IMPROVING SHAPLEY DATA VALUATION FOR RECOMMENDATION SYSTEMS BY DEBIASING POPULARITY

Anonymous authors
Paper under double-blind review

ABSTRACT

Data valuation decides which interactions a recommendation system keeps and acts on, and how revenue is split among content providers. A wrong value therefore propagates into both what users see next and who gets paid. In systems with implicit feedback, the logging policy makes such errors systematic: exposure is missing not at random (MNAR), so a point can look valuable because it was shown often rather than because it is relevant. We formalize this intuition by revealing that *naive Shapley valuation* shifts credit from low-exposure to high-exposure items, and can reverse their order when exposure gaps outweigh relevance gaps. To counter this, we propose *SHADE* (*SHApley with Debiased Estimates*), which corrects each per-point value with debiased and doubly robust estimates before solving the cooperative game with the closed-form KNN-Shapley recursion. Experiments on synthetic MNAR benchmarks and real recommendation datasets confirm that *naive valuation* consistently under-values the long tail, while *SHADE* recovers much of the oracle level and improves value ranking. The correction also assigns value to items that were never exposed, which enables augmentation and collection strategies that naive valuation cannot support.

1 INTRODUCTION

Data valuation assigns credit to individual training points according to their contribution to model utility (Ghorbani & Zou, 2019; Jia et al., 2019b). In recommendation, these values are already acted on. *Data Shapley* is used in production to debug KNN recommenders (Kersbergen et al., 2025) and prune interaction logs (Zhang et al., 2025). The same values also guide many downstream decisions, including which *data to pay for* and which *interactions to collect next*. Once exposure follows these values, however, we need to know that they measure *what we think they measure*. On logged recommendation data, they often do not. Consider two items that are equally relevant to a user, but one is repeatedly surfaced while the other is rarely shown. The first can accumulate far more logged interactions and therefore appear more valuable, even though the two items are equally relevant. The difference reflects the *opportunity to be observed*, which the platform controls, rather than *usefulness*, which is what a data value is intended to capture. Standard valuation applied directly to the log cannot separate these two effects.

The source of this distortion is selective exposure. In implicit-feedback logs, exposure is missing-not-at-random (MNAR): what is shown depends on user behavior and production policies that often favor already-popular content (Schnabel et al., 2016; Saito et al., 2020; Chen et al., 2023). The usual implicit-feedback value of a training point is one when an interaction is logged and zero otherwise. Its expectation is therefore the relevance of the point multiplied by its exposure probability. Our example is a direct consequence: equal relevance but unequal exposure produces unequal expected value. We show that the resulting *Shapley* values inherit this distortion in a precise form. Credit shifts toward frequently exposed head items and away from rarely exposed long-tail items, with an exact rate given by Theorem 5.3. Exposure also changes the ranking. For two adjacent points in the nearest-neighbor ordering, the less relevant point is ranked higher in expectation once the exposure gap exceeds the relevance gap (Equation 7). A sufficiently over-exposed item can therefore outrank a more relevant but rarely shown one.

The problem extends beyond measurement because data values can guide subsequent decisions. Suppose a platform allocates future exposure or a data-collection budget in proportion to values

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.

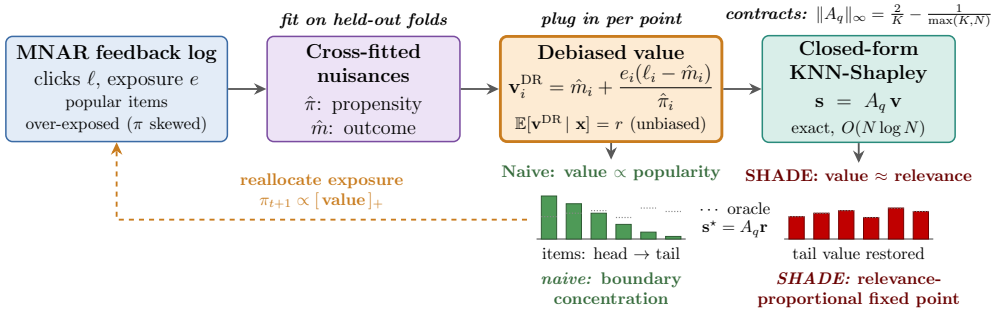

Figure 1: **SHADE overview.** A cross-fitted debiased value enters the unchanged KNN-Shapley recursion. Bottom: naive value follows popularity; SHADE value follows relevance (dotted: oracle). Dashed loop: values guiding future exposure.

estimated from the current log. The resulting feedback is self-reinforcing: items shown more often receive higher measured value and are then shown more often again. Feedback loops of this kind are well documented in recommendation (Chaney et al., 2018; Mansoury et al., 2020); here the loop operates through the valuation layer. Under *naive valuation*, relevance-proportional exposure is not a fixed point, and in expectation the dynamics push the exposure of each item toward its upper or lower limit (Theorem E.6 in Appendix E.10). In simulation, initially unpopular items remain near the exposure floor after fifteen rounds. Popularity-biased valuation therefore creates both a *measurement error*, by assigning credit to the wrong points, and a *control error*, by feeding this distortion into the next log. Our goal is to recover the correct value *level* for every candidate point.

Existing approaches address only part of this problem. Data valuation methods take the observed log as given: they ask how much each point contributes to utility, but they do not model why that point was or was not observed. Debiased recommendation does model exposure through inverse-propensity and doubly-robust estimators (Schnabel et al., 2016; Saito et al., 2020; Wang et al., 2019), but these methods correct the training or evaluation objective of the recommender rather than the credit assigned to individual data points. Restricting valuation to exposed interactions is also insufficient. Although a logged click is an unbiased signal of relevance conditional on exposure, this restriction removes never-exposed candidates from the game and discards much of the long tail. The correction must therefore apply to every candidate, exposed or not, and it must occur before the cooperative game is formed. What is missing is an MNAR correction at the per-point value layer whose guarantees propagate through the Shapley operator.

This is the role of SHADE (SHApley with Debiased Estimates; Figure 1). The closed-form KNN-Shapley recursion of Jia et al. (2019a) was derived for binary label matches, but we show that it returns exact Shapley values for arbitrary fixed per-point scalars (Lemma E.2). This allows the value of each point to be corrected before the recursion runs, while preserving the $O(N \log N)$ cost per query. A natural correction is inverse-propensity weighting, which divides each logged click by its exposure propensity. This estimator is unbiased when the propensity model is correct, but its variance is largest for the rarely exposed points that most need correction, and it still assigns zero to never-exposed candidates. SHADE instead uses a doubly-robust value (Robins et al., 1994): an outcome-model prediction plus a propensity-weighted correction when a label is observed. The value is defined for every candidate, is unbiased if either the propensity or outcome model is correct, and otherwise has bias proportional to the product of the two nuisance errors. Because the KNN-Shapley recursion is a fixed linear map that contracts for $K \geq 2$, these guarantees pass from the per-point values to the resulting Shapley values without amplification. In the feedback setting, the corrected values also restore relevance-proportional exposure as the expected fixed point.

We evaluate SHADE on synthetic MNAR benchmarks with exact oracle valuation and on real recommendation data with randomized-exposure references or natural production exposure. Across these settings, *naive valuation* systematically under-values the long tail, while SHADE recovers the intended value level relative to the available oracle or randomized reference. The correction also assigns value to never-exposed pairs, enabling augmentation and prospective collection strategies that naive valuation cannot support. The experiments identify where this recovery becomes difficult: ranking quality depends on the outcome model, nuisance misspecification can over-correct the popularity–value relationship, and downstream accuracy gains vary across settings. The central

empirical result is that correcting selective exposure restores long-tail data value; downstream gains follow when that corrected value changes which data can be used or collected.

Contributions.

1. **SHADE** (Section 4): we place a cross-fitted doubly-robust value inside the exact KNN-Shapley recursion, preserving its cost while obtaining double-robustness, contraction, and ranking guarantees.
2. **Theory** (Section 5): we characterize the long-tail bias of naive Shapley valuation, derive a rank-inversion condition, and analyze value-guided exposure dynamics.
3. **Experiments** (Section 6): we evaluate the method across synthetic and real recommendation settings, including external baselines, million-scale valuation, and failure cases.

2 Related Work

Four lines of work border this paper (extended discussion: Appendix C). *Learning from corrupted labels.* Our observation model is the SAR-PU setting (Elkan & Noto, 2008; Bekker et al., 2019), written as a corrupted-label kernel (van Rooyen & Williamson, 2018). This literature corrects the *aggregate risk* of a learner, not the value attributed to individual points by a cooperative game. *Data valuation.* Data Shapley and its closed-form KNN special case (Ghorbani & Zou, 2019; Jia et al., 2019a) are now applied to recommendation logs (Kersbergen et al., 2025; Zhang et al., 2025), alongside model-free and validation-free alternatives (Just et al., 2023; Kessler et al., 2025; Xu et al., 2021; Pruthi et al., 2020; Seedat et al., 2022). They differ in how they score a point, but all take the observed utility as given and do not model why a point was observed (Section 6.4). *Debiased recommendation.* IPS and doubly-robust estimators correct the *training or evaluation objective* of a recommender for exposure (Schnabel et al., 2016; Saito et al., 2020; Wang et al., 2019). We apply them to a different quantity, the per-point value that a valuation method consumes. *Feedback loops.* Algorithmic confounding amplifies popularity through the predictions of a retrained *model* (Chaney et al., 2018; Chen et al., 2023); we isolate a loop through the *valuation* signal (Section 5). The gap these lines of work leave is MNAR correction at the per-point value layer, propagated through a cooperative-game valuation operator.

3 Setup: Implicit Feedback as SAR-PU Corruption

Each candidate user–item pair has a latent relevance; exposure gates whether it can appear in the log, so the implicit signal is a selectively observed version of relevance. Let $D = \{(u_i, \mathrm{item}_i, \ell_i, \mathbf{x}_i)\}_{i=1}^{N}$ be the candidate pairs, with $\mathbf{x}_i \in \mathbb{R}^d$ a joint user–item representation and $\ell_i \in \{0, 1\}$ the observed signal (notation in Table 2). Each pair has true binary relevance $y = \ell^{\mathrm{true}}$ with $r(\mathbf{x}) = \mathbb{P}(y{=}1 \mid \mathbf{x})$. The log records (e, ℓ) instead of y: exposure $e \sim \mathrm{Bern}(\pi(\mathbf{x}))$ with $y \perp e \mid \mathbf{x}$ (Assumption 1(ii)), and $\ell = e \cdot y$. All N candidates are players, exposed or not: the never-exposed ones are what make the problem MNAR (Section 4).

Each point is in one of three states, positive P $(e{=}1, \ell{=}1)$, exposed-unclicked C $(e{=}1, \ell{=}0)$, or never-exposed U $(e{=}0)$, generated by the Markov corruption kernel (van Rooyen & Williamson, 2018) $\mathbb{P}(P \mid y{=}1, \mathbf{x}) = \pi(\mathbf{x})$, $\mathbb{P}(C \mid y{=}0, \mathbf{x}) = \pi(\mathbf{x})$, $\mathbb{P}(U \mid y, \mathbf{x}) = 1 - \pi(\mathbf{x})$. This is the SAR-PU labeling mechanism of Bekker et al. (2019); Bekker & Davis (2020), which generalizes SCAR PU learning (Elkan & Noto, 2008). An exposed-unclicked observation is an instance-dependent noisy negative, and merging C with U, as standard negative sampling does, discards that structure. The kernel is parameterized by the propensity $\pi(\mathbf{x})$, which is routinely estimated (Schnabel et al., 2016; Liang et al., 2016), and it is invertible wherever $\pi > 0$ ($\mathbb{E}[e\ell/\pi \mid \mathbf{x}] = r(\mathbf{x})$); the per-point IPS value of Section 4 is its backward correction. The PU literature applies this correction to an aggregate risk, not per point inside a Shapley operator (Section 2).

Queries $\{(\mathbf{x}_q, \ell_q)\}_{q=1}^{Q}$ with known relevance (from a randomized-exposure slice) serve as the test points of KNN-Shapley. An unbiased query set does not remove the bias: queries only *score* coalitions, while the corrupted quantity is the value of each candidate point ($\mathbb{E}[\mathbf{v}_i^{\mathrm{naive}} \mid \mathbf{x}_i] = \pi_i r_i$; Theorem 5.3). All methods in Table 1 receive the same unbiased queries, yet naive reaches only $\tau = 0.205$.

Table 1: Level and rank fidelity on the synthetic MNAR benchmark (λ=2, 8 seeds, all-N players). Among observation-responsive methods, SHADE gives the strongest joint recovery (Section 6). Bold: best per column (nearest oracle target for tail share/Gini; highest τ). The DM ablation, whose τ can exceed that of SHADE, is in Appendix F.2.

value channel	tail share ($\to 0.215$)	Gini ($\to 0.204$)	τ vs. oracle
naive ($e\ell$)	0.002	0.257	0.205
IPS	0.196	**0.221**	0.204
popularity count	0.030	0.434	0.007
density (label-free)	0.193	0.238	−0.016
SHADE	**0.225**	0.175	**0.547**

KNN-Shapley. For a query q and per-point values $\{v_i\}$, define the coalition utility $u_q(S) = \frac{1}{K}\sum_{i\in \mathrm{KNN}_K(S,q)} v_i$, with $\mathrm{KNN}_K(S,q)$ the (at most) K points of S nearest to $\mathbf{x}_q$. For the 0/1 label-match indicator, Jia et al. (2019a) show the Shapley values of this game admit a closed form: sorting by distance to $\mathbf{x}_q$ ascending,

$$s_{(N)} = \frac{v_{(N)}}{N}, \qquad s_{(i)} = s_{(i+1)} + \frac{v_{(i)} - v_{(i+1)}}{K} \cdot \frac{\min(K,i)}{i}. \tag{1}$$

The recursion is exact for any *fixed* per-point scalars v_i (Lemma E.2; we take $N \geq K$, and Appendix E.2 covers $N < K$). What remains open is the scalar that a selectively observed candidate should contribute.

4 Method: Debiased Value Functions

SHADE (*SHApley with Debiased Estimates*) separates statistical correction from cooperative-game attribution. We first estimate what each candidate point would be worth without selective exposure, and then apply the exact KNN-Shapley recursion of Equation 1 unchanged. Because that recursion is exact for arbitrary per-point scalars (Lemma E.2), SHADE is not a new approximate Shapley algorithm: a corrected value is substituted before the existing exact computation runs. SHADE denotes the all-N method with the doubly robust value $\mathbf{v}^{\mathrm{DR}}$ of Equation 2 (ablation labels: Appendix A).

Correcting the observation channel. The *naive* value $\mathbf{v}_i^{\mathrm{naive}} = e_i\ell_i$ counts a logged interaction as one and everything else as zero. Under MNAR exposure, it mixes relevance with exposure: it is biased for r_i by the factor π_i, and the bias propagates to the Shapley values (Theorem 5.3). The *IPS* value $\mathbf{v}_i^{\mathrm{IPS}} = e_i\ell_i/\hat{\pi}_i$ divides each click by its estimated exposure probability. This is the per-point IPS estimator of Schnabel et al. (2016) and the backward correction of the kernel (Section 3). It is unbiased when $\hat{\pi} = \pi$, but its variance is $O(1/\pi_i)$, largest on the tail points it is meant to correct. The *doubly-robust* value adds an outcome model $\hat{m}(\mathbf{x}) \approx r(\mathbf{x})$, fit on exposed points and cross-fitted:

$$\mathbf{v}_i^{\mathrm{DR}} = \hat{m}(\mathbf{x}_i) + \frac{e_i\,(\ell_i - \hat{m}(\mathbf{x}_i))}{\hat{\pi}_i}, \tag{2}$$

the prediction of the model, plus an IPS-weighted correction wherever a label was actually observed. This is the AIPW (augmented inverse-propensity weighting) estimator (Robins et al., 1994) applied per point. It is unbiased if *either* $\hat{\pi} = \pi$ *or* $\hat{m} = r$ (Proposition 5.1) and has second-order error otherwise (Corollary E.3), which is the fallback property SHADE relies on. When $\hat{m}$ is informative, the residual correction can reduce variance relative to pure IPS (small propensities still require clipping). Never-exposed points ($e_i = 0$) receive $\hat{m}(\mathbf{x}_i)$, so the valuation covers all of U.

To isolate what the outcome model contributes, we also report the ablation $\mathbf{v}_i^{\mathrm{DM}} = \hat{m}(\mathbf{x}_i)$ (DM), which is Equation 2 without the residual correction. Under cross-fitting, the score of point i excludes its realized (e_i, ℓ_i), with conditional bias

$$\mathbb{E}[\mathbf{v}_i^{\mathrm{DM}} \mid \mathbf{x}_i] - r_i = \hat{m}(\mathbf{x}_i) - r_i, \tag{3}$$

which is first-order in outcome-model error, with no propensity-model fallback (Remark F.1).

The player universe is part of the estimand. We ask for the value of every candidate point, including those that were never exposed, so all N points are players. This also makes the oracle comparison well-posed: the sort order, and hence the operator of Lemma E.1, depends on $\{\mathbf{x}_i\}, \mathbf{x}_q$ alone. Valuing only exposed points asks a different question and conditions away the MNAR problem: $\mathbb{E}[\ell_i \mid \mathbf{x}_i, e_i{=}1] = r_i$ under Assumption 1(ii) makes naive valuation conditionally unbiased there, and the AIPW augmentation then carries a conditional bias of $(1 - 1/\hat{\pi}_i)(\hat{m}_i - r_i)$ per point, which degrades the exposed-only DR ablation (Section 6). Because a Shapley value is defined relative to its player set, changing the candidate universe changes the game and need not preserve absolute values. Our guarantees condition on a fixed universe selected independently of realized exposure. In the production-log experiments, the candidate pairs are therefore sampled and frozen *before* their realized exposure or click is read (Appendix F.11), and Appendix F.23 varies that choice (conclusion unchanged, effect sizes vary).

Complexity. Nuisance fitting is a one-off cost; each query then costs $O(N \log N)$ (sort-dominated), versus Monte-Carlo permutation Shapley, whose sample requirement grows with N (Ghorbani & Zou, 2019); see Section 6.1. Pseudocode is Algorithm 1 (Appendix D).

5 Theory

All proofs are in Section E, and Appendix E.9 checks each bound numerically.

Assumption 1 (Identifiability and regularity)**.** (i) Positivity: $\pi_i \geq \pi_{\min} > 0$ on the player set; in the algorithm the *estimate* is additionally guarded by clipping $\hat{\pi}_i \leftarrow \max(\hat{\pi}_i, \varepsilon)$; (ii) unconfoundedness: $\ell_i^{\text{true}} \perp e_i \mid \mathbf{x}_i$; (iii) boundedness: $\ell_i \in [0,1]$ and, after clipping, $|v_i| \leq B$; (iv) (e_i, ℓ_i) independent across i given $\{\mathbf{x}_i\}$; (v) cross-fitted nuisances, so that $\hat{m}(\mathbf{x}_i), \hat{\pi}_i$ are fitted without (e_i, ℓ_i) and, by (iv), independent of it, and the sort order depends on $\{\mathbf{x}_i\}, \mathbf{x}_q$ only.

Assumption 1(ii) is an identification condition, not only a nuisance-model condition: double robustness protects against misspecification of one nuisance *given* identification, not against unmeasured exposure–relevance confounding (imposed in Appendix F.22, where recovery degrades as predicted).

The recursion is exact and linear. Recursion Equation 1 returns the *exact* Shapley values of $u_q(S) = \frac{1}{K}\sum_{i \in \mathrm{KNN}_K(S,q)} v_i$ for arbitrary fixed per-point scalars $\{v_i\}$ (Lemma E.2; the generic-scalar statement does not appear in Jia et al., 2019a), so substituting a debiased $\mathbf{v}$ costs no approximation. The recursion is also a fixed linear map, and a contraction for $K \geq 2$. With $a_j := 1/\max(K, j)$, Equation 1 is $\mathbf{s} = A_q \mathbf{v}$ for an explicit lower-triangular A_q depending only on the sort order (Lemma E.1), with

$$\|A_q\|_\infty = \frac{2}{K} - \frac{1}{N}, \qquad \|\hat{\mathbf{s}} - \mathbf{s}^\star\|_\infty \leq \left(\tfrac{2}{K} - \tfrac{1}{N}\right)\|\hat{\mathbf{v}} - \mathbf{v}^\star\|_\infty, \qquad \textstyle\sum_i s_i = u_q(D). \tag{4}$$

Since $\|A_q\|_\infty < 1$ for $K \geq 2$, value error does not accumulate through the recursion; the last identity is efficient. Throughout, $\mathbf{s}^\star = A_q \mathbf{r}$ is the oracle valuation (same player set and sort order); expectations are conditional on $\{\mathbf{x}\} := \{\mathbf{x}_i\}_{i=1}^N$ and $\mathbf{x}_q$ (Assumption 1(v) makes A_q deterministic, so linearity applies). Correcting the per-point value therefore also corrects the Shapley attribution:

Proposition 5.1 (Unbiasedness and double robustness)**.** *Under Assumption 1(i,ii,iv,v), conditionally on the out-of-fold nuisance fit, the pointwise value bias obeys the exact AIPW identity*

$$b_i := \mathbb{E}[\mathbf{v}_i^{\mathrm{DR}} \mid \mathbf{x}_i] - r_i = \left(\hat{m}(\mathbf{x}_i) - r_i\right)\frac{\hat{\pi}_i - \pi_i}{\hat{\pi}_i}, \tag{5}$$

so $\mathbb{E}[\mathbf{s}_i^{\mathrm{DR}} \mid \{\mathbf{x}\}, \mathbf{x}_q] = \mathbf{s}_i^\star$ *if* $\hat{\pi} = \pi$ or $\hat{m} = r$. *In particular, with* $\hat{\pi} = \pi$ *(clipping not binding),* $\mathbb{E}[\mathbf{s}_i^{\mathrm{IPS}} \mid \{\mathbf{x}\}, \mathbf{x}_q] = \mathbf{s}_i^\star$ *for all* i, q*: IPS is the* $\hat{m} \equiv 0$ *special case.*

The pointwise identity is classical AIPW (Robins et al., 1994); what matters is that it passes unchanged through the deterministic all-N operator to the Shapley values. The bias is a *product* of two errors and vanishes when either factor does, which is why $\mathbf{v}^{\mathrm{DR}}$ is the primary value of SHADE. When both nuisances are wrong, pushing Equation 5 through Equation 4 (Corollary E.3) gives, with $\Delta_m := \|\hat{m} - r\|_\infty$, $\Delta_\pi := \|\hat{\pi} - \pi\|_\infty$, for every q, i,

$$\left|\mathbb{E}[\mathbf{s}_i^{\mathrm{DR}}] - \mathbf{s}_i^\star\right| \;\leq\; \|A_q\|_\infty \max_j |b_j| \;\leq\; \left(\tfrac{2}{K} - \tfrac{1}{N}\right)\frac{\Delta_m \Delta_\pi}{\varepsilon} \;=:\; \varepsilon_{\mathrm{b}}. \tag{6}$$

The $\mathbf{v}^{\mathrm{DR}}$ bias is thus *second order*, $O(\Delta_m \Delta_\pi)$; IPS is first order in Δ_π and naive zeroth order. Neither result is specific to the Shapley game:

Corollary 5.2 (Linear solution concepts inherit the correction)**.** *Let $\tilde{u}_q$ be any coalition utility linear in $\mathbf{v}$ with coefficients determined by $\{\mathbf{x}\}, \mathbf{x}_q$ alone (the KNN utility above, its threshold variant), and let ϕ be any solution concept linear in the game (any semivalue: Shapley, Banzhaf, Beta-Shapley). Then $\mathbf{s}^\phi = A_q^\phi \mathbf{v}$ for a deterministic matrix A_q^ϕ, and Proposition 5.1 and Equation 6 hold verbatim with $\|A_q^\phi\|_\infty$ in place of $2/K - 1/N$.*

The exposure distortion enters before any game is defined, so a linear valuation operator inherits whichever bias is present in the value vector, and correcting $\mathbf{v}$ corrects every linear concept at once (Section 6.4); only the contraction constant in Equation 4 is Shapley-specific. The game layer still matters: query aggregation amplifies value-level bias and adds fidelity to corrected values, more than doubling the corrected-vs-naive recovery gap (Appendix F.14). Beyond bias, Proposition E.4 (Appendix E.7) shows that the query-averaged values converge to the oracle in probability under consistent nuisances, $Q \to \infty$, and vanishing query concentration ($\chi^* \to 0$). Under naive valuation, selective exposure redistributes value as follows.

Theorem 5.3 (Long-tail bias of naive valuation)**.** *Under Assumption 1(i,ii,v), $\mathbb{E}[\mathbf{v}_i^{\mathrm{naive}} \mid \mathbf{x}_i] = \pi_i r_i$ exactly. For a query q with point i at rank ρ ($a_\rho = 1/\max(K, \rho)$), the Shapley bias decomposes as*

$$\mathbb{E}[\mathbf{s}_i^{\mathrm{naive}} - \mathbf{s}_i^\star] = -a_\rho(1-\pi_i)r_i \;+\; C_\rho, \quad C_\rho = \sum_{j > \max(\rho, K)} \frac{(1-\pi_{(j)})r_{(j)}}{j(j-1)} \;\geq 0.$$

(a) $\partial_{\pi_i} \mathbb{E}[\mathbf{s}_i^{\mathrm{naive}} - \mathbf{s}_i^\star] = a_\rho r_i > 0$; for two points h, t with $r_h = r_t = r$ at adjacent ranks $\rho, \rho{+}1$, $\mathbb{E}\big[(\mathbf{s}_h^{\mathrm{naive}} - \mathbf{s}_t^{\mathrm{naive}}) - (\mathbf{s}_h^\star - \mathbf{s}_t^\star)\big] = a_\rho\, r\, (\pi_h - \pi_t)$, free of cross-terms. (b) $\mathbb{E}[\mathbf{s}_i^{\mathrm{naive}} - \mathbf{s}_i^\star] \gtrless 0 \iff (1-\pi_i)\, r_i \lessgtr C_\rho / a_\rho$. (c) Under the conditions of Proposition 5.1, $\mathbb{E}[\mathbf{s}_i^{\mathrm{DR}} - \mathbf{s}_i^\star] = 0$ for all π; in general it obeys the Corollary E.3 bound, which is free of π_i.

Part (a) is the exact rate at which value moves from low- to high-exposure points; (b) says each point also receives a share C_ρ of the value that more distant points lose, so high-π points can be overvalued; (c) says that, with either nuisance correct, $\mathbf{v}^{\mathrm{DR}}$ removes the whole π-structure of the bias, not just its magnitude.

Exposure also changes the ordering. By Corollary E.5(a) (Appendix E.10), $\mathbf{s}_i^\star - \mathbf{s}_j^\star > 2\varepsilon_{\mathrm{b}} \Rightarrow \mathbb{E}[\mathbf{s}_i^{\mathrm{DR}}] > \mathbb{E}[\mathbf{s}_j^{\mathrm{DR}}]$: SHADE orders correctly, in expectation, every oracle pair separated by more than $2\varepsilon_{\mathrm{b}}$, and hence *every* oracle-separated pair when either nuisance is correct. *Naive misranking can be checked directly* (Corollary E.5(b)): for a head point h and adjacent-rank tail point t with $r_t > r_h$ but $\pi_h r_h > \pi_t r_t$, the adjacent-rank identity gives exactly

$$\mathbf{s}_h^\star - \mathbf{s}_t^\star = a_\rho(r_h - r_t) < 0, \qquad \mathbb{E}[\mathbf{s}_h^{\mathrm{naive}} - \mathbf{s}_t^{\mathrm{naive}}] = a_\rho(\pi_h r_h - \pi_t r_t) > 0, \tag{7}$$

so the pair is inverted in expectation whenever $\pi_h/\pi_t > r_t/r_h > 1$, that is, whenever the exposure gap exceeds the relevance gap, a condition that depends on (π, r) alone.

Unbiasedness is not ranking quality. The DM ablation has first-order outcome-model bias and no propensity-model fallback, yet it avoids the residual/IPW variance term: under a strong $\hat{m}$ it can reach a higher Kendall τ than SHADE, and it collapses when $\hat{m}$ carries little relevance signal (Remark F.1, Appendix F.2). We therefore do not claim uniform rank dominance for SHADE.

What happens when a platform acts on these values? A canonical action is to reallocate exposure toward measured value: $\pi_{t+1} = \mathrm{clipnorm}(\mathbb{E}[V_t \mid \pi_t])$, where V_t is the per-item mean value and clipnorm enforces the exposure budget and bounds. Theorem E.6 (Appendix E.10; using Proposition 5.1) characterizes these expected dynamics. Under naive valuation, $\mathbb{E}[V \mid \pi] = \pi \odot \mathbf{r}$, so $\pi \mapsto \mathrm{clipnorm}(\pi \odot \mathbf{r})$: $\pi \propto \mathbf{r}$ is not generally a fixed point, and every fixed point places all but at most one item at a box boundary, monotonically in relevance. Under SHADE, if either nuisance is correct, $\mathbb{E}[V \mid \pi] = \mathbf{r}$, so $\pi^\star_{\mathrm{SHADE}} = \mathrm{clipnorm}(\mathbf{r})$ is the unique fixed point, reached in one expected step. With bounded nuisance-tracking error, every perturbed fixed point lies within $O(\Delta_m \Delta_\pi / \epsilon)$ of it in ℓ_1 (Proposition E.7). This result concerns only the expected value channel, not the full KNN-Shapley loop; Section 6.2 simulates the latter to test whether the concentration drives an initially popularity-defined tail toward the exposure floor. It makes no finite-sample ordering claim.

6 EXPERIMENTS

We ask whether correcting selective exposure recovers the intended data value and what limits the recovery (Section 6.1), what changes when values are acted on (Sections 6.2 and 6.3), and whether the benefit is specific to KNN-Shapley (Section 6.4).

Data and metrics. The synthetic MNAR benchmark has latent-factor true relevance r and popularity/relevance-correlated exposure with bias strength λ; there the oracle valuation $\mathbf{s}^\star = A_q\mathbf{r}$ is exact. Five real datasets cover three exposure regimes: real MNAR exposure with unbiased references (Coat, Schnabel et al., 2016; KuaiRand-Pure, Gao et al., 2022b); real outcomes with imposed, known MNAR exposure, which admits the oracle-τ metric (KuaiRec fully-observed block, Gao et al., 2022a; MovieLens-25M, Harper & Konstan, 2015; Amazon Reviews, Hou et al., 2024); and real production exposure at scale (KuaiRec big matrix, $N{=}10^6$ players; level metrics only). Nuisances are cross-fitted logistic models; numbers are mean $\pm$ std over ≥ 5 seeds (Appendix F). *Level fidelity* (tail share, Gini, popularity-decoupling) matters when values are paid out or acted on (Theorem E.6). *Rank fidelity* (Kendall τ) is a bias–variance property, which is why the low-variance DM ablation can rank well without responding to the observed labels (Remark F.1).

6.1 RECOVERING DATA VALUE UNDER MNAR EXPOSURE

With a known oracle, the predicted bias and its correction both appear. On the synthetic benchmark, naive valuation leaves the long tail almost no value, and SHADE restores the oracle share (Table 1); the decile-resolved bias shows the sign flip of Theorem 5.3(b) (Figure 2, Appendix F.12). SHADE also holds τ near 0.55 at every bias strength, against 0.12–0.22 for naive and IPS. The all-N player set is essential, and nonlinear-exposure and real-relevance replications rule out circularity (Appendices F.2, F.16).

The recovery survives real outcomes and scale. The closed form stays exact where prior recommendation-Shapley work needs Monte-Carlo sampling (Kersbergen et al., 2025) or amortized estimators (Zhang et al., 2025), so the full $6{\times}10^6$-point ML-25M pool can be valued on one CPU instance. There, SHADE recovers the oracle ranking better than naive valuation ($\tau = 0.268$ against 0.187, at every seed and scale) and returns a tail that naive valuation under-allocates about 500-fold to the order of the oracle. The fully observed KuaiRec block shows the same pattern (Appendix F.7), and also the limit of the ranking claim: under a strong outcome model the DM ablation nearly matches SHADE (0.225 against 0.235; Remark F.1).

Outcome-model quality limits rank recovery. Amazon Reviews is the clearest case: with $\mathrm{corr}(\hat{m}, r) \approx 0$ the DM ablation collapses ($\tau = 0.011$) and IPS falls *below* naive, while SHADE still recovers tail mass (0.060 vs. oracle 0.072; naive 0.0002). Level recovery thus remains possible where rank recovery is not. Across KuaiRec, ML-25M and Amazon, the true-nuisance ceiling ($\tau = 0.60$–0.90) is far above estimated SHADE (0.20–0.26), and nuisance interpolation localizes the gap to the outcome model: setting $\hat{m}$ to its true value alone recovers it, setting $\hat{\pi}$ alone does not (Figure 6b, Appendix F.20). The remaining rank error thus arises in outcome estimation (Section 6.3).

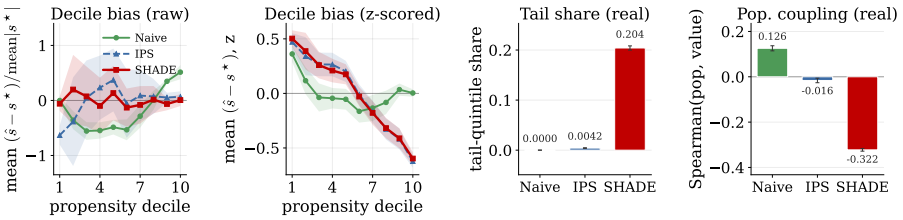

Figure 2: **Bias structure, synthetic and real** ($\pm$std bands). *Panels 1–2* (synthetic, $\lambda{=}2$, 8 seeds): bias by true-propensity decile; naive under-values the mid/tail and over-values the head (Theorem 5.3), IPS corrects the head but not the clipped tail, SHADE is near-flat. *Panels 3–4* (KuaiRec big matrix, $N{=}10^6$, 5 seeds): naive gives the tail 0.0000 and alone rewards exposure; SHADE recovers the tail but over-corrects the coupling.

Does level recovery survive natural exposure? With natural exposure and the candidate universe frozen before it is read (Appendix F.11), naive valuation gives the tail 0.0000 ± 0.0000 on the KuaiRec big matrix and SHADE recovers 0.204 ± 0.004 (Figure 2). The coupling is less favorable: naive alone

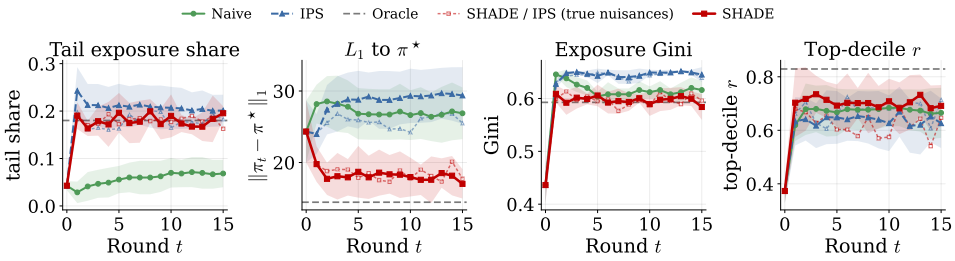

Figure 3: **Valuation-guided exposure over 15 rounds** (5 seeds, ±std). Naive holds the initial-popularity tail quintile near the exposure floor; SHADE converges near $\pi^\star$ in about two rounds; IPS overshoots (large weights on rare clicks). Faint dashed: true nuisances; grey dashed: oracle loop.

rewards exposure (Spearman $+0.13$ vs. a measured ≈ 0 target), but SHADE *over-corrects* to -0.32, consistent with the measured $\hat{m}$ miscalibration acting through Equation 5. On KuaiRand-Pure, naive gives 0.009 ± 0.020 (randomized-slice reference 0.175) and SHADE *overshoots* to 0.299 ± 0.057: the log is so sparse that the residual correction rarely applies, leaving SHADE $\hat{m}$-dominated and the coupling measurement unable to separate methods (Table 16). The natural-log evidence supports recovery of long-tail value levels, not universal decorrelation between popularity and measured value.

Long-tail under-valuation is consistent across datasets. The same tail under-allocation appears on all six datasets, with the corrected methods at or near each available reference (Figure 6a; payments and the exposed/never-exposed split: Appendices F.18, F.17; runtime and memory: Appendix F.7).

6.2 When data values drive future exposure

If a platform allocates future exposure by measured data value, does the value-layer bias survive the full Shapley game and change what is collected? Theorem E.6 covers the expected value channel only, so we simulate the full KNN-Shapley loop with nuisances refit each round (Figure 3; Appendix F.6). The concentration reappears. Naive valuation holds the exposure share of the initial-popularity tail quintile at 0.068 ± 0.029 (reference 0.187); SHADE brings it to 0.195 ± 0.015 within two rounds, indistinguishable from true-nuisance SHADE, and the ordering is unchanged as $\pi_{\min}$, catalog size, and budget vary (Appendix F.19). The theorem guarantees boundary concentration ordered by relevance, not that the initial-popularity tail sits at the lower bound; that part is empirical.

Correcting valuation matters most when it changes what is collected. Reweighting a *fixed* log does not lift the head-dominated *top line* of a real ranker (Appendix F.8); steering *collection* changes the data. With identical iALS rankers (Hu et al., 2008) trained on the log each policy collects, SHADE-collected logs match oracle-collected logs on the tail segment (tail-Recall@20 0.096 ± 0.003 vs. 0.097) while naive-collected logs collapse it (0.002); the top line instead *favors* naive collection, above even the oracle policy, since it rewards the concentration itself. On ML-25M real relevance, the allocation of SHADE tracks the oracle policy, but tail ranking metrics are zero for every policy *including oracle*: the collection behavior transfers, the tail ranking gain does not (Appendix F.9).

6.3 When corrected valuation improves downstream learning

Corrected valuation assigns a value to never-exposed candidates ($\hat{m}$ on U, where the naive value is identically 0; Section 4). Naive valuation has no such channel, which explains which interventions work. Reweighting or pruning a fixed log cannot recover relevant pairs the log never contained: no reweighting-only intervention improves top-line accuracy, and pruning is harmful or null throughout (Tables 7 and 14, Appendix F.8). Augmentation and prospective collection can reach those pairs (Figure 4).

Augmentation from the unexposed set. Adding the top-valued *never-exposed* pairs as pseudo-positives to iALS training (budget 5% of observed clicks) yields significant top-line gains: ML-25M NDCG@20 $+0.0010$ ($\sim +4.1\%$) via direct $\hat{m}$ ranking, and KuaiRec $+0.0005$ ($+5.1\%$, pooled 24 seeds, side-information-enriched $\hat{m}$), matching or exceeding an *oracle* selection (Table 14). The gain comes from the coverage of U by the outcome-model channel, not from Shapley being the best selector there: an $\hat{m}$ refit without IPS weights gives the same gain, and direct $\hat{m}$ ranking beats the full Shapley score (Appendix F.9).

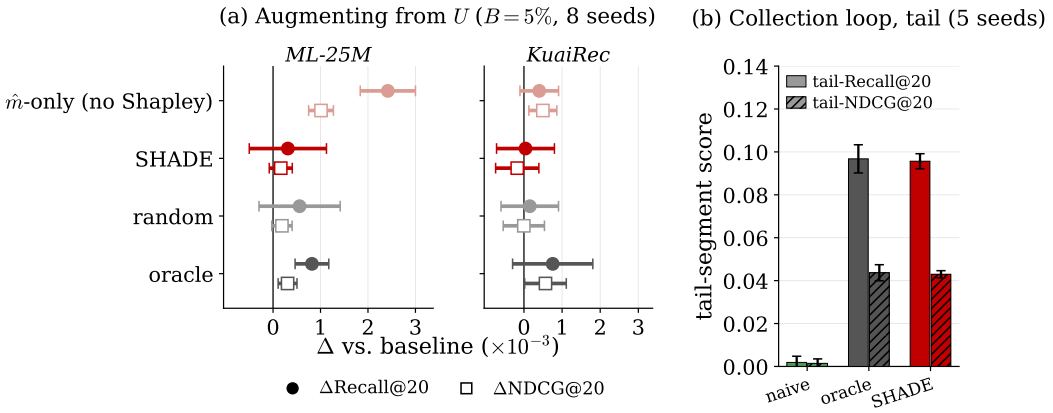

Figure 4: **NDCG@20 and Recall@20 for augmentation and collection** (iALS, uniform-exposure eval, 95% CIs). **(a)** Augmentation from U (naive has no value on U; KuaiRec shows the original 8-seed batch, superseded by the pooled 24-seed analysis in Appendix F.10). **(b)** Collection loop, tail segment (synthetic; solid = tail-Recall@20, hatched = tail-NDCG@20): SHADE-collected logs match oracle-collected logs; naive-collected logs collapse the tail.

Composition and confirmation. Augmentation composes with reweighting. Because the reweighting scheme it builds on (Table 7) was selected on the 16 discovery seeds, we froze the procedure and report its evaluation on 16 disjoint fresh seeds as primary. The gain persists (NDCG@20 $+0.0016$, $+6.6\%$; 95% CI $[+0.0012, +0.0020]$), above reweighting alone, augmentation alone, and a random-augmentation control, with no interaction between the two, so it is not explained by the original selection step (Appendix F.9).

Where the gain stops. The effect persists under an EASE ranker at a smaller size, and heuristics that use no exposure correction match it only in isolated pre-registered settings. On KuaiRand-Pure, a natural MNAR log with a randomized evaluation slice and a frozen *exposure-independent candidate universe*, augmentation is null at every tested budget. Weak outcome-model signal, sparse observed support, and small effective intervention size are intertwined there, so the experiment identifies a persistent null, not its unique mechanism; iterating the augmentation is also a measured negative, so selection is one-shot (Appendix F.10). Debiased valuation therefore creates an additional downstream action; it does not guarantee an accuracy gain on every natural log.

6.4 Is the correction specific to KNN-Shapley?

It is not. Changing the valuation rule while keeping the raw MNAR value leaves the exposure distortion in place. Fourteen external alternatives, from other solution concepts to model-free scores, each in its best-case configuration under the same MNAR protocol, show this (Table 18; Appendices F.13, F.15, F.21): *none combines oracle-level tail share with oracle-ranking recovery*, and the highest external all-N rank correlation is nearly orthogonal to the oracle, against the 0.547 ± 0.020 of SHADE. Correcting the value channel composes with the method: replacing the raw $e \cdot \ell$ value by $\mathbf{v}^{\mathrm{DR}}$ corrects the MNAR distortion for the linear solution concepts, as Corollary 5.2 predicts, and a $\mathbf{v}^{\mathrm{DR}}$ pseudo-outcome target improves TracIn in both settings; the two choices are complementary.

7 Conclusion

Selective exposure contaminates data value before any game is formed: the naive per-point value inherits the logging policy, and the Shapley values built on it move credit toward high-exposure items. SHADE corrects that value layer and keeps the exact closed-form recursion, with double-robustness and ranking guarantees under the stated assumptions. SHADE recovers long-tail value levels in every setting we tested, and acting on corrected values changes which data are collected. Rank recovery is limited by nuisance quality, $\hat{m}$ miscalibration can *over-correct* the popularity–value coupling, and accuracy gains are protocol-specific: the natural-log KuaiRand augmentation remains null.

The guarantees require positivity, a player universe fixed independently of realized exposure, and no unmeasured exposure–relevance confounding given $\mathbf{x}$ (sensitivity in Appendices F.22 and F.23).

ETHICS STATEMENT

This work characterizes and corrects a mechanism by which measured data value is systematically transferred from long-tail to popular items. Where data values are used for curation, acquisition, marketplaces, or creator compensation, uncorrected valuation would structurally underpay contributors of long-tail data; the corrections we propose reduce that structural unfairness, and the feedback-loop analysis shows that acting on uncorrected values would deepen it. The work involves no human subjects; all datasets are public research datasets used under their respective licenses.

REPRODUCIBILITY STATEMENT

The closed-form recursion is verified against brute-force Shapley enumeration to 4×10^{-15} and was replicated by a third, independently written implementation. Every theorem that admits a numerical check is accompanied by one in the code release. All reported numbers were recomputed from the archived result files by an independent verification script ($1{,}445$ checks covering every experiment in the paper), and each matched to its stated precision. For the production logs, these checks also confirm that every player universe was fixed before any exposure lookup. A fresh-seed re-execution of the core synthetic and Coat experiments (seeds 100–107) reproduces their method orderings and headline numbers within seed-to-seed variability; the larger-scale and accuracy experiments are covered by recomputation only. The one quantity that varies materially across seeds is the deepest-tail IPS bias decile, where clipping makes the raw estimate high-variance; this does not change any ordering. The confounding-sensitivity experiment (Appendix F.22) includes unit tests confirming that the latent confounder is excluded from the nuisance and KNN features, that $\gamma{=}0$ recovers the unconfounded generator exactly, and that the marginal exposure rate stays within tolerance across the sweep. The ML-25M composition result was re-run once under a configuration fixed before the run, on 16 seeds disjoint from those used for selection. In the candidate-universe experiment (Appendix F.23), each universe is fixed before exposure is imposed and is shared by all methods and the oracle. All code, experiment and verification scripts, and per-seed outputs are released.

AI USE STATEMENT

We used large language models solely to polish the writing of this paper, including checking typos, grammar, and phrasing. They were not used to generate research ideas or design experiments; all such contributions were made entirely by the authors, who take full responsibility for the content of the paper.

REFERENCES

Jessa Bekker and Jesse Davis. Learning from positive and unlabeled data: A survey. *Machine Learning*, 109, 2020.

Jessa Bekker, Pieter Robberechts, and Jesse Davis. Beyond the selected completely at random assumption for learning from positive and unlabeled data. In *Joint European Conference on Machine Learning and Knowledge Discovery in Databases (ECML-PKDD)*, 2019.

Asia J. Biega, Krishna P. Gummadi, and Gerhard Weikum. Equity of attention: Amortizing individual fairness in rankings. In *Proceedings of the 41st International ACM SIGIR Conference on Research and Development in Information Retrieval (SIGIR)*, 2018.

Allison J. B. Chaney, Brandon M. Stewart, and Barbara E. Engelhardt. How algorithmic confounding in recommendation systems increases homogeneity and decreases utility. In *Proceedings of the 12th ACM Conference on Recommender Systems (RecSys)*, 2018.

Jiawei Chen, Hande Dong, Xiang Wang, Fuli Feng, Meng Wang, and Xiangnan He. Bias and debias in recommender system: A survey and future directions. *ACM Transactions on Information Systems*, 41(3), 2023.

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

Yu-Ting Chou, Gang Niu, Hsuan-Tien Lin, and Masashi Sugiyama. Unbiased risk estimators can mislead: A case study of learning with complementary labels. In *Proceedings of the 37th International Conference on Machine Learning (ICML)*, 2020.

Charles Elkan and Keith Noto. Learning classifiers from only positive and unlabeled data. In *Proceedings of the 14th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining (KDD)*, 2008.

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

Chongming Gao, Shijun Li, Yuan Zhang, Jiawei Chen, Biao Li, Wenqiang Lei, Peng Jiang, and Xiangnan He. KuaiRand: An unbiased sequential recommendation dataset with randomly exposed videos. In *Proceedings of the 31st ACM International Conference on Information and Knowledge Management (CIKM)*, 2022b.

Amirata Ghorbani and James Zou. Data Shapley: Equitable valuation of data for machine learning. In *Proceedings of the 36th International Conference on Machine Learning (ICML)*, 2019.

Siyuan Guo, Lixin Zou, Yiding Liu, Wenwen Ye, Suqi Cheng, Shuaiqiang Wang, Hechang Chen, Dawei Yin, and Yi Chang. Enhanced doubly robust learning for debiasing post-click conversion rate estimation. In *Proceedings of the 44th International ACM SIGIR Conference on Research and Development in Information Retrieval (SIGIR)*, 2021.

F. Maxwell Harper and Joseph A. Konstan. The MovieLens datasets: History and context. *ACM Transactions on Interactive Intelligent Systems*, 5(4), 2015.

Yupeng Hou, Jiacheng Li, Xiangjun Fu, Zhankui He, An Yan, Xiusi Chen, and Julian McAuley. Bridging language and items for retrieval and recommendation: Benchmarking LLMs as semantic encoders. *arXiv preprint arXiv:2403.03952*, 2024.

Yifan Hu, Yehuda Koren, and Chris Volinsky. Collaborative filtering for implicit feedback datasets. In *Proceedings of the 8th IEEE International Conference on Data Mining (ICDM)*, pp. 263–272, 2008.

Takashi Ishida, Gang Niu, Weihua Hu, and Masashi Sugiyama. Learning from complementary labels. In *Advances in Neural Information Processing Systems (NeurIPS)*, 2017.

Ruoxi Jia, David Dao, Boxin Wang, Frances Ann Hubis, Nezihe Merve Gurel, Bo Li, Ce Zhang, Costas J. Spanos, and Dawn Song. Efficient task-specific data valuation for nearest neighbor algorithms. *Proceedings of the VLDB Endowment*, 12(11), 2019a.

Ruoxi Jia, David Dao, Boxin Wang, Frances Ann Hubis, Nick Hynes, Nezihe Merve Gürel, Bo Li, Ce Zhang, Dawn Song, and Costas J. Spanos. Towards efficient data valuation based on the Shapley value. In *Proceedings of the 22nd International Conference on Artificial Intelligence and Statistics (AISTATS)*, 2019b.

Hoang Anh Just, Feiyang Kang, Jiachen T. Wang, Yi Zeng, Myeongseob Ko, Ming Jin, and Ruoxi Jia. LAVA: Data valuation without pre-specified learning algorithms. In *Proceedings of the 11th International Conference on Learning Representations (ICLR)*, 2023.

Masahiro Kato, Shuting Wu, Kodai Kureishi, and Shota Yasui. Automatic debiased learning from positive, unlabeled, and exposure data. *arXiv preprint arXiv:2303.04797*, 2023.

Barrie Kersbergen, Olivier Sprangers, Bojan Karlaš, Maarten de Rijke, and Sebastian Schelter. Scalable data debugging for neighborhood-based recommendation with data Shapley values. In *Proceedings of the 19th ACM Conference on Recommender Systems (RecSys)*, 2025.

Samuel Kessler, Tam Le, and Vu Nguyen. SAVA: Scalable learning-agnostic data valuation. In *Proceedings of the 13th International Conference on Learning Representations (ICLR)*, 2025.

Ryuichi Kiryo, Gang Niu, Marthinus C. du Plessis, and Masashi Sugiyama. Positive-unlabeled learning with non-negative risk estimator. In *Advances in Neural Information Processing Systems (NeurIPS)*, 2017.

Pang Wei Koh and Percy Liang. Understanding black-box predictions via influence functions. In *Proceedings of the 34th International Conference on Machine Learning (ICML)*, 2017.

Yongchan Kwon and James Zou. Beta Shapley: A unified and noise-reduced data valuation framework for machine learning. In *Proceedings of the 25th International Conference on Artificial Intelligence and Statistics (AISTATS)*, 2022.

Yongchan Kwon and James Zou. Data-OOB: Out-of-bag estimate as a simple and efficient data value. In *Proceedings of the 40th International Conference on Machine Learning (ICML)*, 2023.

Haoxuan Li, Chunyuan Zheng, and Peng Wu. StableDR: Stabilized doubly robust learning for recommendation on data missing not at random. In *Proceedings of the 11th International Conference on Learning Representations (ICLR)*, 2023.

Dawen Liang, Laurent Charlin, James McInerney, and David M. Blei. Modeling user exposure in recommendation. In *Proceedings of the 25th International Conference on World Wide Web (WWW)*, 2016.

Ziqi Liu, Kiljae Lee, Yuan Zhang, and Weijing Tang. First-order efficiency for probabilistic value estimation via a statistical viewpoint. *arXiv preprint arXiv:2605.02827*, 2026.

Masoud Mansoury, Himan Abdollahpouri, Mykola Pechenizkiy, Bamshad Mobasher, and Robin Burke. Feedback loop and bias amplification in recommender systems. In *Proceedings of the 29th ACM International Conference on Information and Knowledge Management (CIKM)*, 2020.

Keziah Naggita and Julienne LaChance. A case for data valuation transparency via DValCards. In *Proceedings of the 8th AAAI/ACM Conference on AI, Ethics, and Society (AIES)*, 2025.

Juan C. Perdomo, Tijana Zrnic, Celestine Mendler-Dünner, and Moritz Hardt. Performative prediction. In *Proceedings of the 37th International Conference on Machine Learning (ICML)*, 2020.

Garima Pruthi, Frederick Liu, Mukund Sundararajan, and Satyen Kale. Estimating training data influence by tracing gradient descent. In *Advances in Neural Information Processing Systems (NeurIPS)*, 2020.

James M. Robins, Andrea Rotnitzky, and Lue Ping Zhao. Estimation of regression coefficients when some regressors are not always observed. *Journal of the American Statistical Association*, 89(427), 1994.

Yuta Saito, Suguru Yaginuma, Yuta Nishino, Hayato Sakata, and Kazuhide Nakata. Unbiased recommender learning from missing-not-at-random implicit feedback. In *Proceedings of the 13th International Conference on Web Search and Data Mining (WSDM)*, 2020.

Tobias Schnabel, Adith Swaminathan, Ashudeep Singh, Navin Chandak, and Thorsten Joachims. Recommendations as treatments: Debiasing learning and evaluation. In *Proceedings of the 33rd International Conference on Machine Learning (ICML)*, 2016.

Nabeel Seedat, Jonathan Crabbé, Ioana Bica, and Mihaela van der Schaar. Data-IQ: Characterizing subgroups with heterogeneous outcomes in tabular data. In *Advances in Neural Information Processing Systems (NeurIPS)*, 2022.

Yinan Shen, Ziao Yang, and Hongfu Liu. Validation-induced Shapley shifts: How validation structure distorts data valuation. *arXiv preprint arXiv:2607.03675*, 2026.

Ashudeep Singh and Thorsten Joachims. Fairness of exposure in rankings. In *Proceedings of the 24th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining (KDD)*, 2018.

Harald Steck. Embarrassingly shallow autoencoders for sparse data. In *Proceedings of the World Wide Web Conference (WWW)*, pp. 3251–3257, 2019.

Brendan van Rooyen and Robert C. Williamson. A theory of learning with corrupted labels. *Journal of Machine Learning Research*, 18(228), 2018.

Jiachen T. Wang and Ruoxi Jia. Data Banzhaf: A robust data valuation framework for machine learning. In *Proceedings of the 26th International Conference on Artificial Intelligence and Statistics (AISTATS)*, 2023.

Jiachen T. Wang, Yuqing Zhu, Yu-Xiang Wang, Ruoxi Jia, and Prateek Mittal. Threshold KNN-Shapley: A linear-time and privacy-friendly approach to data valuation. In *Advances in Neural Information Processing Systems (NeurIPS)*, 2023a.

Wei Wang, Takashi Ishida, Yu-Jie Zhang, Gang Niu, and Masashi Sugiyama. Learning with complementary labels revisited: The selected-completely-at-random setting is more practical. In *Proceedings of the 41st International Conference on Machine Learning (ICML)*, 2024.

Wenjie Wang, Fuli Feng, Xiangnan He, Liqiang Nie, and Tat-Seng Chua. Denoising implicit feedback for recommendation. In *Proceedings of the 14th ACM International Conference on Web Search and Data Mining (WSDM)*, pp. 373–381, 2021.

Xiaojie Wang, Rui Zhang, Yu Sun, and Jianzhong Qi. Doubly robust joint learning for recommendation on data missing not at random. In *Proceedings of the 36th International Conference on Machine Learning (ICML)*, 2019.

Zongwei Wang, Min Gao, Wentao Li, Junliang Yu, Linxin Guo, and Hongzhi Yin. Efficient bi-level optimization for recommendation denoising. In *Proceedings of the 29th ACM SIGKDD Conference on Knowledge Discovery and Data Mining (KDD)*, pp. 2502–2511, 2023b.

Xinyi Xu, Zhaoxuan Wu, Chuan Sheng Foo, and Bryan Kian Hsiang Low. Validation free and replication robust volume-based data valuation. In *Advances in Neural Information Processing Systems (NeurIPS)*, 2021.

Longqi Yang, Yin Cui, Yuan Xuan, Chenyang Wang, Serge Belongie, and Deborah Estrin. Unbiased offline recommender evaluation for missing-not-at-random implicit feedback. In *Proceedings of the 12th ACM Conference on Recommender Systems (RecSys)*, 2018.

Jiangxing Yu, Hong Zhu, Chih-Yao Chang, Xinhua Feng, Bowen Yuan, Xiuqiang He, and Zhenhua Dong. Influence function for unbiased recommendation. In *Proceedings of the 43rd International ACM SIGIR Conference on Research and Development in Information Retrieval (SIGIR)*, 2020.

Xiyu Yu, Tongliang Liu, Mingming Gong, and Dacheng Tao. Learning with biased complementary labels. In *Proceedings of the European Conference on Computer Vision (ECCV)*, 2018.

Yansen Zhang, Xiaokun Zhang, Ziqiang Cui, and Chen Ma. Shapley value-driven data pruning for recommender systems. In *Proceedings of the 31st ACM SIGKDD Conference on Knowledge Discovery and Data Mining (KDD)*, 2025.

APPENDIX

A NOTATION

Table 2: Summary of notation used throughout the paper. Vectors in $\mathbb{R}^N$ are bold ($\mathbf{s} = A_q\mathbf{v}$, $\mathbf{r}$, $\mathbf{b}$); feature vectors $\mathbf{x}_i \in \mathbb{R}^d$ are bold; scalar entries carry subscripts (r_i, b_i, s_i).

symbol	meaning
Data and observation model (Section 3)	
N	number of logged candidate points (the all-N player set)
i, (i)	point index; (i) denotes sort rank w.r.t. a query (rank 1 = nearest)
$\mathbf{x}_i \in \mathbb{R}^d$	joint user–item feature representation of point i
u_i, item_i	user and item of point i
$y_i = \ell_i^{\text{true}}$	true (unobserved) binary relevance
$r_i = r(\mathbf{x}_i)$	true relevance rate $\mathbb{P}(y_i{=}1 \mid \mathbf{x}_i) = \mathbb{E}[\ell_i^{\text{true}} \mid \mathbf{x}_i]$
e_i	exposure indicator, $e_i \sim \text{Bern}(\pi_i)$
$\ell_i \in \{0, 1\}$	observed click signal, $\ell_i = e_i \cdot y_i$
$\pi_i = \pi(\mathbf{x}_i)$	true exposure propensity $\mathbb{P}(e_i{=}1 \mid \mathbf{x}_i)$
P, C, U	positive ($e{=}1, \ell{=}1$), exposed-unclicked ($e{=}1, \ell{=}0$), never-exposed ($e{=}0$)
$\{(\mathbf{x}_q, \ell_q)\}_{q=1}^{Q}$	Q validation queries with known relevance (randomized slice)
λ	synthetic exposure bias-strength parameter
Nuisances and value functions (Section 4)	
$\hat{\pi}_i$	estimated (cross-fitted) propensity *after* clipping, $\hat{\pi}_i \leftarrow \max(\hat{\pi}_i, \varepsilon)$
$\hat{m}(\mathbf{x}_i)$	estimated (cross-fitted) outcome model, $\hat{m} \approx r$
ε	propensity clipping level (default 0.05)
$\mathbf{v}_i^{\text{naive}} = e_i\ell_i$	naive per-point value
$\mathbf{v}_i^{\text{IPS}} = e_i\ell_i/\hat{\pi}_i$	IPS per-point value
$\mathbf{v}_i^{\text{DR}} = \hat{m}_i + e_i(\ell_i - \hat{m}_i)/\hat{\pi}_i$	doubly-robust (AIPW) per-point value
$\mathbf{v}_i^{\text{DM}} = \hat{m}(\mathbf{x}_i)$	outcome-model-only direct-method (DM) ablation
B	bound on $\lvert v_i \rvert$ after clipping (Assumption 1(iii))
J	number of cross-fitting folds
KNN-Shapley operator and values (Sections 3 and 5)	
K	number of nearest neighbors
$\text{KNN}_K(S, q)$	the $\le K$ points of coalition S nearest to $\mathbf{x}_q$
$u_q(S)$	coalition utility $\frac{1}{K}\sum_{i \in \text{KNN}_K(S,q)} v_i$
s_i, $s_i^{(q)}$, $\bar{s}_i$	Shapley value of point i; per-query; query-averaged
$\mathbf{s}^{\text{naive}}, \mathbf{s}^{\text{IPS}}, \mathbf{s}^{\text{DR}}, \mathbf{s}^{\text{DM}} = A_q\mathbf{v}^{\text{DM}}$	Shapley values from naive / IPS / DR / DM per-point values
$\mathbf{s}^\star = A_q\mathbf{r}$	oracle Shapley values (recursion on true relevances)
A_q	the linear KNN-Shapley operator, $\mathbf{s} = A_q\mathbf{v}$ (Lemma E.1)
$a_j = 1/\max(K, j)$	diagonal entry / rank weight at rank j
$\lVert A_q \rVert_\infty$	operator ∞-norm $= 2/K - 1/\max(K, N)$
$\rho = \text{rank}_q(i)$	sort rank of point i for query q
Bias, sensitivity, and dynamics (Section 5)	
$b_i = \mathbb{E}[\mathbf{v}_i^{\text{DR}} \mid \mathbf{x}_i] - r_i$	DR value bias given the out-of-fold fit, $(\hat{m}_i - r_i)(\hat{\pi}_i - \pi_i)/\hat{\pi}_i$
$\Delta_m = \lVert \hat{m} - r \rVert_\infty$	outcome-model error
$\Delta_\pi = \lVert \hat{\pi} - \pi \rVert_\infty$	propensity error of the clipped estimate
C_ρ	nonnegative cross-term in the long-tail bias (Theorem 5.3)
ε_{b}	ranking-consistency envelope $\lVert A_q \rVert_\infty \max_j \lvert b_j \rvert$ (Corollary E.5)
T_1, T_2, T_3	label-noise / bias / query-sampling error terms (Proposition E.4)
χ^*	query-concentration factor $\in (0, 1]$ (Proposition E.4)
$\pi_t(j)$	exposure of item j at feedback round t (Theorem E.6)
$\pi_{\min}, \pi_{\max}$	exposure box bounds; budget $\sum_j \pi_t(j) = \mathcal{B}$
$\mathcal{B}$	exposure budget (Theorem E.6); distinct from the value bound B
$\text{clipnorm}(\cdot)$	budget-feasible clip-and-renormalize map (Theorem E.6)
$\mathbb{E}, \mathbb{P}, \mathbb{R}, \mathbb{1}$	expectation, probability, reals, indicator

Labeling conventions. Unless otherwise stated, SHADE denotes our proposed all-N KNN-Shapley method using the doubly robust value $\mathbf{v}^{\text{DR}}$ in Equation 2. We reserve **DR** for the underlying

AIPW estimator $\mathbf{v}^{\mathrm{DR}}$, the induced quantity $\mathbf{s}^{\mathrm{DR}}$, and ablations that are not the full SHADE method (e.g., exposed-only DR, or a $\mathbf{v}^{\mathrm{DR}}$ pseudo-outcome inside another valuation method). All empirical results for our proposed method are labeled SHADE. We reserve **DM** for the KNN-Shapley ablation $\mathbf{s}^{\mathrm{DM}} = A_q \mathbf{v}^{\mathrm{DM}} = A_q \hat{m}$. In augmentation experiments, direct ranking of never-exposed pairs by $\hat{m}$, without the Shapley recursion, is labeled $\hat{m}$-only; its side-information variant is labeled $\hat{m}$-enriched.

B Limitations and future work.

First, the exactness and transfer results (Lemma E.2, Corollary 5.2) cover valuation operators that are *linear* in the per-point values: KNN-Shapley and other semivalues over the same utility. Retraining-based schemes (permutation Data Shapley, Data-OOB, influence functions) absorb exposure bias through exposed-only training gradients, which a corrected value vector cannot reach; extending the correction to that training channel is left for future work. Second, SHADE presupposes a propensity model, an outcome model with usable relevance signal, and a randomized or fully observed slice for queries and references, which confines our real-data evidence to the few public logs that provide one. We leave to future work a closed-loop evaluation in which nuisances are re-estimated, preferences drift and the candidate universe changes, since Theorem E.6 characterizes only the expected value channel and Section 6.2 simulates the rest.

C Extended Related Work

This appendix expands Section 2. For each neighboring work, we state which object it estimates or corrects, and why that object differs from the per-point value of a selectively observed candidate.

Learning from corrupted and positive-unlabeled data. This literature corrects the *risk* of a learner trained on corrupted labels. Our observation model is its SAR-PU setting: in positive-unlabeled (PU) learning, only positives are ever labeled (Elkan & Noto, 2008; Kiryo et al., 2017), and under the "selected at random" assumption, the labeling propensity depends on the instance (Bekker et al., 2019; Bekker & Davis, 2020). We write it as a corrupted-label kernel in the sense of van Rooyen & Williamson (2018). Complementary-label learning is the multiclass relative of the same idea: under the noiseless click model, an exposed-unclicked observation establishes the complementary label (Ishida et al., 2017), and the channel $y \mapsto \{P, C, U\}$ is an instance-dependent transition kernel with $\pi(\mathbf{x})$ in the role of the kernel (cf. the class-dependent kernels of Yu et al., 2018 and the constant-π SCAR-CLL of Wang et al., 2024).

In the binary one-vs-rest setting, the CLL view adds no machinery beyond the SAR-PU one, so we use SAR-PU vocabulary throughout. Estimators in this family rewrite an *aggregate risk* through the kernel alone, and Chou et al. (2020) document the high variance of such unbiased-risk rewrites, which is one reason our primary value is doubly robust and not pure IPS. The object we need is different: a value for each candidate point, with a guarantee that survives a cooperative-game operator, which is where the additional questions of Lemma E.1 and Proposition E.4 arise. Here, the AIPW-style correction (kernel plus outcome model) is applied at the per-point value layer and then propagated through that operator.

Data valuation. Data valuation estimates the contribution of a training point to a utility. Data Shapley (Ghorbani & Zou, 2019; Jia et al., 2019b) and its closed-form KNN special case (Jia et al., 2019a) established the task-specific formulation; Beta Shapley (Kwon & Zou, 2022) and Data Banzhaf (Wang & Jia, 2023) vary the semivalue; influence functions (Koh & Liang, 2017) give a first-order alternative; and utility redesigns that preserve the KNN closed form are established practice (Wang et al., 2023a). Other families avoid the learned model altogether. LAVA and its scalable batched variant SAVA value points by the sensitivity of a class-conditional optimal-transport distance to a clean validation set (Just et al., 2023; Kessler et al., 2025); volume-based valuation scores feature-space diversity with no validation data at all (Xu et al., 2021); TracIn attributes validation-loss reduction along the training trajectory (Pruthi et al., 2020); and Data-IQ characterizes points by training-dynamics confidence (Seedat et al., 2022).

What these methods share is that the utility is computed from the observed data as given: none of them models why a point was or was not observed. We therefore place all four families on our MNAR benchmark in Section 6.4. In implicit-feedback recommendation, Kersbergen et al. (2025)

debug KNN recommenders with Monte-Carlo Data Shapley at scale, and Zhang et al. (2025) prune interactions by Shapley-estimated loss contributions. Both score against on-log, uncorrected utilities, which is the setting whose bias we characterize, and both serve as our naive baselines. Yu et al. (2020) use influence functions for MNAR-motivated reweighting at training time, whereas we produce post-hoc per-point values with guarantees. Three recent papers are adjacent but address other questions. Liu et al. (2026) use an AIPW-form error for sample-efficient Shapley estimation, which reduces variance under a fixed utility and is orthogonal to correcting a utility biased by MNAR exposure. Shen et al. (2026) show that KNN-Shapley values shift with validation-set perturbations and offer normalization heuristics, not a correction with guarantees. Naggita & LaChance (2025) document empirically that standard valuation metrics can undervalue underrepresented data; Theorem 5.3 gives a mechanism for that observation under MNAR exposure, together with a correction.

Debiased recommendation. This literature corrects the *training or evaluation* objective of a recommender for exposure. IPS-based debiasing of MNAR feedback began with Schnabel et al. (2016) and Yang et al. (2018); Saito et al. (2020) treat implicit feedback; Wang et al. (2019) introduced doubly-robust joint learning, with stability refinements under vanishing propensities (Guo et al., 2021; Li et al., 2023). We use the same estimators for a different target, the per-point quantity consumed by a *valuation* method, which calls for guarantees at the level of individual Shapley values and not of an aggregate risk. Closest in machinery, Kato et al. (2023) learn a classifier from positive, unlabeled, and exposure data with a doubly-robust objective; the output is a debiased classifier, not per-point values with Shapley semantics.

Feedback loops and popularity bias. Work on recommendation feedback loops studies amplification through the *model*: algorithmic confounding compounds over retraining (Chaney et al., 2018; Mansoury et al., 2020; Chen et al., 2023). We isolate a loop one level up, through the *valuation*. A platform acting on uncorrected values (the use case of Kersbergen et al., 2025; Zhang et al., 2025) induces exposure-multiplicative, boundary-concentrated expected dynamics even with a perfect relevance model, and corrected values remove the exposure-dependent term from those dynamics (Theorem E.6). Performative prediction (Perdomo et al., 2020) studies equilibria of repeated retraining; Theorem E.6 is instead a decoupling result, in which unbiasedness removes the performative term from the expected map. Fairness-of-exposure methods impose merit-proportional exposure as a constraint (Singh & Joachims, 2018; Biega et al., 2018); here the merit-proportional allocation is the unconstrained fixed point of the expected dynamics, a consequence of correcting the measurement.

D ALGORITHM

Algorithm 1 SHADE: doubly-robust KNN-Shapley

Input: log $\{(\mathbf{x}_i, \ell_i, e_i)\}_{i=1}^N$, queries $\{\mathbf{x}_q\}_{q=1}^Q$ (with relevance r_q if the curation transform is used), neighbors K, folds J, clip level ε
// Cross-fitted nuisances (one-off). Partition $\{1, \dots, N\}$ into folds $I_1, \dots, I_J$
for $j = 1$ **to** J **do**
  $\hat{\pi}_i \leftarrow \hat{\pi}^{(-j)}(\mathbf{x}_i)$, $i \in I_j$, from propensity fit on $\{(\mathbf{x}_k, e_k) : k \notin I_j\}$
  $\hat{m}_i \leftarrow \hat{m}^{(-j)}(\mathbf{x}_i)$, $i \in I_j$, from outcome fit on *exposed* $\{(\mathbf{x}_k, \ell_k) : k \notin I_j,\ e_k{=}1\}$, IPS-weighted by $1/\max(\hat{\pi}_k, \varepsilon)$
end for
$\mathbf{v}_i^{\mathrm{DR}} \leftarrow \hat{m}_i + e_i(\ell_i - \hat{m}_i)/\max(\hat{\pi}_i, \varepsilon)$ for all i (unexposed get $\hat{m}_i$)
(optional, curation) $v_i^q \leftarrow \mathbf{v}_i^{\mathrm{DR}} r_q + (1 - \mathbf{v}_i^{\mathrm{DR}})(1 - r_q)$ agreement transform, Appendix F.8
for each query q **do**
  sort points by $\|\mathbf{x}_i - \mathbf{x}_q\|$ ascending; run recursion Equation 1 on $(\mathbf{v}_{(1)}^{\mathrm{DR}}, \dots, \mathbf{v}_{(N)}^{\mathrm{DR}})$
end for
Return: $s_i = \frac{1}{Q}\sum_q s_i^q$

Algorithm 1 has three stages with different costs. The nuisances are fitted once per log: the propensity model on all points and the outcome model on exposed points only, both with J-fold cross-fitting, so that the values $\hat{\pi}_i$ and $\hat{m}_i$ used for point i come from models trained without its fold. The outcome fit is weighted by the inverse clipped propensities of its training points, which targets the population of candidates and not only the exposed subpopulation. These weights use propensities cross-fitted on the

whole sample, so the outcome model of a fold is not strictly independent of the exposures in that fold, which Assumption 1(v) idealizes. A point can influence its own $\hat{m}_i$ only through the propensity fits behind the weights, a dependence that shrinks as N grows. It does not affect Proposition 5.1, whose conclusion holds for every fit when either nuisance is correct, and nested cross-fitting removes it. The value $\mathbf{v}_i^{\mathrm{DR}}$ is then computed once per point and does not depend on the query. A never-exposed point has $e_i = 0$, so the residual term vanishes and $\mathbf{v}_i^{\mathrm{DR}} = \hat{m}_i$: the outcome model is the only available information about its relevance, and this is what lets the valuation cover U (Section 4). Only the last stage is per query: one sort and one pass of the recursion, $O(N \log N)$. The optional agreement transform is used only in the curation experiments of Appendix F.8.

E PROOFS

Throughout, (i) in parentheses denotes sort rank with respect to the query (rank 1 = nearest, ties broken by a fixed deterministic rule, so the sort order is a deterministic function of $\{\mathbf{x}_i\}, \mathbf{x}_q$), and $a_j := \min(K, j)/(Kj) = 1/\max(K, j)$. Numerical checks of the closed forms are collected in Appendix E.9; they support the proofs and are not part of them.

E.1 PROOF OF LEMMA E.1 (OPERATOR FORM AND CONTRACTION)

Lemma E.1 (Linear-operator form; contraction of the recursion)**.** *Fix $N \geq K$, a query q, and the induced sort order; let $a_j := 1/\max(K, j)$. The map $v \mapsto s$ of Equation 1 is the fixed linear map $\mathbf{s} = A_q \mathbf{v}$ with, in sort order,*

$$(A_q)_{ij} = \begin{cases} a_i, & j = i, \\ a_j - a_{j-1}\ (\leq 0), & j > i, \\ 0, & j < i, \end{cases} \qquad \sum_{j=1}^{N} \left|(A_q)_{ij}\right| = \tfrac{2}{\max(K,i)} - \tfrac{1}{N}, \qquad \sum_{i=1}^{N} (A_q)_{ij} = \tfrac{\mathbb{1}[j \leq K]}{K}.$$

Consequently, writing $\hat{\mathbf{s}} = A_q \hat{\mathbf{v}}$ and $\mathbf{s}^\star = A_q \mathbf{v}^\star$ for any two value vectors $\hat{\mathbf{v}}, \mathbf{v}^\star \in \mathbb{R}^N$,

$$\|A_q\|_\infty = \frac{2}{K} - \frac{1}{N}, \qquad \|\hat{\mathbf{s}} - \mathbf{s}^\star\|_\infty \ \leq\ \left(\frac{2}{K} - \frac{1}{N}\right) \|\hat{\mathbf{v}} - \mathbf{v}^\star\|_\infty, \qquad \sum_{i=1}^{N} s_i = u_q(D).$$

In particular, A_q is a contraction in ℓ_∞ if and only if $K \geq 2$.

Proof. Throughout, $a_j = 1/\max(K, j)$, base case $s_{(N)} = a_N v_{(N)}$ (covers both regimes; Appendix E.2).

Operator form (backward induction; base $s_{(N)} = a_N v_{(N)}$ is row N):

$$s_{(i)} = s_{(i+1)} + a_i(v_{(i)} - v_{(i+1)}) = a_i v_{(i)} + (a_{i+1} - a_i) v_{(i+1)} + \sum_{j=i+2}^{N} (a_j - a_{j-1}) v_{(j)},$$

which is row i of A_q; $a_j - a_{j-1} = 0$ for $j \leq K$ and $= -1/(j(j-1)) < 0$ for $j > K$.

Row mass (off-diagonal entries share one sign; telescoping):

$$\sum_j |A_{ij}| = a_i + \sum_{j>i} (a_{j-1} - a_j) = 2a_i - a_N, \qquad \|A_q\|_\infty = \max_i\, (2a_i - a_N) = \frac{2}{K} - \frac{1}{\max(K, N)},$$

attained at every row $i \leq K$; per row the sharper bound is $2/\max(K, i) - a_N$.

Efficiency (column sums):

$$\sum_i (A_q)_{ij} = a_j + (j-1)(a_j - a_{j-1}) = \begin{cases} 1/K, & j \leq K, \\ \frac{1}{j} + \frac{j-1}{j} - 1 = 0, & j > K, \end{cases} \qquad \sum_i s_i = \frac{1}{K} \sum_{j \leq K} v_{(j)} = u_q(D).$$

Contraction. $K \geq 2 \Rightarrow \|A_q\|_\infty = 2/K - a_N \leq 1 - a_N < 1$. For $K = 2$ the constant is $1 - 1/N$, which is below one for every N but not bounded away from one as N grows; for $K \geq 3$ it is at most $2/3$. For $K = 1$, $\|A_q\|_\infty = 2 - 1/N \geq 1$, so the map is bounded, uniformly in N, but not a contraction. □

E.2 Base-case correction for $N < K$

The recursion as usually stated ($s_{(N)} = v_{(N)}/N$) is exact only for $N \geq K$. For $N < K$ the exact Shapley value of $u_q(S) = \frac{1}{K}\sum_{i\in \mathrm{KNN}_K(S,q)} v_i$ (with $|\mathrm{KNN}_K(S,q)| = \min(K,|S|)$) requires $s_{(N)} = \min(K,N)/(KN)\cdot v_{(N)} = v_{(N)}/K$. The difference is not cosmetic: in brute-force enumeration the uncorrected base case is off by up to 0.88 on random $\mathbf{v}$ when $N < K$, and the corrected one is exact. All statements use $a_N = 1/\max(K,N)$, covering both regimes; the main text assumes $N \geq K$, always true in practice.

E.3 Proof of Lemma E.2 (exactness for arbitrary fixed scalars)

Jia et al. (2019a) prove the closed form for the 0/1 match indicator $v_i = \mathbb{1}[y_i = y_q]$, unweighted-KNN regression, and a piecewise-utility-difference class; the arbitrary-fixed-scalar statement our debiased values require does not appear there. The proof adapts the structure of their argument.

Lemma E.2 (Exactness for generic per-point values)**.** *Fix $N \geq K$. For any fixed scalars $\{v_i\}_{i=1}^N \subset \mathbb{R}$, recursion Equation 1 returns the exact Shapley values of the game $u_q(S) = \frac{1}{K}\sum_{i\in \mathrm{KNN}_K(S,q)} v_i$ (for $N < K$, exactness holds with base case $s_{(N)} = v_{(N)}/K$; Appendix E.2).*

Proof. Work in sort order with the permutation form $s_i = \mathbb{E}_\sigma\big[u(P_\sigma^i \cup \{i\}) - u(P_\sigma^i)\big]$ (σ uniform, P_σ^i the predecessors of i in σ).

Marginal structure. For $S \not\ni (i)$, let $c_i(S) := |\{p \in S : p \text{ strictly closer to } q\}|$ and $w_K(S) :=$ value of the K-th nearest point of S if $|S| \geq K$, else 0 (note $|S| < K \Rightarrow c_i(S) < K$). Then

$$u(S\cup\{(i)\}) - u(S) = \frac{\mathbb{1}[c_i(S) < K]}{K}\big(v_{(i)} - w_K(S)\big),$$

with no constraint on $\mathbf{v}$ beyond coalition-independence.

Base case. $c_N(S) = |S|$ and $|P_\sigma^{(N)}| \sim \mathrm{Unif}\{0,\dots,N-1\}$, so

$$s_{(N)} = \frac{v_{(N)}}{K}\,\mathbb{P}\big(|P_\sigma^{(N)}| < K\big) = \frac{v_{(N)}}{K}\cdot\frac{\min(K,N)}{N} = a_N v_{(N)}.$$

Adjacent-rank difference. Fix $i < N$; couple by the measure-preserving transposition τ swapping $(i) \leftrightarrow (i{+}1)$, so $s_{(i)} - s_{(i+1)} = \mathbb{E}_\sigma[m_{(i)}(\sigma) - m_{(i+1)}(\tau\sigma)]$ (m_p = marginal of p). With $P_0 := P_\sigma^{(i)} \setminus \{(i{+}1)\}$,

$$m_{(i)}(\sigma) - m_{(i+1)}(\tau\sigma) = u(P_0\cup\{(i)\}) - u(P_0\cup\{(i{+}1)\}) = \frac{\mathbb{1}[c_i(P_0) < K]}{K}\big(v_{(i)} - v_{(i+1)}\big):$$

if $(i{+}1) \notin P_\sigma^{(i)}$ both marginals subtract the common $u(P_0)$; if $(i{+}1) \in P_\sigma^{(i)}$ the leading terms $u(P_0\cup\{(i),(i{+}1)\})$ cancel; in either case sort-adjacency gives $c_i(P_0) = c_{i+1}(P_0)$ and an identical evicted value. The position of (i) among the i smallest ranks is uniform, so

$$\mathbb{P}\big(c_i(P_0) < K\big) = \frac{\min(K,i)}{i}, \qquad s_{(i)} - s_{(i+1)} = \frac{v_{(i)} - v_{(i+1)}}{K}\cdot\frac{\min(K,i)}{i},$$

which is Equation 1; backward induction from the base case completes the proof, and Appendix E.1 gives $\mathbf{s} = A_q\mathbf{v}$. □

The proof nowhere uses $v_i \in \{0,1\}$, nonnegativity, or boundedness, only that v_i is a fixed scalar independent of coalition membership. Negative values raise no difficulty, since the marginal-contribution formula holds for signed v. $\mathbf{v}^{\mathrm{naive}}, \mathbf{v}^{\mathrm{IPS}}, \mathbf{v}^{\mathrm{DR}}$ all qualify: each is computed once from $(\mathbf{x}_i, \ell_i, e_i)$ and the cross-fitted nuisances (Assumption 1(v)), never from coalitions.

E.4 Proof of Proposition 5.1 (unbiasedness and double robustness)

Proposition 5.1 (Unbiasedness and double robustness)**.** *Under Assumption 1(i,ii,iv,v), conditionally on the out-of-fold nuisance fit, the pointwise value bias obeys the exact AIPW identity*

$$b_i := \mathbb{E}[\mathbf{v}_i^{\mathrm{DR}} \mid \mathbf{x}_i] - r_i = \big(\hat{m}(\mathbf{x}_i) - r_i\big)\,\frac{\hat{\pi}_i - \pi_i}{\hat{\pi}_i}, \tag{5}$$

so $\mathbb{E}[\mathbf{s}_i^{\mathrm{DR}} \mid \{\mathbf{x}\}, \mathbf{x}_q] = \mathbf{s}_i^\star$ *if* $\hat{\pi} = \pi$ or $\hat{m} = r$. *In particular, with* $\hat{\pi} = \pi$ *(clipping not binding),* $\mathbb{E}[\mathbf{s}_i^{\mathrm{IPS}} \mid \{\mathbf{x}\}, \mathbf{x}_q] = \mathbf{s}_i^\star$ *for all* i, q*: IPS is the* $\hat{m} \equiv 0$ *special case.*

Proof. Let $\mathcal{F}_{-i}$ be the out-of-fold sample on which the nuisance values of point i are fitted. By Assumption 1(iv,v) it is independent of (e_i, ℓ_i) given $\{\mathbf{x}\}$, so conditionally on $(\{\mathbf{x}\}, \mathcal{F}_{-i})$ the numbers $\hat{m}(\mathbf{x}_i)$ and $\hat{\pi}_i$ are fixed, while $\mathbb{E}[e_i \mid \cdot] = \pi_i$ and, since $\ell_i = e_i y_i$ gives $e_i \ell_i = \ell_i$, $\mathbb{E}[e_i \ell_i \mid \cdot] = \pi_i r_i$ (Assumption 1(ii)). The division is well defined because the propensity entering the estimator is clipped, $\hat{\pi}_i \geq \varepsilon$. Writing $\mathbb{E}[\,\cdot \mid \mathbf{x}_i]$ for this conditional expectation,

$$b_i := \mathbb{E}[\mathbf{v}_i^{\mathrm{DR}} \mid \mathbf{x}_i] - r_i = \hat{m}(\mathbf{x}_i) + \frac{\pi_i r_i - \pi_i \hat{m}(\mathbf{x}_i)}{\hat{\pi}_i} - r_i = \big(\hat{m}(\mathbf{x}_i) - r_i\big)\frac{\hat{\pi}_i - \pi_i}{\hat{\pi}_i},$$

which is Equation 5: a product of the two nuisance errors, with $b_i = 0$ for every realization of the fit if $\hat{\pi}_i = \pi_i$ or $\hat{m}(\mathbf{x}_i) = r_i$. Under either condition the tower property gives $\mathbb{E}[\mathbf{v}_i^{\mathrm{DR}} \mid \{\mathbf{x}\}] = r_i$ for each i, and determinism of A_q (below) gives

$$\mathbb{E}[\mathbf{s}^{\mathrm{DR}} \mid \{\mathbf{x}\}, \mathbf{x}_q] = A_q\, \mathbb{E}[\mathbf{v}^{\mathrm{DR}} \mid \{\mathbf{x}\}] = A_q \mathbf{r} = \mathbf{s}^\star.$$

The IPS statement is the instance $\hat{m} \equiv 0$: $\mathbf{v}^{\mathrm{DR}}\big|_{\hat{m}\equiv 0} = \mathbf{v}^{\mathrm{IPS}}$ and Equation 5 reduces to $b_i = -r_i(\hat{\pi}_i - \pi_i)/\hat{\pi}_i$, zero at $\hat{\pi} = \pi$. The expectation of the vector $\mathbf{v}^{\mathrm{DR}}$ is taken coordinate by coordinate: different points use different fits, and we never condition on all fits jointly, since a model fitted on a fold that contains point i is not independent of (e_i, ℓ_i). □

Why linearity is legitimate (and where it would break). If the sort order or the player set depended on the realized (e, ℓ), A_q would be random and correlated with $\mathbf{v}$, and $\mathbb{E}[A_q \mathbf{v}] \neq A_q \mathbb{E}[\mathbf{v}]$. Both are excluded by construction: the player set is all N points, and ranks are functions of $\{\mathbf{x}_i\}, \mathbf{x}_q$ with deterministic tie-breaking. With an exposed-only player set, the set, hence A_q, becomes a random function of e, so this proof no longer applies; that alone does not show that an exposed-only estimator is biased, and the conditional bias it does carry is computed in Section 4. This is why the all-N decision is part of the result. With true propensities clipped at ε, the bias of $\mathbf{v}_i^{\mathrm{IPS}}$ is $r_i(\pi_i/\max(\pi_i, \varepsilon) - 1)$, nonzero only where $\pi_i < \varepsilon$, and it propagates through A_q with the constants of Lemma E.1.

Identity Equation 5 underlies the theory section: Proposition 5.1 reads off its zeros, Corollary E.3 its magnitude. It needs no positivity of the true π_i for the algebra, but if $\pi_i = 0$ (structurally unexposable point) correctness requires $\hat{m}(\mathbf{x}_i) = r_i$ at that point, since the $\hat{\pi}$ route is unavailable; the $\mathbf{v}^{\mathrm{DR}}$ estimator reduces to the direct method there, the standard causal-support limitation, which we make explicit.

E.5 Proof of Corollary E.3

The result bounds the Shapley-value bias when both nuisances are wrong: it is at most the product of the two nuisance errors, scaled by the operator norm. It is the classical second-order (product-of-errors) property of AIPW estimators (Robins et al., 1994; Chernozhukov et al., 2018) composed with Lemma E.1, and we accordingly state it as a corollary, not a theorem.

Corollary E.3 (Misspecification sensitivity)**.** *Under Assumption 1 with clipping at* ε*, for every* q *and every* i *at rank* ρ *in the induced sort order,*

$$\big|\mathbb{E}[\mathbf{s}_i^{\mathrm{DR}}] - \mathbf{s}_i^\star\big| \;\leq\; \Big(\tfrac{2}{\max(K,\rho)} - \tfrac{1}{N}\Big) \max_j |b_j| \;\leq\; \Big(\tfrac{2}{K} - \tfrac{1}{N}\Big)\frac{\Delta_m \Delta_\pi}{\varepsilon},$$

where b_j *is the product-form bias Equation 5,* $\Delta_m := \|\hat{m} - r\|_\infty$, $\Delta_\pi := \|\hat{\pi} - \pi\|_\infty$*, and* $\hat{\pi}$ *denotes the propensity that enters the estimator, that is, the estimate after clipping at* ε *(Assumption 1(i)). For random fits,* Δ_m *and* Δ_π *are almost-sure bounds.*

Proof. By Lemmas E.1 and E.2 and determinism of A_q,

$$\mathbb{E}[\mathbf{s}^{\mathrm{DR}}] - \mathbf{s}^\star = A_q\big(\mathbb{E}[\mathbf{v}^{\mathrm{DR}}] - \mathbf{r}\big) = A_q \mathbb{E}[\mathbf{b}], |(A_q \mathbf{b})_i| \leq \big(2a_{\mathrm{rank}_q(i)} - a_N\big) \max_j |b_j| \leq \Big(\frac{2}{K} - \frac{1}{N}\Big) \max_j |b_j|,$$

where $|\mathbb{E}[b_j]| \leq \sup |b_j|$ and, by Equation 5 with $\hat{\pi}_j \geq \varepsilon$, $|b_j| \leq \Delta_m \Delta_\pi / \varepsilon$. The clipping convention matters. If the raw estimate equals a true propensity $\pi_j < \varepsilon$, the used propensity is ε and $b_j =$

$(\hat{m}(\mathbf{x}_j) - r_j)(\varepsilon - \pi_j)/\varepsilon$, which the bound covers because $\Delta_\pi \geq |\varepsilon - \pi_j|$ by definition; with Δ_π computed from the raw estimate, the bound would fail at such points. □

With cross-fitted nuisances and $\mathbf{x}_j$ as draws from a covariate distribution, Cauchy–Schwarz on Equation 5 gives the population version $\mathbb{E}_x|b| \leq \|\hat{m} - r\|_{L_2}\|(\hat{\pi} - \pi)/\hat{\pi}\|_{L_2}$, the familiar product-of-L_2-rates condition (both nuisances $o(n^{-1/4}) \Rightarrow$ bias $o(n^{-1/2})$) (Chernozhukov et al., 2018). We state the sup-norm version in the main text because the valuation target is per-point.

E.6 Proof of Corollary 5.2 (linear solution concepts)

Corollary 5.2 (Linear solution concepts inherit the correction)**.** *Let $\tilde{u}_q$ be any coalition utility linear in $\mathbf{v}$ with coefficients determined by $\{\mathbf{x}\}, \mathbf{x}_q$ alone (the KNN utility above, its threshold variant), and let ϕ be any solution concept linear in the game (any semivalue: Shapley, Banzhaf, Beta-Shapley). Then $\mathbf{s}^\phi = A_q^\phi \mathbf{v}$ for a deterministic matrix A_q^ϕ, and Proposition 5.1 and Equation 6 hold verbatim with $\|A_q^\phi\|_\infty$ in place of $2/K - 1/N$.*

Proof. Linearity of $\tilde{u}_q$ in $\mathbf{v}$: $\tilde{u}_q(S) = \sum_j c_{q,j}(S)\, v_j$ with $c_{q,j}(S)$ functions of $\{\mathbf{x}\}, \mathbf{x}_q$ alone (KNN: $c_{q,j}(S) = \frac{1}{K}\mathbb{1}[j \in \mathrm{KNN}_K(S, q)]$). Composing with a semivalue $\phi_i(u) = \sum_{S \subseteq D\setminus\{i\}} w_{|S|}\big(u(S \cup \{i\}) - u(S)\big)$ (fixed size weights $w_{|S|}$):

$$s_i^\phi = \sum_j (A_q^\phi)_{ij}\, v_j, \qquad (A_q^\phi)_{ij} = \sum_{S \subseteq D\setminus\{i\}} w_{|S|}\big(c_{q,j}(S \cup \{i\}) - c_{q,j}(S)\big),$$

with A_q^ϕ a function of $\{\mathbf{x}\}, \mathbf{x}_q$ (and the weights) only, hence deterministic under Assumption 1(v). Conditioning as in Appendix E.4, with $\mathbf{b}$ the AIPW bias Equation 5:

$$\mathbb{E}[\mathbf{s}^{\phi,\mathrm{DR}} \mid \{\mathbf{x}\}, \mathbf{x}_q] = A_q^\phi(\mathbf{r} + \mathbf{b}), \qquad \big|\mathbb{E}[s_i^{\phi,\mathrm{DR}}] - s_i^{\phi,\star}\big| \leq \|A_q^\phi\|_\infty \max_j |b_j| \leq \|A_q^\phi\|_\infty \frac{\Delta_m \Delta_\pi}{\varepsilon},$$

i.e. Proposition 5.1 for ϕ ($\mathbf{b} = \mathbf{0}$ under either correct nuisance) and Equation 6 with the operator norm of the concept itself. Substituting $\mathbf{v} = \mathbf{v}^{\mathrm{naive}}$,

$$\mathbb{E}[\mathbf{s}^{\phi,\mathrm{naive}}] - \mathbf{s}^{\phi,\star} = -A_q^\phi\big((\mathbf{1} - \pi) \odot \mathbf{r}\big),$$

the value-level statement of Theorem 5.3; its rank-ρ decomposition and efficiency ($\sum_i s_i = u_q(D)$) are Shapley-specific and are not claimed. □

What is generic here is the passage of the value-level identity through a deterministic linear operator. The constant $2/K - 1/N$, the rank-ρ decomposition of Theorem 5.3 and efficiency belong to KNN-Shapley and are not inherited by other concepts.

E.7 Statement and proof of the concentration result

The following is the full statement summarized in Section 5.

Proposition E.4 (Concentration and consistency)**.** *Under Assumption 1, with the nuisance functions held fixed (as under sample splitting with a separate nuisance sample), let $\bar{s}_i = \frac{1}{Q}\sum_q s_i^{(q)}$, let $\bar{s}_i^\star := \mathbb{E}_q[s_i^{\mathrm{oracle},(q)}]$ be the population oracle target, and let $\chi^* := \max_j \frac{1}{Q}\sum_q K/\max(K, \mathrm{rank}_q(j)) \in (0, 1]$ measure query concentration. The error decomposes into label noise, bias, and query sampling,*

$$\bar{s}_i - \bar{s}_i^\star = \underbrace{\bar{s}_i - \mathbb{E}_v[\bar{s}_i \mid q_{1:Q}]}_{T_1} + \underbrace{\mathbb{E}_v[\bar{s}_i \mid q_{1:Q}] - \tfrac{1}{Q}\textstyle\sum_q s_i^{\mathrm{oracle},(q)}}_{T_2} + \underbrace{\tfrac{1}{Q}\textstyle\sum_q s_i^{\mathrm{oracle},(q)} - \bar{s}_i^\star}_{T_3},$$

with each term bounded with probability $\geq 1 - \delta$: $|T_1| \leq \frac{2B}{K}\sqrt{\chi^* \ln(2/\delta)}$ (label/exposure noise is shared across queries *and shrinks only under query diversity, $\chi^* \to 0$, not in Q); $|T_2| \leq (2/K - 1/N)\max_j |b_j|$, zero under Proposition 5.1; $|T_3| \leq (2/K - 1/N)\|\mathbf{r}\|_\infty\sqrt{2\ln(2/\delta)/Q}$. With consistent cross-fitted nuisances, $Q \to \infty$, and $\chi^* \to 0$, $\bar{s}_i \to \bar{s}_i^\star$ in probability.*

Proof. The draws (e, ℓ) are shared by all queries (every query reads the same $\mathbf{v}$), so Hoeffding over queries alone is invalid for the label noise; the two randomness sources are treated separately. With $\tilde{A}_{ij} := \frac{1}{Q}\sum_q |A^{(q)}_{\mathrm{rank}_q(i),\mathrm{rank}_q(j)}|$ and every entry of column c of A_q bounded by $1/\max(K, c)$,

$$\tilde{A}_{ij} \le \frac{\chi_j}{K}, \qquad \sum_j \tilde{A}_{ij} \le \frac{2}{K} - \frac{1}{N}, \qquad \kappa_i := \sum_j \tilde{A}^2_{ij} \le \frac{2\chi^*}{K^2}.$$

Each with probability $\ge 1 - \delta$:

$$|T_1| \le B\sqrt{2\kappa_i \ln(2/\delta)} \le \frac{2B}{K}\sqrt{\chi^* \ln(2/\delta)},$$

by Hoeffding over the N independent draws (Assumption 1(iv));

$$|T_2| = \left|\frac{1}{Q}\sum_q (A^{(q)}\mathbf{b})_i\right| \le \left(\frac{2}{K} - \frac{1}{N}\right)\max_j |b_j|,$$

deterministic given the nuisances, and zero under Proposition 5.1;

$$|T_3| \le \left(\frac{2}{K} - \frac{1}{N}\right)\|\mathbf{r}\|_\infty \sqrt{\frac{2\ln(2/\delta)}{Q}},$$

by Hoeffding over the Q i.i.d. queries; values bounded by $\|A_q\|_\infty \|\mathbf{r}\|_\infty$).

Union bound over the three terms; consistency follows from $\max_j |b_j| \to_p 0$ (either nuisance consistent, Equation 5), $Q \to \infty$, $\chi^* \to 0$. T_1 does not shrink in Q, only in query diversity. Since $\chi^* \ge K/N$ always, $\chi^* \to 0$ also requires $N/K \to \infty$. As an illustration, not a result we prove: if the rank of each point is uniformly distributed across queries, then $\chi_j = O((K/N)(1 + \ln(N/K)))$ and $T_1 = O(B\sqrt{\ln(N/K)/(KN)})$. For a fixed finite query set the target is $\frac{1}{Q}\sum_q s_i^{\mathrm{oracle},(q)}$ itself and T_3 is absent. □

Two caveats on T_1. The Hoeffding step treats the values v_j as independent, which is exact for fixed nuisances; under J-fold cross-fitting the v_j share fold-level dependence through the fitted nuisances, which this bound does not track. Second, concentration statements in the KNN-Shapley literature implicitly average over test points as if per-query values were independent; they are not when training labels are shared. The χ^* factor is the cost of stating this correctly: label noise averages out only to the extent that different queries consult different neighbors. T_1 is also the variance channel through which "unbiased $\neq$ best ranking" operates (Remark F.1): the $\mathbf{v}^{\mathrm{DR}}$ correction raises the value bound from $B = 1$ for naive values to as much as $B = 1/\varepsilon$ (since $\hat{m} \in [0, 1]$ and $\hat{\pi} \ge \varepsilon$ give $|\mathbf{v}^{\mathrm{DR}}_i| \le 1/\varepsilon$) in exchange for $T_2 = 0$, and whether the trade improves rank correlation depends on κ_i.

E.8 Proof of Theorem 5.3

Theorem 5.3 (Long-tail bias of naive valuation)**.** *Under Assumption 1(i,ii,v), $\mathbb{E}[\mathbf{v}^{\mathrm{naive}}_i \mid \mathbf{x}_i] = \pi_i r_i$ exactly. For a query q with point i at rank ρ ($a_\rho = 1/\max(K, \rho)$), the Shapley bias decomposes as*

$$\mathbb{E}[\mathbf{s}^{\mathrm{naive}}_i - \mathbf{s}^\star_i] = -a_\rho(1 - \pi_i)r_i \;+\; C_\rho, \quad C_\rho = \sum_{j > \max(\rho, K)} \frac{(1 - \pi_{(j)})r_{(j)}}{j(j-1)} \;\ge 0.$$

(a) $\partial_{\pi_i}\mathbb{E}[\mathbf{s}^{\mathrm{naive}}_i - \mathbf{s}^\star_i] = a_\rho r_i > 0$; for two points h, t with $r_h = r_t = r$ at adjacent ranks $\rho, \rho{+}1$, $\mathbb{E}\big[(\mathbf{s}^{\mathrm{naive}}_h - \mathbf{s}^{\mathrm{naive}}_t) - (\mathbf{s}^\star_h - \mathbf{s}^\star_t)\big] = a_\rho\, r\,(\pi_h - \pi_t)$, free of cross-terms. (b) $\mathbb{E}[\mathbf{s}^{\mathrm{naive}}_i - \mathbf{s}^\star_i] \gtrless 0 \iff (1 - \pi_i)\, r_i \lessgtr C_\rho/a_\rho$. (c) Under the conditions of Proposition 5.1, $\mathbb{E}[\mathbf{s}^{\mathrm{DR}}_i - \mathbf{s}^\star_i] = 0$ for all π; in general it obeys the Corollary E.3 bound, which is free of π_i.

Proof. Assumption 1(ii) gives $\mathbb{E}[\mathbf{v}^{\mathrm{naive}}_i \mid \mathbf{x}_i] = \pi_i r_i$ exactly, so $b^{\mathrm{naive}}_i = -(1 - \pi_i)r_i \le 0$.

Decomposition. $\mathbb{E}[\mathbf{s}^{\mathrm{naive}}] - \mathbf{s}^\star = A_q \mathbf{b}^{\mathrm{naive}}$; row ρ of Lemma E.1 has diagonal a_ρ and off-diagonal entries $-1/(j(j-1))$ for $j > \max(\rho, K)$ (else 0), so

$$\mathbb{E}[\mathbf{s}^{\mathrm{naive}}_i - \mathbf{s}^\star_i] = -a_\rho(1 - \pi_i)r_i + C_\rho, \qquad C_\rho = \sum_{j > \max(\rho, K)} \frac{(1 - \pi_{(j)})r_{(j)}}{j(j-1)} \in \big[0, (a_\rho - 1/N)\beta_\infty\big],$$

with $\beta_\infty := \max_j (1-\pi_j) r_j \le 1$ and $\sum_{j>\max(\rho,K)} \frac{1}{j(j-1)} = a_\rho - \frac{1}{N}$ (telescoping); each negative entry multiplies a nonpositive bias, hence $C_\rho \ge 0$.

(a). The decomposition is affine in π_i with slope $a_\rho r_i > 0$ (C_ρ excludes rank ρ). For equal relevance r at adjacent ranks $\rho, \rho{+}1$ (for non-adjacent pairs the intermediate points contribute cross-terms), the recursion step $s_{(\rho)} - s_{(\rho+1)} = a_\rho (v_{(\rho)} - v_{(\rho+1)})$ is cross-term-free, so exactly

$$\mathbb{E}[\mathbf{s}_h^{\text{naive}} - \mathbf{s}_t^{\text{naive}}] = a_\rho \, r \, (\pi_{\text{head}} - \pi_{\text{tail}}), \qquad \mathbf{s}_h^\star - \mathbf{s}_t^\star = 0.$$

(b). C_ρ is π_i-free, so $\mathbb{E}[\mathbf{s}_i^{\text{naive}} - \mathbf{s}_i^\star] \gtrless 0 \iff (1-\pi_i) r_i \lessgtr C_\rho / a_\rho$. Efficiency (Lemma E.1) forces

$$\sum_i \mathbb{E}[\mathbf{s}_i^{\text{naive}} - \mathbf{s}_i^\star] = \frac{1}{K} \sum_{j \le K} b_{(j)} \le 0 :$$

The total is understated even as head entries individually exceed their oracle values. A tail point ($\pi_i \le \vartheta$) is net-undervalued whenever $(1-\vartheta) r_i > (1 - 1/(a_\rho N)) \beta_\infty$.

(c). Query aggregation replaces a_ρ by $\bar{a}_i = \chi_i / K$. Under Proposition 5.1, both terms vanish for every π. When both nuisances are wrong, the exact pointwise bias is the product form of Equation 5, which in general still depends on π_i through $\hat{\pi}_i - \pi_i$. What Corollary E.3 provides is an envelope that does not involve π_i. The long-tail structure of the bias is therefore removed when either nuisance is correct and bounded, and not removed otherwise. □

A caveat on the cross-term. The informal claim "naive undervalues tail items" is true for raw values and true relatively (statement (a), cross-term-free), but the absolute per-point Shapley bias can have either sign, and at the head it is positive: the empirical $+0.51$ head-decile bias (Figure 2) is predicted, not anomalous. A version of the theorem claiming uniformly negative bias would be falsified by our own experiment.

E.9 Numerical verification of the statistical claims

The checks below are diagnostics that accompany the proofs; none of them is used as an argument. *Operator and exactness.* A_q reconstructed column-by-column by feeding unit vectors through the recursion matches the closed form to $3{\times}10^{-17}$ over $N \in \{10, 10^2, 2{\times}10^3\} \times K \in \{1, 5, 20\}$; the norm and the attaining rows match exactly. The recursion agrees with brute-force Shapley over all 2^N coalitions with arbitrary signed Gaussian $\mathbf{v}$, $(N, K) \in \{(8,1), \ldots, (8,8), (10,4)\}$, with max discrepancy $4{\times}10^{-15}$. *Linear solution concepts.* Brute-force Shapley, Banzhaf, and Beta$(4,1)$ semivalues of the KNN game at $N = 10$ are linear in $\mathbf{v}$ to machine precision; the explicitly constructed A_q^ϕ reproduces each concept, plugging the closed-form $\mathbb{E}[\mathbf{v}^{\text{DR}}]$ with one correct nuisance recovers $A_q^\phi \mathbf{r}$ exactly, and the two-wrong-nuisance bias respects the $\|A_q^\phi\|_\infty \Delta_m \Delta_\pi / \varepsilon$ bound. *Ranking and feedback.* In the expected value-channel map of Theorem E.6, naive converges to the bang-bang fixed point and SHADE reaches $\text{clipnorm}(r)$ in one step; the two ranking conditions of Corollary E.5 produce zero misrankings and zero false positives over random instances.

Bias statements. These are checked by Monte Carlo ($N = 200$ fixed points in $\mathbb{R}^2$, smooth $r(\mathbf{x}) \in [0.05, 0.95]$, MNAR propensities correlated with r; $Q = 30$ queries, $K = 5$; aggregated operator from the closed form, $\|\bar{A}\|_\infty = 0.086$ vs. the per-query bound 0.395; $R = 2000$ replications of (e, ℓ), MC standard error $\approx 9{\times}10^{-4}$; the complete set of configurations is in the code release). Max |empirical s-bias| vs. predicted: naive $6.2{\times}10^{-3}$ vs. $6.1{\times}10^{-3}$ (Theorem 5.3 matched to $6{\times}10^{-4}$); IPS with true π, SHADE with both nuisances true, and SHADE with *either* nuisance deliberately wrong all land at 1.1–$1.9{\times}10^{-3}$ against a prediction of 0 (consistent with zero at this MC error; both fallback directions of Proposition 5.1 realized); SHADE with *both* wrong $3.6{\times}10^{-3}$ vs. $3.5{\times}10^{-3}$ (matches Equation 5 to 10^{-3}); DM with the same wrong $\hat{m}$ $9.9{\times}10^{-3}$, first-order in Δ_m, exact.

Note the ordering DM > naive > both-wrong DR (with $\Delta_m = 0.5$, $\Delta_\pi = 0.15$): the quadratic product-of-errors structure is visible in the numbers. Theorem 5.3 checks: the own/cross decomposition is exact to machine precision; own term ≤ 0 and cross-term ≥ 0 for all 200 points; naive s-bias increasing in π_i, SHADE bias flat in π ($\le 5{\times}10^{-5}$ per quintile); under strong MNAR ($\pi_{\text{head}} = 0.97$ vs. $\pi_{\text{tail}} = 0.08$) the exact operator calculation gives positive total naive bias for 59/59 head points and negative for 140/141 tail points.

E.10 Proofs for Corollary E.5 and Theorem E.6

Part (a) of the corollary says that corrected valuation orders, in expectation, any pair the oracle separates by more than twice the bias envelope. It is a triangle inequality on the Corollary E.3 envelope, stated only to read that envelope as an ordering guarantee; it is not a statement about the realized ordering of a single draw. Part (b), the checkable condition for naive valuation, is the substantive part: for two points at adjacent ranks, whenever the exposure advantage of the head point exceeds the relevance advantage of the tail point, naive valuation inverts the pair in expectation, and the condition depends on (π, r) alone.

Corollary E.5 (Ranking: SHADE consistency, checkable naive misranking)**.** ***(a)*** *Let* $\varepsilon_{\mathrm{b}} := \|A_q\|_\infty \max_j |b_j| \le \|A_q\|_\infty \Delta_m \Delta_\pi / \varepsilon$ *be the sensitivity envelope of Corollary E.3 (*b_j *the pointwise* $\mathbf{v}^{\mathrm{DR}}$ *bias,* $\|A_q\|_\infty = 2/K - 1/\max(K, N)$ *by Lemma E.1). If two points satisfy* $\mathbf{s}^\star_i - \mathbf{s}^\star_j > 2\varepsilon_{\mathrm{b}}$*, then* $\mathbb{E}[\mathbf{s}^{\mathrm{DR}}_i \mid \{\mathbf{x}\}, \mathbf{x}_q] > \mathbb{E}[\mathbf{s}^{\mathrm{DR}}_j \mid \{\mathbf{x}\}, \mathbf{x}_q]$*, for any bias pattern inside the Corollary E.3 envelope (with* $\varepsilon_{\mathrm{b}} = 0$*, hence for every oracle-separated pair, under either nuisance correct). Aggregated over a fixed set of* Q *queries, with* $\mathbf{s}^\star$ *replaced by the query-averaged oracle and under the fixed-nuisance condition of Proposition E.4, the realized averages satisfy* $\bar{s}_i > \bar{s}_j$ *with probability at least* $1 - \delta$ *once the separation exceeds twice the high-probability envelope* η_δ *of the proof.* ***(b)*** *Fix a query and let a head point* h *at rank* ρ *and a tail point* t *at the adjacent rank* $\rho + 1$ *satisfy* $r_t > r_h$ *(oracle ranks* t *above* h*) and* $\pi_h r_h > \pi_t r_t$*. Then, exactly,*

$$\mathbf{s}^\star_h - \mathbf{s}^\star_t = a_\rho(r_h - r_t) < 0, \qquad \mathbb{E}[\mathbf{s}^{\mathrm{naive}}_h - \mathbf{s}^{\mathrm{naive}}_t] = a_\rho(\pi_h r_h - \pi_t r_t) > 0,$$

so naive valuation inverts the pair in expectation; since $r_t > r_h$*, the hypothesis is equivalent to* $\pi_h/\pi_t > r_t/r_h > 1$*. Under the part-(a) condition SHADE preserves the pair.*

Proof. (a). Corollary E.3 gives $|\mathbb{E}[\mathbf{s}^{\mathrm{DR}}_k] - \mathbf{s}^\star_k| \le \varepsilon_{\mathrm{b}}$ for $k \in \{i, j\}$, so

$$\mathbf{s}^\star_i - \mathbf{s}^\star_j > 2\varepsilon_{\mathrm{b}} \implies \mathbb{E}[\mathbf{s}^{\mathrm{DR}}_i] \ge \mathbf{s}^\star_i - \varepsilon_{\mathrm{b}} > \mathbf{s}^\star_j + \varepsilon_{\mathrm{b}} \ge \mathbb{E}[\mathbf{s}^{\mathrm{DR}}_j].$$

For the aggregated version, set

$$\eta_\delta := \Big(\frac{2}{K} - \frac{1}{\max(K, N)}\Big) \max_j |b_j| + B\sqrt{2\max(\kappa_i, \kappa_j)\ln(4/\delta)}, \qquad \kappa_i, \kappa_j \le \frac{2\chi^*}{K^2},$$

(the T_2 and T_1 terms of Proposition E.4); a union bound over the two points gives $|\bar{s}_k - \mathbf{s}^\star_k| \le \eta_\delta$, $k \in \{i, j\}$, simultaneously with probability $\ge 1 - \delta$, and the same implication applies to $\bar{s}$.

(b). The recursion step is the cross-term-free adjacent-rank identity $s_{(\rho)} - s_{(\rho+1)} = a_\rho(v_{(\rho)} - v_{(\rho+1)})$; substituting $\mathbf{v} = \mathbf{r}$ and $\mathbf{v} = \mathbb{E}[\mathbf{v}^{\mathrm{naive}}] = \pi \odot \mathbf{r}$ gives exactly the two displays of the statement. No closed-form condition is claimed for non-adjacent pairs; there $A_q(\pi \odot \mathbf{r})$ has to be evaluated directly. □

The feedback-loop setting (Section 5): a platform reallocates exposure toward the data measured as valuable, the canonical action on data values. At round t, item j has exposure propensity $\pi_t(j) \in [\pi_{\min}, \pi_{\max}]$ under a fixed budget $\sum_j \pi_t(j) = \mathcal{B}$; exposures and labels follow Section 3; per-point values ($\mathbf{v}^{\mathrm{naive}}$, $\mathbf{v}^{\mathrm{IPS}}$, or $\mathbf{v}^{\mathrm{DR}}$) are aggregated per item as $V_t(j)$, the mean of the values of the candidate points of item j. These points share the propensity $\pi_t(j)$, and r_j denotes their mean relevance, so that $\mathbb{E}[V_t(j) \mid \pi_t] = \pi_t(j)\, r_j$ under naive valuation by the first identity of Theorem 5.3. When every item has the same number of candidate points, sums give the same dynamics, because clipnorm is invariant to a common rescaling. The corresponding full-method policies are naive, IPS, and SHADE. The platform reallocates $\pi_{t+1}(j) \propto \mathcal{B}\,[V_t(j)]_+$, clipped to the box. We assume $0 < \pi_{\min} < \pi_{\max}$, a feasible budget $M\pi_{\min} \le \mathcal{B} \le M\pi_{\max}$ for M items, and $r_j > 0$, which make clipnorm well defined. The theorem below isolates the expected value-channel dynamics. The map applies clipnorm to the expected item values; it is a mean-field description, not the expectation of the stochastic update. Per-point values are aggregated per item before exposure is reallocated, while the full KNN-Shapley reallocation used experimentally is treated separately (Section 6.2, Appendix F.6).

Theorem E.6 (Naive valuation concentrates exposure; SHADE restores the relevance-proportional allocation)**.** *Consider the expected exposure dynamics* $\pi_{t+1} = \mathrm{clipnorm}(\mathbb{E}[V_t \mid \pi_t])$*, where* $\mathrm{clipnorm}(w)_j = \mathrm{clip}(cw_j, \pi_{\min}, \pi_{\max})$ *and* $c > 0$ *is chosen so that* $\sum_j \pi_j = \mathcal{B}$*.*

(a) Naive concentrates. *Suppose $\mathbb{E}[V \mid \pi] = \pi \odot \mathbf{r}$ and the r_j are distinct. Then the relevance-proportional allocation $\pi^\star \propto \mathbf{r}$ is not, in general, a fixed point. For any two items i, j that remain interior after an update, $\frac{\pi_{t+1}(i)}{\pi_{t+1}(j)} = \frac{\pi_t(i) r_i}{\pi_t(j) r_j}$. Every fixed point has $\pi^\star(j) \in \{\pi_{\min}, \pi_{\max}\}$ for all but at most one item j, and the boundary assignment is monotone in relevance r: lower-relevance items occupy the lower boundary before higher-relevance items.*

(b) SHADE restores. *If $\hat{\pi} = \pi$ or $\hat{m} = r$, then $\mathbb{E}[V \mid \pi] = r$ for every π, and $\pi^\star_{\mathrm{SHADE}} = \mathrm{clipnorm}(\mathbf{r})$ is the unique fixed point, reached in one expected step from any initial allocation.*

(c) Conditional subset share. *Let $L(\pi^\star) := \{j : \pi^\star(j) = \pi_{\min}\}$ denote the lower-bound set of a naive fixed point. For any pre-specified subset $T \subseteq L(\pi^\star)$, $\frac{\sum_{j \in T} \pi^\star(j)}{\mathcal{B}} = \frac{|T| \pi_{\min}}{\mathcal{B}}$. At the SHADE fixed point the same subset receives $\frac{\sum_{j \in T} \mathrm{clipnorm}(r)_j}{\mathcal{B}}$. No claim is made that a tail defined by initial popularity must belong to $L(\pi^\star)$; that relationship is evaluated empirically in Section 6.2.*

Proof of Theorem E.6 (feedback loop). (a). On the interior, $\mathrm{clipnorm}(w)_j = c w_j$ for the common normalizing constant c, so under $\mathbb{E}[V \mid \pi] = \pi \odot \mathbf{r}$,

$$\frac{\pi_{t+1}(i)}{\pi_{t+1}(j)} = \frac{c\, \pi_t(i) r_i}{c\, \pi_t(j) r_j} = \frac{\pi_t(i) r_i}{\pi_t(j) r_j}.$$

Hence the relative exposure of two interior items is multiplied by r_i / r_j at each update. At a fixed point, an interior item j must satisfy $\pi^\star(j) = c\, \pi^\star(j) r_j$, and therefore $c r_j = 1$. Because the r_j are distinct, at most one item can satisfy this interior fixed-point condition.

Here $\pi^\star(j) \geq \pi_{\min} > 0$ is used to divide. For an item fixed at the lower boundary, $\mathrm{clip}(c\, \pi_{\min} r_j, \pi_{\min}, \pi_{\max}) = \pi_{\min}$ requires $c r_j \leq 1$. Likewise, an item fixed at the upper boundary requires $c r_j \geq 1$. Thus the boundary assignment is thresholded, and hence monotone, in r: lower-relevance items occupy the lower boundary and higher-relevance items the upper boundary, with at most one interior item satisfying $c r_j = 1$.

Finally, if $\pi^\star \propto \mathbf{r}$ and all coordinates are interior, then $\mathrm{clipnorm}(\pi^\star \odot \mathbf{r}) \propto \mathbf{r}^2$, which is not proportional to r unless the relevant coordinates of r are equal. Thus the relevance-proportional allocation is not in general a fixed point of the naive dynamics.

(b). Under either correct nuisance, Proposition 5.1 gives $\mathbb{E}[V \mid \pi] = \mathbf{r}$, independent of the current exposure allocation. Therefore $\pi_{t+1} = \mathrm{clipnorm}(\mathbf{r})$ for every π_t. The map is constant, so its only fixed point is its value $\pi^\star_{\mathrm{SHADE}} = \mathrm{clipnorm}(\mathbf{r})$, which is well defined under the feasibility conditions above and is reached in one step of the expected dynamics.

(c). If $T \subseteq L(\pi^\star)$, then by definition $\pi^\star(j) = \pi_{\min}$ for every $j \in T$, and therefore $\frac{\sum_{j \in T} \pi^\star(j)}{\mathcal{B}} = \frac{|T| \pi_{\min}}{\mathcal{B}}$. Under SHADE, part (b) gives $\pi^\star_{\mathrm{SHADE}} = \mathrm{clipnorm}(\mathbf{r})$, so the corresponding share is $\frac{\sum_{j \in T} \mathrm{clipnorm}(r)_j}{\mathcal{B}}$. The theorem does not identify a subset defined by initial popularity with $L(\pi^\star)$. □

Proposition E.7 (Robustness of the SHADE fixed point under estimated nuisances)**.** *In the setting of Theorem E.6, suppose the nuisances are refit each round with uniform tracking error against the moving exposure, $\Delta_\pi := \sup_t \|\hat{\pi}_t - \pi_t\|_\infty$ (a modeling assumption on the refit protocol, not a consequence of the estimator), that the induced value bias $b(\pi) := \mathbb{E}[V \mid \pi] - r$ is continuous in π, and*

$$\varepsilon_V := \sup_\pi \|\mathbb{E}[V \mid \pi] - r\|_\infty \;\leq\; \Delta_m \Delta_\pi / \varepsilon \;<\; r_{\min}.$$

Then a fixed point of the perturbed map $\pi \mapsto \mathrm{clipnorm}(r + b(\pi))$ exists, and every fixed point $\tilde{\pi}$ satisfies $\|\tilde{\pi} - \pi^\star_{\mathrm{SHADE}}\|_1 \leq L\, \Delta_m \Delta_\pi / \varepsilon$, where L is the $\ell_\infty \to \ell_1$ Lipschitz constant of clipnorm *on $\{w : \min_j w_j \geq r_{\min} - \varepsilon_V\}$ (absorbing the catalog size).*

Proof. Write $\mathbb{E}[V \mid \pi] = \mathbf{r} + b(\pi)$, $\|b(\pi)\|_\infty \leq \varepsilon_V$ (the Corollary E.3 envelope, round-by-round under the tracking assumption). Existence: the budget box $\{\pi \in [\pi_{\min}, \pi_{\max}]^M : \sum_j \pi_j = \mathcal{B}\}$ is compact, convex and nonempty under the feasibility condition, clipnorm is continuous on strictly positive weight vectors, and b is continuous by assumption, so the perturbed map is a continuous self-map of the box and Brouwer applies. clipnorm is scale-invariant, hence not globally Lipschitz, which

is why the statement is localized. On the compact set $\{w : r_{\min} - \varepsilon_V \le w_j \le 1 + \varepsilon_V\}$ it is continuous and piecewise smooth: on each piece the interior coordinates equal $c\, w_j$ with $c = R/\sum_{k\,\text{interior}} w_k$ for a constant R, and the denominator is at least $r_{\min} - \varepsilon_V > 0$, so the derivatives are bounded and L is finite (it grows with the catalog size M). Hence any fixed point $\tilde{\pi} = \mathrm{clipnorm}(\mathbf{r} + b(\tilde{\pi}))$ satisfies

$$\|\tilde{\pi} - \pi^{\star}_{\mathrm{SHADE}}\|_1 = \|\mathrm{clipnorm}(\mathbf{r} + b(\tilde{\pi})) - \mathrm{clipnorm}(\mathbf{r})\|_1 \le L\, \|b(\tilde{\pi})\|_\infty \le L\, \Delta_m \Delta_\pi / \varepsilon.$$

□

Numerical checks of the fixed-point structure and of the two ranking conditions are reported in Appendix E.9.

F Additional Experimental Details

F.1 Synthetic MNAR benchmark and shared hyperparameters

The synthetic benchmark is the one setting in which every quantity of the theory is known, so the predicted bias and its correction can be compared with an exact oracle. Each point carries latent-factor features $\mathbf{x}_i \in \mathbb{R}^{16}$. True relevance $r_i = r(\mathbf{x}_i) \in (0, 1)$ is a smooth function of the latent factors, so the features are informative for r without revealing it. The exposure policy sets π_i through a softmax correlated with popularity and relevance, with bias-strength parameter λ. At $\lambda = 0$ exposure is uniform and low (mean 0.35); at $\lambda = 3$ it is strongly skewed, with propensities reaching 10^{-3}. The oracle valuation $\mathbf{s}^\star = A_q \mathbf{r}$ is computed exactly from the closed form on the same player set and sort order as every competing method, so the methods differ only in the per-point values they pass to the recursion. Table 3 collects the shared hyperparameters. Seeds are configurable through environment variables, and the script defaults reproduce the released result files exactly.

Two supplementary measurements complete the picture of Section 6.1. Where clipping binds, the pooled tail bias of IPS reaches -0.63 in the lowest deciles (-0.39 in decile 2), whereas the deepest-decile bias of SHADE is -0.06. The Gini coefficient of item value mass is 0.257 for naive, 0.221 for IPS and 0.175 for SHADE, against 0.204 for the oracle. Naive valuation over-concentrates value on head items, and SHADE flattens the distribution slightly more than the oracle does.

Table 3: Shared hyperparameters of the valuation experiments (dataset-specific values in Table 5).

players / queries / K	$N = 3{,}000$ (synthetic default), $Q = 20$ oracle queries, $K = 10$
label regimes	*exposure-only*: $\ell = e \cdot \mathrm{Bern}(r)$, so $\mathbb{E}[\ell \mid e{=}1] = r$; *examination bias*: $\eta = 2$, $\mathbb{E}[\ell \mid e{=}1] = \theta r$ with rank-dependent θ
nuisances	cross-fitted logistic regression, $J = 5$ folds; $\hat{\pi}$: e on features + popularity + interactions; $\hat{m}$: ℓ on features among exposed points, complementary fold
misspecification	delete popularity/interaction features ("bad" nuisances); temperature-scale $\hat{\pi}$, $T \in [0.5, 8]$
clipping	$\varepsilon = 0.05$ default; sweep $[0.02, 0.5]$
seeds	≥ 5 per stochastic experiment (default 8), mean $\pm$ std

F.2 Rank recovery and robustness to misspecification (synthetic benchmark)

Remark F.1 (DM is an outcome-model ablation; bias is not ranking). The DM ablation uses $\mathbf{v}_i^{\mathrm{DM}} = \hat{m}(\mathbf{x}_i)$, whose conditional bias under cross-fitting is $\hat{m}(\mathbf{x}_i) - r_i$ (Equation 3): first-order outcome-model error and no propensity-model fallback. DM avoids the residual/IPW variance term, so in our strong-$\hat{m}$ regimes it can reach a higher Kendall τ than SHADE; when $\hat{m}$ carries little relevance signal it can collapse instead (Section 6.1). We therefore do not claim uniform rank dominance for SHADE; its advantage is observation-responsive valuation with double robustness and stronger level fidelity under MNAR.

The synthetic results summarized in Section 6 are reported here in full. They address three points: how well each method recovers the oracle ranking, whether the unbiasedness of the corrected values is visible under repeated sampling, and whether the comparison depends on a logistic exposure model that a logistic $\hat{\pi}$ can invert.

Table 4: Player-set ablation on the synthetic MNAR benchmark. Kendall τ versus the corresponding oracle under weak ($\lambda = 0$) and strong ($\lambda = 3$) exposure bias. Restricting valuation to exposed-only players makes naive valuation conditionally unbiased, while doubly robust valuation loses the all-N player-set advantage predicted in Section 4. Bold indicates the best result within each player-set block.

method (player set)	$\lambda = 0$	$\lambda = 3$
naive (all N)	0.116 ± 0.014	0.219 ± 0.010
IPS (all N)	0.115 ± 0.015	0.219 ± 0.014
SHADE (all N)	$\mathbf{0.556 \pm 0.015}$	$\mathbf{0.549 \pm 0.024}$
naive (exposed only)	$\mathbf{0.371 \pm 0.014}$	$\mathbf{0.353 \pm 0.021}$
IPS (exposed only)	0.367 ± 0.013	0.241 ± 0.032
DR (exposed only)	0.236 ± 0.015	0.306 ± 0.026

On rank fidelity (all-N players, N=3,000, K=10), SHADE sustains Kendall $\tau \approx 0.55$ across all bias strengths λ, while naive and IPS sit at 0.12–0.22. The outcome-model-only DM ablation attains the highest Kendall τ in this synthetic setting. This is the bias–variance behavior anticipated in Remark F.1: when $\hat{m}$ is strong, the low-variance model score can rank well despite lacking pointwise outcome responsiveness and a propensity-model fallback. It is not a uniform advantage. The relative Shapley bias of DM remains 0.99 under label resampling, and DM collapses when $\hat{m}$ is uninformative (Amazon, Section 6.1). The gap between SHADE and naive is also not specific to K=10: over $K \in \{1, \ldots, 50\}$ it stays at $\approx$+0.33.

The choice that matters most is the player set. Restricting the game to exposed points reverses the ordering, as Section 4 predicts: naive valuation is conditionally unbiased there and reaches $\tau \approx 0.36$, and the exposed-only DR ablation drops below it (full numbers in the player-set ablation below).

Unbiasedness was examined directly over 400 label resamples at fixed (X, r, π) and true propensities. The relative Shapley bias of naive plateaus at 0.66 and that of DM at 0.99. The bias of IPS and SHADE instead falls as $1/\sqrt{R}$ to 0.35–0.36 (0.27 with both nuisances true); for SHADE this holds even with a deliberately wrong $\hat{m} \equiv 0.5$, as double robustness predicts. The cost is a 5–8$\times$ per-replicate variance premium. These are the full results behind Remark F.1.

In the 2$\times$2 nuisance-quality grid, SHADE dominates IPS whenever $\hat{m}$ is good (0.55 vs. 0.20). When $\hat{m}$ is bad, clipped SHADE falls to 0.13, below the 0.20 of IPS: true propensities reach 10^{-3}, so ε=0.05 caps the correction.

A semi-synthetic design invites a circularity concern, namely that a logistic $\hat{\pi}$ merely inverts a logistic-generated exposure. Two replications address it. A nonlinear-by-construction exposure model leaves the synthetic ordering unchanged (SHADE 0.545 $\gg$ naive 0.150). On real KuaiRec relevance, SHADE beats naive in 5/5 seeds under all three exposure regimes (logistic; nonlinear, outside the class of $\hat{\pi}$; and $\hat{\pi}$ denied the popularity feature), and the tail share of naive collapses to 0.03–0.07 against an oracle share of 0.253 in every regime (protocol in Appendix F.16).

F.3 Player-set ablation

Restricting the game to exposed players does more than discard data. It changes the estimand: the question becomes the value of a point among the points that were shown, and conditioning on exposure removes the MNAR problem that the correction is meant to address (Section 4). Table 4 gives the full numbers of the exposed-only ablation of Appendix F.2 (Kendall τ vs. oracle, exposure-only regime; bold = best per column within each player-set block).

On exposed-only players naive valuation does not degrade with λ at all, and the exposed-only DR ablation sits below naive at every λ. This is the empirical form of the conditional-unbiasedness and negative-coefficient mechanisms of Section 4. It is not an argument for valuing exposed points only, because the exposed-only game assigns no value to the never-exposed candidates that the all-N estimand covers. The mild increase of all-N naive with λ (0.12 $\rightarrow$ 0.22) is a floor effect. At $\lambda = 0$ exposure is uniform and low, so 65% of players carry $v = 0$ regardless of relevance; naive valuation is uniformly poor on the all-N game, not increasingly so.

F.4 COAT PROTOCOL

Coat (Schnabel et al., 2016) pairs real MNAR exposure with an unbiased reference. It has 290 users $\times$ 300 items, 6,960 MNAR training ratings and 4,640 uniformly-random test ratings; ratings are binarized at ≥ 4. Propensities come from a cross-fitted logistic regression on user and item features plus item popularity (exposure AUC 0.653; estimated propensities in $[0.03, 0.39]$). No true relevance exists for the training pairs, so Kendall τ vs. an oracle is not computable on Coat. The MCAR test slice serves instead as real oracle queries for the pruning and tail-share tasks, over 5 seeds.

The pruning summary of Appendix F.8 rests on the following measurements. Synthetic pruning is flat up to 40% (AUC 0.84). Coat pruning is mildly negative (0.57–0.58 vs. random 0.598), since the mild propensities of Coat leave little to correct.

F.5 INDEPENDENT VERIFICATION AND REPRODUCIBILITY

The reproducibility bookkeeping for all of Appendix F is collected here, so that the later subsections can concentrate on design and interpretation. Full per-seed tables, the experiment and verification scripts, per-family baseline settings, the correctness tests of each baseline implementation, and the registered predictions of every pre-registered experiment with their resolutions are in the code release. That statement is not repeated below.

Recomputation. Every reported number was recomputed from the archived result files by a verification pass whose scripts are separate from the experiment code. The pass comprises 135 recomputation checks on the core suite, 221 on the accuracy suite, 316 on the extension suite (including the pooled 24-seed KuaiRec analysis), 300 on the external-baseline suite, 252 on the augmentation-controls suite, 70 on the value-vs-Shapley-layer suite (Appendix F.14), 57 on the natural-production-exposure suite (Appendix F.11), 42 on the confounding-sensitivity suite (Appendix F.22), and 52 on the fresh-seed-confirmation and candidate-universe suite (Appendices F.9, F.23). These 1,445 checks in nine suites reproduce every reported number to its stated precision.

Fresh seeds. Recomputation shows that the numbers follow from the result files; it does not show that those files are typical draws. The core synthetic/Coat suite was therefore re-executed from scratch on fresh seeds (seeds 100–107). The rerun reproduces the method orderings and reported numbers of that suite within seed variability. The only materially seed-sensitive quantity is the deepest-tail IPS bias decile, where clipping inflates the variance of the raw estimate. The larger-scale and accuracy suites are covered by the recomputation checks only. The frozen-configuration confirmation of the composition result on disjoint seeds is a separate, experiment-specific check (Appendix F.9).

Independent implementations. The recursion-vs-brute-force test was replicated by a third, independently written implementation. TMC Monte-Carlo Shapley converges to the recursion (relative ℓ_2 error 0.026 at 4,000 permutations, N=200). Influence functions cannot be applied to the KNN utility directly, because it is not smooth in model parameters. That family is covered instead by the logistic-surrogate influence baseline of Section 6.4 (protocol in Appendix F.21), alongside LOO, random, Data-OOB (Kwon & Zou, 2023), and uncorrected KNN-Shapley.

Deterministic replay and shared random streams. Several downstream experiments extend an earlier one with further interventions. Each extension reuses the per-seed random streams of its parent experiment, so the drawn exposure, the clicks and the candidate sets are identical, and every intervention shared with the parent must return the numbers of the parent. This is what allows contrasts across those experiments to be paired within seed. The shared interventions reproduce bit-for-bit in the hybrid-reweighting, reweighting-decomposition and augmentation-control experiments, and in the unaugmented baseline, mean-hybrid and SHADE replications inside the composition discovery run. The ranker-family swap keeps the drawn logs and the augmented sets identical per seed. The full-scale seed-0 run of the natural-log KuaiRand-Pure augmentation experiment, executed twice, is bit-for-bit identical. Two reruns crossed machines or library versions. The enriched-outcome-model experiment agrees with its parent to cross-machine BLAS tolerance ($\sim 3\times10^{-5}$ on KuaiRec). In the iterated-augmentation experiment the baseline replicates bit-for-bit and the one-shot intervention to $\sim 3\times10^{-4}$; this library-version drift is 12–30$\times$ below the effect, and the paired effect is exact. The coupling calibration of Appendix F.11 reuses the frozen candidate universes, queries, and per-seed random stream of the KuaiRand-Pure valuation experiment exactly: the per-seed universe hashes are identical and the reported couplings reproduce to max $|\Delta| < 10^{-16}$.

Frozen candidate universes. In the natural-log experiments (KuaiRec big matrix and KuaiRand-Pure) the candidate IDs are hashed when the universe is frozen, before any exposure is looked up, and the hashes are released per seed. For the KuaiRec big matrix the released tests also assert that the IDs are invariant to permuting, zeroing, or replacing the exposure data, and that no code path samples exposed and unexposed pairs separately. The sensitivity experiment of Appendix F.23 records a SHA-256 hash for each frozen universe. KuaiRand-Pure is the official distribution, md5-verified. For the composition confirmation, the complete configuration was hashed before any fresh seed ran; the configuration hash and all outputs are included in the release.

Compute environment. The ML-25M $N{=}6{\times}10^6$ replication ran as a SageMaker training job (16 vCPU / 128 GB, CPU-only, pure NumPy/scikit-learn). The pruning and reweighting experiments and the composition discovery run used the same job type. The augmentation-control experiments on both datasets ran on a 32-vCPU machine, and the enriched-outcome-model, second-ranker-family and iterated-augmentation experiments on a CPU-only machine (32 vCPU, 249 GB). The composition confirmation ran on a different machine and BLAS library than the discovery run. Runtime and memory measurements of the valuation itself are part of the scaling results in Appendix F.7.

F.6 Feedback-loop simulation protocol

Theorem E.6 is stated at the item-aggregated value channel. This simulation instead reallocates exposure from full KNN-Shapley scores: per-point Shapley values are computed each round, aggregated per item, and then acted on. It therefore tests whether the amplification mechanism of the theorem persists through the game layer; it is not a direct instantiation of the theorem.

The setting has $M = 100$ items $\times$ $n = 30$ points each, with item relevance r_j drawn independently of initial popularity, so initial popularity carries no information about relevance. The loop runs $T = 15$ rounds under budget $\mathcal{B} = 30$ (mean exposure 0.3), with $\pi \in [\pi_{\min}, \pi_{\max}] = [0.01, 0.9]$ and $\pi_0 \propto$ popularity. Each round draws $e \sim \mathrm{Bern}(\pi_t)$ and $\ell = e \cdot \mathrm{Bern}(r)$, computes per-point KNN-Shapley values (all-N players, clipping at $\pi_{\min}$), aggregates them per item as the mean, and reallocates $\pi_{t+1} \propto \mathcal{B}[V_t]_+$ clipped to the box; bisection meets the budget exactly after clipping. The methods compared are naive; IPS and SHADE with cross-fitted nuisances refit each round; IPS and SHADE with true nuisances; and the oracle-valuation loop ($v = r$). The reference π^* is the clipped proportional allocation of $\mathcal{B}$ to r. Each round records the tail-quintile (by initial popularity) exposure share, the item-exposure Gini, $\|\pi_t - \pi^*\|_1$, and the mean relevance of the top exposure decile, over 5 seeds.

All learned methods place comparably relevant items at the top of the exposure ranking (0.63–0.71 mean top-decile relevance, within one std). The failure of naive is thus specifically tail under-allocation, which head-centric metrics do not register. The base-configuration L_1 distances to $\pi^\star$ behind Section 6.2 are 17.0 ± 2.1 for SHADE (indistinguishable from true-nuisance SHADE) and 29.4 ± 4.0 for IPS, the worst of the three. The IPS result reflects click-sampling variance and is specific to this configuration (Appendix F.19).

F.7 Real-data protocols: KuaiRec, MovieLens-25M, Amazon

All three constructions follow the same design, so results are comparable across datasets. Features and popularity are fit on data disjoint from the player/query pool. MNAR exposure is imposed as logit $\pi = c_0 + 2(\rho + 0.7z)$, where ρ is standardized log popularity and z is standardized outcome confounding (mean exposure 0.30). Labels follow the bern channel $\ell = e \cdot \mathrm{Bern}(r)$, with $K = 10$ and clip 0.05. Because the exposure is imposed on real outcomes and therefore known, the oracle τ is computable on these datasets. Table 5 summarizes the constructions; the notes below record only what the results depend on.

KuaiRec 2.0 (Gao et al., 2022a). The fully-observed *small* matrix (99.6% observed watch ratios) supplies ground-truth outcomes. Features and popularity are fit on the *big* matrix only, with zero pair overlap (verified programmatically). By design, z enters the exposure but not the nuisance features. Identification (Assumption 1) still holds, while the z-component of π is unlearnable for $\hat{\pi}$: a deliberate Δ_π stress inside the Corollary E.3 envelope, not unmeasured confounding (Appendix F.16 sharpens it; the dedicated violation of Assumption 1(ii) is Appendix F.22). Two label channels are used. The *bern* channel, $r = \mathrm{clip}(wr/2, 0, 1)$, carries the reported results. The *real* channel

Table 5: Real-data experimental constructions (per seed; all official distributions).

dataset	pool	features	N	queries	seeds	nuisance quality
KuaiRec 2.0	1,411 × 3,327 block	d=16 SVD (big matrix)	30k	500 MCAR	5	exp. AUC 0.80
ML-25M	162k × 59k	d=32 SVD (split A)	10^6	200 MCAR	5	0.807; $\hat{m}$ AUC 0.61
ML-25M (ceiling)	as above	as above	6×10^6	200 MCAR	5	identical
Amazon (Video_Games)	2.77M × 137k	same construction	10^5	200 MCAR	5	SVD expl. var. 0.058
Amazon (Books, Movies)	category replications	same	10^5	200 MCAR	5 each	–

($r = \mathbb{1}[wr \geq 1]$, $\ell = e \cdot r$ deterministic) is a caveat channel: exposed labels are then ground truth, which makes exposed-only methods trivially perfect at $\tau = 1$ and rewards IPS over SHADE. On the bern channel of the fully observed block ($N = 30{,}000$; real outcomes, imposed MNAR) SHADE recovers the oracle ranking best, at $\tau = 0.235$ against 0.184 for naive; IPS lies between the two and Data-OOB comes last (Table 6). Naive valuation also under-allocates the tail, a $\sim 4\times$ under-share relative to the oracle. The DM ablation reaches 0.225, the strong-$\hat{m}$ regime of Remark F.1. IPS does not keep its place above naive at scale: at the ML-25M ceiling its value is $\tau = 0.179$, below naive.

MovieLens-25M (Harper & Konstan, 2015). Relevance is the real graded rating, $r = \mathrm{clip}((\text{rating} - 0.5)/4.5, 0, 1)$. The construction is designed to exclude leakage. In a 50/50 user split, split A fits the item factors and ρ. The ratings of each split-B user are then divided per user into a *profile* half (user embedding) and an *experiment* half (the player/query pool), so experiment pairs never touch any feature. Data-OOB runs on a documented 10^5 exposed subsample, because its bagging ensemble is infeasible at $\sim 3\times10^5$ exposed points. The outcome model has $\mathrm{corr}(\hat{m}, r) \approx 0.32$. At $N{=}10^6$ SHADE again leads, at $\tau = 0.264$ against 0.188 for naive, and IPS now falls below naive; the ordering is sign-consistent in every seed (Table 6). The tail share of naive collapses to 0.0001, while SHADE lands on the oracle share. The $N{=}6\times10^6$ replication is the no-leakage pool ceiling: the pool holds $\sim$6.18M leakage-free pairs, so $N = 10^7$ is not attainable. It has identical nuisance quality, and its results appear in Section 6.1 (compute environment in Appendix F.5).

Amazon Reviews 2023 (Hou et al., 2024) (Video_Games, 0-core ratings-only, 4.56M ratings). The relevant failure mode here is feature poverty at density 10^{-5}: SVD explained variance is 0.058 and $\mathrm{corr}(\hat{m}, r) \approx 0.00$. This is the source of the main-text sparsity findings. The Books and Movies_and_TV replications reproduce the pattern. Naive valuation leaves the tail almost no value, and SHADE recovers most of the oracle share, 0.169 of 0.203 on Books (both categories in Table 6). Figure 6(a) plots the Books run.

The bars of Figure 6(a) are means ± std over 8 seeds (synthetic $\lambda{=}2$) or 5 seeds (KuaiRec bern channel, KuaiRand, ML-25M, Amazon-Books). Coat is a single deterministic run with no oracle, since no true relevance exists for its training pairs. The KuaiRand dash is the randomized-slice *reference* (0.175), not an oracle. The ML-25M bars come from the $N{=}6\times10^6$ pool-ceiling run, whose oracle tail share is itself small.

Scale. The ordering is unchanged at the 6×10^6 ceiling: SHADE reaches $\tau = 0.268$ against 0.187 for naive, with IPS below naive, and the signs agree at every seed from $N{=}10^6$ up (Table 6a). The tail under-allocation of naive sharpens with scale to $\sim 500\times$ below the oracle share, while SHADE brings the tail back to the order of the oracle (at or above it in every seed; recovered, not exact).

The end-to-end scaling numbers (Section 6.1) come from single-seed runs of the ML-25M procedure at three sizes, with wall-clock time and peak RSS recorded (Table 6b). The full procedure takes 671 s at the largest size, and nuisance estimation accounts for 80–90% of the time. The recursion alone is linear in N; its laptop anchor at $N{=}10^6$ is in the table. Memory scales as queries$\times N$, so the 200-query end-to-end run sets the peak at 13 GB, and the 6×10^6 replication still fits one CPU instance.

F.8 A REAL RECOMMENDER TRAINED ON CURATED DATA

The valuation results say which interactions carry value. They do not say whether curating a log by those values helps a recommender trained on it. To examine this, we fit an implicit-feedback iALS recommender (Hu et al., 2008) (pure NumPy) on MovieLens-25M under imposed MNAR exposure ($N = 10^6$/seed, 5,953 k-core items, 8 seeds). The log is curated by pruning or by confidence reweighting, and the ranker is evaluated on a held-out uniform-exposure slice (protocol in the next

Table 6: Imposed-exposure real-data constructions of Table 5. *(a)* Kendall τ vs. the oracle and bottom-quintile tail value share (5 seeds, means; the std of the KuaiRec SHADE entry is shown, and ML-25M stds at $N{=}10^6$ are ≤ 0.004). On KuaiRec, Data-OOB reaches $\tau = 0.111$. Bold: highest τ, and the tail share closest to the oracle. Entries not reported in this appendix are marked –. *(b)* Cost of the ML-25M procedure: single-seed wall-clock time and peak RSS.

		Kendall τ			tail value share		
(a) construction	N	SHADE	IPS	naive	naive	SHADE	oracle
KuaiRec (bern)	30k	$\mathbf{0.235 \pm 0.002}$	0.213	0.184	0.067	–	0.253
ML-25M	10^6	**0.264**	0.180	0.188	0.0001	**0.0108**	0.0104
ML-25M (ceiling)	$6{\times}10^6$	**0.268**	0.179	0.187	$9{\times}10^{-6}$	**0.0057**	0.0045
Amazon Books	10^5	–	–	–	0.0004	0.169	0.203
Amazon Movies_and_TV	10^5	–	–	–	0.0005	0.107	0.144

(b) quantity	setting	value
full procedure, wall-clock	$N \in \{10^5, 10^6, 3{\times}10^6\}$	19/284/671 s
recursion alone, wall-clock	$N{=}10^6$, $Q{=}8$, laptop	1.24 ± 0.04 s
recursion alone, peak RSS	same run	1.1 GB
end-to-end, peak RSS	200 queries	13 GB

subsection). Curating a fixed log trades a small accuracy cost for tail coverage; no intervention lifts the top line and the tail together.

Pruning. Pruning degrades Recall@20 monotonically and falls below random pruning (at $q{=}0.3$: SHADE/IPS 0.039 vs. random 0.053). The reason is that the lowest-valued points are informative tail interactions. The same mild pruning negative appears on the 7k real points of Coat (Appendix F.4).

Reweighting (per-method results in Table 7). At 16 seeds no reweighting scheme gains on both axes. Every value with tail reach costs top-line accuracy: naive reweighting changes Recall@20 by -0.0070 ($p{=}2{\times}10^{-8}$), whereas a random-weight control is a true null. Correction cuts the cost 2–4$\times$ and carries 3.5–5$\times$ the tail coverage of naive (IPS 0.0069, SHADE 0.0046 vs. 0.0013, $p{\leq}0.0004$), so naive is worst on both axes. The schemes that do gain have zero tail coverage. They are Data-OOB ($+0.0004$ NDCG@20) and, of three a-priori Data-OOB/SHADE hybrids, an equal-weight mean that attains the best top line of any reweighting scheme (Recall@20 $+0.0022$). A pre-registered decomposition locates the cost in the sharp item-level up-weights that produce the tail reach. We conclude that reweighting is not the accuracy channel of SHADE. On a fixed MNAR log, no reweighting can recover relevance signal that was never logged. The accuracy gains of SHADE come from channels that a fixed log does not offer: augmenting from the never-exposed set (Section 6.3) and steering what gets collected (Section 6.2).

F.9 Downstream ranker protocol

The experiment reuses the ML-25M no-leakage construction of Appendix F.7 with two changes, both needed to make a real tail measurable. The first is a k-core item filter (≥ 100 pool appearances). Without it the bottom popularity quintile is $\sim$31k items appearing $< 5\times$ in a 10^6 sample, which floors tail-coverage at 0 for every method including an oracle. The second is a lowered exposure skew ($\lambda : 2 \to 1$). Under it the bottom quintile receives $\sim 0.7\%$ of training clicks, still $\sim 30\times$ under a uniform share, so the tail is severely under-exposed yet clickable; at $\lambda = 2$ it receives 0. The filter is applied at analysis time, and the underlying derived dataset is unchanged. Each seed draws exposure and clicks (mean exposure 0.300, $\sim$235k clicks), computes all-N values, curates, and trains. Ranker, evaluation, and intervention settings are collected in Table 8. Tail-coverage@20 is the fraction of top-20 recommendations in the bottom popularity quintile, and recommendation Gini is the concentration of exposure across items.

The reweighting claims rest on a 16-seed rerun, sized for power, that adds tail-segment Recall/NDCG@20 (relevant sets restricted to bottom-quintile items, $\sim$456 qualifying test users per seed). These tail-segment metrics are floor-degenerate on the fixed log: every method scores 0.0000 except a single IPS seed at 0.002. This is consistent with reweighting being unable to add tail signal

Table 7: **Fixed-log reweighting** (ML-25M, 16 seeds; iALS, uniform-exposure eval; paired-t vs. the no-intervention baseline, NDCG@20 0.0243; every nonzero entry differs from baseline at $p<0.01$ except the tail coverage of naive, $p<0.05$, and the random weights, not significant). Reweighting is not the accuracy channel of SHADE (Appendix F.8): no value gains on both axes; values with tail reach cost the top line (correction cuts the cost 2–4×), and the values that gain have zero tail coverage. Bold: best per column.

value used	ΔNDCG@20	tail coverage@20
naive	−0.0030	0.0013[♭]
IPS	−0.0014	**0.0069**
Data-OOB	+0.0004	0
mean(OOB, SHADE)	**+0.0008**[♮]	0
random	+0.0001	0
SHADE	−0.0021	0.0046

[♭]The tail coverage of naive is tested vs. the no-intervention baseline (the convention of the table); the "worst on both axes" reading compares it to IPS/SHADE, whose coverage is 3.5–5× larger ($p\leq 0.0004$ vs. naive).
[♮]Table entries are ΔNDCG@20; the mean(OOB, SHADE) gain quoted in Appendix F.8 (+0.0022) is Recall@20.

Table 8: Key hyperparameters of the downstream-ranker experiments (protocol-specific deviations in the paragraphs of this subsection).

universe	ML-25M k-core ($\geq$ 100 pool appearances): 5,953 items / 5.83M pairs; exposure skew $\lambda = 1$; $N = 10^6$ per seed
ranker	iALS (Hu et al., 2008): 32 factors, 12 ALS sweeps, confidence $c_{ui} = 1 + 40\, w_{ui}$ ($w_{ui} = 1$ for a click under pruning; min-max-normalized value under reweighting), ridge 0.1, closed-form NumPy updates; second family: EASE, $\lambda_{\text{EASE}} = 5000$
evaluation	held-out uniformly-exposed slice, 50k pairs, relevance threshold 0.7
augmentation	top-B of U, $B \in \{1, 5, 10\}\%$ of observed positives (frontier sweeps to 20%), $w_{\text{aug}} \in \{0.25, 1\}$; pre-registered settings ML-25M B=5%/w_{aug}=0.25, KuaiRec B=5%/w_{aug}=1
collection loops	synthetic: T=10 rounds, 200k impressions/round, budget $\mathcal{B} = 0.3M$; ML-25M: T=8, 250k/round; $\pi \in [0.01, 0.9]$
seeds	pruning grids 8; reweighting/hybrid/composition 16; augmentation 8 (KuaiRec pooled 24); collection loops 5

that was never collected. Significance is paired-t over seeds against the matched no-curation / random baseline (compute environment in Appendix F.5).

Collection-loop ranker (Section 6.2). The fixed-log results leave open whether corrected valuation helps once it decides what is collected. This experiment runs the loop of Section 6.2 at 2,000 users × 1,000 items (loop settings in Table 8). Candidate impressions are shared across policies, nuisances are refit every round, and the policies are naive, IPS, SHADE, static and oracle. The accumulated clicked log of each policy (∼272k events) trains the same iALS ranker, which is evaluated on a fixed MCAR slice (relevance threshold 0.8; tail = bottom initial-popularity quintile) at checkpoints $t \in \{2, 5, 10\}$ over 5 seeds.

The exposure trajectories reproduce the dynamics of Figure 3 inside this run: naive tail exposure stays pinned just above its floor, and SHADE reaches 0.195 from round 2, close to the 0.200 of $\pi^\star$ (Table 9a). On the tail segment, the ranker trained on the SHADE-collected log matches the one trained on the oracle-collected log: tail-Recall@20 is 0.096 against 0.097, and neither tail metric differs significantly. Naive collection leaves both tail metrics at the floor, significantly below SHADE in all 5 seeds (Table 9c). The static never-reallocate policy sits 5–6× below SHADE on tail metrics, so the gain comes from corrected reallocation specifically. The top line behaves differently. Top-line Recall/NDCG@20 favors the naive policy (NDCG@20 0.119 against 0.068 for SHADE), above even the oracle policy. IPS attains the best single tail-vs-top-line point, but it gets there by overshooting its converged tail exposure (0.257 vs. 0.200) with near-naive concentration of its recommendations.

Collection loop on real relevance (Section 6.2). The same mechanics are applied to the ML-25M k-core universe, with real graded r and real SVD features, to see which part of the synthetic result

transfers. The loop runs $T{=}8$ rounds of 250k real-pair candidate impressions, disjoint from the eval slice. The propensity clip is 0.05, the real-data convention (the 0.01 of the synthetic loop was its known design floor). The relevance threshold is 0.7, the checkpoints are $t \in \{2, 5, 8\}$, and there are 5 seeds. The reference $\pi^\star$ is the clipped-proportional allocation to per-item mean relevance, with tail share 0.193.

The allocation behavior transfers. The tail exposure share of SHADE rises to 0.265–0.273 by round 2 and tracks the oracle policy at every round, and its tail click share is well above that of naive (Table 9b). Naive flattens at 0.104–0.109, well below $\pi^\star{=}0.193$. The ranker trained on the SHADE log stays slightly below the one trained on the oracle log: the gaps are -0.004 on Recall@20 and -0.002 on NDCG@20, both significant, and the HR@10 gap is not significant (Table 9c). The head-reward caveat is stronger here. The static policy attains the best top line of all five (NDCG@20 0.0349, above the 0.0342 of naive), and tail factors do not enter a top-20 among 5,953 candidates for any policy.

Table 9: Collection-loop ranker experiments (iALS trained on the log each policy collects; MCAR evaluation slice; 5 seeds). *(a)* Synthetic loop, 2,000 users $\times$ 1,000 items; the exposure floor is 0.0067 and the recommendation Gini of IPS is 0.597. *(b)* ML-25M k-core loop, where the tail-segment ranker metrics are floor-degenerate for every policy. Tail exposure is the bottom-quintile exposure share, given as a range where it moves over rounds (the IPS entry is its converged value); $\pi^\star$ is the reference allocation. *(c)* Paired t-tests over seeds. Entries not reported in this appendix are marked –.

(a) policy	tail exposure	tail-Recall@20	tail-NDCG@20	Recall@20	NDCG@20
naive	0.008–0.011	0.002	0.001	0.154	0.119
IPS	0.257	0.159	–	–	0.095
SHADE	0.195	0.096 ± 0.003	0.043	0.094	0.068
oracle ($v{=}r$)	–	0.097	0.044	–	–
reference $\pi^\star$	0.200				

(b) policy	tail exposure	tail click share	NDCG@20
naive	0.104–0.109	0.0067	0.0342
SHADE	$0.088 \rightarrow 0.265$–0.273	0.043	–
oracle ($v{=}r$)	0.258–0.264	–	–
static	–	–	0.0349
reference $\pi^\star$	0.193		

(c) contrast	metric	estimate	p
synthetic, SHADE vs. oracle policy	tail-Recall@20, tail-NDCG@20	–	≥ 0.11
synthetic, SHADE vs. naive	tail-Recall@20, tail-NDCG@20	5/5 seeds	$\leq 3.2{\times}10^{-6}$
ML-25M, SHADE − oracle policy	Recall@20	-0.004	0.005
ML-25M, SHADE − oracle policy	NDCG@20	-0.002	0.014
ML-25M, SHADE − oracle policy	HR@10	-0.003	not significant

Hybrid reweighting (Appendix F.8). Data-OOB gains on the top line without tail coverage, and SHADE does the opposite, so it is natural to check whether a combination keeps both. The 16-seed reweighting experiment was rerun with three pre-registered Data-OOB/SHADE hybrids on clicked positives: the elementwise product; an equal-weight mean (fixed 0.5, no tuning); and a gate that gives bottom-quintile item positives ($\sim$0.6–0.7% of clicks) the SHADE weight. The methods shared with the 16-seed run reproduce it exactly (Appendix F.5), so all tests across methods are paired within seed. No hybrid keeps both properties. The product inherits the full cost of SHADE with the tail coverage of SHADE. The gate is indistinguishable from Data-OOB. The mean attains the best top line of any reweighting scheme (Recall@20 $+0.0022$ vs. baseline, and significantly above Data-OOB itself; Table 10) yet has zero tail coverage at all 16 seeds.

Reweighting-cost decomposition (Appendix F.8). To locate the source of the reweighting cost, the same 16-seed experiment was run with five pre-registered surgical variants of the SHADE weights over item-popularity deciles: a mean-one rescale (mass removed), within-decile rank-normalization (tilt removed, ordering kept), a per-decile mean broadcast (ordering removed, tilt kept), $\hat{m}$ as weight, and a floor $0.5 + 0.5\,w$. None has a tuning parameter. The results reverse the natural guess that the

popularity tilt causes the cost. Broadcasting the per-decile mean SHADE weight, which keeps the tilt and destroys the ordering, turns the cost of SHADE into a small significant gain (NDCG@20 $+0.0006$, statistically at the level of Data-OOB). Its Recall@20 gain is the best of any reweighting scheme and significantly above that of Data-OOB (Table 10). Keeping the ordering without the tilt retains the cost (-0.0017, significant). Rescaling to mean one is catastrophic (-0.0079 NDCG@20), which shows that the mass shrinkage of the min–max normalization was silently protective. Every softened form has zero tail coverage: the sharp item-level up-weights that carry the cost are also what produce the tail reach.

Augmentation (Section 6.3). Augmentation uses the one thing a corrected value has and a naive value lacks, namely a score on never-exposed pairs. Unexposed pairs (e=0; $\sim$700k of 10^6 on ML-25M, KuaiRec bern construction with N=30k) are ranked by (i) the SHADE score, (ii) $\hat{m}$ directly, (iii) random, or (iv) true relevance (oracle). On U the per-point $\mathbf{v}^{\mathrm{DR}}$ value equals $\hat{m}$, so option (ii) ablates the recursion step. This $\hat{m}$-only selection is distinct from the DM ablation of Remark F.1: DM applies the KNN-Shapley recursion A_q to $\mathbf{v}^{\mathrm{DM}} = \hat{m}$, whereas $\hat{m}$-only selection ranks U directly by $\hat{m}$ and intentionally omits the recursion. The top-B pairs enter the iALS confidence matrix as pseudo-positives at w_{aug} (grid and pre-registered settings in Table 8; 8 seeds each, the Table 14 KuaiRec row pools 24). Naive valuation has no counterpart: its value is identically zero on U.

All tests are paired-t vs. the unaugmented baseline of the same run; the baseline levels and the test statistics are collected in Table 10. The two main results of Section 6.3 are both significant: $\hat{m}$-only selection on ML-25M, and $\hat{m}$-enriched selection on KuaiRec with its pooled 24 seeds. A KuaiRec sweep $m_a = (1-a)\hat{m} + a\,r$, $a \in \{0, 0.5, 1\}$, finds the gain flat in a: at $\mathrm{corr}(\hat{m}, r) \approx 0.63$ the outcome model already saturates the small oracle headroom. The a=1 selection is identical to the oracle selection on every seed. On secondary metrics, the ML-25M $\hat{m}$-only selection also gains Recall@20. The ML-25M NDCG@20 gain of the oracle selection is $+0.0003$, and $\hat{m}$-only significantly exceeds it, a top-line head-preference artifact, not selection perfection. Random augmentation is null or marginal on both datasets.

Mechanism and headroom details (Section 6.3). The oracle-exceedance on ML-25M reflects the head preference of the top-line metric, not selection perfection. $\hat{m}$ concentrates on confidently-relevant head pairs, which help a head-heavy evaluation slice most; this is the same head reward that inflates the naive collection policy in Section 6.2. Ranking U by $\hat{m}$ directly beats the full SHADE score, significantly so on KuaiRec (Table 10): on U the per-point value $\mathbf{v}^{\mathrm{DR}}$ equals $\hat{m}$, and the query-anchored recursion dilutes it by feature-space proximity. The claim is therefore corrected-vs-naive (the corrected value function covers U; that of naive does not), not Shapley-vs-$\hat{m}$. The headroom is structurally modest. Exposure that loads on quality depletes U of relevant pairs (KuaiRec: $\sim$6.5% of U relevant vs. 13.7% marginally), so augmentation gains are single-digit-percent.

The IPS weighting inside $\hat{m}$ is not what produces the gain either. Refitting $\hat{m}$ without its IPS sample weights, on the same random streams (Appendix F.5), ties the corrected selection at every setting on both datasets: the NDCG@20 differences are negligible and not significant, and the two selections overlap at 0.93–0.97. Composition with the full SHADE weights (instead of the OOB–SHADE mean) turns the reweighting cost of SHADE net-positive on Recall@20 ($+0.0026$, significant) but collapses the tail coverage of SHADE ($0.0046 \rightarrow 0.0002$). Head-leaning pseudo-positives push the up-weighted tail items back out of the top-20, so even the composed procedure remains subject to the cost–coverage coupling of Appendix F.8.

Augmentation controls and practitioner baselines (Section 6.3). A gain from $\hat{m}$-selected pseudo-positives is informative only if simpler selectors do not produce it. The two augmentation experiments were therefore rerun on the same random streams (Appendix F.5) with further selection rules: $\hat{m}$ refit without IPS weights (the correction ablation), item popularity, Laplace-smoothed item CTR, self-training (the scores of the baseline iALS itself on U), item–item cosine, ascending cross-fitted $\hat{\pi}$ (blind exploration), and random. Every rule is applied to the same drawn log, budget and ranker within a seed, so the comparison with correction-free heuristics is paired and on equal terms. The budget frontier sweeps $B \in \{1, 2.5, 5, 10, 20\}\%$ at w_{aug}=1. Per-seed diagnostics record the corr of each score with true relevance on U and the top-B overlaps between selections (the 0.88 $\hat{m}$–exploration overlap on KuaiRec and 0.000 on ML-25M discussed in Appendix F.10).

The frontier is what separates the selectors (Table 11). On ML-25M self-training stays at $\sim$+3.7% from B=5% to B=10% and falls significantly behind from B=10% on, while $\hat{m}$ selection grows

Table 10: Test statistics of the fixed-log interventions (iALS, uniform-exposure evaluation; paired t-tests over seeds; estimates are differences in the named metric). Reweighting variants: ML-25M, 16 seeds, against the no-intervention baseline unless another reference is named. Augmentation: 8 seeds per dataset against the unaugmented baseline (NDCG@20 0.0244 ML-25M, 0.0106 KuaiRec), except the pooled 24-seed KuaiRec row and the composition with the full SHADE weights, which uses 16 seeds (baseline 0.0243). Estimates marked * are repeated from Table 14; "–" means that only the test is reported.

intervention	reference	metric	estimate	p
Hybrid reweighting and cost decomposition (ML-25M)				
mean(OOB, SHADE)	baseline	Recall@20	$+0.0022$	3×10^{-5}
mean(OOB, SHADE)	Data-OOB	Recall@20	–	0.011
per-decile mean broadcast	baseline	NDCG@20	$+0.0006$	0.036
per-decile mean broadcast	baseline	Recall@20	$+0.0037$	–
per-decile mean broadcast	Data-OOB	Recall@20	–	2×10^{-4}
within-decile rank-normalization	baseline	NDCG@20	-0.0017	4×10^{-7}
mean-one rescale	baseline	NDCG@20	-0.0079	–
Augmentation from U, ML-25M				
$\hat{m}$-only	baseline	NDCG@20	$+0.0010^{*}$	4×10^{-5}
$\hat{m}$-only	baseline	Recall@20	$+0.0024$	–
oracle selection	baseline	NDCG@20	$+0.0003$	–
$\hat{m}$-only	oracle selection	NDCG@20	–	0.002
SHADE weights + $\hat{m}$-only	baseline	Recall@20	$+0.0026$	10^{-4}
SHADE weights + $\hat{m}$-only	SHADE weights alone	tail coverage@20	$0.0046 \to 0.0002$	0.002
Augmentation from U, KuaiRec and both datasets				
$\hat{m}$-enriched, KuaiRec (pooled 24)	baseline	NDCG@20	$+0.0005^{*}$	0.0021
$\hat{m}$-only, KuaiRec	full SHADE score	NDCG@20	–	0.003
random, both datasets	baseline	NDCG@20	–	≥ 0.054
$\hat{m}$ without IPS weights, both	$\hat{m}$-only	NDCG@20	$\lvert\Delta\rvert \leq 2\times10^{-5}$	≥ 0.84

monotonically to $+10.6\%$ at $B{=}20\%$. On KuaiRec the top line of $\hat{m}$ selection rolls over at $B{\approx}5\%$, consistent with the $\sim 6.5\%$ oracle-relevance ceiling of the selected pool. Smoothed CTR keeps growing there but stays tail-neutral-to-negative throughout. Exploration decays until it ends below the baseline and significantly below $\hat{m}$ selection at the largest budget, and the tail frontier of $\hat{m}$ selection compounds to $+280\%$.

Natural-log replication with randomized evaluation (Appendix F.10). The two augmentation results above rest on imposed exposure. KuaiRand-Pure allows the same intervention to be tested on a natural MNAR log: it provides a normal-policy log ($12{,}171 \times 7{,}583$) and a separate uniformly randomized intervention slice for evaluation. So that the player set does not depend on realized exposure, the protocol first samples and freezes an *exposure-independent candidate universe* of $N{=}150{,}000$ user–item pairs per seed ($\sim$uniform per item, period-1 feature-period pairs excluded) before consulting pair-level normal-policy exposure (candidate hashes in Appendix F.5). Only afterward are candidates assigned to

$$P = \{i : e_i{=}1,\ \ell_i{=}1\}, \qquad C = \{i : e_i{=}1,\ \ell_i{=}0\}, \qquad U = \{i : e_i{=}0\}$$

from the normal-policy log. The exposed/unexposed composition is thus induced by the natural logging policy and is not fixed through exposure-stratified sampling. Across seeds the frozen universe is almost entirely never-exposed, with a realized exposure rate of 0.0015 ± 0.0001 (set sizes in Table 12). This sparsity is expected. Candidate sampling does not condition on realized exposure, so the observed fraction reflects the sparsity of the production logging policy itself, and the construction aligns the empirical candidate universe with the fixed-player-set condition used by the valuation theory. At the pre-registered 5% budget the integer selection rule $\max(1, \mathrm{round}(0.05\,|P|))$ therefore adds only a handful of pseudo-positive pairs per seed.

The randomized intervention slice is used only for evaluation and reference quantities. It does not determine candidate membership or augmentation selection, and all interventions share the same frozen universe within each seed, which preserves paired comparisons. The iALS ranker, nuisances,

Table 11: Correction-free selection rules against $\hat{m}$ selection at the B=5% setting and along the budget frontier, which is swept at w_{aug}=1 (8 seeds per dataset). Entries are the relative change in NDCG@20 against the unaugmented baseline; p-values are paired t-tests over seeds of each rule against $\hat{m}$ selection at the same budget, except the $\hat{m}$ row at B=20%, which is tested against the baseline. Overlap is the top-B overlap with the $\hat{m}$ selection.

selection rule	budget B	rel. ΔNDCG@20	p	overlap
		ML-25M		
$\hat{m}$ selection	20%	+10.6%	3×10^{-5}	
self-training	5–10%	∼+3.7%		
self-training	10%	–	1.5×10^{-4}	
self-training	20%	–	3.7×10^{-3}	
smoothed CTR, popularity	5%	below $\hat{m}$	≤0.011	
low-$\hat{\pi}$ exploration	5%	≈ random		0.000
		KuaiRec		
$\hat{m}$ selection	5% (peak)	+5.0%		
smoothed CTR	5%	+10.5%	0.045	
smoothed CTR	20%	+16%		
low-$\hat{\pi}$ exploration	5%	matches $\hat{m}$		0.88
low-$\hat{\pi}$ exploration	20%	−3.6%	0.0043	

selection rules, budgets, and diagnostics are those of the experiments above; the unbiased relevance reference is per-item random-slice CTR (7,467 items, ≥10 impressions). The eval slice is fixed across seeds, so paired contrasts are valid while the seed spread under-represents eval-sampling noise. The effective evaluation set is small, 90.6 ± 10.6 users per seed: randomized-slice users holding at least one clicked evaluation pair and at least one training positive.

Table 12: Natural-log augmentation on KuaiRand-Pure: the frozen exposure-independent universe, the randomized evaluation, and the tests (8 seeds, mean ± std over seeds; paired t-tests against the unaugmented baseline; iALS). The randomized slice serves only evaluation and reference quantities.

quantity	value
Frozen candidate universe (per seed)	
candidate pairs N	150,000
clicked positives P	102 ± 11
exposed-unclicked pairs C	123 ± 10
never-exposed pairs U	$149{,}775 \pm 16$
realized exposure rate	0.0015 ± 0.0001
pseudo-positives added at the pre-registered 5% budget	5.1 ± 0.8
Randomized evaluation	
evaluation users per seed	90.6 ± 10.6 (range 77–107)
baseline NDCG@20	0.0022 ± 0.0010
$\text{corr}(\hat{m}, \text{CTR}_{\text{rand}} \mid U)$	0.02 ± 0.13
top-set overlap, $\hat{m}$ selection and low-$\hat{\pi}$ exploration	0.094
Tests	
$\hat{m}$-only selection at the pre-registered setting, ΔNDCG@20	+0.0001 (p=0.65)
smallest p over the 1–20% sweep, any rule and metric	p=0.085 (exploration)
$\hat{m}$ selection, all budgets	$p \geq 0.22$
smoothed CTR, all budgets	$p \geq 0.18$

Cross-channel composition (Section 6.3): discovery and frozen-configuration confirmation. Augmentation and reweighting act on different parts of the log, the first on never-exposed pairs and the second on observed positives, so the two can be combined. The composed intervention reweights observed positives and adds $\hat{m}$-selected pseudo-positives. Pseudo-positives always enter at w_{aug} regardless of the reweighting scheme (the reweight maps only observed positives). The

interventions compared are the unaugmented baseline, the mean hybrid and SHADE reweighting (each replicating its earlier run exactly, Appendix F.5), $\hat{m}$-augmentation alone, each reweighting composed with $\hat{m}$-augmentation, and a mean-hybrid + random-augmentation control. Additivity is a paired t on the per-seed interaction $(\Delta_{\text{composed}} - \Delta_{\text{reweight}} - \Delta_{\text{augment}})$.

Discovery (seeds 0–15*).* The 16-seed hybrid experiment, extended with the augmentation step at the pre-registered ML-25M setting, found NDCG@20 $+0.0019$ ($+7.7\%$), positive in every seed, a Recall@20 gain of similar relative size, and a null reweight×augment interaction (Table 13). The order of selection matters for how this estimate should be read. The mean(OOB, SHADE) parent was identified as the best reweighting scheme in the hybrid experiment (its win was not predicted there), on the same 16 seeds. The composition experiment was then registered with that winner, with the set of interventions fixed before the composed run and no correction beyond this disclosure (the composed effect exceeds a Bonferroni bound over the family of interventions by orders of magnitude). Because the parent was selected on the seeds that produced the estimate, the discovery run alone carries a post-selection concern.

Confirmation (seeds 200–215*).* To remove that concern, the complete procedure was frozen before any fresh seed ran: candidate pool and derived features, iALS settings, nuisances and clipping, the reweighting normalization, budget B=5%, w_{aug}=0.25, and the full fixed set of interventions (configuration hash in Appendix F.5). It was then evaluated once on 16 fresh seeds disjoint from the discovery/selection seeds and from every other seed set in this paper, with no further tuning and the pre-specified criterion "positive paired mean ΔNDCG@20 with 95% CI excluding zero." The criterion is met. Composed − baseline NDCG@20 is $+0.0016$ ($+6.6\%$ of its own confirmatory baseline), positive in every fresh seed, with a CI that excludes zero; the paired t-test and a sign-flip permutation test agree, and Recall@20 rises as well (Table 13). The composed intervention significantly exceeds the reweighting parent, the augmentation parent, and the random-augmentation control, whose own gain is statistically that of the reweighting parent. The interaction is null, consistent with additive composition, and the zero tail coverage reproduces. The confirmatory point estimate is 14% below the discovery estimate, which lies inside the confirmatory CI.

The independent confirmation resolves the parent-selection concern for the composed-effect claim. The fresh-seed estimate is the primary composition result of the paper, and the discovery run is retained as discovery evidence only. The two seed sets are never pooled.

Table 13: Cross-channel composition on ML-25M (mean(OOB, SHADE) reweighting composed with $\hat{m}$ augmentation; iALS; 16 seeds per column; paired t-tests over seeds unless noted). The confirmation column is the frozen-configuration run on fresh seeds and is the primary result; the two columns are never pooled. Entries are differences in NDCG@20 unless a metric is named; relative sizes are against the baseline of the same run.

contrast	discovery (seeds 0–15)	confirmation (seeds 200–215)
composed − baseline	$+0.0019$ ($+7.7\%$)	$+0.0016$ ($+6.6\%$)
95% paired CI	–	$[+0.0012, +0.0020]$
paired t-test	$p=2\times10^{-9}$	$p=1.3\times10^{-7}$
sign-flip permutation, 200k resamples	–	$p=3\times10^{-5}$
seeds positive	16/16	16/16
composed − baseline, Recall@20	$+0.0050$ ($+8.4\%$)	$+0.0041$ (15/16 seeds)
composed − reweighting parent	–	$+0.0009$ ($p=1\times10^{-6}$)
composed − augmentation parent	–	$+0.0006$ ($p=6\times10^{-4}$)
composed − random-augmentation control	–	$+0.0008$ ($p=9\times10^{-5}$)
that control − reweighting parent	–	$p=0.46$
reweight×augment interaction	$p\geq0.38$	-0.0000 ± 0.0002 at 95%, $p=0.65$

Enriched outcome model (Appendix F.10). The nuisance-interpolation results of Appendix F.20 locate the rank-recovery gap in $\hat{m}$, so we examined whether better features for $\hat{m}$ convert into downstream gains. The augmentation experiments were rerun with three feature sets for the same cross-fitted IPS-weighted logistic $\hat{m}$ (same $\hat{\pi}$): an SVD-only control, SVD + side information, and side information alone. The ML-25M side features are a tag-genome truncated SVD (64 dims, 72% variance), per-user mean genome profiles from the profile half (no pool-half leakage), and 20 genre indicators. All are external 2019-frozen content descriptors that cannot leak the drawn exposure or

clicks (309/213 features vs. the 97-dim control). The KuaiRec side features are item category multi-hot, per-item daily-engagement statistics, user demographics, and big-matrix category-affinity profiles (219/171 vs. 49). Everything else follows the composition (ML-25M) and KuaiRec augmentation protocols verbatim at the pre-registered settings and on the same random streams (Appendix F.5). Each feature set reports held-out AUC/log-loss, $\mathrm{corr}(\hat{m}, r)$ on all/exposed/never-exposed pairs, the true-relevance rate of the selected top-B, and the full set of interventions. ML-25M uses 8 seeds. KuaiRec uses 8 plus a pre-registered confirmatory extension of 16 seeds, registered before any extension seed ran and with no optional stopping. The pooled 24-seed analysis is the reported one, with both head-to-head contrasts named as the primary tests.

Second ranker family (Appendix F.10). To check that the augmentation gain is not a property of iALS, the ML-25M augmentation protocol was repeated with the ranker swapped to EASE (Steck, 2019) (one $5{,}953 \times 5{,}953$ closed-form solve). The grid, seeds, and random streams are exactly those of the iALS experiment. The drawn logs and the augmented sets are therefore identical per seed, and the ranker family is the only change. Pseudo-positives enter at w_{aug} in both the Gram matrix and the prediction-time user vector. Elementwise-weighted EASE has no closed form, so this is the standard cheap analogue of iALS confidence, a documented deviation. $\lambda_{\mathrm{EASE}} = 5000$ was calibrated once on the seed-0 baseline before any augmented run (metrics plateau on $[5000, 20000]$). Cross-ranker statistics pair the per-seed delta of each intervention against its own baseline, over 8 seeds.

Iterated augmentation (Appendix F.10). If one round of augmentation helps, repeating it with the retrained ranker as selector might be expected to help more. The ML-25M augmentation experiment at its pre-registered setting was therefore run with $T{=}4$ self-training rounds at matched total budget: each round scores the remaining unexposed pairs with the current ranker, adds the top-b_t as pseudo-positives, and retrains. The interventions are one-shot augmentation (replication accuracy in Appendix F.5), uncorrected iteration, a guarded iteration, and a full-budget-per-round control that separates iteration from budget splitting. The guard allocates the round budget across popularity quintiles $\propto$ the share of $\sum \hat{m}$ attributable to each quintile over the remaining unexposed pairs, so the ranker is trusted only within strata. Per-round diagnostics are the head-quintile share, the mean true relevance and relevant fraction of each batch, and $\mathrm{corr}(\text{selection score}, r)$ before selection. A synthetic threshold companion sweeps MNAR severity with oracle-blend selectors $(1{-}\gamma)\, z(\text{ranker}) + \gamma\, z(r)$, an instrument to span the correlation axis and not a deployable method, and pools per-round $(\mathrm{corr}, \Delta\text{NDCG@20})$ pairs to locate the crossover.

The detailed results, summarized in Appendix F.10, are as follows. The NDCG@20 change against one-shot is small and not significant, alongside the Recall@20 loss (Table 15). Retraining on the head-biased pseudo-positives of each round degrades the next selection monotonically (mean batch relevance $0.753 \to 0.723$). The guard pins the head share of the selection at 0.63 vs. 0.9998 uncorrected, but it costs top-line accuracy on this head-heavy evaluation, ending significantly below one-shot on NDCG@20, and brings no measurable tail benefit, since the tail segment floors for every intervention. Controlling the diagnostic therefore yields no accuracy gain here, which we report as found. All results use 8 seeds.

F.10 Accuracy experiments: heuristics, scope, ranker family, enrichment, iteration

This subsection analyzes the results summarized in Section 6.3; the protocols are in the preceding subsection. Its purpose is to mark where the augmentation gain holds and where it stops.

Correction-free heuristics match the gain at isolated settings, not across budgets and datasets. Practitioner alternatives (self-training, item–item CF similarity, smoothed item CTR, blind low-$\hat{\pi}$ exploration, random; 8 seeds, both datasets, protocols above) behave differently on the two datasets at the $B{=}5\%$ setting (Table 11). On KuaiRec, smoothed CTR significantly beats the top line of $\hat{m}$ selection ($+10.5\%$ vs. $+5.0\%$ NDCG@20) while staying tail-neutral-to-negative, and low-$\hat{\pi}$ exploration reproduces the entire tail lift and top line of $\hat{m}$ selection, with which its top-B set largely coincides. On ML-25M the same heuristics fail: exploration is $\approx$ random and shares nothing with the $\hat{m}$ selection, and CTR/popularity fall significantly below it. What separates the selectors is the budget frontier. On ML-25M the outcome channel grows monotonically to $+10.6\%$ at $B{=}20\%$ and is the only selector that keeps growing there; on KuaiRec only its tail frontier never rolls over. No

Table 14: **NDCG@20 across intervention channels** (iALS, uniform-exposure eval; paired-t vs. the no-intervention baseline of each channel; per-block seeds, baselines, and reweighting parents in Table 7 and Appendix F; exact p-values in Appendices F.9–F.10 and the footnotes). Tail-NDCG@20 is floor-degenerate for the fixed-log ML-25M interventions (those rows report tail *coverage*@20) and for real-data collection policies (–). Bold: best per column within each fixed-log block; the collection block is unbolded because its top line rewards the head-concentrated distortion itself ([‡]). The composition block reports the independent frozen-configuration confirmation on 16 fresh seeds, disjoint from the discovery/selection seeds; the discovery estimates ($+0.0019$, $+7.7\%$) are in Appendix F.9.

channel	value used	ΔNDCG@20	tail signal	vs. oracle selection
Fixed log: augmenting from U (naive $\equiv 0$ on U: no naive selection)				
augment, ML-25M	random	$+0.0002$	cov 0	below
augment, ML-25M	SHADE	$+0.0002$	cov 0	below
augment, ML-25M	$\hat{m}$-only	$+\mathbf{0.0010}$ ($+4.1\%$)	cov 0	exceeds[†]
augment, KuaiRec	random	0.0000	–	below
augment, KuaiRec	SHADE	-0.0002	–	below
augment, KuaiRec	$\hat{m}$-enriched[¶]	$+\mathbf{0.0005}$ ($+5.1\%$)	tail-rec $+\mathbf{0.0046}$	matches
Fixed log, composed: reweight × augment (ML-25M; frozen-configuration confirmation, 16 fresh seeds)				
compose	mean+random aug	$+0.0008$[§]	cov 0	–
compose	SHADE+$\hat{m}$-only aug	$+0.0005$	cov **0.0002**	–
compose	mean(OOB,SHADE)+$\hat{m}$ aug.	$+\mathbf{0.0016}$ ($+6.6\%$)	cov 0	–
Collection loop: ranker on the collected log (5 seeds; abs. NDCG@20, not Δ)				
collect, synth.	naive	0.1185	tail-NDCG 0.001	above[‡]
collect, synth.	oracle (v=r)	0.0666	tail-NDCG 0.044	=
collect, synth.	SHADE	0.0681	tail-NDCG 0.043	matches ($p \geq 0.11$)
collect, ML-25M	naive	0.0342	–	above[‡]
collect, ML-25M	oracle (v=r)	0.0251	–	=
collect, ML-25M	SHADE	0.0231	–	tracks (-0.002, p=0.014)

[†] $+0.0010$ vs. oracle $+0.0003$ at the quoted w_{aug}=0.25 setting (p=0.002); see the head-preference caveat in the text. [‡] top-line NDCG rewards the head-concentrated distortion itself: naive beats even perfect-knowledge collection. [§] statistically the gain of the reweighting parent (vs. mean(OOB, SHADE) alone p=0.46); the $\hat{m}$-composed intervention beats the random-composed one by $+0.0008$ ($p = 9\times10^{-5}$). [¶] pooled 24-seed analysis, side-information-enriched $\hat{m}$ (Appendix F.10): the enrichment carries the significance; the earlier SVD-only 8-seed estimate ($+0.0005$, p=0.015) was a favorable draw, not significant at 24 seeds ($+0.0003$, p=0.074).

single heuristic delivers head and tail on both datasets. The advantage of the outcome channel lies in budget-scaling and cross-dataset generality, not in uniqueness at one setting.

Scope: the natural-log augmentation result remains null. KuaiRand-Pure tests whether the augmentation gain extends to a natural MNAR normal-policy log, evaluated on the randomized intervention slice of the platform. Under the frozen-universe construction the augmentation channel remains null. Across the full 1–20% budget sweep no tested budget produces a statistically reliable improvement for any selection rule on any metric (Table 12). At the pre-registered setting, $\hat{m}$-only selection changes NDCG@20 by $+0.0001$, far from significance. A downward smoothed-CTR effect observed before the universe was frozen in this way is no longer significant at any budget.

The experiment does not support the reading that this null is "a scale failure, not a scorer failure." Outcome-model relevance ranking is itself weak and uncertain in this block: $\text{corr}(\hat{m}, \text{CTR}_{\text{rand}} \mid U) = 0.02 \pm 0.13$, and the top-set overlap between $\hat{m}$ selection and low-propensity exploration is 0.094. At the same time, the frozen 150,000-pair universe contains only 102 ± 11 clicked positives and 123 ± 10 exposed-unclicked pairs per seed, so the pre-registered 5% intervention adds only 5.1 ± 0.8 pseudo-positive pairs. The experiment therefore identifies a persistent null result, not its unique mechanism: weak outcome-model signal, sparse observed support, and small effective intervention size are intertwined in this setting. Accordingly, the positive augmentation gains of Section 6.3 remain scoped to the two simulated-exposure protocols. The natural-log result is evidence against a universal end-to-end accuracy claim, not evidence that corrected valuation must improve a recommender whenever natural exposure is present.

The gain survives a change of ranker family; its relative size does not. Swapping iALS for EASE (Steck, 2019), with the drawn logs and augmented sets held identical so that the ranker is the only change, gives a 45% stronger baseline (NDCG@20 0.0354 vs. 0.0244). The $\hat{m}$-augmentation gain survives on it, at $+0.0006$, significant and positive in every seed, and the full ordering of interventions is replicated (Table 15). It attenuates significantly, to $\sim 0.6\times$ the absolute delta of iALS, i.e. +1.8–2.6% relative on the higher baseline. The +4–5% and +6.6% (confirmatory composed) relative figures are thus iALS-specific; the direction, significance, and ordering of the channel are not. The ML-25M tail floor persists under EASE, so it is a data/exposure property, not a ranker artifact.

Better outcome-model features help only where selection binds. Feeding the same logistic learner real side information (ML-25M: tag-genome SVD + genres; KuaiRec: categories, engagement statistics, user features) raises $\mathrm{corr}(\hat{m}, r)$ on never-exposed pairs from 0.418 to 0.504 on ML-25M and by a similar margin on KuaiRec, significantly in every case (Table 15). The nuisance-ceiling lever of Section 6.1 is real. It converts to NDCG only where selection binds. On KuaiRec (24 seeds: 8 plus a pre-registered 16-seed confirmatory extension; ~ 0.28 of selected pseudo-positives relevant) the enriched selection significantly beats the SVD-only control head-to-head ($+0.0003$ NDCG@20) and the baseline by $+0.0005$ (+5.1%). At 24 seeds the un-enriched control alone is not significant. Its 8-seed batch was a favorable draw, so the Table 14 KuaiRec row reports the pooled enriched result; the enrichment therefore carries the significance. ML-25M shows no conversion, for a measured reason. The SVD-only selection is already 88% oracle-relevant, and enrichment cannot raise it; the small change it produces goes in the wrong direction. Enriched and control selections are statistically identical, and the composed gain does not grow (+7.06% vs. +6.95%). Budget and evaluation geometry bind there, not $\hat{m}$ information.

Iterating the augmentation is a measured negative. Retrain-and-reselect rounds end at or below one-shot at matched total budget (Recall@20 -0.0007, significant; 8 seeds), and budget splitting alone is not the cause (Table 15). Iteration brings a failure mode of its own. The MNAR-trained ranker is a $\sim 4\times$ worse relevance model on never-exposed pairs than $\hat{m}$ (corr with r: 0.10 vs. 0.42), and its uncorrected selection saturates at 99.98% head-quintile in the first round. This is an empirical confirmation-bias effect qualitatively analogous to the amplification mechanism of Theorem E.6, but not covered by that theorem (the pseudo-label/retraining loop is outside its formal scope). A synthetic threshold study puts break-even selector quality at $\mathrm{corr}(\mathrm{score}, r) \gtrsim 0.35$, about the level of the one-shot selector itself, which the trained ranker misses by an order of magnitude. An exposure-corrected guard controls the selection diagnostic but yields no accuracy gain (protocol above). One-shot $\hat{m}$ selection is therefore the design, not a simplification.

F.11 Real-exposure experiment protocols

The imposed-exposure experiments know the propensities by construction. The two experiments here instead use exposure as it occurred on a production platform. Because the guarantees condition on a player universe selected independently of realized exposure (Section 4), both experiments freeze the candidate universe before any exposure is read.

KuaiRec big matrix (12.5M interactions, 7,176 users $\times$ 10,728 videos; the production log itself, exposure NOT imposed). Users are split into disjoint halves. The embedding/query half fits d=16 SVD factors and supplies 500 exposed query pairs per seed; the player half supplies the candidate universe. For each seed we first sample and freeze a universe of $N = 10^6$ distinct (player-half user, video) candidate pairs, approximately uniform per item over the full grid, without consulting pair-level production exposure (hashes and invariance tests in Appendix F.5). Only after the universe is frozen do we assign $e_i = 1$ iff the pair occurs in the production log and $\ell_i = \mathbb{1}[wr \geq 1]$ when exposed. The exposed/unexposed proportions are thus induced by the log and not fixed by sampling quotas: realized exposure is 0.134 and the realized click rate 0.048 per seed. $\hat{\pi}$ is a cross-fitted logistic model on the real exposure (AUC 0.921); 5 seeds. No oracle τ exists, because true r is unknown off-log. The reported metrics are the tail-quintile share of positive item value, the item-value Gini, and Spearman(item log-popularity, item value). The tail-share contrast of Section 6.1 (naive vs. SHADE) is significant at p=4×10^{-8} here and at p=10^{-4} on KuaiRand-Pure. The strongly negative coupling of DM (-0.51) is its own distortion, not a win.

KuaiRand-Pure (Gao et al., 2022b) follows the same pre-exposure design. Per seed, a universe of $N = 50$k candidate pairs over split-B users $\times$ all 7,583 items is frozen, $\sim$uniform per item, excluding

Table 15: Statistics of the three follow-up experiments on the augmentation gain (iALS unless EASE is named; paired t-tests over seeds; estimates are differences in the named metric). The EASE rows are at the pre-registered ML-25M setting unless w_{aug} is given. – means that only the estimate or only the test is reported.

contrast	metric	estimate	p
Second ranker family (EASE, ML-25M, 8 seeds)			
$\hat{m}$ augmentation − baseline	NDCG@20	$+0.0006$ (8/8 seeds)	3×10^{-5}
same at $w_{\text{aug}}=1$	NDCG@20	$+0.0009$	–
gain under EASE relative to iALS	NDCG@20	$\sim0.6\times$	0.0015
Enriched outcome model, KuaiRec (pooled 24 seeds)			
SVD-only to enriched features	$\text{corr}(\hat{m}, r)$ on U	$0.439 \to 0.497$–0.524	$<10^{-6}$
enriched − SVD-only selection	NDCG@20	$+0.0003$ (19/24 seeds)	0.0022
95% CI		$[+0.0001, +0.0004]$	
enriched selection − baseline	NDCG@20	$+0.0005$ ($+5.1\%$)	0.0021
SVD-only selection − baseline	NDCG@20	$+0.0003$	0.074
Enriched outcome model, ML-25M (8 seeds)			
SVD-only to enriched features	$\text{corr}(\hat{m}, r)$ on U	$0.418 \to 0.504$	$<10^{-6}$
SVD-only to enriched features	oracle-relevant share	wrong direction	0.07
enriched − SVD-only selection	NDCG@20	–	0.32
composed gain, enriched vs. SVD-only	rel. NDCG@20	$+7.06\%$ vs. $+6.95\%$	0.80
Iterated augmentation (ML-25M, 8 seeds, matched total budget)			
iterated − one-shot	Recall@20	-0.0007	0.028
iterated − one-shot	NDCG@20	-0.0002	0.096
full-budget-per-round control vs. iterated	NDCG@20	–	0.67
guarded − one-shot	NDCG@20	-0.0004	0.020
mean batch relevance, first to last round	–	$0.753 \to 0.723$	2×10^{-6}

only the period-1 pairs of each user (the feature period whose clicks build the user embeddings; an x-side filter). The randomized intervention slice plays no role in player membership. It supplies only the unbiased tail reference (0.175 of randomized click mass) and 500 MCAR queries per seed. Exposure is then read from the period-2 natural log: realized exposure is 0.0014 (56–86 exposed pairs per seed) and the click rate 0.0007, so the experiment inherits the true sparsity of the log. Full oracle-τ is not identifiable, since player overlap with the random slice is $\sim0.6\%$; it is reported as coverage, not τ. The same slice measures $\hat{m}$ miscalibration directly (predicted 0.460 vs. randomized true 0.175, AUC 0.514). With ~70 exposed training points the outcome model carries little relevance signal, which is the binding constraint on this experiment.

Coupling calibration. A popularity–value coupling of zero is not the right target when relevance itself correlates with popularity. The couplings of Section 6.1 are therefore scored against measured unbiased targets (Table 16). The calibration reuses the frozen candidate universes and queries of the KuaiRand-Pure experiment (replay details in Appendix F.5), and all couplings are Spearman over the same item pool (items with ≥ 10 randomized impressions; 5 seeds). The oracle-value target is KNN-Shapley on the same players/queries with unbiased item-CTR values, which is the value concept of the paper itself ($+0.245 \pm 0.002$; the raw popularity–CTR coupling of the randomized slice is $+0.399$). A pair-level learned oracle gives $+0.758$, but popularity is one of its features and item aggregation cancels its noise; it is a knowingly inflated upper bound and is not used. At the natural sparsity of the log every method undershoots the target (p=0.001–0.007, under in 5/5 seeds each), and the paired |gap| contrast between naive and SHADE is not significant ($+0.084$, p=0.58). The rank-coupling comparison does not discriminate the methods on this log, and we report it as such without adjudicating.

On the KuaiRec side, the block targets are ≈ 0 (relevance vs. big-matrix popularity -0.024 binary / -0.021 graded; exposure-free oracle-value couplings -0.046 ± 0.006 / -0.132 ± 0.009). The big-matrix IPS coupling is -0.016, between the $+0.126$ of naive and the -0.322 of SHADE. Against the targets IPS lands nearest, naive overshoots to $+0.13$, and SHADE over-corrects to -0.32; the DM ablation (-0.51) overshoots further in the negative direction. This comparison is directional only:

the 3,327 items of the block are a popular-skewed subset of the production log, so no significance is claimed.

Table 16: Coupling calibration on KuaiRand-Pure (Spearman(popularity, item value), 5 seeds; gaps vs. the oracle-value target $+0.245 \pm 0.002$; the raw popularity–CTR coupling on the randomized slice is $+0.399$). The valuation player universe is sampled and frozen before pair-level normal-policy exposure is read; the randomized intervention slice is used only to form the exposure-free calibration target. At the natural exposure of the log (0.0014) every method undershoots the target and no pairwise method contrast is significant (naive vs. SHADE |gap|: p=0.58); the comparison is underpowered here, reported as such. DM is included as an outcome-model diagnostic, not as one of the main corrected valuation methods. Bold: smallest |gap| among the main valuation methods (nominal only, given the non-significant contrasts).

method	coupling	gap vs. target	
naive	-0.271 ± 0.137	-0.516 ± 0.137	p=0.0011, under in 5/5 seeds
IPS	-0.271 ± 0.137	-0.515 ± 0.137	p=0.0011, under in 5/5 seeds
SHADE	$\mathbf{-0.187 \pm 0.187}$	$\mathbf{-0.432 \pm 0.187}$	p=0.0067, under in 5/5 seeds
DM (outcome-model ablation)	-0.186 ± 0.190	-0.430 ± 0.190	(diagnostic; p=0.0072)

F.12 Decile-level bias structure (Section 6.1)

Figure 2 shows the decile-resolved counterpart of Theorem 5.3 summarized in Section 6.1. Binning by true-propensity decile (λ=2), naive under-values the mid/tail by up to -0.56 and over-values the head by up to $+0.51$ (the sign flip of Theorem 5.3(b)). IPS corrects the head but not the clipped tail, and SHADE stays flat ($|\text{bias}| \leq 0.2$). On real Coat exposure the tail popularity quintile gets 8.8% of positive value mass under naive vs. 16–18% under SHADE/IPS, half its share.

F.13 Solution-concept ablation protocol

This subsection gives the protocol of the solution-concept comparison reported in Appendix F.21. Each concept is run on the raw and on the corrected value channel, to see whether the tail under-allocation and its correction depend on the game. The setting is synthetic MNAR at λ=2, N=2,000, Q=300, 5 seeds; the bound on N comes from the budget of the MC semivalue estimators, not from SHADE. Data Banzhaf uses the MSR paired estimator with 4,000 uniform coalition draws. Beta$(4, 1)$-Shapley uses Beta-binomial size sampling with $20N$ draws. Its implementation is validated by the exact identity Beta$(1, 1) \equiv$ Shapley (max deviation 10^{-12} at N=8 brute force) and by MC convergence checks. Threshold-KNN-Shapley uses a within-radius additive utility (radius = 5% distance quantile) in closed form. Influence functions use the analytic inverse-Hessian for L2 logistic regression on observed labels (exposed-only, like Data-OOB), validated against actual leave-one-out retraining (rank correlation > 0.8). Every concept is run on two value channels, raw $e \cdot \ell$ and $\mathbf{v}^{\mathrm{DR}}$, and the tail-quintile share and τ vs. the all-N oracle are reported for both. The ablation has known weak points: the τ of Banzhaf is poor in both channels at this MC budget, and influence functions are nearly orthogonal to the oracle ranking ($\tau = 0.015$). In wall-clock terms at N=2,000, the exact recursion values all points in 0.12 s vs. 17/249 s for the Banzhaf/Beta MC estimators; the closed form is what makes the $\mathbf{v}^{\mathrm{DR}}$-channel correction deployable across concepts.

F.14 Value layer vs. Shapley layer (Corollary 5.2)

One may ask whether the game does real work, or whether the recovery is already complete in the debiased per-point values. To separate the two, each method is scored at three levels on the main configuration (Appendix F.1: λ=2, N=3,000, Q=300, K=10, all-N players, 8 seeds): at the value layer (per-point $\hat{v}$ vs. the oracle value $\mathbf{r}$), at the Shapley layer (query-averaged $\hat{s}$ vs. $\mathbf{s}^\star$), and cross-layer ($\tau(\hat{v}, \mathbf{s}^\star)$: the raw value predicting the target of the game itself).

Against the common target $\mathbf{s}^\star$, the recursion adds fidelity for the corrected value $\mathbf{v}^{\mathrm{DR}}$ ($\tau(\mathbf{v}^{\mathrm{DR}}, \mathbf{s}^\star) = 0.407 \pm 0.013$ vs. $\tau(\mathbf{s}^{\mathrm{DR}}, \mathbf{s}^\star) = 0.547 \pm 0.020$) and destroys it for biased ones (naive $0.329 \to 0.205$; IPS $0.303 \to 0.204$): query-locality aggregation compounds value-level bias. The DM ablation moves $0.590 \to 0.762$. The large τ gain of DM here reflects strong outcome-model signal and low

variance; it does not alter its first-order dependence on $\hat{m} - r$. Consequently, the corrected-over-naive gap more than doubles through the SHADE recursion. At the value layer, each method against its own-layer oracle, it is 0.510 ± 0.013 vs. 0.383 ± 0.012 (gap 0.13); at the Shapley layer it is 0.547 vs. 0.205 (gap 0.34). Tail under-allocation, by contrast, originates in the value layer (naive tail-quintile share 0.005 vs. oracle 0.107 before any game; SHADE 0.111; these are shares of per-point value mass, not the Shapley-layer shares of Table 1), and the game neither creates nor corrects it, consistent with Corollary 5.2 and with its K-independence (Appendix F.2). A caveat on measurement: value-layer τ for the sparse binary channels (naive/IPS: most entries exactly 0) leans on τ_b tie handling, so cross-layer comparisons are cleanest on the common-target column. The analysis also covers a single configuration, the main one.

F.15 Denoising-baseline protocol (Section 6.4)

Denoising methods also score logged interactions, which makes them a natural external comparison. The setting is the synthetic MNAR setting of the solution-concept ablation (λ=2, N=2,000, 5 seeds), with an exposed-only player set for comparability with Data-OOB. Both denoising methods are reduced to the NumPy logistic backbone used throughout this section; the reductions are surrogates of the published methods and are labeled as such.

T-CE (ADT; Wang et al., 2021). The "memorization gap" that separates clean from noisy interactions is realized by *cross-fitting* a per-point clean score $\exp(-\mathrm{CE}_i)$. This is the probability that an L2-logistic outcome model (same family as $\hat{m}$) assigns to the observed label of i out-of-fold, so a mislabeled point cannot fit its own wrong label. The top loss quantile is then hard-dropped (the dynamic drop-rate schedule of ADT collapsed to its converged plateau). **BOD** (Wang et al., 2023b). A bi-level loop (5 outer steps) uses an L2-logistic inner solver refit under per-point sample weights and a one-step meta-gradient update $w_i \leftarrow w_i + \mathrm{lr} \cdot \langle g_{\mathrm{val}}, g_i \rangle$, which aligns the train gradient of each point with a clean-validation gradient; the value is the learned weight. This reproduces the mechanism of BOD, not its exact matrix-factorization update, and the logistic backbone recovers no oracle ranking ($\tau \approx 0$), which we report without tuning. The released correctness tests verify that injected label flips receive lower clean scores than true labels with high probability, and both baselines are deterministic given the seed.

The full numbers behind Section 6.4 are oracle τ T-CE 0.164, BOD -0.028 vs. SHADE 0.522, and tail-quintile share 0.09/0.08 vs. SHADE 0.181 (oracle 0.179 in this N=2,000, 5-seed setting). The difference is structural. Denoising scores which interactions to trust; it does not recover how much each is worth, nor value the never-exposed log.

F.16 Anti-circularity on real relevance

The concern addressed here is circularity: a logistic $\hat{\pi}$ applied to logistic-generated exposure may succeed for reasons that would not hold elsewhere. The test uses the KuaiRec small block as in the KuaiRec construction of Appendix F.7 (N = 30k, real graded relevance, 5 seeds) and varies the imposed exposure across three regimes on identical outcomes. Regime (i), *logistic*, is the baseline, in which the class of $\hat{\pi}$ can represent π. Regime (ii), *nonlinear*, sets $\mathrm{logit}\,\pi = c_0 + \lambda(1.5\mathbb{1}[\rho > 0.5] - \mathbb{1}[\rho < -0.5] + \rho z)$, outside the logistic class; this is confirmed by AUC$(e; \hat{\pi})$ dropping $0.797 \to 0.778$ and mean $|\hat{\pi} - \pi|$ rising $0.104 \to 0.146$. Regime (iii), *dropfeat*, keeps logistic exposure but denies $\hat{\pi}$ the popularity feature ρ, the bias driver. SHADE beats naive in 5/5 seeds in all three regimes (τ: 0.238/0.181/0.234 vs. 0.184/0.123/0.184), and the tail collapse of naive (0.03–0.07 vs. oracle 0.253) is regime-invariant. In these regimes the advantage does not depend on $\hat{\pi}$ being able to represent the exposure mechanism. One caveat: under nonlinear exposure the DM ablation beats SHADE on τ (0.220 vs. 0.181). This is the low-variance outcome-model regime described in Remark F.1, not a uniform DM advantage.

F.17 Tail-share decomposition: exposed vs. never-exposed (Section 6.1)

A matching tail share does not by itself show that value sits on the same points as under the oracle. This decomposition (synthetic MNAR, N=3,000, λ=2, 5 seeds) splits the bottom-quintile value share of SHADE by exposure status. On these five seeds the share is 0.176 for SHADE and 0.201 for the oracle; Table 1 reports 0.225 and 0.215 on eight seeds, a difference within the seed-to-seed spread

of the SHADE share (± 0.061). Exposed points contribute 0.083 ± 0.041 ($\sim$45% of the tail mass, from only 4.9% of tail points being exposed; the $e(\ell - \hat{m})/\hat{\pi}$ correction upweights that minority sharply). Never-exposed points contribute 0.093, which undershoots the never-exposed tail mass of the oracle (0.183; the oracle places $\sim$91% of its tail mass on never-exposed points); the exposed-point correction (0.083 vs. oracle 0.018) makes up most of that shortfall, though not all of it. The tail share of SHADE is thus assembled differently from that of the oracle, with more of it carried by the few exposed tail points.

F.18 Payment-fairness simulation (Section 6.1)

When data values determine payments to contributors, the tail bias becomes a payment bias. In this simulation (synthetic MNAR, $N = 3{,}000$, $\lambda \in \{1, 2, 3\}$, 5 seeds each) a fixed budget is paid to items in proportion to $[\text{item value}]_+$, where the item value is the sum of the Shapley values of its points (all-N players). The fair target is the oracle-value payout. The metrics are the tail-quintile payout share vs. oracle, the tail underpayment factor (oracle share / method share), the misallocated budget fraction ($\frac{1}{2}L_1$ between payout vectors), and Spearman(popularity, payout). At λ=2 naive underpays the tail 55×, with +0.847 coupling and 52% of the budget misallocated. IPS gives 1.65× and +0.24; SHADE gives 0.88× and −0.05. DM has the lowest misallocation (0.21) but pays by prediction (Remark F.1). IPS is unstable in λ: it overpays at λ=1 and collapses to 4.2× underpay at λ=3 (click-sampling variance). Figure 5 plots both metrics across the λ sweep.

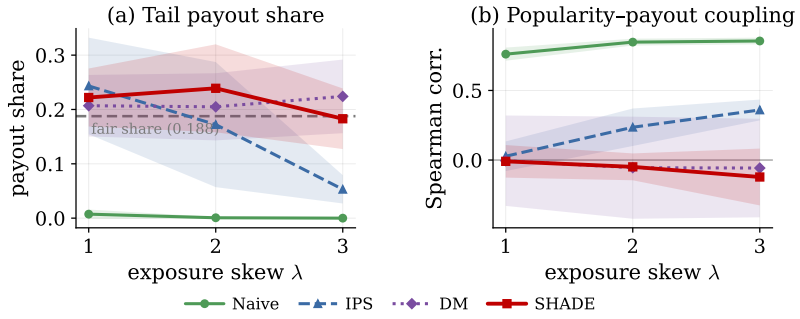

Figure 5: **Payment fairness across exposure skew** (synthetic MNAR, N=3,000, 5 seeds per λ, ±std bands). *(a)* Tail-quintile payout share vs. λ; grey dashed line: the fair (oracle-payout) share 0.188, an item-level payout share on these seeds and not the point-level tail share of Table 1. Naive pays the tail essentially nothing at every skew (share $\to 0$ at λ=3); IPS decays with λ (click-sampling variance); SHADE and the DM ablation stay near the fair share in this synthetic setting; the favorable level of DM here depends entirely on outcome-model calibration and has no propensity-model fallback (contrast the Amazon collapse in Section 6.1). *(b)* Spearman(popularity, payout): the payout of naive tracks popularity (+0.76 to +0.86); SHADE stays near zero to mildly negative.

F.19 Feedback-loop stress test (Section 6.2)

The feedback result of Section 6.2 is reported at one configuration, so we check that it does not depend on that choice. The simulation loop of Appendix F.6 is rerun with one setting varied at a time and the others at their defaults. The swept settings are the exposure floor $\pi_{\min}$, the number of items M, and the budget per item $\mathcal{B}/M$ (grid in Table 17), with 5 seeds per configuration and converged (round-T) allocations reported. Naive has the lowest popularity-tail exposure share in every configuration, often lying near the numerical floor $n_{\text{tail}}\pi_{\min}/\mathcal{B}$ and rising with $\pi_{\min}$ as that floor does. This floor proximity is an empirical property of the simulated initial-popularity tail; Theorem E.6 characterizes the relevance-ordered boundary concentration but does not assert that the popularity-defined tail must equal the lower-bound set. SHADE reaches 0.13–0.20 vs. oracle 0.16–0.20, within one standard deviation of the fair reference in 11 of 12 configurations and far above naive in every one. The IPS-worst finding of the main loop does not replicate across the sweep, where IPS is above naive everywhere; it is specific to that configuration (Appendix F.6).

Table 17: Feedback-loop stress test: swept settings (one varied at a time, the others at the defaults of Appendix F.6) and the converged popularity-tail exposure share of each policy across the configurations (5 seeds per configuration).

swept setting	values
exposure floor $\pi_{\min}$	$\{0.005, 0.01, 0.02, 0.05, 0.1\}$
number of items M	$\{50, 100, 200, 400\}$
budget per item $\mathcal{B}/M$	$\{0.15, 0.3, 0.5\}$
policy	**converged tail exposure share**
naive	lowest in every configuration; mean 0.094 at $\pi_{\min}$=0.005 $\rightarrow$ 0.142 at 0.1
IPS	0.13–0.22
SHADE	0.13–0.20
oracle	0.16–0.20

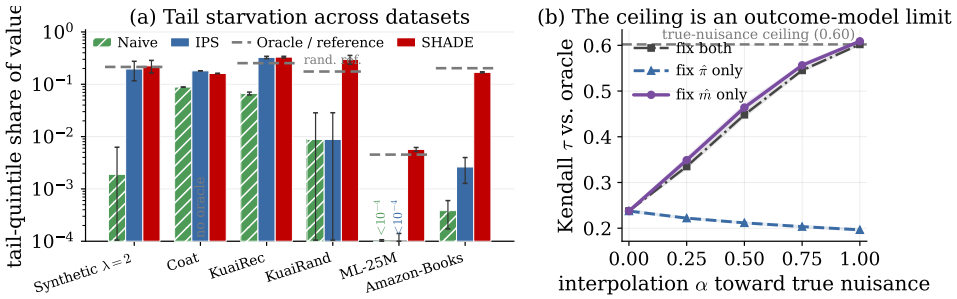

Figure 6: **Tail under-allocation is pervasive; the correction is $\hat{m}$-limited.** *(a)* Tail value share on every dataset tested (log axis, floored at 10^{-4}; grey dashes: oracle/reference shares; caveats in Appendix F.7). *(b)* Nuisance interpolation (KuaiRec, 5 seeds): only the $\hat{m}$ path climbs to the true-nuisance ceiling.

F.20 Nuisance-interpolation decomposition (Section 6.1)

Estimated SHADE sits far below its true-nuisance ceiling on real data (Section 6.1). This decomposition attributes the gap to one nuisance or the other. It uses the KuaiRec small block with features fixed at the d=16 default, so the oracle Shapley and the true-nuisance SHADE ceiling are fixed endpoints; 5 seeds. Nuisances are interpolated toward truth, $m_\alpha = (1-\alpha)\hat{m} + \alpha r$ and $\operatorname{logit} \pi_\alpha = (1-\alpha)\operatorname{logit} \hat{\pi} + \alpha \operatorname{logit} \pi$, in three sweeps (outcome only, propensity only, both) plus the 2×2 corners. The corners are $(\hat{\pi}, \hat{m})$ 0.238; $(\pi, \hat{m})$ 0.196; $(\hat{\pi}, r)$ 0.609; (π, r) 0.602. Setting $\hat{m}$ to its true value alone recovers $+102\%$ of the gap, and setting $\hat{\pi}$ alone -11% (tightening $\hat{\pi}$ without a better $\hat{m}$ removes the variance-buffering effect of the outcome model, consistent with double robustness). Along all three sweeps τ collapses onto the $\Delta_m \Delta_\pi$ product (Corollary E.3).

Three further measurements bear on whether the ceiling is a matter of information or of model capacity. The companion feature sweep (SVD dimension $d \in \{4, \ldots, 64\}$ feeding both nuisances, learner fixed) shows that richer features of the same type raise τ absolutely (0.210 $\rightarrow$ 0.264) but leave the recovered fraction flat ($\sim$40%), because the oracle geometry rises in lockstep. A stronger off-the-shelf learner (gradient boosting for $\hat{m}$) is worse on synthetic (τ 0.547 $\rightarrow$ 0.448, 8/8 seeds; it overfits the $\sim 10^3$ exposed points) and yields essentially no change on KuaiRec (0.238 $\rightarrow$ 0.250, closing $\sim$3% of the gap to the true-nuisance ceiling). A variance-optimal (MRDR-style) outcome loss shifts KuaiRec τ by only $+0.007$ (5/5 seeds) while hurting on synthetic. These are three independent confirmations that the ceiling is informational, not a capacity limit.

F.21 External baselines: results and protocols (Section 6.4)

Production-Shapley designs and exposed-only alternatives. A direct numeric comparison with Kersbergen et al. (2025); Zhang et al. (2025) is not possible: they report ranking metrics on a biased holdout, whereas we report τ against an exposure-free oracle. We therefore reimplement their shared

design, an uncorrected on-log utility over exposed-only players, which coincides with naive valuation there (verified to equality). That game is tail-blind by construction. Its tail share matches its own exposed-only oracle (0.086 vs. 0.087), because the game has no tail mass to allocate, and it reaches only $\tau = 0.359$ against that oracle (SHADE: 0.547 against the all-log oracle).

The corruption is not specific to the Shapley game. Data Banzhaf (Wang & Jia, 2023), Beta$(4, 1)$-Shapley (Kwon & Zou, 2022), and Threshold-KNN-Shapley (Wang et al., 2023a) on the raw $e{\cdot}\ell$ channel all under-allocate the tail (0.007–0.051 vs. oracle 0.179 in the N=2,000 setting of Appendix F.13), and the $\mathbf{v}^{\mathrm{DR}}$ value channel corrects every concept at once (0.17–0.19); the bias enters through $\mathbb{E}[v \mid \pi] = \pi r$, before any game is defined, as Corollary 5.2 predicts (protocol in Appendix F.13).

Exposed-only alternatives do not escape it either. Data-OOB (Kwon & Zou, 2023) needs an observed label per point, so it cannot value the never-exposed 65–93% of the log, and it is worst-in-class on τ and tail share. The denoising baselines T-CE (ADT; Wang et al., 2021) and BOD (Wang et al., 2023b) and influence functions (Koh & Liang, 2017) trail SHADE badly on oracle τ and under-allocate the tail. IPS-correcting only the validation loss of an influence function, with true propensities, recovers almost none of the tail share (0.070 vs. oracle 0.238). The correction must act on the per-point value, not the evaluation utility (Appendix F.15; the influence paragraph below).

Model-free and validation-free families (Table 18; each in its best-case form; LAVA/SAVA receive the clean randomized validation slice their assumption asks for). *LAVA* (Just et al., 2023) is anti-correlated with the oracle: its label-distance term reads the exposure skew itself as distribution shift and down-values observed clicks wholesale. *SAVA* (Kessler et al., 2025) reproduces LAVA on our data (Spearman 0.999/0.967), so it is tabulated and not discussed separately. *Volume* (Xu et al., 2021) (label- and validation-free) is the cleanest control. It has tail share at the oracle level with a ranking nearly orthogonal to the oracle, so the tail correction of SHADE (oracle share together with $\tau = 0.547$) is not a diversity effect. *TracIn* (Pruthi et al., 2020) is not competitive, but it yields one positive result: swapping the target of the validation gradient for the $\mathbf{v}^{\mathrm{DR}}$ pseudo-outcome moves τ from -0.133 to -0.030 (synthetic) and -0.148 to $+0.033$ (KuaiRec), so the correction composes directionally across valuation families. *Data-IQ* (Seedat et al., 2022) lands in the neighborhood of the denoising family ($\tau = 0.225$, below naive-exposed 0.359). Across all four families the pattern holds: no external baseline combines oracle-level tail share with oracle-ranking recovery.

Table 18: **External valuation baselines on the MNAR benchmark** (synthetic λ=2, N=3,000, 8 seeds, mean±std). The τ of each method is against the oracle KNN-Shapley of its own player set, never compared across sets. Bold: best per column within each block.

method	family	τ vs. own-set oracle	tail share
All-N players (vs. all-N oracle; tail share → 0.215)			
naive	on-log value	0.205 ± 0.009	0.002 ± 0.004
LAVA	OT	-0.075 ± 0.016	0.180 ± 0.048
SAVA	OT (batched)	-0.070 ± 0.013	0.177 ± 0.046
Volume	leverage	0.035 ± 0.020	$\mathbf{0.209 \pm 0.009}$
SHADE	AIPW value	$\mathbf{0.547 \pm 0.020}$	0.225 ± 0.061
Exposed-only players (vs. exposed-only oracle; tail share → 0.087)			
naive (exp.)	on-log value	$\mathbf{0.359 \pm 0.020}$	$\mathbf{0.086 \pm 0.019}$
TracIn	grad. trajectory	-0.133 ± 0.013	0.072 ± 0.009
TracIn ($\mathbf{v}^{\mathrm{DR}}$ target)	grad. trajectory	-0.030 ± 0.027	0.052 ± 0.028
Data-IQ	training dynamics	0.225 ± 0.034	0.044 ± 0.018

Protocols. All four families share the synthetic MNAR setting of the suite (λ=2, N=3,000, Q=300, K=10, clip 0.05, 8 seeds), Coat (5 seeds), and, for TracIn/Data-IQ/Volume, the KuaiRec bern channel (N=30k, 5 seeds). Every method uses the standard $[z_u, w_i, z_u w_i, \rho]$ feature vectors, and each variant is scored against the oracle of its own player set. Per-family settings follow the published defaults of each method. LAVA/SAVA use log-domain Sinkhorn, tested against the exact OT LP, with SAVA batched at 1,024. Volume uses a documented per-point leverage form of the published per-*set* score. TracIn is TracInCP over 20 logistic-SGD checkpoints, validated on a held-out logged slice. Data-IQ

uses GBM confidence over 100 staged checkpoints, as published. TracIn ($\mathbf{v}^{\mathrm{DR}}$ target) differs from TracIn only in targeting the $\mathbf{v}_q^{\mathrm{DR}}$ pseudo-outcome $\hat{m}_q + e_q(\ell_q - \hat{m}_q)/\hat{\pi}_q$.

Several diagnostics explain the behavior of individual baselines. Zeroing the label-distance term of LAVA moves its τ from -0.08 toward -0.03, so the anti-correlation of Section 6.4 is carried by the label cost reading the exposure skew as distribution shift. The $\sim$82k all-N pool of Coat exceeds a dense OT cost matrix, so each seed uses a shared 6,000-pair subsample. The batching of SAVA changes nothing (ablation flat, τ -0.065 to -0.070 from K_t=12 to 3; an N=20,000 run keeps -0.070 ± 0.014). The validation gradient of TracIn targets the logged clicks, and its τ stays ≈ 0 even with oracle validation targets, which makes this a family property and not a validation-bias artifact.

The KuaiRec numbers behind Section 6.4 and the Coat tail shares are in Table 19. On KuaiRec the external baselines all stay below the exposed-only references, naive-exposed at 0.385 and exposed-only SHADE at 0.359. Data-IQ is the highest of them, Volume comes next because its leverage picks up the popularity feature ρ, TracIn is negative, and TracIn with the $\mathbf{v}^{\mathrm{DR}}$ target turns positive on average but with high seed variance. The Coat tail shares (all-N block) come from two constructions, each internally consistent: the full pool with a single deterministic split, on which Volume is run, and the shared 5-seed 6,000-pair subsample, on which LAVA/SAVA are run. In each construction, the external baseline has the largest tail share, above SHADE and well above naive; LAVA/SAVA reach theirs at Gini 0.106 ± 0.005. The nominally-best Coat tail share of LAVA/SAVA comes with the flattest item-value Gini of the methods compared on that subsample (the 0.106 above; Gini is not tabulated in Table 18). It reflects indifference to labels, not tail discovery.

Table 19: External baselines on real data (mean $\pm$ std over 5 seeds where a std is shown). KuaiRec bern channel, N=30k: Kendall τ of each method against the oracle of its own player set; the naive and SHADE entries are the exposed-only variants. Coat: all-N tail value share on the full pool (single deterministic split) and on the shared 6,000-pair subsample. Methods not run in a construction are marked "–".

	KuaiRec	Coat tail share	
method	τ	full pool	subsample
Volume	0.116 ± 0.003	0.195	–
Data-IQ	0.146 ± 0.010	–	–
TracIn	-0.148 ± 0.010	–	–
TracIn ($\mathbf{v}^{\mathrm{DR}}$ target)	$+0.033 \pm 0.146$	–	–
LAVA/SAVA	–	–	0.196 ± 0.03
naive	0.385	0.016	0.081 ± 0.027
SHADE	0.359	0.126	0.159 ± 0.028

Influence-function baseline (validation channel). Influence functions permit a sharper test, because the exposure correction can be placed in the evaluation utility while the per-point training signal is left uncorrected. Influence values (up-weighting form) are computed on an L2-logistic surrogate fit to the exposed points (5 seeds; synthetic $N = 3{,}000$, KuaiRec $N = 30$k): $\mathrm{IF}(i) = g_{\mathrm{val}}^{\top} H^{-1} g_i$ with closed-form Hessian, validated against leave-one-out retraining ($\rho > 0.97$, sign agreement > 0.9). Validation is a held-out logged slice (≈ 300 exposed points), scored with raw CE and, deliberately favorably, with IPS-weighted CE under true propensities, so the residual tail under-allocation quoted above cannot be attributed to propensity estimation. Influence values under-allocate the tail even more than naive (KuaiRec tail share 0.027 vs. naive 0.074, oracle 0.238). Exposure bias enters through the exposed-only training gradients, which validation reweighting cannot reach, and an unexposed point has no gradient at all (the Data-OOB/denoising limitation).

F.22 Sensitivity to unmeasured exposure–relevance confounding

If exposure and relevance share an unobserved cause after conditioning on $\mathbf{x}$, neither IPS nor SHADE is unbiased in general. To separate nuisance misspecification from failure of identification, this experiment deliberately violates Assumption 1(ii). On the main synthetic benchmark (λ=2, N=3,000, Q=300, K=10, ε=0.05, 5-fold cross-fitting, 8 seeds; all settings unchanged), each pair also draws a latent $h_i \sim \mathcal{N}(0,1)$ that shifts both true logits,

$$\mathrm{logit}(r_i) \mathrel{+}= \gamma\, h_i, \qquad \mathrm{logit}(\pi_i) \mathrel{+}= \gamma\, h_i,$$

for $\gamma \in \{0, 0.25, 0.5, 1\}$. The latent h_i is withheld from $\mathbf{x}_i$ entirely: it is seen by neither nuisance learner nor the KNN representation, so h is an unobserved common cause of exposure and relevance. The exposure intercept is recalibrated at each γ, so the marginal exposure rate stays approximately fixed, and exposure density is not the confound. The generator is unit-tested in the released suite. At γ=0 it reproduces the identified main benchmark exactly, because h is drawn from a separate random stream. Across the sweep, the realized mean exposure stays within 0.003 of that of the benchmark (0.347–0.350 vs. 0.350).

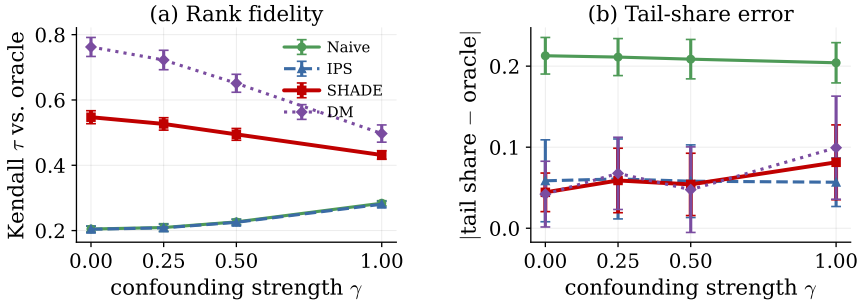

Figure 7: **Sensitivity to unmeasured confounding** (8 seeds, ±std). A latent $h \sim \mathcal{N}(0, 1)$ enters both $\text{logit}(r)$ and $\text{logit}(\pi)$ at strength γ but is unobserved by both nuisance models and the KNN features; the exposure intercept is recalibrated per γ. *(a)* Kendall τ vs. the oracle valuation; *(b)* absolute error of the tail-quintile value share against that of the oracle.

Rank fidelity of every corrected method degrades monotonically in γ (Figure 7; endpoints in Table 20). The τ of SHADE falls from 0.547 to 0.431 across the sweep, and the DM ablation falls fastest. Naive and IPS rise instead: under confounding, exposure itself becomes informative about relevance, which helps the uncorrected ranking. The SHADE−naive τ gap therefore shrinks from +0.342 to +0.148, and it stays significant at every γ. The tail-share error of SHADE is non-monotone at intermediate γ and roughly doubles overall, as it increasingly overshoots an oracle share that itself drifts downward. At γ=1 it is no longer distinguishable from plain IPS on this metric, and on item-value Gini error IPS is closer to the oracle than SHADE at $\gamma \geq 0.5$. The tail under-allocation of naive is unaffected (≈0.21 error throughout); it is structural, not confounding-driven. The mechanism is visible in the nuisance diagnostics: the mean errors $|\hat{m} - r|$ and $|\hat{\pi} - \pi|$ both grow with γ, since the h-component of both true functions is unlearnable from $\mathbf{x}$.

Table 20: Endpoints of the unmeasured-confounding sweep (main synthetic benchmark, 8 seeds; mean, with ± std except for the τ gap, whose ± is a 95% CI). At the intermediate strengths γ=0.25 and γ=0.5 the τ of SHADE is 0.527 and 0.495. The SHADE−naive gap is significant at every γ (paired t-test, $p \leq 3.7\times10^{-8}$). Bold: highest τ among the observation-responsive methods; as in Table 1, the DM ablation is listed for reference and can exceed it (Remark F.1).

quantity	γ=0	γ=1
Kendall τ, SHADE	**0.547 ± 0.020**	**0.431 ± 0.013**
Kendall τ, DM ablation	0.762	0.497
Kendall τ, naive	0.205	0.284
Kendall τ, IPS	0.204	0.281
SHADE−naive τ gap	+0.342 ± 0.014	+0.148 ± 0.014
tail-share error, SHADE	0.044 ± 0.024	0.081 ± 0.046
tail-share error, SHADE−IPS	–	+0.025 ± 0.040 (p=0.19)
oracle tail share	0.215	0.204
mean $\lvert\hat{m} - r\rvert$	0.081	0.169
mean $\lvert\hat{\pi} - \pi\rvert$	0.030	0.104

The experiment is a sensitivity analysis, not an extension of the double-robust guarantee. Once $y \perp e \mid \mathbf{x}$ fails, Proposition 5.1 no longer identifies the oracle value, and SHADE is biased here. Empirically, SHADE degrades as the latent confounding strengthens while the uncorrected methods

improve, so the corrected-vs-naive gap narrows without closing on this grid. Any relative advantage that remains at nonzero γ is an empirical property of this stress test, not an unbiasedness result.

F.23 Sensitivity to the candidate player universe

Shapley values are defined relative to a player set, so changing the candidate universe changes the cooperative game itself. We therefore do not test invariance of absolute values across universes, and none is claimed. The question is whether the qualitative correction conclusion (naive versus SHADE agreement with the same-universe oracle) is stable to reasonable exposure-independent changes in the player universe. This is distinct from the frozen-universe condition of the natural-log experiments (Appendix F.11): the fixed-player condition concerns independence from realized exposure; it does not make the choice of fixed player universe immaterial.

The experiment uses the fully observed KuaiRec block with imposed MNAR exposure, as in the KuaiRec construction of Appendix F.7 (bern channel $\ell = e \cdot \mathrm{Bern}(r)$; λ=2, γ=0.7, mean exposure 0.3; cross-fitted 5-fold nuisances, ε=0.05 clipping, K=10). Per seed, 500 MCAR oracle query pairs are drawn and frozen, shared by every configuration of that seed and excluded from every universe. For each seed, each universe size $N \in \{10{,}000, 20{,}000, 30{,}000, 60{,}000\}$ and each of two exposure-independent sampling designs (uniform pair sampling and item-stratified sampling with ≈uniform per-item quotas), we first sample and freeze the candidate pairs (hashes in Appendix F.5) and then impose the same MNAR exposure realization for every valuation method. Oracle, naive, IPS, and SHADE are therefore compared within the identical cooperative game for that universe; 8 seeds, 64 frozen universes in total.

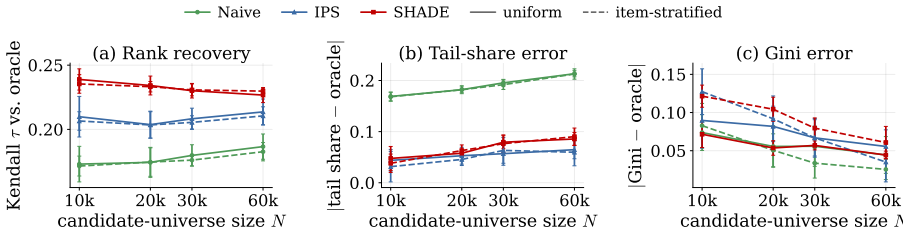

Figure 8: **Candidate-player-universe sensitivity on fully observed KuaiRec** (8 seeds, ±std; solid = uniform pair sampling, dashed = item-stratified). Candidates are sampled before imposed exposure, and all methods are evaluated against the oracle of the same frozen universe. Changing N changes the cooperative game; we therefore compare within-universe rank recovery *(a)* and tail-share/Gini errors *(b,c)*, not absolute Shapley values across universes.

The primary conclusion is stable in every tested universe (Figure 8). SHADE beats naive on Kendall τ vs. the same-universe oracle at every size and under both designs (SHADE 0.227–0.239 vs. naive 0.171–0.187; paired $\Delta\tau$ +0.040 to +0.066, $p \le 3\times10^{-5}$ in each configuration, and 64/64 seed-universe pairs), with IPS in between. SHADE also reduces the tail-share error by 0.115–0.130 in every configuration (naive 0.168–0.213 vs. SHADE 0.038–0.090; 64/64 pairs). Two sensitivities should be noted. First, magnitudes drift with N. The τ margin narrows modestly (+0.066 → +0.040 under uniform sampling) and the tail-share error of SHADE grows (≈0.04 → 0.09), an increasing overshoot of the same-universe oracle share; IPS attains the smallest tail-share error in most configurations. Second, the secondary Gini metric reverses. The item-value Gini of naive is nominally closer to that of the same-universe oracle than that of SHADE (significantly so under item-stratified sampling, $p \le 0.03$ at every N; statistically tied under uniform sampling, $p \ge 0.87$; SHADE closer in only 16/64 pairs). The concentrated value mass of naive happens to match the oracle Gini level of this block even while its tail share and ranking are badly wrong.

The main correction conclusion persists across the tested exposure-independent universes, but its magnitude is sensitive to the universe size, and one secondary metric (Gini error) is design-dependent and can favor naive. The τ of each universe measures within-universe ranking recovery, which is why values may be summarized across N; the absolute values themselves change with the cooperative game and are never compared across universes. Candidate-universe choice should therefore be treated as part of the valuation design, not as an innocuous implementation detail.