on the same parameter as the kernel. Theorem 4.1 gives a fixed-measure statement for the Beta family, with ordered sign complexity as the additional condition. We do not evaluate the empirical findings or mitigation methods in that work.

8 Limitations and Scope

The analysis assumes a fixed generator–verifier law and conditionally i.i.d. candidates. The utility curve averages over the task distribution and therefore does not control worst-task behavior. Correlated or adaptive rollouts, changing generators, joint breadth–depth allocation, and unbounded utilities fall outside this model. The smooth witnesses used for the finite-horizon counterexample do not have Hölder norms bounded uniformly in H, and we do not claim that the constructed pattern occurs in deployed models. The recovery-testing result addresses the sign of $J_H - J_2$ in a fixed audit experiment, not the optimality of arbitrary adaptive stopping rules. Likewise, the minimax results estimate one fixed J_N under squared error; the finite breadth-selection corollary gives simultaneous confidence control but does not solve the minimax problem of learning the optimal breadth. The audit designs preserve the task distribution. General and smooth rank-space results assume observed ranks, with raw-score reference-data costs treated separately. Without bounded leverage, Theorem 6.1 gives upper and lower bounds rather than an exact rate. Adaptation to unknown smoothness and smooth auditing with unknown task CDFs remain open.

9 Conclusion

A first decline in Best-of-N utility does not establish that larger N is worse. After transforming scores to task-conditional ranks, the full utility curve becomes a family of Beta-weighted averages of one calibration profile. The ordered sign complexity of that profile gives a sharp bound on reversals, while smoothness alone does not. The same representation also clarifies the evaluation problem: passive labels can require order H samples to detect a recovery near breadth H, whereas labels collected from the selected winners can remove this dependence in label count, at the cost of additional generation. Thus an observed local peak and evidence that no later recovery exists should be treated as separate statistical questions.

Reproducibility Statement

Complete proofs, the explicit smooth recovery pair, least-favorable constructions, the sharpness argument, and the analytic counterexample appear in Sections A and D to I. The figures and numerical checks are reproducible with the accompanying `plot_theory.py`, `verify_math.py`, and `verify_smooth_recovery.py`; Section F.1 gives the exact finite product. The figures evaluate analytic formulas and rate proxies. The quadrature script reports numerical error estimates; the smooth-profile perturbation bound is proved in Section E.

AI Usage Disclosure

The research idea originated with the author, who designed the core mathematical framework, including the task-conditional rank and Beta-moment representation, calibration profiles, and minimax auditing formulation. Most theorem and lemma statements were developed by the author. Large language models (LLMs) assisted with the formulation and refinement of some theorems and lemmas and with the construction and revision of some proof arguments, including the finite-horizon recovery and recovery-testing results. They were also used for repeated proof audits and to identify potential gaps. The author performed independent verification in addition to these LLM-assisted checks. LLMs were used extensively for related-work research and to generate code for numerical checks and figures. Most of the manuscript text was drafted using LLMs; the author proofread, corrected, and revised it. LLMs also assisted with LaTeX preparation. The author takes responsibility for the final text, mathematical claims, proofs, references, and code.

REFERENCES

Charlie Snell, Jaehoon Lee, Kelvin Xu, and Aviral Kumar. Scaling LLM test-time compute optimally can be more effective than scaling model parameters. *arXiv:2408.03314*, 2024.

Ahmad Beirami, Alekh Agarwal, Jonathan Berant, Alexander D'Amour, Jacob Eisenstein, Chirag Nagpal, and Ananda Theertha Suresh. Theoretical guarantees on the Best-of-N alignment policy. In *Proceedings of the 42nd International Conference on Machine Learning*, volume 267 of *Proceedings of Machine Learning Research*, pages 3580–3602, 2025.

Ananth Balashankar, Ziteng Sun, Jonathan Berant, Jacob Eisenstein, Michael Collins, Adrian Hutter, Jong Lee, Chirag Nagpal, Flavien Prost, Aradhana Sinha, Ananda Theertha Suresh, and Ahmad Beirami. InfAlign: Inference-aware language model alignment. In *Proceedings of the 42nd International Conference on Machine Learning*, volume 267 of *Proceedings of Machine Learning Research*, pages 2646–2672, 2025.

Audrey Huang, Adam Block, Qinghua Liu, Nan Jiang, Akshay Krishnamurthy, and Dylan J. Foster. Is Best-of-N the best of them? Coverage, scaling, and optimality in inference-time alignment. In *Proceedings of the 42nd International Conference on Machine Learning*, volume 267 of *Proceedings of Machine Learning Research*, pages 25075–25126, 2025.

Baihe Huang, Shanda Li, Tianhao Wu, Yiming Yang, Ameet Talwalkar, Kannan Ramchandran, Michael I. Jordan, and Jiantao Jiao. Sample complexity and representation ability of test-time scaling paradigms. In *International Conference on Learning Representations*, 2026.

Florian E. Dorner, Yatong Chen, André F. Cruz, and Fanny Yang. ROC-n-reroll: How verifier imperfection affects test-time scaling. In *International Conference on Learning Representations*, 2026.

Hadi Khalaf, Claudio Mayrink Verdun, Alex Oesterling, Himabindu Lakkaraju, and Flavio du Pin Calmon. Inference-time reward hacking in large language models. In *Advances in Neural Information Processing Systems 38 (NeurIPS 2025)*, 2025.

Zhuohao Yu, Zhiwei Steven Wu, and Adam Block. From curiosity to caution: Mitigating reward hacking for Best-of-N with pessimism. In *International Conference on Learning Representations*, 2026.

Ved Sriraman and Adam Block. Revisiting the (Sub)Optimality of Best-of-N for Inference-Time Alignment. In *Proceedings of the Thirty-Ninth Conference on Learning Theory*, volume 336 of *Proceedings of Machine Learning Research*, pages 5980–6028, 2026.

Leo Gao, John Schulman, and Jacob Hilton. Scaling laws for reward model overoptimization. In *Proceedings of the 40th International Conference on Machine Learning*, volume 202 of *Proceedings of Machine Learning Research*, pages 10835–10866, 2023.

Joar Skalse, Nikolaus H. R. Howe, Dmitrii Krasheninnikov, and David Krueger. Defining and characterizing reward hacking. In *Advances in Neural Information Processing Systems*, 2022.

Yaqi Duan, Zeyu Jia, and Mengdi Wang. Minimax-optimal off-policy evaluation with linear function approximation. In *Proceedings of the 37th International Conference on Machine Learning*, volume 119 of *Proceedings of Machine Learning Research*, pages 2701–2709, 2020.

Nickos Papadatos. On sequences of expected maxima and expected ranges. *Journal of Applied Probability*, 54(4):1144–1166, 2017.

H. A. David and J. Galambos. The asymptotic theory of concomitants of order statistics. *Journal of Applied Probability*, 11(4):762–770, 1974.

Xinlei Wang, Lynne Stokes, Johan Lim, and Min Chen. Concomitants of multivariate order statistics with application to judgment poststratification. *Journal of the American Statistical Association*, 101(476):1693–1704, 2006.

R. W. Nahhas, D. A. Wolfe, and H. Chen. Ranked set sampling: Cost and optimal set size. *Biometrics*, 58(4):964–971, 2002.

Wayne Yuan Gao. Minimax linear estimation at a boundary point. *Journal of Multivariate Analysis*, 165:262–269, 2018.

Felix Hausdorff. Summationsmethoden und Momentfolgen. I. *Mathematische Zeitschrift*, 9:74–109, 1921.

David V. Widder. *The Laplace Transform*. Princeton University Press, 1941.

Samuel Karlin. *Total Positivity, Volume I*. Stanford University Press, 1968.

Allan Pinkus. *Totally Positive Matrices*. Cambridge University Press, 2010.

Nicholas H. Bingham, Charles M. Goldie, and Jozef L. Teugels. *Regular Variation*. Cambridge University Press, 1987.

Alexandre B. Tsybakov. *Introduction to Nonparametric Estimation*. Springer, 2009.

Charles J. Stone. Optimal global rates of convergence for nonparametric regression. *The Annals of Statistics*, 10(4):1040–1053, 1982.

Gianluca Detommaso, Martin Andres Bertran, Riccardo Fogliato, and Aaron Roth. Multicalibration for confidence scoring in LLMs. In *Proceedings of the 41st International Conference on Machine Learning*, volume 235 of *Proceedings of Machine Learning Research*, pages 10624–10641, 2024.

Anthony E. Brockwell. Universal residuals: A multivariate transformation. *Statistics & Probability Letters*, 77(14):1473–1478, 2007.

Jean-Bernard Lasserre. The moment problem with bounded density. *arXiv:math/0607463v2*, 2007.

Yu-Xiang Wang, Alekh Agarwal, and Miroslav Dudík. Optimal and adaptive off-policy evaluation in contextual bandits. In *Proceedings of the 34th International Conference on Machine Learning*, volume 70 of *Proceedings of Machine Learning Research*, pages 3589–3597, 2017.

Cong Ma, Banghua Zhu, Jiantao Jiao, and Martin J. Wainwright. Minimax off-policy evaluation for multi-armed bandits. *IEEE Transactions on Information Theory*, 68(8):5314–5339, 2022.

A Proofs for Conditional Quantile Geometry

Theorem A.1 (Task-conditional reduction and invariance)**.** *Equation* (2) *holds for arbitrary task heterogeneity under conditional i.i.d. sampling. It is invariant under every jointly measurable task-dependent transformation $S \mapsto g_Z(S)$ with g_z strictly increasing. Every result below therefore applies to the average conditional profile $\bar{\psi}$, but the curve identifies neither the distribution of task profiles nor taskwise risks beyond their average.*

Remark A.2 (Non-identification under task heterogeneity). Even the complete sequence $(J_N)_{N\geq 1}$ cannot recover worst-task behavior. For two equally likely tasks, any bounded perturbation h for which $\bar{\psi} \pm h \in [0,1]$ produces task profiles $\psi_1 = \bar{\psi} + h$ and $\psi_2 = \bar{\psi} - h$ with the same population curve as the homogeneous profile $\bar{\psi}$, while the two taskwise curves generally differ. Thus task-conditional normalization is necessary for the reduction, but the resulting curve still describes only an average over tasks and does not guarantee taskwise control.

Marginal score ranks need not preserve within-task selection geometry. The next proposition concerns a specific inferential error: pooling heterogeneous raw scores, applying one marginal CDF, estimating one pooled calibration profile, and then inserting that profile into the homogeneous Beta formula. It is not a claim about procedures that compute ranks separately within each task or prompt.

Proposition A.3 (Pooled normalization can reverse the conclusion)**.** *There are two equally likely tasks whose verifier ranks utility perfectly within each task and whose true curve is*

$$J_N = \frac{3N+1}{4(N+1)}, \tag{22}$$

which is strictly increasing, while applying the homogeneous transform to the marginal score quantile gives

$$\widetilde{J}_N = \frac{N-1}{2(N+1)} + 2^{-N}. \tag{23}$$

In particular $J_2 > J_1$ but $\widetilde{J}_2 < \widetilde{J}_1$.

Proof of Theorem A.1. Conditionally on $Z = z$, the probability integral transform gives $X_i = F_z(S_i) \overset{\text{i.i.d.}}{\sim} \mathrm{Unif}(0,1)$ and preserves the within-task score order. Therefore the winning conditional quantile has density $p_N(x) = Nx^{N-1}$ and

$$\mathbb{E}[U_{\hat{i}_N} \mid Z = z] = N \int_0^1 \psi_z(x) x^{N-1}\, dx.$$

Taking expectation over Z and applying Fubini's theorem proves Equation (2). If S is replaced by $g_z(S)$ with g_z strictly increasing, then

$$F_{g_z(S)|Z=z}(g_z(S)) = F_z(S)$$

almost surely, so X, ψ_z, and the curve are unchanged. □

Proof of Proposition A.3. Let $Z \in \{A, B\}$ be uniform and let $R \sim \mathrm{Unif}(0,1)$ conditional on either task. In task A, set

$$S = R, \qquad U = \frac{1+R}{2};$$

in task B, set

$$S = 2 + R, \qquad U = \frac{R}{2}.$$

In both tasks, score and utility are strictly increasing functions of R, hence the verifier ranks perfectly. If $R_{(N)}$ is the maximum of N uniforms, then $\mathbb{E}R_{(N)} = N/(N+1)$, and averaging the two conditional utilities gives

$$J_N = \frac{1}{2}\left(\frac{1}{2} + \frac{1}{2}\frac{N}{N+1}\right) + \frac{1}{2}\left(\frac{1}{2}\frac{N}{N+1}\right) = \frac{3N+1}{4(N+1)}.$$

This is strictly increasing.

The marginal score supports of the two tasks are disjoint. If $u = F_S(S)$ is the marginal score quantile, then the pooled conditional utility is

$$\widetilde{\psi}(u) = \begin{cases} \frac{1}{2} + u, & 0 \le u < \frac{1}{2}, \\ u - \frac{1}{2}, & \frac{1}{2} \le u \le 1. \end{cases}$$

Applying the homogeneous transform incorrectly yields

$$\begin{aligned} \widetilde{J}_N &= N \int_0^{1/2} \left(\frac{1}{2} + u\right) u^{N-1}\, du + N \int_{1/2}^1 \left(u - \frac{1}{2}\right) u^{N-1}\, du \\ &= \frac{N-1}{2(N+1)} + 2^{-N}. \end{aligned}$$

Substitution of $N = 1, 2$ gives the claimed reversal. □

Justification of Remark 2.2. Conditionally on $S = s$ and $Z = z$, X is uniform on the interval $[F_z(s^-), F_z(s)]$. These intervals partition $[0, 1]$ up to endpoints, and their lengths equal the corresponding score masses, so $X \mid Z = z$ is uniform. For two candidates, $S_i < S_j$ implies $X_i < X_j$; when $S_i = S_j$, their order is determined by the independent uniforms. Hence maximizing X is exactly randomized tie breaking among maximal scores. □

B Complete Feasibility and Identification

Let $\mathfrak{J}_1$ be the set of all realizable sequences $(J_N)_{N\ge1}$.

Theorem B.1 (Complete realizability)**.** *For a real sequence $j = (j_N)_{N\ge1}$ define*

$$a_n = \frac{j_{n+1}}{n+1}, \qquad b_n = \frac{1 - j_{n+1}}{n+1}.$$

Then $j \in \mathfrak{J}_1$ if and only if both (a_n) and (b_n) are completely monotone. Equivalently, for every $n, k \ge 0$,

$$(-1)^k \mathsf{D}^k \left(\frac{j_{n+1}}{n+1}\right) \ge 0, \qquad (-1)^k \mathsf{D}^k \left(\frac{1 - j_{n+1}}{n+1}\right) \ge 0. \tag{24}$$

When feasible, the representing profile is unique Lebesgue-a.e., and the same curve is realizable with $U \in \{0, 1\}$.

This is the classical bounded-density Hausdorff criterion applied to Best-of-N (Hausdorff, 1921); equivalently, it is the specialization of Lasserre (2007, Theorem 3(c)) to $[0, 1]$ with Lebesgue reference measure and density bound one. Necessity follows by integrating $x^n(1 - x)^k$ against $\psi\, dx$ and $(1 - \psi)\, dx$. Conversely, moment determinacy forces the two representing measures to sum to Lebesgue measure, so the first has a density in $[0, 1]$. The task-conditional reduction therefore turns these classical moment inequalities into constraints on the entire utility curve. A finite subset provides only necessary conditions: the rejection rule in Corollary D.4 can detect violations in noisy estimates, but failure to reject establishes neither finite-prefix extendibility nor feasibility of the unknown infinite curve.

The continuous-breadth Bernstein–Widder analogue and the extreme-point geometry are stated and proved in Section D and Corollary D.2; neither adds assumptions to Theorem B.1.

C Endpoint Geometry and Power Laws

The same kernel localizes near one: if $Q_N = 1 - X_{(N)}$, then $\mathbb{P}(Q_N > y) = (1 - y)^N \le e^{-Ny}$ for $0 \le y \le 1$. When the endpoint limit exists, put $u_\infty = \psi(1-)$ and $g(y) = u_\infty - \psi(1 - y)$.

Theorem C.1 (Abelian transfer)**.** *Assume ψ is bounded, the endpoint limit $u_\infty = \psi(1-)$ exists, and, for $a, \gamma > 0$ and a slowly varying ℓ,*

$$g(y) \sim a y^\gamma \ell(1/y) \qquad (y \downarrow 0).$$

Then

$$u_\infty - J_N \sim a\Gamma(1 + \gamma) N^{-\gamma} \ell(N). \tag{25}$$

Under these assumptions, the breadth exponent and leading constant follow from the endpoint gap. This is an Abelian implication from the profile to the curve; no Tauberian converse is asserted. The proof in Section G uses Potter bounds and uniform integrability, not weak convergence alone.

D Continuous Breadth and Convex Geometry

For real $t > 0$ define $J(t) = t \int_0^1 \psi(x) x^{t-1} dx$.

Theorem D.1 (Continuous-breadth realizability)**.** *A function $j : (0, \infty) \to \mathbb{R}$ has this representation for a measurable $0 \le \psi \le 1$ if and only if*

$$t \mapsto j(t)/t, \qquad t \mapsto (1 - j(t))/t \tag{26}$$

are completely monotone. The profile is unique almost everywhere.

Corollary D.2 (Extreme points)**.** *The feasible set $\mathfrak{J}_1$ is convex. Its extreme points are exactly the curves generated by $\psi = \mathbf{1}_A$ for measurable $A \subseteq [0, 1]$, modulo null sets.*

Corollary D.3 (Identification versus audit ill-conditioning)**.** *If two bounded profiles generate the same J_N for every integer $N \ge 1$, then they agree almost everywhere. Nevertheless, passive estimation of a single high-breadth moment requires order N labels for a fixed sufficiently small root-MSE without further regularity (Corollary 6.3).*

Proof of Lemma 2.1. Condition on $Z = z$ and let $X_i = F_z(S_i)$. Continuity of the conditional score law gives i.i.d. uniform X_i, and the winner has quantile $M_N = \max_i X_i$ with density $N x^{N-1}$. Conditional on $(X_{1:N}, Z = z)$, independence gives

$$\mathbb{E}[U_{\hat{\imath}_N} \mid X_{1:N}, Z = z] = \psi_z(M_N).$$

Integrating and then averaging over Z yields Equation (2). Conversely, for any measurable $\psi \in [0, 1]$, draw X uniformly and then U conditionally Bernoulli with success probability $\psi(X)$. Taking $S = X$ realizes the profile and preserves ranking. □

Proof of Theorem B.1. Suppose first that $j_N = J_N(\psi)$ for a measurable $0 \le \psi \le 1$. Then

$$a_n = \int_0^1 x^n \psi(x)\, dx, \qquad b_n = \int_0^1 x^n (1 - \psi(x))\, dx.$$

Therefore, for every $n, k \ge 0$,

$$(-1)^k \mathsf{D}^k a_n = \int_0^1 x^n (1 - x)^k \psi(x)\, dx \ge 0,$$
$$(-1)^k \mathsf{D}^k b_n = \int_0^1 x^n (1 - x)^k (1 - \psi(x))\, dx \ge 0.$$

Thus both sequences are completely monotone.

Conversely, assume (a_n) and (b_n) are completely monotone. Hausdorff's moment theorem (Hausdorff, 1921) gives finite positive Borel measures μ and ν on $[0, 1]$ such that

$$a_n = \int_0^1 x^n\, d\mu(x), \qquad b_n = \int_0^1 x^n\, d\nu(x).$$

Their sum satisfies

$$\int_0^1 x^n\, d(\mu + \nu)(x) = a_n + b_n = \frac{1}{n+1} = \int_0^1 x^n\, dx.$$

The Hausdorff moment problem on the compact interval is determinate, so

$$\mu + \nu = \lambda,$$

where λ is Lebesgue measure. Hence $0 \le \mu \le \lambda$. By the Radon–Nikodym theorem, $d\mu = \psi\, d\lambda$ for a measurable $0 \le \psi \le 1$. Then

$$j_{n+1} = (n + 1) a_n = (n + 1) \int_0^1 x^n \psi(x)\, dx,$$

which is exactly the transform. Uniqueness follows from determinacy: if two profiles have the same curve, their induced measures have the same moments and are equal. Binary realization follows from Lemma 2.1. □

Corollary D.4 (Finite-sample feasibility rejection certificate)**.** *Suppose estimates $\widehat{j}_1, \dots, \widehat{j}_H$ satisfy, with probability at least $1-\delta$,*

$$|\widehat{j}_N - j_N| \le \epsilon_N, \qquad N = 1, \dots, H. \tag{27}$$

Set $\widehat{a}_n = \widehat{j}_{n+1}/(n+1)$, $\widehat{b}_n = (1-\widehat{j}_{n+1})/(n+1)$, and for $n+k+1 \le H$ define

$$r_{n,k} = \sum_{r=0}^{k} \binom{k}{r} \frac{\epsilon_{n+r+1}}{n+r+1}. \tag{28}$$

If, for any admissible (n,k),

$$(-1)^k \mathsf{D}^k \widehat{a}_n < -r_{n,k} \quad \textit{or} \quad (-1)^k \mathsf{D}^k \widehat{b}_n < -r_{n,k}, \tag{29}$$

then the null hypothesis that $(j_N)_{N\ge1}$ is a realizable bounded-utility curve is rejected with type-I error at most δ. In particular, if $\widehat{j}_N$ is the average of m_N independent bounded labels from actual N-way winners, one may take

$$\epsilon_N = \sqrt{\frac{\log(2H/\delta)}{2m_N}}. \tag{30}$$

Failure to reject only certifies that no violation is resolved at the stated confidence radii; it does not establish feasibility of the unknown full sequence.

Proof of Corollary D.4. On the simultaneous event (27), the kth forward-difference identity and the triangle inequality give

$$\left|\mathsf{D}^k \widehat{a}_n - \mathsf{D}^k a_n\right| \le \sum_{r=0}^{k} \binom{k}{r} \frac{|\widehat{j}_{n+r+1} - j_{n+r+1}|}{n+r+1} \le r_{n,k}.$$

The same bound holds for $\widehat{b}_n - b_n$, because the deterministic term $1/(n+1)$ cancels. Under the realizability null, Theorem B.1 gives $(-1)^k \mathsf{D}^k a_n \ge 0$ and $(-1)^k \mathsf{D}^k b_n \ge 0$ for every admissible pair. Therefore (29) cannot occur on the simultaneous event, proving type-I error at most δ. For winner-label averages, Hoeffding's inequality and a union bound over $N = 1, \dots, H$ yield (27) with the radii in (30). □

Proof of Theorem D.1. If $j(t) = J(t;\psi)$, the substitution $x = e^{-y}$ gives

$$\frac{j(t)}{t} = \int_0^\infty e^{-ty} \psi(e^{-y})\, dy, \qquad \frac{1-j(t)}{t} = \int_0^\infty e^{-ty} [1 - \psi(e^{-y})]\, dy.$$

Both are Laplace transforms of positive measures and hence completely monotone.

Conversely, the Bernstein–Widder theorem (Widder, 1941) gives positive Borel measures μ and ν on $[0,\infty)$ whose Laplace transforms are the two functions in Equation (26). Their sum has transform $1/t$, the Laplace transform of Lebesgue measure. Uniqueness of Laplace transforms for locally finite positive measures gives $\mu + \nu = \lambda$, so $d\mu = f\, d\lambda$ for a measurable $0 \le f \le 1$. Set $\psi(x) = f(-\log x)$ for $x \in (0,1]$ and define $\psi(0)$ arbitrarily. Reversing the substitution proves the representation. The same uniqueness argument gives uniqueness almost everywhere. □

Proof of Corollary D.2. The transform is linear, so $\mathfrak{J}_1$ is convex. It is injective by Corollary D.3. The extreme points of the order interval $\{\psi \in L^\infty : 0 \le \psi \le 1\}$ are precisely the indicator functions: if ψ takes values strictly between zero and one on a set of positive measure, it admits a nontrivial symmetric perturbation there; conversely an indicator cannot be expressed as a nontrivial average of two functions in the interval. Injective affine maps preserve extreme points. □

E Proofs for Finite-Horizon Recovery and Testing

An explicit smooth transition. Let $a(t) = e^{-1/t}$ for $t > 0$ and $a(t) = 0$ for $t \le 0$, and set

$$T(t) = \frac{a(t)}{a(t) + a(1-t)}.$$

Then $T \in C^\infty(\mathbb{R})$, $T = 0$ on $(-\infty, 0]$, $T = 1$ on $[1, \infty)$, and $T' > 0$ on $(0, 1)$. For $b = 1 - 1/H$ and $\eta = 1/(10^4 H)$ define

$$\psi_{0,H}(x) = 0.1 + 0.8T((x - 0.4)/\eta) - 0.7T((x - 0.75)/\eta), \tag{31}$$
$$\psi_{1,H}(x) = \psi_{0,H}(x) + 0.7T((x - b)/\eta). \tag{32}$$

For $H \ge 64$ the transition intervals are disjoint and contained in $(0, 1)$. Both profiles take values in $[0.1, 0.9]$, and they agree on $[0, b]$. Their derivative signs occur in respectively two and three ordered nonzero intervals. Consequently $\operatorname{osc}(d\psi_{0,H}) = 1$ and $\operatorname{osc}(d\psi_{1,H}) = 2$.

Proof of Theorem 3.1 and the gap assertions of Theorem 5.1. The corresponding unsmoothed profiles have curves

$$J_N^0 = 0.2 - 0.8(0.4)^N + 0.7(0.75)^N, \qquad J_N^1 = J_N^0 + 0.7(1 - b^N).$$

Thus J_N^1 is (4), and

$$G_1^0 = 0.06075, \qquad G_2^0 = -0.0216375, \qquad G_N^1 = G_N^0 + \frac{0.7}{H} b^N.$$

For $H \ge 64$, $G_1^1 \ge 0.06075$ and

$$G_2^1 \le -0.0216375 + 0.7/64 = -0.0107.$$

Furthermore, $J_2^0 = 0.46575$ and

$$J_2^1 = 0.46575 + 1.4/H - 0.7/H^2 \le 0.487625.$$

Since $b^H \le e^{-1} < 0.368$ and $0.8(0.4)^{64} < 10^{-20}$,

$$J_H^1 > 0.9 - 0.7(0.368) - 10^{-20} = 0.6424 - 10^{-20},$$

so $J_H^1 - J_2^1 > 0.154775 - 10^{-20}$. In the other model,

$$J_H^0 - J_2^0 \le 0.2 + 0.7(0.75)^{64} - 0.46575 < -0.2657.$$

These numerical inequalities follow directly from the displayed powers; the accompanying script also checks their rational parts exactly.

The L^1 change from smoothing is at most 2.2η for either profile. As $p_N(x) \le N$, for every $N \le H$,

$$|J_N(\psi_{i,H}) - J_N^i| \le 2.2H\eta = 0.00022.$$

Each difference of two utilities changes by at most 0.00044. Hence both smoothed models satisfy $J_2 - J_1 \ge 0.06$ and $J_2 - J_3 \ge 0.01$, while their horizon gaps have magnitude at least 0.15 with the asserted opposite signs. Set $\psi_H = \psi_{1,H}$ and realize it using uniform scores and Bernoulli utility via Lemma 2.1. This fixes the law for every breadth.

The positive first gain and negative second gain give one reversal. Since $J_H > J_2$, at least one gain G_N with $3 \le N < H$ is positive, giving a second reversal. The bound in Theorem 4.1 and $\operatorname{osc}(d\psi_{1,H}) = 2$ exclude any further reversals in the complete sequence. The definition of τ_H now gives the claimed missed-utility bound. □

Proof of the testing bounds in Theorem 5.1. For m i.i.d. observations let $E = \{X_j \le b \text{ for all } j\}$. Under either model $P(E) = b^{Km}$, because the CDF of p_K at b is b^K. The restrictions of the two data laws to E coincide, including labels, since the profiles agree there. For a randomized test, its conditional probabilities of returning zero and one sum to one. Therefore

$$P_0(T = 1, E) + P_1(T = 0, E) = b^{Km}.$$

The larger unconditional error is at least $b^{Km}/2$, proving (11). Extra observations of unlabeled ranks or scores do not alter this argument in the constructed uniform-score models: their law is identical under both hypotheses, independent of the labeled samples. For $K = 1$, Bernoulli's inequality gives $b^m \geq 1 - m/H$, proving the $3/8$ statement. If both errors are at most δ, then

$$m \geq \frac{\log(1/(2\delta))}{-K \log(1 - 1/H)} \geq \frac{H-1}{K} \log \frac{1}{2\delta},$$

using $-\log(1 - 1/H) \leq 1/(H-1)$.

For a matching pair-specific upper test, put $a = b + \eta$, $u = 1 - a = (1 - 10^{-4})/H$, and $\pi = 1 - a^K$. Since $Ku \leq 1$, the first two terms of the inclusion–exclusion bound give $\pi \geq Ku - \binom{K}{2} u^2 \geq Ku/2 \geq K/(3H)$. On $[a, 1]$, the label means are exactly 0.2 and 0.9. Classify by the sign of the average of $Z_j = \mathbf{1}\{X_j \geq a\}(Y_j - 0.55)$, negative for model zero. Under the two models, $\mathbb{E}Z_j = \mp 0.35\pi$, $\mathrm{Var}(Z_j) \leq 0.55^2\pi$, and the range of Z_j has length one. Bernstein's inequality bounds either error by

$$\exp\left\{ -\frac{m(0.35\pi)^2}{2(0.55^2\pi + 0.35\pi/3)} \right\} \leq e^{-m\pi/8} \leq e^{-mK/(24H)}.$$

The necessary and sufficient label counts therefore agree up to universal factors for $0 < \delta \leq 1/4$; the displayed constants are not claimed optimal.

For the rank-free upper bound, let $Y_{H,j}$ and $Y_{2,j}$ be independently generated and labeled winners, and put $D_j = Y_{H,j} - Y_{2,j} \in [-1, 1]$. The D_j are i.i.d. with mean $\Delta = J_H - J_2$. Classify by the sign of $r^{-1} \sum_j D_j$, assigning a zero arbitrarily. A wrong sign requires a one-sided deviation of at least $|\Delta| \geq d$. Hoeffding's inequality on an interval of length two gives error at most $e^{-rd^2/2}$ under either sign. The construction uses $2r$ labels and $r(H+2)$ independently generated candidates. Neither score ranks nor either profile need be known for this test. □

Scope of the testing result. The smooth laws depend on H, but for each fixed H the law itself is unchanged as N ranges over the positive integers. The utility gap is uniform for $H \geq 64$, whereas the derivative bounds are not. The lower bound applies to the fixed-K, i.i.d. audit experiment, even when the two candidate profiles are known. It does not rule out every adaptive breadth strategy, nor does it give a global stopping certificate from arbitrary finite data. Finite-prefix extrapolation and fixed-horizon recovery are also different questions: the two early curves here need not agree exactly, and the lower bound concerns sampled labels rather than noiseless access to an entire prefix.

F Proofs for Sharp Reversal Bounds

Lemma F.1 (Strict positivity of the generalized Vandermonde kernel)**.** *If $t_1 < \cdots < t_r$ and $0 < x_1 < \cdots < x_r < 1$, then* $\det[x_j^{t_i}]_{i,j=1}^r > 0$.

Proof. Put $y_j = \log x_j$. A nonzero linear combination of r exponentials $e^{t_i y}$ has at most $r - 1$ distinct zeros. Indeed, divide by $e^{t_1 y}$ and apply Rolle's theorem: its derivative is a combination of at most $r - 1$ exponentials, so induction proves the claim (a constant combination is immediate). A zero collocation determinant would give a nonzero combination vanishing at all r points, a contradiction. The determinant therefore has constant sign on the connected domain $y_1 < \cdots < y_r$. Letting the points coalesce identifies this sign with the Wronskian

$$e^{(t_1 + \cdots + t_r)y} \prod_{1 \leq i < j \leq r} (t_j - t_i) > 0,$$

because the coalescence Vandermonde and factorial factors are positive. This proves strict positivity. □

Lemma F.2 (Finite-matrix variation diminution)**.** *If a real matrix A is totally nonnegative, then* $\mathsf{S}^-(Aw) \leq \mathsf{S}^-(w)$ *for every real vector w, with zeros removed in both vectors.*

Proof. This is the classical finite-matrix variation-diminishing characterization of total nonnegativity (Karlin, 1968) as stated in Pinkus (2010, Theorem 3.4). This is the only external total-positivity result needed in the signed-measure argument below. □

Lemma F.3 (Ordered signed-measure variation diminution)**.** *Let $\mathcal{T}$ and $\mathcal{X}$ be ordered compact intervals and let $K : \mathcal{T} \times \mathcal{X} \to \mathbb{R}$ be continuous. Suppose every collocation matrix $[K(t_i, x_j)]$ at increasing points is totally nonnegative. For a finite signed Borel measure ν on $\mathcal{X}$, define $F(t) = \int K(t, x)d\nu(x)$. Then*

$$\mathsf{S}^-(F) \le \operatorname{osc}(\nu).$$

The conclusion also holds when $\mathcal{T}$ is a discrete ordered subset; only finitely many rows are needed at a time.

Proof. The assertion is trivial if $\operatorname{osc}(\nu) = \infty$. Otherwise take a minimal decomposition (5). Fix increasing $t_1 < \cdots < t_d$. Approximate each positive component μ_j weakly by a positive atomic measure carried by its ordered interval I_j. Concatenate the atoms in increasing order and combine atoms at equal abscissae. The resulting ordered weight vector has at most $\operatorname{osc}(\nu)$ sign changes. Its transform at (t_i) is a totally nonnegative collocation matrix times that vector, so Lemma F.2 gives at most $\operatorname{osc}(\nu)$ output sign changes.

Continuity of $K(t_i, \cdot)$ makes the atomic transforms converge coordinatewise to $(F(t_i))_{i=1}^d$. If the limit vector, after deleting zeros, had more sign changes, a subsequence of its nonzero coordinates would retain those strict signs under all sufficiently close approximations, a contradiction. Taking the supremum over finite increasing samples proves the claim. The same proof applies to a discrete row set. □

Proof of Theorem 4.1. Because $tx^{t-1}dx = d(x^t)$, Stieltjes integration by parts gives

$$J(t) = \psi(1) - \int_{[0,1]} x^t \, d\psi(x).$$

Here $\psi(1) = \psi(1-)$ and the extended Stieltjes measure has no endpoint atoms. Integration by parts is first applied on $[\epsilon, 1-\epsilon]$ and then passed to the endpoints using boundedness of ψ and finiteness of $|d\psi|$. Subtracting at consecutive integers gives Equation (6). Moreover, $\sup_{0<x<1} x^t|\log x| = (et)^{-1}$. Differentiation under the finite signed measure is therefore justified on compact t-intervals and gives Equation (7).

By Lemma F.1, x^t is strictly totally positive on the interior. The gain kernel $K_G(t, x) = x^t(1 - x)$ and derivative kernel $K_D(t, x) = x^t(-\log x)$ are positive column rescalings there and extend continuously to $[0, 1]$ with zero endpoint columns; hence all their collocation matrices are totally nonnegative. Applying Lemma F.3 directly to the fixed signed measure $d\psi$ gives both bounds in Equation (8). This avoids treating $(-\log x)d\psi$ as a globally finite measure near zero and makes the endpoint cases explicit.

For sharpness fix $k \ge 0$, integers $1 \le N_0 < \cdots < N_k$, and points $0 < x_0 < \cdots < x_k < 1$. Let $A_{ij} = x_j^{N_i}$ and $y_i = (-1)^i$. By Lemma F.1, A is invertible; let $w = A^{-1}y$ and $\nu_0 = \sum_{j=0}^k w_j \delta_{x_j}$. Its selected moments are y_i, so its moment transform has at least k sign changes. Variation diminution forces the ordered nonzero weights to have exactly k sign changes: fewer would contradict the selected output, while more are impossible for $k + 1$ entries. In particular, all $k + 1$ weights are nonzero.

Replace each atom by a C^∞ bump of the same signed mass, with pairwise disjoint supports in sufficiently small neighborhoods of the x_j. Moment continuity preserves the signs at $N_0, \ldots, N_k$ and at the corresponding real points, while the ordered bump density has exactly k sign changes across its $k+1$ ordered nonzero bumps. Denote it by g. The full discrete transform and the continuous transform $t \mapsto \int x^t g(x)dx$ therefore have at least k and, by the upper bound, at most k sign changes. Choose supports away from one and set $\psi'(x) = g(x)/(1 - x)$. After integration, multiply by a sufficiently small positive constant and add $1/2$. The resulting C^∞ profile lies strictly in $(0, 1)$, has $\operatorname{osc}(d\psi) = k$, and its complete gain sequence has exactly k sign changes. For sharpness of the continuous bound, repeat the construction with arbitrary real $0 < t_0 < \cdots < t_k$ and set $\psi'(x) = g(x)/(-\log x)$ on supports contained in a compact subinterval of $(0, 1)$. Then $J'(t) = \int x^t g(x)dx$ has exactly k sign changes. □

Proof of Corollary 4.2. The first statement is immediate from Theorem 4.1. If $d\psi \geq 0$, then $G_N = \int x^N(1-x)d\psi$ is a Hausdorff moment sequence of a positive measure, so $(-1)^k \mathsf{D}^k G_N \geq 0$. For $N \geq 2$, Cauchy–Schwarz gives $G_N^2 \leq G_{N-1}G_{N+1}$; this is log-convexity in the nonnegative-sequence sense, including the identically zero case. □

Proof of Proposition 4.3. The reward lies in $[(1-\varepsilon)/2, (1+\varepsilon)/2] \subset (0,1)$ and $r(1) = 1/2$. It is absolutely continuous and of bounded variation because its derivative is $O((1-x)^{\rho-1})$, which is integrable for every $\rho > 0$. If $\rho > 1$, direct differentiation gives $r'(0^+) = -\varepsilon\rho/2$ and $r'(1^-) = 0$, so r extends to a C^1 function on $[0,1]$. With $\alpha = \rho + i\omega$, the beta integral gives

$$J(t) = \frac{1}{2} + \frac{\varepsilon}{2}\,\mathrm{Re}\left[\Gamma(1+\alpha)\frac{\Gamma(t+1)}{\Gamma(t+1+\alpha)}\right].$$

Let $A = (\varepsilon/2)|\Gamma(1+\alpha)|$ and $\phi = \arg\Gamma(1+\alpha)$. The fixed-parameter expansion

$$\frac{\Gamma(t+1)}{\Gamma(t+1+\alpha)} = t^{-\alpha}\left(1 - \frac{\alpha(1+\alpha)}{2t} + O(t^{-2})\right)$$

and the identity $H'(t) = H(t)\{\mathrm{digamma}(t+1) - \mathrm{digamma}(t+1+\alpha)\}$, where $H(t) = \Gamma(1+\alpha)\Gamma(t+1)/\Gamma(t+1+\alpha)$ and $\mathrm{digamma} = \Gamma'/\Gamma$, justify the derivative expansion using $\mathrm{digamma}(t+1) - \mathrm{digamma}(t+1+\alpha) = -\alpha/t + O(t^{-2})$. Hence

$$J(t) = \frac{1}{2} + At^{-\rho}\cos(\phi - \omega\log t) + O(t^{-\rho-1}), \tag{33}$$

$$J'(t) = At^{-\rho-1}\{-\rho\cos(\phi - \omega\log t) + \omega\sin(\phi - \omega\log t)\} + O(t^{-\rho-2}). \tag{34}$$

The bracket is a nonzero sinusoid. Choose compact subintervals of one logarithmic period on which it is bounded respectively above and below by opposite nonzero constants. Their translates by $2\pi/\omega$ give alternating signs for all sufficiently large t because the remainder is one order smaller. Continuity then gives infinitely many local extrema.

For the integer conclusion no differentiation of an asymptotic remainder is needed. With $H(t) = \Gamma(1+\alpha)\Gamma(t+1)/\Gamma(t+1+\alpha)$, the Gamma recurrence gives exactly

$$H(N+1) - H(N) = -\frac{\alpha}{N+1+\alpha}H(N).$$

Applying the Gamma-ratio expansion to this identity gives

$$J_{N+1} - J_N = AN^{-\rho-1}\{-\rho\cos(\phi - \omega\log N) + \omega\sin(\phi - \omega\log N)\} + O(N^{-\rho-2}).$$

Each translated fixed-sign logarithmic interval becomes a multiplicative interval whose Euclidean length diverges, hence contains an integer for all sufficiently large translates. The remainder cannot change the selected signs, so the integer gain sequence reverses infinitely often.

Finally, for $t_2 > t_1$,

$$\frac{p_{t_2}(x)}{p_{t_1}(x)} = \frac{t_2}{t_1}x^{t_2-t_1}$$

is strictly increasing, proving strict MLR and strict TP_2; and $\partial_t \log p_t(x) = t^{-1} + \log x$ is continuous and strictly increasing in x. □

F.1 Scope of the TP_2 comparison

The comparison is to Theorem 1 and Corollary 3 in the NeurIPS 2025 proceedings paper and arXiv v2 of Khalaf et al. (2025). The stated hypotheses are common support, strict TP_2, a score continuous and strictly increasing in the outcome, and a bounded nonnegative reward. Equation (9) is positive and bounded; p_t is positive on common support $(0,1)$; the displayed likelihood ratio proves strict MLR/TP_2; and $\partial_t \log p_t(x) = t^{-1} + \log x$ satisfies the score condition. The conclusion fails in both continuous and integer breadth senses. For $\rho > 1$, the endpoint derivatives are $r'(0^+) < 0$ and $r'(1^-) = 0$, matching the displayed reward-derivative inequalities in the "immediate decline" row, while the expected reward still reverses infinitely often. This comparison concerns those displayed inequalities; we do not assume the preceding endpoint lemma's conclusions as additional hypotheses. The direct counterexample to Theorem 1 is independent of that endpoint classification.

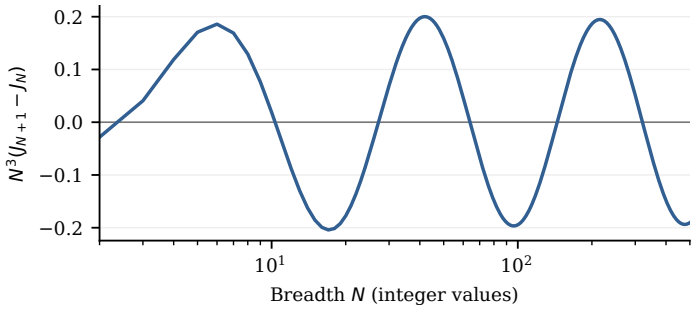

Figure 2: Integer gains for the positive C^1 log-periodic profile in Proposition 4.3, with $\varepsilon = 1/2$, $\rho = 2$, and $\omega = 4$. The N^3 factor exposes signs of decaying gains; it is a rescaling, not the magnitude of the utility change. Values come from the exact formula, checked against the finite product above.

The underlying issue is structural. For a fixed reward,

$$J'(t) = \int p_t(x) r(x) \partial_t \log p_t(x)\, dx.$$

Although the integrand has a simple pointwise sign pattern for each fixed t, the signed input itself varies with t. Classical variation diminution controls $t \mapsto \int K(t,x) d\nu(x)$ for one fixed ν. Stieltjes integration by parts in Theorem 4.1 restores that valid form for the Beta family. For rewards with $\mathrm{osc}(dr) \le 1$, including the toy reward $x^p(1-x)$ used in the cited work, Corollary 4.2 gives the corresponding valid one-extremum guarantee.

Numerical illustration. The oscillation is already visible for ordinary parameter values and does not rely on a finely tuned numerical example. For example, take $\varepsilon = 1/2$, $\rho = 2$, and $\omega = 4$. For integer N, the Gamma ratio in (10) equals $\prod_{k=1}^{N} k/(k+\rho+i\omega)$. Evaluating this finite product gives

$$G_2 \approx -3.54 \times 10^{-3}, \quad G_3 \approx 1.51 \times 10^{-3}, \quad G_{10} \approx 1.88 \times 10^{-5}, \quad G_{11} \approx -2.80 \times 10^{-5},$$

and later $G_{27} < 0 < G_{28}$ and $G_{63} > 0 > G_{64}$. These values illustrate the closed-form construction; the analytic argument above proves the existence of infinitely many reversals.

G Proof of the Abelian Transfer

Proof of Theorem C.1. Let $M_N = \max_{i \le N} X_i$ and $Q_N = 1 - M_N$. Then

$$u_\infty - J_N = \mathbb{E} g(Q_N).$$

Moreover, $NQ_N \Rightarrow E$ for $E \sim \mathrm{Exp}(1)$, and for every $p > 0$,

$$\mathbb{E} Q_N^p = \frac{\Gamma(1+p)\Gamma(N+1)}{\Gamma(N+1+p)} \sim \Gamma(1+p) N^{-p}.$$

Choose $\eta \in (0, \gamma)$ and $\delta > 0$ sufficiently small that g is positive and Potter bounds (Bingham et al., 1987) apply on $(0, \delta]$. On $\{Q_N \le \delta\}$,

$$\frac{g(Q_N)}{g(1/N)} \le C\left[(NQ_N)^{\gamma-\eta} + (NQ_N)^{\gamma+\eta}\right].$$

For some $r > 1$, raising the right-hand side to the rth power and using the displayed moment formula with exponents $r(\gamma \pm \eta)$ gives a uniformly bounded rth moment. Hence the truncated ratios are uniformly integrable. Regular variation together with $NQ_N \Rightarrow E$ then gives

$$\frac{g(Q_N)}{g(1/N)} \mathbf{1}\{Q_N \le \delta\} \Rightarrow E^\gamma$$

and convergence of expectations. On the complement,

$$\frac{\mathbb{E}[|g(Q_N)| \mathbf{1}\{Q_N > \delta\}]}{g(1/N)} \le \frac{2\|\psi\|_\infty (1-\delta)^N}{g(1/N)} \longrightarrow 0,$$

because $g(1/N)$ is regularly varying and exponential decay dominates its reciprocal. Therefore

$$\frac{\mathbb{E}g(Q_N)}{g(1/N)} \to \mathbb{E}E^\gamma = \Gamma(1+\gamma).$$

Since $g(1/N) \sim aN^{-\gamma}\ell(N)$, the conclusion follows. □

H Additional Audit Statements

Bounded leverage is a genuine restriction. For any $B_0 > 2$, choose a measurable set E with $\int_E p_N = B_0^{-2}$ and define $q_{B_0} = p_N/B_0$ on E and $q_{B_0} = c_{B_0} p_N$ on E^c, where $c_{B_0} = (1 - B_0^{-3})/(1 - B_0^{-2})$. Then q_{B_0} is a density,

$$B_N(q_{B_0}) = B_0, \qquad A_N(q_{B_0}) = B_0^{-1} + \frac{(1 - B_0^{-2})^2}{1 - B_0^{-3}} = 1 + O(B_0^{-1}).$$

Hence $B_N/A_N \to \infty$ even though the second-moment overlap remains bounded. The bounded-leverage condition is therefore not automatic. This construction does not show that importance weighting is suboptimal for the particular design; when leverage is unbounded, Proposition H.3 instead gives an upper bound in terms of the full weight tail.

H.1 Observation models and scope of comparison

Result or ingredient	Observation model or mathematical object	Role in this paper
Bounded-density moments	A measure dominated by a fixed reference measure; Lasserre (2007, Theorem 3(c))	Exact specialization after task-conditional rank reduction; not a new moment theorem.
Classical variation diminution	One fixed signed input and a totally positive kernel	Stieltjes integration by parts gives the fixed input for both breadth differences and derivatives; smooth interpolation proves sharpness.
Wang et al. (2017)	Finite actions, contextual observations; bounded conditional means/variances and moment/sample-size conditions	Closely related overlap lower bounds. We keep the bounded-label experiment and the all-m, N constants explicit.
Ma et al. (2022)	Finite-action bounded-reward OPE; known behavior policy	A stronger full minimax characterization in its setting. We give upper and lower bounds for general rank designs, which match under bounded leverage.
Winner auditing	Independent actual winners and their utility labels	Rank-free implementation of the fixed-breadth $1/m$ rate.
Passive / partial targeting	Rank–label pairs under 1 or p_K	The general and smooth rate calculations assume observed ranks. Raw-score reuse is covered by Theorem H.1 and Proposition H.2, not by an asserted unknown-rank smoothness theorem.

The table separates classical ingredients from results that are specific to the present observation models. The finite-horizon recovery and sharp reversal theorems concern the population curve. The recovery-testing theorem uses the same smooth construction to obtain matching fixed-K label complexity, while the general audit results describe how resolution changes with the sampling design.

H.2 Auditing from raw scores

The next construction addresses the case in which a passive labeled set already exists and new labels cannot be acquired. Task-matched unlabeled scores are used to recover conditional ranks. When all data can be collected from scratch, direct winner labeling remains the simpler benchmark.

Theorem H.1 (Raw-score rank auditing). *Let $(Z_j, S_j, Y_j)_{j=1}^m$ be i.i.d. passive triples from the fixed target task law, with continuous conditional score CDFs. For each j, conditional on Z_j, draw $N-1$ i.i.d. reference scores $S'_{j,\ell}$ independently of (S_j, Y_j) and of all other observations and reference pools. Define, for integers $m, N \geq 1$,*

$$\widetilde{J}_N^{\mathrm{rank}} = \frac{N}{m} \sum_{j=1}^{m} Y_j \prod_{\ell=1}^{N-1} \mathbf{1}\{S'_{j,\ell} \leq S_j\}. \tag{35}$$

Then $\mathbb{E}\widetilde{J}_N^{\mathrm{rank}} = J_N$ without knowing F_Z, and its clipped version has worst-case risk at most $1 \wedge N/m$. With probability at least $1-\delta$,

$$|\widetilde{J}_N^{\mathrm{rank}} - J_N| \leq \sqrt{\frac{2N\log(2/\delta)}{m}} + \frac{N\log(2/\delta)}{3m}. \tag{36}$$

For $N = 1$, the empty product equals one. For discrete scores, compare independently randomized score–uniform pairs lexicographically, using the same auxiliary uniform whenever a score is reused. The rates hold uniformly over the bounded conditional label laws. Unlabeled references remove knowledge of the CDF, not the passive label lower bound.

Proposition H.2 (Shared-reference rank auditing). *Fix a homogeneous task law with continuous score CDF and integers $m \geq 1$, $N \geq 2$, $M \geq N-1$. Let $(S_j, Y_j)_{j=1}^m$ be i.i.d. passive pairs and let $S'_1, \ldots, S'_M$ be one independent i.i.d. reference pool from the same score law. With $R_M(v) = \sum_{r=1}^M \mathbf{1}\{S'_r \leq v\}$ and*

$$\widehat{w}_{M,N}(v) = N \frac{\binom{R_M(v)}{N-1}}{\binom{M}{N-1}}, \qquad \widetilde{J}_{N,M}^{\mathrm{pool}} = \frac{1}{m} \sum_{j=1}^{m} Y_j \widehat{w}_{M,N}(S_j),$$

where the numerator is zero if $R_M(v) < N-1$, define $\widehat{J}_{N,M}^{\mathrm{pool}} = \Pi_{[0,1]}\{\widetilde{J}_{N,M}^{\mathrm{pool}}\}$. The un-clipped estimator is unbiased, and the clipped estimator satisfies

$$\sup_{\psi \in \Psi_{\mathrm{all}}} \sup_{P \in \mathfrak{P}(\psi,1)} \mathbb{E}(\widehat{J}_{N,M}^{\mathrm{pool}} - J_N)^2 \leq 1 \wedge \left(\frac{N}{m} + \frac{4N}{M}\right). \tag{37}$$

For $N = 1$, use the passive label average; no reference pool is needed and its risk is at most $1/m$. For discrete scores the lexicographic convention above applies. Expectation in (37) includes both the independent reference pool and the labeled pairs; on the quantile scale their one-pair law is P. Sharing the reference pool reduces generation from $m(N-1)$ fresh comparisons to one pool of size M, which can be useful when passive labels already exist. A single pool shared across different task laws is generally biased because conditional ranks, not pooled ranks, are the invariant coordinate.

H.3 Truncation and decision certificates

Proposition H.3 (Truncated importance weighting under infinite overlap). *For $w = p_N/q$, $w_\tau = w \wedge \tau$, $b_\tau = \mathbb{E}_q[(w-\tau)_+]$, and $v_\tau = \mathbb{E}_q[w_\tau^2]$, define*

$$\widetilde{J}_\tau = \frac{1}{m} \sum_{j=1}^{m} w_\tau(X_j) Y_j, \qquad \widehat{J}_\tau = \Pi_{[0,1]}\{\widetilde{J}_\tau\}. \tag{38}$$

Then

$$\sup_{\psi \in \Psi_{\mathrm{all}}} \sup_{P \in \mathfrak{P}(\psi,q)} \mathbb{E}_{P^{\otimes m}}(\widehat{J}_\tau - J_N)^2 \leq b_\tau^2 + \frac{v_\tau}{m}. \tag{39}$$

The bound remains informative when $A_N(q) = \infty$. In that regime the exact minimax rate depends on the full tail of the importance weights, which we do not characterize here.

Corollary H.4 (Certified finite breadth selection). *Let $\mathcal{B} \subset \mathbb{N}$ be finite. Suppose estimates $(\widehat{J}_N)_{N \in \mathcal{B}}$ and deterministic radii $(r_N)_{N \in \mathcal{B}}$ satisfy, with probability at least $1-\delta$,*

$$|\widehat{J}_N - J_N| \leq r_N \qquad \textit{simultaneously for all } N \in \mathcal{B}.$$

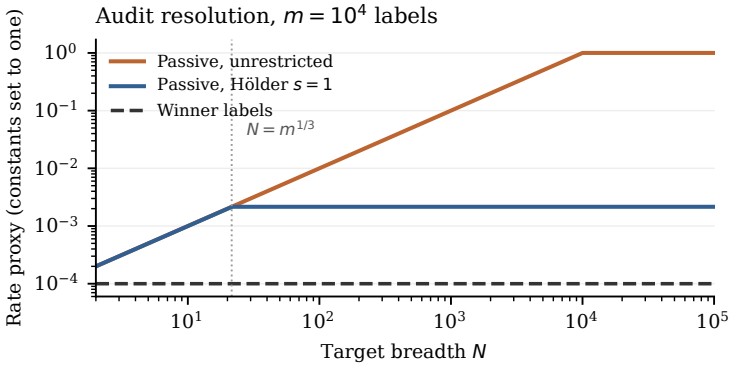

Figure 3: **Fixed-breadth audit resolution.** Rate functions from Corollary 6.3 and Theorem 6.5, with $m = 10^4$ and $s = 1$. Constants are set to one; these are theoretical rate proxies, not exact minimax risks or empirical MSEs.

Choose

$$\widehat{N} \in \arg\max_{N \in \mathcal{B}} (\widehat{J}_N - r_N), \qquad N^\star \in \arg\max_{N \in \mathcal{B}} J_N.$$

Then, on the same event,

$$J_{N^\star} - J_{\widehat{N}} \le 2 r_{N^\star} \le 2 \max_{N \in \mathcal{B}} r_N. \tag{40}$$

Thus any simultaneous confidence certificate yields a finite-sample guarantee for selection over the finite set $\mathcal{B}$; no unimodality assumption is needed.

Corollary H.5 (Passive audit sample complexity)**.** *Under Theorem 6.5, the smallest passive sample size achieving minimax root-MSE at most $\epsilon \in (0, 1/4)$ satisfies*

$$m^*_{N,s}(\epsilon) \asymp_{s,L} \min\{N\epsilon^{-2}, \epsilon^{-(2s+1)/s}\}. \tag{41}$$

I Proofs of the Minimax Audit Rates

Write $\widetilde{J}^{\mathrm{IW}}_{N,q} = m^{-1} \sum_{j=1}^m w(X_j) Y_j$ for the un-clipped importance estimator.

Proof of Theorem 6.1. Write $w = p_N/q$, $A = \mathbb{E}_q w^2$, and $B = \|w\|_\infty$. Since $\mathbb{E}_q w = 1$ and $w \le B$, one has $1 \le A \le B$. The un-clipped importance estimator is unbiased and

$$\mathrm{Var}\left(\frac{1}{m} \sum_{j=1}^m w(X_j) Y_j \right) \le \frac{\mathbb{E}_q[w(X)^2 Y]}{m} \le \frac{A}{m}.$$

Projection onto $[0, 1]$ cannot increase squared error, proving the upper bound together with the trivial risk bound one.

For the lower bound let $h = w/B$. Then $0 \le h \le 1$,

$$\int p_N h = \frac{A}{B}, \qquad \int h^2 q = \frac{A}{B^2}.$$

Restricting to the Bernoulli-label subclass, let $a = \frac{1}{8} \min\{1, B/\sqrt{mA}\}$ and $\psi_\pm = 1/2 \pm ah$. The profiles lie in $[3/8, 5/8]$ and their target separation is $\delta_J = 2aA/B$. For $0 \le u \le 1/8$, the exact Bernoulli divergence is $2u \log\{(1+2u)/(1-2u)\} \le 16u^2$. Therefore

$$\mathrm{KL}(P_+^{\otimes m} \| P_-^{\otimes m}) \le 16 m a^2 \int h^2 q = 16 m a^2 A / B^2 \le 1/4.$$

Pinsker's inequality gives total variation at most $1/\sqrt{8} < 1/2$. The nearest-target testing reduction (Tsybakov, 2009, Section 2.2) bounds the maximum squared-error risk below by $\delta_J^2(1 - \mathrm{TV})/8 \ge \delta_J^2/16$. Consequently,

$$\mathcal{R}_{m,N}(\Psi_{\mathrm{all}}; q) \ge \frac{a^2 A^2}{4B^2} = \frac{1}{256} \left(\frac{A}{m} \wedge \frac{A^2}{B^2} \right).$$

Together with the preceding variance calculation, this proves the stated constants. □

Proof of Corollary 6.2. If $B \le \kappa A$, then $A^2/B^2 \ge \kappa^{-2}$. Combining this with Theorem 6.1 gives the lower bound $c_\kappa(1 \wedge A/m)$ and the same theorem gives the matching upper bound. □

Proof of Proposition H.3. For $w_\tau = w \wedge \tau$, the estimator has absolute bias at most

$$\mathbb{E}_q[(w - w_\tau)\psi] \le \mathbb{E}_q[(w - \tau)_+] = b_\tau$$

and variance at most v_τ/m. Projection onto $[0, 1]$ cannot increase squared error, so the displayed bound follows. This argument remains valid when $A = \infty$. □

Proof of Corollary 6.3. Cauchy–Schwarz gives

$$A_N(q) = \int \frac{p_N^2}{q} \ge \frac{(\int p_N)^2}{\int q} = 1,$$

with equality exactly when p_N/q is constant q-almost everywhere, equivalently $q = p_N$. For $q = p_N$, $A_N = B_N = 1$. No design can improve the $1/m$ order: within the Bernoulli-label subclass, the constant alternatives $\psi_\pm \equiv 1/2 \pm c/\sqrt{m}$ have target separation $2c/\sqrt{m}$ and m-sample KL bounded by a universal constant under every q. For $q \equiv 1$,

$$A_N(1) = \int_0^1 N^2 x^{2N-2}\, dx = \frac{N^2}{2N-1}, \qquad B_N(1) = N \le 2A_N(1).$$

Apply Corollary 6.2. □

Proof of Theorem 6.4. A K-way score winner has quantile density $q = p_K(x) = Kx^{K-1}$. For $K \le N$,

$$A_N(p_K) = \int_0^1 \frac{N^2 x^{2N-2}}{Kx^{K-1}}\, dx = \frac{N^2}{K(2N-K)},$$

and

$$B_N(p_K) = \operatorname*{ess\,sup}_{x\in[0,1]} \frac{N}{K} x^{N-K} = \frac{N}{K}.$$

Hence $B_N(p_K)/A_N(p_K) = (2N - K)/N \le 2$, so Corollary 6.2 gives

$$\mathcal{R}^{\mathrm{cost}}_{C,N}(K) \asymp 1 \wedge \frac{N^2}{m_K K(2N-K)}.$$

For $K \in \mathcal{K}_C$, writing $x_K = C/(c_y + Kc_g) \ge 1$ gives $x_K/2 \le \lfloor x_K \rfloor \le x_K$. Consequently the floor-aware rate proxy and the continuous-budget proxy in (18) differ by at most a universal factor, including after clipping at one.

Normalize by c_g and consider the unsaturated continuous proxy

$$f(K) = \frac{K + \lambda}{K(2N-K)}, \qquad 0 < K < N.$$

Its derivative has the sign of $K^2 + 2\lambda K - 2\lambda N$. If $\lambda = 0$, the proxy is increasing and the constrained optimum is $K = 1$. If $\lambda > 0$, it decreases up to the unique positive root $\sqrt{\lambda^2 + 2\lambda N} - \lambda$ and increases thereafter. Clipping this root to the feasible interval $[1, K_{\max}]$ and checking the adjacent integers therefore minimizes the continuous-budget proxy over $\mathcal{K}_C$; if every proxy value is saturated at one, every feasible K is a minimizer. Combining the uniform equivalence between the two proxies with the two-sided minimax equivalence proves (20). The example in Remark I.1 shows why this argument does not identify an exact floor-aware optimizer. □

Remark I.1 (Why exact rounding is not valid). The floor in m_K can move the minimizer of the floor-aware rate proxy away from the nearest integer to $K^\dagger_{\mathrm{cont}}$. For example, take $N = 3$, $c_y = 1.5c_g$, and $C = 9c_g$. Then $K^\dagger_{\mathrm{cont}} \approx 1.854$, whereas

$$\frac{N^2}{m_K K(2N-K)} = 0.6,\ 0.5625,\ 0.5 \quad \text{for } K = 1, 2, 3,$$

respectively. The closed form remains rate-optimal by Theorem 6.4, but it is not an exact optimizer of the finite-budget proxy.

Proof of Theorem H.1. Let

$$I_j = \prod_{\ell=1}^{N-1} \mathbf{1}\{S'_{j,\ell} \le S_j\}.$$

The stated conditional independence from the full labeled pair implies $\mathbb{E}[I_j \mid Z_j, S_j, Y_j] = F_{Z_j}(S_j)^{N-1}$. In particular, continuity of the task-specific score law gives

$$\mathbb{E}[I_j \mid Z_j, S_j] = F_{Z_j}(S_j)^{N-1}.$$

Writing $X_j = F_{Z_j}(S_j)$ and using $\mathbb{E}[Y_j \mid X_j, Z_j] = \psi_{Z_j}(X_j)$ therefore yields

$$\mathbb{E}[NY_jI_j] = N\mathbb{E}[\psi_{Z_j}(X_j)X_j^{N-1}] = J_N.$$

Thus the un-clipped estimator is unbiased. Since $0 \le Y_j \le 1$ and $I_j \in \{0,1\}$,

$$\mathbb{E}[(NY_jI_j)^2] \le N^2\mathbb{E}[Y_jI_j] = NJ_N \le N.$$

Independence across j gives variance at most N/m, and projection onto $[0,1]$ cannot increase squared error. This proves the risk bound in Theorem H.1. Bernstein's inequality, using summand range N and variance proxy N, gives Equation (36).

For the matching lower bound, restrict the model to the continuous score law $S \sim \mathrm{Unif}(0,1)$. Under this submodel, every additional unlabeled score has a distribution independent of ψ and hence contributes zero likelihood ratio between profile alternatives. The passive-design lower bound in Theorem 6.1 with $q \equiv 1$ therefore continues to hold even after adjoining arbitrarily many unlabeled scores. Since $A_N(1) \asymp N$, the lower bound is $c(1 \wedge N/m)$. □

Proof of Proposition H.2. For the fixed homogeneous task law, conditionally on an audited score v, the binomial-combination identity gives

$$\mathbb{E}\left[\frac{\binom{R_M(v)}{N-1}}{\binom{M}{N-1}}\middle| v\right] = F_S(v)^{N-1}.$$

Thus the un-clipped estimator is unbiased. Let $\mathcal{S}$ denote the shared reference pool and put $W_j = Y_j\widehat{w}_{M,N}(S_j)$. Conditional on $\mathcal{S}$, the W_j are i.i.d. Since $0 \le \widehat{w}_{M,N} \le N$ and $0 \le Y_j \le 1$, $W_j^2 \le NW_j$. Taking expectations and using $\mathbb{E}W_j = J_N \le 1$ yields

$$\mathbb{E}\{\mathrm{Var}(\widetilde{J}^{\mathrm{pool}}_{N,M} \mid \mathcal{S})\} \le \frac{N}{m}.$$

It remains to control the variance of the conditional mean. On the score-quantile scale, that mean is the order-$(N-1)$ U-statistic with symmetric kernel

$$h(u_1, \dots, u_{N-1}) = N\int_{\max_r u_r}^1 \psi(x)\,dx.$$

If $V_{(N-1)}$ is the maximum of $N-1$ independent uniforms, then

$$0 \le h \le N(1 - V_{(N-1)}), \qquad \mathbb{E}h^2 \le N^2\mathbb{E}(1 - V_{(N-1)})^2 = \frac{2N}{N+1} \le 2.$$

Let $r = \lfloor M/(N-1) \rfloor$. A random permutation of the reference sample forms r disjoint blocks of size $N-1$; the U-statistic is the conditional expectation, given the sample, of the average of the r independent block kernels. Rao–Blackwellization therefore gives

$$\mathrm{Var}\{\mathbb{E}(\widetilde{J}^{\mathrm{pool}}_{N,M} \mid \mathcal{S})\} \le \frac{\mathrm{Var}(h)}{r} \le \frac{4N}{M},$$

where $\lfloor x \rfloor \ge x/2$ for $x \ge 1$ was used. The law of total variance proves Equation (37); clipping cannot increase squared error. □

Corollary I.2 (Importance-weighted finite-sample certificate)**.** *Let $A = A_N(q)$ and $B = B_N(q) < \infty$. The un-clipped estimator obeys, with probability at least $1-\delta$,*

$$|\widetilde{J}^{\mathrm{IW}}_{N,q} - J_N| \le \sqrt{\frac{2A\log(2/\delta)}{m}} + \frac{B\log(2/\delta)}{3m}. \tag{42}$$

Proof of Corollary I.2. Let $T_j = w(X_j)Y_j$. Then $0 \le T_j \le B$, $\mathbb{E}T_j = J_N$, and $\mathrm{Var}(T_j) \le \mathbb{E}T_j^2 \le A$. For either sign and $0 < \lambda < 3/B$, the bounded-summand mgf bound is

$$\log \mathbb{E} \exp\left\{ \pm\lambda \sum_{j=1}^{m} (T_j - \mathbb{E}T_j) \right\} \le \frac{mA\lambda^2}{2(1 - B\lambda/3)}.$$

Taking $u = \log(2/\delta)$ and $\lambda = \sqrt{2u/(mA)}/\{1 + (B/3)\sqrt{2u/(mA)}\}$ in the Chernoff bound gives the threshold $\sqrt{2mAu} + Bu/3$ for the sum. A union bound over the two signs proves Equation (42), including the displayed $1/3$ coefficient. □

Proof of Corollary H.4. On the simultaneous coverage event,

$$J_{\widehat{N}} \ge \widehat{J}_{\widehat{N}} - r_{\widehat{N}} \ge \widehat{J}_{N^\star} - r_{N^\star} \ge J_{N^\star} - 2r_{N^\star},$$

where the middle inequality is the definition of $\widehat{N}$. Rearranging proves Equation (40). □

We next prove Theorem 6.5. Let $r = \lceil s \rceil - 1$, $v(t) = (1, t, \ldots, t^r)^\top$, and

$$G = \int_0^1 v(t)v(t)^\top \, dt.$$

For bandwidth $h \in (0, 1]$, define

$$\widehat{\beta}_h = G^{-1} \frac{1}{mh} \sum_{j=1}^{m} v\left(\frac{1 - X_j}{h}\right) Y_j \mathbf{1}\{X_j \ge 1 - h\}. \tag{43}$$

Let $Q_N = 1 - M_N$ for $M_N \sim \mathrm{Beta}(N, 1)$ and

$$c_{N,h,k} = \mathbb{E}[(Q_N/h)^k], \qquad k = 0, \ldots, r, \qquad c_{N,h} = (c_{N,h,0}, \ldots, c_{N,h,r})^\top,$$

and define

$$\widetilde{J}^{\mathrm{LP}}_{N,h} = c_{N,h}^\top \widehat{\beta}_h, \qquad \widehat{J}^{\mathrm{LP}}_{N,h} = \Pi_{[0,1]}\{\widetilde{J}^{\mathrm{LP}}_{N,h}\}. \tag{44}$$

Lemma I.3 (Localized polynomial upper bound)**.** *If $\psi \in \Psi_s(L)$ and $Nh \ge 1$, then*

$$\sup_{P \in \mathfrak{P}(\psi, 1)} \mathbb{E}_{P^{\otimes m}}[(\widehat{J}^{\mathrm{LP}}_{N,h} - J_N(\psi))^2] \le C_{s,L} \left(h^{2s} + \frac{1}{mh} \right). \tag{45}$$

Proof. Taylor expansion at one gives, for all $y \in [0, 1]$,

$$\psi(1 - y) = \sum_{k=0}^{r} b_k (y/h)^k + R(y), \qquad |R(y)| \le C_{s,L} y^s,$$

where $b_k = (-h)^k \psi^{(k)}(1)/k!$. On the fitting interval $y \le h$, the expectation of Equation (43) equals b plus a vector of norm $O(h^s)$. Moreover,

$$c_{N,h,k} = h^{-k} \frac{\Gamma(k+1)\Gamma(N+1)}{\Gamma(N+k+1)} \le C_k (Nh)^{-k} \le C_k.$$

The local-fitting bias is therefore $O(h^s)$, while the target Taylor remainder is

$$\mathbb{E}|R(Q_N)| \le C_{s,L} \mathbb{E} Q_N^s \le C_{s,L} N^{-s} \le C_{s,L} h^s.$$

The summand in Equation (43) has second moment $O(h^{-1})$: it is supported on a set of probability h and carries the factor h^{-1}. Uniformly over every conditional label law in $\mathfrak{P}(\psi, 1)$, boundedness of Y controls this second moment. Since G^{-1} and $c_{N,h}$ are uniformly bounded when $Nh \ge 1$,

$$\mathrm{Var}(\widetilde{J}^{\mathrm{LP}}_{N,h}) \le \frac{C_s}{mh}.$$

Combining squared bias and variance proves the bound for the un-clipped estimator; projection onto $[0, 1]$ cannot increase squared error, proving the lemma. □

Upper bound in Theorem 6.5. If $N \le m^{1/(2s+1)}$, use the passive importance estimator from Theorem 6.1, whose risk is at most CN/m. If $N > m^{1/(2s+1)}$, take $h = m^{-1/(2s+1)}$. Then $Nh \ge 1$, and Lemma I.3 gives risk at most $Cm^{-2s/(2s+1)}$. □

Lower bound in Theorem 6.5. Choose $\varphi \in C^\infty([0,\infty))$ with $0 \le \varphi \le 1$, equal to one on $[0, 1/2]$ and zero on $[1, \infty)$. Its Hölder-s norm on the half-line is bounded by a constant depending only on s. Smoothness at zero is one-sided. Set

$$h = \max\left\{\frac{1}{N}, m^{-1/(2s+1)}\right\}, \qquad a = c_0 \min\{h^s, (mh)^{-1/2}\},$$

and define

$$b = \min\{1/4, L/4\}, \qquad \psi_\pm(x) = b \pm a\varphi\left(\frac{1-x}{h}\right).$$

Use Bernoulli labels with these conditional means. Let $C_s \ge 1$ bound the cutoff's derivatives through order r and its order-r Hölder seminorm, and choose

$$0 < c_0 \le \min\{b/2, L/(2C_s), \sqrt{b}/8\}.$$

The kth derivative of the perturbation is bounded by $C_s a h^{-k}$ for $0 \le k \le r$, and its highest-derivative Hölder seminorm by $C_s a h^{-s}$. Since $h \le 1$ and $ah^{-s} \le c_0$, all defining bounds of $\Psi_s(L)$ hold: the zeroth-order bound is at most $b + C_s c_0 \le 3L/4$, and the other bounds are at most $L/2$. Both profiles lie in $[b/2, 3b/2] \subset (0, 3/8]$. Thus no lower threshold on L is needed. The Beta weight of $[1-h/2, 1]$ is

$$1 - (1-h/2)^N \ge 1 - e^{-Nh/2} \ge 1 - e^{-1/2},$$

so the target separation is at least $2(1-e^{-1/2})a$. The Bernoulli inequality $\mathrm{kl}(p,q) \le (p-q)^2/\{q(1-q)\}$, with $q = \psi_- \ge b/2$ and $q \le b \le 1/4$, bounds the m-sample KL divergence by

$$\frac{16ma^2}{b}\int_0^1 \varphi^2\left(\frac{1-x}{h}\right)dx \le \frac{16ma^2h}{b} \le \frac{16c_0^2}{b} \le \frac{1}{4}.$$

Le Cam's lemma (Tsybakov, 2009, Section 2.2) gives risk at least ca^2. If $N \le m^{1/(2s+1)}$, then $h = 1/N$ and $a^2 \asymp_{s,L} N/m$. If $N \ge m^{1/(2s+1)}$, then $h = m^{-1/(2s+1)}$ and $a^2 \asymp_{s,L} m^{-2s/(2s+1)}$. These choices also apply when $m = 1$ or $N = 1$. Constants may depend on L; no uniform lower constant as $L \downarrow 0$ is asserted. □

Proof of Corollary H.5. By Theorem 6.5, up to constants depending only on s, L, squared error at most ϵ^2 is attainable when at least one of

$$\frac{N}{m} \lesssim \epsilon^2, \qquad m^{-2s/(2s+1)} \lesssim \epsilon^2$$

holds. Solving the two inequalities and taking the smaller sample size gives the upper bound in Equation (41); the matching minimax lower bound follows from the same rate equivalence. □