GSRT: GRANULARITY-SELECTED RESIDUAL TRANSFER FOR SPARSE PRODUCTION TARGETS IN TWO-TOWER RETRIEVAL

Anonymous authors
Paper under double-blind review

ABSTRACT

Rare target events invite transfer from abundant auxiliary feedback, but the granularity of that transfer is an open choice. One global coefficient ignores heterogeneity, while per-user coefficients overfit sparse labels. We study this choice for transferring consumption feedback to sparse creator actions in two-tower retrieval. Under a local risk approximation we characterize the supportable granularity by an information budget, the heterogeneity of the transfer relation times the Fisher information the target labels carry about the transfer coefficient. Above a criterion, group-level coefficients have lower predicted risk than a global one, and a detection lower bound shows that far below it no procedure can exploit group differences. Guided by this analysis we propose Granularity-Selected Residual Transfer (GSRT), which adds a source-derived residual to a frozen target-only predictor, selects among no transfer and global, group and instance-level coefficients by cross-fitted held-out loss, and serves transfer only when an evidence gate rejects no transfer. Because the coefficient is user-side, the score keeps an exact inner-product factorization, so one approximate-nearest-neighbour index per target serves it. On a 4,800-cell controlled benchmark the empirical boundary falls between the two thresholds and the gate cuts cells with negative transfer from 30% to 5%. On production logs of a short-video platform, where the target model sees fewer than 100 positives, selecting the coefficient for ranking and calibrating probabilities separately raises recall inside a 1% exposure quota by 0.062 without increasing Brier loss, and the gain vanishes as the target model is given more labels. On public logs where no transferable signal is detected, GSRT returns no transfer with no significant loss while forced transfer loses.

1 INTRODUCTION

Recommender systems on user-generated-content platforms allocate attention to existing content, and the supply of that content depends on the same users deciding to create. Production-oriented retrieval therefore has to predict rare creator actions, such as uploads and entries into creation tools, from orders of magnitude fewer positives than consumption provides. The two behaviours are related but distinct. At the anonymised short-video platform that motivates this work, three online tests raised trend-content plays by 0.354–1.774% while upload-author counts rose by at most 0.159% (Appendix H.1). A consumption predictor thus carries information about production without measuring readiness to produce. This paper asks how much of that information a production predictor should use, and at what resolution, from one transfer coefficient shared by all users to one per user.

Neither end of that range is safe. A single coefficient discards whatever heterogeneity exists, and a coefficient per user spends a few hundred positives on estimates that are mostly noise. Knowing the true structure would not settle the choice either, because every additional coefficient is paid for in target labels that the rare event does not supply, so a relation that varies by user may still be affordable only at group resolution. Which structure holds is hard to establish in any case, since the variable that governs it, such as readiness, intention, or attainability, is often unobserved. Shared-expert architectures (Ma et al., 2018a; Tang et al., 2020; Li et al., 2023) learn where to share inside a jointly trained network, but their objectives do not say at which resolution the routing can be estimated, and auxiliary gradients can oppose the sparse target's (Du et al., 2018; Yu et al., 2020; Wang et al., 2019).

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.

Retrieval adds a constraint of its own. Candidate-generation scores must factorize into independently computed query and item vectors for ANN search (Covington et al., 2016; Malkov & Yashunin, 2020), so a controller that uses user–item cross features cannot be served.

We treat the choice as model selection under limited evidence. GSRT (Granularity-Selected Residual Transfer) keeps a target-only production predictor and adds a consumption residual weighted by a user-side coefficient, chosen from a hierarchy that runs from no transfer through one global value and pooled group values to an instance-level gate. How fine a rule the labels can support is governed by an *information budget* $h^2\mathcal{I}$, the heterogeneity of the relation times the information the target rows carry about the coefficient. We evaluate on two million users of that platform's production logs, on a controlled benchmark whose governing variable is known, and on public logs. Our main contributions are:

- **An information budget for resolution.** Under a local quadratic surrogate with near-constant event probability, group coefficients have lower predicted risk than one global coefficient when $h^2\mathcal{I} \gtrsim G - 1$, and in a two-group detection problem with $h^2\mathcal{I} \leq \frac{1}{4}$ no procedure recovers more than half of their advantage. The criterion is the same for the Brier score and the log loss (Section 3.3).
- **GSRT.** The procedure selects the resolution on held-out users, serves transfer only on significant evidence, and folds the chosen user-side coefficient into the query vector so that one ANN index per target serves it exactly (Sections 3.2 and 3.4).
- **Ordering and calibration selected apart.** Because the folded coefficient alone fixes the retrieved set, ordering and calibration can be selected by different losses. On the platform's production logs, at an upload rate of 7.0×10^{-5}, the squared loss selects a *negative* coefficient in 59 of 60 cells, while selecting the coefficient for ranking raises recall inside a global 1% exposure quota by 0.062 on four disjoint slices of the training stream (Section 4.5).
- **The boundary and the gate, measured.** On a 4,800-cell benchmark with known ground truth the empirical boundary falls between the two thresholds, and the evidence gate cuts cells with negative gain from 30% to 5% (Sections 4.1–4.4).

On public short-video logs, where no transferable click signal is detected, GSRT returns no transfer with no significant loss while forced transfer loses (Section 4.6). For a platform, GSRT turns sparse creator labels into a measurable gain wherever they can support one, declines transfer where they cannot, and runs inside existing two-tower retrieval.

2 Related Work

Model selection under transfer. The question of how much an auxiliary source should influence a target learner has a statistical answer in the bias–variance trade-off of borrowing (Caruana, 1997; Maurer et al., 2016) and a practical one in held-out selection (Cawley & Talbot, 2010), and sample-variance penalization gives data-dependent bounds for choosing among candidates (Maurer & Pontil, 2009). Negative transfer has been characterized as source–target mismatch (Wang et al., 2019) and countered by selecting source observations (Lu et al., 2013), grouping tasks (Fifty et al., 2021), weighting auxiliary gradients by similarity (Du et al., 2018), or balancing losses (Liu et al., 2019). Hypothesis transfer learning chooses how far to lean on a source hypothesis from few target labels and bounds the risk of that choice (Kuzborskij & Orabona, 2013; Aghbalou & Staerman, 2023). Transfer in high-dimensional regression detects which sources are informative, with minimax rates in the contrast between source and target (Li et al., 2022; Tian & Feng, 2023; Hanneke & Kpotufe, 2019). Pretest estimators (Bancroft, 1944) and partial pooling (Efron & Morris, 1973; Gelman & Hill, 2007) are the classical forms of our evidence gate and of our group shrinkage. None of these asks at what *resolution* the target labels can support a transfer rule, which is the question we study.

Adaptive sharing in multi-task networks. Adaptive mixtures of experts (Jacobs et al., 1991), cross-stitch units (Misra et al., 2016), MMoE (Ma et al., 2018a), PLE (Tang et al., 2020), AdaTT (Li et al., 2023), STEM (Su et al., 2024), and multitask video ranking (Zhao et al., 2019) learn where sharing happens inside a jointly trained network. MetaBalance (He et al., 2022), DRGrad (Liu et al., 2025), PCGrad (Yu et al., 2020), CAGrad (Liu et al., 2021), and multi-objective formulations (Şener & Koltun, 2018) change how auxiliary gradients enter optimization. They choose a routing parameterization or an optimizer rather than the resolution that a few hundred positives can support, and a gate

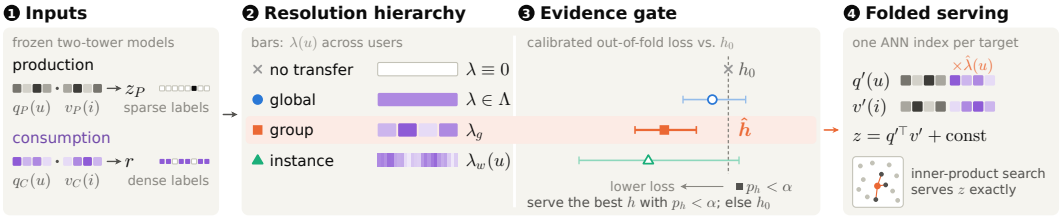


Figure 1: GSRT. From left, the frozen target-only and consumption two-tower models supply the base logit b and the residual r. The candidate hierarchy runs from no transfer through a global coefficient and pooled group coefficients to an instance gate. Cross-fitted out-of-fold losses are compared with no transfer by the evidence gate (schematic), and the selected user-side coefficient is folded into the query so that one ANN index per target serves the score (Sections 3.2 and 3.4).

on user–item cross features cannot be indexed, although two-tower variants can (Section 4.6). On the public logs we compare with two-tower MMoE and PLE under the same protocol (Section 4.6).

Sparse targets, calibration, and retrieval. ESMM borrows from denser funnel events when the event sequence is mandatory (Ma et al., 2018b), which production actions are not, and cross-domain recommendation transfers representations between domains that share users (Zhu et al., 2021). Rare-event regression needs probability correction (King & Zeng, 2001), delayed outcomes make recent negatives premature (Chapelle, 2014), focal loss reweights without calibrating (Lin et al., 2017), downsampling shifts posterior odds (He et al., 2014), and nonuniform retention needs a log-odds correction (Wang et al., 2021). Two-tower retrieval adds item-frequency bias from in-batch negatives (Yi et al., 2019; Yang et al., 2020), and InfoNCE learns a density-ratio critic rather than an event probability (van den Oord et al., 2018).

Creator ecosystems. Recommendation changes what producers make (Jagadeesan et al., 2023; Hron et al., 2023; Perdomo et al., 2020), and the closest industrial work co-optimizes generation and consumption in ranking (Zhang et al., 2024).

3 Method and Theory

3.1 Problem

Let $u \in \mathcal{U}$ index users, i items, C a dense consumption task, and P_k, $k = 1, \dots, K$, sparse production targets (Appendix A lists the notation). A target-only two-tower model scores $z_{P_k}(u,i) = q_{P_k}(u)^\top v_{P_k}(i) + b_k$ and predicts $\sigma(z_{P_k})$. Candidate generation requires v_{P_k} to be independent of u so that one ANN index per target can be built offline (Covington et al., 2016). The objective is expected Brier risk $R(\hat{p}) = \mathbb{E}[(\hat{p} - p^\star)^2]$ on the unresampled target distribution, with $p^\star = P(Y_{P_k} = 1 \mid u, i)$ the observable Bayes probability, the conditional probability given everything the predictor sees. This is the quantity a retrieval stage is built on. The downstream policy question, a user-level production outcome under a consumption guardrail, is left to the randomized test specified in Appendix H.5.

3.2 Residual Transfer with a Selected Resolution

GSRT starts from a fitted target-only logit z_{P_k} and a consumption-derived residual $r_k(u,i)$, the consumption model's bilinear score without its intercept, and scores

$$z_k^{\mathrm{GSRT}}(u,i) = z_{P_k}(u,i) + \lambda_k(u)\, r_k(u,i) + c_k \tag{1}$$

with a user-side coefficient $\lambda_k(u)$ and a calibration offset c_k. The coefficient trades the variance of the sparse target fit against the discrepancy between the two tasks (Proposition B.1, Appendix B.1), and both vary across users. Setting $\lambda_k \equiv 0$ recovers the target-only model, so every family below contains no transfer as a special case, and deciding how much to borrow becomes a choice of resolution. Write $b = z_{P_k}$, $r = r_k$, $\Lambda = \{-1, -0.9, \dots, 1\}$, and let $\mathcal{D}_{\mathrm{ctrl}}$, $\mathcal{D}_{\mathrm{sel}}$, $\mathcal{D}_{\mathrm{cal}}$ be disjoint sets of target users. GSRT fits four families on a set $\mathcal{D}$ of controller users (one of the two folds $\mathcal{D}_{\mathrm{ctrl}}$, $\mathcal{D}_{\mathrm{sel}}$, or their union):

- *No transfer* h_0: $\lambda \equiv 0$.
- *Global*: $\lambda^{\mathrm{glob}} = \arg\min_{\lambda \in \Lambda} \frac{1}{|\mathcal{D}|} \sum_{(x,y)\in\mathcal{D}} \big(\sigma(b + \lambda r) - y\big)^2$.

- *Group, pooled by information*: with λ_g^{loc} the grid minimizer within group g, $\mathcal{I}_g = \sum_{t \in g} \hat{p}_t(1 - \hat{p}_t)\, r_t^2$ the group's information for the coefficient ($\hat{p}_t = \sigma(b_t)$), and $\bar{\iota}$ the information per positive over all groups,

$$\lambda_g^{(s)} = \lambda^{\mathrm{glob}} + \frac{\mathcal{I}_g}{\mathcal{I}_g + s\,\bar{\iota}}\big(\lambda_g^{\mathrm{loc}} - \lambda^{\mathrm{glob}}\big), \qquad s \in \{5, 20, 100\}. \tag{2}$$

  When information is proportional to positives this is the familiar count-based pooling $n_g^+/(n_g^+ + s)$, and when it is not, groups with the same positives but weaker residuals are pooled harder (Section 3.3).
- *Instance*: a bounded linear gate $\lambda_w(u) = \tanh\big(w^\top[1; u; \mathrm{onehot}(g)]\big)$ with w minimizing the controller Brier loss plus $\rho\|w\|^2$, $\rho \in \{0.1, 1, 10\}$.

The hierarchy is $\mathcal{H}_k = \{h_0, \lambda^{\mathrm{glob}}, \lambda^{(5)}, \lambda^{(20)}, \lambda^{(100)}, \lambda_{w_{0.1}}, \lambda_{w_1}, \lambda_{w_{10}}\}$ (Fig. 1, with exact objectives and settings in Appendix E.2 and Table 3). Every candidate $h \in \mathcal{H}_k$ first receives its own offset c_h on $\mathcal{D}_{\mathrm{cal}}$ so that its mean prediction matches the observed rate. Selection is then cross-fitted and gated. Every family is fitted on each of the two folds $\mathcal{D}_{\mathrm{ctrl}}$, $\mathcal{D}_{\mathrm{sel}}$ and scored on the other, so that every controller user j contributes one out-of-fold Brier loss $\ell_h(j)$ per candidate (the mean of $\big(y - \sigma(b + \lambda_h(u)r + c_h)\big)^2$ over the user's rows). With $\bar{\ell}_h$ the mean out-of-fold loss and p_h the one-sided paired t test of $\ell_h - \ell_{h_0}$ over users, GSRT serves

$$\widehat{h}_k = \arg\min\big\{\bar{\ell}_h :\ h = h_0 \ \text{ or } \ p_h < \alpha\big\}, \qquad \alpha = 0.05, \tag{3}$$

that is, the candidate with the smallest out-of-fold loss among those that beat no transfer significantly, or no transfer if none does. We call the condition $p_h < \alpha$ the evidence gate. Here α is a per-candidate level, and the conservative rule below is its family-wise counterpart. The served candidate is then refitted on both folds and recalibrated, so selection and serving share the candidate but not its fitted coefficients. Three rules on the same losses serve as comparators: the ungated argmin $\arg\min_h \bar{\ell}_h$, the one-standard-error rule (the simplest candidate within one paired standard error of the best), and a conservative rule that minimizes the Bonferroni-adjusted one-sided upper confidence limit of the paired excess over h_0. A certified variant, the t certificate, repeats the test on certification users disjoint from every other role and serves $\widehat{h}_k$ only when the one-sided $1 - \delta$ upper limit of its paired excess over h_0 is nonpositive ($\delta = 0.05$, Appendix D).

3.3 The Information Budget of Resolution

Let $z^\star$ be a logit of $p^\star$, $\varepsilon = z^\star - b$ the ideal correction, and $w = \sigma'(z^\star)^2$. A second-order expansion of the Brier risk of $\sigma(b + \lambda(X)r)$ around $z^\star$ is the weighted quadratic surrogate $Q(\lambda) = \mathbb{E}[w(\varepsilon - \lambda(X)r)^2]$ (Appendix B.4). The statements below concern Q, and the benchmark measures how closely they track the exact risk.

Proposition 3.1 (Weighted resolution). *Let $\varepsilon = \alpha(G)\, r + \nu$ with $\mathbb{E}[w\nu \mid X, G, r] = 0$, where X are the controller's inputs and G the variable governing the relation, and let $m_g = \mathbb{E}[wr^2 \mid G = g]$. (a) If G is a function of X with values g of probability π_g, the best constant coefficient is $\bar{\lambda} = \sum_g \pi_g m_g \alpha_g / \sum_g \pi_g m_g$ and its excess surrogate risk over the best group controller is $A = \sum_g \pi_g m_g (\alpha_g - \bar{\lambda})^2$. (b) If $G = H$ with H independent of (X, r), every controller $\lambda(X)$ satisfies $Q(\lambda) - \min Q \geq \mathbb{E}[wr^2]\,\mathrm{Var}(\alpha(H))$, with equality only for the constant $\mathbb{E}[\alpha(H)]$.*

Part (a) measures the loss of a single coefficient by the residual-weighted spread of the group relations. Part (b) says that when the governing variable is hidden, no controller on the observed inputs has a first-order advantage over the constant. A second-order term remains, because the observable Bayes probability $\mathbb{E}_H[\sigma(b + \alpha(H)r)]$ is a mixture, but it stays below the estimation cost of every non-constant family throughout the benchmark (Section 4). A hidden governing variable is therefore a regime in which the right answer is one coefficient or none.

Proposition 3.2 (Cost of resolution). *Let a family λ_θ, $\theta \in \mathbb{R}^d$, be fitted by minimizing the empirical Brier loss on n i.i.d. controller rows, with $\theta^\star$ the interior population minimizer and $\phi = \partial\lambda_\theta/\partial\theta$ at $\theta^\star$. Under the standard regularity conditions of M-estimation, $\mathbb{E}[R(\hat{\theta})] - R(\theta^\star) = \frac{1}{2n}\,\mathrm{tr}(H^{-1}V) + o(1/n)$, with H the population Hessian and V the covariance of the per-row gradient at $\theta^\star$. In the Gauss–Newton approximation, exact when the family contains $z^\star$, $H = 2\mathbb{E}[\sigma'^2 r^2 \phi\phi^\top]$ and*

$V = 4\mathbb{E}[p^\star(1-p^\star)\sigma'^2 r^2 \phi\phi^\top]$, *and the predicted costs are* $C_{\mathrm{glob}} = \frac{1}{n}\mathbb{E}[p^\star(1-p^\star)\sigma'^2 r^2]/\mathbb{E}[\sigma'^2 r^2]$ *for the global family and* $C_{\mathrm{grp}} = \frac{1}{n}\sum_g \mathbb{E}[p^\star(1-p^\star)\sigma'^2 r^2 \mathbf{1}_g]/\mathbb{E}[\sigma'^2 r^2 \mathbf{1}_g]$ *for unpooled group coefficients, so that with near-constant* $p^\star$ *each additional coefficient costs about one Bernoulli variance* $\bar{p}(1-\bar{p})/n$.

The Gauss–Newton forms are the computable ones, and the benchmark reports them next to the realized excess of every family (Appendices B.5 and F). Combining the two propositions gives the criterion (Appendix B.6).

Corollary 3.3 (Supportable resolution)**.** *Group resolution has lower predicted risk than a global coefficient iff* $A > C_{\mathrm{grp}} - C_{\mathrm{glob}}$. *With near-constant* $p^\star$, *the coefficient information* $\mathcal{I} = \sum_{\mathrm{rows}} p^\star(1-p^\star)r^2$ *and the residual-weighted heterogeneity* $h^2 = \sum_g \pi_g \mathbb{E}[r^2 \mid g](\alpha_g - \bar{\lambda})^2/\mathbb{E}[r^2]$, *the criterion for* G *groups reads*

$$h^2\,\mathcal{I} \;\gtrsim\; G-1. \tag{4}$$

Eq. (4) is the classical comparison between separate and pooled estimates of G normal means (Efron & Morris, 1973), where separate estimates have lower total risk when the spread of the means exceeds $(G-1)/G$ times the sampling variance of each estimate. Here the spread is h^2 and, for groups of equal information, the sampling variance of each group coefficient is $G/\mathcal{I}$. The positive count enters Eq. (4) only through $\mathcal{I} \approx n\bar{p}(1-\bar{p})\mathbb{E}[r^2]$, so the same positives support finer resolution when the residual is stronger and coarser resolution when it is weaker. With near-constant $p^\star$ the threshold does not depend on the scoring rule. Under log loss the curvature weight of Proposition 3.1 changes from σ'^2 to σ', the information equality makes each additional coefficient cost $1/(2n)$ instead of $\bar{p}(1-\bar{p})/n$, and the two changes cancel (Appendix B.7). $\mathcal{I}$ is the Fisher information the target rows carry about the coefficient. Selecting among the $|\mathcal{H}_k|$ calibrated candidates on n out-of-fold users adds at most $O(\sqrt{\log|\mathcal{H}_k|/n})$ (Appendix D). A converse holds for a detection surrogate whose weights are fixed at the global logit. Below the same budget, up to a constant, no procedure can exploit the finer resolution.

Theorem 3.4 (Detection lower bound)**.** *Consider two groups of equal mass with base and residual fixed,* N *target rows with labels* $y_t \sim \mathrm{Bern}(\sigma(b_t + \alpha_{g_t} r_t))$, *and the two hypotheses* $\alpha = (\bar{\lambda}+h, \bar{\lambda}-h)$ *and* $\alpha = (\bar{\lambda}-h, \bar{\lambda}+h)$. *Let* $w^0 = \sigma'(b+\bar{\lambda}r)^2$ *be the surrogate weight at the global logit,* $m_g^0 = \mathbb{E}[w^0 r^2 \mid G=g]$, $\bar{m}^0 = \sum_g \pi_g m_g^0$, *and* $Q_s^0(\lambda) = \sum_g \pi_g m_g^0(\lambda_g - \alpha_g^s)^2$ *the detection surrogate of a group-constant controller under hypothesis* s *(a general* $\lambda(X)$ *enters through its* $w^0 r^2$*-weighted group means). For any procedure that outputs a user-side controller* $\hat{\lambda}$ *from the data,*

$$\max_s \mathbb{E}_s\big[Q_s^0(\hat{\lambda})\big] \;\geq\; \big(1 - h\sqrt{\mathcal{I}_h}\big)\, h^2 \bar{m}^0, \qquad \mathcal{I}_h = \sum_{t=1}^{N} \sigma'(b_t + \bar{\lambda} r_t)\, r_t^2\, e^{h|r_t|}. \tag{5}$$

For the truth-weighted surrogate Q_s *of Proposition 3.1 the bound holds with an extra factor* e^{-2hB} *when* $|r| \leq B$ *(Appendix B.8).*

Under Q^0 the constant $\bar{\lambda}$ has excess exactly $h^2\bar{m}^0$ under both hypotheses, so Eq. (5) says that when $h^2\mathcal{I}_h \leq \frac{1}{4}$ no procedure, including one that knows the base, the residual, and $\bar{\lambda}$, closes more than half of the gap between the constant and group resolution. As $h \to 0$, $\mathcal{I}_h$ tends to the information $\mathcal{I}$ of all target rows the procedure observes. The bound is a two-point worst case (parallelogram law and Pinsker's inequality, Appendix B.8). Sharpening its constant towards $G-1$ for G groups would need a multi-hypothesis argument that we do not pursue. Together with Eq. (4) the boundary of supportable group resolution is $h^2\mathcal{I} \asymp 1$ up to the constants $G-1$ above and $\frac{1}{4}$ below, a band that Section 4 tests. The budget is not a quantity GSRT estimates. The procedure selects by held-out loss, and the budget predicts when that selection can reach group resolution, which Section 4.2 tests against the achieved share of the group family's advantage.

3.4 ANN-Preserving Serving

Proposition 3.5 (Exact folding)**.** *Suppose* $z_{P_k}(u,i) = q_{P_k}(u)^\top v_{P_k}(i) + b_k$, $r_k(u,i) = q_{C,k}(u)^\top v_{C,k}(i)$, *and* λ_k *depends only on* u. *Then*

$$z_k^{\mathrm{GSRT}}(u,i) - b_k - c_k = q'_k(u)^\top v'_k(i), \qquad q'_k(u) = \begin{bmatrix} q_{P_k}(u) \\ \lambda_k(u)\, q_{C,k}(u) \end{bmatrix}, \quad v'_k(i) = \begin{bmatrix} v_{P_k}(i) \\ v_{C,k}(i) \end{bmatrix}. \tag{6}$$

Table 1: Headline regime of the benchmark ($\pi = 1\%$, 30,000 target users, $h = 0.8$): expected Brier gain over no transfer per million, 20-seed mean with 95% Student-t interval. Every method appears in Table 5.

Method	Group world	Continuous world	Hidden world
Uniform ($\lambda = 1$)	-751.5 [-800.4, -702.5]	-623.5 [-676.5, -570.5]	-739.2 [-801.7, -676.7]
Global blend	-0.4 [-4.9, 4.2]	-2.5 [-9.7, 4.6]	1.2 [-4.0, 6.4]
Information pooling	143.4 [136.2, 150.7]	-3.2 [-10.7, 4.3]	-0.5 [-5.3, 4.4]
Instance gate	140.1 [132.5, 147.6]	196.0 [182.2, 209.7]	-4.2 [-10.6, 2.3]
Cross-fitted argmin (ungated)	143.0 [135.8, 150.1]	196.0 [182.2, 209.7]	-0.6 [-2.0, 0.8]
GSRT (gated, $\alpha = 0.05$)	143.0 [135.8, 150.1]	196.0 [182.2, 209.7]	-0.1 [-0.3, 0.1]
Capacity control	30.6 [21.6, 39.6]	1.2 [-1.2, 3.5]	-2.1 [-6.7, 2.4]

The item vector v'_k does not depend on the user, so a maximum-inner-product index built from v'_k serves the GSRT score exactly (Malkov & Yashunin, 2020). The identity extends to any user-side matrix acting on $q_{C,k}$, so serving does not require the coefficient to be a scalar. The cost is a longer item vector, the target and consumption item embeddings side by side, and one scaling of the consumption query per user (cosine indices are covered in Appendix B.3). Restricting the coefficient to the user side keeps the whole hierarchy of Section 3.2 servable, whereas a controller on user–item cross features cannot be indexed and can only re-rank a set it did not choose. Calibration under the stratified negative sampling of production streams likewise runs after retrieval (Appendix C). Both the benchmark and the public-data study check Eq. (6) numerically in every cell.

Ordering and calibration separate. The index ranks items by $q'_k(u)^\top v'_k(i)$, and an increasing transformation of the score leaves that ranking unchanged. Precision and recall at every quota therefore depend on λ_k alone, and the probabilities can be repaired afterwards by a monotone map that never touches the index. Where the target is sparse enough that the squared loss is dominated by calibration, the coefficient becomes selectable only through this separation (Section 4.5).

4 Experiments

Controlled benchmark. The benchmark generates its own rows, so its governing variable is known and every prediction of Section 3.3 can be checked cell by cell, which no production log allows. The target's relation to the shared score is set by an observed group, by a continuous feature, by a hidden group, or by that hidden group seen through a proxy that is wrong with probability q, which attenuates the heterogeneity by $(1 - q)^2$. Heterogeneity is $h \in \{0, 0.2, \ldots, 0.8\}$, prevalence $\pi \in \{1\%, 0.1\%\}$, and the target users run from 3,000 to 100,000. Scale worlds multiply every user feature by $\omega \in \{0.5, 1, 2\}$ at fixed prevalence, so cells with the same positives carry ω^2 times the information, and a matched world gives equal positives with a sixteenfold ratio of information. Appendix E.1 states the generating process in full.

Protocol and metric. All methods share one frozen ridge-logistic base and residual, and are compared against each fixed family of Section 3.2 served alone, against unpooled, count-based and empirical-Bayes pooling, a hard gate, uniform transfer with $\lambda \equiv 1$, and source-only replacement (Appendix E.2). A first campaign on seeds 101–120 fixed the data-generating process and the comparator set and measured what ungated selection costs (negative gain in 31% of its 4,000 cells). The gate and the t certificate were then specified on development seeds and the protocol frozen. The confirmatory campaign reported here is a regime map of 4,800 cells and a scale sweep of 800 cells, run once on fresh seeds 201–220 with the first campaign's rule alongside (Appendix F). The primary metric is the expected Brier gain over no transfer on 10,000 independent test users, per million, with Student-t intervals across seeds and Holm adjustment over the predeclared family of contrasts (Appendix E.5).

Hypotheses. Five predictions of Section 3.3, denoted H1–H5 below, and two diagnostics were fixed with the protocol before the confirmatory campaign ran, and Appendix E.4 states them in full. Table 1 gives the headline regime, and the complete tables are in Appendix F.

4.1 The Regime Map

Fig. 2 shows which resolution wins in each configuration at $\pi = 1\%$. Each square averages 20 seeds, the lines use each column's measured information (Appendix E.3), and marker size shows how often GSRT picked the winner. In the observable-group world the winner changes from a coarse family

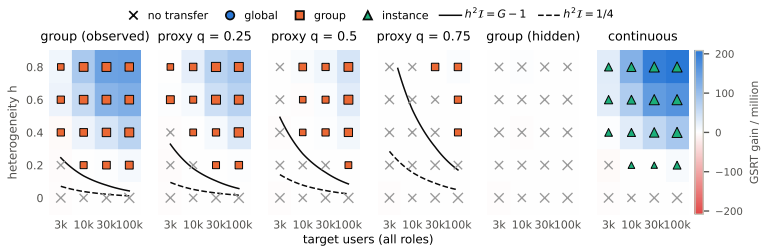


Figure 2: Regime map at $\pi = 1\%$. Colour: GSRT gain over no transfer. Marker: best fixed family. Solid and dashed lines: criterion and detection bound (Section 3.3).

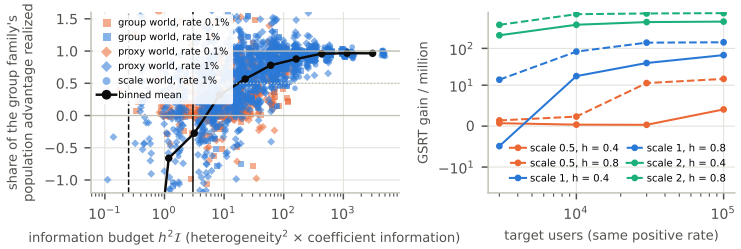


Figure 3: Left: share of the group family's advantage realized by information pooling, against the budget $h^2\mathcal{I}$ (2,122 cells). Right: GSRT gain in the scale worlds, which have equal positives and ω^2 times the information.

(no transfer, occasionally one global coefficient) to group resolution as h or the number of target users grows, and the change happens between the two lines. No configuration below the dashed detection bound is won by a group controller, and every configuration above the solid criterion is, which confirms H1. At $\pi = 0.1\%$ the information budget exceeds the criterion only in the top corner of the map (Fig. 6). The same lines, attenuated by $(1-q)^2$, order the proxy worlds. The group family still wins most of the map at $q = 0.25$, only the configurations with the most users and the largest heterogeneity at $q = 0.5$, and almost none at $q = 0.75$. In the hidden world ($q = 1$) no family beats no transfer, as Proposition 3.1(b) predicts to first order. The continuous world has no group-level relation. There the instance gate wins the high-information configurations and nothing wins the low-information ones. Over the 4,800 cells, the family predicted by Corollary 3.3 from population risk plus estimation cost coincides with the empirically best family in 56% of cells and is within 5 per million of it in 92% (costs in Table 11). Table 1 lists the headline regime ($\pi = 1\%$, 30,000 users, $h = 0.8$), where GSRT gains 143.0 [135.8, 150.1] per million in the group world and 196.0 [182.2, 209.7] in the continuous world and coincides with the best fixed family in every world, while uniform transfer loses hundreds per million everywhere and a tuned global coefficient helps only where heterogeneity is absent.

4.2 The Information Law

Corollary 3.3 and Theorem 3.4 predict that positives affect whether group resolution can be exploited only through the budget $h^2\mathcal{I}$. Fig. 3 tests this on the 2,122 cells of the regime map and the scale sweep that have an observable group relation and a positive population advantage. The median share of the group family's advantage realized by information pooling is -1.44 below the detection bound, 0.00 between the bound and the criterion, 0.32 just above the criterion, and 0.91 above $h^2\mathcal{I} = 30$ (Table 12). The binned mean crosses one half near $h^2\mathcal{I} \approx 10$, and both prevalences and all feature scales fall on the same curve. In the scale worlds (right panel and Table 10), at 10,000 users and $h = 0.8$, the same positive rate yields a GSRT gain of 2.3 [0.4, 4.2] per million at $\omega = 0.5$ and 847.1 [793.2, 900.9] at $\omega = 2$, and the group family wins 35% of the $\omega = 2$ cells against 15% at $\omega = 0.5$. The boundary moves by about a factor of four in users per doubling of ω, as ω^2 predicts, which confirms H2. Below the detection bound the share is negative, so pursuing group resolution costs more than it returns. Inside the band it is 0.00, as Corollary 3.3 requires, since a global coefficient has the lower predicted risk there and nothing is left to collect. Only above the criterion does the share become positive and grow, to 0.91 rather than 1 because a procedure also pays the selection

regret of Section 4.4. The predeclared matched-world diagnostic isolates the other half of the same question, the weighting inside a resolution. Count and information pooling coincide within noise in the headline group world (-0.4 [-1.2, 0.4] per million), and in the matched world, built expressly so that equal positive counts carry a sixteenfold ratio of information, the aggregate difference over its 1% cells is 1.6 [0.6, 2.7] per million in favour of the simpler count weighting (Table 8). This is a second-order effect beside the choice of resolution, which moves the same cells by tens of per million. The budget therefore governs the resolution itself rather than the weighting that implements it.

4.3 Observability, Rules, and Controls

Fig. 5 (left, Appendix F) traces the gain as the governing variable becomes harder to observe. At 30,000 users the GSRT gain falls from 143.0 per million with the observed group to 66.7 at $q = 0.25$, 13.1 at $q = 0.5$, 1.6 at $q = 0.75$, and -0.1 in the hidden world, and it falls monotonically at every number of target users (H3). More users raise the gain where the group is partly observed, from 1.6 to 40.2 per million at $q = 0.5$ between 3,000 and 100,000 users, but in the hidden world no number of users gives a gain distinguishable from zero. In the headline regimes GSRT serves information pooling in the group world and the gate in the continuous world, and it exceeds global blending by 143.4 [135.5, 151.2] and 198.5 [187.8, 209.2] per million, both surviving Holm correction (right panel, Table 4). Replacing the residual by the target model's own score gives a controller that can only rescale the base, so whatever it collects is pure recalibration. That control is worth 30.6 per million in the headline group world, a shuffled-source residual collects 4.2, and a random direction of the same norm 47.4. Any gain measured under a squared loss with a free calibration offset, which is how the benchmark scores every method, contains a component of this kind, and the strongest of the three controls bounds it at 47.4. GSRT exceeds the capacity control by 112.3 [100.8, 123.9] per million, the one of the three contrasts inside the predeclared primary family, at Holm-adjusted $p < 0.001$, and the other two by 138.8 and 95.6 as secondary contrasts (Table 4). At least two thirds of the headline gain is therefore attributable to the source direction itself. Under the ranking selection of Section 4.5, where an ordering leaves nothing to recalibrate, no control passes the evidence gate in any cell.

4.4 Selection Cost and the Evidence Gate

Across the 4,800 regime-map cells, GSRT is within 5 per million of the best calibrated candidate in 77% of cells, with median regret 0.3 and mean 7.3 per million (Fig. 4, Table 6). Without the gate, the cross-fitted argmin has negative gain in 1429 cells (30%, with mean -19.3 and worst -468.5 per million), and in 88% of them no candidate gains more than 5 per million. The gate changes the served candidate in 50% of cells, nearly always to no transfer. Every measure of downside improves together, which confirms H4. The negative cells fall from 1429 to 243, their mean severity from -19.3 to -8.9 per million, the tail below -50 from 158 cells to 11, and the worst cell from -468.5 to -313.0, while the mean gain rises from 8.3 to 10.6 per million and 100% of the ungated gain is kept in both headline worlds. The cells the gate declines are overwhelmingly cells where no candidate had anything to offer. The one-standard-error rule keeps a marginally higher mean gain with twice as many negative cells, and the conservative rule has fewer negative cells at a lower mean gain (Table 6). The gate sits between them, and its level was fixed on the development seeds before the campaign. The t certificate certifies 360 of the 1208 cells in which GSRT serves transfer, gaining 113.7 per million on average at 30,000 users against 35.9 for the declined ones, while the empirical-Bernstein certificate certifies 0 cells at these sample sizes (Proposition D.1, Appendix D). H5 holds on both clauses, so all five predictions hold. On public logs the gate serves transfer in 42 of 160 cells without manufacturing a gain (Section 4.6), and on production logs it serves 58 of 120 upload cells while no residual-replacement control passes it in any cell (Section 4.5).

4.5 Ranking at Production Sparsity

The benchmark scores probabilities at prevalences of 0.1% and 1%, whereas retrieval is judged by what it places inside a fixed quota, for targets about two orders of magnitude rarer. Appendix I studies this setting on the production logs of the anonymised platform of Appendix H (upload rate 7.0×10^{-5}), with a target model fitted on 67–80 positives. At this sparsity squared loss is dominated by calibration. With an intercept as its only calibration device, it selects a negative coefficient, which

shrinks the overdispersed base, in 59 of 60 cells, and after a monotone repair it selects a positive one in only 39 (Table 42). Section 3.4 allows another choice. We select the global and group coefficients for ranking by placement values, the fraction of negatives that each positive outranks (DeLong et al., 1988), whose cross-fitted mean is the AUC. Groups are pooled by their positive counts, and a binned isotonic map repairs the probabilities. Placement values give a positive coefficient in 60 of the cells. With group coefficients this rule raises AUC by 0.023 [0.009, 0.038] and recall inside a 1% quota (the positives among the top 1% of test rows across all users, a global exposure quota rather than a per-user cutoff) by 0.062 [0.037, 0.088] across four disjoint slices of the production training stream, and its Brier gain on held-out labels over the equally repaired base is positive in 60 of 60 cells (0.007 [0.005, 0.009] per million; Table 41), so the ranking gain is not paid for in calibration. Along the base axis the selected coefficient falls monotonically from 0.54 to 0.29, 0.06 and -0.01 as the base is trained on 1M, 3M, 10M and 100M rows, the share of seeds in which the gate admits transfer falls with it from 5/5 to 0/5, and by 100M rows every gain is indistinguishable from zero (Table 38). This is the direction the surrogate predicts, since its optimal constant coefficient $\mathbb{E}[w\varepsilon r]/\mathbb{E}[wr^2]$ shrinks with the base's own estimation error (Appendix B.1), and the selector follows it over a hundredfold range of training rows. The study is one of adapting a frozen, exported base. The controller sees 359 upload positives against the scarce base's 68, and where those labels can refit the base instead, the 3M-row base of Table 38 already ranks better than the adapted 1M-row base, so the gain belongs to targets whose base cannot be refitted on the controller window. Along the controller axis a window four times longer raises the controller positives from 359 to 1,424 and widens the range of the served group coefficients from 0.39–0.94 to 0.07–0.98 (Table 40), so more information buys finer resolution, as Corollary 3.3 predicts. The residual-replacement controls never pass the evidence gate here (0 of 45 cells; Table 39), because an ordering leaves nothing to recalibrate. Resolution pays only where the grouping carries the relation. In AUC within the author life-cycle groups, group coefficients beat one coefficient by 0.0089 [0.0063, 0.0115] when the grouping is real and by -0.0003 [-0.0010, 0.0003] when it is a random label.

4.6 Public Logs

KuaiRand-Pure (Gao et al., 2022) is a public short-video log whose *random*-exposure period serves as the target distribution, with a dense click label and sparse like, comment, forward, and follow labels. We apply the benchmark's protocol over 160 cells with a linear two-tower base and click residual (Appendix G). At the full label budget nothing transfers. Uniform transfer loses 70.2 per million on likes, while GSRT stays between 0.0 and 4.4 over 16 target-distribution configurations with no interval excluding zero, and a neural residual and jointly trained MMoE and PLE comparators find nothing either (Appendix G.5). Cutting the target labels to a tenth while keeping the click source abundant raises GSRT to 9.3 [0.5, 18.1] per million on likes, but an item-feature audit locates that gain in released statistics that aggregate the target actions themselves, and removing them returns GSRT to 0.21 [-0.29, 0.70] (Appendix G.6). We therefore credit no transfer to the click here. Under both feature sets the selector behaves as designed, returning no transfer with no detectable loss where no transferable click signal is found, while forced transfer pays 70.2 per million for the same decision.

5 Conclusion

A sparse production target can borrow from dense consumption feedback only at a resolution its labels can support, and that resolution is set by an information budget, the heterogeneity of the transfer relation times the Fisher information the labels carry about the coefficient. Above the criterion a group rule has lower predicted risk than one global coefficient, and below the detection bound no procedure recovers more than half of its advantage. The benchmark places the empirical boundary between the two thresholds, the production logs turn a selected coefficient into recall inside an exposure quota at an upload rate of 7.0×10^{-5} when the base cannot be refitted, and the public logs show the selector declining transfer with no detectable loss. Two properties carry this past the setting studied here. With near-constant event probability the budget belongs to the labels rather than to the scoring rule, since the Brier score and the log loss give the same threshold on the same Fisher information. The folded coefficient is served by an ordinary index, and because it alone fixes the ordering, the loss that selects it and the map that calibrates it need not agree. Deriving the threshold's constant for a ranking criterion, and running the user-randomized test of Appendix H.5, are the natural next steps.

AI USE STATEMENT

We used generative AI tools only to polish author-written text (grammar, wording, and flow). We did not use them for research ideation, methodology or experimental design, theoretical derivation and proofs, code writing, data analysis, result interpretation, or literature search and citation generation. The remaining required disclosure tasks are not applicable. The authors checked every suggested edit for unchanged technical meaning and rejected or rewrote edits that failed. We take responsibility for the final content of this work, including text, claims, or artifacts produced with the aid of generative AI.

ETHICS STATEMENT

Feedback and adaptation. GSRT scores the probability of a production action to inform exposure allocation, and the paper does not model how users and creators respond once such a predictor shapes what they see. Exposure allocated on predicted production changes the production it predicts (Perdomo et al., 2020; Jagadeesan et al., 2023; Hron et al., 2023), and a coefficient selected on last week's logs inherits that drift. The evidence gate and target calibration are offline safeguards rather than guarantees under deployment. The user-randomized test of Appendix H.5, with a prespecified production endpoint and a consumption guardrail, measures that response, and collective adaptation is left to future work.

Broader impacts. A production predictor that borrows from consumption can raise creator supply, but it can also intensify engagement optimization, push users toward labour-like activity, amplify social comparison, or concentrate exposure among established creators. The last risk appears in our own results. With more controller labels the served group coefficients spread until the newest authors receive almost no transfer (Appendix I), a statistically right decision that concentrates exposure if left unaudited. Deployments should report served coefficients and outcomes by creator group, treat minors and users vulnerable to coercive comparison separately, and monitor creator opportunity, well-being, and supply concentration. Self-efficacy, intention, and attainability motivate heterogeneous transfer here and should not be read off embeddings without validated measurement (Bandura, 2001; Ajzen, 1991).

Safeguards and release considerations. The public study uses KuaiRand-Pure under CC BY-SA 4.0 without redistributing it, and the ranking study uses production logs of the anonymised platform under its data governance, analysed offline and reported as aggregates. The released score-table subset carries integer identifiers, a day index, frozen scores, labels, grouping variables, and standardized user features, with direct identifiers removed and negatives subsampled as its manifest documents. Material from the industrial record is published de-identified, without platform names, product identifiers, people, dashboards, watermarks, or internal links, and one source image containing personal identifiers was excluded. Before any production-oriented change to exposure we recommend the evidence gate, the optional certificate, target calibration, per-group audits of served coefficients, opt-out and review policies, and a randomized test with a consumption guardrail.

REPRODUCIBILITY STATEMENT

The supplementary material contains `code/README.md`, which maps every table to a script and its flags. `code/benchmark_campaign.py` and `benchmark_dgp.py`, together with the campaign configurations and the frozen protocols, rerun the controlled benchmark of Sections 4.1–4.4 and Appendix F. `code/gsrt_core.py` runs every campaign of the public-data study (Appendix G) on KuaiRand-Pure (Gao et al., 2022). `code/gsrt_from_scores.py` and `code/gsrt_rank_aware.py` form the pipeline of Appendix I, and `gsrt_subset/` is a de-identified subset of one score-table export on which that pipeline runs end to end. Each run records its configuration, hashes, and per-cell outputs, from which all reported statistics, tables, and figures are regenerated. Both benchmark campaigns were run once after their protocols were frozen, and Appendix E restates the data-generating process and hyperparameters, so the benchmark and the public-data study are reproducible in full from public data and the released code. The full-scale production-log results of Section 4.5 are computed on internal logs under the platform's data governance and are rerunnable there but not from the public release. The industrial numbers of Appendix H are transcriptions of an anonymised source (Appendix H.6).

REFERENCES

Anass Aghbalou and Guillaume Staerman. Hypothesis transfer learning with surrogate classification losses: Generalization bounds through algorithmic stability. In *Proceedings of the 40th International Conference on Machine Learning (ICML)*, 2023.

Icek Ajzen. The theory of planned behavior. *Organizational Behavior and Human Decision Processes*, 1991. doi: 10.1016/0749-5978(91)90020-T.

T. A. Bancroft. On biases in estimation due to the use of preliminary tests of significance. *The Annals of Mathematical Statistics*, 15(2):190–204, 1944.

Albert Bandura. Social cognitive theory of mass communication. *Media Psychology*, 2001. doi: 10.1207/s1532785xmep0303_03.

Yoshua Bengio and Yves Grandvalet. No unbiased estimator of the variance of k-fold cross-validation. *Journal of Machine Learning Research*, 5:1089–1105, 2004.

Rich Caruana. Multitask learning. *Machine Learning*, 1997. doi: 10.1023/a:1007379606734.

Gavin C. Cawley and Nicola L. C. Talbot. On over-fitting in model selection and subsequent selection bias in performance evaluation. *Journal of Machine Learning Research*, 11:2079–2107, 2010. URL `https://www.jmlr.org/papers/v11/cawley10a.html`.

Olivier Chapelle. Modeling delayed feedback in display advertising. In *Proceedings of the 20th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining (KDD '14)*, 2014. doi: 10.1145/2623330.2623634.

Paul Covington, Jay Adams, and Emre Sargin. Deep neural networks for youtube recommendations. In *Proceedings of the 10th ACM Conference on Recommender Systems (RecSys '16)*, 2016. doi: 10.1145/2959100.2959190.

Elizabeth R. DeLong, David M. DeLong, and Daniel L. Clarke-Pearson. Comparing the areas under two or more correlated receiver operating characteristic curves: A nonparametric approach. *Biometrics*, 44(3):837–845, 1988. doi: 10.2307/2531595.

Yunshu Du, Wojciech M. Czarnecki, Siddhant M. Jayakumar, Mehrdad Farajtabar, Razvan Pascanu, and Balaji Lakshminarayanan. Adapting auxiliary losses using gradient similarity. arXiv preprint arXiv:1812.02224, 2018.

Bradley Efron and Carl Morris. Stein's estimation rule and its competitors—an empirical bayes approach. *Journal of the American Statistical Association*, 68(341):117–130, 1973. doi: 10.1080/01621459.1973.10481350.

Christopher Fifty, Ehsan Amid, Zhe Zhao, Tianhe Yu, Rohan Anil, and Chelsea Finn. Efficiently identifying task groupings for multi-task learning. In *Advances in Neural Information Processing Systems 34 (NeurIPS 2021)*, 2021.

Chongming Gao, Shijun Li, Yuan Zhang, Jiawei Chen, Biao Li, Wenqiang Lei, Peng Jiang, and Xiangnan He. Kuairand: An unbiased sequential recommendation dataset with randomly exposed videos. In *Proceedings of the 31st ACM International Conference on Information & Knowledge Management (CIKM '22)*, pp. 3953–3957. ACM, 2022. doi: 10.1145/3511808.3557624.

Andrew Gelman and Jennifer Hill. *Data Analysis Using Regression and Multilevel/Hierarchical Models*. Cambridge University Press, 2007.

Chuan Guo, Geoff Pleiss, Yu Sun, and Kilian Q. Weinberger. On calibration of modern neural networks. In *Proceedings of the 34th International Conference on Machine Learning (ICML)*, volume 70, pp. 1321–1330, 2017.

Steve Hanneke and Samory Kpotufe. On the value of target data in transfer learning. In *Advances in Neural Information Processing Systems 32 (NeurIPS)*, 2019.

Xinran He, Junfeng Pan, Ou Jin, Tianbing Xu, Bo Liu, Tao Xu, Yanxin Shi, Antoine Atallah, Ralf Herbrich, Stuart Bowers, and Joaquin Quiñonero Candela. Practical lessons from predicting clicks on ads at facebook. In *Proceedings of the Eighth International Workshop on Data Mining for Online Advertising (ADKDD '14)*, 2014. doi: 10.1145/2648584.2648589.

Yun He, Xue Feng, Cheng Cheng, Geng Ji, Yunsong Guo, and James Caverlee. Metabalance: Improving multi-task recommendations via adapting gradient magnitudes of auxiliary tasks. In *Proceedings of the ACM Web Conference 2022 (WWW '22)*, 2022. doi: 10.1145/3485447.3512093.

Jiri Hron, Karl Krauth, Michael I. Jordan, Niki Kilbertus, and Sarah Dean. Modeling content creator incentives on algorithm-curated platforms. In *International Conference on Learning Representations (ICLR 2023)*, 2023.

Robert A. Jacobs, Michael I. Jordan, Steven J. Nowlan, and Geoffrey E. Hinton. Adaptive mixtures of local experts. *Neural Computation*, 1991. doi: 10.1162/neco.1991.3.1.79.

Meena Jagadeesan, Nikhil Garg, and Jacob Steinhardt. Supply-side equilibria in recommender systems. In *Advances in Neural Information Processing Systems 36 (NeurIPS 2023)*, 2023. doi: 10.52202/075280-0642.

Gary King and Langche Zeng. Logistic regression in rare events data. *Political Analysis*, 2001. doi: 10.18637/jss.v008.i02.

Ilja Kuzborskij and Francesco Orabona. Stability and hypothesis transfer learning. In *Proceedings of the 30th International Conference on Machine Learning (ICML)*, pp. 942–950, 2013.

Danwei Li, Zhengyu Zhang, Siyang Yuan, Mingze Gao, Weilin Zhang, Chaofei Yang, Xi Liu, and Jiyan Yang. Adatt: Adaptive task-to-task fusion network for multitask learning in recommendations. In *Proceedings of the 29th ACM SIGKDD Conference on Knowledge Discovery and Data Mining (KDD '23)*, 2023. doi: 10.1145/3580305.3599769.

Sai Li, T. Tony Cai, and Hongzhe Li. Transfer learning for high-dimensional linear regression: Prediction, estimation and minimax optimality. *Journal of the Royal Statistical Society Series B*, 84(1):149–173, 2022.

Tsung-Yi Lin, Priya Goyal, Ross Girshick, Kaiming He, and Piotr Dollár. Focal loss for dense object detection. In *Proceedings of the IEEE International Conference on Computer Vision (ICCV)*, 2017. doi: 10.1109/iccv.2017.324.

Bo Liu, Xingchao Liu, Xiaojie Jin, Peter Stone, and Qiang Liu. Conflict-averse gradient descent for multi-task learning. In *Advances in Neural Information Processing Systems 34 (NeurIPS 2021)*, 2021.

Shengchao Liu, Yingyu Liang, and Anthony Gitter. Loss-balanced task weighting to reduce negative transfer in multi-task learning. In *Proceedings of the AAAI Conference on Artificial Intelligence 33*, volume 33, pp. 9977–9978, 2019. doi: 10.1609/aaai.v33i01.33019977.

Yuguang Liu, Yiyun Miao, and Luyao Xia. Direct routing gradient (drgrad): A personalized information surgery for multi-task learning (mtl) recommendations. In *Proceedings of the AAAI Conference on Artificial Intelligence 39*, 2025. doi: 10.1609/aaai.v39i12.33333.

Zhongqi Lu, Erheng Zhong, Lili Zhao, Evan Wei Xiang, Weike Pan, and Qiang Yang. Selective transfer learning for cross domain recommendation. In *Proceedings of the 2013 SIAM International Conference on Data Mining (SDM)*, pp. 641–649, 2013. doi: 10.1137/1.9781611972832.71.

Jiaqi Ma, Zhe Zhao, Xinyang Yi, Jilin Chen, Lichan Hong, and Ed H. Chi. Modeling task relationships in multi-task learning with multi-gate mixture-of-experts. In *Proceedings of the 24th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining (KDD '18)*, 2018a. doi: 10.1145/3219819.3220007.

Xiao Ma, Liqin Zhao, Guan Huang, Zhi Wang, Zelin Hu, Xiaoqiang Zhu, and Kun Gai. Entire space multi-task model: An effective approach for estimating post-click conversion rate. In *Proceedings of the 41st International ACM SIGIR Conference on Research and Development in Information Retrieval (SIGIR '18)*, 2018b. doi: 10.1145/3209978.3210104.

Yu. A. Malkov and D. A. Yashunin. Efficient and robust approximate nearest neighbor search using hierarchical navigable small world graphs. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 42(4):824–836, 2020. doi: 10.1109/tpami.2018.2889473.

Andreas Maurer and Massimiliano Pontil. Empirical bernstein bounds and sample variance penalization. In *Proceedings of the 22nd Annual Conference on Learning Theory (COLT)*, 2009.

Andreas Maurer, Massimiliano Pontil, and Bernardino Romera-Paredes. The benefit of multitask representation learning. *Journal of Machine Learning Research*, 2016.

Ishan Misra, Abhinav Shrivastava, Abhinav Gupta, and Martial Hebert. Cross-stitch networks for multi-task learning. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2016. doi: 10.1109/cvpr.2016.433.

Xinkun Nie and Stefan Wager. Quasi-oracle estimation of heterogeneous treatment effects. *Biometrika*, 108(2):299–319, 2021. doi: 10.1093/biomet/asaa076.

Juan C. Perdomo, Tijana Zrnic, Celestine Mendler-Dünner, and Moritz Hardt. Performative prediction. In *Proceedings of the 37th International Conference on Machine Learning (ICML)*, volume 119, pp. 7599–7609, 2020.

John C. Platt. Probabilistic outputs for support vector machines and comparisons to regularized likelihood methods. In Alexander J. Smola, Peter L. Bartlett, Bernhard Schölkopf, and Dale Schuurmans (eds.), *Advances in Large Margin Classifiers*, pp. 61–74. MIT Press, 1999.

Ozan Şener and Vladlen Koltun. Multi-task learning as multi-objective optimization. In *Advances in Neural Information Processing Systems 31 (NeurIPS 2018)*, 2018.

Liangcai Su, Junwei Pan, Ximei Wang, Xi Xiao, Shijie Quan, Xihua Chen, and Jie Jiang. Stem: Unleashing the power of embeddings for multi-task recommendation. In *Proceedings of the AAAI Conference on Artificial Intelligence 38*, 2024. doi: 10.1609/aaai.v38i8.28749.

Hongyan Tang, Junning Liu, Ming Zhao, and Xudong Gong. Progressive layered extraction (ple): A novel multi-task learning (mtl) model for personalized recommendations. In *Proceedings of the 14th ACM Conference on Recommender Systems (RecSys '20)*, 2020. doi: 10.1145/3383313.3412236.

Ye Tian and Yang Feng. Transfer learning under high-dimensional generalized linear models. *Journal of the American Statistical Association*, 118(544):2684–2697, 2023.

Aäron van den Oord, Yazhe Li, and Oriol Vinyals. Representation learning with contrastive predictive coding. arXiv preprint arXiv:1807.03748, 2018.

A. W. van der Vaart. *Asymptotic Statistics*. Cambridge Series in Statistical and Probabilistic Mathematics. Cambridge University Press, 1998. doi: 10.1017/CBO9780511802256.

HaiYing Wang, Aonan Zhang, and Chong Wang. Nonuniform negative sampling and log odds correction with rare events data. In *Advances in Neural Information Processing Systems 34 (NeurIPS 2021)*, 2021.

Zi-Rui Wang, Zihang Dai, Barnabás Póczos, and Jaime Carbonell. Characterizing and avoiding negative transfer. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, 2019. doi: 10.1109/CVPR.2019.01155.

Ji Yang, Xinyang Yi, Derek Zhiyuan Cheng, Lichan Hong, Yang Li, Simon Xiaoming Wang, Taibai Xu, and Ed H. Chi. Mixed negative sampling for learning two-tower neural networks in recommendations. In *Companion Proceedings of the Web Conference 2020*, 2020. doi: 10.1145/3366424.3386195.

Xinyang Yi, Ji Yang, Lichan Hong, Derek Zhiyuan Cheng, Lukasz Heldt, Aditee Kumthekar, Zhe Zhao, Li Wei, and Ed Chi. Sampling-bias-corrected neural modeling for large corpus item recommendations. In *Proceedings of the 13th ACM Conference on Recommender Systems (RecSys '19)*, 2019. doi: 10.1145/3298689.3346996.

Tianhe Yu, Saurabh Kumar, Abhishek Gupta, Sergey Levine, Karol Hausman, and Chelsea Finn. Gradient surgery for multi-task learning. In *Advances in Neural Information Processing Systems 33 (NeurIPS 2020)*, volume 33, pp. 5824–5836, 2020.

Bianca Zadrozny and Charles Elkan. Transforming classifier scores into accurate multiclass probability estimates. In *Proceedings of the 8th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining (KDD)*, pp. 694–699. ACM, 2002. doi: 10.1145/775047.775151.

Zhen Zhang, Qingyun Liu, Yuening Li, Sourabh Bansod, Mingyan Gao, Yaping Zhang, Zhe Zhao, Lichan Hong, Ed H. Chi, Shuchao Bi, and Liang Liu. Co-optimize content generation and consumption in a large scale video recommendation system. In *Proceedings of the 18th ACM Conference on Recommender Systems (RecSys '24)*, pp. 762–764, 2024. doi: 10.1145/3640457.3688033.

Zhe Zhao, Lichan Hong, Li Wei, Jilin Chen, Aniruddh Nath, Shawn Andrews, Aditee Kumthekar, Maheswaran Sathiamoorthy, Xinyang Yi, and Ed Chi. Recommending what video to watch next: A multitask ranking system. In *Proceedings of the 13th ACM Conference on Recommender Systems (RecSys '19)*, 2019. doi: 10.1145/3298689.3346997.

Feng Zhu, Yan Wang, Chaochao Chen, Jingjie Zhou, Longfei Li, and Guanfeng Liu. Cross-domain recommendation: Challenges, progress, and prospects. In *Proceedings of the 30th International Joint Conference on Artificial Intelligence (IJCAI)*, pp. 4721–4728, 2021.

A NOTATION

Table 2: Notation used in Section 3.

Symbol	Meaning
u, i, g	user, item, observed user group (G groups, masses π_g)
C, P_k	dense consumption task, sparse production target $k = 1, \dots, K$
$q_{P_k}(u)$, $v_{P_k}(i)$, b_k	target-only query vector, item vector, intercept
$q_{C,k}(u)$, $v_{C,k}(i)$	consumption query and item vectors defining the residual $r_k(u, i)$
z_{P_k}, r_k, $\lambda_k(u)$, c_k	target-only logit, consumption residual, user-side coefficient, calibration offset, Eq. (1)
$p^\star$, $z^\star$, ε, w	observable Bayes probability, its logit, ideal correction $z^\star - b$, Brier weight $\sigma'(z^\star)^2$
$Q(\lambda)$, $\alpha(G)$, m_g, A	weighted surrogate risk, governing relation, $\mathbb{E}[wr^2 \mid G = g]$, excess of the best constant, Proposition 3.1
H, V, ϕ, C_{glob}, C_{grp}	Hessian and score variance of a family, its parameter gradient, estimation costs, Proposition 3.2
$\mathcal{I}$, $\mathcal{I}_g$, $\bar{\iota}$, h^2	coefficient information $\sum p(1-p)r^2$, its group part, information per positive, residual-weighted heterogeneity, Eqs. (2) and (4)
$\mathcal{I}_h$, $\bar{m}$, $P_\pm$	information with the factor $e^{h\lvert r\rvert}$, $\sum_g \pi_g m_g$, the two hypotheses of Theorem 3.4
Λ, λ^{glob}, λ_g^{loc}, $\lambda_g^{(s)}$, s	coefficient grid, global and local grid values, information-pooled coefficient, pooling strength, Eq. (2)
$\lambda_w(u)$, ρ	instance gate and its ridge penalty
$\mathcal{D}_{\mathrm{ctrl}}$, $\mathcal{D}_{\mathrm{cal}}$, $\mathcal{D}_{\mathrm{sel}}$, $\mathcal{D}_{\mathrm{cert}}$	controller-fitting, calibration, selection, and certification users
$\mathcal{H}_k$, h_0, $\widehat{h}_k$	candidate hierarchy, no-transfer candidate, selected candidate, Eq. (3)
e_j, $\widehat{V}$, δ, U	per-user excess Brier, its sample variance, confidence level, empirical-Bernstein bound, Eq. (7)
θ_P, θ_C, Δ, v_P, v_C, c	target and source parameters, discrepancy, estimation variances and covariance, Proposition B.1
g_P, g_C, η	production and consumption gradients, step size, Appendix B.2
P, N, HN, EN, U	production positives, negatives, hard and easy negatives, sampling user group
$s_y(x)$, $p(x)$, $p_s(x)$	retention probability, target-distribution and retained-sample posteriors, Proposition C.1
ω, q	feature scale of the scale worlds, proxy noise of the proxy worlds

B PROOFS AND DERIVATIONS

B.1 OPTIMAL BORROWING WEIGHT

Proposition B.1 (Optimal borrowing weight)**.** *Let $\widehat{\theta}_P$ and $\widehat{\theta}_C$ be unbiased estimates of θ_P and $\theta_C = \theta_P + \Delta$ with variances v_P, v_C and error covariance c, and let $\widetilde{\theta}_P(\lambda) = (1-\lambda)\widehat{\theta}_P + \lambda\widehat{\theta}_C$. Then* $\mathrm{MSE}(\lambda) = (1-\lambda)^2 v_P + \lambda^2(v_C + \Delta^2) + 2\lambda(1-\lambda)c$*, minimized at $\lambda^\star = (v_P - c)/(v_P + v_C + \Delta^2 - 2c)$ whenever the denominator is positive, and for independent errors $\lambda^\star = v_P/(v_P + v_C + \Delta^2)$.*

The optimum grows with target uncertainty and shrinks with discrepancy, and both vary across users, which is why GSRT makes the coefficient user-side and lets the target evidence set its resolution.

Proof. Let $e_P = \widehat{\theta}_P - \theta_P$ and $e_C = \widehat{\theta}_C - \theta_C$, so that $\mathbb{E}[e_P] = \mathbb{E}[e_C] = 0$, $\mathbb{E}[e_P^2] = v_P$, $\mathbb{E}[e_C^2] = v_C$, and $\mathbb{E}[e_P e_C] = c$. The transferred estimator satisfies

$$\widetilde{\theta}_P(\lambda) - \theta_P = (1-\lambda)e_P + \lambda e_C + \lambda\Delta.$$

Squaring and taking expectations, the cross terms with Δ vanish by unbiasedness and

$$\mathrm{MSE}(\lambda) = (1-\lambda)^2 v_P + \lambda^2(v_C + \Delta^2) + 2\lambda(1-\lambda)c.$$

The derivative $-2(1-\lambda)v_P + 2\lambda(v_C + \Delta^2) + 2(1-2\lambda)c$ vanishes at

$$\lambda^\star = \frac{v_P - c}{v_P + v_C + \Delta^2 - 2c},$$

and the second derivative $2(v_P + v_C + \Delta^2 - 2c)$ is positive exactly when the denominator is positive, so the stationary point is the minimizer. Setting $c = 0$ gives the independent case. Dependence can therefore increase or decrease the optimal amount of borrowing. The expression governs a convex blend of two estimates (Caruana, 1997; Maurer et al., 2016). □

Vector case. For vector parameters with error covariances V_P, V_C, cross-covariance C, and discrepancy Δ, the same expansion gives, for a matrix blend Λ, $\mathbb{E}\|\widetilde{\theta}_P(\Lambda) - \theta_P\|^2 = \operatorname{tr}\big[(I - \Lambda)V_P(I - \Lambda)^\top + \Lambda(V_C + \Delta\Delta^\top)\Lambda^\top + (I - \Lambda)C\Lambda^\top + \Lambda C^\top(I - \Lambda)^\top\big]$, minimized at $\Lambda^\star = (V_P - C)(V_P + V_C + \Delta\Delta^\top - C - C^\top)^{-1}$ when the inverse exists. The scalar rule of the main text is the case $\Lambda = \lambda I$. Any user-side matrix keeps the folding identity of Proposition 3.5 exact, a matrix acting on the item side does not, and the scalar of the main text is the lowest-complexity member of the foldable class.

Additive form. GSRT adds the source score itself, $b + \lambda r$, rather than blending two estimates of one parameter, and the blend above is the special case in which the residual is the source-minus-target direction. For the additive form the relevant optimum is that of the surrogate of Appendix B.4, whose best constant coefficient is $\lambda^\star = \mathbb{E}[w\varepsilon r]/\mathbb{E}[wr^2]$ with $\varepsilon = z^\star - b - c$, the correction the frozen base still needs. A base fitted on few positives is shrunk and noisy, so ε retains a component along $z^\star$ that the source direction partly explains and $\lambda^\star > 0$, and as the base is fitted on more rows, $\varepsilon \to 0$ and $\lambda^\star \to 0$. This is the monotonicity that Section 4.5 observes. The proposition above gives the same direction for the blend but not the value of the additive optimum.

B.2 First-Order Effect of an Auxiliary Update

For a twice-differentiable production loss L_P with gradient $g_P = \nabla_\theta L_P(\theta)$ and a consumption gradient $g_C = \nabla_\theta L_C(\theta)$, Taylor's theorem gives

$$L_P(\theta - \eta g_C) = L_P(\theta) - \eta\, g_P^\top g_C + \frac{\eta^2}{2}\, g_C^\top \nabla^2 L_P(\widetilde{\theta})\, g_C$$

for some $\widetilde{\theta}$ on the segment between θ and $\theta - \eta g_C$. Positive alignment lowers the production loss to first order and negative alignment raises it for a sufficiently small step. Global averages of $g_P^\top g_C$ can conceal opposing signs across production targets, user populations, and layers (Du et al., 2018; Fifty et al., 2021). PCGrad and CAGrad modify conflicting updates directly (Yu et al., 2020; Liu et al., 2021), and MetaBalance and DRGrad adapt auxiliary influence in recommendation (He et al., 2022; Liu et al., 2025). The expansion holds step by step, which is the level at which auxiliary gradients help or conflict.

B.3 Proof of Proposition 3.5 (Exact Folding)

With $q'_k(u) = [q_{P_k}(u); \lambda_k(u) q_{C,k}(u)]$ and $v'_k(i) = [v_{P_k}(i); v_{C,k}(i)]$, block-wise expansion of the inner product gives

$$q'_k(u)^\top v'_k(i) = q_{P_k}(u)^\top v_{P_k}(i) + \lambda_k(u)\, q_{C,k}(u)^\top v_{C,k}(i) = z_{P_k}(u, i) - b_k + \lambda_k(u) r_k(u, i),$$

which equals $z_k^{\mathrm{GSRT}}(u, i) - b_k - c_k$ by Eq. (1). Because λ_k depends only on u, every factor of v'_k is a function of i alone, so the item side can be indexed before any query arrives, and the constants $b_k + c_k$ are rank-preserving within a target. In the benchmark the base and residual are bilinear in $u \odot i$, so the folded query is $[\, b_0 + c_k;\ u \odot (\beta_P + \lambda(u)\beta_C)\,]$ and the document is $[1;\ i]$. Section 4 reports the numerical check. □

Index types and memory. Folding is exact for maximum-inner-product search. A cosine index normalizes the folded document $v'_k(i)$, whose norm differs from that of $v_{P_k}(i)$ and varies with the item, so the ordering it serves is that of $q'^\top_k v'_k / \|v'_k\|$. The standard reduction of inner-product search to cosine search, which augments every document with $\sqrt{M^2 - \|v'_k\|^2}$, restores Eq. (6) exactly, so both index types are covered. The consumption block $v_{C,k}(i)$ is the same vector for every target that borrows from the same consumption head, so an implementation that supports shared storage across the K indices pays for it once.

B.4 Proof of Proposition 3.1 (Weighted Resolution) and the Surrogate

The surrogate. Let $z^\star$ be a logit of the observable Bayes probability, $p^\star = \sigma(z^\star)$, and $z = b + \lambda(X)r + c$ a served logit. With $\delta = z - z^\star$, $\sigma(z) - p^\star = \sigma'(z^\star)\delta + \frac{1}{2}\sigma''(\xi)\delta^2$ for some ξ between z and $z^\star$, so $\mathbb{E}[(\sigma(z) - p^\star)^2] = \mathbb{E}[\sigma'(z^\star)^2\delta^2] + O(\mathbb{E}|\delta|^3)$ because $|\sigma''| \le \frac{1}{6\sqrt{3}}$ and $|\sigma'| \le \frac{1}{4}$. Writing $\varepsilon = z^\star - b - c$ gives $\delta = \lambda(X)r - \varepsilon$ and the weighted surrogate $Q(\lambda) = \mathbb{E}[w(\varepsilon - \lambda(X)r)^2]$ with $w = \sigma'(z^\star)^2$. The offset c shifts ε by a constant and is absorbed into ν below.

Proof of (a). Under $\varepsilon = \alpha(G)r + \nu$ with $\mathbb{E}[w\nu \mid X, G, r] = 0$ and G a function of X, for any $\lambda(X)$,

$$Q(\lambda) = \mathbb{E}\big[w(\alpha(G) - \lambda(X))^2 r^2\big] + \mathbb{E}[w\nu^2],$$

since the cross term vanishes by the conditional moment condition. Restricting to controllers constant within groups, $\lambda(X) = \lambda_G$, the first term is $\sum_g \pi_g m_g(\alpha_g - \lambda_g)^2$ with $m_g = \mathbb{E}[wr^2 \mid G = g]$, minimized at $\lambda_g = \alpha_g$ (value zero). For a constant λ, it equals $\sum_g \pi_g m_g(\alpha_g - \lambda)^2$, a quadratic in λ minimized at $\bar\lambda = \sum_g \pi_g m_g \alpha_g / \sum_g \pi_g m_g$ with minimum $\sum_g \pi_g m_g(\alpha_g - \bar\lambda)^2$, which is the stated excess A. A controller that varies within a group cannot do better than the group-wise optimum α_g because $\alpha(G)$ is constant given X in that group. □

Proof of (b). With $G = H$ independent of (X, r) and $\mathbb{E}[w\nu \mid X, H, r] = 0$,

$$\begin{aligned} Q(\lambda) &= \mathbb{E}\big[wr^2(\alpha(H) - \lambda(X))^2\big] + \mathbb{E}[w\nu^2] \\ &= \mathbb{E}\big[wr^2\big]\,\mathrm{Var}(\alpha(H)) + \mathbb{E}\big[wr^2(\mathbb{E}[\alpha(H)] - \lambda(X))^2\big] + \mathbb{E}[w\nu^2], \end{aligned}$$

where the decomposition uses $\mathbb{E}[wr^2(\alpha(H) - \mathbb{E}\alpha)(\mathbb{E}\alpha - \lambda(X))] = 0$, which holds because H is independent of (X, r) and hence of w, r, and $\lambda(X)$ (here w is evaluated at the observable logit and does not depend on H). The middle term is nonnegative and vanishes only for $\lambda(X) = \mathbb{E}[\alpha(H)]$ almost surely on $\{wr^2 > 0\}$. □

The order of the statement. The observable Bayes probability is $\bar p = \mathbb{E}_H[\sigma(b + \alpha(H)r + c)]$, whose logit equals $b + \mathbb{E}[\alpha]r + c$ to first order in $\mathrm{Var}(\alpha(H))r^2$, so the exact Brier-optimal user-side coefficient departs from the constant only at second order. Part (b) is therefore a first-order statement, and the second-order remainder is what a procedure would have to detect to beat the constant. In the benchmark's hidden world that remainder is smaller than the estimation cost of every non-constant family (Section 4), which is why no family beats no transfer there in any cell.

B.5 Proof of Proposition 3.2 (Cost of Resolution)

Let $\ell_\theta(x, y) = (\sigma(b + \lambda_\theta(x)r) - y)^2$ and $L(\theta) = \mathbb{E}\ell_\theta$, with $\theta^\star$ the population minimizer and $\hat\theta$ the minimizer of the empirical loss on n i.i.d. rows. Under the standard regularity conditions for M-estimation (twice differentiable loss with integrable envelopes, $\theta^\star$ interior with nonsingular Hessian), $\sqrt{n}(\hat\theta - \theta^\star) \to N(0, H^{-1}VH^{-1})$ with $H = \nabla^2 L(\theta^\star)$ and $V = \mathrm{Var}(\nabla\ell_{\theta^\star})$, and the second-order expansion $L(\hat\theta) - L(\theta^\star) = \frac{1}{2}(\hat\theta - \theta^\star)^\top H(\hat\theta - \theta^\star) + o_p(\|\hat\theta - \theta^\star\|^2)$ gives $\mathbb{E}[L(\hat\theta) - L(\theta^\star)] = \frac{1}{2n}\,\mathrm{tr}(H^{-1}V) + o(1/n)$ (van der Vaart, 1998, Ch. 5). It remains to compute H and V. With $p = \sigma(b + \lambda_\theta r)$ and $\phi = \partial\lambda_\theta/\partial\theta$,

$$\nabla\ell_\theta = 2(p - y)\,\sigma'(b + \lambda_\theta r)\,r\,\phi, \qquad \nabla^2\ell_\theta = 2\sigma'^2 r^2\phi\phi^\top + 2(p - y)\,\partial_\theta\big(\sigma' r\phi\big).$$

At $\theta^\star$, $\mathbb{E}[(p - y) \mid x] = p - p^\star$, so the second term of the Hessian has expectation $2\mathbb{E}[(p - p^\star)\partial_\theta(\sigma' r\phi)]$, which vanishes when the family contains the truth and is otherwise a misspecification term that enters the $1/n$ coefficient of the cost. For the variance, $\mathrm{Var}(\nabla\ell \mid x) = 4\sigma'^2 r^2\phi\phi^\top\,\mathrm{Var}(y \mid x) = 4p^\star(1 - p^\star)\sigma'^2 r^2\phi\phi^\top$, and V is its expectation plus the variance of $\mathbb{E}[\nabla\ell \mid x] = 2(p - p^\star)\sigma' r\phi$, which is zero at $\theta^\star$ for a well-specified family. The Gauss–Newton forms of the proposition, $H = 2\mathbb{E}[\sigma'^2 r^2\phi\phi^\top]$ and $V = 4\mathbb{E}[p^\star(1 - p^\star)\sigma'^2 r^2\phi\phi^\top]$, are these expressions with both misspecification terms dropped. They are exact for a family that contains $z^\star$, which the group family does under group heterogeneity, and an approximation for the global family, which does not. The general expansion holds with the full H and V. The proposition states the Gauss–Newton forms because these are the ones the benchmark computes, and it reports them next to the realized excess of every family, so the size of the approximation is measured rather than assumed (Appendix F). For a scalar family $\phi = 1$ and $\frac{1}{2}\,\mathrm{tr}(H^{-1}V) = \mathbb{E}[p^\star(1 - p^\star)\sigma'^2 r^2]/\mathbb{E}[\sigma'^2 r^2]$. For unpooled group coefficients the parameters are

variation-independent across groups, H and V are diagonal with entries restricted to the group's rows, and the trace is the stated sum. When $p^\star$ is nearly constant, each ratio is $\bar{p}(1-\bar{p})$ and the cost of d coefficients is $d\,\bar{p}(1-\bar{p})/n$. For the tanh gate, $\phi = (1-\lambda^2)[1; u; \mathrm{onehot}(g)]$ and the trace is evaluated numerically. The intercept and the one-hot block are collinear, so a pseudo-inverse is used and the ridge penalty of the estimator is ignored, which makes the predicted cost of the gate conservative. □

The benchmark computes H and V on the controller rows with the true probabilities and reports the predicted cost next to the realized excess of each family (Appendix F).

B.6 Proof of Corollary 3.3 (Supportable Resolution)

By Propositions 3.1(a) and 3.2, the predicted surrogate risk of the best global coefficient exceeds that of the best group controller by $A = \sum_g \pi_g m_g(\alpha_g - \bar{\lambda})^2$, and the estimation costs are C_{glob} and C_{grp}, so group resolution is predicted to win iff $A > C_{\mathrm{grp}} - C_{\mathrm{glob}}$. With $p^\star \approx \bar{p}$ throughout, $w \approx \bar{p}^2(1-\bar{p})^2$, $m_g \approx \bar{p}^2(1-\bar{p})^2\mathbb{E}[r^2 \mid g]$, and $C_{\mathrm{grp}} - C_{\mathrm{glob}} \approx (G-1)\bar{p}(1-\bar{p})/n$. Substituting and dividing by $\bar{p}(1-\bar{p})\mathbb{E}[r^2]$,

$$A > C_{\mathrm{grp}} - C_{\mathrm{glob}} \iff \frac{\sum_g \pi_g \mathbb{E}[r^2 \mid g](\alpha_g - \bar{\lambda})^2}{\mathbb{E}[r^2]} \cdot n\bar{p}(1-\bar{p})\mathbb{E}[r^2] > G-1 \iff h^2\mathcal{I} > G-1,$$

using $\mathcal{I} = \sum_{\mathrm{rows}} p^\star(1-p^\star)r^2 \approx n\bar{p}(1-\bar{p})\mathbb{E}[r^2]$. □

B.7 The Same Criterion under Log Loss

Corollary 3.3 is stated for the Brier risk that Section 3.1 evaluates. The same threshold, with the same h^2 and the same $\mathcal{I}$, governs log-loss risk.

Population side. Let $R_{\log}(\lambda) = \mathbb{E}[-y\log\sigma(z) - (1-y)\log(1-\sigma(z))]$ with $z = b + \lambda(X)r + c$. Its excess over the risk at $z^\star$ is $\mathbb{E}\big[\mathrm{KL}(\mathrm{Bern}(p^\star) \,\|\, \mathrm{Bern}(\sigma(z)))\big]$, and by the Taylor form established in Step 3 of Appendix B.8, $\mathrm{KL} = \frac{1}{2}\sigma'(\xi)(z-z^\star)^2$ for some ξ between $z^\star$ and z. To second order the excess is therefore the weighted quadratic

$$Q_{\log}(\lambda) = \tfrac{1}{2}\,\mathbb{E}\big[\sigma'(z^\star)\,(\varepsilon - \lambda(X)r)^2\big],$$

which is the surrogate of Proposition 3.1 with the weight $w = \sigma'(z^\star)^2$ replaced by $\frac{1}{2}\sigma'(z^\star)$. Both parts of that proposition are proved for a generic nonnegative weight, using only the conditional moment condition $\mathbb{E}[w\nu \mid X, G, r] = 0$ and the group masses $m_g = \mathbb{E}[wr^2 \mid G = g]$, so they hold verbatim once the moment condition is restated for the new weight. The excess of the best constant over the best group controller is then $A_{\log} = \frac{1}{2}\sum_g \pi_g \mathbb{E}[\sigma' r^2 \mid g](\alpha_g - \bar{\lambda})^2$.

Cost side. For the log loss, $\nabla_\theta \ell = (\sigma(z) - y)\, r\, \phi$ and $\nabla^2_\theta \ell = \sigma'(z) r^2 \phi\phi^\top + (\sigma(z) - y)\,\partial_\theta(r\phi)$. At a well-specified interior $\theta^\star$ the second Hessian term has mean zero, so $H = \mathbb{E}[\sigma' r^2 \phi\phi^\top]$, and

$$V = \mathbb{E}\big[(\sigma(z) - y)^2 r^2 \phi\phi^\top\big] = \mathbb{E}\big[p^\star(1-p^\star) r^2 \phi\phi^\top\big] = H,$$

the information equality. Proposition 3.2 then gives $\frac{1}{2n}\,\mathrm{tr}(H^{-1}V) = d/(2n)$ for a family with d parameters, so each additional coefficient costs $1/(2n)$, free of $\bar{p}(1-\bar{p})$. As in the Brier case of Appendix B.6, the group family contains the truth and its cost is exact, while a global coefficient under group heterogeneity is misspecified and takes the same well-specified approximation, so $C_{\mathrm{grp}} - C_{\mathrm{glob}} \approx (G-1)/(2n)$.

The criterion. Group resolution has lower predicted log-loss risk iff $A_{\log} > (G-1)/(2n)$, that is iff $\sum_g \pi_g \mathbb{E}[\sigma' r^2 \mid g](\alpha_g - \bar{\lambda})^2 > (G-1)/n$, the factors $\frac{1}{2}$ cancelling. With $p^\star \approx \bar{p}$ throughout, $\sigma' \approx \bar{p}(1-\bar{p})$, the two weightings define the same $\bar{\lambda}$ and the same h^2, and the left side is $\bar{p}(1-\bar{p})\mathbb{E}[r^2]h^2$. Multiplying by n,

$$h^2\,\mathcal{I} > G-1, \qquad \mathcal{I} = n\bar{p}(1-\bar{p})\mathbb{E}[r^2] = \sum_{\mathrm{rows}} p^\star(1-p^\star)r^2,$$

which is Eq. (4). □

Why the two changes cancel. Under the Brier risk the curvature weight is σ'^2 and each coefficient costs $\bar{p}(1-\bar{p})/n$, while under the log loss the weight is σ' and each coefficient costs $1/(2n)$. Both sides of the comparison carry one factor of $\bar{p}(1-\bar{p})$ less in the log-loss case, so the ratio that defines the criterion is unchanged. What survives is $\mathcal{I} = \sum_{\text{rows}} p^\star(1-p^\star)r^2$, the Fisher information the target rows carry about a scalar coefficient on r. The boundary of supportable resolution is therefore a property of the labels and the residual rather than of the scoring rule. A platform that fits and evaluates with cross-entropy meets the same boundary as one that uses the Brier score, and the benchmark of Section 4, which measures it under the Brier score, measures a quantity common to both. When $p^\star$ varies across groups the two weightings define different best constants and different h^2, and the criterion becomes specific to the scoring rule.

B.8 Proof of Theorem 3.4 (Detection Lower Bound)

Fix the design (base logits b_t, residuals r_t, groups g_t) and let $P_\pm$ be the label distributions under $\alpha^\pm = (\bar{\lambda} \pm h, \bar{\lambda} \mp h)$. The surrogate weights are fixed at the global logit $z_t = b_t + \bar{\lambda} r_t$, so that $w_t^0 = \sigma'(z_t)^2$ and $m_g^0 = \mathbb{E}[w^0 r^2 \mid G = g]$ are the same under both hypotheses, which is what Step 1 needs. For a controller λ that is constant within groups, $Q_s^0(\lambda) = \sum_g \pi_g m_g^0 (\lambda_g - \alpha_g^s)^2$, and for a general user-side controller one replaces λ_g by its $w^0 r^2$-weighted group mean, which by Jensen can only decrease Q_s^0, so the bound for group-constant controllers implies the bound for all.

Step 1 (parallelogram). For every real λ_g, $(\lambda_g - \alpha_g^+)^2 + (\lambda_g - \alpha_g^-)^2 \geq \frac{1}{2}(\alpha_g^+ - \alpha_g^-)^2 = 2h^2$. Since the same m_g^0 weight both excesses, $Q_+^0(\lambda) + Q_-^0(\lambda) \geq 2h^2 \sum_g \pi_g m_g^0 = 2h^2 \bar{m}^0$ for every λ.

Step 2 (Le Cam). For any estimator $\hat{\lambda}$ (a measurable function of the labels),

$$\max_s \mathbb{E}_s[Q_s^0(\hat{\lambda})] \geq \tfrac{1}{2}\big(\mathbb{E}_+[Q_+^0(\hat{\lambda})] + \mathbb{E}_-[Q_-^0(\hat{\lambda})]\big) \geq \tfrac{1}{2} \int \min(dP_+, dP_-) \inf_\lambda [Q_+^0(\lambda) + Q_-^0(\lambda)]$$
$$\geq (1 - \mathrm{TV}(P_+, P_-))\, h^2 \bar{m}^0,$$

using $\int \min(dP_+, dP_-) = 1 - \mathrm{TV}$.

Step 3 (Pinsker and an exact second-order form). Labels are independent given the design, so $\mathrm{KL}(P_+ \| P_-) = \sum_t \mathrm{KL}\big(\mathrm{Bern}(\sigma(z_t + s_t h r_t)) \,\|\, \mathrm{Bern}(\sigma(z_t - s_t h r_t))\big)$ with $z_t = b_t + \bar{\lambda} r_t$ and $s_t = \pm 1$ the sign of the group. For fixed z_1, the function $f(z_2) = \mathrm{KL}(\mathrm{Bern}(\sigma(z_1)) \| \mathrm{Bern}(\sigma(z_2))) = \sigma(z_1)(z_1 - z_2) + \log \frac{1+e^{z_2}}{1+e^{z_1}}$ satisfies $f(z_1) = 0$, $f'(z_2) = \sigma(z_2) - \sigma(z_1)$ so $f'(z_1) = 0$, and $f''(z_2) = \sigma'(z_2)$, so by Taylor's theorem with Lagrange remainder, $f(z_2) = \frac{1}{2}\sigma'(\xi)(z_1 - z_2)^2$ for some ξ between z_1 and z_2. Since $(\log \sigma')' = 1 - 2\sigma \in (-1, 1)$, $\sigma'(\xi) \leq \sigma'(z_t) e^{|\xi - z_t|} \leq \sigma'(z_t) e^{h|r_t|}$, because both z_1 and z_2 lie within $h|r_t|$ of z_t. With $(z_1 - z_2)^2 = 4h^2 r_t^2$,

$$\mathrm{KL}(P_+ \| P_-) \leq \sum_t 2h^2 \sigma'(z_t) r_t^2 e^{h|r_t|} = 2h^2 \mathcal{I}_h, \qquad \mathrm{TV}(P_+, P_-) \leq \sqrt{\mathrm{KL}/2} \leq h\sqrt{\mathcal{I}_h}.$$

Combining the three steps gives Eq. (5).

Step 4 (truth-weighted surrogate). Under hypothesis s the surrogate of Proposition 3.1 has weights $w_{s,t} = \sigma'(b_t + \alpha_{g_t}^s r_t)^2$, whose argument differs from z_t by $h|r_t|$. Since $|(\log \sigma'^2)'| = 2|1 - 2\sigma| < 2$, $e^{-2h|r_t|} w_t^0 \leq w_{s,t} \leq e^{2h|r_t|} w_t^0$, so $Q_s(\lambda) \geq e^{-2hB} Q_s^0(\lambda)$ for every controller when $|r_t| \leq B$, and Eq. (5) holds for Q_s with the factor e^{-2hB}. □

Remarks. (i) The bound is nontrivial only when $h\sqrt{\mathcal{I}_h} < 1$, and this is its purpose. It certifies that below the information budget no procedure can exploit the finer resolution, and above it Corollary 3.3 takes over. (ii) $\mathcal{I}_h \leq e^{hB}\mathcal{I}$ when $|r| \leq B$, and $\mathcal{I}_h \to \mathcal{I}$ as $h \to 0$. Here $\mathcal{I}$ counts every target row the procedure observes, not only the controller rows. (iii) A G-point family would be needed to sharpen the constant towards the $G - 1$ of Eq. (4), and we prove only the two-point bound. (iv) A numerical check with the benchmark's grid estimator (Appendix F), which evaluates Q^0 with the weights at the global logit, found no violation of Eq. (5) over $N \in \{2000, 8000, 32000\}$ and $h \in \{0.05, 0.1, 0.2, 0.4\}$. (v) The weights must not depend on the hypothesis. With the truth-weighted m_g of Proposition 3.1, Step 1 fails. For two rows with $b = -8$, $r = 1$, $\bar{\lambda} = 0$, $h = 0.2$, the weights $m_\pm = \sigma'(b \pm h)^2$ swap between the hypotheses, and the label-free constant $c = h(m_+ - m_-)/(m_+ + m_-) \approx 0.076$ has truth-weighted excess $\approx 4.16 \times 10^{-9}$ under both

hypotheses, below $(1 - h\sqrt{\mathcal{I}_h})h^2\bar{m} \approx 4.83 \times 10^{-9}$. On the same design Q^0 obeys Eq. (5), and the factor e^{-2hB} of Step 4 restores the truth-weighted bound. (vi) Under Q^0 the constant λ is the minimax constant, with excess $h^2\bar{m}^0$ under both hypotheses, while the best global coefficient under a single hypothesis has excess $A = h^2\bar{m}^0 \cdot 4m_1^0 m_2^0/(m_1^0 + m_2^0)^2 \leq h^2\bar{m}^0$, with equality when the two groups carry equal weighted mass, so "half of the gap" in the main text is measured from the minimax constant.

B.9 Proof of Proposition C.1 (Retention Log-Odds Correction)

By Bayes' rule, $p_s(x) = P(Y = 1 \mid X = x, S = 1) = s_1(x)p(x)/\big(s_1(x)p(x) + s_0(x)(1 - p(x))\big)$. Writing the corresponding expression for $Y = 0$ and taking the ratio,

$$\frac{p_s(x)}{1 - p_s(x)} = \frac{s_1(x)}{s_0(x)} \cdot \frac{p(x)}{1 - p(x)},$$

and taking logarithms gives $\operatorname{logit} p(x) = \operatorname{logit} p_s(x) + \log\big(s_0(x)/s_1(x)\big)$. If $s_0(x)/s_1(x)$ is constant on a stratum, the correction is a constant shift of the logit on that stratum, which leaves the ordering of scores within the stratum unchanged. □

B.10 Proof of Corollary C.2 (Label-Dependent Retention)

Let $\rho(U, y_C, y_P) = P(S = 1 \mid x, Y_C = y_C, Y_P = y_P)$ depend on x only through the sampling group U. Then

$$s_{y_P}(x) = P(S = 1 \mid x, Y_P = y_P) = \sum_{y_C \in \{0,1\}} \rho(U, y_C, y_P)\, P(Y_C = y_C \mid x, Y_P = y_P),$$

which depends on x through $P(Y_C \mid x, Y_P)$ unless $\rho(U, 0, y_P) = \rho(U, 1, y_P)$, and substituting into Proposition C.1 gives the first claim. When every positive is retained, $s_1(x) = 1$, and with $\rho_H = \rho(U, 1, 0)$ for hard negatives and $\rho_E = \rho(U, 0, 0)$ for easy negatives, $s_0(x) = \rho_E + (\rho_H - \rho_E)P(Y_C = 1 \mid x, Y_P = 0)$, which is the displayed formula. Because $P(Y_C = 1 \mid x, Y_P = 0) = \big(p_C(x) - P(Y_C = 1, Y_P = 1 \mid x)\big)/(1 - p(x))$ and $p(x)$ is of order 10^{-3} or smaller for production targets, replacing it by the consumption posterior $p_C(x)$ changes the corrected logit by $O(p(x))$, so the consumption head that supplies r_k also supplies the correction at serving time. The correction is increasing in $p_C(x)$ when $\rho_H > \rho_E$, so hard-negative enrichment makes the retained-sample score understate the target-distribution probability most for consumption-likely items and reorders candidates across consumption levels. □

C Calibration under Stratified Negative Sampling

Production training streams retain positives and negatives at rates set within a user group U, with hard negatives (consumption-positive) and easy negatives (consumption-negative) retained at different rates (Appendix H.2). The following statements, kept from the analysis of the industrial pipeline, say what such retention does to the served probability and why the correction must run after retrieval (proofs in Appendices B.9 and B.10).

Proposition C.1 (Retention log-odds correction)**.** *Let $S = 1$ indicate that an observation is retained, $s_y(x) = P(S = 1 \mid X = x, Y = y) > 0$, $p(x) = P(Y = 1 \mid X = x)$, and $p_s(x) = P(Y = 1 \mid X = x, S = 1)$. Then* $\operatorname{logit} p(x) = \operatorname{logit} p_s(x) + \log\big(s_0(x)/s_1(x)\big)$. *If the retention ratio is constant within a stratum, the correction is a constant logit shift within that stratum.*

Corollary C.2 (Label-dependent retention)**.** *If $P(S = 1 \mid x, Y_C, Y_P) = \rho(U, Y_C, Y_P)$ depends on x only through the sampling group U, then $s_0(x)/s_1(x)$ varies with x through $P(Y_C \mid x, Y_P)$ and is constant within U only if the hard- and easy-negative retention rates coincide. When all positives are retained,* $\operatorname{logit} p(x) = \operatorname{logit} p_s(x) + \log\big(\rho_E + (\rho_H - \rho_E)\, P(Y_C = 1 \mid x, Y_P = 0)\big)$.

The correction recovers target-distribution odds under known or consistently estimated retention (Wang et al., 2021) and changes calibration, not the ordering within a stratum, so no monotone correction can repair a non-monotone served score (Appendix H.3). Under label-dependent retention the corrected score needs the consumption posterior, which the head that supplies r_k provides up to

$O(p(x))$ (Appendix B.10). Because the correction depends on the user–item pair, it cannot be folded into Eq. (6) and is a re-ranking correction. Hard-negative enrichment ($\rho_H > \rho_E$) understates the probability most for consumption-likely items and reorders candidates. GSRT's own calibration is the intercept of Section 3.2, fitted on the target distribution, which the public-data study shows to matter more than any transfer decision (Appendix G).

D Certification and the Cost of Selection

A certified variant of GSRT compares the served calibrated candidate with no transfer on certification users disjoint from every other role. Let $e_j \in [-1, 1]$, $j = 1, \ldots, n$, be the per-user mean excess Brier loss of the served candidate over no transfer, $\bar{e}$ its mean and s its standard deviation. Two certificates are analysed.

Empirical-Bernstein certificate (first campaign). For variables of range R the bound (Maurer & Pontil, 2009)

$$U = \bar{e} + \sqrt{\frac{2\widehat{V}\log(2/\delta)}{n}} + \frac{7R\log(2/\delta)}{3(n-1)} \tag{7}$$

holds with probability at least $1 - \delta$, with $\widehat{V} = s^2$, and the certified predictor is the served one if $U \leq 0$ and no transfer otherwise. The range enters a priori rather than through the data. A difference of Brier losses between two predictors whose probabilities are unconstrained lies in $[-1, 1]$, so $R = 2$ throughout this paper.

Proposition D.1 (Certification sample requirement)**.** *With tolerance zero, $U \leq 0$ requires $n - 1 \geq 7R\log(2/\delta)/(3\Delta)$, where $\Delta = -\bar{e} > 0$ is the observed mean improvement in Brier units. At $R = 2$, certifying an improvement of 126 per million at $\delta = 0.05$ needs at least 1.37×10^5 certification users.*

Proof. The range term $7R\log(2/\delta)/(3(n-1))$ is positive and does not shrink with the sample variance, so $U \leq 0$ requires $\bar{e} + 7R\log(2/\delta)/(3(n-1)) \leq 0$, i.e., $n - 1 \geq 7R\log(2/\delta)/(3\Delta)$. With $R = 2$, $\Delta = 1.26 \times 10^{-4}$ and $\log(40) = 3.689$ this is $n - 1 \geq 1.37 \times 10^5$. □

The certification users of every benchmark cell number between 450 and 15,000, an order of magnitude below what Proposition D.1 requires at $R = 2$, so this certificate abstains at these sample sizes and returned no transfer in every cell of every run, in the first campaign and in this one (0 certified cells), exactly as the proposition predicts. That is a property of the bound rather than of the gains. Its variance term is $\sqrt{2\log(2/\delta)} \approx 2.7$ times the standard error of $\bar{e}$, about 14 per million at a standard error near 5, while its range term alone is about 3,800 per million at $n = 4{,}500$. The range term is the one that binds, and R enters it a priori rather than through the data. The realized e_j have a range near 0.05, forty times smaller than the $R = 2$ that unconstrained Brier differences admit, so a protocol that capped the served probability would lower the requirement of Proposition D.1 in proportion. Proposition D.1 is therefore the quantitative reason a distribution-free certificate is the wrong instrument at these sample sizes, and the benchmark certifies with the t test below.

The t certificate. The certificate of Section 3.2 is the one-sided paired t test of $H_0 : \mathbb{E}[e] \geq 0$ at level $\delta = 0.05$. The served candidate is certified when $\bar{e} + t_{1-\delta,\,n-1}\, s/\sqrt{n} \leq 0$ and the certification users carry at least 30 positive events (below that the certificate abstains rather than rely on the normal approximation of the user-level mean). User-level means are bounded and independent, so the asymptotic level-δ approximation is the standard one. Over the regime map it certifies 360 of the 1208 cells in which GSRT serves a transfer candidate (0 of 221 at 3,000 users, 23 of 271 at 10,000, 95 of 286 at 30,000, 242 of 430 at 100,000). The certified cells gain 113.7 per million on average at 30,000 users against 35.9 for the uncertified ones, and 4 of the 360 have negative expected gain, a realized error rate of 1.1% against the nominal 5% (Table 7, Fig. 4). Certification is therefore an available strengthening of the gate rather than a replacement for it. It concentrates on the large gains, at 85.2 [51.4, 119.0] per million in the headline group world against 143.0 [135.8, 150.1] for GSRT, and a deployment that must justify every served coefficient can pay that price.

Cost of selection. Selecting among K calibrated candidates by an out-of-fold Brier loss $\widehat{R}$ costs at most $R(\widehat{h}) - \min_h R(h) \le 2\sup_h |\widehat{R}(h) - R(h)|$. With n controller users and user-level losses in $[0,1]$, Hoeffding's inequality bounds the supremum by $\sqrt{\log(2K/\delta)/(2n)}$ with probability $1-\delta$, and the empirical-Bernstein version replaces the range by the standard deviation of the paired differences, which is smaller by two to three orders of magnitude here because the candidates differ on few rows. The evidence gate tests each of the seven transfer candidates against h_0 at level α separately, so α is a per-candidate level. The conservative rule of Section 3.2 supplies the family-wise version by Bonferroni adjustment, and the benchmark runs both (Table 6). The paired t statistics treat cross-fitted losses as independent across users, the usual approximation for two-fold cross-validation (Bengio & Grandvalet, 2004). Because the gate can move the served candidate in either direction relative to the best one, the benchmark measures the realized regret directly rather than bounding it (Section 4.4), and finds GSRT within 5 per million of the best calibrated candidate in 77% of cells.

E Benchmark, Controllers, and Protocol

E.1 Data-Generating Process

Every user u has a feature vector drawn uniformly from $[-\sqrt{3}, \sqrt{3}]^d$ with $d = 4$, an observed group g drawn uniformly from four groups, an independent hidden group drawn the same way, and four items whose features are independent standard normal. Interactions are represented by $u \odot i$. The shared direction is $e_s = d^{-1/2}\mathbf{1}$ and the private direction e_p is the unit vector proportional to $(0,1,2,3) - 1.5$, orthogonal to e_s. The source (consumption) logit is $1.2\,(u \odot i)^\top e_s$ and the private score $1.2\,(u \odot i)^\top e_p$. The target logit is $a \cdot \mathrm{source} + \sqrt{1-a^2} \cdot \mathrm{private}$ with $a = \mathrm{clip}(0.2 + h\varrho, -1, 1)$, where ϱ depends on the world:

- *group*: ϱ takes the four equally spaced values in $[-1,1]$ indexed by the observed group.
- *continuous*: $\varrho = \tanh(1.5\,u_1)$.
- *hidden*: ϱ is indexed by the hidden group, which is independent of everything a learner sees.
- *proxy* q: ϱ is indexed by the hidden group, and the observed group equals the hidden group with probability $1-q$ and a uniformly random label otherwise ($q = 0$ is the group world, $q = 1$ the hidden world).
- *scale* ω: the group world with every user feature multiplied by ω before scores are formed and the target intercept re-solved to keep the positive rate unchanged, so that cells with the same positives carry ω^2 times the coefficient information.
- *matched*: the group world with the user features of observed group g multiplied by $\omega_g \in (0.4, 0.4, 1.6, 1.6)$ and a per-group target intercept solved so that every group has the same positive rate, which gives the groups equal positives and a sixteenfold ratio of coefficient information.

Intercepts are found by root finding on an independent design population of 30,000 users, before any learner runs, so that the source prevalence is 0.2 and the target prevalence is π, and labels are Bernoulli draws. Each split uses its own random stream, so source, training, controller-fitting, selection, calibration, certification, and test users are independent. In the hidden and proxy worlds the benchmark stores both the privileged probability, which conditions on the hidden group, and the observable Bayes probability, which averages it over $P(\text{hidden} \mid \text{observed group})$. Expected Brier risk is identical under the two by the law of total variance, and only the decomposition into reducible and irreducible parts differs. Cells with $h = 0$ are run in the group world only, since the other worlds coincide there. No true probability, coefficient, or hidden label enters any learner, and true probabilities are used only for the E0 checks, the theory quantities of Appendix F, and the expected-Brier evaluation on the test split.

E.2 Predictors and Controller Objectives

The base predictor is a logistic regression on $[1;\, u \odot i]$ fitted on the target training users with an ℓ_2 penalty of 1.0 on the slopes and a Jeffreys-style pseudocount on the intercept. The residual is the slope vector of the same regression fitted on the source users with the source label, with its intercept discarded, so that $r(u,i) = (u \odot i)^\top \beta_C$. Both are frozen before any controller is fitted. The global and local grid searches minimize the mean squared error between $\sigma(b + \lambda r)$ and the observed label

over the 21-point grid on the controller-fitting rows (all rows for the global value, the rows of group g for λ_g^{loc}, falling back to the global value for an empty group). Information-weighted pooling applies Eq. (2) with $\mathcal{I}_g = \sum_{t \in g} \hat{p}_t(1-\hat{p}_t) r_t^2$, $\hat{p}_t = \sigma(b_t)$, and $\bar{\iota}$ the total information divided by the total positives of the controller rows. Count-weighted pooling replaces $\mathcal{I}_g/\bar{\iota}$ by the positive count n_g^+. Empirical-Bayes pooling uses the weights $\tau^2/(\tau^2 + v_g)$ with v_g the sandwich variance of λ_g^{loc} and τ^2 the method-of-moments estimate of the between-group variance, and has no tuning constant. Unpooled group weights use λ_g^{loc} directly, nonnegative pooling clips each count-pooled candidate at zero, and the hard gate thresholds each count-pooled candidate at $\tau \in \{0, 0.1, 0.25, 0.5\}$. The instance gate minimizes the controller Brier loss plus $\rho\|w\|^2$ by L-BFGS-B from a zero start with $\rho \in \{0.1, 1, 10\}$. The no-transfer candidate is $\lambda \equiv 0$, uniform transfer is $\lambda \equiv 1$, and source-only replacement uses the base intercept plus the residual with the base slope dropped.

Calibration and cross-fitted selection. Every candidate, including the fixed comparators, receives an intercept offset fitted on the calibration users so that its mean predicted probability equals the smoothed observed rate. The controller-fitting users (25%) and the selection users (15%) are two folds. Every family is fitted on each fold, calibrated, and scored on the other fold, so that each of the 40% controller users contributes one out-of-fold user-level Brier loss per candidate. The served candidate of every rule, and every fixed comparator, is refitted on the union of the two folds and calibrated again. Each fixed comparator's hyperparameter (s, τ, or ρ) is chosen by the same out-of-fold losses within its family.

Selection rules. GSRT is the evidence-gated rule of Eq. (3). With e_h the user-level paired out-of-fold excess of candidate h over no transfer, a candidate is eligible when the one-sided paired t test of $\mathbb{E}[e_h] < 0$ has $p < \alpha = 0.05$, and the eligible candidate with the smallest mean out-of-fold loss is served (no transfer when none is eligible). The rule is also computed at $\alpha \in \{0.1, 0.2\}$. The ungated cross-fitted argmin serves the candidate with the smallest mean out-of-fold loss. The one-standard-error rule computes, for every candidate, the user-level paired difference of out-of-fold Brier loss to the best candidate and returns the simplest candidate (in the order no transfer, global, pooling with $s = 100, 20, 5$, gate with $\rho = 10, 1, 0.1$) whose mean difference is at most one standard error. The conservative rule returns the candidate minimizing the one-sided upper confidence limit $\bar{e}_h + t_{1-\delta/7,\, n-1}\, s_h/\sqrt{n}$ of the paired excess (Bonferroni over the seven transfer candidates, with no transfer at limit zero). The empirical-Bernstein rule of the first campaign returns the candidate minimizing $\bar{e}_h + \sqrt{2\widehat{V}_h \log(2K/\delta)/n} + 7R\log(2K/\delta)/(3(n-1))$ with $K = |\mathcal{H}| = 8$, $R = 2$, and $\delta = 0.05$. The split-selection rule of the first campaign fits every family on the controller-fitting users only and serves the argmin of the calibrated Brier loss on the selection users, with the one-standard-error variant on the same split.

Residual-replacement controls. Three controls rerun the complete pipeline, families, calibration, and selection, with the residual replaced by (i) the slope vector of a source model fitted on a random permutation of the source labels, (ii) a random direction with the norm of the true residual, and (iii) the base model's own slope vector, so that the controller can only rescale the target model's score (capacity control).

Representability of the generating relations. The instance gate's input $[1; u; \mathrm{onehot}(g)]$ contains both the observed group and u_1, so $\tanh(w^\top[\cdot])$ represents the continuous-world relation exactly and the group-world relation through the one-hot block. The group families represent the group-world relation exactly and the continuous-world relation only through group-wise averages, which are zero by symmetry. No family represents the hidden-world relation, and in the proxy worlds the group families see the relation attenuated by $(1-q)$.

E.3 Splits, Certification, and Theory Quantities

The total target-user budget is divided 30/25/15/15/15% into training, controller-fitting, selection, calibration, and certification users. Certification computes, on the certification users, the per-user mean of the observed-label Brier difference between the served predictor and no transfer and applies the one-sided paired t test of Appendix D with $\delta = 0.05$ and tolerance zero, provided the certification users carry at least 30 positive events, and the empirical-Bernstein bound of the first campaign is

applied to the same differences as an ablation. For every cell the benchmark also computes, from the fitted base and residual and the true observable probabilities of the test rows, the population best risk of each family (no transfer, the global family on a 201-point grid, group coefficients by coordinate descent on the same grid, and the gate fitted to the true probabilities), and, from the controller rows and their true probabilities, the estimation cost of each family by the sandwich formula of Proposition 3.2. The predicted family minimizes population risk plus estimation cost. The empirical best family is the fixed calibrated candidate of each family with the lowest test risk. The candidate oracle is the member of $\mathcal{H}$ with the lowest test risk, and the selection regret is the test risk of the selected candidate minus the oracle's.

E.4 Predeclared Hypotheses and Diagnostics

Five predictions of Section 3.3 were fixed with the protocol before the confirmatory campaign was run, and are referred to as H1–H5 in Section 4.

- H1 The empirical global/group boundary lies in the band $\frac{1}{4} \leq h^2\mathcal{I} \leq G - 1$ (Corollary 3.3, Theorem 3.4).
- H2 At a fixed positive count the boundary moves with $\mathcal{I}$, that is with ω^2 in the scale worlds.
- H3 The attainable gain decreases monotonically in the proxy noise q and vanishes in the hidden world.
- H4 The evidence gate cuts the negative cells of ungated selection by at least half and keeps at least 90% of the headline gains.
- H5 The t certificate certifies where the served gain is large relative to its standard error and abstains elsewhere, and the empirical-Bernstein certificate abstains everywhere.

Two diagnostics were predeclared with them and are reported alongside. The residual-replacement controls of Appendix E.2 measure what the pipeline collects when the source direction carries no information, and the matched world of Appendix E.1 is built so that information pooling and count pooling weight the groups differently at equal positive counts. Section 4.3 and Section 4.2 report them.

E.5 Evaluation and Statistics

The test split contains 10,000 users (40,000 interactions). For each method the expected Brier loss $\mathbb{E}[(\hat{p} - p^\star)^2 + p^\star(1 - p^\star)]$ is averaged over test rows, and the primary metric is the no-transfer loss minus the method loss, multiplied by 10^6. Every test user contributes four interactions, so this average over rows equals the average over users of their mean gains, and a paired contrast equals the difference of two primary metrics (on public logs the two weightings differ, as Appendix G.3 explains). Per-cell intervals resample users with replacement (300 replicates). Across seeds, means and two-sided Student-t intervals are formed from the cell values. Paired contrasts are formed on identical test users within each seed and tested with a two-sided one-sample t test on the paired differences across seeds. Holm's step-down adjustment is applied to the predeclared primary family (GSRT against global blend, count and information pooling, empirical Bayes, the instance gate, the one-standard-error rule, and the capacity control) within each confirmatory run. Development used seeds 41, 53, and 67 on a reduced grid. The first confirmatory campaign used seeds 101–120, and the second, reported in the main text, uses seeds 201–220. Each was run once, after its protocol, candidate set, and contrast family were frozen. Every cell of the confirmatory runs completed.

Table 3: Hyperparameters and settings of the campaigns.

Component	Setting
Feature dimension, groups, items per user	4, 4, 4
User and item features	Uniform on $[-\sqrt{3}, \sqrt{3}]^4$ (times ω in scale worlds); standard normal
Signal scale, base relation, heterogeneity	1.2; 0.2; $h \in \{0, 0.2, 0.4, 0.6, 0.8\}$ (regime map), $\{0, 0.4, 0.8\}$ (scale)
Source prevalence, target prevalence	0.2; $\pi \in \{0.01, 0.001\}$ (regime map), 0.01 (scale)
Worlds	group, continuous, hidden, proxy $q \in \{0.25, 0.5, 0.75\}$, matched (second campaign); scale $\omega \in \{0.5, 1, 2\}$
Design population for intercepts	30,000 users
Source users, target budget, test users	15,000; $\{3000, 10{,}000, 30{,}000, 100{,}000\}$; 10,000
Target split fractions	train 0.30, controller 0.25, selection 0.15, calibration 0.15, certification 0.15
Base and residual fits	Ridge-logistic, penalty 1.0 on slopes, L-BFGS-B
Coefficient grid	21 points, $\{-1, -0.9, \ldots, 1\}$ (201 points for population optima)
Pooling strengths, gate penalties, hard thresholds	$\{5, 20, 100\}$; $\{0.1, 1, 10\}$; $\{0, 0.1, 0.25, 0.5\}$
Gate level, certification level, tolerance, certification positives	$\alpha = 0.05$ (also 0.1, 0.2); $\delta = 0.05$; 0; ≥ 30
Bootstrap replicates	300
Seeds	development 41, 53, 67; first campaign 101–120; second campaign 201–220
Environment	Python 3.12, NumPy 2.2–2.5, SciPy 1.18, single BLAS thread per worker

F Complete Results of the Benchmark Campaigns

This appendix reports the second campaign in full. It comprises the regime map (seven worlds, two prevalences, four target budgets, five heterogeneity levels, 20 seeds) and the information-scale sweep (three feature scales and the shared world, 20 seeds), both run once on seeds 201–220 after the protocol of Appendix E was frozen, together with the development run used to fix it (seeds 41, 53, 67). Every cell completed, one of them (proxy $q = 0.75$, 100,000 users, seed 213) after a solver restart on the same objective, as the run manifest records. Tables report 20-seed means with 95% Student-t intervals, and the primary metric is the expected Brier gain over no transfer per million predictions on 10,000 independent test users. The numerical check of Theorem 3.4 (Appendix B.8) and the per-cell theory quantities are stored with the run outputs.

Table 4: Paired contrasts in the headline regimes: GSRT minus comparator per million (20-seed mean, 95% interval) and Holm-adjusted p over the predeclared family of 1669 contrasts. "Secondary" rows lie outside that family and are unadjusted, and "identical" means that the predictions coincide with GSRT's in every seed.

GSRT minus	Group world	Holm p	Continuous world	Holm p
Global blend	143.4 [135.5, 151.2]	< 0.001	198.5 [187.8, 209.2]	< 0.001
Information pooling	-0.5 [-1.2, 0.3]	1.000	199.2 [188.5, 209.8]	< 0.001
Empirical Bayes	-0.4 [-1.1, 0.4]	1.000	201.4 [190.3, 212.4]	< 0.001
Instance gate	2.9 [0.9, 4.9]	1.000	identical	–
One-SE rule	2.1 [0.1, 4.0]	1.000	0.1 [-2.1, 2.3]	1.000
Capacity control	112.3 [100.8, 123.9]	< 0.001	194.8 [180.5, 209.0]	< 0.001
Shuffled-source control	138.8 [131.2, 146.4]	secondary	192.8 [179.1, 206.4]	secondary
Random-direction control	95.6 [72.2, 118.9]	secondary	152.8 [118.3, 187.3]	secondary

The first campaign. The same data-generating process, worlds (without the matched world), budgets, prevalences, comparators and controls were first run on seeds 101–120 with the split-selection rule (controllers fitted on the 25% controller-fitting users, the argmin chosen on the 15% selection users, no gate) and the empirical-Bernstein certificate. Its findings reproduce on the fresh seeds of the second campaign (headline gains 126.3 [118.4, 134.2] per million in the group world and 188.5 [175.2, 201.9] in the continuous world), so the regime map itself is confirmed twice on disjoint seeds. It also measured the two quantities that the rules of Section 3.2 were designed around. The

first is what ungated selection costs on the downside, which was negative gain in 1250 of its 4,000 cells (31%, mean -21.3 per million, worst -365.4, 143 cells below -50, and 88% of the negative cells are cells where no candidate gains more than 5 per million). The second is the sample size a distribution-free certificate needs, which Proposition D.1 then quantified. Cross-fitted, evidence-gated selection and the t certificate were specified on the development seeds, the gate level was fixed there ($\alpha = 0.05$, the smallest of $\{0.05, 0.1, 0.2\}$ meeting the downside target on development, while $\alpha = 0.1$ had a marginally smaller development regret but left the worst loss unchanged), and the second campaign was run once. The first campaign's rule and certificate run in every cell of the second campaign as the split-selection and empirical-Bernstein rows, and its complete outputs are archived with the code.

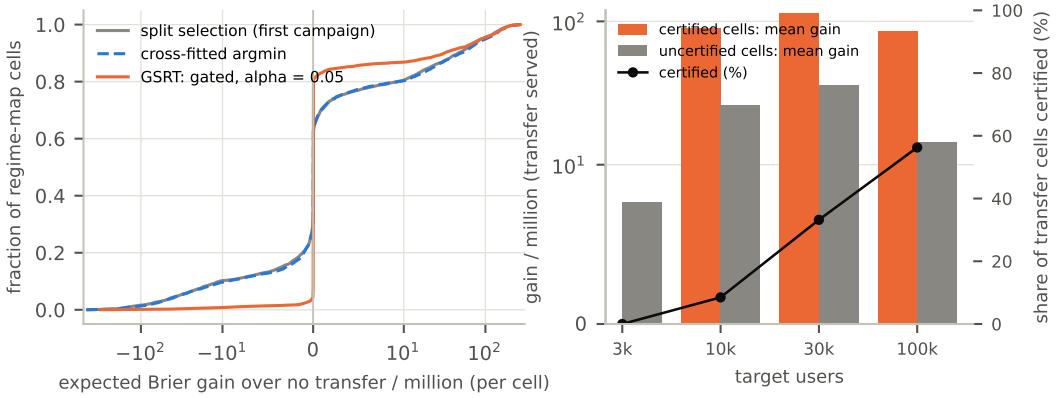


Figure 4: Selection rules and certification over the 4,800 regime-map cells (Section 4.4). Left: distribution of the per-cell gain of the first campaign's split-selection rule, the ungated cross-fitted argmin, and GSRT (symmetric log axis). Right: certification by budget, where bars are the mean gain of the cells the t certificate certifies and declines, and the line is the share of transfer cells certified.

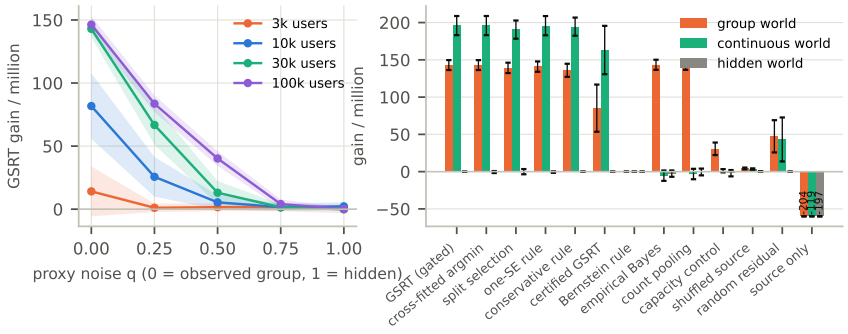


Figure 5: Headline setting, $\pi = 1\%$ and $h = 0.8$ (Section 4.3). Left: GSRT gain against the proxy noise q (bands are 95% intervals). Right: selection rules, pooling comparators, and residual-replacement controls at 30,000 users (bars below -60 clipped).

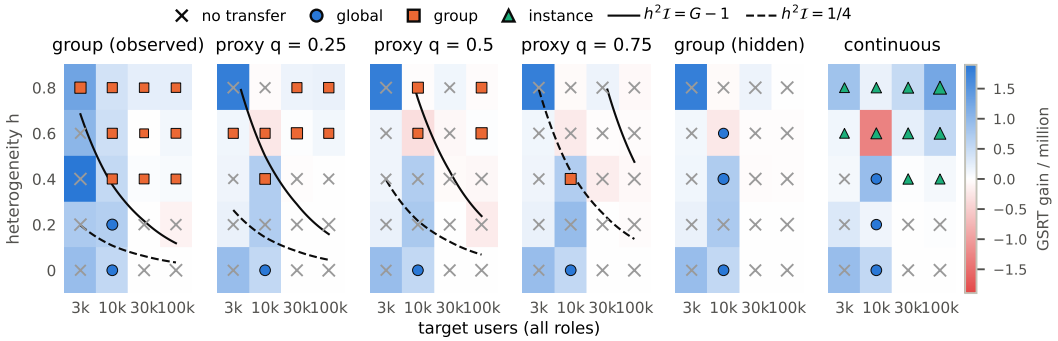


Figure 6: Regime map at $\pi = 0.1\%$ (20 seeds per cell), with the conventions of Fig. 2. At this prevalence the information budget exceeds the criterion only at the largest budgets and heterogeneities.

Table 5: Headline regimes ($\pi = 1\%$, 30,000 target users, $h = 0.8$): expected Brier gain over no transfer per million, 20-seed mean with 95% Student-t interval. No transfer is exactly zero in every cell. GSRT is the evidence-gated cross-fitted rule of Section 3.2, and the split-selection row is the rule of the first campaign.

Method	Group	Continuous	Hidden	Proxy $q = 0.5$	Matched
Uniform ($\lambda = 1$)	−751.5 [−800.4, −702.5]	−623.5 [−676.5, −570.5]	−739.2 [−801.7, −676.7]	−739.2 [−801.7, −676.7]	−1773.4 [−1890.3, −1656.4]
Source only	−204.3 [−215.3, −193.3]	−119.1 [−128.1, −110.0]	−196.8 [−208.7, −184.9]	−196.8 [−208.7, −184.9]	−238.5 [−261.9, −215.1]
Global blend	−0.4 [−4.9, 4.2]	−2.5 [−9.7, 4.6]	1.2 [−4.0, 6.4]	1.2 [−4.0, 6.4]	4.4 [−1.4, 10.3]
Group weights	145.0 [137.8, 152.2]	−8.0 [−16.5, 0.5]	−7.4 [−12.5, −2.3]	37.0 [28.0, 46.1]	1.6 [−12.0, 15.3]
Count pooling	143.3 [136.1, 150.6]	−3.2 [−10.7, 4.3]	−0.5 [−5.3, 4.4]	35.8 [28.5, 43.1]	4.0 [−8.0, 15.9]
Information pooling	143.4 [136.2, 150.7]	−3.2 [−10.7, 4.3]	−0.5 [−5.3, 4.4]	35.7 [28.5, 43.0]	0.6 [−11.6, 12.9]
Empirical Bayes	143.3 [136.1, 150.6]	−5.4 [−12.9, 2.0]	−2.7 [−7.2, 1.9]	35.4 [27.2, 43.6]	0.4 [−13.3, 14.1]
Instance gate	140.1 [132.5, 147.6]	196.0 [182.2, 209.7]	−4.2 [−10.6, 2.3]	21.5 [11.6, 31.5]	−6.0 [−17.8, 5.8]
GSRT (gated, $\alpha = 0.05$)	143.0 [135.8, 150.1]	196.0 [182.2, 209.7]	−0.1 [−0.3, 0.1]	13.1 [3.1, 23.0]	2.5 [−3.0, 8.0]
Cross-fitted argmin	143.0 [135.8, 150.1]	196.0 [182.2, 209.7]	−0.6 [−2.0, 0.8]	27.2 [18.5, 35.9]	1.6 [−9.3, 12.5]
Split selection (v4 rule)	139.2 [131.8, 146.6]	190.5 [177.5, 203.6]	−0.1 [−3.8, 3.7]	29.4 [21.0, 37.8]	2.9 [−9.1, 14.9]
One-SE rule	140.9 [133.5, 148.3]	195.8 [182.1, 209.6]	−0.6 [−1.7, 0.5]	17.8 [7.7, 27.9]	−0.2 [−7.8, 7.5]
Conservative rule	136.0 [126.6, 145.3]	194.5 [181.5, 207.5]	0.0 [0.0, 0.0]	3.1 [−3.4, 9.6]	1.7 [−2.5, 5.9]
Empirical-Bernstein rule	0.0 [0.0, 0.0]	0.0 [0.0, 0.0]	0.0 [0.0, 0.0]	0.0 [0.0, 0.0]	0.0 [0.0, 0.0]
Certified GSRT (t)	85.2 [51.4, 119.0]	163.1 [128.5, 197.8]	0.0 [0.0, 0.0]	5.6 [−1.6, 12.8]	−0.0 [−0.0, 0.0]
Certified GSRT (EB)	0.0 [0.0, 0.0]	0.0 [0.0, 0.0]	0.0 [0.0, 0.0]	0.0 [0.0, 0.0]	0.0 [0.0, 0.0]
Capacity control	30.6 [21.6, 39.6]	1.2 [−1.2, 3.5]	−2.1 [−6.7, 2.4]	0.9 [−4.9, 6.7]	10.6 [2.0, 19.2]
Shuffled source	4.2 [2.8, 5.5]	3.2 [1.9, 4.5]	0.3 [−0.0, 0.6]	1.2 [0.4, 2.0]	8.1 [4.6, 11.6]
Random residual	47.4 [24.3, 70.5]	43.2 [11.7, 74.7]	0.0 [−0.0, 0.0]	2.6 [−1.7, 6.8]	33.2 [−7.0, 73.4]

Table 6: Selection rules over the regime-map cells: mean gain per million, cells with negative gain, mean of the negative cells, worst cell, cells below -50 per million, and mean and median selection regret against the best calibrated candidate. The gated rule at $\alpha = 0.05$ is GSRT, and the split-selection rule is the first campaign's.

Rule	Mean gain	Negative cells	Mean of negatives	Worst	< -50	Regret, mean	Regret, median
GSRT (gated, $\alpha = 0.05$)	10.6	243	−8.9	−313.0	11	7.3	0.3
Gated, $\alpha = 0.1$	10.7	448	−12.1	−313.0	28	7.2	0.3
Gated, $\alpha = 0.2$	9.7	828	−17.3	−468.5	82	8.2	0.4
Cross-fitted argmin	8.3	1429	−19.3	−468.5	158	9.6	0.6
Split selection (v4 rule)	7.1	1473	−19.9	−443.5	163	11.7	0.8
One-SE rule	10.8	462	−13.4	−313.0	33	7.1	0.3
Split one-SE rule	8.8	452	−12.4	−391.2	28	9.1	0.6
Conservative rule	8.7	43	−3.0	−42.5	0	9.1	0.5
Empirical-Bernstein rule	0.0	0	0.0	0.0	0	17.9	0.8
Certified GSRT (t)	6.9	4	−0.4	−0.8	0	10.9	0.5

Table 7: Certification over the regime map by target budget: cells in which GSRT served a transfer candidate, cells among them whose certification users carry at least 30 positives (where the t certificate applies), cells certified by the t certificate and by the empirical-Bernstein certificate, and the mean gain of the served candidate in certified and in uncertified cells (per million).

Users	Transfer served	Eligible	Certified (t)	Certified (EB)	Gain, certified	Gain, uncertified	Certified GSRT, mean
3k	221	0	0	0	–	7.6	0.0
10k	271	108	23	0	88.6	26.0	1.7
30k	286	221	95	0	113.7	35.9	9.0
100k	430	430	242	0	84.3	14.2	17.0
all	1208	759	360	0	92.3	20.8	6.9

Table 8: Matched world ($\pi = 1\%$, equal positives per group, $16\times$ different information): gains of information pooling, count pooling and GSRT per million (20-seed means), and the paired contrast information minus count pooling with its 95% interval.

Users	h	Information pooling	Count pooling	GSRT	Information − count
3k	0.2	-1.1	-3.8	6.6	$2.7\,[-0.9, 6.4]$
3k	0.4	21.2	23.5	9.3	$-2.2\,[-7.4, 2.9]$
3k	0.6	6.3	12.0	-5.6	$-5.6\,[-16.2, 5.0]$
3k	0.8	-36.3	-23.7	-11.7	$-12.6\,[-29.0, 3.8]$
10k	0.2	-0.7	-1.6	0.0	$0.9\,[-0.7, 2.6]$
10k	0.4	1.1	0.5	0.0	$0.6\,[-1.3, 2.5]$
10k	0.6	15.2	16.9	0.0	$-1.7\,[-3.4, -0.0]$
10k	0.8	-11.4	-7.6	0.0	$-3.8\,[-6.3, -1.4]$
30k	0.2	5.9	6.4	3.6	$-0.4\,[-1.4, 0.5]$
30k	0.4	20.6	21.0	8.9	$-0.3\,[-0.9, 0.2]$
30k	0.6	26.7	28.3	17.9	$-1.6\,[-2.7, -0.5]$
30k	0.8	0.6	4.0	2.5	$-3.3\,[-5.3, -1.4]$
100k	0.2	8.5	8.4	3.9	$0.1\,[-0.1, 0.3]$
100k	0.4	24.7	25.5	22.9	$-0.8\,[-1.3, -0.2]$
100k	0.6	34.8	36.2	31.7	$-1.3\,[-2.1, -0.6]$
100k	0.8	6.5	8.2	4.0	$-1.7\,[-2.5, -1.0]$

Table 9: Regime map: GSRT gain over no transfer per million (20-seed mean), the empirically best fixed family (n = no transfer, g = global, G = group, i = instance, by majority over seeds), and the family predicted by Corollary 3.3 (in parentheses), for every world, prevalence, target budget, and heterogeneity. Cells with $h = 0$ are shared by all worlds.

World, rate	Users	$h = 0$	$h = 0.2$	$h = 0.4$	$h = 0.6$	$h = 0.8$
group, 1%	3k	−2.8 n (n)	−4.9 n (n)	−4.9 G (n)	17.8 G (G)	14.1 G (G)
group, 1%	10k	−0.2 n (n)	0.6 G (n)	17.7 G (G)	56.6 G (G)	81.8 G (G)
group, 1%	30k	0.2 n (n)	2.8 G (G)	40.0 G (G)	115.2 G (G)	143.0 G (G)
group, 1%	100k	0.0 n (n)	8.4 G (G)	65.6 G (G)	124.1 G (G)	146.3 G (G)
group, 0.1%	3k	0.9 n (n)	0.7 n (n)	1.9 n (n)	1.0 n (n)	1.3 G (n)
group, 0.1%	10k	0.5 g (n)	0.7 g (n)	0.5 G (n)	0.4 G (n)	0.4 G (n)
group, 0.1%	30k	0.0 n (n)	0.0 n (n)	0.0 G (n)	−0.0 G (n)	0.2 G (n)
group, 0.1%	100k	0.0 n (n)	−0.1 n (n)	0.1 G (n)	0.1 G (G)	0.2 G (G)
continuous, 1%	3k	−2.8 n (n)	−5.2 n (n)	3.1 i (n)	27.0 i (n)	28.4 i (i)
continuous, 1%	10k	−0.2 n (n)	−0.8 i (n)	28.3 i (i)	77.1 i (i)	118.9 i (i)
continuous, 1%	30k	0.2 n (n)	0.0 i (i)	61.5 i (i)	152.5 i (i)	196.0 i (i)
continuous, 1%	100k	0.0 n (n)	7.8 i (i)	85.5 i (i)	165.4 i (i)	208.2 i (i)
continuous, 0.1%	3k	0.9 n (n)	0.3 n (n)	0.1 i (n)	0.1 i (n)	0.8 i (n)
continuous, 0.1%	10k	0.5 g (n)	0.1 g (n)	0.9 g (n)	−1.3 i (n)	0.2 i (n)
continuous, 0.1%	30k	0.0 n (n)	0.0 n (n)	0.0 i (n)	0.2 i (n)	0.5 i (n)
continuous, 0.1%	100k	0.0 n (n)	0.0 n (n)	0.1 i (n)	0.5 i (n)	1.2 i (i)
hidden, 1%	3k	−2.8 n (n)	−0.5 n (n)	0.0 n (n)	0.0 n (n)	0.2 n (n)
hidden, 1%	10k	−0.2 n (n)	0.0 n (n)	−2.6 n (n)	0.0 n (n)	2.2 n (n)
hidden, 1%	30k	0.2 n (n)	−0.7 n (n)	−0.1 n (n)	0.0 n (n)	−0.1 n (g)
hidden, 1%	100k	0.0 n (n)	0.0 n (n)	−0.0 n (n)	−0.3 n (n)	0.1 n (g)
hidden, 0.1%	3k	0.9 n (n)	0.2 n (n)	0.1 n (n)	0.1 n (n)	1.8 n (n)
hidden, 0.1%	10k	0.5 g (n)	0.7 n (n)	0.7 g (n)	−0.2 g (n)	−0.0 n (n)
hidden, 0.1%	30k	0.0 n (n)	−0.0 n (n)	−0.0 n (n)	−0.0 n (n)	0.2 n (n)
hidden, 0.1%	100k	0.0 n (n)	0.0 n (n)	−0.0 n (n)	−0.0 n (n)	−0.0 n (n)
proxy $q = 0.25$, 1%	3k	−2.8 n (n)	−4.3 n (n)	−1.6 n (n)	2.9 G (n)	1.1 G (G)
proxy $q = 0.25$, 1%	10k	−0.2 n (n)	0.0 n (n)	1.0 G (G)	8.4 G (G)	25.7 G (G)
proxy $q = 0.25$, 1%	30k	0.2 n (n)	−0.7 G (G)	8.4 G (G)	48.0 G (G)	66.7 G (G)
proxy $q = 0.25$, 1%	100k	0.0 n (n)	2.6 G (G)	34.0 G (G)	68.6 G (G)	83.6 G (G)
proxy $q = 0.25$, 0.1%	3k	0.9 n (n)	0.2 n (n)	0.1 G (n)	0.1 G (n)	1.8 n (n)
proxy $q = 0.25$, 0.1%	10k	0.5 g (n)	0.7 n (n)	0.7 G (n)	−0.2 G (n)	−0.0 G (n)
proxy $q = 0.25$, 0.1%	30k	0.0 n (n)	−0.0 n (n)	−0.0 n (n)	0.0 G (n)	0.2 G (n)
proxy $q = 0.25$, 0.1%	100k	0.0 n (n)	0.0 n (n)	0.0 G (n)	0.1 G (n)	0.2 G (G)
proxy $q = 0.5$, 1%	3k	−2.8 n (n)	−0.5 n (n)	−2.8 n (n)	0.0 n (n)	1.6 n (n)
proxy $q = 0.5$, 1%	10k	−0.2 n (n)	0.0 n (n)	0.0 G (n)	−2.0 G (G)	5.4 G (G)
proxy $q = 0.5$, 1%	30k	0.2 n (n)	−0.4 n (n)	1.6 G (G)	4.7 G (G)	13.1 G (G)
proxy $q = 0.5$, 1%	100k	0.0 n (n)	0.4 G (G)	11.5 G (G)	28.6 G (G)	40.2 G (G)
proxy $q = 0.5$, 0.1%	3k	0.9 n (n)	0.2 n (n)	0.1 n (n)	0.1 G (n)	1.8 n (n)
proxy $q = 0.5$, 0.1%	10k	0.5 g (n)	0.7 n (n)	0.7 n (n)	−0.3 G (n)	−0.0 G (n)
proxy $q = 0.5$, 0.1%	30k	0.0 n (n)	−0.0 n (n)	−0.0 n (n)	−0.1 n (n)	0.1 n (n)
proxy $q = 0.5$, 0.1%	100k	0.0 n (n)	−0.2 n (n)	−0.1 n (n)	0.0 G (n)	−0.0 G (n)
proxy $q = 0.75$, 1%	3k	−2.8 n (n)	0.0 n (n)	0.0 n (n)	0.0 n (n)	1.4 n (n)
proxy $q = 0.75$, 1%	10k	−0.2 n (n)	−1.6 n (n)	0.0 n (n)	−0.0 n (n)	1.6 n (n)
proxy $q = 0.75$, 1%	30k	0.2 n (n)	−1.3 n (n)	−0.7 n (n)	−0.3 n (n)	1.6 G (G)
proxy $q = 0.75$, 1%	100k	0.0 n (n)	0.0 n (n)	−0.1 G (G)	1.2 G (G)	4.1 G (G)
proxy $q = 0.75$, 0.1%	3k	0.9 n (n)	0.2 n (n)	0.1 n (n)	0.1 n (n)	1.8 n (n)
proxy $q = 0.75$, 0.1%	10k	0.5 g (n)	0.9 n (n)	0.7 G (n)	−0.2 g (n)	−0.1 n (n)
proxy $q = 0.75$, 0.1%	30k	0.0 n (n)	−0.0 n (n)	−0.2 n (n)	−0.0 n (n)	0.2 n (n)
proxy $q = 0.75$, 0.1%	100k	0.0 n (n)	−0.0 n (n)	−0.1 n (n)	−0.0 n (n)	−0.1 n (n)
matched, 1%	3k	−2.8 n (n)	6.6 n (n)	9.3 G (g)	−5.6 G (n)	−11.7 n (n)
matched, 1%	10k	−0.2 n (n)	0.0 n (n)	0.0 G (G)	0.0 G (G)	0.0 n (n)
matched, 1%	30k	0.2 n (n)	3.6 G (G)	8.9 G (G)	17.9 G (G)	2.5 G (G)
matched, 1%	100k	0.0 n (n)	3.9 G (G)	22.9 G (G)	31.7 i (G)	4.0 i (G)
matched, 0.1%	3k	0.9 n (n)	−1.3 n (n)	−1.9 n (n)	−6.6 n (n)	13.1 n (n)
matched, 0.1%	10k	0.5 g (n)	0.1 n (n)	0.9 n (n)	0.9 n (n)	−4.1 n (n)
matched, 0.1%	30k	0.0 n (n)	−1.3 n (n)	−1.2 n (g)	0.2 n (n)	−0.0 n (n)
matched, 0.1%	100k	0.0 n (n)	0.5 G (g)	1.0 G (g)	1.5 G (g)	0.7 n (i)

Table 10: Same positives, different information (scale worlds, $\pi = 1\%$): GSRT gain per million (20-seed mean with 95% interval), the winning family, and the mean controller information $\mathcal{I}$ (positive-equivalents summed over groups).

Scale ω, h	Users	GSRT gain	Best family	$\mathcal{I}$ (pos.-equiv.)
scale $\omega = 0.5$, $h = 0$	3k	0.1 [−0.1, 0.3]	none	46
scale $\omega = 0.5$, $h = 0$	10k	−0.0 [−0.4, 0.3]	global	158
scale $\omega = 0.5$, $h = 0$	30k	−0.1 [−0.3, 0.1]	none	481
scale $\omega = 0.5$, $h = 0$	100k	−0.0 [−0.0, 0.0]	none	1603
scale $\omega = 0.5$, $h = 0.4$	3k	0.7 [−0.4, 1.9]	instance	46
scale $\omega = 0.5$, $h = 0.4$	10k	0.4 [−0.6, 1.3]	group	157
scale $\omega = 0.5$, $h = 0.4$	30k	0.3 [−0.1, 0.8]	group	477
scale $\omega = 0.5$, $h = 0.4$	100k	4.1 [3.3, 4.8]	group	1601
scale $\omega = 0.5$, $h = 0.8$	3k	1.4 [−0.3, 3.1]	group	45
scale $\omega = 0.5$, $h = 0.8$	10k	2.3 [0.4, 4.2]	group	159
scale $\omega = 0.5$, $h = 0.8$	30k	11.4 [9.1, 13.7]	group	479
scale $\omega = 0.5$, $h = 0.8$	100k	14.8 [14.3, 15.3]	group	1606
scale $\omega = 1$, $h = 0$	3k	−2.8 [−8.7, 3.1]	none	47
scale $\omega = 1$, $h = 0$	10k	−0.2 [−2.4, 2.0]	none	161
scale $\omega = 1$, $h = 0$	30k	0.2 [−0.2, 0.5]	none	476
scale $\omega = 1$, $h = 0$	100k	0.0 [0.0, 0.0]	none	1606
scale $\omega = 1$, $h = 0.4$	3k	−4.9 [−15.1, 5.3]	group	47
scale $\omega = 1$, $h = 0.4$	10k	17.7 [5.3, 30.0]	group	160
scale $\omega = 1$, $h = 0.4$	30k	40.0 [26.8, 53.3]	group	479
scale $\omega = 1$, $h = 0.4$	100k	65.6 [63.1, 68.1]	group	1610
scale $\omega = 1$, $h = 0.8$	3k	14.1 [−7.0, 35.3]	group	46
scale $\omega = 1$, $h = 0.8$	10k	81.8 [54.0, 109.6]	group	163
scale $\omega = 1$, $h = 0.8$	30k	143.0 [135.8, 150.1]	group	477
scale $\omega = 1$, $h = 0.8$	100k	146.3 [139.1, 153.6]	group	1616
scale $\omega = 2$, $h = 0$	3k	−6.3 [−19.5, 6.9]	none	49
scale $\omega = 2$, $h = 0$	10k	−0.1 [−0.3, 0.1]	none	163
scale $\omega = 2$, $h = 0$	30k	−0.9 [−2.8, 1.0]	none	483
scale $\omega = 2$, $h = 0$	100k	0.0 [0.0, 0.0]	none	1610
scale $\omega = 2$, $h = 0.4$	3k	226.4 [116.2, 336.6]	group	49
scale $\omega = 2$, $h = 0.4$	10k	438.0 [384.7, 491.3]	group	162
scale $\omega = 2$, $h = 0.4$	30k	517.3 [500.9, 533.6]	group	486
scale $\omega = 2$, $h = 0.4$	100k	533.1 [517.6, 548.5]	instance	1613
scale $\omega = 2$, $h = 0.8$	3k	435.6 [263.9, 607.2]	group	46
scale $\omega = 2$, $h = 0.8$	10k	847.1 [793.2, 900.9]	instance	164
scale $\omega = 2$, $h = 0.8$	30k	886.0 [847.8, 924.3]	instance	484
scale $\omega = 2$, $h = 0.8$	100k	904.4 [859.8, 949.0]	instance	1613

Table 11: Predicted (Proposition 3.2) versus realized estimation cost per family, per million, means over cells of the regime map by prevalence and budget. Realized cost is the test excess of the calibrated fixed candidate over its population risk.

Rate	Users	Global, pred.	Global, real.	Group, pred.	Group, real.	Instance, pred.	Instance, real.
1%	3k	22.6	42.2	80.8	67.0	171.4	109.5
1%	10k	7.4	14.6	27.6	30.4	59.9	52.1
1%	30k	2.5	5.9	9.5	12.4	20.9	19.9
1%	100k	0.8	3.3	2.9	5.5	6.4	7.2
0.1%	3k	4.6	6.2	12.9	6.2	26.0	7.7
0.1%	10k	3.1	3.1	9.4	3.9	18.9	7.5
0.1%	30k	1.3	1.4	4.2	2.3	8.8	4.1
0.1%	100k	0.5	0.8	1.6	1.4	3.5	2.5

Table 12: Information law: share of the group family's population advantage realized by information pooling over global blending, binned by the information budget $h^2\mathcal{I}$ (regime-map and scale cells with an observable group relation and a population advantage of at least one per million, excluding the matched world because its groups differ in scale).

$h^2\mathcal{I}$	Cells	Mean share	Median share
$< 1/4$	12	−14.47	−1.44
$[1/4,\ 3)$	304	−1.34	0.00
$[3,\ 30)$	859	0.32	0.32
≥ 30	947	0.84	0.91

Table 13: Paired contrasts in the headline regimes ($\pi = 1\%$, 30,000 users, $h = 0.8$): GSRT minus comparator per million (20-seed mean, 95% interval) and Holm-adjusted p over the predeclared primary family of 1669 testable contrasts across the whole regime map. Secondary contrasts (rules that share GSRT's candidates) are reported without adjustment.

World	GSRT minus	Δ [95% CI]	Holm p
group	global blend	143.4 [135.5, 151.2]	0.000
group	count pooling	−0.4 [−1.2, 0.4]	1.000
group	information pooling	−0.5 [−1.2, 0.3]	1.000
group	empirical Bayes	−0.4 [−1.1, 0.4]	1.000
group	instance gate	2.9 [0.9, 4.9]	1.000
group	one-SE rule	2.1 [0.1, 4.0]	1.000
group	capacity control	112.3 [100.8, 123.9]	0.000
group	cross-fitted argmin	identical	–
group	split selection (v4 rule)	3.8 [0.3, 7.2]	secondary
group	conservative rule	7.0 [1.4, 12.6]	secondary
group	empirical-Bernstein rule	143.0 [135.8, 150.1]	secondary
group	certified GSRT	57.8 [23.5, 92.1]	secondary
group	GSRT on the count hierarchy	0.1 [0.0, 0.2]	secondary
group	shuffled-source control	138.8 [131.2, 146.4]	secondary
group	random-direction control	95.6 [72.2, 118.9]	secondary
continuous	global blend	198.5 [187.8, 209.2]	0.000
continuous	count pooling	199.2 [188.6, 209.8]	0.000
continuous	information pooling	199.2 [188.5, 209.8]	0.000
continuous	empirical Bayes	201.4 [190.3, 212.4]	0.000
continuous	instance gate	identical	–
continuous	one-SE rule	0.1 [−2.1, 2.3]	1.000
continuous	capacity control	194.8 [180.5, 209.0]	0.000
continuous	cross-fitted argmin	identical	–
continuous	split selection (v4 rule)	5.4 [−0.3, 11.2]	secondary
continuous	conservative rule	1.4 [−1.5, 4.4]	secondary
continuous	empirical-Bernstein rule	196.0 [182.2, 209.7]	secondary
continuous	certified GSRT	32.8 [−5.4, 71.0]	secondary
continuous	GSRT on the count hierarchy	identical	–
continuous	shuffled-source control	192.8 [179.1, 206.4]	secondary
continuous	random-direction control	152.8 [118.3, 187.3]	secondary
hidden	global blend	−1.3 [−6.5, 3.9]	1.000
hidden	count pooling	0.4 [−4.5, 5.2]	1.000
hidden	information pooling	0.4 [−4.5, 5.2]	1.000
hidden	empirical Bayes	2.6 [−2.0, 7.2]	1.000
hidden	instance gate	4.1 [−2.4, 10.5]	1.000
hidden	one-SE rule	0.5 [−0.6, 1.6]	1.000
hidden	capacity control	2.0 [−2.5, 6.6]	1.000
hidden	cross-fitted argmin	0.5 [−0.9, 1.9]	secondary
hidden	split selection (v4 rule)	−0.0 [−3.7, 3.6]	secondary
hidden	conservative rule	−0.1 [−0.3, 0.1]	secondary
hidden	empirical-Bernstein rule	−0.1 [−0.3, 0.1]	secondary
hidden	certified GSRT	−0.1 [−0.3, 0.1]	secondary
hidden	GSRT on the count hierarchy	identical	–
hidden	shuffled-source control	−0.4 [−0.7, −0.0]	secondary
hidden	random-direction control	−0.1 [−0.3, 0.1]	secondary
proxy $q = 0.5$	global blend	11.9 [0.0, 23.7]	1.000
proxy $q = 0.5$	count pooling	−22.7 [−31.7, −13.7]	0.066
proxy $q = 0.5$	information pooling	−22.7 [−31.6, −13.8]	0.060
proxy $q = 0.5$	empirical Bayes	−22.4 [−32.4, −12.3]	0.266
proxy $q = 0.5$	instance gate	−8.5 [−16.9, −0.1]	1.000
proxy $q = 0.5$	one-SE rule	−4.8 [−10.4, 0.9]	1.000
proxy $q = 0.5$	capacity control	12.1 [−0.1, 24.3]	1.000
proxy $q = 0.5$	cross-fitted argmin	−14.1 [−21.3, −6.9]	secondary
proxy $q = 0.5$	split selection (v4 rule)	−16.3 [−24.6, −8.1]	secondary
proxy $q = 0.5$	conservative rule	9.9 [1.9, 18.0]	secondary
proxy $q = 0.5$	empirical-Bernstein rule	13.1 [3.1, 23.0]	secondary
proxy $q = 0.5$	certified GSRT	7.4 [−0.7, 15.6]	secondary
proxy $q = 0.5$	GSRT on the count hierarchy	0.9 [−1.0, 2.9]	secondary
proxy $q = 0.5$	shuffled-source control	11.8 [1.8, 21.9]	secondary
proxy $q = 0.5$	random-direction control	10.5 [−1.0, 22.0]	secondary
matched	global blend	−1.9 [−6.7, 2.8]	1.000
matched	count pooling	−1.5 [−11.8, 8.9]	1.000
matched	information pooling	1.9 [−9.0, 12.7]	1.000
matched	empirical Bayes	2.1 [−10.1, 14.2]	1.000
matched	instance gate	8.5 [−1.9, 18.9]	1.000
matched	one-SE rule	2.7 [−3.7, 9.0]	1.000
matched	capacity control	−8.1 [−16.8, 0.5]	1.000
matched	cross-fitted argmin	0.9 [−8.5, 10.2]	secondary
matched	split selection (v4 rule)	−0.4 [−11.0, 10.2]	secondary
matched	conservative rule	0.8 [−0.5, 2.1]	secondary
matched	empirical-Bernstein rule	2.5 [−3.0, 8.0]	secondary
matched	certified GSRT	2.5 [−3.0, 8.0]	secondary
matched	GSRT on the count hierarchy	0.1 [−1.1, 1.4]	secondary
matched	shuffled-source control	−5.6 [−12.3, 1.1]	secondary
matched	random-direction control	−30.7 [−70.4, 9.0]	secondary

G PUBLIC-DATA STUDY ON KUAIRAND-PURE

G.1 DATA AND TASK

KuaiRand-Pure (Gao et al., 2022) is a public log of a short-video platform released under CC BY-SA 4.0. It contains 27285 users, 7583 videos, 1141112 interactions collected under the platform's standard exposure policy from 8 to 21 April, and 1186059 interactions collected from 22 April to 8 May under *random* exposure, in which the videos shown to a user were drawn uniformly from the pool. Each interaction carries a dense consumption label, the click, and several sparse participation labels, and we use like, comment, forward, and follow as sparse targets. Their prevalence in the random-exposure logs is 0.48%, 0.035%, 0.034%, and 0.026% against 17.6% for the click (1.87%, 0.26%, 0.10%, 0.10%, and 46.3% in the standard-exposure logs), so the targets are two to three orders of magnitude sparser than the source, as in the industrial setting, and the exposure policy changes every rate several-fold. The four targets are participation actions directed at content or creators rather than uploads, so the study exercises the transfer mechanism on participation targets at comparable sparsity. The random-exposure logs are the target distribution of Section 3.1, since they are free of the exposure policy that generated the training stream and later in time. Evaluation uses them exclusively, on users that never appear in any fitting role.

G.2 FEATURES, BASE MODELS, AND RESIDUAL

User inputs are one-hot encodings of the released categorical features (activity degree, live-streamer and video-author flags, follower, fan, friend, and registration-age ranges, and thirteen anonymised one-hot features of small cardinality) plus standardized logarithms of the follower, fan, friend, and registration-day counts. Item inputs are one-hot encodings of video type, upload type, and music type, a multi-hot encoding of the 46 content tags, the standardized logarithm of the duration, and standardized logarithms of sixteen released engagement statistics. The base predictor is a linear two-tower logistic model $z(u, i) = (W_u x_u)^\top (W_i x_i) + b$ of dimension 32 fitted by Adam with binary cross-entropy on the standard-exposure logs of the training users. The residual is the same architecture fitted on the click label of the same users with its intercept dropped. Both are frozen before any controller is fitted, as in the benchmark. The item tower carries content and statistic features rather than a per-video identity block, which on seed 101's selection split raised the random-exposure log loss of both base models by fitting the exposure policy. The training settings (dimension 32, four epochs, learning rate 10^{-2}, weight decay 10^{-4}) were fixed on that split before the campaign and never changed. The target-only models reach test AUCs of 0.646 (like), 0.671 (comment), 0.575 (forward), and 0.718 (follow) under random exposure. The click model's own accuracy drops across exposure policies, one of the shifts the study measures directly.

G.3 PROTOCOL

Users are the sampling unit. Each seed holds out 20% of users as test users and divides the rest, the target-user pool, 30/25/15/15/15% into training, controller-fitting, selection, calibration, and certification users, the fractions of the benchmark. The base and residual models always use the standard-exposure logs of the training users. The four controller roles draw their rows either from the random-exposure logs of their own users, the target-distribution setting, or from the standard-exposure logs of the same users, the training-stream setting that a team without exploration traffic would have to use. A budget of 25% subsamples the target-user pool to emulate label scarcity. The observed group of the group families is the activity degree (9 levels). A group-removed control replaces it by a random label and removes it from every input, the real-data analogue of the hidden-heterogeneity world. Controller families (with information-weighted pooling and the explicit no-transfer candidate), calibration of every candidate before selection, selection, certification, the folding check, across-seed Student-t intervals, paired contrasts, and Holm adjustment are the benchmark's (Appendices E.2–E.5), with the observed Brier loss replacing the expected Brier loss because true probabilities are unknown. Five seeds, four targets, two budgets, two worlds, and two selection sources give 160 cells. Selection is the cross-fitted rule of Section 3.2 with the evidence gate at $\alpha = 0.05$, and the first campaign's split-selection rule runs alongside in every cell. The folding identity held exactly in every cell. Transfer is served in 42 of the cells, the t certificate accepts 11 of them, and the empirical-Bernstein certificate accepts 0, as Proposition D.1 leads one to expect at these sample sizes.

Estimands. Unlike the benchmark's, the test users here differ in their number of interactions, so weighting users equally and weighting interactions equally define different targets. Every gain in this appendix weights test interactions equally, as the no-transfer loss minus the method loss averaged over the test rows of a cell. A paired contrast is the difference of two such gains on the same test rows, and every interval shown is a Student-t interval over seeds. Selection, the evidence gate, and the certificate weight users equally instead. They average each user's losses first and test across users, because users are the sampling unit of every split. Selection and reporting therefore use different weightings here, whereas in the benchmark they coincide.

G.4 RESULTS

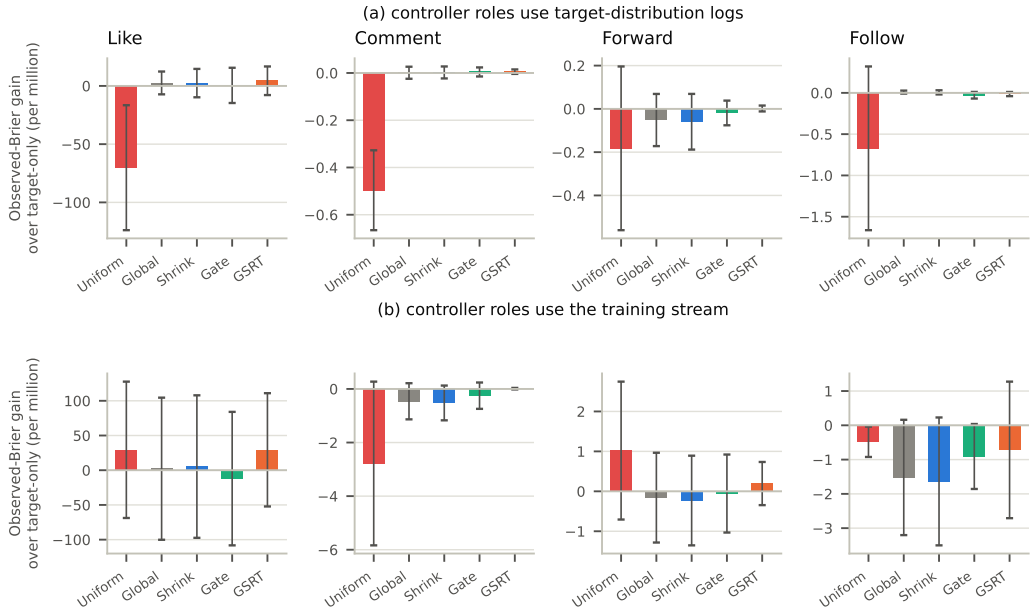


Figure 7: KuaiRand-Pure, observed group, full budget: observed-Brier gain over target-only on the random-exposure logs of held-out users (five seeds, with bars for means and whiskers for 95% Student-t intervals). (a) Controller roles use target-distribution logs. (b) They use the training stream. Note the different vertical scales, as the training-stream setting is far noisier.

Table 14 gives every method in the primary configuration, Table 15 the effect of the selection source, Table 16 the quarter budget and the group-removed control, and Table 17 the paired contrasts. Three findings hold in every configuration. With target-distribution controller rows, uniform transfer never helps and harms likes and comments significantly (70.2 and 0.50 per million), and no selective controller, GSRT included, has an interval above zero. Over the 16 target-distribution configurations GSRT's mean ranges from 0.0 to 4.4 per million with intervals above zero in 0 and below zero in 0. With the explicit no-transfer candidate and calibration before selection, the selector returns no transfer or a candidate indistinguishable from it, without manufacturing a gain. The choice of controller rows dominates. With training-stream rows the intercept alone raises the target-only log loss by 20.8% (like), 34.9% (comment), 18.7% (forward), and 9.6% (follow), and every transfer estimate becomes erratic (GSRT's means range from -124.0 to 29.4 over the stream configurations, uniform transfer on likes at quarter budget is 286.6 $[-295.2, 868.3]$, and the interval of every configuration still covers zero).

Table 14: Primary configuration (observed group, target-distribution controller rows, full budget): observed-Brier gain over target-only per million, five-seed mean with 95% Student-t interval. Target only is exactly zero. Certified GSRT serves the selected rule where the t certificate accepts it and no transfer elsewhere, and the empirical-Bernstein certificate accepts no cell.

Method	Like	Comment	Forward	Follow
Uniform transfer	-70.2 [-123.9, -16.6]	-0.50 [-0.67, -0.33]	-0.18 [-0.56, 0.20]	-0.67 [-1.66, 0.32]
Global blend	2.6 [-7.2, 12.4]	0.00 [-0.02, 0.03]	-0.05 [-0.17, 0.07]	0.01 [-0.01, 0.03]
Group weights	2.5 [-9.8, 14.7]	0.00 [-0.02, 0.03]	-0.06 [-0.19, 0.07]	-0.01 [-0.05, 0.04]
Count pooling	2.4 [-9.8, 14.6]	0.00 [-0.02, 0.03]	-0.06 [-0.19, 0.07]	0.01 [-0.02, 0.03]
Information pooling	2.4 [-9.8, 14.6]	0.00 [-0.02, 0.03]	-0.06 [-0.19, 0.07]	0.01 [-0.02, 0.03]
Nonnegative pooling	2.0 [-10.4, 14.5]	-0.01 [-0.02, 0.01]	-0.06 [-0.19, 0.07]	-0.00 [-0.01, 0.01]
Hard gate	-1.11 [-4.21, 1.98]	0.00 [0.00, 0.00]	-0.07 [-0.21, 0.06]	-0.17 [-0.64, 0.30]
Instance gate	0.5 [-14.7, 15.6]	0.00 [-0.01, 0.02]	-0.02 [-0.08, 0.04]	-0.03 [-0.07, 0.01]
GSRT, count hierarchy	4.4 [-7.9, 16.7]	0.01 [-0.00, 0.02]	0.00 [-0.01, 0.02]	-0.01 [-0.04, 0.01]
GSRT	4.4 [-7.9, 16.7]	0.01 [-0.00, 0.01]	0.00 [-0.01, 0.02]	-0.01 [-0.04, 0.01]
Certified GSRT	4.4 [-7.9, 16.7]	0.00 [0.00, 0.00]	-0.00 [-0.01, 0.00]	0.00 [0.00, 0.00]

Table 15: Effect of the controller-row source (observed group, full budget). Target-only log loss with the intercept fitted on target-distribution rows or on the training stream, and the gains of GSRT and of global blending under each source (per million, five-seed mean with 95% interval).

Target	Log loss, target rows	Log loss, stream rows	Increase (%)	GSRT, target rows	GSRT, stream rows	Global blend, stream rows
Like	0.0287	0.0347	20.8	4.4 [-7.9, 16.7]	29.4 [-52.2, 111.0]	2.2 [-100.2, 104.5]
Comment	0.00284	0.00384	34.9	0.01 [-0.00, 0.01]	0.01 [-0.02, 0.04]	-0.46 [-1.13, 0.21]
Forward	0.00330	0.00392	18.7	0.00 [-0.01, 0.02]	0.19 [-0.35, 0.73]	-0.16 [-1.28, 0.97]
Follow	0.00263	0.00288	9.6	-0.01 [-0.04, 0.01]	-0.72 [-2.71, 1.27]	-1.52 [-3.20, 0.16]

Table 16: Quarter budget and group-removed control (target-distribution controller rows): gains per million, five-seed mean with 95% interval.

Target	Uniform, 25% budget	GSRT, 25% budget	Uniform, group removed	GSRT, group removed	GSRT, group removed, 25%
Like	-85.4 [-109.4, -61.4]	0.00 [0.00, 0.00]	-60.8 [-99.7, -22.0]	0.00 [0.00, 0.00]	1.25 [-2.21, 4.71]
Comment	-1.37 [-2.76, 0.01]	0.01 [-0.01, 0.02]	-0.53 [-0.95, -0.11]	0.00 [-0.00, 0.01]	-0.01 [-0.04, 0.02]
Forward	-0.66 [-1.92, 0.61]	0.01 [-0.01, 0.03]	-0.12 [-0.29, 0.04]	0.00 [-0.00, 0.01]	0.05 [-0.05, 0.15]
Follow	-1.27 [-3.05, 0.52]	0.02 [-0.04, 0.08]	-0.45 [-1.13, 0.23]	0.00 [0.00, 0.00]	0.00 [0.00, 0.00]

Table 17: Paired contrasts in the primary configuration (differences of the gains of Table 14 on the same test rows, per million, with positive favouring GSRT), two-sided paired t tests over five seeds, and Holm-adjusted p over the 12 testable contrasts.

Target	GSRT minus	Δ [95% CI]	p	Holm
Like	global blend	1.86 [-1.44, 5.16]	0.19	1.00
Like	information pooling	2.00 [0.58, 3.42]	0.02	0.21
Like	instance gate	3.97 [0.76, 7.18]	0.03	0.29
Comment	global blend	0.00 [-0.02, 0.03]	0.60	1.00
Comment	information pooling	0.00 [-0.02, 0.03]	0.70	1.00
Comment	instance gate	0.00 [-0.02, 0.02]	0.86	1.00
Forward	global blend	0.05 [-0.06, 0.17]	0.27	1.00
Forward	information pooling	0.06 [-0.06, 0.19]	0.24	1.00
Forward	instance gate	0.02 [-0.03, 0.07]	0.35	1.00
Follow	global blend	-0.02 [-0.06, 0.01]	0.14	1.00
Follow	information pooling	-0.02 [-0.07, 0.02]	0.28	1.00
Follow	instance gate	0.01 [-0.03, 0.06]	0.40	1.00

Other dense sources. Replacing the click by long views or by profile entries as the dense label, with the same protocol in the primary configuration, changes nothing. With profile entries, the closest funnel predecessor of a follow, GSRT gains 0.05 [-0.03, 0.13] per million on follows and 3.96 [-7.03, 14.94] on likes, and uniform transfer loses 61.81 [8.57, 115.04] on likes. With long views, GSRT gains 4.43 [-7.87, 16.73] on likes and 0.00 [-0.00, 0.01] on follows.

G.5 Neural Residual and Jointly Trained Comparators

Two extension runs use the primary configuration (observed group, target-distribution controller and calibration rows) at the full and the quarter budget, with five seeds and four targets each (40 cells per run). Their hyperparameters were fixed before the runs by a probe on the selection split of seed 101 (random-exposure log loss), and the probe outputs are stored with the runs.

Neural residual. An MLP two-tower click model (one hidden layer of width 64 per tower, output dimension 32, 4 epochs, Adam with learning rate 10^{-3} and weight decay 10^{-4}) is fitted on the click label of the training users and supplies the residual in place of the linear click model. The base stays linear, so the comparison isolates the consumption representation. On the probe split the MLP predicts clicks on the random-exposure rows with log loss 0.569 and AUC 0.569, against 0.573 and 0.568 for the linear model, a modestly richer representation of the same signal. Table 18 gives the results. GSRT selects no transfer in 78% of the cells and stays within a few per million of it everywhere. Its means range from 0.00 to 4.67 per million over the 8 configurations, with intervals above zero in 0 and below zero in 0 of them, while uniform transfer of the neural residual loses -58.0 $[-101.5, -14.4]$ per million on likes and -0.33 $[-0.49, -0.17]$ on comments. A better click model does not make the click a better source for these targets.

Jointly trained comparators. Three neural two-tower models are fitted on the same training users: a single-task MLP two-tower model on the target label (neural target-only), a two-tower MMoE with 4 experts per tower and one softmax gate per task, and a two-tower PLE with 2 shared and 1 task-specific experts per task. The MMoE and the PLE are trained jointly on the click and the target label with summed binary cross-entropy. Each task's score is a dot product of a user-only and an item-only embedding plus an intercept, so every comparator is servable from one ANN index per target exactly as GSRT is. Epochs and learning rates were chosen per class by the probe (neural target-only 2 epochs, MMoE 8, PLE 2, with learning rates 10^{-3}, 10^{-3}, and 3×10^{-3}). The MMoE and the PLE have about 69 thousand parameters each, against 17 thousand for the single-task model. Every comparator receives the intercept calibration of the candidates on the calibration users and is scored on the same test rows, and GSRT with the linear residual is recomputed in the same run so that the contrasts are paired. Tables 19 and 20 give the results. On comments, forwards, and follows every comparator stays within a fraction of one per million of the linear target-only model, as GSRT does. On likes, the target with the most positives, the single-task neural model gains 34.2 $[-28.7, 97.1]$ per million at full budget, a capacity effect of the MLP towers that involves no transfer. The PLE matches it (34.5 $[-35.2, 104.1]$), and the MMoE, whose experts serve both tasks, changes the like gain to 4.0 $[-80.1, 88.2]$, which is -30.2 $[-55.9, -4.5]$ per million relative to the single-task model. Over the 20 testable contrasts of the primary configuration, 0 survive Holm correction (smallest adjusted p 0.46). Over the 24 comparator–configuration pairs, 7 intervals lie above zero and 0 below. The intervals above zero are the like and comment gains of the single-task neural model and of the comparators that match it. These are gains of MLP towers over a linear base, which appear without any joint training, rather than gains of transfer.

Capacity-matched comparison. Nothing in GSRT depends on the base being linear. A third run therefore gives GSRT the neural target-only model as its base (same architecture, setting, and seed as the comparator) and the MLP click model as its residual, and refits the MMoE and the PLE in the same run, so that every method is scored against one shared neural base (Tables 21 and 22). Relative to that base, GSRT selects the base itself on likes at full budget in every seed (gain 0.00 [0.00, 0.00] per million), and its means range from 0.00 to 0.00 over 8 configurations, with intervals above zero in 0 and below zero in 0. The MMoE's gain on likes is -32.1 $[-59.8, -4.5]$, because its shared experts cost the sparse head part of what the single-task towers had learned. The PLE, whose task-specific experts can ignore the click, stays within noise of the base (-1.7 $[-12.7, 9.4]$). Of the 8 paired contrasts between GSRT and the two comparators, 0 survive Holm correction (smallest adjusted p 0.26). On one base, selective transfer and joint training arrive at the same place, since there is no consumption signal here to use and a model loses only by sharing regardless.

Table 18: Neural click residual: an MLP two-tower click model (hidden width 64, 4 epochs) replaces the linear click model that supplies the residual (observed group, target-distribution controller rows). Observed-Brier gain over target-only per million, five-seed mean with 95% Student-t interval.

Method	Like	Comment	Forward	Follow
Uniform transfer	-58.0 [-101.5, -14.4]	-0.33 [-0.49, -0.17]	-0.17 [-0.50, 0.16]	-0.71 [-1.93, 0.50]
Global blend	2.0 [-9.3, 13.3]	0.00 [-0.02, 0.02]	-0.05 [-0.17, 0.07]	0.00 [-0.01, 0.02]
Information pooling	2.0 [-9.7, 13.6]	-0.00 [-0.02, 0.02]	-0.06 [-0.18, 0.06]	0.01 [-0.01, 0.02]
Instance gate	0.5 [-15.8, 16.9]	-0.01 [-0.03, 0.01]	-0.02 [-0.07, 0.03]	-0.01 [-0.06, 0.03]
GSRT	4.7 [-8.3, 17.6]	0.00 [-0.00, 0.00]	-0.00 [-0.01, 0.01]	0.00 [-0.01, 0.02]
Uniform transfer, 25% budget	-60.1 [-89.4, -30.8]	-1.04 [-2.14, 0.06]	-0.43 [-1.28, 0.42]	-1.11 [-2.77, 0.55]
Global blend, 25% budget	-10.4 [-21.7, 1.0]	-0.11 [-0.32, 0.10]	-0.07 [-0.21, 0.06]	-0.02 [-0.25, 0.21]
Information pooling, 25% budget	-12.0 [-24.1, 0.1]	-0.12 [-0.35, 0.11]	-0.08 [-0.21, 0.06]	-0.02 [-0.27, 0.22]
Instance gate, 25% budget	-4.4 [-10.0, 1.2]	-0.07 [-0.27, 0.12]	0.00 [-0.03, 0.03]	0.03 [-0.07, 0.13]
GSRT, 25% budget	0.00 [0.00, 0.00]	0.01 [-0.01, 0.02]	0.01 [-0.01, 0.03]	0.00 [0.00, 0.00]

Table 19: Jointly trained two-tower comparators (observed group, target-distribution calibration rows): a single-task MLP two-tower model, a two-tower MMoE with 4 experts per tower, and a two-tower PLE with 2 shared and 1 task-specific experts per task, all fitted on the click and target labels of the training users and calibrated like every candidate. Observed-Brier gain over the linear target-only model per million, five-seed mean with 95% interval, with GSRT (linear residual) recomputed in the same run.

Method	Like	Comment	Forward	Follow
Neural target-only	34.2 [-28.7, 97.1]	0.12 [0.02, 0.23]	0.12 [-0.01, 0.25]	0.17 [-0.21, 0.54]
MMoE (joint)	4.0 [-80.1, 88.2]	0.11 [0.00, 0.21]	0.08 [-0.04, 0.20]	0.04 [-0.36, 0.45]
PLE (joint)	34.5 [-35.2, 104.1]	0.09 [0.00, 0.17]	0.13 [0.01, 0.25]	0.14 [-0.22, 0.50]
GSRT	4.4 [-7.9, 16.7]	0.01 [-0.00, 0.01]	0.00 [-0.01, 0.02]	-0.01 [-0.04, 0.01]
Neural target-only, 25% budget	56.5 [22.4, 90.6]	0.71 [-0.52, 1.93]	0.28 [-0.13, 0.70]	0.22 [-0.19, 0.63]
MMoE (joint), 25% budget	46.4 [20.3, 72.5]	0.66 [-0.53, 1.84]	0.26 [-0.14, 0.65]	0.21 [-0.18, 0.60]
PLE (joint), 25% budget	55.9 [25.2, 86.5]	0.70 [-0.51, 1.92]	0.28 [-0.13, 0.70]	0.21 [-0.19, 0.62]
GSRT, 25% budget	0.00 [0.00, 0.00]	0.01 [-0.01, 0.02]	0.01 [-0.01, 0.03]	0.02 [-0.04, 0.08]

Table 20: Paired contrasts against the jointly trained comparators (full budget, differences of interaction-weighted gains on the same test rows per million, with positive favouring the first method), two-sided paired t tests over five seeds, and Holm-adjusted p over the 20 testable contrasts.

Target	Contrast	Δ [95% CI]	p	Holm
Like	GSRT minus MMoE	0.40 [-71.75, 72.55]	0.99	1.00
Like	GSRT minus PLE	-30.04 [-87.45, 27.37]	0.22	1.00
Like	GSRT minus neural target-only	-29.79 [-80.40, 20.82]	0.18	1.00
Like	MMoE minus neural target-only	-30.19 [-55.88, -4.51]	0.03	0.59
Like	PLE minus neural target-only	0.25 [-7.26, 7.76]	0.93	1.00
Comment	GSRT minus MMoE	-0.10 [-0.21, 0.01]	0.06	1.00
Comment	GSRT minus PLE	-0.08 [-0.17, 0.01]	0.06	1.00
Comment	GSRT minus neural target-only	-0.12 [-0.23, -0.01]	0.04	0.73
Comment	MMoE minus neural target-only	-0.02 [-0.07, 0.04]	0.46	1.00
Comment	PLE minus neural target-only	-0.04 [-0.06, -0.01]	0.02	0.46
Forward	GSRT minus MMoE	-0.08 [-0.21, 0.05]	0.15	1.00
Forward	GSRT minus PLE	-0.13 [-0.25, -0.01]	0.04	0.73
Forward	GSRT minus neural target-only	-0.12 [-0.25, 0.01]	0.07	1.00
Forward	MMoE minus neural target-only	-0.04 [-0.09, 0.01]	0.10	1.00
Forward	PLE minus neural target-only	0.01 [-0.04, 0.05]	0.65	1.00
Follow	GSRT minus MMoE	-0.06 [-0.49, 0.37]	0.73	1.00
Follow	GSRT minus PLE	-0.16 [-0.53, 0.22]	0.31	1.00
Follow	GSRT minus neural target-only	-0.18 [-0.57, 0.21]	0.26	1.00
Follow	MMoE minus neural target-only	-0.12 [-0.47, 0.22]	0.38	1.00
Follow	PLE minus neural target-only	-0.03 [-0.09, 0.03]	0.30	1.00

Table 21: Capacity-matched comparison: GSRT with the neural target-only model of Table 19 as its base and the MLP click model of Table 18 as its residual, against the jointly trained comparators fitted in the same run. Observed-Brier gain over the shared neural target-only base per million, five-seed mean with 95% interval (the base itself is exactly zero).

Method	Like	Comment	Forward	Follow
MMoE (joint)	-32.1 [-59.8, -4.5]	-0.02 [-0.07, 0.04]	-0.04 [-0.09, 0.01]	-0.12 [-0.43, 0.19]
PLE (joint)	-1.7 [-12.7, 9.4]	-0.04 [-0.08, 0.01]	0.01 [-0.04, 0.05]	-0.02 [-0.04, 0.01]
GSRT, neural base	0.00 [0.00, 0.00]	0.00 [0.00, 0.00]	-0.00 [-0.00, 0.00]	0.00 [0.00, 0.00]
MMoE (joint), 25% budget	-6.4 [-12.0, -0.8]	-0.06 [-0.16, 0.04]	-0.02 [-0.03, -0.01]	-0.01 [-0.04, 0.02]
PLE (joint), 25% budget	3.0 [-4.4, 10.5]	-0.01 [-0.03, 0.00]	0.00 [-0.01, 0.02]	-0.00 [-0.01, 0.01]
GSRT, neural base, 25% budget	0.00 [0.00, 0.00]	0.00 [0.00, 0.00]	0.00 [0.00, 0.00]	0.00 [0.00, 0.00]

Table 22: Paired contrasts of the capacity-matched run (full budget, differences of interaction-weighted gains on the same test rows per million, with positive favouring GSRT), two-sided paired t tests over five seeds, and Holm-adjusted p over the 8 testable contrasts.

Target	Contrast	Δ [95% CI]	p	Holm
Like	GSRT minus MMoE	32.11 [4.45, 59.76]	0.03	0.26
Like	GSRT minus PLE	1.67 [-9.38, 12.71]	0.70	1.00
Comment	GSRT minus MMoE	0.02 [-0.04, 0.07]	0.49	1.00
Comment	GSRT minus PLE	0.04 [-0.01, 0.08]	0.11	0.68
Forward	GSRT minus MMoE	0.04 [-0.01, 0.09]	0.10	0.68
Forward	GSRT minus PLE	-0.01 [-0.05, 0.04]	0.61	1.00
Follow	GSRT minus MMoE	0.12 [-0.19, 0.43]	0.35	1.00
Follow	GSRT minus PLE	0.02 [-0.01, 0.04]	0.10	0.68

G.6 Budget Campaign: Scarce Target Labels, Abundant Source Labels

The campaigns above fit the click residual on the training users of the target budget, so that a smaller budget also shrinks the source data. The benchmark's premise is the opposite, with scarce target labels and abundant consumption labels. The budget campaign therefore fits the click residual on the training users and on every user outside the target roles and the test users (the un-subsampled remainder of the pool), keeps controller, selection, calibration, certification and test users out of the residual's sample, and cuts the target budget to 25, 10, and 5% of the users. The full budget is included for continuity (at the full budget there is no remainder and the residual is the one above). The rules are those of the second benchmark campaign, with cross-fitted selection over the controller-fitting and selection users, the evidence gate at the protocol's level $\alpha = 0.05$, the t certificate on the certification users, and the first campaign's split-selection rule alongside. The campaign uses twenty seeds, four targets, the observed group, and target-distribution controller rows, and the source-users setting of the first campaign is rerun at the two smallest budgets as an ablation. It was run in two parts, five seeds (101–105) first and then fifteen more seeds (106–120) with the same script and settings, in a single extension declared before it ran and not enlarged afterwards, and every table reports the twenty seeds together. With twenty seeds the 10% like interval excludes zero and the 5% interval does not (16.6 [-1.7, 34.8] per million), so only the 10% budget is read as a signal anywhere in this paper. Tables 23–27 report every number, and Section 4.6 summarizes them.

Item-feature ablation. The released item statistics include monthly counts of likes, comments, follows and shares, the same actions that serve as targets. The campaign was therefore rerun with those four counts removed from the item tower, and with static content features only (Table 27). With the released statistics 3 of the 6 small-budget intervals on likes, forwards and follows are above zero. Without the four counts only 2 are, both on follows at a fraction of one per million, and with static features none is. The target-only model improves as the features are removed (like AUC at 10% of 0.575, 0.591, and 0.636), because the behavioural aggregates hurt a target model fitted on a few hundred positives, and the residual repairs part of that damage. The ablation therefore identifies exactly what the small-budget gain is made of, and we credit it to those features rather than to the click, which is the conservative reading in both directions. Under every feature set GSRT either serves a candidate that helps or returns no transfer, and never pays for the decision.

Table 23: Budget campaign on KuaiRand-Pure (click residual fitted on the training users and on every user outside the target roles, observed group, target-distribution controller rows): observed-Brier gain over target-only per million, 20-seed mean with 95% Student-t interval, by target budget. Target only is exactly zero. The bottom block gives GSRT's gain as a percentage of the target-only Brier loss and the training positives.

Budget	Method	Like	Comment	Forward	Follow
100%	Uniform transfer	−62.0 [−73.6, −50.4]	−0.44 [−0.56, −0.32]	−0.05 [−0.17, 0.06]	−0.35 [−0.70, 0.00]
	Global blend	−0.11 [−2.03, 1.82]	0.01 [−0.01, 0.03]	0.03 [−0.04, 0.10]	0.03 [−0.00, 0.06]
	Information pooling	−0.82 [−3.19, 1.55]	0.01 [−0.01, 0.03]	0.03 [−0.04, 0.10]	0.03 [−0.00, 0.06]
	Instance gate	−2.5 [−5.3, 0.4]	0.01 [−0.00, 0.03]	0.05 [−0.03, 0.12]	−0.01 [−0.04, 0.01]
	Split selection (v4 rule)	0.95 [−1.48, 3.38]	0.01 [−0.01, 0.02]	0.05 [−0.03, 0.12]	0.01 [−0.01, 0.03]
	Cross-fitted argmin	0.90 [−1.49, 3.29]	0.02 [0.00, 0.03]	0.05 [−0.02, 0.12]	−0.01 [−0.03, 0.01]
	GSRT	1.18 [−1.13, 3.49]	0.01 [−0.00, 0.02]	0.05 [−0.02, 0.12]	0.00 [−0.01, 0.02]
	Certified GSRT	1.11 [−1.21, 3.42]	0.00 [0.00, 0.00]	0.04 [−0.03, 0.12]	0.00 [0.00, 0.00]
25%	Uniform transfer	−68.3 [−92.3, −44.2]	−0.94 [−1.41, −0.46]	−0.65 [−1.13, −0.17]	−1.19 [−2.14, −0.25]
	Global blend	−1.49 [−3.59, 0.61]	−0.00 [−0.07, 0.07]	−0.02 [−0.18, 0.14]	0.17 [−0.15, 0.49]
	Information pooling	−2.28 [−4.31, −0.26]	0.00 [−0.07, 0.07]	−0.03 [−0.19, 0.14]	0.11 [−0.10, 0.32]
	Instance gate	1.1 [−4.8, 7.0]	0.03 [−0.03, 0.10]	0.16 [−0.07, 0.38]	0.19 [−0.15, 0.54]
	Split selection (v4 rule)	2.2 [−3.5, 7.8]	0.05 [−0.01, 0.10]	0.11 [−0.09, 0.31]	0.03 [0.00, 0.06]
	Cross-fitted argmin	1.5 [−4.3, 7.3]	0.05 [−0.01, 0.11]	0.17 [−0.04, 0.38]	0.20 [−0.15, 0.54]
	GSRT	2.6 [−2.9, 8.2]	0.03 [−0.02, 0.08]	0.05 [−0.02, 0.12]	0.02 [−0.00, 0.05]
	Certified GSRT	2.6 [−2.9, 8.2]	0.00 [0.00, 0.00]	0.00 [0.00, 0.00]	0.00 [0.00, 0.00]
10%	Uniform transfer	−29.9 [−51.6, −8.3]	0.01 [−0.41, 0.43]	0.55 [0.02, 1.07]	0.45 [−0.64, 1.54]
	Global blend	5.1 [−5.0, 15.1]	0.28 [−0.02, 0.58]	0.66 [0.21, 1.11]	0.98 [0.22, 1.74]
	Information pooling	4.6 [−5.6, 14.9]	0.29 [−0.02, 0.59]	0.65 [0.20, 1.11]	0.99 [0.23, 1.76]
	Instance gate	11.1 [0.9, 21.2]	0.32 [0.02, 0.61]	0.69 [0.25, 1.12]	1.08 [0.34, 1.83]
	Split selection (v4 rule)	10.2 [0.6, 19.9]	0.11 [−0.08, 0.31]	0.61 [0.15, 1.07]	1.09 [0.35, 1.84]
	Cross-fitted argmin	10.3 [1.6, 19.0]	0.25 [−0.02, 0.52]	0.69 [0.24, 1.14]	1.03 [0.25, 1.80]
	GSRT	9.3 [0.5, 18.1]	0.10 [−0.10, 0.31]	0.50 [0.10, 0.90]	1.07 [0.30, 1.84]
	Certified GSRT	0.91 [−1.00, 2.82]	0.00 [0.00, 0.00]	0.00 [0.00, 0.00]	0.00 [0.00, 0.00]
5%	Uniform transfer	−2.4 [−28.9, 24.1]	−0.99 [−2.31, 0.33]	−0.14 [−0.87, 0.59]	0.62 [−0.82, 2.07]
	Global blend	16.3 [−2.0, 34.5]	0.11 [−0.92, 1.14]	0.22 [−0.32, 0.76]	1.07 [−0.17, 2.31]
	Information pooling	16.2 [−2.2, 34.5]	0.10 [−0.94, 1.13]	0.21 [−0.34, 0.76]	1.07 [−0.17, 2.31]
	Instance gate	16.9 [−1.4, 35.3]	0.72 [−0.89, 2.32]	0.35 [−0.12, 0.81]	1.06 [−0.16, 2.28]
	Split selection (v4 rule)	17.6 [−0.6, 35.8]	0.75 [−0.77, 2.27]	0.31 [−0.18, 0.81]	1.07 [−0.15, 2.30]
	Cross-fitted argmin	17.1 [−1.0, 35.1]	0.74 [−0.87, 2.35]	0.34 [−0.16, 0.85]	1.08 [−0.15, 2.32]
	GSRT	16.6 [−1.7, 34.8]	0.56 [−0.33, 1.45]	0.31 [−0.17, 0.79]	1.06 [−0.18, 2.30]
	Certified GSRT	0.91 [−0.99, 2.81]	0.00 [0.00, 0.00]	0.00 [0.00, 0.00]	0.00 [0.00, 0.00]
100%	GSRT, % of target-only Brier	0.02 [−0.02, 0.07]	0.00 [−0.00, 0.00]	0.01 [−0.00, 0.03]	0.00 [−0.01, 0.01]
25%	GSRT, % of target-only Brier	0.05 [−0.06, 0.16]	0.01 [−0.01, 0.02]	0.01 [−0.00, 0.03]	0.01 [−0.00, 0.02]
10%	GSRT, % of target-only Brier	0.19 [0.01, 0.36]	0.03 [−0.03, 0.09]	0.15 [0.03, 0.26]	0.35 [0.11, 0.60]
5%	GSRT, % of target-only Brier	0.35 [−0.03, 0.72]	0.16 [−0.08, 0.40]	0.08 [−0.04, 0.21]	0.37 [−0.09, 0.84]
100%	Training positives (mean)	5118	703	282	276
25%	Training positives (mean)	1260	175	74	69
10%	Training positives (mean)	487	70	29	26
5%	Training positives (mean)	260	38	15	13

Table 24: Selection and certification in the budget campaign: seeds (of 20) in which GSRT served a transfer candidate, in which the t certificate certified it, and in which the empirical-Bernstein certificate did, together with the candidate served most often.

Budget	Target	Transfer served	Certified (t)	Certified (EB)	Most served
100%	Like	2	1	0	target only
100%	Comment	4	0	0	target only
100%	Forward	11	5	0	target only
100%	Follow	3	0	0	target only
25%	Like	1	1	0	target only
25%	Comment	3	0	0	target only
25%	Forward	5	0	0	target only
25%	Follow	5	0	0	target only
10%	Like	8	1	0	target only
10%	Comment	5	0	0	target only
10%	Forward	12	0	0	target only
10%	Follow	11	0	0	target only
5%	Like	6	1	0	target only
5%	Comment	10	0	0	target only
5%	Forward	8	0	0	target only
5%	Follow	11	0	0	target only

Table 25: Paired contrasts in the budget campaign (GSRT minus comparator, as differences of the gains of Table 23 on the same test rows, per million, 20-seed mean with 95% interval), two-sided paired t tests and Holm-adjusted p over the 64 testable contrasts.

Budget	Target	GSRT minus	Δ [95% CI]	p	Holm
100%	Like	global blend	1.29 [0.55, 2.02]	0.002	0.101
100%	Like	split selection	0.23 [−0.18, 0.65]	0.256	1.000
100%	Like	instance gate	3.65 [2.73, 4.58]	0.000	0.000
100%	Like	information pooling	2.00 [1.43, 2.57]	0.000	0.000
100%	Comment	global blend	−0.01 [−0.02, 0.01]	0.428	1.000
100%	Comment	split selection	0.00 [−0.01, 0.01]	0.908	1.000
100%	Comment	instance gate	−0.01 [−0.02, 0.01]	0.361	1.000
100%	Comment	information pooling	−0.01 [−0.02, 0.01]	0.487	1.000
100%	Forward	global blend	0.02 [−0.01, 0.04]	0.224	1.000
100%	Forward	split selection	0.01 [−0.00, 0.01]	0.072	1.000
100%	Forward	instance gate	0.00 [−0.01, 0.02]	0.581	1.000
100%	Forward	information pooling	0.02 [−0.01, 0.05]	0.140	1.000
100%	Follow	global blend	−0.03 [−0.05, −0.01]	0.014	0.871
100%	Follow	split selection	−0.01 [−0.02, 0.01]	0.457	1.000
100%	Follow	instance gate	0.02 [−0.01, 0.05]	0.226	1.000
100%	Follow	information pooling	−0.02 [−0.04, −0.00]	0.038	1.000
25%	Like	global blend	4.1 [−0.3, 8.5]	0.064	1.000
25%	Like	split selection	0.46 [−0.44, 1.35]	0.299	1.000
25%	Like	instance gate	1.55 [0.15, 2.94]	0.032	1.000
25%	Like	information pooling	4.9 [0.4, 9.4]	0.035	1.000
25%	Comment	global blend	0.03 [−0.02, 0.08]	0.196	1.000
25%	Comment	split selection	−0.02 [−0.04, 0.01]	0.149	1.000
25%	Comment	instance gate	−0.01 [−0.05, 0.04]	0.791	1.000
25%	Comment	information pooling	0.03 [−0.02, 0.07]	0.257	1.000
25%	Forward	global blend	0.07 [−0.07, 0.21]	0.310	1.000
25%	Forward	split selection	−0.06 [−0.25, 0.13]	0.499	1.000
25%	Forward	instance gate	−0.11 [−0.33, 0.12]	0.329	1.000
25%	Forward	information pooling	0.07 [−0.07, 0.22]	0.293	1.000
25%	Follow	global blend	−0.14 [−0.46, 0.17]	0.354	1.000
25%	Follow	split selection	−0.01 [−0.03, 0.01]	0.355	1.000
25%	Follow	instance gate	−0.17 [−0.52, 0.18]	0.327	1.000
25%	Follow	information pooling	−0.08 [−0.29, 0.12]	0.404	1.000
10%	Like	global blend	4.2 [0.1, 8.3]	0.045	1.000
10%	Like	split selection	−0.93 [−4.39, 2.54]	0.581	1.000
10%	Like	instance gate	−1.8 [−5.1, 1.6]	0.284	1.000
10%	Like	information pooling	4.6 [0.4, 8.9]	0.033	1.000
10%	Comment	global blend	−0.18 [−0.42, 0.07]	0.157	1.000
10%	Comment	split selection	−0.01 [−0.03, 0.01]	0.297	1.000
10%	Comment	instance gate	−0.21 [−0.45, 0.02]	0.075	1.000
10%	Comment	information pooling	−0.18 [−0.44, 0.08]	0.159	1.000
10%	Forward	global blend	−0.16 [−0.44, 0.12]	0.253	1.000
10%	Forward	split selection	−0.11 [−0.42, 0.20]	0.461	1.000
10%	Forward	instance gate	−0.19 [−0.46, 0.09]	0.167	1.000
10%	Forward	information pooling	−0.16 [−0.44, 0.12]	0.259	1.000
10%	Follow	global blend	0.09 [−0.11, 0.28]	0.361	1.000
10%	Follow	split selection	−0.02 [−0.08, 0.03]	0.363	1.000
10%	Follow	instance gate	−0.01 [−0.07, 0.05]	0.664	1.000
10%	Follow	information pooling	0.08 [−0.12, 0.27]	0.409	1.000
5%	Like	global blend	0.29 [−2.01, 2.58]	0.797	1.000
5%	Like	split selection	−1.03 [−2.15, 0.09]	0.069	1.000
5%	Like	instance gate	−0.37 [−2.59, 1.86]	0.735	1.000
5%	Like	information pooling	0.38 [−2.07, 2.82]	0.751	1.000
5%	Comment	global blend	0.45 [−0.07, 0.97]	0.085	1.000
5%	Comment	split selection	−0.19 [−0.84, 0.46]	0.552	1.000
5%	Comment	instance gate	−0.16 [−0.90, 0.58]	0.660	1.000
5%	Comment	information pooling	0.46 [−0.06, 0.98]	0.077	1.000
5%	Forward	global blend	0.09 [−0.10, 0.28]	0.329	1.000
5%	Forward	split selection	−0.00 [−0.14, 0.14]	0.975	1.000
5%	Forward	instance gate	−0.03 [−0.16, 0.10]	0.597	1.000
5%	Forward	information pooling	0.10 [−0.10, 0.31]	0.314	1.000
5%	Follow	global blend	−0.01 [−0.05, 0.03]	0.503	1.000
5%	Follow	split selection	−0.01 [−0.05, 0.03]	0.486	1.000
5%	Follow	instance gate	−0.00 [−0.05, 0.04]	0.898	1.000
5%	Follow	information pooling	−0.01 [−0.05, 0.02]	0.463	1.000

Table 26: Source-data ablation at the small budgets: the click residual fitted on the training users only (so that it shrinks with the target budget, the first campaign's setting) against the residual fitted on every user outside the target roles. GSRT gain per million, 20-seed mean with 95% interval.

Budget	Residual fitted on	Like	Comment	Forward	Follow
10%	training users only	9.0 [1.1, 16.9]	0.20 [−0.13, 0.52]	0.36 [−0.02, 0.74]	0.79 [0.13, 1.45]
10%	all users outside the roles	9.3 [0.5, 18.1]	0.10 [−0.10, 0.31]	0.50 [0.10, 0.90]	1.07 [0.30, 1.84]
5%	training users only	6.0 [−3.9, 16.0]	0.33 [−0.13, 0.79]	0.37 [−0.11, 0.85]	0.53 [−0.35, 1.41]
5%	all users outside the roles	16.6 [−1.7, 34.8]	0.56 [−0.33, 1.45]	0.31 [−0.17, 0.79]	1.06 [−0.18, 2.30]

Table 27: Item-feature ablation of the budget campaign (same seeds, splits and rule). The released item statistics include monthly counts of the same actions that serve as targets. The rows remove those four counts, and then every behavioural statistic. Target-only AUC on likes (random-exposure test rows) and GSRT gain per million by target, 20-seed mean with 95% Student-t interval. Comments have no interval above zero under any feature set.

Budget	Item features	Item dimension	Like AUC	Like	Forward	Follow
10%	All released statistics	86	0.575	9.29 [0.47, 18.12]	0.50 [0.10, 0.90]	1.07 [0.30, 1.84]
10%	Without the four target-label counts	82	0.591	2.14 [−0.59, 4.86]	0.26 [−0.01, 0.52]	0.27 [0.08, 0.45]
10%	Static content features only	70	0.636	0.21 [−0.29, 0.70]	0.06 [−0.02, 0.14]	0.31 [−0.13, 0.76]
5%	All released statistics	86	0.541	16.55 [−1.66, 34.76]	0.31 [−0.17, 0.79]	1.06 [−0.18, 2.30]
5%	Without the four target-label counts	82	0.563	2.53 [−0.21, 5.27]	0.17 [−0.04, 0.39]	0.47 [0.02, 0.91]
5%	Static content features only	70	0.621	0.65 [−0.75, 2.04]	0.26 [−0.04, 0.55]	0.13 [−0.07, 0.33]

What the study establishes. On these logs the selector does not invent a gain when the residual carries nothing, and forced transfer is the only controller that loses substantially. This holds for a linear and for a neural click residual, and jointly trained shared-expert models of the MMoE and PLE kind find no signal that the selector misses. On data with a genuine random-exposure period it further establishes that the training-stream shift of Corollary C.2 is the first-order effect a production team would face. With the target labels cut to a tenth GSRT lifts the like predictor by 9.3 [0.5, 18.1] per million over twenty seeds, and the feature ablation attributes that lift to item statistics which aggregate the target labels rather than to the click (Appendix G.6, Table 27). At the full and 25% budgets the click signal a participation target could borrow is absent, which is the regime in which the criterion of Section 3.3 calls for no transfer and in which GSRT returns it. The production logs of Appendix I supply the opposite regime, where the budget is present and the same rule serves a coefficient that gains.

H Industrial Record and Validation Protocol

The design of GSRT was motivated by an anonymised industrial record from a short-video platform. Every number in this appendix is transcribed from that source and checked against its text layer and rendered pages, with the source's displayed precision retained (Appendix H.6). The record supplies the four facts GSRT is built around: earlier online tests that moved consumption far more than production (Appendix H.1), sparse targets under stratified negative sampling (Appendix H.2), served scores that cannot be trusted as probabilities (Appendix H.3), and a bundled deployment whose components cannot be separated after the fact (Appendix H.4). It also fixes the validation protocol of Appendix H.5. The platform's own logs enter this paper separately, in the offline ranking study of Appendix I, and this appendix reports only what the source document states.

H.1 Prior Online Tests

Table 28: Three earlier online tests reported by the source (relative changes in percent). "Strict" denotes strict-caliber trend plays, and efficiency tiers are the source's own categories.

Test	Stage	Days	Play metrics (%)	Creator metrics (%)
Retrieval shaping	retrieval	4	strict +1.774; high-efficiency strict +3.173	upload authors +0.119
Ranking dynamic dual channel	ranking	6	strict +0.354; high-efficiency strict +0.924	creator DAU +0.049; uploaded works −3.858
Re-ranking generator	re-ranking	13	strict +0.51; high/mid/low efficiency +1.447/+0.847/+0.246	upload authors +0.159

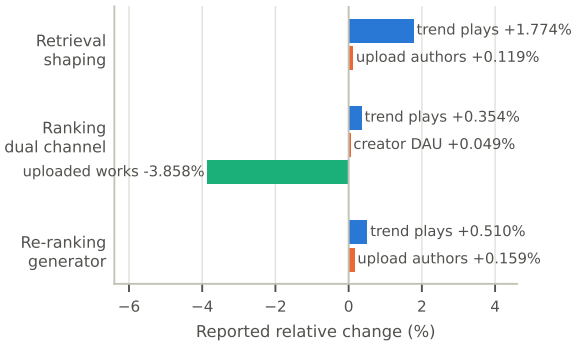


Figure 8: Three earlier online tests reported by the industrial source (Table 28). Plays rose in every test, while creator metrics moved far less.

An exposure initiative raised the share of trend content in total plays from 0.2% to 0.4% and then met a conversion bottleneck. Three earlier online tests raised trend-content plays by 0.354–1.774% while upload-author counts rose by at most 0.159%, and one ranking test lost 3.858% of uploaded works (Table 28, Fig. 8). The source annotates the first and third tests as converting plays into creator activity at a very low rate and the second as raising plays while lowering the number of uploaded works. These annotations rather than the numbers alone motivated a production-specific retrieval path. Play metrics respond far more readily than creator metrics, so any production endpoint must be a user-level outcome.

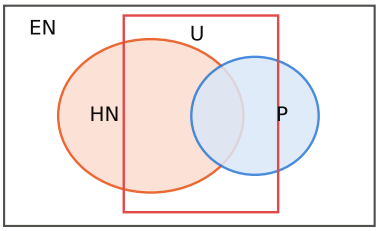
H.2 Sparsity of Production Targets

Table 29: Positive rates by production target and sampling scope, as displayed by the source.

Target	Current	Filter removed	Core target users	Core non-target users	Social core target users	Social core non-target users
Strict-caliber upload	0.00030	0.000021	0.000034	1.6272E−08	0.000033	4.4098E−08
Second-hop collection	0.0024	0.00065	0.0014	0.000012	0.0016	0.000019
Music one-hop entry	0.0048	0.0028	0.0049	0.0013	0.0053	0.0015
Effect-template one-hop entry	0.00030	0.000063	0.00010	1.6191E−07	0.000083	9.1471E−07



(a) Positive rates by target and sampling scope (log scale).



(b) Sampler definitions: positives P, hard negatives HN, easy negatives EN, user group U.

Figure 9: Industrial sparsity and sampling definitions, redrawn from the source.

Positive rates span five orders of magnitude across targets and sampling scopes, from 4.8×10^{-3} to 1.6×10^{-8} (Table 29), and prevalence depends on both the action and the selected population, so a per-target controller is necessary at the formulation level even though the benchmark evaluates one production target at a time. The sampler defines $P = \bigcup_k P_k$ and $N = HN \cup EN$ and sets per-target positive shares and the $HN : EN$ ratio within a user group U (Fig. 9), so hard (consumption-positive) and easy negatives are retained at different rates within a user group, which is the setting of Appendix C.

H.3 Score-Trust Diagnostics

Fig. 10 sets the served score against realized strict trend plays (Tables 30 and 31). In natural traffic the four buckets with scores at or above 0.7 contained 0, 0, 1, and 13 strict trend plays while the score-zero bucket held 29.60% of them. The production dual channel shows the same non-monotone shape, with higher realized ratios in the mid-range than in the top buckets. Precision and recall therefore fail together, and by Proposition C.1 no monotone correction can repair the ordering. Exposure guarantees also shape which items reach the dual channel, so its buckets describe a constrained channel. Negatives unseen in training scored higher than training-distribution negatives (Table 32, Fig. 11), and seeding videos receive far more plays across the whole distribution (Table 33), so item popularity, exposure, and observed production outcomes are tightly coupled in the logs and memorized exposure is a live explanation. A non-resampled temporal test set is needed to tell transferable behaviour from memorized exposure. These observations fix the four requirements GSRT is built to meet.

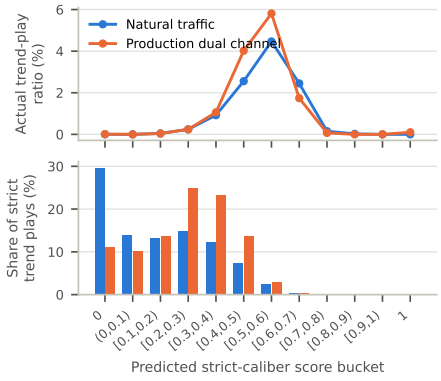


Figure 10: Served score versus realized strict trend plays (Tables 30–31).

Table 30: Natural-traffic score buckets for strict trend plays (single-column traffic outside the dual channel). The source marks the trend-play shares from $[0.4, 0.5)$ down to 0 as missed recall.

Score bucket	Total plays	Play share (%)	Strict trend plays	Trend-play share (%)	Actual ratio (%)
1	25914	0.00	0	0.00	0.000
[0.9,1)	7274	0.00	0	0.00	0.000
[0.8,0.9)	4107	0.00	1	0.00	0.024
[0.7,0.8)	8568	0.00	13	0.01	0.152
[0.6,0.7)	17957	0.00	440	0.46	2.450
[0.5,0.6)	54384	0.00	2428	2.54	4.465
[0.4,0.5)	277568	0.01	7101	7.44	2.558
[0.3,0.4)	1272145	0.03	11741	12.30	0.923
[0.2,0.3)	5833784	0.15	14073	14.74	0.241
[0.1,0.2)	27506336	0.72	12562	13.15	0.046
(0,0.1)	2768660452	72.47	13237	13.86	0.000
0	651809470	17.06	28263	29.60	0.004
Null	364977168	9.55	5635	5.90	0.002
Total	3820455127	100.00	95494	100.00	0.002

Table 31: Production-dual-channel score buckets for strict trend plays. The source marks the actual-ratio cells from $[0.6, 0.7)$ to $[0.3, 0.4)$ as the precision region, and the dual channel has no null bucket.

Score bucket	Total plays	Play share (%)	Strict trend plays	Trend-play share (%)	Actual ratio (%)
1	2075	0.00	2	0.01	0.10
[0.9,1)	130	0.00	0	0.00	0.00
[0.8,0.9)	511	0.00	0	0.00	0.00
[0.7,0.8)	1478	0.00	1	0.01	0.07
[0.6,0.7)	3048	0.00	53	0.32	1.74
[0.5,0.6)	8089	0.01	470	2.86	5.81
[0.4,0.5)	55853	0.06	2248	13.66	4.02
[0.3,0.4)	359111	0.36	3821	23.23	1.06
[0.2,0.3)	1736152	1.72	4101	24.93	0.24
[0.1,0.2)	5047676	5.00	2243	13.63	0.04
(0,0.1)	80347888	79.57	1682	10.22	0.00
0	13409487	13.28	1830	11.12	0.01
Total	100971498	100.00	16451	100.00	0.02

Table 32: Score quantiles for pure negatives (unseen in training) and non-pure negatives (training distribution), as displayed by the source.

Target	Negative type	Mean	p10	p25	p50	p75	p90	p95	p99
Upload attribution	Pure	0.006635	0.000007	0.000007	0.000007	0.001647	0.013782	0.032389	0.132562
Upload attribution	Non-pure	0.004168	0.000005	0.000005	0.000005	0.000314	0.009599	0.018563	0.076268
Second-hop collection	Pure	0.023329	0.000009	0.000405	0.006113	0.021186	0.059791	0.110195	0.27206
Second-hop collection	Non-pure	0.016316	0.000009	0.000239	0.004317	0.01516	0.039628	0.070678	0.200267

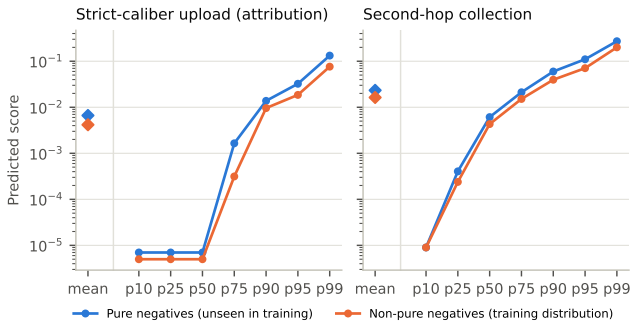


Figure 11: Score quantiles for pure and non-pure negatives, with diamonds denoting means. Redrawn from Table 32.

Table 33: Per-video play counts reported by the source.

Videos	min	max	mean	p10	p25	p50	p75	p90	p95	p97.5	p99	p99.5	p99.9
All videos	0	81270039	21589	0	1	38	476	10554	56811	157985	486876	883207	2433271
Seeding (trend) videos	0	59840252	116800	7	49	2499	46064	216219	493479	913579	1897330	3149660	8388051

H.4 The Deployed Shared/Specific-Expert System

The industrial reference system, SSE, replaced the per-target towers with user-side production, shared, and consumption experts and item-side per-target, shared-production, shared-global, and consumption experts combined per task by a gate and an MLP (MMoE/PLE family; Ma et al., 2018a; Tang et al., 2020). SSE has no residual controller, no held-out selection of granularity, and no reported calibration stage, so it fixes the resolution of sharing by architecture where GSRT selects it from the labels. The two compose rather than compete. Because r_k can be the bilinear score of an SSE consumption head, a platform running such a system can add the selected coefficient on top of it without changing the towers.

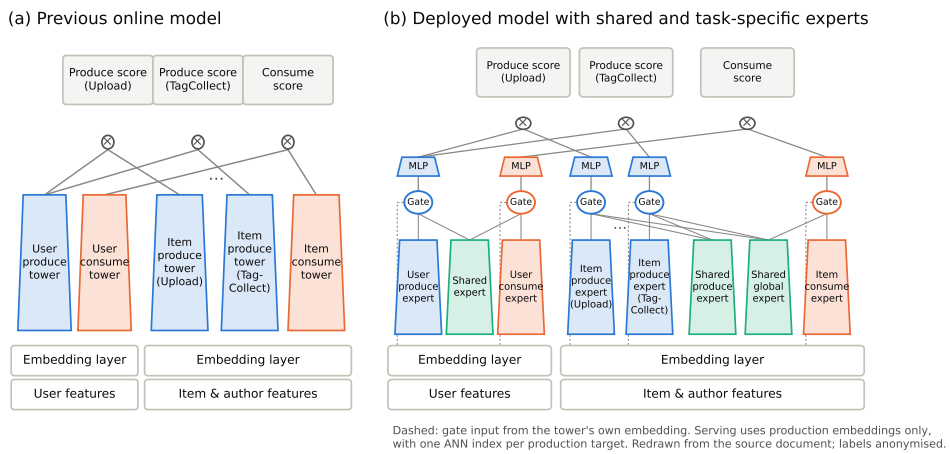


Figure 12: Deployed SSE architecture as described by the source: separate user produce/consume towers and per-target item towers (left) were replaced by user-side production, shared, and consumption experts and item-side per-target, shared-production, shared-global, and consumption experts combined by per-task gates and MLPs (right). Redrawn with anonymised labels.

The user inputs of SSE include profile variables, production statistics, and historical behaviour, and item inputs include item profiles, item statistics, and creator profiles. Training used the four production targets and effective play with BCE and InfoNCE, stratified sampling by the definitions of Appendix H.2, multiple epochs, and repeated-round dropout. Serving retained the production embeddings, built one ANN index per target, and assigned target-specific quotas. Pre-ranking and ranking exposure guarantees were deployed alongside. Table 35 separates the deployed components from the proposed method and names the evidence executed for each.

Table 34 lists the online changes reported for the bundled deployment. All are positive. The components shipped together, so the record reports their combined effect on the industrial system whose constraints GSRT is designed for (Table 35).

Table 34: Online changes reported for the bundled SSE deployment, whose components shipped together. The source states the figures at the precision shown.

Reported metric	Relative change
APP duration	$+0.089\%$
Per-capita APP usage duration	$+0.089\%$
Total consumption VV	$+0.074\%$
Social-type core creator DAU	$+0.391\%$

Table 35: Evidence mapping between the deployed reference system and GSRT.

Component	Deployed SSE	GSRT	Executed evidence
Shared and task-specific experts	Yes	Not required	Bundled industrial deployment
Consumption residual controller	Not reported	Core mechanism	Controlled benchmark, public-data study, and production-log ranking study (Sections 4 and 4.5)
Held-out granularity selection	Not reported	Core mechanism	Controlled benchmark, public-data study, and production-log ranking study (Sections 4 and 4.5)
Target-distribution calibration	Not specified	Pipeline stage	Algebraic identity; intercept offset in the study
User-only ANN folding	Production embeddings served	Required	Numerical identity check in every cell
Risk certification	Not reported	Optional	t certificate over the regime map (Table 7); sample requirement of Proposition D.1
Exposure guarantees	Yes	Excluded	Bundled in the industrial deployment; separated by the protocol of Appendix H.5

H.5 Validation Protocol

The source argues that a learned gate can learn a zero weight for consumption whenever consumption is unhelpful, but the benchmark shows that this holds only when the governing variable is among the gate's inputs (Proposition 3.1), and that searching over gates costs accuracy when no heterogeneity exists (Section 4). Validation should therefore score on a frozen, time-separated, non-resampled target distribution. It should compare no transfer, global, group, instance, and fixed shared-expert conditions under identical fitting windows, label budgets, and candidate sets, with paired user-level inference, multiplicity adjustment, and certification splits sized by Proposition D.1. It should evaluate calibration apart from ranking (Corollary C.2), and it should end with a user-randomized test whose primary endpoint is whether a user produces within a prespecified window, with consumption as a guardrail and a design that addresses interference between users through shared exposure. Table 36 lists the questions and the design each requires, and the paragraphs below state the protocol in full.

Target-distribution evaluation. Validation should begin with a frozen, time-separated, non-resampled evaluation distribution. Each retained training observation should carry its user group, production target, hard- or easy-negative stratum, and retention probability, and the served score must be identified as a BCE probability, a transformed score, an InfoNCE similarity, or a combination. The first model comparison should use capacity-matched target-only and consumption-assisted towers across target–population cells, so that the gain can be related to target information and source–target discrepancy as Proposition B.1 predicts.

Conditional transfer and calibration. A controller comparison should include no transfer, global blending, group shrinkage, instance gating, and the deployed fixed expert system under identical fitting windows, label budgets, and candidate sets, reporting proper scoring rules, target-specific retrieval recall, fixed-quota precision, and subgroup results with paired user-level inference and multiplicity adjustment across production targets. Calibration should be evaluated independently of ranking, since neural scorers are miscalibrated even without resampling (Guo et al., 2017). A global retention correction, a per-stratum correction (Proposition C.1), and the label-dependent correction of Corollary C.2 can be compared with the uncorrected score using Brier loss, log loss, calibration curves, and quota decisions. Persistence of the high-score anomaly after correction would direct attention to score semantics, label definitions, temporal shift, item exposure, or memorization. BCE and InfoNCE should receive separate roles, contrastive training for large-corpus retrieval and an independently evaluated BCE head for probabilities (van den Oord et al., 2018; Yi et al., 2019).

Randomized policy evaluation. Online evaluation should randomize users and separate sample processing, transfer architecture, and exposure guarantees, with a prespecified power analysis (minimum

detectable effect on the production endpoint) and a design that addresses interference between users through shared exposure, since a user-level SUTVA assumption is not automatic on a recommender platform. Its primary production endpoint should be whether a user produces within a prespecified window, rather than an exposure-level sum dominated by already active creators, and consumption should be a prespecified guardrail or non-inferiority criterion. A subsequent incremental-effect study should define treatment as a randomized recommendation policy and estimate heterogeneous effects from pretreatment features (Nie & Wager, 2021), comparing propensity ranking with incremental-effect ranking to determine whether the model identifies established creators or users whose behaviour changes under the policy.

Fig. 13 shows the framework behind this protocol, from a supporting research report, in which a randomized recommendation policy, predictive sharing between consumption and production, and long-term consumption and creator-supply objectives are separate questions.

Table 36: Validation questions and the design each requires.

ID	Scientific question	Required design
E0	Are scores evaluated on the target distribution?	Frozen temporal test set, retention probabilities, explicit score semantics
E1	Does transfer gain vary with scarcity and discrepancy?	Capacity-matched target-only and assisted models across target–population cells
E2	Where does local negative transfer occur?	Target-by-group-by-layer gradient-alignment analysis
E3	Does conditional transfer beat fixed sharing?	No-transfer, global, group, instance, and fixed-expert conditions with paired tests
E4	Does correction improve score trust?	Brier, log loss, calibration curves, ranking, and fixed-quota precision
E5	What causes high-score, low-outcome anomalies?	Calibration, score semantics, exposure, and seen-versus-unseen item analysis
E6	Should BCE and InfoNCE have separate roles?	Retrieval-quality and probability-quality ablation
E7	Which component changes online outcomes?	Decomposed user-randomized test with a production endpoint and consumption guardrail
E8	Who is induced to produce?	Randomized exploration and heterogeneous treatment-effect estimation

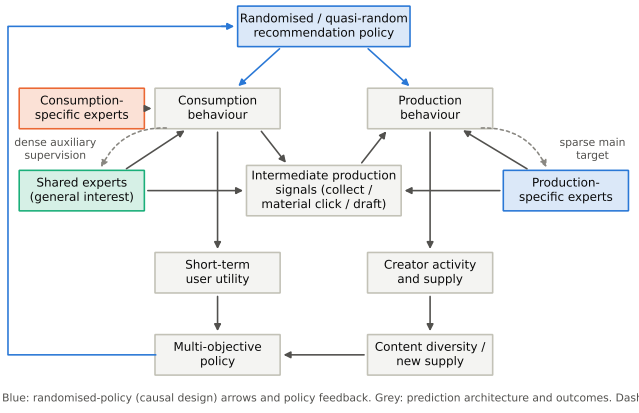


Figure 13: Framework from the supporting research report separating a randomized recommendation policy, predictive sharing between consumption and production, and long-term consumption and creator-supply objectives.

H.6 Evidence Provenance

Table 37: Provenance of every evidence type used in the manuscript.

Evidence	Source	Transformation	Location
Prior online interventions	Industrial source	Redrawn and anonymised	Fig. 8; Table 28
Positive rates and sampler	Industrial source	Direct transcription; duplicated union index corrected	Table 29; Fig. 9
Score buckets	Industrial source (embedded image)	Visual transcription, arithmetic check, redrawn	Fig. 10; Tables 30–31
Negative quantiles, play distribution	Industrial source	Direct transcription	Tables 32–33
SSE architecture and deployment metrics	Industrial source	Redrawn and anonymised; direct transcription	Fig. 12; Table 34
Derivations	This paper	Proofs in Appendix B; numerical check of Theorem 3.4	Section 3
Benchmark results	v4 campaign (data-generating process of Appendix E; regime map, scale sweep, development run)	Seed means, intervals, paired contrasts, Holm adjustment, regret, theory quantities	Section 4; Appendix F
Public-data results	KuaiRand-Pure (Gao et al., 2022), CC BY-SA 4.0	User-disjoint roles, random-exposure test, seed statistics, item-feature ablation	Section 4.6; Appendix G
Ranking at production sparsity	Production logs of the anonymised platform (offline export; four disjoint slices of the training stream)	Placement-selected coefficient, monotone repair, residual-replacement controls, role-seed and slice statistics	Section 4.5; Appendix I
ANN identity and E0 checks	v4 campaign	Direct-versus-folded comparison in every cell	Section 3.4; Appendix E.5
Prediction and causal framework	Supporting research report	Redrawn conceptual framework	Fig. 13
Validation protocol	This paper	Research design; E1 and E4 executed offline on production logs (Appendix I), the user-randomized test not executed	Appendix H.5

Industrial diagnostics and deployment changes are attributed to the anonymised source, benchmark results to the archived campaigns, mathematical identities to derivation, and the validation protocol to research design. The ranking results of Appendix I are offline measurements on the platform's own logs, while the online figures of Appendix H.4 belong to the industrial deployment that motivated GSRT.

I Production Logs at Upload Sparsity: Selecting the Coefficient for Ranking

The benchmark of Section 4 is scored by expected Brier loss. A retrieval stage is judged by what it places inside a fixed quota, at positive rates two orders of magnitude below the benchmark's (Table 29), and a Brier gain need not be a ranking gain. This appendix runs questions E1 and E4 of the validation protocol (Table 36) at production sparsity, on the production logs of the anonymised platform of Appendix H, with its label definitions, its segment variables, and its own exposure frequency. An offline export of the frozen model scores and the labels is what makes the question answerable, because the coefficient can then be selected and scored on held-out days without retraining or redeploying the production model. True probabilities are unknown here, as in any log, so every loss is measured on observed labels and every comparison is against the frozen base under the same monotone repair. The study is repeated on four disjoint slices of the training stream, each with its own frozen models, so that a finding is not one fit of one base model. What it establishes is a property

of the selection procedure at production sparsity, evaluated on held-out days under the platform's own exposure policy. The user-randomized design that turns such a comparison into a policy claim is specified in Appendix H.5.

Data. The export covers 2 million users in 8 author life-cycle levels whose strict-upload rates span three orders of magnitude, with an overall rate of 7.0×10^{-5}. The target is the platform's strict-caliber upload label of Table 29 and the source is the consumption stream of the same rows. The frozen base is the linear two-tower model of the public-data study (Appendix G) with the segment among its user inputs, fitted by case-control sampling on nested prefixes of one training stream, so that the scarcity of the base model is an axis of the design (E1), and the frozen consumption model reaches AUC 0.730 on consumption. Controllers are fitted on a seven-day window and every method is scored on the next seven days, on users disjoint from all fitting roles, with the Brier loss computed on the observed labels of those days. Splits, roles, budgets and the residual are those of the main text.

Squared loss is dominated by calibration. A base fitted on a few dozen positives is overdispersed, and at this positive rate the squared loss is governed by the level and spread of probabilities near zero, not by their order. With an intercept as the only calibration device, as in the main text, squared loss selects a negative coefficient in 59 of the 60 scarce-base cells of the 4 slices (mean -0.41; Table 42), because on a score correlated with the base a negative coefficient shrinks the logits. When every coefficient on the grid is first given a two-parameter monotone repair (Platt, 1999), the squared loss is nearly flat in the coefficient (the whole grid spans 0.11 per million of a loss near 65) and its choice is unstable in sign (positive in 39 cells, mean 0.38), while placement values select a positive coefficient in 60 (mean 0.52). The coefficient has to be released from calibration before the source signal becomes selectable, and then selected by a loss that sees order.

Ordering and calibration can be chosen separately. For a user-side coefficient the served probability $\sigma(b+\lambda r+c)$ is increasing in $z = b+\lambda r$, so a monotone map of z leaves the ordering unchanged, and with it AUC, precision and recall at every quota, and the set retrieved under Proposition 3.5. We therefore choose λ for ranking and repair the probabilities afterwards. The coefficient maximizes the cross-fitted AUC, computed from the placement values $v_i = \Pr(z_i > z_{\text{neg}})$ of the positives against a fixed sample of negatives (DeLong et al., 1988). The evidence gate is the paired one-sided t test of Section 3.2 applied to the out-of-fold placement gains of the positives at $\alpha = 0.05$, group coefficients are shrunk toward the global one by their positive counts, and a binned isotonic map (Zadrozny & Elkan, 2002) fitted on the calibration users turns z into a probability. Retrieval needs z alone, so serving is unchanged. Every comparison below is against the base given the same monotone repair, so that the repair's own gain is not credited to transfer.

Results. At the scarce base placement selects a coefficient of 0.54 where squared loss selects -0.40 (Table 38). AUC rises by 0.0224 [0.0103, 0.0345] and recall inside the 1% quota, the share of test positives among the top 1% of test rows across all users, by 0.062 [0.043, 0.081] from 0.163, a relative gain of 38%, while the Brier gain on held-out labels over the repaired base is 0.008 [0.007, 0.010] per million, so the ranking gain is not paid for in calibration. The choice of repair is what buys that last property, and it is a free choice because the ranking columns do not depend on it. A binned isotonic map keeps the Brier gain positive, whereas a two-parameter repair of the same coefficients gives -0.21 [-0.36, -0.07] per million, and both are reported. Over the 9 configurations of the upload target (two grouping variables and the removed-group control, at three controller budgets) the AUC interval lies above zero in 9 and the recall interval in 9. With more training rows for the base, the coefficient, the share of seeds in which transfer is served, and every gain go to zero. The source is worth borrowing where the target labels could not pay for the target model, which is the criterion of Section 3.3 read on production data.

Controls. Table 39 repeats the residual-replacement controls of Section 4 under placement selection. With the consumption score the gate fires in 9/15 cells of the life-cycle grouping and AUC rises by 0.0200 [0.0080, 0.0320]. With a consumption model refitted on permuted labels, with a random direction of the same mean and norm, and with the base model's own score, it fires in 0/15, 0/15 and 0/15, and no control gains. The paired differences in ΔAUC between the consumption score and the three controls are 0.0202 [0.0091, 0.0313], 0.0241 [0.0156, 0.0325] and 0.0200 [0.0080, 0.0320]. The capacity control is zero by construction. A direction that is a monotone function of the base score

inside a group leaves every within-group placement value unchanged, so all coefficients tie and the served ordering is the base model's. Under squared loss the same controls gain (Section 4), because any added direction lets the controller recalibrate, whereas an ordering offers no such slack. The gain is a gain inside levels. The pair-weighted AUC computed within life-cycle levels rises by 0.0408 [0.0280, 0.0535] under group coefficients and by 0.0318 [0.0199, 0.0437] under one coefficient, a paired difference of 0.0089 [0.0063, 0.0115], and with the group replaced by a random label the difference is -0.0003 [-0.0010, 0.0003]. Group resolution therefore pays only where the grouping carries the relation.

Independent slices of the stream. A base model fitted on a few dozen positives is itself a random object, so the study is repeated on 4 disjoint slices of the training stream, on each of which the users, the training rows and every frozen model are drawn afresh, and the intervals below therefore cover that variability rather than only the draw of user roles (Table 41). The squared-loss and the placement coefficients have opposite signs on 4 of the 4 slices, and the interval over role seeds lies above zero on 4 slices for AUC and on 4 for recall. Counting each slice once, AUC rises by 0.023 [0.009, 0.038], the AUC within levels by 0.051 [0.026, 0.075] and recall inside the quota by 0.062 [0.037, 0.088], against -0.004 [-0.007, -0.002] in AUC for the random direction. The finding is therefore a property of the procedure at this sparsity and not of one fit of one base model.

Controller window and group resolution. Lengthening the controller window from 12M to 48M rows, with the same frozen models, raises the controller positives from 359 to 1,424 (Table 40). The gains keep their size and their intervals narrow. The served group coefficients spread. They range from 0.39 to 0.94 under the short window, where shrinkage holds the sparse levels near the global coefficient, and from 0.07 to 0.98 under the long one, where the newest authors receive almost no transfer and the established ones nearly all of it. Finer resolution is again bought with more positives.

The gate and the quota. The gate buys protection in proportion to how weak the relation is. Over the scarce-base cells of all 4 slices, the placement rule served without any gate lowers recall in 0 and AUC in 1 of the 120 upload cells, of which the gate serves 58. On uploads, where the relation to consumption is strong, there is little to protect, and the gate withholds the cells whose evidence is thin. On the camera target, whose relation to consumption is weaker, the ungated rule lowers recall in 8 and AUC in 44 of 120 cells (worst -0.038 and -0.022), and the gate serves 18, cutting the AUC losses by more than four fifths. With the base trained on 100M rows, where the criterion of Section 3.3 calls for no transfer, the ungated rule loses recall in 23 of 90 cells of the first slice (worst -0.033) and the gate serves 2.

Which quantity the gate should test is a design choice the same data settles. Selecting the coefficient on recall inside the quota, the quantity a deployment is judged by, selects 0.90 and gains 0.028 [-0.002, 0.057], against 0.062 [0.043, 0.081] for placement values. A 1% quota is cleared by a few dozen controller positives, so its maximizer follows noise, and a gate on it serves only 16 of the camera cells. Placement values, which use every positive–negative pair, are the statistic with the evidence behind them, and recall inside the quota is the quantity to demand evidence for rather than the one to maximize. Where the two part, as on the camera target, whose ungated AUC moves by 0.000 [-0.007, 0.008] while its quota recall moves by 0.031 [0.024, 0.037], a deployment can run the gate at the level its own downside tolerance implies, and the campaign reports $\alpha = 0.05$ throughout.

What this establishes. At production sparsity the squared loss is dominated by calibration and spends the coefficient on shrinkage. Separating ordering from calibration, which the folding identity of Proposition 3.5 permits, lets a placement-selected coefficient raise ranking at no cost in calibration, and question E4 should therefore report ranking next to the Brier loss. The finding holds on four disjoint slices of the production training stream, tracks the budget of Section 3.3 as the base model is given more rows, and survives every residual-replacement control, none of which passes the gate in any cell. It is an offline comparison on logs collected under the incumbent exposure policy, which is why it can be answered without redeploying the production model, and Appendix H.5 specifies the user-randomized run that turns it into a policy claim.

Table 38: Production logs of the anonymised platform, strict-upload target, author life cycle as the group, full controller budget, five seeds (mean [95% Student-t interval]). The base model is trained on nested prefixes of the production training stream. *Squared loss*: the global-blend candidate of the main text, one coefficient minimizing the squared loss on the controller users. *Placement*: group coefficients chosen by cross-fitted placement values, with a monotone repair of the probabilities. *Served* counts the seeds in which the evidence gate on the placement gains admits the candidate at $\alpha = 0.05$. Differences are those of the placement-selected candidate against the frozen base, and the last column is the Brier gain on held-out labels per million over the base given the same repair.

Base trained on	Positives	Base AUC	Squared loss		Placement				
			λ	ΔAUC	$\bar{\lambda}$	Served	ΔAUC	Δrecall@1%	Brier vs. repaired base
1M rows	68	0.823	-0.40	-0.039 [-0.058, -0.021]	0.54	5/5	0.022 [0.010, 0.034]	0.062 [0.043, 0.081]	0.008 [0.007, 0.010]
3M rows	215	0.912	-0.11	-0.002 [-0.005, 0.001]	0.29	2/5	0.001 [-0.002, 0.005]	0.064 [0.026, 0.103]	0.007 [0.005, 0.009]
10M rows	728	0.927	-0.35	-0.005 [-0.008, -0.002]	0.06	1/5	0.002 [0.000, 0.004]	0.021 [0.015, 0.027]	0.004 [0.002, 0.005]
100M rows	7,109	0.931	-0.10	-0.001 [-0.001, 0.000]	-0.01	0/5	0.001 [-0.001, 0.003]	0.006 [-0.008, 0.021]	0.001 [-0.001, 0.003]

Table 39: The residual-replacement controls of Section 4 under placement selection (production logs, base trained on 1M rows, strict-upload target, author life cycle as the group). A seed's value is the mean over the three controller budgets, and entries are means over five seeds [95% Student-t interval]. *Served* counts the cells in which the placement gate fired. *Within level* is the pair-weighted AUC computed inside life-cycle levels, which moving whole levels past each other cannot raise. The last column is the paired difference in ΔAUC between the consumption score and the control.

Residual direction	Served	ΔAUC	ΔAUC within level	Δrecall@1%	Consumption $-$ control
Consumption score	9/15	0.0200 [0.0080, 0.0320]	0.0408 [0.0280, 0.0535]	0.057 [0.040, 0.073]	–
Refitted on permuted labels	0/15	-0.0002 [-0.0012, 0.0008]	0.0001 [-0.0013, 0.0014]	-0.005 [-0.010, 0.000]	0.0202 [0.0091, 0.0313]
Random direction, same norm	0/15	-0.0041 [-0.0118, 0.0036]	-0.0043 [-0.0120, 0.0035]	-0.012 [-0.023, -0.001]	0.0241 [0.0156, 0.0325]
Base model's own score	0/15	0.0000 [0.0000, 0.0000]	0.0000 [0.0000, 0.0000]	0.000 [0.000, 0.000]	0.0200 [0.0080, 0.0320]

Table 40: Controller window and group resolution (production logs, base trained on 1M rows, strict-upload target, author life cycle, full budget, five seeds). The two exports share the frozen models and differ only in the length of the controller window. *Positives* are the upload events the controller sees, and $\lambda_{\min}$, $\lambda_{\max}$ the smallest and largest of the eight served group coefficients, all averaged over seeds.

Window	Positives	$\lambda_{\min}$	$\lambda_{\max}$	ΔAUC	Δrecall@1%
12M rows	359	0.39	0.94	0.0224 [0.0103, 0.0345]	0.062 [0.043, 0.081]
48M rows	1,424	0.07	0.98	0.0202 [0.0118, 0.0286]	0.072 [0.056, 0.087]

Table 41: Disjoint slices of the production training stream (base trained on 1M rows, strict-upload target, author life cycle as the group). Each slice supplies its own users, items, training rows and frozen models. Rows: a seed's value is the mean over the three controller budgets, entries are means over five role seeds [95% Student-t interval]. *All*: each slice counts once (mean and Student-t interval over slices). *Random* is ΔAUC with the consumption score replaced by a random direction of the same mean and norm. *Brier*: Brier gain on held-out labels over the base given the same binned repair, per million.

Slice	Positives	Base AUC	λ sq. loss	$\bar{\lambda}$ placement	Served	ΔAUC	ΔAUC within level	Δrecall@1%	Brier	Random
7	68	0.823	-0.40	0.52	9/15	0.020 [0.008, 0.032]	0.041 [0.028, 0.054]	0.057 [0.040, 0.073]	0.008 [0.008, 0.009]	-0.004 [-0.012, 0.004]
8	79	0.829	-0.41	0.69	9/15	0.037 [0.024, 0.050]	0.073 [0.054, 0.092]	0.070 [0.059, 0.081]	0.006 [0.004, 0.008]	-0.006 [-0.010, -0.002]
9	67	0.806	-0.56	0.47	5/15	0.018 [0.010, 0.026]	0.044 [0.027, 0.061]	0.043 [0.030, 0.055]	0.006 [0.003, 0.009]	-0.005 [-0.014, 0.004]
10	80	0.822	-0.29	0.38	5/15	0.019 [0.011, 0.027]	0.044 [0.033, 0.056]	0.080 [0.050, 0.110]	0.008 [0.007, 0.009]	-0.002 [-0.004, -0.001]
All			-0.41	0.52		0.023 [0.009, 0.038]	0.051 [0.026, 0.075]	0.062 [0.037, 0.088]	0.007 [0.005, 0.009]	-0.004 [-0.007, -0.002]

Table 42: What each selection loss does with the coefficient (production logs, base trained on 1M rows, strict-upload target, author life cycle, three controller budgets, five role seeds per slice). Mean selected global coefficient and the number of cells with the stated sign. *Squared loss*: an intercept is the only calibration device, as in the main text. *After repair*: squared loss of the probabilities after a two-parameter monotone repair fitted for every coefficient on the grid.

Slice	Cells	Squared loss		After repair		Placement	
		λ	negative	λ	positive	λ	positive
7	15	-0.40	15/15	0.20	8/15	0.52	15/15
8	15	-0.41	15/15	0.52	13/15	0.69	15/15
9	15	-0.56	14/15	0.28	8/15	0.47	15/15
10	15	-0.29	15/15	0.50	10/15	0.38	15/15