RISK-CONTROLLED PERSONALIZED POPULARITY ALIGNMENT FOR FROZEN COLLABORATIVE FILTERING

Anonymous authors
Paper under double-blind review

ABSTRACT

Popularity correction can improve agreement with user histories while reducing ranking utility. We study when to apply such edits to frozen recommenders under explicit utility-loss budgets. Our interface separates candidate construction, user ordering and risk authorization. Given the resulting candidate lists, independent calibration selects the largest score prefix whose positive-part harms for Recall and NDCG satisfy their respective budgets. Rejected edits return the original list exactly. Conformal risk control bounds marginal expected harm under exchangeability. We instantiate the interface with RCPA-CF, our risk-controlled popularity-alignment method, and compare CP and TTEN-style corrections. Evaluation covers three datasets, four backbones and five seeds, with disjoint calibration and evaluation users. Under the conservative RCPA setting, learned ordering reduces mean popularity calibration error (PCE) by 1.839%, versus 1.654% for gain-only and 1.626% for regret-only ordering. All three orderings meet both realized budgets in 56/60 runs; this is a batch diagnostic rather than a probability guarantee. Ordering gains concentrate in Gowalla BPR-MF and LightGCN. Historical calibration-pool replays show an additional benefit over a change-based heuristic in selected settings. Further analyses relate candidate choice and alignment metrics to retained benefit and user-level loss distributions. The interface thus enables selective popularity correction under explicit utility-loss budgets.

1 INTRODUCTION

Better popularity alignment does not always mean a more useful recommendation list. For a frozen recommender, the practical question is therefore *which users should receive a correction under a limited utility-loss budget?* This question arises because collaborative filtering inherits skewed exposure from observational data: popular items can dominate recommendations and reinforce feedback loops (Mansoury et al., 2020; Abdollahpouri et al., 2019b; Chen et al., 2023). Yet a uniform long-tail bonus can over-correct users who prefer popular items. Popularity calibration offers a personalized objective by aligning each list with its user's history (Steck, 2018; Abdollahpouri et al., 2021; Schirmer et al., 2026). Even so, better alignment can come at the expense of ranking utility, making selective application essential when losses must be limited.

Constructing a correction and deciding where to apply it are different problems. Two edits with similar alignment gains may have very different relevance costs, while the users with the largest gains need not incur the smallest losses. Applying the same correction strength to everyone does not resolve this allocation problem. Moreover, average accuracy can hide individual losses by offsetting them with other users' improvements. We therefore measure each user's positive-part utility loss before averaging and ask how much alignment benefit can be retained under explicit Recall and NDCG harm budgets. This formulation makes the cost of applying a correction part of the decision itself.

We introduce a modular risk-authorization interface that separates candidate construction, user ordering and calibration. A fixed correction generates an edited list for each user; a fixed rule or learned harm predictor then orders users by estimated harm relative to alignment benefit. Independent calibration selects the largest prefix whose corrected risk estimates satisfy both budgets, and users outside that prefix keep their original lists exactly. Our primary candidate construction is RCPA-CF, our risk-controlled popularity-alignment method, which adjusts popularity in either direction within the frozen candidate pool. The interface separates candidate construction from risk authorization, so

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.

CP- and TTEN-style corrections can also be evaluated under the same protocol to isolate candidate-construction effects from learned ordering.

The contribution is a decision framework for allocating popularity corrections under explicit utility-loss constraints. Conformal risk control supplies marginal expected-harm guarantees under exchangeability, while benefit–cost ordering determines which alignment benefits survive calibration. A fractional allocation oracle and a fixed-threshold stability analysis explain the ordering principle and its assumptions. This separation also guides evaluation: holding the candidate fixed isolates ordering effects, whereas holding the authorization rule fixed exposes differences between corrections. Because the guarantee concerns expected harm, we additionally report realized batch risks and losses among users who receive edits.

Across three datasets, four frozen backbones and five seeds, learned ordering retains a mean PCE reduction of 1.839%, compared with 1.654% for gain-only and 1.626% for regret-only ordering. All three orderings meet both realized budgets in 56/60 runs. The additional benefit concentrates in Gowalla BPR-MF and LightGCN, where it persists across repeated user partitions; simpler orders remain competitive elsewhere. Candidate comparisons and metric checks further separate the benefit of constructing an edit from the benefit of deciding who receives it. Together, these analyses identify where learned allocation is useful within the existing-user protocol, without requiring changes to the frozen recommender.

2 Related Work

Personalized popularity calibration. Popularity calibration aligns recommendation composition with user histories rather than uniformly promoting unpopular items (Steck, 2018; Abdollahpouri et al., 2021; Yavru et al., 2026). This user-centered view recognizes that the appropriate popularity mix varies across users. Recent sequential approaches combine popularity quantile calibration with activation steering (Schirmer et al., 2026), extending calibration beyond list re-ranking. Across these approaches, historical preferences provide a reference for alignment, while ranking accuracy remains a separate evaluation objective.

Training-time and post-hoc correction. Popularity interventions can act during model training or after scores are frozen. Training-time methods revise objectives or representations (Zhang et al., 2021; Ning et al., 2024; Cai et al., 2024; Liu et al., 2023; 2026); others address graph propagation and popularity shifts (Wang et al., 2024; Zhang et al., 2023a;b; Islam et al., 2026a). Post-hoc approaches instead modify scores or lists after training: personalized re-ranking and list-level calibration adjust item order directly (Abdollahpouri et al., 2019a; 2021; Gulsoy et al., 2025), whereas PPD perturbs embeddings (Islam et al., 2026b) and TTEN attenuates item norms (Kim et al., 2023). These interventions construct different candidate edits; our frozen-backbone interface instead controls who receives an edit, allowing candidate scope and output order to be analyzed separately.

Selective prediction and risk control. Selective prediction trades coverage for conditional risk by allowing a predictor to abstain (El-Yaniv & Wiener, 2010; Geifman & El-Yaniv, 2017; Franc et al., 2023). Conformal risk control instead calibrates a fixed family using held-out outcomes (Bates et al., 2021; Angelopoulos et al., 2024; Overman et al., 2024). For recommendation, Angelopoulos et al. (2023) calibrate pre-trained rankers and optimize diversity under false-discovery-rate constraints. Group-conditional extensions further refine the population covered by a risk bound (Martinez Gil et al., 2024). These lines of work distinguish the quality of a selection score from the validity of calibration, while the chosen loss and conditioning population determine the meaning of a risk bound.

Exposure objectives and evaluation. Popularity alignment is related to exposure fairness, but the objectives differ: exposure-based approaches study how attention is distributed across items or providers (Singh & Joachims, 2018; Biega et al., 2018; Klimashevskaia et al., 2024). Work on popularity bias further emphasizes that users and other stakeholders can favor different outcomes (Boratto et al., 2021; Kowald et al., 2020; Chen et al., 2023). Since historical profiles reflect both preferences and prior exposure, agreement with those profiles captures only one aspect of recommendation quality. Ranking utility and exposure measures are complementary; user satisfaction and creator welfare require direct evidence beyond implicit-feedback rankings (Klimashevskaia et al., 2023).

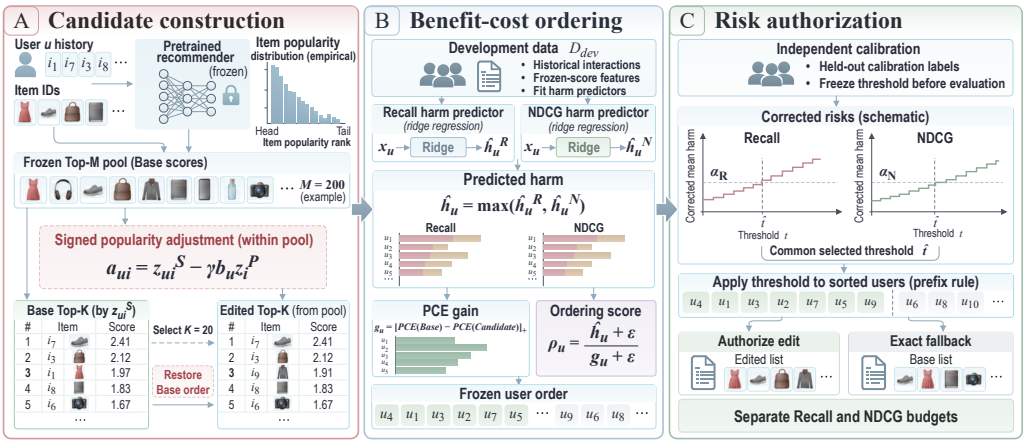

Figure 1: Candidate edits and risk authorization.

3 Modular Risk Authorization

Figure 1 summarizes the three stages of the interface. RCPA first constructs a signed edit within the frozen Top-M candidate pool, after which ordering and calibration determine whether to apply it. Alternative candidates share the same endpoint losses and calibration rule, with all development outputs, fitted models and tie rules fixed before calibration and evaluation.

3.1 Problem Formulation

Let $\mathcal{E}_{\mathrm{tr}} \subseteq \mathcal{U} \times \mathcal{I}$ be the training graph, $\mathcal{H}_u$ user u's history, and $d_i = |\{u : (u,i) \in \mathcal{E}_{\mathrm{tr}}\}|$ item popularity. A frozen backbone B provides unseen-item scores, a Top-M candidate set $\mathcal{C}_u^M$, and the ordered Base Top-K list π_u^B; every edit remains inside $\mathcal{C}_u^M$. Each target observation is $Z_u = (X_u, Y_u)$. The inputs X_u contain only deployment-observable history and frozen scores. Held-out relevance outcomes Y_u are used to evaluate utility. The main experiment uses disjoint calibration and evaluation subsets of existing users on all three datasets. Their training histories and learned ID representations are retained. Section 4.1 specifies the common partition rule.

Following Schirmer et al. (Schirmer et al., 2026), define $\mathcal{T} = \{0, 0.1, \ldots, 1\}$ and let $q_{u,\tau}(\pi)$ be the empirical τ-quantile of item degrees in list π. With $\widehat{F}_u$ the empirical CDF of degrees in $\mathcal{H}_u$, user-level Popularity Calibration Error is

$$\mathrm{PCE}_u(\pi) = |\mathcal{T}|^{-1} \sum_{\tau \in \mathcal{T}} \big(\tau - \widehat{F}_u(q_{u,\tau}(\pi))\big)^2. \tag{1}$$

As a quantile–CDF score, PCE can remain positive when discrete degree distributions coincide, because ties affect the CDF values. A lower score therefore need not imply a smaller distributional distance. Appendix R gives a counterexample and a same-candidate metric audit; Appendix Q examines corrections on a common raw-score baseline. Coverage, tail exposure, mean log-popularity, and Gini remain complementary diagnostics. The empirical inverse-CDF and tie conventions are specified in Appendix C.

Let q_R and q_N denote per-user Recall@20 and NDCG@20, respectively, and write $[v]_+ = \max(v, 0)$. For $m \in \{R, N\}$, define one-sided same-backbone harm

$$\ell_m(u; \pi) = [q_m(u, \pi_u^B) - q_m(u, \pi_u)]_+ \in [0, 1]. \tag{2}$$

Thus the endpoint bounds are $B_R = B_N = 1$. The objective is to reduce mean PCE subject to $\mathbb{E}[\ell_m(U; \pi)] \le \alpha_m$ for both endpoints. This is a population-average claim whose expectation is over the calibration sample and one fresh protocol-induced target unit under the stated exchangeability contract. Section 3.4 specifies this averaging unit. The budget $\alpha_m = 0.001$ is 0.1 absolute percentage points on the $[0, 1]$ utility scale, not 0.1% relative to Base. Averaging positive-part losses prevents utility gains for one user from canceling another user's harm; the development net-drop criterion instead averages signed losses.

3.2 BIDIRECTIONAL CANDIDATE CORRECTION

Let $d^B_{u,(0.5)}$ be the median item degree in π^B_u and define the signed Base mismatch $b_u = \widehat{F}_u(d^B_{u,(0.5)}) - 0.5$. The median gives a simple signed direction without fitting a calibration curve. An 11-quantile alternative has dataset-dependent benefits (Appendix F). Positive b_u requests movement toward the tail and negative b_u toward the head. We standardize $\log(1 + d_i)$ over the complete item catalog to obtain z^p_i, and standardize the finite Base scores separately inside each user's Top-M pool to obtain z^s_{ui}; both use population variance. A zero catalog popularity variance is treated as an error, while a zero per-user relevance scale is floored at the score dtype's positive epsilon. We then score candidates by

$$a_{ui}(\gamma) = z^s_{ui} - \gamma b_u z^p_i, \qquad i \in \mathcal{C}^M_u. \tag{3}$$

For a nontrivial edit, the policy selects the K largest adjusted scores inside $\mathcal{C}^M_u$ and restores Base-score order within the selected set. This confines the edit to membership and popularity composition; position-sensitive effects are reported in Appendix F. Equal-score and bypass cases are handled explicitly (Appendix C). Standardization gives a common dimensionless coefficient but does not equalize margins or edit rates across backbones. Proof details are in Appendix A.

3.3 BENEFIT–COST ORDERING

Let X denote deployment-observable information, and let $H_m \in [0, 1]$ be the candidate edit's one-sided harm for endpoint $m \in \mathcal{M} = \{R, N\}$, and

$$\mu_m(x) = \mathbb{E}[H_m \mid X = x], \qquad g(x) = [\mathrm{PCE}(\pi^B_x) - \mathrm{PCE}(\pi^\gamma_x)]_+. \tag{4}$$

Here $g_u = g(X_u)$ is the positive part of the signed, label-free PCE difference $\Delta_u = \mathrm{PCE}_u(\pi^B_u) - \mathrm{PCE}_u(\pi^\gamma_u)$. The oracle counts positive gains; reported PCE changes retain the sign of Δ_u. The fractional multi-budget oracle, with measurable authorization function $a = a(X) \in [0, 1]$, is

$$V_{\mathrm{multi}} = \sup_{0 \le a \le 1} \mathbb{E}[ag] \quad \text{s.t.} \quad \mathbb{E}[a\mu_m] \le \alpha_m \quad (m \in \mathcal{M}). \tag{5}$$

The exact oracle uses dual-weighted conditional harms; its LP characterization is given in Appendix A.1. The budget-normalized envelope is

$$V_{\mathrm{env}} = \sup_{0 \le a \le 1} \mathbb{E}[ag] \quad \text{s.t.} \quad \mathbb{E}\left[a \max_m \{\mu_m/\alpha_m\}\right] \le 1, \tag{6}$$

and lies between one-half and the full two-endpoint oracle value. This compares fractional policies with true conditional harms and boundary randomization; it does not bound the benefit delivered by the fitted score with its separately calibrated threshold. The oracle is therefore a reference for allocation, not a claim that the learned order is optimal.

We use ridge as a compact, reproducible ordering model with 14 fixed features, rather than selecting a model class on calibration outcomes. Its role is to allocate benefit; the risk guarantee also applies to simpler frozen orders. We fit development-only predictors $\widehat{h}^R_u, \widehat{h}^N_u \in [0, 1]$, set $\widehat{h}_u = \max_m \widehat{h}^m_u$, and order lower scores first using

$$\bar{\rho}_u = \begin{cases} (\widehat{h}_u + \varepsilon)/(g_u + \varepsilon), & g_u > 0, \\ C + \widehat{h}_u, & g_u \le 0, \end{cases} \qquad \rho_u = \bar{\rho}_u + 10^{-9} r^{\mathrm{regret}}_u, \tag{7}$$

where $\varepsilon = 10^{-8}$ and $C = 10^6$ are fixed numerical constants. The stabilizer regularizes small gains; the sentinel assigns a finite score to non-positive gains. The regret $r^{\mathrm{regret}}_u = [\bar{s}^B_u - \bar{s}^\gamma_u]_+/\sigma_u$ uses mean frozen scores of the two lists and the floored within-pool score scale σ_u from Section 3.2. The additive term is part of the score, rather than a lexicographic tie rule. Because positive-branch ratios can exceed this finite sentinel, the sentinel does not force every positive-gain edit ahead of every non-positive-gain edit. The primary experiments have $\alpha_R = \alpha_N = 10^{-3}$ and use this fixed scoring rule. The unequal-budget sensitivity analysis uses $\bar{\rho}^{(\alpha)}_u = \max_m\{(\widehat{h}^m_u + \varepsilon_h)/\alpha_m\}/(g_u + \varepsilon_g)$ on the positive-gain branch. Its non-positive-gain sentinel is a finite parameter-only bound above every positive-branch score, independent of the scoring batch. The six-budget sensitivity analysis, including both unequal-budget pairs, is reported in Appendix G. The equal-budget branch retains

the scale in Eq. (7). Recalibrated sensitivity checks find unchanged endpoint harms across the tested constants and tie rules (Appendix O.1). The features summarize user history, candidate alignment, score margin, list changes and frozen-model regret (Table 6). Features, coefficients, γ, budgets, and score rules are fixed before calibration labels are read. A threshold authorizes the complete finite-score atom $\{u : \rho_u \le t\}$ and never randomizes within a boundary tie. Candidate Top-K ties are a separate implementation issue discussed in Appendix C.

One further option is to retain only authorized edits with strictly positive signed PCE gain. This weakly improves PCE alignment without increasing either harm, so the original risk bound still holds. Appendix R proves the property and reports filtered development replays separately from the predeclared primary policy.

What learning changes. Learning affects how much benefit a calibrated score prefix retains. At a fixed threshold, the conditional stability result in Appendix A.1 confines disagreements to a score boundary band under uniform prediction-error and positive-gain assumptions. The empirical comparison evaluates the separately calibrated case where each ordering receives its own calibration threshold. The fixed-threshold bound does not control that threshold difference.

3.4 Multi-Endpoint Calibration and Exact Fallback

After the user order is frozen, an independent split calibrates the largest budget-feasible prefix. For $Z = (X, Y)$ define

$$H_m(Z) = [q_m(Z, \pi^B) - q_m(Z, \pi^\gamma)]_+ \in [0, B_m], \quad B_R = B_N = 1, \tag{8}$$

$$A_t(X) = \mathbb{I}[\rho(X) \le t], \qquad L_m(Z; t) = A_t(X) H_m(Z). \tag{9}$$

Untreated units incur zero loss relative to the same backbone.

Assumption 1. *Conditional on development outputs, calibration units $Z_1, \dots, Z_n$ and a fresh target unit are exchangeable. The candidate policy, complete score, tie rule and budgets are fixed before calibration labels are read. The budgets satisfy $\alpha_m \ge B_m/(n+1) = 1/(n+1)$.*

Even the empty prefix carries this $1/(n+1)$ finite-sample correction floor; a smaller requested budget is infeasible for the corrected-prefix search. The implementation then returns Base for every user, with exactly zero harm, rather than claiming that this fallback satisfies the corrected empirical inequality.

Using the distinct calibration scores together with an empty-prefix sentinel, compute

$$\widehat{R}^+_m(t) = \frac{\sum_{i=1}^n L_m(Z_i; t) + B_m}{n+1}, \tag{10}$$

$$\widehat{t}_m = \max\{t : \widehat{R}^+_m(t) \le \alpha_m\}, \qquad \widehat{t} = \min_m \widehat{t}_m. \tag{11}$$

Nestedness makes corrected risk non-decreasing, so $\widehat{t}$ is the largest common-score prefix satisfying all corrected budgets on the realized sample.

Theorem 1. *Under Assumption 1, for every endpoint m,*

$$\mathbb{E}[L_m(Z_{n+1}; \widehat{t})] \le \alpha_m. \tag{12}$$

The expectation is over calibration and one fresh target unit, conditional on the frozen development outputs; the result holds for any finite frozen score.

This result specializes multi-risk CRC (Angelopoulos et al., 2024, Section 4.3) to marginal expected harm over calibration and one target unit. Corrected calibration constraints hold samplewise, while evaluation-batch means can fluctuate; both are reported. Fixed group partitions also permit group-specific thresholds when within-group laws remain stable (Appendix A).

In implementation, each backbone retains its native scores and seen-item mask. The wrapper adds no trainable parameters and requires one user sort and at most $n+1$ calibration-prefix scans per batch; adapter identity and tie checks are detailed in Appendices B and C.

4 EXPERIMENTS

4.1 EXPERIMENTAL SETUP AND MAIN RESULTS

We evaluate Gowalla, Yelp2018 and Amazon-Book with a common existing-user protocol. Each dataset uses four frozen backbones—BPR-MF, LightGCN, XSimGCL and LightCCF—and five seeds (2020–2024), following standard collaborative-filtering evaluation settings (Wang et al., 2019; He et al., 2020; Rendle et al., 2009; Yu et al., 2024; Zhang et al., 2025). Training/validation interactions are unchanged. A fixed random user partition (seed 2027) divides each held-out interaction pool into disjoint calibration/evaluation groups: 14,929/14,929, 15,834/15,834 and 26,322/26,321 users, respectively. Both stages exclude training and validation items. Candidate-specific predictors are fitted on validation data. Calibration thresholds are saved before evaluation labels are loaded. All settings use $M = 200$, $K = 20$ and $\alpha_R = \alpha_N = .001$, with predeclared primary $\gamma = .0175$ and sensitivity $\gamma = .1$. We treat these budgets as explicit deployment choices: 10^{-3} corresponds to 0.1 absolute percentage points of population-average positive-part harm for each endpoint. The sensitivity grid in Appendix G shows how the tradeoff changes when practitioners choose a different tolerance. The expected-harm bound requires Assumption 1. Appendix O details the unified evaluation and Appendix C specifies the implementation and data roles.

Table 1 gives all 12 primary settings under learned harm/gain ordering. Base R/N reports mean Recall@20/NDCG@20 on the $[0, 1]$ scale. Base PCE, paired absolute reductions and positive-part harms use 10^{-3} units (budget 1). PCE reduction is a percentage; $\pm$ is the five-seed SD on one fixed split. Within counts seeds meeting both realized budgets; bold marks an above-budget mean.

Table 1: Unified-protocol primary results.

Backbone	Base R/N	PCE B/Δ	PCE red.	H_R	H_N	Within
Gowalla						
BPR-MF	0.1614/0.0938	58.739/1.973	3.358±0.327	0.741±0.361	0.305±0.113	3/5
LightGCN	0.1962/0.1175	125.949/3.178	2.523±0.181	**1.020±0.215**	0.637±0.105	3/5
XSimGCL	0.2032/0.1242	103.037/0.786	0.763±0.030	0.222±0.074	0.068±0.020	5/5
LightCCF	0.1947/0.1187	53.128/0.705	1.327±0.041	0.396±0.101	0.118±0.031	5/5
Yelp2018						
BPR-MF	0.0619/0.0385	50.173/1.921	3.829±0.111	0.554±0.033	0.524±0.031	5/5
LightGCN	0.0907/0.0576	121.532/1.877	1.547±0.105	0.479±0.096	0.316±0.025	5/5
XSimGCL	0.1055/0.0682	111.554/0.511	0.458±0.022	0.120±0.046	0.044±0.014	5/5
LightCCF	0.0944/0.0610	107.062/0.975	0.903±0.082	0.182±0.054	0.135±0.059	5/5
Amazon-Book						
BPR-MF	0.0775/0.0491	82.319/3.482	4.230±0.055	0.674±0.057	0.874±0.037	5/5
LightGCN	0.0960/0.0620	148.083/2.369	1.600±0.024	0.489±0.082	0.424±0.032	5/5
XSimGCL	0.1106/0.0738	133.873/0.723	0.540±0.027	0.146±0.047	0.054±0.014	5/5
LightCCF	0.1113/0.0754	60.925/0.607	0.996±0.041	0.193±0.069	0.070±0.023	5/5

At the predeclared strength $\gamma = .0175$, mean PCE reduction is 1.839%, versus 1.654% for gain-only and 1.626% for regret-only. All three orders meet both realized budgets in 56/60 runs; this is a batch diagnostic, not a probability guarantee. The ungated primary candidate gives 1.939% reduction but meets both budgets in 50/60 runs. The four budget exceedances under learned ordering are Gowalla BPR-MF seeds 2022/2023 and LightGCN seeds 2021/2022, all on Recall. Their Recall harms are $1.093, 1.174, 1.362, 1.024 \times 10^{-3}$, respectively, against a budget of 10^{-3}. At $\gamma = .1$, learned ordering gives 5.727% reduction and 43/60 satisfaction, versus 4.802% and 46/60 for gain-only: more alignment benefit does not imply higher realized budget satisfaction.

4.2 ORDERING UNDER REPEATED USER PARTITIONS

The repeated-partition ordering analysis uses the calibration pools specified in Appendix D. Six fixed orders share Gowalla candidates and 100 partitions per seed: 5,000 calibration users select each complete-score-group threshold; the remaining 9,929 users evaluate it. Users recur across partitions, so this within-pool resampling is dependent. Table 2 reports the six orders and learned-minus-regret PCE reduction (percent and percentage points, respectively). Means $\pm$ SD summarize five seed means, each averaging 100 shared partitions. Harms use 10^{-3} units; the Admit/change column reports

learned authorization/list-change percentages, and Within is the percentage of batches meeting both budgets.

Table 2: Shared-pool ordering on Gowalla.

Order	BPR-MF	LightGCN	XSimGCL	LightCCF
Random	1.537±0.229	1.184±0.110	0.735±0.035	1.360±0.038
PCE gain	2.003±0.311	1.051±0.222	0.736±0.035	1.369±0.032
Regret	1.748±0.155	1.219±0.125	0.731±0.035	1.355±0.037
Predicted harm	2.828±0.240	2.059±0.149	0.734±0.035	1.359±0.036
Predicted harm/gain	3.233±0.254	2.352±0.096	0.736±0.035	1.366±0.033
Regret/gain	2.415±0.200	1.632±0.137	0.735±0.035	1.366±0.033

Backbone	Paired gain	Learned H_R/H_N	Admit/change	Within
BPR-MF	+1.485±0.303	0.870/0.360	10.13/10.13	67.6
LightGCN	+1.133±0.171	0.877/0.522	13.51/13.35	69.4
XSimGCL	+0.005±0.001	0.302/0.089	99.98/6.64	100.0
LightCCF	+0.011±0.008	0.317/0.103	99.97/6.58	100.0

Learned ordering improves over regret-only in all five seed means for BPR-MF/LightGCN (97.6%/99.4% of paired partitions). Differences on XSimGCL/LightCCF are negligible, and learned ordering is not best among all six orders for these backbones. Mean harms and batchwise satisfaction remain distinct (Table 2; full results in Appendix D). To interpret this allocation benefit, Appendix K compares it with an outcome-aware reference. This empirical diagnostic is separate from the ordering-stability theorem, which assumes a uniform conditional-mean prediction-error bound.

4.3 Alternative candidate corrections

Table 3 evaluates fixed Top-200 candidate pools on LightGCN under the unified protocol, with Base-score order restored after selection. The candidate corrections include CP (Abdollahpouri et al., 2021) and TTEN (Kim et al., 2023); G/Y/A denote the datasets. PCE reduction is percent under gain-only; learned gain is the paired learned-minus-gain-only difference in percentage points. Harms use 10^{-3} units, Changed is the percentage of altered Top-20 memberships, and $\pm$ is five-seed SD. The weakest positive CP/TTEN strengths are controls fixed before evaluation.

Table 3: Unified-protocol candidate comparisons.

	Candidate	PCE red.	H_R/H_N	Changed	Learned gain	Learned H_R/H_N
G	RCPA	1.237±0.364	1.271/0.836	2.65	+1.286±0.384	1.020/0.637
G	TTEN .01	0.951±0.023	0.573/0.531	13.33	-0.006±0.011	0.572/0.527
G	TTEN .05	4.006±0.401	1.012/0.372	19.54	-0.006±0.160	0.984/0.370
G	TTEN .25	4.325±0.918	0.815/0.312	5.30	+1.461±0.582	0.911/0.392
G	CP .0005	4.041±0.555	0.950/0.782	9.62	+0.245±0.166	0.965/0.704
Y	RCPA	1.550±0.099	0.477/0.314	14.33	-0.004±0.007	0.479/0.316
Y	TTEN .01	1.088±0.070	0.334/0.222	13.29	-0.002±0.011	0.333/0.220
Y	TTEN .05	4.922±0.392	0.971/0.385	29.48	+0.058±0.056	1.008/0.394
Y	TTEN .25	5.742±0.422	0.866/0.337	8.06	+0.624±0.089	0.845/0.378
Y	CP .0005	5.287±0.454	0.977/0.606	19.50	+0.052±0.082	0.959/0.570
A	RCPA	1.602±0.025	0.490/0.425	17.53	-0.002±0.002	0.489/0.424
A	TTEN .01	0.975±0.036	0.363/0.350	15.46	-0.007±0.012	0.363/0.351
A	TTEN .05	4.940±0.214	1.000/0.464	37.56	+0.013±0.063	1.002/0.452
A	TTEN .25	7.290±0.422	1.085/0.445	11.67	+0.723±0.294	0.979/0.456
A	CP .0005	6.329±0.192	1.041/0.837	19.07	+0.160±0.115	1.051/0.747

The comparison grid contains Base and 34 non-Base candidates: two RCPA strengths, 22 CP weight/order variants and ten Top-M TTEN exponent/order variants. Each has a separate validation-only 14-feature ridge fit (penalty 10), and each order is separately calibrated on the same users. All 137 outcomes per run are retained. The TTEN-style comparison is restricted to the common Top-M pool. Realized risks and changed-list rates vary at equal nominal budgets, so candidate comparisons report all three outcomes alongside PCE.

Because the primary protocol freezes backbone scores, these comparisons test the interface's separation of candidate construction from authorization; training-time or embedding-level methods are treated as transfer diagnostics rather than primary candidate comparisons.

Supplementary common-pool comparisons cover global/native alternatives, 100 shared partitions and construction ablations; their grid and protocol are specified in Appendix N. Their complete-grid audit finds four Gowalla CP variants improving on RCPA under gain-only ordering without increasing either harm or changed-list rate in all five seed means (Appendix H). These comparisons describe the recorded candidate implementations. Finite-precision raw/sigmoid ranking differences can affect the TTEN baseline; Appendix Q gives a common-raw-score comparison. Restoration preserves the selected membership while retaining the backbone ranking within the selected set; the appendix quantifies its position-sensitive alignment tradeoff. Construction ablations appear in Appendix F. A further common-raw, 24-candidate grid retains conservative RCPA as a non-dominated mean tradeoff at all nine dataset–budget combinations; stronger RCPA is dominated in two (Appendix Q).

4.4 Ordering gains, robustness and cost

Figure 2 compares the primary learned order with both simple orders on the unified protocol. The left panel displays the two material Gowalla gains on the percentage-point scale, with each paired seed shown individually. The right panel magnifies the ten near-zero settings by 10^3 and summarizes their five paired seeds with mean $\pm$ SD; the zero line and signs are unchanged. G/Y/A denote the datasets and MF/GCN/XSim/CCF the backbones. Method settings and comparison strengths are fixed before evaluation.

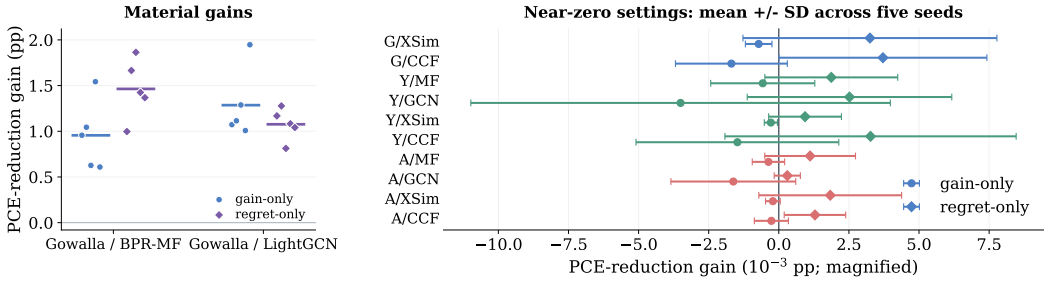

Figure 2: Paired ordering gains; near-zero settings use a magnified scale.

The six-budget resampling analysis is reported in Appendices G and H. Paired seed intervals appear in Appendix O. Historical group-mixture experiments reduce observed budget exceedance with group-specific thresholds while retaining much of the pooled benefit (Appendix M). They vary group proportions with within-group laws fixed; temporal drift remains outside this experiment. Calibration-pool diagnostics report authorization rates, harm incidence and positive-loss quantiles for edited users (Appendix J). The partition audit below uses the unified Gowalla user pool.

Metric and predictor checks. As a distributional cross-check, historical fixed RCPA lists show learned/gain-only W_1 reductions of .9454/.5723% on Gowalla, .8933/.8939% on Yelp2018 and 1.1415/1.1398% on Amazon-Book (Appendix Q). This cross-check preserves the Gowalla ordering advantage under a distribution distance, while the other differences remain small. These and the common-score W_1 ordering control (Appendix S) reuse historical pools. Predictor checks reveal a related distinction: although change fraction is a strong harm-classification signal, learned ordering retains more PCE benefit in Gowalla MF/LightGCN calibration-pool replay (Table 30). Thus classification AUROC alone does not determine the benefit retained after separate calibration. The fixed-policy partition audit in Figure 3 uses five specified splits of the same Gowalla user pool, four backbones and five model seeds. Each split has 14,929 users per stage. The original split is included; repeated users make the 100 cells dependent. All thresholds are calibrated separately with frozen predictors. Each dot is one cell; black ticks are equal-cell means. The dashed line is the population Recall budget. Right-panel means condition on authorized changes to Top-20 membership. Both realized budgets are met in 79/100 cells: BPR-MF 14/25, LightGCN 15/25, and XSimGCL/LightCCF 25/25 each. All 21 exceedances concern Recall; NDCG meets its budget in every cell. All corrected

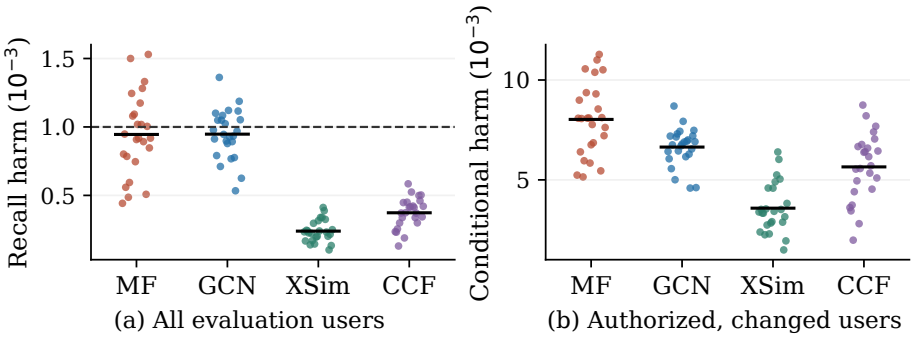

Figure 3: Gowalla risk across five user partitions.

calibration risks meet both budgets. BPR-MF/LightGCN select policies close to the calibration boundary, with mean corrected Recall harms $.976/.981 \times 10^{-3}$. The backbones with repeated exceedances have little calibration slack. Appendix P examines this association.

Computationally, the predictor occupies 464 bytes in float64, while the complete wrapper costs .04494–.04632 ms/user and at most 2.8 MiB extra peak allocation. Its cached-Base batch overhead is 88.58–94.69%; Appendices N and I separate fitting from wrapper costs. Applicability also depends on the available edits: the KuaiRec diagnostic changes only .36% of lists and its PCE interval crosses zero. PAAC/DR-GNN and PPD transfer diagnostics likewise expose protocol-dependent limitations (Appendix L). Since corrections remain inside Top-M, retrieval quality bounds the available edits, while the formal risk interpretation depends on the calibration assumptions.

Table 4: Ordering value across Gowalla partitions.

	Gain-only		Regret-only		Learned	
Backbone	PCE red.	Within	PCE red.	Within	PCE red.	Within
BPR-MF	2.544	15/25	1.969	12/25	3.554	14/25
LightGCN	0.879	14/25	1.532	14/25	2.483	15/25
XSimGCL	0.757	25/25	0.754	25/25	0.756	25/25
LightCCF	1.336	25/25	1.331	25/25	1.335	25/25

Table 4 reports the separately calibrated simple orders retained in the same audit. PCE reduction is percent, averaged over 25 cells per backbone; Within counts cells meeting both realized budgets. Learned ordering gains 1.010/1.604 PCE-reduction points over gain-only on BPR-MF/LightGCN. Budget satisfaction changes from 15/25 to 14/25 and from 14/25 to 15/25, respectively. Its benefit is improved allocation of alignment gains under selective authorization, with no consistent advantage in batchwise budget satisfaction. On XSimGCL/LightCCF, gain-only retains a small advantage.

Variation across settings. Primary learned-order gains concentrate in Gowalla BPR-MF and LightGCN. The other ten learned-minus-gain-only means range from $-.003503$ to $-.000211$ PCE-reduction points. Candidate strength introduces a separate tradeoff: CP and TTEN can provide larger gains, with different harms and change rates (Table 3). Deployment choices therefore require development-stage selection followed by calibration of the frozen policy.

5 Conclusion

We presented a modular risk-authorization interface, instantiated with RCPA-CF, for allocating popularity corrections under explicit utility-loss budgets. Separating candidate construction, ordering, and calibration distinguishes correction quality from allocation benefit. The conservative RCPA policy yields 1.839% mean PCE reduction and 56/60 realized dual-budget satisfaction under the common existing-user protocol. Learned ordering adds benefit in Gowalla BPR-MF and LightGCN but is not universally superior; historical replays and complementary diagnostics delimit when it helps. Under frozen-policy and exchangeability assumptions, CRC provides marginal expected-harm control for selective correction.

ETHICS STATEMENT

This work studies exposure and popularity alignment in public, anonymized implicit-feedback datasets. Popularity alignment and demographic fairness are distinct objectives. The datasets do not provide demographic attributes for evaluating group protection. Deployment can still alter creator and item exposure; accordingly, we report catalog and user-alignment metrics together with ranking harm and exact fallback behavior. No human-subject data were newly collected for this study.

REPRODUCIBILITY STATEMENT

The appendices specify data partitions, frozen candidate construction, calibration rules, proof assumptions and complete seed-level comparisons. To make comparisons reproducible, the budget-conditioned audit fixes user partitions across methods and budgets, preserves original score identities and reports every setting. Cached-input replay requires no model fitting, while separate direct-sum checks verify selected prefixes and target metrics. File hashes identify the data, fitted models and outputs for each evaluation protocol. The supplementary materials provide the corresponding numerical summaries and audit artifacts for inspecting and replaying the reported comparisons. They are intended for audit replay rather than full retraining.

AI USE STATEMENT

Generative AI tools assisted with literature organization, methodological reasoning and proof revision. The authors are responsible for the accuracy, originality, and integrity of the final manuscript, code, and supplementary material, including all material produced with AI assistance.

REFERENCES

Himan Abdollahpouri, Robin Burke, and Bamshad Mobasher. Managing popularity bias in recommender systems with personalized re-ranking. In *Proceedings of the 32nd International Florida Artificial Intelligence Research Society Conference (FLAIRS)*, pp. 413–418. AAAI Press, 2019a.

Himan Abdollahpouri, Masoud Mansoury, Robin Burke, and Bamshad Mobasher. The unfairness of popularity bias in recommendation. In *Proceedings of the Workshop on Recommendation in Multi-Stakeholder Environments at the 13th ACM Conference on Recommender Systems (RMSE at RecSys)*, 2019b.

Himan Abdollahpouri, Masoud Mansoury, Robin Burke, Bamshad Mobasher, and Edward C. Malthouse. User-centered evaluation of popularity bias in recommender systems. In *Proceedings of the 29th ACM Conference on User Modeling, Adaptation and Personalization (UMAP)*, pp. 119–129, 2021.

Anastasios N. Angelopoulos, Karl Krauth, Stephen Bates, Yixin Wang, and Michael I. Jordan. Recommendation systems with distribution-free reliability guarantees. In *Proceedings of the Twelfth Symposium on Conformal and Probabilistic Prediction with Applications (COPA)*, volume 204 of *Proceedings of Machine Learning Research*, pp. 175–193, 2023.

Anastasios N. Angelopoulos, Stephen Bates, Adam Fisch, Lihua Lei, and Tal Schuster. Conformal risk control. In *Proceedings of the Twelfth International Conference on Learning Representations (ICLR)*, 2024.

Stephen Bates, Anastasios N. Angelopoulos, Lihua Lei, Jitendra Malik, and Michael I. Jordan. Distribution-free, risk-controlling prediction sets. *Journal of the ACM*, 68(6):43:1–43:34, 2021.

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)*, pp. 405–414, 2018.

Ludovico Boratto, Gianni Fenu, and Mirko Marras. Connecting user and item perspectives in popularity debiasing for collaborative recommendation. *Information Processing & Management*, 58(1):102387, 2021.

Miaomiao Cai, Lei Chen, Yifan Wang, Haoyue Bai, Peijie Sun, Le Wu, Min Zhang, and Meng Wang. Popularity-aware alignment and contrast for mitigating popularity bias. In *Proceedings of the 30th ACM SIGKDD Conference on Knowledge Discovery and Data Mining (KDD)*, pp. 187–198, 2024.

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):67:1–67:39, 2023.

Ran El-Yaniv and Yair Wiener. On the foundations of noise-free selective classification. *Journal of Machine Learning Research*, 11(53):1605–1641, 2010.

Vojtěch Franc, Daniel Průša, and Václav Voráček. Optimal strategies for reject option classifiers. *Journal of Machine Learning Research*, 24(11):1–49, 2023.

Yonatan Geifman and Ran El-Yaniv. Selective classification for deep neural networks. In *Proceedings of the 31st Conference on Neural Information Processing Systems (NeurIPS)*, volume 30, 2017.

Mert Gulsoy, Emre Yalcin, Yucel Tacli, and Alper Bilge. DUoR: Dynamic user-oriented re-ranking calibration strategy for popularity bias treatment of recommendation algorithms. *International Journal of Human-Computer Studies*, 203:103578, 2025.

Xiangnan He, Kuan Deng, Xiang Wang, Yan Li, Yongdong Zhang, and Meng Wang. LightGCN: Simplifying and powering graph convolution network for recommendation. In *Proceedings of the 43rd International ACM SIGIR Conference on Research and Development in Information Retrieval (SIGIR)*, pp. 639–648, 2020.

Md Aminul Islam, Ahmed Sayeed Faruk, Sourav Medya, and Elena Zheleva. Debiasing message passing to mitigate popularity bias in GNN-based collaborative filtering. *arXiv preprint arXiv:2605.11145*, 2026a.

Md Aminul Islam, Elena Zheleva, and Ren Wang. Post-hoc popularity bias correction in GNN-based collaborative filtering. In *Proceedings of the ACM Web Conference 2026 (WWW)*, pp. 6818–6829, 2026b.

Dain Kim, Jinhyeok Park, and Dongwoo Kim. Test-time embedding normalization for popularity bias mitigation. In *Proceedings of the 32nd ACM International Conference on Information and Knowledge Management (CIKM)*, pp. 4023–4027, 2023.

Anastasiia Klimashevskaia, Mehdi Elahi, Dietmar Jannach, Lars Skjærven, Astrid Tessem, and Christoph Trattner. Evaluating the effects of calibrated popularity bias mitigation: A field study. In *Proceedings of the 17th ACM Conference on Recommender Systems (RecSys)*, pp. 1084–1089, 2023.

Anastasiia Klimashevskaia, Dietmar Jannach, Mehdi Elahi, and Christoph Trattner. A survey on popularity bias in recommender systems. *User Modeling and User-Adapted Interaction*, 34(5): 1777–1834, 2024.

Dominik Kowald, Markus Schedl, and Elisabeth Lex. The unfairness of popularity bias in music recommendation: A reproducibility study. In *Proceedings of the 42nd European Conference on Information Retrieval (ECIR)*, volume 12036 of *Lecture Notes in Computer Science*, pp. 35–42, 2020.

Lingfeng Liu, Yixin Song, Dazhong Shen, Bing Yin, Hao Li, Yanyong Zhang, and Chao Wang. Rethinking popularity bias in collaborative filtering via analytical vector decomposition. In *Proceedings of the 32nd ACM SIGKDD Conference on Knowledge Discovery and Data Mining V.1 (KDD)*, pp. 879–890, 2026.

Yuanhao Liu, Qi Cao, Huawei Shen, Yunfan Wu, Shuchang Tao, and Xueqi Cheng. Popularity debiasing from exposure to interaction in collaborative filtering. In *Proceedings of the 46th International ACM SIGIR Conference on Research and Development in Information Retrieval (SIGIR)*, pp. 1801–1805, 2023.

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)*, pp. 2145–2148, 2020.

Natalia Martinez Gil, Dhaval Patel, Chandra Reddy, Giri Ganapavarapu, Roman Vaculin, and Jayant Kalagnanam. Identifying homogeneous and interpretable groups for conformal prediction. In *Proceedings of the Fortieth Conference on Uncertainty in Artificial Intelligence (UAI)*, volume 244 of *Proceedings of Machine Learning Research*, pp. 2471–2485, 2024.

Wentao Ning, Reynold Cheng, Xiao Yan, Ben Kao, Nan Huo, Nur Al Hasan Haldar, and Bo Tang. Debiasing recommendation with personal popularity. In *Proceedings of the ACM Web Conference 2024 (WWW)*, pp. 3400–3409, 2024.

William Overman, Jacqueline Jil Vallon, and Mohsen Bayati. Aligning model properties via conformal risk control. In *Proceedings of the 38th Conference on Neural Information Processing Systems (NeurIPS)*, volume 37, 2024.

Steffen Rendle, Christoph Freudenthaler, Zeno Gantner, and Lars Schmidt-Thieme. BPR: Bayesian personalized ranking from implicit feedback. In *Proceedings of the Twenty-Fifth Conference on Uncertainty in Artificial Intelligence (UAI)*, pp. 452–461. AUAI Press, 2009.

Mona Schirmer, Anton Thielmann, Pola Schwöbel, Thomas Martynec, Giuseppe Di Benedetto, Ben London, and Yannik Stein. Aligning recommendations with user popularity preferences. In *Proceedings of the 2026 ACM Conference on Fairness, Accountability, and Transparency (FAccT)*, pp. 4773–4793, 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)*, pp. 2219–2228, 2018.

Harald Steck. Calibrated recommendations. In *Proceedings of the 12th ACM Conference on Recommender Systems (RecSys)*, pp. 154–162, 2018.

Bohao Wang, Jiawei Chen, Changdong Li, Sheng Zhou, Qihao Shi, Yang Gao, Yan Feng, Chun Chen, and Can Wang. Distributionally robust graph-based recommendation system. In *Proceedings of the ACM Web Conference 2024 (WWW)*, pp. 3777–3788, 2024.

Xiang Wang, Xiangnan He, Meng Wang, Fuli Feng, and Tat-Seng Chua. Neural graph collaborative filtering. In *Proceedings of the 42nd International ACM SIGIR Conference on Research and Development in Information Retrieval (SIGIR)*, pp. 165–174, 2019.

Ismail Burak Yavru, Emre Yalcin, Alper Bilge, and Sevcan Yilmaz Gunduz. SUPER: Smart user-centric popularity exposure reduction for fair and diverse recommendations. *IEEE Access*, 14: 37582–37605, 2026.

Junliang Yu, Xin Xia, Tong Chen, Lizhen Cui, Nguyen Quoc Viet Hung, and Hongzhi Yin. XSimGCL: Towards extremely simple graph contrastive learning for recommendation. *IEEE Transactions on Knowledge and Data Engineering*, 36(2):913–926, 2024.

An Zhang, Jingnan Zheng, Xiang Wang, Yancheng Yuan, and Tat-Seng Chua. Invariant collaborative filtering to popularity distribution shift. In *Proceedings of the ACM Web Conference 2023 (WWW)*, pp. 1240–1251, 2023a.

Yang Zhang, Fuli Feng, Xiangnan He, Tianxin Wei, Chonggang Song, Guohui Ling, and Yongdong Zhang. Causal intervention for leveraging popularity bias in recommendation. In *Proceedings of the 44th International ACM SIGIR Conference on Research and Development in Information Retrieval (SIGIR)*, pp. 11–20, 2021.

Yifei Zhang, Hao Zhu, Yankai Chen, Zixing Song, Piotr Koniusz, and Irwin King. Mitigating the popularity bias of graph collaborative filtering: A dimensional collapse perspective. In *Proceedings of the 37th Conference on Neural Information Processing Systems (NeurIPS)*, volume 36, 2023b.

Yu Zhang, Yiwen Zhang, Yi Zhang, Lei Sang, and Yun Yang. Unveiling contrastive learning's capability of neighborhood aggregation for collaborative filtering. In *Proceedings of the 48th International ACM SIGIR Conference on Research and Development in Information Retrieval (SIGIR)*, pp. 1985–1994, 2025.

A Proofs and scope of the theoretical statements

A.1 Allocation statements and learned-order stability

The following results distinguish outcome-aware fractional optimization from the deployed learned order. Definitions of the benefit, conditional harms, and allocation objectives are given in Section 3.3.

Theorem 2 (Exact multi-budget oracle)**.** *Whenever Eq.* (5) *has primal and dual optima with no duality gap, some $\lambda_m^* \geq 0$ yield an optimal rule*

$$a^*_{\mathrm{multi}}(x) = \mathbb{I}[g(x) > \textstyle\sum_m \lambda_m^* \mu_m(x)] + \xi(x)\mathbb{I}[g(x) = \textstyle\sum_m \lambda_m^* \mu_m(x)], \tag{13}$$

where $\xi \in [0,1]$ allocates only boundary mass.

Theorem 3 (Conservative max-envelope allocation)**.** *For non-negative integrable g, c_α, an optimum of Eq.* (6) *thresholds increasing $r(x) = c_\alpha(x)/g(x)$, with boundary randomization if needed:*

$$a^*(x) = \mathbb{I}[g(x) > \lambda^* c_\alpha(x)] + \xi \mathbb{I}[g(x) = \lambda^* c_\alpha(x)]. \tag{14}$$

Moreover,

$$|\mathcal{M}|^{-1} V_{\mathrm{multi}} \leq V_{\mathrm{env}} \leq V_{\mathrm{multi}}; \tag{15}$$

for two endpoints the envelope retains at least half of the exact fractional oracle benefit.

Theorem 4 (Boundary-localized robustness of the learned order)**.** *On the positive-gain branch let $c(x) = \max_m \mu_m(x)$ and $r(x) = c(x)/g(x)$. Suppose $0 < g_{\min} \leq g \leq g_{\max}$, $0 \leq c \leq c_{\max}$, and $|\widehat{h} - c| \leq \delta$. For the full score $\rho = \bar{\rho} + e$, assume $|e| \leq \tau$ and define*

$$\eta = \frac{\delta}{g_{\min} + \varepsilon} + \frac{\varepsilon(g_{\max} + c_{\max})}{g_{\min}(g_{\min} + \varepsilon)} + \tau. \tag{16}$$

Then $|\rho - r| \leq \eta$ and, for every threshold t,

$$\{r \leq t - \eta\} \subseteq \{\rho \leq t\} \subseteq \{r \leq t + \eta\}. \tag{17}$$

The learned and oracle threshold policies therefore differ in benefit by at most

$$\mathbb{E}[g(X)\mathbb{I}(|r(X) - t| \leq \eta)]. \tag{18}$$

If $\Pr(|r(X) - t| \leq s) \leq D s^\beta$ for a range containing $s = \eta$, the gap is at most $g_{\max} D \eta^\beta$.

For unequal budgets, the sandwich argument uses the normalized envelope $c_\alpha = \max_m \mu_m/\alpha_m$ and $\bar{\rho}^{(\alpha)}$. The error radius then includes separate contributions from ε_h and ε_g. The common scale cancels for equal budgets. The specified perturbation has $\tau = 10^{-9} r_{\max}^{\mathrm{regret}}$ if regret is bounded by $r_{\max}^{\mathrm{regret}}$ on the target domain. This result assumes uniform prediction error and the stated margin condition. Empirical harm correlations alone are insufficient to verify these assumptions. The theorem compares policies at a common fixed threshold. For different fixed thresholds t_L, t_O, the disagreement band instead has radius $\eta + |t_L - t_O|$ around t_O, by the same triangle inequality. Separately calibrated thresholds require an additional bound on $|t_L - t_O|$ for comparison with the oracle. The risk-control theorem is independent of these assumptions and applies to any frozen finite score.

A.2 Risk interpretation and fixed-group extension

For the joint calibration–target draw, define the scalar treatment probability $p = \Pr(A_{\widehat{t}}(X) = 1) > 0$. Then

$$\mathbb{E}[L_m] = p\mathbb{E}[H_m \mid A_{\widehat{t}} = 1], \tag{19}$$

$$\Pr(A_{\widehat{t}} = 1, H_m > \zeta) \leq \alpha_m/\zeta. \tag{20}$$

For an observable group partition $G = \phi(X)$ frozen independently of groupwise calibration labels, let $\widehat{t}_{g,m}$ be the endpoint cutoff calibrated within group g and deploy

$$\widehat{t}_g = \min_m \widehat{t}_{g,m}, \qquad A_{\mathrm{vec}}(X) = \mathbb{I}\{\rho(X) \le \widehat{t}_{\phi(X)}\}. \tag{21}$$

Theorem 5 in Appendix A gives endpoint-wise control under any change in group proportions when within-group laws remain fixed. An implementation can use independent label-free pilot users to freeze ϕ and disjoint calibration users for risk fitting.

Theorem 5 (Robustness to fixed-group mixture proportions)**.** *Condition on a partition ϕ frozen without groupwise calibration labels. If calibration and target units are exchangeable within every group g and $\alpha_m \ge B_m/(n_g+1)$, then*

$$\mathbb{E}[L_m(Z_g; \widehat{t}_g)] \le \alpha_m \quad \textit{for every } g, m. \tag{22}$$

Therefore any target that changes only group weights ω while preserving within-group laws satisfies

$$\sum_g \omega_g \mathbb{E}[L_m(Z_g; \widehat{t}_g)] \le \alpha_m. \tag{23}$$

The partition and within-group distributions must remain fixed.

A.3 Structural details of the candidate correction

Differentiating Eq. (3) with respect to z_i^p gives $-\gamma b_u$, establishing the signed response. For a nontrivial edit, candidate containment holds because selection is performed only over $\mathcal{C}_u^M$. At $\gamma = 0$ or $M = K$, the implementation returns the independently computed Base list verbatim. This bypass preserves exact identity even with score ties. A separate Top-K operation could select different tied items despite monotone standardization.

For completeness, suppose a size-K set S omits candidate i but contains candidate j with $a_{ui} > a_{uj}$. Replacing j by i strictly increases $\sum_{k\in S} a_{uk}$. Repeating the exchange yields the K largest adjusted scores. The final Base-score sort changes only the permutation and preserves strict Base-score inequalities. Equal-score pairs may change their relative order. These observations are construction checks rather than a statistical utility guarantee.

A.4 Proof of Theorem 2

For a finite population with masses p_j, Eq. (5) is a linear program in treatment probabilities $a_j \in [0,1]$. Its Lagrangian for multipliers $\lambda_m \ge 0$ is

$$\sum_m \lambda_m \alpha_m + \sum_j p_j a_j \left(g_j - \sum_m \lambda_m \mu_{m,j} \right). \tag{24}$$

Linear-programming strong duality supplies optimal multipliers λ^*. Pointwise maximization over $a_j \in [0,1]$ sets $a_j = 1$ when $g_j > \sum_m \lambda_m^* \mu_{m,j}$, sets $a_j = 0$ under the reverse strict inequality, and permits any value on equality. Complementary slackness and boundary allocation choose the equality values for the active budgets. This is Eq. (13). The same argument applies to any distribution for which the stated primal–dual regularity holds.

A.5 Proof of Theorem 3

Write $c = c_\alpha$. If $\mathbb{E}[c(X)\mathbb{I}(g(X) > 0)] \le 1$, treating every positive-gain unit is feasible and optimal because $g \ge 0$. Consider the nontrivial case in the theorem. As λ increases, the cost of the strict superlevel set $\{g > \lambda c\}$ is non-increasing. Choosing a boundary value λ^* and, when needed, randomizing on $\{g = \lambda^* c\}$ therefore gives

$$\mathbb{E}[a^*(X)c(X)] = 1. \tag{25}$$

Formally, take the quantile of g/c under the finite cost-weighted measure $c(x)dP(x)$. Include positive-gain points with $c = 0$ first, since they consume no budget. A common randomization probability on the boundary fills any atom exactly.

For any feasible rule a, pointwise maximization of the Lagrangian integrand gives

$$a(x)(g(x) - \lambda^* c(x)) \le a^*(x)(g(x) - \lambda^* c(x)). \tag{26}$$

Taking expectations, using feasibility of a, and then Eq. (25),

$$\mathbb{E}[ag] \le \mathbb{E}[a(g - \lambda^* c)] + \lambda^* \tag{27}$$
$$\le \mathbb{E}[a^*(g - \lambda^* c)] + \lambda^* = \mathbb{E}[a^* g]. \tag{28}$$

Thus a^* solves Eq. (6). Dividing $g > \lambda^* c$ by positive g yields the increasing $r = c/g$ order. Only an atom at equality requires randomization.

It remains to prove Eq. (15). Every envelope-feasible rule is multi-budget feasible because $\mu_m/\alpha_m \le c_\alpha$ pointwise, so $V_{\text{env}} \le V_{\text{multi}}$. Conversely, for any multi-budget-feasible a,

$$\mathbb{E}\left[\frac{a(X)}{|\mathcal{M}|} c_\alpha(X)\right] \le \frac{1}{|\mathcal{M}|} \sum_{m \in \mathcal{M}} \frac{\mathbb{E}[a(X)\mu_m(X)]}{\alpha_m} \le 1. \tag{29}$$

Thus $a/|\mathcal{M}|$ is envelope feasible and retains $1/|\mathcal{M}|$ of its benefit. Taking the supremum over multi-budget feasible rules gives the lower bound.

A.6 Proof of Theorem 4

On the positive-gain branch, first separate prediction and stabilization error:

$$\left|\frac{\widehat{h} + \varepsilon}{g + \varepsilon} - \frac{c}{g}\right| \le \frac{|\widehat{h} - c|}{g + \varepsilon} + \left|\frac{c + \varepsilon}{g + \varepsilon} - \frac{c}{g}\right| \tag{30}$$
$$= \frac{|\widehat{h} - c|}{g + \varepsilon} + \frac{\varepsilon |g - c|}{g(g + \varepsilon)} \tag{31}$$
$$\le \frac{\delta}{g_{\min} + \varepsilon} + \frac{\varepsilon(g_{\max} + c_{\max})}{g_{\min}(g_{\min} + \varepsilon)} = \eta - \tau. \tag{32}$$

Adding $|e| \le \tau$ gives $|\rho - r| \le \eta$. If $r(x) \le t - \eta$, the uniform bound implies $\rho(x) \le t$; if $\rho(x) \le t$, it implies $r(x) \le t + \eta$. This proves Eq. (17). The symmetric difference of the learned and oracle treatment sets is consequently contained in $\{|r - t| \le \eta\}$. Multiplying its indicator by non-negative g and taking expectations proves Eq. (18). The margin-rate statement follows from $g \le g_{\max}$. For fixed t_L, t_O, $|(\rho - t_L) - (r - t_O)| \le \eta + |t_L - t_O|$ proves the stated different-threshold extension. No bound on $|t_L - t_O|$ is asserted.

A.7 Proof of Theorem 1

We give the argument in the score-prefix orientation implemented by the gate. Equal numerical scores enter as one group. For endpoint m, first ignore the other endpoints and let $\widehat{t}_m$ be the calibration threshold from Eq. (11). Its corrected feasibility means

$$\frac{\sum_{i=1}^n L_m(Z_i; \widehat{t}_m) + B_m}{n + 1} \le \alpha_m. \tag{33}$$

After adjoining Z_{n+1}, the loss at the same threshold is at most B_m. Therefore Eq. (33) implies

$$\frac{1}{n+1} \sum_{i=1}^{n+1} L_m(Z_i; \widehat{t}_m) \le \alpha_m. \tag{34}$$

Now group all $n + 1$ units by score. Let t_m° denote the largest prefix whose full-sample mean loss for endpoint m is at most α_m. The policy induced by $\widehat{t}_m$ on these $n + 1$ units is a feasible prefix by Eq. (34); hence it cannot extend beyond t_m°. Monotonicity gives

$$L_m(Z_{n+1}; \widehat{t}_m) \le L_m(Z_{n+1}; t_m^\circ). \tag{35}$$

The full-sample selector t_m° is invariant to permutations of the $n+1$ exchangeable units. Conditional on their unordered multiset, each unit is equally likely to occupy position $n+1$. Consequently,

$$\mathbb{E}[L_m(Z_{n+1}; t_m^\circ) \mid \{Z_1, \ldots, Z_{n+1}\}] = \frac{1}{n+1} \sum_{i=1}^{n+1} L_m(Z_i; t_m^\circ) \tag{36}$$

$$\leq \alpha_m. \tag{37}$$

Taking expectations and using Eq. (35) proves $\mathbb{E}[L_m(Z_{n+1}; \widehat{t}_m)] \leq \alpha_m$. This is the finite-score specialization of the CRC exchangeability argument (Angelopoulos et al., 2024).

Finally, the deployed threshold satisfies $\widehat{t} \leq \widehat{t}_m$ for every endpoint. Pointwise monotonicity therefore gives $L_m(Z; \widehat{t}) \leq L_m(Z; \widehat{t}_m)$ for every Z, proving all inequalities in Eq. (12). Pointwise domination gives each expectation bound directly, without a multiplicity adjustment.

A.8 Proof of Theorem 5

Fix the pilot information and hence the partition ϕ. For a particular group g and endpoint m, apply Theorem 1 to the n_g exchangeable calibration units in that group. The resulting threshold $\widehat{t}_{g,m}$ satisfies

$$\mathbb{E}[L_m(Z_{g,n_g+1}; \widehat{t}_{g,m})] \leq \alpha_m. \tag{38}$$

By construction, $\widehat{t}_g \leq \widehat{t}_{g,m}$. Nested treatment sets and non-negative harms therefore imply the pointwise domination

$$L_m(z; \widehat{t}_g) \leq L_m(z; \widehat{t}_{g,m}) \tag{39}$$

for every target unit z. Taking expectations proves Eq. (22). If the target changes only the group weights, the law of total expectation gives its risk as $\sum_g \omega_g \mathbb{E}[L_m(Z_g; \widehat{t}_g)]$. Every term is at most α_m and the non-negative weights sum to one, which proves Eq. (23). The proof uses no relationship between different groups, but it does require each within-group target law to match its calibrated law.

A.9 Score agnosticism and interpretation

Each corrected empirical risk is non-decreasing along the common score order because a longer prefix adds only non-negative harms. Endpoint m is therefore feasible exactly through $\widehat{t}_m$, so their intersection is feasible exactly through $\min_m \widehat{t}_m$. This proves samplewise maximality.

The proof of Theorem 1 never compares $\widehat{h}_u$ with $H_m(u)$ and never assumes that the score order is optimal. It requires a fixed policy, nested authorization sets, bounded losses and exchangeable units. The guarantee therefore holds for any fixed score. A poor score can still produce low benefit or authorize very few edits.

For the random calibrated threshold, define the scalar $p = \Pr(A_{\widehat{t}}(X) = 1) > 0$ under the joint calibration–target draw, conditional on frozen development outputs. It is not the treatment rate conditional on one realized threshold. The law of total expectation and exact fallback give

$$\mathbb{E}[L_m(Z; \widehat{t})] = p\mathbb{E}[H_m(Z) \mid A_{\widehat{t}}(X) = 1], \tag{40}$$

so an average-risk bound implies only $\mathbb{E}[H_m \mid A_{\widehat{t}} = 1] \leq \alpha_m/p$, not a conditional bound of α_m. Likewise, Markov's inequality yields Eq. (20); it does not turn the result into a per-user certificate or a calibration-conditional high-probability guarantee.

Evaluation population. All three main datasets use disjoint calibration and evaluation groups of existing users, with training histories and backbone representations held fixed. The theorem averages over calibration and a target unit under Assumption 1. The same partition unit is used across datasets, and exchangeability is required conditional on fixed development outputs. In contrast, the supplementary resampling studies characterize variation conditional on their specified calibration pools.

Table 5: Native scorer contracts.

Backbone	Frozen inference score	Family
BPR-MF	$\sigma(\mathbf{p}_u^\top \mathbf{q}_i)$	factorization
LightGCN	$\sigma((\mathbf{e}_u^*)^\top \mathbf{e}_i^*)$	graph propagation
XSimGCL	clean graph-view embedding dot product	graph contrastive
LightCCF	$\sigma((\mathbf{e}_u^*)^\top \mathbf{e}_i^*)$	neighborhood contrast

B BACKBONE-SPECIFIC IMPLEMENTATION

Table 5 lists the native scoring functions. Each comparison uses the same backbone and seed.

Each adapter loads the native checkpoint and computes item scores with the model's own inference procedure. It then excludes seen items using the stage-specific mask. Here σ is the sigmoid; XSimGCL uses its clean unperturbed dot product without this activation. Although sigmoid is monotone, its magnitudes affect the subsequent relevance standardization, so these activation conventions are part of the frozen numerical interface. A regression test compares the native and wrapped score vectors and Top-20/Top-40 outputs at zero strength. Base and RCPA then reuse the same checkpoint and candidate pool. Each backbone retains its own scorer and numerical conventions throughout the paired comparisons.

C IMPLEMENTATION AND EVALUATION DETAILS

C.1 FIXED SETTINGS AND DATA ROLES

The primary correction strength is $\gamma = .0175$; $\gamma = .1$ is a separate sensitivity point. We use $K = 20$, $M = 200$, 11 popularity quantiles, ridge penalty 10 and positive-part Recall/NDCG budgets of .001. These settings were fixed after development analysis and are recorded with the fitted parameters in the model files. The subsequent stages have distinct roles: validation outcomes fit candidate-specific predictors, calibration outcomes determine thresholds, and evaluation outcomes measure the resulting policies.

Common evaluation. The main results use $K = 20$, five seeds, full-ranking evaluation with stage-specific seen-item masks, training-degree popularity and 11 quantile points for PCE. Supplementary resampling and metric sensitivity studies specify their sampling units and metric variants separately. The candidate size is $M = 200$. For each list quantile τ, the empirical inverse-CDF index is $\max\{\lceil \tau K \rceil - 1, 0\}$ after sorting recommendation degrees. The history CDF includes items tied at the queried degree. Every evaluated user has at least one training interaction. One-item histories use the same empirical CDF without interpolation. Empty histories are rejected.

History, masking and normalization. Training interactions alone define item degree, the history CDF and history features. Validation recommendations exclude training items. Both calibration and evaluation recommendations exclude training and validation items under the unified protocol. Calibration and evaluation users are disjoint, and both stages use the same availability rule. Excluded items are not added to the history used to compute PCE.

Log-degree is standardized across the full catalog using population standard deviation. A constant popularity vector raises an error. Relevance is standardized separately within each user's finite Top-M pool, with population variance and a scale floor at the score dtype's machine epsilon. These moments exclude non-finite scores. Fewer than 20 finite unseen candidates raises an error. These are distinct from the fit-fold feature normalization used by ridge.

Authorization and risk units. Authorization counts users admitted by the score threshold, whereas effective change counts users whose Top-20 membership changes. Since an authorized list can retain its original membership, the two rates need not coincide. Selective authorization and personalized item correction are thus distinct. The one-sided population loss includes zero for every fallback user. A budget of 10^{-3} permits a population-average loss of 0.1 absolute percentage points. Loss among treated users can exceed this population budget. Improvements for other users do not cancel this loss. Development net-drop checks average signed losses. Calibration constraints include the finite-sample

correction, while evaluation reports realized positive-part harm. The lower bound $1/(n+1)$ is specific to the corrected-prefix procedure; a deterministic all-Base policy retains zero harm even below that floor.

Loss bounds and small calibration samples. The bound $B_R = B_N = 1$ follows from the range of normalized Recall and NDCG. A smaller bound requires a valid support restriction for both calibration and deployment, specified in advance. Replacing B_m by an empirical maximum to improve feasibility does not preserve the stated theorem. With $B_m = 1$ and $\alpha_m = 10^{-3}$, the corrected-prefix rule requires at least 999 calibration units even for the empty prefix; authorizing useful edits may require more data. Below this floor, returning Base still ensures zero harm. In either case, application-specific utility tolerances require separate assessment.

Adaptive correction strengths. An extension could learn a deployment-observable mapping $\gamma(X)$ using only development information. Fix the complete candidate policy and score before independent calibration. Each user's candidate harm then remains fixed as the threshold varies, and $A_t = \mathbb{I}[\rho(X) \le t]$ remains nested. The score-agnostic argument can then apply under the same exchangeability and bounded-loss assumptions. The new policy requires calibration on its own losses. Choosing γ using calibration outcomes, or varying the candidate policy with t, is outside the stated frozen-policy proof. Its benefit and computational cost require empirical evaluation.

Candidate and gate ties. For $M = K$ or $\gamma = 0$, the implementation reuses the frozen ordered Base list verbatim. Otherwise PyTorch `topk` chooses the adjusted-score set and a second `topk` restores Base-score order. Software, hardware and deterministic settings are fixed. The order of exact candidate-score ties can differ across devices. Gate-score ties are distinct: threshold selection considers each exact finite-score atom as a whole and authorizes precisely $\rho_u \le t$, with no boundary randomization. We report the number of unique score atoms, the tied-user fraction, and the largest atom for every ordering.

Learned allocation implementation. Table 6 specifies each input. Ridge fitting standardizes features using only the fitting data. Standard deviations at most 10^{-12} are replaced by one. The intercept is unpenalized, and endpoint predictions are clipped to $[0, 1]$. Five-fold development diagnostics use fold seed 2026 and predictions fitted outside each holdout fold. The deployed predictor is instead the frozen full-development fit; calibration does not refit its coefficients. Penalty 10 is fixed, with $\{0, 1, 10, 100\}$ used only for the separate development sensitivity analysis. Table 6 reports the 14 predictor inputs, computed from training histories, frozen scores and candidate lists. Let P_B, P_C denote Base and corrected PCE. Their signed median CDF mismatches are b_B, b_C. Let c denote the changed-item fraction and r the nonnegative normalized Base-score regret. Finally, ℓ is the corrected-minus-Base difference in mean $\log(1 + d_i)$. No evaluation relevance labels enter these inputs.

Table 6: Predictor feature definitions.

Index	Input	Source
1	$\log(1 + \lvert\mathcal{H}_u\rvert)$	Training history
2	b_B	History and Base list
3	$\lvert b_B\rvert$	History and Base list
4	$c = 1 - \lvert B_u \cap C_u\rvert/20$	Candidate lists
5	$r = [(\bar{s}_B - \bar{s}_C)/\sigma_{u,M}]_+$	Frozen Top-M scores
6	P_B	History and Base list
7	P_C	History and corrected list
8	$P_B - P_C$	History and candidate lists
9	$[P_B - P_C]_+$	History and candidate lists
10	b_B (second implementation input)	History and Base list
11	b_C	History and corrected list
12	ℓ	Training degrees and lists
13	$\lvert\ell\rvert$	Training degrees and lists
14	rc	Frozen scores and lists

Inputs 2 and 10 are identical under the 11-quantile definition: both use the history CDF at the Base list's empirical median minus 0.5. The frozen specification retains both. They encode the same signal, with ridge regularization distributing weight between them. Across all 60 frozen fits, their centers and scales agree exactly, and the maximum absolute difference between their endpoint coefficients is 1.72×10^{-14}. PCE differences and their component values are also algebraically related. Thus individual coefficient magnitudes should not be read as feature importance.

Table 7 reports sign stability of standardized coefficients in the 60 frozen full-development predictors. Entries count settings (out of 12) with all five seed coefficients positive/negative; coefficients within 10^{-12} of zero count as neither. The coefficients describe conditional associations among correlated inputs.

Table 7: Predictor coefficient sign stability.

Input index	Recall: positive/negative	NDCG: positive/negative
1	0/5	10/0
2	1/0	2/0
3	0/2	1/2
4	12/0	12/0
5	6/1	4/3
6	1/1	2/1
7	0/2	1/2
8	3/1	3/1
9	1/1	1/2
10	1/0	2/0
11	0/1	0/2
12	2/0	1/1
13	1/4	0/4
14	4/4	3/4

Table 7 summarizes signs within each setting across five seeds. Complete coefficient values, centers, scales and per-setting means/SDs accompany the results. This audit uses the actual frozen full-development models, while the cross-fitted prediction-error analysis remains a distinct development diagnostic. Neither establishes a uniform error bound on conditional expected harm.

Each endpoint fit has 14 coefficients and an intercept: 30 fitted coefficients in total, plus 28 feature-normalization scalars. Their raw double-precision numerical state occupies 464 bytes, excluding software and serialization overhead. For $d = 14$ inputs and n fitting users, forming and solving the ridge systems costs $O(nd^2 + d^3)$; applying the two fitted linear models costs $O(d)$ per user after feature construction. These arithmetic/storage counts exclude list generation and feature construction and are not end-to-end latency measurements.

Backbone configurations. BPR-MF is implemented as zero propagation layers in the shared embedding codepath. LightGCN uses three propagation layers and 64-dimensional embeddings. XSimGCL uses three graph layers, contrastive layer 1, $\lambda = 0.2$, perturbation scale $\epsilon = 0.2$, and temperature 0.15. LightCCF uses a LightGCN encoder, three layers, 64-dimensional embeddings, $\lambda_{ssl} = 5$, and temperature 0.22 on all three datasets. PAAC and DR-GNN use separate same-path control implementations in the transfer diagnostics of Appendix L.

Reproducible analysis. Tables are generated from seed-level and aggregate numerical results, with development, calibration, and protected-test outcomes kept separate. Checksums identify model checkpoints, data protocols, and numerical outputs. The accompanying structured results retain the values underlying each table.

Saved evaluation state. Each unified run saves its calibrated thresholds before loading evaluation labels and records the validation-model identity. All 137 candidate/order outcomes are retained per run. The comparison grid is descriptive; evaluation outcomes do not select the reported primary policy.

D INDEPENDENTLY CALIBRATED ORDERING AUDIT

Table 8 reports independently calibrated orders on the Gowalla calibration pool: five training seeds, 100 shared random partitions per seed (5,000 calibration, 9,929 evaluation users). PCE reductions and paired differences are in percent and percentage points. H_R/H_N are evaluation harms in 10^{-3} units; Within is the percentage of the 500 evaluation batches meeting both budgets. Means first average partitions within seed; $\pm$ is SD of five seed means, not a confidence interval.

Table 8: Independently calibrated ordering on Gowalla.

Backbone	Learned PCE red.	Regret PCE red.	Paired difference	Learned H_R/H_N	Within (%)
BPR-MF	3.233±0.254	1.748±0.155	+1.485±0.303	0.870/0.360	67.6
LightGCN	2.352±0.096	1.219±0.125	+1.133±0.171	0.877/0.522	69.4
XSimGCL	0.736±0.035	0.731±0.035	+0.005±0.001	0.302/0.089	100.0
LightCCF	1.366±0.033	1.355±0.037	+0.011±0.008	0.317/0.103	100.0

This conditional audit uses all four Gowalla backbones and five seeds. The 100 user partitions are shared across backbones, seeds and orders. The random comparator uses a fixed user-identity hash order; variation across partitions does not represent repeated random-order draws. The predictor is fitted on the separate development set and then held fixed across partitions. Both thresholds use calibration outcomes only. The 12,000 policy evaluations share users and are statistically dependent. The complete release includes every threshold and paired comparison, per-seed partition SDs and 2.5/97.5 percentiles. Those percentiles describe variation across partitions of the fixed pool. Table 9 reports all six independently calibrated orders, retaining every backbone and seed. Values are means over five seed means, each averaging 100 shared partitions. PCE reduction is relative (%); harms are in 10^{-3} units; authorization, changed-list and Within columns are percentages. Full per-seed means, partition SDs and quantiles accompany the results.

Table 9: Complete independently calibrated ordering results.

Backbone	Order	PCE red.	H_R	H_N	Authorized	Changed	Within
BPR-MF	Random	1.537	0.901	0.403	36.412	7.937	61.400
BPR-MF	PCE gain	2.003	0.895	0.398	4.135	4.135	60.600
BPR-MF	Regret	1.748	0.852	0.419	89.153	12.029	70.200
BPR-MF	Predicted harm	2.828	0.855	0.427	92.611	14.228	70.000
BPR-MF	Predicted harm/gain	3.233	0.870	0.360	10.130	10.130	67.600
BPR-MF	Regret/gain	2.415	0.871	0.365	9.926	9.926	68.600
LightGCN	Random	1.184	0.879	0.609	41.172	9.597	65.600
LightGCN	PCE gain	1.051	0.910	0.567	3.089	2.935	59.200
LightGCN	Regret	1.219	0.851	0.668	90.042	13.232	68.200
LightGCN	Predicted harm	2.059	0.861	0.591	95.700	18.890	68.600
LightGCN	Predicted harm/gain	2.352	0.877	0.522	13.508	13.353	69.400
LightGCN	Regret/gain	1.632	0.878	0.556	9.176	9.022	67.200
XSimGCL	Random	0.735	0.305	0.091	99.980	6.661	100.000
XSimGCL	PCE gain	0.736	0.303	0.090	99.980	6.643	100.000
XSimGCL	Regret	0.731	0.305	0.091	99.980	6.643	100.000
XSimGCL	Predicted harm	0.734	0.304	0.090	99.979	6.642	100.000
XSimGCL	Predicted harm/gain	0.736	0.302	0.089	99.980	6.643	100.000
XSimGCL	Regret/gain	0.735	0.301	0.089	99.981	6.644	100.000
LightCCF	Random	1.360	0.316	0.102	99.632	6.580	100.000
LightCCF	PCE gain	1.369	0.317	0.103	99.973	6.578	100.000
LightCCF	Regret	1.355	0.303	0.099	99.972	6.576	100.000
LightCCF	Predicted harm	1.359	0.317	0.103	99.966	6.570	100.000
LightCCF	Predicted harm/gain	1.366	0.317	0.103	99.973	6.578	100.000
LightCCF	Regret/gain	1.366	0.317	0.103	99.973	6.577	100.000

E ALIGNMENT-METRIC ROBUSTNESS

All three datasets and five LightGCN seeds use their saved calibration rankings. Training degrees and histories define every alternative metric. CP-group JSD uses the source CP grouping: descending item popularity, approximately 20/60/20 percent of training interaction mass for head/mid/tail. User-history and list histograms are normalized separately; JSD uses base-2 logarithms with zero-probability terms contributing zero and no additive smoothing. This grouping is distinct from rank-quantile exposure strata. The position-weighted diagnostic replaces the list histogram's equal weights by

normalized $1/\log_2(k+1)$ at rank k, retaining the unweighted history histogram. It measures position-weighted popularity alignment. PCE uses the same empirical inverse-CDF convention at 5, 11 and 21 evenly spaced quantiles. The 11-point values reproduce the saved per-user PCE to 10^{-12}. The complete per-seed results also retain the uniform candidate policy and the primary policy's PCE-improved, worsened and unchanged outcome groups, with their population sizes and signed utility changes/positive-part harms. These groups are defined after outcomes are observed. Table 10 reports five-seed means for the fixed primary LightGCN policy on calibration users. Base values and paired absolute reductions are scaled by 10^3; SD is over five seed differences. Relative reductions are means of paired percentages. Metrics have different definitions and their scales should not be combined.

Table 10: Popularity-alignment metric robustness.

Dataset	Metric	Base	Absolute reduction	Relative (%)	Positive seeds
Gowalla	PCE, 5 quantiles	125.028	3.144±0.092	2.514	5/5
Gowalla	PCE, 11 quantiles	126.253	3.118±0.144	2.469	5/5
Gowalla	PCE, 21 quantiles	125.395	2.910±0.148	2.320	5/5
Gowalla	CP-group JSD	216.361	4.715±0.209	2.179	5/5
Gowalla	Position-weighted JSD	247.207	3.802±0.202	1.538	5/5
Yelp2018	PCE, 5 quantiles	119.105	1.968±0.033	1.655	5/5
Yelp2018	PCE, 11 quantiles	121.917	1.894±0.031	1.555	5/5
Yelp2018	PCE, 21 quantiles	121.587	1.852±0.029	1.525	5/5
Yelp2018	CP-group JSD	204.749	2.657±0.041	1.299	5/5
Yelp2018	Position-weighted JSD	223.550	1.965±0.033	0.880	5/5
Amazon-Book	PCE, 5 quantiles	146.905	2.532±0.031	1.724	5/5
Amazon-Book	PCE, 11 quantiles	148.519	2.381±0.027	1.603	5/5
Amazon-Book	PCE, 21 quantiles	147.719	2.329±0.022	1.577	5/5
Amazon-Book	CP-group JSD	238.491	3.551±0.070	1.489	5/5
Amazon-Book	Position-weighted JSD	257.354	2.567±0.052	0.998	5/5

F One-Factor Construction Ablations

Table 11 reports one-factor construction ablations on frozen LightGCN, with common gain-only authorization. Values average 100 partitions within each of five seeds; ± denotes SD of seed means. PCE reduction, Changed and Within are percentages; harms are in 10^{-3} units. Curve and robust variants match development-user perturbation energy, not realized risk or separately optimized strength.

Table 11: Complete candidate-construction ablations.

Dataset	Candidate	PCE red.	H_R	H_N	Changed	Within
Gowalla	RCPA, restored	1.051±0.222	0.910±0.015	0.567±0.031	2.93	59.20
Gowalla	RCPA, native	0.864±0.173	0.830±0.028	0.697±0.051	1.82	63.80
Gowalla	Full-quantile mean	0.966±0.177	0.914±0.018	0.548±0.045	2.41	61.40
Gowalla	Robust relevance scale	1.001±0.107	0.926±0.014	0.544±0.025	2.94	59.40
Yelp2018	RCPA, restored	1.557±0.069	0.525±0.059	0.318±0.026	14.13	100.00
Yelp2018	RCPA, native	1.558±0.069	0.487±0.058	0.501±0.077	12.97	97.20
Yelp2018	Full-quantile mean	1.580±0.061	0.556±0.055	0.333±0.029	14.34	100.00
Yelp2018	Robust relevance scale	1.596±0.040	0.538±0.071	0.320±0.027	14.47	100.00
Amazon-Book	RCPA, restored	1.613±0.021	0.426±0.035	0.406±0.021	17.59	100.00
Amazon-Book	RCPA, native	1.619±0.022	0.375±0.029	0.447±0.020	15.24	99.40
Amazon-Book	Full-quantile mean	2.004±0.036	0.512±0.039	0.435±0.020	21.63	100.00
Amazon-Book	Robust relevance scale	1.870±0.013	0.485±0.034	0.426±0.018	20.22	100.00

The reference uses median bias $b_u = F_u(Q_{u,0.5}) - 0.5$, standardized Top-$M$ relevance, and restored Base order. The first ablation uses native adjusted-score order. The second replaces b_u by $\bar{b}_u = 11^{-1}\sum_{j=0}^{10}[F_u(Q_{u,j/10}) - j/10]$. The third replaces the relevance standard deviation by $\mathrm{IQR}/1.349$. To isolate the effect of relevance scaling, the robust variant leaves catalog popularity standardization unchanged, and both relevance scales use the same machine-epsilon floor. Each ablation therefore changes one factor while holding the others fixed.

For curve and robust variants, a single positive multiplier per dataset/seed matches the reference's mean squared, within-candidate-centered raw-score perturbation on validation users. Specifically, if s_u is the relevance scale and z_i the fixed catalog popularity score, define $E = \sum_u M^{-1}\sum_{i\in C_u}[s_u b_u(z_i - \bar{z}_{C_u})]^2$. The alternative multiplier is $\sqrt{E_{\mathrm{ref}}/E_{\mathrm{alt}}}$. This uses only frozen

scores and training histories, not validation outcome labels or calibration outcomes. The original $\gamma = 0.0175$ is retained after this scale conversion. This matches a development perturbation magnitude, not realized risk, changed-list rate, or independently optimized performance. Degenerate energy is rejected rather than assigned an outcome-dependent fallback.

All 15 dataset/seed combinations retain the original candidates exactly. For fixed selected sets, native and restored RCPA have identical PCE and Recall; after separate risk calibration their authorized users may differ. The same 100-partition gain-only protocol yields 6,000 policy evaluations for the four constructions. Full-quantile PCE differences relative to the reference are negative in all five Gowalla seeds, but positive in all five Yelp2018 and Amazon-Book seeds. The Amazon-Book increment is 0.390 ± 0.031 percentage points (mean and SD across seed differences). The benefit of the fuller bias summary therefore varies by dataset.

Position-sensitive restoration tradeoff. With each selected set held fixed, mean reductions in position-weighted CP-group JSD (scaled by 10^3) are $5.101/1.965/2.567$ for restored RCPA and $9.998/3.749/4.976$ for native adjusted-score order on Gowalla/Yelp2018/Amazon-Book. After separate gain-only calibration, they are $1.709/1.969/2.589$ and $1.648/2.587/3.172$, respectively. Thus restoration trades some position-weighted popularity alignment for lower mean NDCG harm; it does not improve every metric. This JSD diagnostic uses the training-history CP groups and normalized $1/\log_2(k+1)$ list weights defined in Appendix E. Fixed-set PCE and Recall identity is checked for every user across all 15 runs.

Budget units and conditioning. An absolute positive-part budget of 10^{-3} applies to the average normalized metric loss over target units, not a relative percentage decrease or a number of clicks. Under the expectation contract, 1,000 target units have expected summed positive-part loss at most one normalized metric unit. This is not an application-specific tolerance recommendation. In the calibration audit, mean Recall/NDCG harms are $.000445/.000296$ for all users, $.001554/.000837$ among authorized users and $.004014/.001993$ among set-changed users. These denominators define different quantities; conditioning on authorization or a changed item set is not covered by the population-average bound. Position changes can affect NDCG even when item membership is unchanged.

G BUDGET-CONDITIONED ORDERING AUDIT

This analysis uses fixed calibration pools and predictors. Six orders are compared on 100 identity-matched partitions per seed, with 5,000 users for calibration and the remainder for evaluation. Target outcomes never select thresholds. Equal-budget scores preserve the original formula; unequal-budget harm/gain scores use the frozen budget-normalized rule. The harm-only comparator normalizes each clipped predicted harm plus its stabilizer by the corresponding budget before taking the maximum. At the matching budgets, thresholds, harms and PCE agree with the ordering analysis in Appendix D.

Table 12 reports learned-minus-comparator PCE reduction in percentage points, as five-seed means and SD after averaging partitions within each seed. Budgets use 10^{-3} units. Pressure is the corrected full-calibration-pool harm divided by budget, maximized over endpoints. The .10-point/four-positive-seed screen is a descriptive, pre-analysis criterion, not a significance or business-value threshold. Qualifying budget points span eight settings against regret-only, six against regret/gain and three against gain-only. Results at different budgets share users. A mean pressure below one can coexist with selective individual partitions or target score ranges. No monotone relationship or learned-policy performance guarantee follows from this diagnostic. All 72 learned-order setting–budget means satisfy both realized budgets; maximum mean Recall and NDCG utilization are .937 and .771, respectively. However, the lowest mean fraction of evaluation batches meeting both budgets is only 55.8%, consistent with a marginal expectation rather than per-batch guarantee.

Table 12: Budget-conditioned ordering increments.

Setting	Budget	Pressure	vs. regret	vs. regret/gain	vs. gain
Gowalla MF	0.5/0.5	5.05	$+1.812 \pm 0.401$	$+0.816 \pm 0.350$	$+1.621 \pm 0.332$
Gowalla MF	1/1	2.52	$+1.485 \pm 0.303$	$+0.818 \pm 0.345$	$+1.230 \pm 0.268$
Gowalla MF	2/2	1.26	$+0.558 \pm 0.418$	$+0.265 \pm 0.220$	$+0.092 \pm 0.047$
Gowalla MF	5/5	0.50	$+0.015 \pm 0.004$	$+0.001 \pm 0.003$	-0.003 ± 0.002

Setting	Budget	Pressure	vs. regret	vs. regret/gain	vs. gain
Gowalla MF	1/0.5	2.62	$+2.342 \pm 0.178$	$+0.907 \pm 0.309$	$+1.542 \pm 0.262$
Gowalla MF	0.5/1	5.05	$+1.772 \pm 0.406$	$+0.818 \pm 0.350$	$+1.623 \pm 0.331$
Gowalla GCN	0.5/0.5	4.90	$+1.397 \pm 0.172$	$+0.672 \pm 0.220$	$+1.171 \pm 0.184$
Gowalla GCN	1/1	2.45	$+1.133 \pm 0.171$	$+0.720 \pm 0.088$	$+1.301 \pm 0.168$
Gowalla GCN	2/2	1.23	$+0.266 \pm 0.064$	$+0.161 \pm 0.096$	$+0.159 \pm 0.106$
Gowalla GCN	5/5	0.49	$+0.022 \pm 0.004$	$+0.015 \pm 0.002$	-0.007 ± 0.002
Gowalla GCN	1/0.5	3.35	$+2.010 \pm 0.143$	$+1.152 \pm 0.320$	$+1.573 \pm 0.123$
Gowalla GCN	0.5/1	4.90	$+0.990 \pm 0.200$	$+0.577 \pm 0.224$	$+1.153 \pm 0.173$
Gowalla XSim	0.5/0.5	0.98	$+0.123 \pm 0.098$	$+0.012 \pm 0.028$	$+0.003 \pm 0.013$
Gowalla XSim	1/1	0.49	$+0.005 \pm 0.001$	$+0.000 \pm 0.001$	-0.001 ± 0.000
Gowalla XSim	2/2	0.25	$+0.005 \pm 0.001$	$+0.000 \pm 0.001$	-0.001 ± 0.000
Gowalla XSim	5/5	0.10	$+0.005 \pm 0.001$	$+0.000 \pm 0.001$	-0.001 ± 0.000
Gowalla XSim	1/0.5	0.58	$+0.005 \pm 0.001$	$+0.000 \pm 0.001$	-0.001 ± 0.000
Gowalla XSim	0.5/1	0.98	$+0.123 \pm 0.098$	$+0.012 \pm 0.028$	$+0.003 \pm 0.013$
Gowalla CCF	0.5/0.5	1.01	$+0.253 \pm 0.166$	$+0.084 \pm 0.060$	-0.005 ± 0.011
Gowalla CCF	1/1	0.50	$+0.011 \pm 0.008$	$+0.000 \pm 0.000$	-0.003 ± 0.001
Gowalla CCF	2/2	0.25	$+0.009 \pm 0.004$	$+0.000 \pm 0.000$	-0.003 ± 0.001
Gowalla CCF	5/5	0.10	$+0.009 \pm 0.004$	$+0.000 \pm 0.000$	-0.003 ± 0.001
Gowalla CCF	1/0.5	0.61	$+0.011 \pm 0.008$	$+0.000 \pm 0.000$	-0.003 ± 0.001
Gowalla CCF	0.5/1	1.01	$+0.253 \pm 0.166$	$+0.084 \pm 0.060$	-0.005 ± 0.011
Yelp2018 MF	0.5/0.5	1.68	$+2.825 \pm 0.308$	$+0.364 \pm 0.191$	$+0.010 \pm 0.254$
Yelp2018 MF	1/1	0.84	$+0.079 \pm 0.037$	$+0.018 \pm 0.011$	-0.002 ± 0.003
Yelp2018 MF	2/2	0.42	$+0.007 \pm 0.002$	$+0.001 \pm 0.001$	-0.002 ± 0.001
Yelp2018 MF	5/5	0.17	$+0.007 \pm 0.002$	$+0.001 \pm 0.001$	-0.002 ± 0.001
Yelp2018 MF	1/0.5	1.46	$+3.645 \pm 0.110$	$+0.483 \pm 0.149$	$+0.032 \pm 0.015$
Yelp2018 MF	0.5/1	1.65	$+2.172 \pm 0.669$	$+0.361 \pm 0.196$	$+0.010 \pm 0.262$
Yelp2018 GCN	0.5/0.5	1.45	$+0.529 \pm 0.172$	$+0.146 \pm 0.038$	-0.007 ± 0.028
Yelp2018 GCN	1/1	0.73	$+0.020 \pm 0.012$	$+0.003 \pm 0.003$	-0.001 ± 0.001
Yelp2018 GCN	2/2	0.36	$+0.008 \pm 0.003$	$+0.002 \pm 0.003$	-0.001 ± 0.001
Yelp2018 GCN	5/5	0.15	$+0.008 \pm 0.003$	$+0.002 \pm 0.003$	-0.001 ± 0.001
Yelp2018 GCN	1/0.5	1.05	$+0.368 \pm 0.188$	$+0.049 \pm 0.015$	$+0.016 \pm 0.006$
Yelp2018 GCN	0.5/1	1.45	$+0.501 \pm 0.160$	$+0.154 \pm 0.038$	$+0.004 \pm 0.021$
Yelp2018 XSim	0.5/0.5	0.67	$+0.015 \pm 0.013$	-0.001 ± 0.002	-0.001 ± 0.001
Yelp2018 XSim	1/1	0.34	$+0.006 \pm 0.001$	-0.000 ± 0.000	-0.000 ± 0.000
Yelp2018 XSim	2/2	0.17	$+0.006 \pm 0.001$	-0.000 ± 0.000	-0.000 ± 0.000
Yelp2018 XSim	5/5	0.07	$+0.006 \pm 0.001$	-0.000 ± 0.000	-0.000 ± 0.000
Yelp2018 XSim	1/0.5	0.50	$+0.006 \pm 0.001$	-0.000 ± 0.000	-0.000 ± 0.000
Yelp2018 XSim	0.5/1	0.67	$+0.015 \pm 0.013$	-0.001 ± 0.002	-0.001 ± 0.001
Yelp2018 CCF	0.5/0.5	0.86	$+0.086 \pm 0.072$	$+0.013 \pm 0.009$	-0.009 ± 0.011
Yelp2018 CCF	1/1	0.43	$+0.016 \pm 0.008$	$+0.010 \pm 0.010$	-0.001 ± 0.001
Yelp2018 CCF	2/2	0.21	$+0.016 \pm 0.008$	$+0.010 \pm 0.010$	-0.001 ± 0.001
Yelp2018 CCF	5/5	0.09	$+0.016 \pm 0.008$	$+0.010 \pm 0.010$	-0.001 ± 0.001
Yelp2018 CCF	1/0.5	0.68	$+0.024 \pm 0.015$	$+0.010 \pm 0.010$	-0.001 ± 0.001
Yelp2018 CCF	0.5/1	0.85	$+0.088 \pm 0.075$	$+0.015 \pm 0.012$	-0.007 ± 0.007
Amazon-Book MF	0.5/0.5	2.16	$+2.954 \pm 0.086$	$+0.694 \pm 0.056$	$+0.378 \pm 0.085$
Amazon-Book MF	1/1	1.08	$+2.611 \pm 0.401$	$+0.030 \pm 0.023$	-0.006 ± 0.005
Amazon-Book MF	2/2	0.54	$+0.005 \pm 0.001$	$+0.001 \pm 0.000$	-0.001 ± 0.001
Amazon-Book MF	5/5	0.22	$+0.005 \pm 0.001$	$+0.001 \pm 0.000$	-0.001 ± 0.001
Amazon-Book MF	1/0.5	2.15	$+3.317 \pm 0.102$	$+1.054 \pm 0.073$	$+0.728 \pm 0.075$
Amazon-Book MF	0.5/1	1.74	$+3.121 \pm 0.104$	$+0.162 \pm 0.145$	-0.114 ± 0.071
Amazon-Book GCN	0.5/0.5	1.33	$+1.043 \pm 0.039$	$+0.190 \pm 0.072$	-0.036 ± 0.026
Amazon-Book GCN	1/1	0.66	$+0.011 \pm 0.004$	$+0.004 \pm 0.001$	-0.006 ± 0.002
Amazon-Book GCN	2/2	0.33	$+0.006 \pm 0.001$	$+0.004 \pm 0.001$	-0.006 ± 0.002
Amazon-Book GCN	5/5	0.13	$+0.006 \pm 0.001$	$+0.004 \pm 0.001$	-0.006 ± 0.002
Amazon-Book GCN	1/0.5	1.21	$+1.150 \pm 0.054$	$+0.187 \pm 0.061$	$+0.025 \pm 0.010$
Amazon-Book GCN	0.5/1	1.26	$+0.467 \pm 0.052$	$+0.142 \pm 0.064$	-0.034 ± 0.023
Amazon-Book XSim	0.5/0.5	0.70	$+0.022 \pm 0.014$	$+0.003 \pm 0.003$	-0.000 ± 0.000
Amazon-Book XSim	1/1	0.35	$+0.004 \pm 0.001$	$+0.000 \pm 0.000$	-0.000 ± 0.000

Setting	Budget	Pressure	vs. regret	vs. regret/gain	vs. gain
Amazon-Book XSim	2/2	0.18	$+0.004 \pm 0.001$	$+0.000 \pm 0.000$	-0.000 ± 0.000
Amazon-Book XSim	5/5	0.07	$+0.004 \pm 0.001$	$+0.000 \pm 0.000$	-0.000 ± 0.000
Amazon-Book XSim	1/0.5	0.51	$+0.004 \pm 0.001$	$+0.000 \pm 0.000$	-0.000 ± 0.000
Amazon-Book XSim	0.5/1	0.70	$+0.022 \pm 0.014$	$+0.003 \pm 0.003$	-0.000 ± 0.000
Amazon-Book CCF	0.5/0.5	0.69	$+0.023 \pm 0.015$	-0.001 ± 0.003	-0.004 ± 0.004
Amazon-Book CCF	1/1	0.34	$+0.009 \pm 0.001$	$+0.000 \pm 0.000$	-0.002 ± 0.000
Amazon-Book CCF	2/2	0.17	$+0.009 \pm 0.001$	$+0.000 \pm 0.000$	-0.002 ± 0.000
Amazon-Book CCF	5/5	0.07	$+0.009 \pm 0.001$	$+0.000 \pm 0.000$	-0.002 ± 0.000
Amazon-Book CCF	1/0.5	0.51	$+0.009 \pm 0.001$	$+0.000 \pm 0.000$	-0.002 ± 0.000
Amazon-Book CCF	0.5/1	0.69	$+0.023 \pm 0.015$	-0.001 ± 0.003	-0.004 ± 0.004

H CONDITIONAL SEED UNCERTAINTY AND CANDIDATE COMPARISONS

For every setting, budget and comparator, we first average the 100 paired partition differences within each seed. Five such values determine the mean, SD and nominal Student-t interval, with four degrees of freedom. Under independent, approximately normal seed-level differences, these intervals describe training-seed variability for the fixed pool and protocol. They exclude uncertainty from data collection or analysis design, and all 216 contrasts are reported. Intervals are unadjusted for multiple comparisons and do not replace the multiplicity analysis of the primary frozen evaluation. Budget points share users and are correlated. Table 13 gives intervals in PCE percentage points; G/Y/A mean Gowalla/Yelp2018/Amazon-Book and MF/GCN/XSim/CCF denote the four backbones. Budget units are 10^{-3}. Figure 4 shows all six budgets. Each panel contains the 72 setting–budget points.

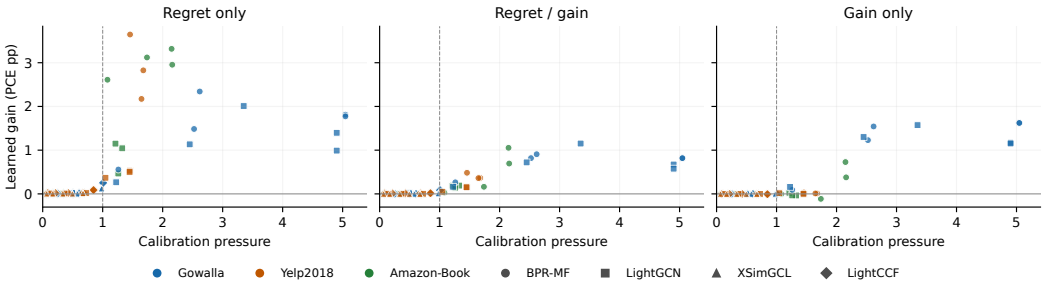

Figure 4: Budget pressure and ordering value.

Table 13: Conditional seed uncertainty.

Setting	Budget	vs. regret	vs. regret/gain	vs. gain
G MF	0.5/0.5	$[+1.315, +2.309]$	$[+0.382, +1.251]$	$[+1.209, +2.033]$
G MF	1/1	$[+1.108, +1.861]$	$[+0.390, +1.247]$	$[+0.896, +1.563]$
G MF	2/2	$[+0.038, +1.077]$	$[-0.008, +0.538]$	$[+0.034, +0.150]$
G MF	5/5	$[+0.011, +0.019]$	$[-0.002, +0.005]$	$[-0.006, -0.000]$
G MF	1/0.5	$[+2.121, +2.562]$	$[+0.524, +1.291]$	$[+1.216, +1.868]$
G MF	0.5/1	$[+1.268, +2.277]$	$[+0.383, +1.252]$	$[+1.212, +2.033]$
G GCN	0.5/0.5	$[+1.183, +1.611]$	$[+0.399, +0.945]$	$[+0.943, +1.400]$
G GCN	1/1	$[+0.921, +1.346]$	$[+0.611, +0.830]$	$[+1.093, +1.510]$
G GCN	2/2	$[+0.187, +0.345]$	$[+0.042, +0.280]$	$[+0.028, +0.291]$
G GCN	5/5	$[+0.017, +0.027]$	$[+0.012, +0.017]$	$[-0.009, -0.005]$
G GCN	1/0.5	$[+1.832, +2.187]$	$[+0.755, +1.549]$	$[+1.421, +1.726]$
G GCN	0.5/1	$[+0.742, +1.239]$	$[+0.299, +0.854]$	$[+0.939, +1.368]$
G XSim	0.5/0.5	$[+0.001, +0.245]$	$[-0.023, +0.048]$	$[-0.012, +0.019]$
G XSim	1/1	$[+0.003, +0.007]$	$[-0.000, +0.001]$	$[-0.001, +0.000]$
G XSim	2/2	$[+0.003, +0.007]$	$[-0.000, +0.001]$	$[-0.001, +0.000]$
G XSim	5/5	$[+0.003, +0.007]$	$[-0.000, +0.001]$	$[-0.001, +0.000]$
G XSim	1/0.5	$[+0.003, +0.007]$	$[-0.000, +0.001]$	$[-0.001, +0.000]$
G XSim	0.5/1	$[+0.001, +0.245]$	$[-0.023, +0.048]$	$[-0.012, +0.019]$

Setting	Budget	vs. regret	vs. regret/gain	vs. gain
G CCF	0.5/0.5	$[+0.046, +0.459]$	$[+0.010, +0.158]$	$[-0.018, +0.009]$
G CCF	1/1	$[+0.001, +0.021]$	$[-0.000, +0.000]$	$[-0.004, -0.002]$
G CCF	2/2	$[+0.004, +0.014]$	$[-0.000, +0.000]$	$[-0.004, -0.002]$
G CCF	5/5	$[+0.004, +0.014]$	$[-0.000, +0.000]$	$[-0.004, -0.002]$
G CCF	1/0.5	$[+0.001, +0.021]$	$[-0.000, +0.000]$	$[-0.004, -0.002]$
G CCF	0.5/1	$[+0.046, +0.459]$	$[+0.010, +0.158]$	$[-0.018, +0.009]$
Y MF	0.5/0.5	$[+2.443, +3.208]$	$[+0.127, +0.602]$	$[-0.306, +0.326]$
Y MF	1/1	$[+0.033, +0.125]$	$[+0.005, +0.032]$	$[-0.006, +0.001]$
Y MF	2/2	$[+0.005, +0.010]$	$[-0.000, +0.002]$	$[-0.003, -0.002]$
Y MF	5/5	$[+0.005, +0.010]$	$[-0.000, +0.002]$	$[-0.003, -0.002]$
Y MF	1/0.5	$[+3.509, +3.782]$	$[+0.298, +0.667]$	$[+0.014, +0.051]$
Y MF	0.5/1	$[+1.341, +3.003]$	$[+0.118, +0.604]$	$[-0.315, +0.336]$
Y GCN	0.5/0.5	$[+0.316, +0.743]$	$[+0.099, +0.194]$	$[-0.042, +0.029]$
Y GCN	1/1	$[+0.005, +0.034]$	$[-0.001, +0.006]$	$[-0.002, +0.000]$
Y GCN	2/2	$[+0.004, +0.012]$	$[-0.001, +0.006]$	$[-0.002, +0.000]$
Y GCN	5/5	$[+0.004, +0.012]$	$[-0.001, +0.006]$	$[-0.002, +0.000]$
Y GCN	1/0.5	$[+0.135, +0.601]$	$[+0.030, +0.068]$	$[+0.009, +0.023]$
Y GCN	0.5/1	$[+0.302, +0.699]$	$[+0.106, +0.202]$	$[-0.022, +0.030]$
Y XSim	0.5/0.5	$[-0.001, +0.031]$	$[-0.003, +0.002]$	$[-0.001, +0.000]$
Y XSim	1/1	$[+0.005, +0.007]$	$[-0.000, +0.000]$	$[-0.001, -0.000]$
Y XSim	2/2	$[+0.005, +0.007]$	$[-0.000, +0.000]$	$[-0.001, -0.000]$
Y XSim	5/5	$[+0.005, +0.007]$	$[-0.000, +0.000]$	$[-0.001, -0.000]$
Y XSim	1/0.5	$[+0.005, +0.007]$	$[-0.000, +0.000]$	$[-0.001, -0.000]$
Y XSim	0.5/1	$[-0.001, +0.031]$	$[-0.003, +0.002]$	$[-0.001, +0.000]$
Y CCF	0.5/0.5	$[-0.003, +0.175]$	$[+0.001, +0.025]$	$[-0.023, +0.005]$
Y CCF	1/1	$[+0.006, +0.027]$	$[-0.002, +0.022]$	$[-0.003, +0.000]$
Y CCF	2/2	$[+0.006, +0.027]$	$[-0.002, +0.022]$	$[-0.003, +0.000]$
Y CCF	5/5	$[+0.006, +0.027]$	$[-0.002, +0.022]$	$[-0.003, +0.000]$
Y CCF	1/0.5	$[+0.004, +0.043]$	$[-0.002, +0.022]$	$[-0.003, +0.000]$
Y CCF	0.5/1	$[-0.005, +0.181]$	$[+0.000, +0.031]$	$[-0.015, +0.002]$
A MF	0.5/0.5	$[+2.848, +3.061]$	$[+0.625, +0.764]$	$[+0.272, +0.484]$
A MF	1/1	$[+2.113, +3.108]$	$[+0.000, +0.059]$	$[-0.012, -0.000]$
A MF	2/2	$[+0.004, +0.006]$	$[+0.000, +0.001]$	$[-0.001, -0.000]$
A MF	5/5	$[+0.004, +0.006]$	$[+0.000, +0.001]$	$[-0.001, -0.000]$
A MF	1/0.5	$[+3.191, +3.444]$	$[+0.963, +1.145]$	$[+0.635, +0.821]$
A MF	0.5/1	$[+2.992, +3.249]$	$[-0.017, +0.342]$	$[-0.202, -0.025]$
A GCN	0.5/0.5	$[+0.995, +1.092]$	$[+0.101, +0.280]$	$[-0.068, -0.004]$
A GCN	1/1	$[+0.006, +0.015]$	$[+0.003, +0.006]$	$[-0.008, -0.003]$
A GCN	2/2	$[+0.005, +0.008]$	$[+0.003, +0.006]$	$[-0.008, -0.003]$
A GCN	5/5	$[+0.005, +0.008]$	$[+0.003, +0.006]$	$[-0.008, -0.003]$
A GCN	1/0.5	$[+1.082, +1.217]$	$[+0.112, +0.263]$	$[+0.013, +0.038]$
A GCN	0.5/1	$[+0.402, +0.532]$	$[+0.063, +0.222]$	$[-0.063, -0.006]$
A XSim	0.5/0.5	$[+0.004, +0.039]$	$[-0.001, +0.007]$	$[-0.001, +0.000]$
A XSim	1/1	$[+0.003, +0.005]$	$[-0.000, +0.000]$	$[-0.001, -0.000]$
A XSim	2/2	$[+0.003, +0.005]$	$[-0.000, +0.000]$	$[-0.001, -0.000]$
A XSim	5/5	$[+0.003, +0.005]$	$[-0.000, +0.000]$	$[-0.001, -0.000]$
A XSim	1/0.5	$[+0.003, +0.005]$	$[-0.000, +0.000]$	$[-0.001, -0.000]$
A XSim	0.5/1	$[+0.004, +0.039]$	$[-0.001, +0.007]$	$[-0.001, +0.000]$
A CCF	0.5/0.5	$[+0.004, +0.041]$	$[-0.005, +0.003]$	$[-0.008, +0.001]$
A CCF	1/1	$[+0.007, +0.010]$	$[-0.000, +0.000]$	$[-0.002, -0.002]$
A CCF	2/2	$[+0.007, +0.010]$	$[-0.000, +0.000]$	$[-0.002, -0.002]$
A CCF	5/5	$[+0.007, +0.010]$	$[-0.000, +0.000]$	$[-0.002, -0.002]$
A CCF	1/0.5	$[+0.007, +0.010]$	$[-0.000, +0.000]$	$[-0.002, -0.002]$
A CCF	0.5/1	$[+0.004, +0.041]$	$[-0.005, +0.003]$	$[-0.008, +0.001]$

Joint candidate constraints. We additionally compare the complete 49-candidate grid within each dataset and ordering rule. We compare mean PCE reduction, Recall harm and NDCG harm, adding

the changed-list rate when required. An alternative must match or improve RCPA on every criterion and strictly improve at least one (numerical tolerance 10^{-12}). No candidate parameter is optimized in this descriptive audit. Of 48 alternatives on Gowalla, six satisfy all four mean comparisons under gain-only ordering; four also satisfy them separately in every seed. These are restored Top-M CP candidates with weights .035, .05, .075 and .1. Under learned ordering eight satisfy the mean comparisons, but none does so in every seed. On Yelp2018 and Amazon-Book, no alternative satisfies all four mean comparisons; RCPA does not satisfy the reverse comparisons against any alternative either. Lack of dominance establishes neither superiority nor a matched point: policies may occupy different tradeoffs. These inequalities compare observed means within the specified pools. Native/restored and global/Top-M variants remain distinct candidates in the complete grid.

I IMPLEMENTATION COST

Predictor fitting and wrapper-cost measurements use validation-only inputs and are separate from protected evaluation. The fitting-cost table is reported in Appendix N; the following tables report online wrapper overhead and peak additional allocation.

Table 14: End-to-end wrapper cost on Gowalla.

Backbone	Base	Full	Δ	Full kuser/s	Rel. overhead	Full peak	Δ peak
BPR-MF	0.02382	0.04527	0.02145	22.09	+90.02%	388.3	1.31
LightGCN	0.02388	0.04502	0.02115	22.21	+88.58%	388.3	1.31
XSimGCL	0.02343	0.04494	0.02151	22.25	+91.78%	309.9	2.79
LightCCF	0.02379	0.04632	0.02253	21.59	+94.69%	388.3	1.31

Table 15: Synchronized runtime component shares.

Backbone	Score	Top-M	History	Rerank	Features	Gate	Fallback
BPR-MF	48.91	4.79	15.38	2.66	28.02	0.19	0.06
LightGCN	49.50	4.82	15.32	2.78	27.32	0.20	0.07
XSimGCL	49.66	4.70	15.26	2.78	27.29	0.19	0.12
LightCCF	48.04	4.70	15.87	2.68	28.40	0.19	0.12
Mean	49.03	4.75	15.46	2.73	27.76	0.19	0.09

J CALIBRATION DIAGNOSTICS AND USER OUTCOMES

These supplementary measurements use the development calibration pools, rather than the disjoint evaluation groups of the unified main experiment. They describe the distribution of benefits and losses conditional on those pools. Table 16 averages five seeds within each dataset–backbone setting and then weights the 12 settings equally. Budgets are in units of 10^{-3}; PCE reduction and admission rates are percentages. In panel B, "Changed" conditions on Top-20 membership changes. The top-decile share is the fraction of positive PCE gain received by the highest-gain tenth of the indicated population. Table 17 reports positive-part harm. Incidence is the percentage with harm above 10^{-12};

Table 16: Budget sensitivity and user outcomes.

A. Budget sensitivity

α_R/α_N	PCE red.	Admitted
0.5/0.5	1.672	65.90
1/1	1.841	85.41
2/2	1.944	91.71
5/5	1.941	100.00
1/0.5	1.807	77.98
0.5/1	1.676	65.95

B. User outcomes

Outcome	All	Changed
PCE improves	10.14	88.90
PCE worsens	0.16	1.66
PCE unchanged	89.70	9.45
Top-decile share	95.16	31.89
Recall harm	0.248	1.980
NDCG harm	0.783	3.390

means include zero-harm users within each named group and retain the original metric units. Activity groups use training-history length and preference groups use training-item popularity. Changing rank positions without changing membership can affect NDCG while leaving Recall unchanged. These

conditional means characterize how loss is concentrated; the population risk bound applies to its specified population.

Table 17: Calibration utility-harm distributions.

Group	Recall incidence	Recall mean	NDCG incidence	NDCG mean
All users	0.248	0.000445	0.783	0.000296
Authorized	0.643	0.001554	1.464	0.000837
Unauthorized	0.000	0.000000	0.000	0.000000
Set-changed	1.980	0.004014	3.390	0.001993
Set-unchanged	0.000	0.000000	0.348	0.000049
Low	0.071	0.000433	0.096	0.000135
Medium	0.080	0.000359	0.125	0.000118
High	0.144	0.000401	0.296	0.000175
Very High	0.710	0.000568	2.667	0.000762
Low-pop preference	0.304	0.000596	0.980	0.000416
Mid-pop preference	0.243	0.000357	0.793	0.000264
High-pop preference	0.198	0.000381	0.575	0.000208

Tail losses after authorization. Table 18 uses the frozen learned policy's calibration-pool replay (Appendix O.1). Incidence conditions on authorized membership changes; $q^+_{.5}/q^+_{.9}$ further condition on positive loss. Entries average 20 job-specific statistics per dataset (four backbones, five seeds). Quantiles use native $[0, 1]$ units and linear interpolation, without pooling users across jobs. Most changed lists have zero observed endpoint harm; a small subset bears the losses. The supplement retains authorized-user counts, unconditional quantiles and maxima.

Table 18: Harm incidence and positive-loss quantiles among authorized membership changes.

	Recall			NDCG		
Dataset	Incidence (%)	$q^+_{.5}$	$q^+_{.9}$	Incidence (%)	$q^+_{.5}$	$q^+_{.9}$
Gowalla	2.08	0.211	0.686	3.07	0.070	0.172
Yelp2018	2.03	0.132	0.347	3.36	0.053	0.116
Amazon-Book	1.97	0.128	0.341	4.09	0.053	0.108

K FINITE-POPULATION ALLOCATION DIAGNOSTIC

Table 19 compares the outcome-aware two-budget fractional allocation optimum with the feasible maximum-normalized-harm envelope. Each of the 15 blocks uses one LightGCN seed and one dataset's fixed calibration pool. Both selectors access the same realized gains and harms. The ratio is envelope gain divided by exact optimum; PCE entries are relative reduction percentages. Confidence intervals summarize the five seed blocks per dataset (15 when pooled); "Both active" counts optima with both constraints binding. The empirical ratios exceed the general one-half lower bound, but depend on these finite populations. Outcome-aware allocation is a diagnostic reference for the ordering analysis, whereas independent calibration evaluates policies selected without those outcomes.

Table 19: Finite-population oracle and envelope comparison.

Scope	γ	Blocks	Env./exact (%)	95% CI	Min. (%)	PCE (Env./exact)	Both active
All datasets	0.0175	15	99.88	[99.78, 99.98]	99.51	2.014 / 2.017	0/15
All datasets	0.1	15	99.72	[99.58, 99.86]	99.33	10.059 / 10.093	5/15
Gowalla	0.0175	5	99.63	[99.54, 99.72]	99.51	2.863 / 2.873	0/5
Gowalla	0.1	5	99.39	[99.34, 99.43]	99.33	13.543 / 13.627	5/5
Yelp2018	0.0175	5	100.00	[100.00, 100.00]	100.00	1.559 / 1.559	0/5
Yelp2018	0.1	5	99.91	[99.90, 99.91]	99.90	8.364 / 8.372	0/5
Amazon-Book	0.0175	5	100.00	[100.00, 100.00]	100.00	1.620 / 1.620	0/5
Amazon-Book	0.1	5	99.87	[99.87, 99.88]	99.87	8.271 / 8.281	0/5

Table 20: KuaiRec boundary audit.

Stage and policy	Users/seed	Change (%)	ΔR	ΔN	ΔPCE [95% CI]	Status
Dev. selected $q = 1, \gamma = 0.025$	200	5.00	-0.107	-0.262	-0.580 [-0.985, -0.176]	selection only
Cal. uniform $q = 1$ diagnostic	1,000	4.56	+0.018	-0.121	-0.450 [-0.577, -0.324]	uncertified
Cal. formal $q = 1, \gamma = 0.025$ gate	1,000	0.36	+0.024	+0.052	-0.109 [-0.340, +0.121]	criteria unmet

L EXTERNAL-DATA AND COMPARATOR DIAGNOSTICS

KuaiRec. The small matrix contains 1,411 users and the stated 3,327-item universe. The split assigns 200 users to development, 1,000 to calibration and 211 to protected evaluation; the latter outcomes remain unread. Five LightGCN checkpoints share the same candidate and masking rules. A bottom-q edit preserves the first $20 - q$ Base positions. Development considers $q \in \{1, 2, 3\}$ and $\gamma \in \{0, .0175, .025, .05, .1, .2\}$, selecting $q = 1, \gamma = .025$ by the fixed smallest-eligible-action rule. In Table 20, metric deltas and paired 95% t intervals use 10^{-3} units. The frozen gate changes 0.36% of lists, with treated counts $[0, 2, 14, 1, 1]$ across seeds. Its PCE interval crosses zero (two-sided sign-flip $p = .125$). All five calibration risk checks pass, while the specified coverage and improvement criteria fail. The KuaiRec results are calibration-cohort diagnostics; protected evaluation outcomes are excluded.

Training-time methods. Table 21 compares paired Full and Control implementations, reporting relative changes. PAAC has five seeds each on Gowalla and Yelp2018 and a single Amazon-Book seed. DR-GNN transfers the DRO weighting component on one Gowalla seed. The transferred objectives and evaluation paths differ from the source papers' complete protocols; these comparisons measure behavior under the stated implementation. PAAC improves exposure while degrading ranking on Gowalla and Yelp2018; the one-seed Amazon-Book result has the opposite coverage tradeoff.

Table 21: Training-time debiasing protocol audits.

Audit	Dataset	Seeds	ΔR	ΔN	ΔCov.	Scope
PAAC Full vs Control	Gowalla	5	-25.77%	-23.82%	+24.09%	protocol diagnostic
PAAC Full vs Control	Yelp2018	5	-37.55%	-34.09%	+38.97%	protocol diagnostic
DR-GNN DRO vs Control	Gowalla	1	-27.58%	-26.25%	-14.84%	protected core port
PAAC Full vs Control	Amazon-Book	1	+2.57%	+9.32%	-5.24%	partial; not pooled

Direct correction feasibility. Table 22 counts LightGCN seeds satisfying both absolute development net-utility budgets in the direct comparator audit. These net-utility checks differ from positive-part harm calibration. The PPD transfer fails this feasibility screen in all three datasets, whereas nonzero CP passes on Amazon-Book at $\lambda = .001$. Candidate-specific calibration in Appendix N separately assesses corrections behind a common risk gate.

M HISTORICAL CALIBRATION AND BOUNDARY DIAGNOSTICS

These diagnostics complement the unified evaluation by examining historical development calibration pools. Since repeated partitions reuse users, they characterize variation within those pools rather than independently confirming the fitted models.

Prediction near the authorization boundary. Frozen predictors use 100 shared Gowalla partitions per seed. Within each evaluation partition, we inspect at most 5% of users immediately inside and outside the learned-score threshold. Identity-based truncation is diagnostic only; deployed score atoms remain indivisible. Table 23 reports candidate outcomes, including rejected edits. Both selective backbones have larger gains and smaller harms inside the boundary. The other two have very few outside users. This supports a local ordering interpretation without assuming accurate individual-harm estimates or a validated model-selection rule. Gains and harms use 10^{-3} absolute units. They are conditional candidate outcomes, not population policy harms.

Sample size and finite-sample correction. Historical LightGCN policies at $\gamma = .0175$ and dual budgets .001 are held fixed. Each of five seeds uses 100 disjoint fitting/target partitions per sample

Table 22: Development feasibility of direct comparators.

Dataset	RCPA-CF	CP (nonzero)	PPD transfer
Gowalla	5/5	0/5	0/5
Yelp2018	5/5	0/5	0/5
Amazon-Book	5/5	5/5 ($\lambda = 0.001$)	0/5

Table 23: Candidate outcomes near historical authorization boundaries.

Backbone	Side	Users	Signed PCE gain	Recall harm	NDCG harm
bpr-mf	inside	496.4	14.573	10.617	4.393
bpr-mf	outside	497.0	8.630	13.150	5.678
lightgcn	inside	497.0	17.473	9.177	5.904
lightgcn	outside	496.2	11.739	16.551	10.444
xsimgcl	inside	497.0	-0.050	1.448	0.397
xsimgcl	outside	2.0	-2.805	14.438	5.029
lightccf	inside	497.0	-0.090	0.952	0.249
lightccf	outside	2.6	-0.900	0.036	0.012

size. We compare CRC correction with naive empirical thresholding. Table 24 shows their joint benefit–exceedance tradeoff. Near the correction floor, CRC can sharply reduce both editing and observed exceedance. Increasing sample size changes this tradeoff; it does not monotonically improve batch satisfaction. Gains are absolute PCE reductions, not percentages.

Table 24: Historical CRC versus naive empirical thresholds.

Dataset	Calibration n	Exceed (%, CRC/naive)	PCE gain (10^{-3}, CRC/naive)
amazon-book	1,000	0.0/0.0	1.094/2.385
amazon-book	2,000	0.0/0.0	2.317/2.394
amazon-book	5,000	0.0/0.0	2.387/2.388
amazon-book	10,000	0.0/0.0	2.383/2.383
gowalla	1,000	0.4/54.0	0.699/3.039
gowalla	2,000	18.0/55.4	2.441/3.157
gowalla	5,000	33.4/51.8	2.975/3.174
gowalla	10,000	37.2/46.6	3.073/3.176
yelp2018	1,000	0.0/0.0	0.924/1.882
yelp2018	2,000	0.0/0.0	1.780/1.893
yelp2018	5,000	0.0/0.0	1.895/1.897
yelp2018	10,000	0.0/0.0	1.895/1.895

Fixed-group mixture changes. A separate historical LightGCN audit uses $\gamma = .1$, a 2,000-user label-free pilot to freeze mismatch groups, and 10,000 disjoint risk-calibration users. Five seeds and 100 partitions per seed cover each dataset. Under observed/high-mismatch-heavy mixtures, pooled thresholds exceed either budget in 34.0/60.7% of target partitions. Group-specific thresholds reduce these rates to 12.7/16.0%, retaining 79.3/71.8% of pooled PCE reduction. A common cutoff satisfying every group is more conservative: exceedance is 2.2/6.2%, with PCE reduction 3.213/3.680%, versus 4.243/4.311% for group-specific thresholds. These are equal-dataset descriptive means. The design varies group proportions while holding within-group laws fixed. It neither tests temporal drift nor establishes high-probability or arbitrary-shift control. Complete mixture rows accompany the evidence supplement.

N CANDIDATE-SPECIFIC LEARNED ORDERING AND FITTING COST

The TTEN comparisons in this section retain their original scoring and ordering conventions. Finite-precision sigmoid scores can merge distinct raw scores into ties, so their zero-strength rankings need not reproduce the recorded Base. Candidate differences can therefore include a Base-ordering effect. Appendix Q provides a separate common-raw-score analysis.

Controlled protocol. The supplementary calibration-pool comparison uses 49 fixed candidates on three LightGCN datasets and five seeds. TTEN exponents are $p \in \{0, .01, .025, .05, .1, .25\}$, each with global/Top-$M$ scope and native/Base-restored order. CP uses Top-M candidates and weights $\{0, .0005, .001, .0025, .005, .0075, .01, .02, .035, .05, .075, .1\}$ with both output orders. The RCPA reference has $\gamma = .0175$. For each of 100 shared partitions, 5,000 users calibrate each threshold and the remaining calibration-pool users evaluate it. This supplementary pool differs from the unified evaluation groups in Appendix O. Each candidate receives the same 14 features, two ridge predictors and regularization strength 10. Feature standardization and positive-part Recall/NDCG harm targets use validation users only. Thus each comparison uses a candidate-specific validation fit without training on calibration outcomes.

For each of 100 shared user partitions, ordering is fixed before threshold selection, using the exported predictions and the fixed benefit/risk score. Learned and gain-only policies use the same candidate lists, calibration users, evaluation users and two nominal budgets of 0.001. Complete tied-score groups are authorized together, with exact Base fallback. This yields 73,500 paired policy evaluations. Independent recomputation checks maximal feasible complete-group prefixes and all four reported target metrics. Repeated partitions share users and are not independent experimental replications; tables first average partitions within seed and then summarize five seed means. Variation across partitions is conditional on these calibration pools.

Interpretation. For the RCPA candidate, learning raises Gowalla PCE reduction from 1.051% to 2.352%, with positive paired differences in all five seeds. The small negative differences on Yelp2018 and Amazon-Book show why this module is optional. Gains also transfer to competing candidates: restored CP at $\lambda = 0.0005$ gains 0.604 ± 0.237, 0.110 ± 0.160, and 0.345 ± 0.087 percentage points on Gowalla, Yelp2018 and Amazon-Book. These fixed comparison points illustrate modularity, not an optimized winner. Across the complete grid, 103 dataset–candidate mean differences are positive, 38 negative and six zero. These counts summarize correlated comparisons. The tables report paired differences and both realized harms, which can differ at equal nominal budgets.

Measured computational scope. The 735 fits each estimate both endpoint predictors, including feature standardization and in-sample diagnostics. Three additional fits per candidate measure timing variability; coefficients agree within 10^{-12} and repeats do not select a model. Timings were recorded on Linux x86-64 with 76 PyTorch CPU threads; the CPU model was not recorded. The timings apply to this execution environment. Ranking generation, relevance-regret computation and feature extraction are separate from the fit and scoring times. The numerical model state remains 30 coefficients and 28 standardization constants (464 bytes in float64), excluding container overhead. Total runtime also includes candidate generation and feature construction. Table 25 reports observed cost of fitting both endpoint predictors for one candidate. Each dataset includes 49 candidates across five seeds. Fit includes standardization and in-sample diagnostics; feature extraction is separate. Repeat is the median across per-model medians of three additional fits. Times describe this execution environment, not historical frozen-model training.

Table 25: Candidate-specific predictor fitting cost.

Dataset	Fit median (ms)	Fit range (ms)	Repeat (ms)	Score (μs/user)
Gowalla	7.81	6.41–17.66	6.22	0.10
Yelp2018	8.70	7.09–66.31	9.41	0.10
Amazon-Book	12.53	11.19–42.86	14.87	0.09

Table 26 summarizes all 49 candidates per dataset. Entries use paired learned-minus-gain-only PCE-reduction percentage points. Counts and ranges use the published three-decimal grid precision; near-zero entries should not be interpreted as exact equality. Complete rows, including both harms and seed SDs, are retained in the accompanying `R49_evidence_supplement.zip`. The full-precision analysis above remains the source for the 103/38/6 sign counts. Appendix H retains the joint candidate-dominance findings, including the stronger Gowalla CP controls.

Table 26: Complete-grid ordering effects by candidate family.

Dataset	Family	Candidates	Positive/negative/zero	Min.	Max.
G	RCPA	1	1/0/0	+1.301	+1.301
G	TTEN	24	12/12/0	-0.010	+1.380
G	CP	24	22/0/2	+0.000	+2.036
Y	RCPA	1	0/1/0	-0.001	-0.001
Y	TTEN	24	12/12/0	-0.008	+1.167
Y	CP	24	22/0/2	+0.000	+1.932
A	RCPA	1	0/1/0	-0.006	-0.006
A	TTEN	24	12/12/0	-0.014	+0.795
A	CP	24	22/0/2	+0.000	+2.725

O UNIFIED EVALUATION PROTOCOL

TTEN retains its historical scoring conventions, including the raw/sigmoid tie issue described in Appendix N. Appendix Q supplies the common-raw comparison.

All three datasets use the same existing-user protocol in this reevaluation. Training and validation interactions, identifiers and frozen backbones are retained. Within each held-out interaction pool, users are partitioned with a fixed random seed (2027) into disjoint calibration and evaluation groups. The calibration/evaluation user counts are 14,929/14,929 for Gowalla, 15,834/15,834 for Yelp2018 and 26,322/26,321 for Amazon-Book. Both stages exclude training and validation interactions from recommendation. The datasets informed method development before this evaluation protocol was fixed. Predictors and policy settings remain fixed during the reported evaluation. The expected-harm theorem additionally assumes exchangeability conditional on development outputs.

For each dataset, backbone and seed, candidate-specific ridge models are fitted only on validation labels. Their parameters are frozen before recalibration. Each of 34 non-Base candidates has gain-only, regret-only and learned harm/gain orders, yielding 102 separately calibrated thresholds. Thresholds are saved before evaluation labels are loaded. We retain Base and each candidate's ungated and three gated outcomes: 137 comparisons per run, 8,220 across three datasets, four backbones and five seeds. All candidates share Base Top-200; the TTEN-style comparison is not a full global reproduction.

Table 27: Unified-protocol RCPA results.

Strength	Ordering	PCE red.	H_R	H_N	Within
0.0175	Ungated	1.939	0.723	0.470	50/60
0.0175	Gain-only	1.654	0.462	0.320	56/60
0.0175	Regret-only	1.626	0.444	0.318	56/60
0.0175	Learned harm/gain	1.839	0.435	0.297	56/60
0.1	Ungated	9.115	3.236	1.611	16/60
0.1	Gain-only	4.802	0.880	0.415	46/60
0.1	Regret-only	3.182	0.912	0.438	41/60
0.1	Learned harm/gain	5.727	0.923	0.393	43/60

Table 27 reports the two strengths fixed before reevaluation. PCE reduction is a relative percentage, averaged equally over 60 runs; positive-part harms use 10^{-3} units, with budget 1 for each endpoint. Within counts evaluation runs satisfying both realized budgets. Runs share datasets and candidate variants, so grid rows are dependent.

At the conservative strength, learned ordering improves mean PCE reduction from 1.654% (gain-only) and 1.626% (regret-only) to 1.839%, while all three orders meet both realized budgets in 56 of 60 runs. As the seed-paired analysis below shows, this average improvement is concentrated in two Gowalla settings. At strength .1, learned ordering improves PCE reduction to 5.727% but meets both budgets in 43/60 runs, versus 46/60 for gain-only. Greater average alignment benefit therefore does not establish a better realized safety rate. Averaging across all 34 candidates gives a different estimand: the learned order's 76.47% grid satisfaction rate is not the conservative RCPA rate or a deployment recommendation.

Seed-paired primary ordering effects. Table 28 gives learned-minus-gain-only PCE differences at the predeclared strength. Two-sided 95% Student-t intervals use five paired seeds. These post-hoc nominal intervals condition on the fixed split, lack multiplicity correction and exclude design-selection and new-population uncertainty. Gains concentrate in the two Gowalla settings; other differences are near zero. Authorized and changed rates use all evaluation users as denominator.

Table 28: Primary learned-minus-gain-only PCE differences and authorization diagnostics.

Dataset	Backbone	PCE diff. (pp) [95% CI]	Admit (%)	Changed (%)	Within
Gowalla	BPR-MF	$+0.95569\ [+0.48245, +1.42893]$	10.4	10.4	3/5
Gowalla	LightGCN	$+1.28588\ [+0.80939, +1.76238]$	14.4	14.4	3/5
Gowalla	XSimGCL	$-0.00072\ [-0.00131, -0.00013]$	100.0	6.7	5/5
Gowalla	LightCCF	$-0.00169\ [-0.00417, +0.00079]$	100.0	6.6	5/5
Yelp2018	BPR-MF	$-0.00057\ [-0.00288, +0.00173]$	100.0	18.4	5/5
Yelp2018	LightGCN	$-0.00350\ [-0.01279, +0.00579]$	100.0	14.3	5/5
Yelp2018	XSimGCL	$-0.00028\ [-0.00058, +0.00001]$	100.0	4.4	5/5
Yelp2018	LightCCF	$-0.00148\ [-0.00597, +0.00301]$	100.0	7.1	5/5
Amazon-Book	BPR-MF	$-0.00037\ [-0.00109, +0.00035]$	100.0	23.3	5/5
Amazon-Book	LightGCN	$-0.00162\ [-0.00439, +0.00114]$	100.0	17.5	5/5
Amazon-Book	XSimGCL	$-0.00021\ [-0.00054, +0.00012]$	100.0	6.2	5/5
Amazon-Book	LightCCF	$-0.00026\ [-0.00102, +0.00049]$	100.0	5.3	5/5

O.1 Risk realizations and notation

Table 29 details the four conservative learned-order runs exceeding a realized budget. Each evaluates 14,929 Gowalla users. Calibration uses the finite-sample correction in Eq. (10); all four corrected calibration risks are below the budget. Evaluation Recall harms exceed it by 2.36–36.19%. NDCG meets its budget in these runs. This comparison is descriptive on the frozen partition and does not identify a cause or estimate a failure probability across independent datasets.

Table 29: Primary Recall-budget exceedances.

Model / seed	Cal. corrected	Eval. harm	Excess (%)	Harmed users	Changed harm
MF / 2022	0.985	1.093	9.32	70	8.099
MF / 2023	0.978	1.174	17.44	56	10.386
GCN / 2021	0.985	1.362	36.19	69	8.696
GCN / 2022	0.980	1.024	2.36	60	7.432

The harm columns use 10^{-3} units, with population budget 1. Harmed users counts positive Recall loss; Changed harm conditions on changed Top-20 membership. Changed-user means are 7.4–10.4 times the population budget. In each run, at least one user loses all held-out Recall (loss 1 on the observed labels); this metric alone does not measure that user's overall experience. These diagnostics motivate distinguishing population-level authorization from protection conditional on receiving an edit.

Current predictor ranking diagnostic. At $\gamma = .0175$, we audit frozen validation-fitted models on archived calibration labels. Model hashes match all 60 records. We invert the archived score formula to recover $\max(\widehat{h}_R, \widehat{h}_N)$, check forward reconstruction, and round to 10^{-8} to suppress sentinel-subtraction error. Throughout this diagnostic, the policy and thresholds remain fixed.

All-user harm-occurrence AUROC ranges from .751 to .973 across setting–endpoint means, but includes many zero-harm, unedited users. Table 30 restricts attention to changed lists with positive gain. We form gain quintiles without splitting ties and weight within-bin AUROC by the number of positive–negative pairs. All four signals use a fixed increasing-harm direction. This descriptive stratification does not establish causal or conditional independence claims.

Table 30: Harm-occurrence AUROC among changed, positive-gain candidates, stratified by gain quintiles. Entries average five seeds and show Recall/NDCG.

Data	Backbone	Predicted harm	Change fraction	Regret	Gain
G	MF	0.563/0.492	0.611/0.574	0.521/0.439	0.550/0.534
G	GCN	0.687/0.638	0.724/0.691	0.474/0.391	0.579/0.551
G	XSim	0.449/0.449	0.499/0.499	0.453/0.453	0.502/0.502
G	CCF	0.515/0.517	0.498/0.498	0.464/0.471	0.487/0.488
Y	MF	0.585/0.594	0.578/0.584	0.488/0.423	0.521/0.512
Y	GCN	0.731/0.758	0.654/0.642	0.390/0.354	0.554/0.522
Y	XSim	0.533/0.526	0.499/0.499	0.448/0.446	0.510/0.512
Y	CCF	0.683/0.679	0.625/0.628	0.318/0.299	0.601/0.580
A	MF	0.753/0.780	0.691/0.642	0.427/0.408	0.566/0.547
A	GCN	0.750/0.817	0.682/0.653	0.329/0.275	0.565/0.539
A	XSim	0.549/0.549	0.499/0.499	0.479/0.479	0.509/0.509
A	CCF	0.521/0.524	0.499/0.499	0.474/0.476	0.519/0.516

The learned signal improves on simple signals in several settings, but change fraction is stronger on Gowalla MF/LightGCN. A separate three-fold replay rotates calibration and target roles within each archived calibration pool, with fixed dual budgets .001. On Gowalla MF/LightGCN, learned ordering yields 3.228/2.319% PCE reduction, versus 3.114/2.145% for change-per-gain and 1.870/.831% for gain-only. The learned-minus-change-per-gain benefit is positive in all five seeds for both backbones. Discrimination alone therefore does not determine separately calibrated policy benefit. The complete 60-job replay retains both harms and all four orders; these means do not establish equal-risk superiority.

These are retrospective diagnostics with held-out labels, not unseen users or independent cohorts. The original models and policies remain fixed. `R51_mechanism_diagnostic.zip` contains all seed and stratum results; `R50_predictor_diagnostic.zip` retains the earlier unrestricted audit. Combined predicted harm is a ranking signal, not a calibrated endpoint estimate.

Table 31: Core notation.

Symbol	Meaning
B, π_u^B	Frozen backbone and its ordered Base recommendation list
B_m	Upper bound 1 on endpoint m's positive-part harm
q_R, q_N	Per-user Recall@20 and NDCG@20
$q_{u,\tau}$	Empirical degree quantile in a recommended list
Δ_u, g_u	Signed PCE reduction and its positive part, respectively
$a_{ui}(\gamma)$	Adjusted item score for candidate selection
$a(X), A_t(X)$	Fractional oracle authorization and implemented threshold authorization
$\mu_m, \widehat{h}_u^m$	True conditional expected harm and its fitted prediction
r_u^{regret}	Nonnegative loss of mean frozen score, normalized by within-pool scale
ρ_u, t	Frozen user ordering score and calibrated threshold
ε, C	Fixed score constants 10^{-8} and 10^6

Finite-score branch audit. The finite C does not force positive gains ahead of all non-positive gains. For $\widehat{h} = 1$ and $g = 10^{-12}$, the positive-branch ratio is about $9.999 \times 10^7 > C + 1$. The additive regret term can also distinguish close scores beyond exact ties. Across 60 primary runs, 1,904 of 1,141,680 user–run rows worsen PCE (0.167%); 132,201 change membership, giving 1.44% worsened among changed rows. Five positive-gain scores exceed $C = 10^6$. Their gains are 1.73×10^{-18}–6.94×10^{-18}, and all are rejected by the deployed thresholds. Repeated users across seeds and backbones make these descriptive counts.

Recalibrated numerical sensitivity. Sixty frozen-model calibration pools use three disjoint folds rotating calibration and target roles. Each user is evaluated once per policy. Score inversion exactly reconstructs the archived scores and baseline replay. Seven one-factor alternatives use $\varepsilon \in \{10^{-9}, 10^{-7}\}$, $C \in \{10^4, 10^8\}$, or regret coefficients 0, 10^{-10} and 10^{-8}. The eighth uses lexicographic (unperturbed score, regret) ordering with complete equal-key atoms. The candidate and dual budgets .001 stay fixed; thresholds are recalibrated.

Mean PCE reduction is 1.82273–1.82275% versus baseline 1.82275%. The largest absolute seed-level change is $.001193$ percentage points; authorization changes for at most $.01895\%$ of a job's target users. Both endpoint harms and budget indicators remain identical in all 180 target folds, of which 164 meet both budgets. These historical replays support stability over the tested range without retuning the primary policy. `R54_closing_diagnostics.zip` retains all 1,800 fold-policy rows, including reconstruction controls.

P GOWALLA PARTITION ROBUSTNESS

The audit fixes the primary RCPA strength $\gamma = .0175$, learned harm/gain order, validation-fitted predictors, frozen backbones and both budgets at $.001$. Partition seeds 2027, 2029, 2031, 2033 and 2035 each divide the same 29,858 users into 14,929 calibration and 14,929 evaluation users. All 100 backbone–model-seed–partition cells are retained. Seed 2027 reproduces all 20 original Gowalla primary rows; the other four partitions add 80 cells. Users and models recur across partitions, so 79/100 is a descriptive frequency, with dependence that precludes an ordinary independent-binomial confidence interval.

For every cell, thresholds were recomputed from calibration arrays, all saved summary rows were reconstructed from evaluation arrays, and saved data and model hashes were checked. All 100 corrected calibration risks meet both budgets. Evaluation Recall exceeds its budget in 21 cells; NDCG meets its budget in all 100 cells. Across partitions, dual-budget satisfaction is 16/20, 17/20, 14/20, 15/20 and 17/20. The 2027 cells overlap the primary experiment and must not be added to it as independent evidence.

Table 32: Partition audit by backbone.

Backbone	Within	Cal. H_R^+	Eval. H_R	Eval. H_N	Changed (%)	Cond. H_R
BPR-MF	14/25	0.976	0.945	0.386	11.60	8.027
LightGCN	15/25	0.981	0.947	0.577	14.19	6.643
XSimGCL	25/25	0.350	0.239	0.073	6.69	3.583
LightCCF	25/25	0.402	0.373	0.115	6.61	5.650

Harms use 10^{-3} units; the population budget is 1. Cal. H_R^+ includes the finite-sample correction. Cond. H_R averages the per-cell mean over users whose candidate is authorized and whose Top-20 membership changes after gating. Rejected candidates return exactly to Base and are excluded from this conditional population. All displayed means weight the 25 cells equally.

BPR-MF/LightGCN operate close to the corrected calibration budget (.976/.981), with evaluation Recall ranges .442–1.530/.534–1.362. XSimGCL/LightCCF have more slack (corrected means .350/.402) and lower maxima (.412/.585). The association is compatible with finite-sample fluctuation; these dependent partitions cannot diagnose a causal mechanism or distribution shift.

The ten largest Recall losses account for 52.0/50.7% of total harm in BPR-MF/LightGCN, versus 97.2/85.3% in XSimGCL/LightCCF. Concentration alone therefore does not explain exceedance: the latter models have lower total harm. The complete 100-cell matrix is retained in `R51_mechanism_diagnostic.zip`; Figure 3 displays all cells.

Q COMMON RAW-SCORE AND WASSERSTEIN AUDIT

To isolate score-path effects, this historical calibration-pool audit uses raw dot products and item-ID ties, with Top-200 supplying candidates and Top-20 defining Base. Under this common scorer, zero-strength TTEN reproduces Base and positive-strength TTEN stays within Top-200. The audit complements the unified evaluation using a separate historical pool.

We measure W_1 between empirical training-history and list distributions of $\log(1 + d_i)$. Identical distributions have zero distance. Relative W_1 and PCE reductions use different denominators.

Exact-tail keeps the first 19 items and minimizes W_1 over position-20 replacements within Top-200, including the original item. Gains at most 10^{-12} return Base. The control instead picks the lowest-

degree eligible tail item with the same guard. Both use gain-only authorization and dual budgets .001.

Three disjoint folds rotate calibration/target roles, evaluating each user once. Across three datasets, four backbones and five seeds, exact-tail/control mean reductions are 4.813/4.669%. Exact-tail leads in seven of twelve settings, with different harms. Both budgets hold in 116/180 folds and 58/60 pooled runs. The pooled Recall exceedances are Gowalla XSimGCL seed 2024 (.001182829) and Yelp2018 BPR-MF seed 2022 (.001028040). Pooling combines three thresholds.

Exact-tail random/gain ordering gives 2.960/4.813% reduction, Recall harm .905/.907 and NDCG harm .309/.312 (10^{-3} units). Gain ordering improves benefit in all twelve settings, but increases neither harm in only five. Relative to the low-degree control, gains average .591, .054, -.013 and -.055 percentage points on LightCCF, XSimGCL, BPR-MF and LightGCN, respectively. Minimizing candidate distance therefore need not improve benefit after calibration. These reused-pool comparisons are diagnostic.

Historical score-path discrepancy. A LightGCN seed-2020 export found 177/320/1,190 zero-strength membership changes on Gowalla/Yelp2018/Amazon-Book. All 1,687 records have equal floating-point sigmoid scores but unequal raw scores. This identifies a mismatch mechanism, not its effect on positive-strength comparisons. The common-raw audit defines a separate Base; historical results retain their implementation identity.

Same-candidate metric sensitivity. To isolate metric sensitivity, we recompute only W_1 for frozen RCPA lists across three datasets, four backbones, five seeds and 100 partitions per job. Ordering and harm calibration remain fixed. Users are disjoint within partitions but recur across repetitions, so Table 33 averages partitions within job before weighting backbones and seeds equally.

The learned-order advantage is concentrated in Gowalla, with tiny learned/gain-only differences elsewhere. Amazon-Book regret-only instead combines smaller harms with less distance reduction, illustrating why these means alone establish neither significance nor equal-risk superiority.

Table 33: Historical same-candidate W_1 reduction (%, higher is better).

Dataset	Learned	Gain-only	Regret-only
Gowalla	0.9454	0.5723	0.7300
Yelp2018	0.8933	0.8939	0.8827
Amazon-Book	1.1415	1.1398	0.7472

Matched candidate grid. A historical LightGCN replay fixes raw Base, Top-200, item-ID ties, restored order and gain-only authorization. Three datasets, five seeds and three rotating calibration/target folds use budgets .0005, .001 and .002. The 24 candidates comprise 2 RCPA, 17 CP and 5 TTEN variants. Six high CP weights were added uniformly after the initial grid produced little editing; all results are retained without selecting a deployment policy.

Table 34 reports five-seed means; harms use 10^{-3} units. "Dominated" counts budgets with an alternative improving at least one of PCE benefit and the two harms without worsening another. Conservative RCPA is non-dominated in all nine dataset–budget means; strong RCPA is dominated on Gowalla at .0005 and .001. Adding changed-list rate leaves these classifications unchanged, but neither criterion implies optimality. All 216 means and 3,240 fold-policy records are retained in `R51_matched_candidates.zip`, with three- and four-criterion checks in `R57_review_evidence.zip`.

R Discrete objectives and positive-gain filtering

A discrete counterexample. Suppose all historical and Base items have degree d. Their empirical distributions coincide, but the right-continuous historical CDF gives $\widehat{F}(d) = 1$ and PCE is 0.35. Replacing half of the 20 Base items by degree-$d-1$ items yields PCE 0.0772727 under the empirical inverse-CDF convention. Thus a smaller PCE can accompany a departure from an exactly matching

Table 34: RCPA in the common-raw grid; dominance uses benefit and two harms. Displayed budgets are .001.

Dataset	γ	PCE ↓(%)	H_R	H_N	Changed (%)	Dominated
Gowalla	0.0175	0.722	0.227	0.075	7.28	0/3
Gowalla	0.1	3.632	0.901	0.282	20.92	2/3
Yelp2018	0.0175	0.541	0.173	0.061	5.33	0/3
Yelp2018	0.1	3.033	0.850	0.307	25.22	0/3
Amazon-Book	0.0175	0.513	0.130	0.048	6.31	0/3
Amazon-Book	0.1	2.985	0.723	0.269	32.05	0/3

distribution. The median-CDF direction signal also equals 0.5 for the matching Base, so ties affect both the objective and the correction direction.

Evaluation of the same historical candidates. The original LightGCN candidates and frozen PCE-based orders are evaluated by log-degree W_1 on historical calibration pools. Five seeds each use 100 shared partitions. Learned/gain-only/regret-based reductions are $1.2254/0.4233/0.7038\%$ on Gowalla, $0.8521/0.8521/0.8515\%$ on Yelp2018, and $1.0715/1.0729/1.0715\%$ on Amazon-Book. This LightGCN diagnostic retains the Gowalla ordering benefit; it is distinct from the four-backbone unified evaluation. Candidate user–seed records with better PCE but worse W_1 account for $0.2505/0.0183/0.0277\%$, respectively. Such disagreement need not arise from degree ties.

Fixed-threshold filtering. Let $\Delta(X) = \mathrm{PCE}(\pi^B) - \mathrm{PCE}(\pi^C)$ and retain the original calibrated threshold. Define

$$A'_t(X) = A_t(X)\mathbf{1}\{\Delta(X) > 0\}.$$

For every observation and each endpoint, $A'_t H_m \le A_t H_m$ because $H_m \ge 0$. Also $A'_t\Delta \ge A_t\Delta$: removed edits have nonpositive signed gain. Taking expectations preserves any valid risk upper bound for the original policy. The filter uses histories and lists, without relevance labels. This pointwise statement concerns the specified objective and holds without recalibrating the threshold.

Table 35 holds the learned-order thresholds fixed. Means average 100 partitions per seed, then five seeds. Reductions are relative percentages with metric-specific Base denominators; harms use 10^{-3} units. Both rates use all target users as denominator: Admit includes unchanged lists, whereas Changed counts only authorized membership changes. The filtered development policies do not replace primary results.

Table 35: Fixed-threshold positive-PCE filtering.

Dataset	Policy	PCE ↓	W1 ↓	H_R	H_N	Admit	Changed
Gowalla	Original	2.35245	1.22543	0.8769	0.5221	13.51	13.35
Gowalla	Positive PCE	2.35246	1.22515	0.8767	0.5219	13.35	13.35
Yelp2018	Original	1.55614	0.85207	0.5254	0.3178	97.29	14.16
Yelp2018	Positive PCE	1.55782	0.80777	0.4726	0.2568	12.60	12.60
Amazon-Book	Original	1.60759	1.07148	0.4252	0.4049	99.76	17.59
Amazon-Book	Positive PCE	1.61936	1.00100	0.3748	0.2985	15.25	15.25

Filtering improves PCE and reduces both harms, but also reduces mean W1 benefit on all three datasets. This tradeoff reflects the objective-specific nature of the pointwise benefit guarantee. Intersecting positive-PCE and positive-W1 masks can likewise remove edits that benefit one metric, so aggregate improvement on both is not assured.

Unchanged membership accounts for 83.13/82.18 percentage points of the authorization-rate drops on Yelp2018/Amazon-Book. Changed-list rates fall by only 1.56/2.34 points, versus 0.00683 on Gowalla. These calculations use the same 15 caches and 100 partitions per seed, reproducing original authorization means to 10^{-12}.

Observed degree atoms. The repeated-degree fraction $1 - |\{d_i : i \in \mathcal{H}_u\}|/n_u$ averages 16.26%, 11.70% and 17.82% over users and five seeds on Gowalla, Yelp2018 and Amazon-Book. A CDF jump occurs at the Base median degree in 8.56%, 7.51% and 10.97% of user–seed records, respectively.

These frequencies measure exposure to discrete atoms, not the fraction of metric improvements caused by ties.

S COMMON-SCORE ALLOCATION WITH A WASSERSTEIN OBJECTIVE

This exploratory control tests allocation when the benefit is a distribution distance rather than PCE. It reuses historical LightGCN validation and calibration pools on three datasets and five seeds. Common raw Top-200 scores and item-ID tie-breaking define Base and all edits. Candidates are raw RCPA at .0175 and the exact-tail edit from Appendix Q. The existing .1 RCPA strength was added in a declared follow-up after the conservative candidate showed little selection; it was not part of the initial grid.

All policies retain only strictly positive log-degree W_1 gains. Gain-only, change-per-gain, regret-per-gain and learned harm-per-gain orders share three rotating, user-disjoint calibration/target folds and budgets .0005, .001 and .002. Each user is evaluated once per policy. Separate ridge harm predictors, with penalty 10, are fitted on validation outcomes only. Their eight inputs are log-history size, mean historical log-degree, Base and candidate W_1, signed gain, changed fraction, position-change fraction and normalized raw-score regret. These are diagnostic predictors, not replacements for the primary 14-feature models.

Table 36 compares learned ordering with change-per-gain at .001. Differences first pool disjoint target folds within each seed, then average five seeds. Alignment entries are differences in relative-reduction percentage points; positive values favor learning. Harm differences use 10^{-3} units, where negative values favor learning. PCE and position-weighted CP-group JSD are recomputed for the same authorized lists, using the definitions in Appendix E.

Table 36: Learned minus change-per-gain under a common W_1 objective.

Candidate	Data	ΔW_1	ΔPCE	ΔJSD	ΔH_R	ΔH_N	Positive seeds
RCPA .0175	A	0.0000	0.0000	0.0000	0.0000	0.0000	0/5
RCPA .0175	G	0.0000	0.0000	0.0000	0.0000	0.0000	0/5
RCPA .0175	Y	0.0000	0.0000	0.0000	0.0000	0.0000	1/5
RCPA .1	A	0.0000	0.0000	0.0000	0.0000	0.0000	0/5
RCPA .1	G	0.0382	0.0627	0.0742	0.0792	0.0221	3/5
RCPA .1	Y	0.0007	0.0013	0.0021	0.0058	0.0026	2/5
Exact-tail	A	0.2895	1.7587	0.9794	-0.0054	0.0323	5/5
Exact-tail	G	0.2158	1.8178	0.7125	0.0513	0.0394	5/5
Exact-tail	Y	0.2243	1.6194	0.8281	0.0273	0.0407	4/5

Here G/Y/A denote Gowalla/Yelp2018/Amazon-Book. Positive seeds count strictly positive W_1 differences above 10^{-12}; a rounded zero can therefore have a positive count. Conservative RCPA leaves little room for ordering to matter. At strength .1, the small mean gains on Gowalla and Yelp2018 accompany higher mean harms. Exact-tail retains additional benefit, but its changed fraction is always $1/20$ for an edit, so change-per-gain and gain-only induce the same order on positive gains. This comparison does not demonstrate an advantage over a distinct change-based heuristic.

The complete three-budget results, absolute outcomes, conditional harms and all seed differences are retained in `R57_review_evidence.zip`. For example, exact-tail's Gowalla learned-minus-change W_1 difference becomes $-.0340$ points at budget .002. Thus increased retained benefit is a budget-dependent tradeoff, not consistent joint-risk dominance. The shared score path resolves numerical comparability within this control, but neither these reused pools nor the diagnostic predictors provide independent confirmation of the primary experiment or repair its historical TTEN path.