THE COST OF SAFE SKIPPING: ADAPTIVE ROUTING WITH PAID FEEDBACK

Anonymous authors
Paper under double-blind review

ABSTRACT

Retrieval across providers with different search costs must decide both where to search and when to stop. Public routing and feedback-based stopping already address parts of this problem; we ask what reordering from paid returns adds once both are available. The key distinction is between finding the complete top-k and deciding to stop, since routing can affect these costs differently. We study this interaction with a neural maximum predictor, residual updates from paid returns, cost-sensitive acquisition and a learned stopping head. On 8.84 million MS MARCO passages, a frozen rank-eight policy reduces mean work by 6.59% relative to fixed order with the same predictor and stopping head. Its independent finite-pool omission-risk upper bound is 3.84% at a 5% target. Replay of 8,192 test queries locates the saving after result completion. Matched acquisition and threshold controls connect the gain to cost prioritization and stopping calibration, with one-shot cost ordering accounting for much of the reduction. On source-pure providers, earlier completion is offset by later searches. Warm neural execution retains the semantic throughput gain, while parallel static search has lower single-query latency.

1 INTRODUCTION

Retrieval from partitioned collections must decide which providers to search and when to stop. Different provider sizes and costs make search count an incomplete measure of efficiency. Stopping too early can omit a global top-k document, so the policy must control both charged work and omissions.

Public information already supplies useful routing signals: RAGRoute predicts source relevance (Guerraoui et al., 2025), Optimist uses score moments (Bruch et al., 2025), and AMIPS learns support functions (Olausson et al., 2026). ConANN uses intermediate results to adapt search effort (Horchidan et al., 2025). Correlated costly-search problems also distinguish fixed-order adaptive stopping from fully adaptive acquisition (Chawla et al., 2020; 2023).

The question is *what adaptive provider order adds when the predictor is shared and stopping already uses paid returns*. A lower work count alone cannot answer it: one-shot cost prioritization can change the route without repeated updates, while the stopping threshold determines how much of that route is executed.

Result completion does not itself tell the policy to stop. An unopened provider may still appear competitive, so finding the answer earlier need not lower total work. We separate work before and after completion on successful queries, retaining failed-query work separately.

We study this interaction with a neural maximum predictor whose residual state updates after each paid return. Dynamic acquisition ranks remaining providers by competitiveness per cost, while a learned head decides when to stop. Matched actors share their predictor, stopping head and work constraints; an independent finite-pool test evaluates each complete frozen policy.

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.

Contributions.

1. A cost-aware paid-feedback controller with fixed/dynamic comparisons that share the predictor, stopping head and work constraints.
2. Same-policy replay that separates work before and after completion, with acquisition, threshold and training controls to test explanations for the saving.
3. Independent confirmation on 8.84M passages, with a 6.59% mean-work saving for the frozen rank-eight policy, and warm execution measurements that distinguish throughput from single-query latency.

2 RETRIEVAL WITH PAID OBSERVATIONS

Let $D_1, \dots, D_M$ be disjoint providers, $T_k(q)$ the canonical global top-k, and $Y_k(q)$ its provider set. Each indexed item (passage or full document) is termed a document. Scores are inner products of normalized vectors; ties use ascending global document ID. A full search of provider i costs $c_i > 0$ and returns its exact local top-k. For searched set $F_\pi(q)$ and returned set $R_\pi(q)$, define

$$W_\pi(q) = \sum_{i \in F_\pi(q)} c_i, \qquad L_\pi(q) = \mathbf{1}\{T_k(q) \not\subseteq R_\pi(q)\}. \tag{1}$$

All paid results are retained. Because every global top-k document belongs to its provider's local top-k, exact local search and canonical merging give $L_\pi = \mathbf{1}\{Y_k \not\subseteq F_\pi\}$. Primary experiments use $k = 10$ and $c_i = |D_i|$; a separate diagnostic varies virtual prices. Document work measures the charged search volume, not monetary cost, communication or total FLOPs.

Before search, the policy knows the query, public source/centroid scores, prices and trained parameters. Its transcript H_t adds only already-paid returns. We seek low expected work with omission risk at most $\alpha = 0.05$ and the pathwise limit

$$W_\pi(q) \le B(q) = \min\{B_0(q), C\}, \qquad B_0(q) = \sum_{j=1}^{r_0} c_{\pi_c(q,j)}. \tag{2}$$

Here π_c is the public centroid order; TUNE sets its reference rank r_0 and the common cap C. Exhausting the budget returns the current result, with an omission recorded if incomplete. No exhaustive fallback bypasses the charge. Statistical acceptance concerns the finite query pool specified below.

3 ROUTING AND STOPPING WITH PAID FEEDBACK

3.1 PUBLIC PREDICTION AND RESIDUAL CONDITIONING

The maximum predictor estimates $v_i(q) = \max_{d \in D_i} s(q, d)$ before search. We use the authors' ICNN-based SupportNet architecture (Olausson et al., 2026). It fits three quarters of TRAIN; the remaining quarter estimates residual mean b and covariance S. Both fixed and dynamic policies initialize $\mu_0 = g_\theta(q) + b$. The full residual model is

$$\Sigma_0 = 0.8S + 0.2\,\mathrm{diag}(S) + 10^{-8} I. \tag{3}$$

For rank r, the alternative $\Sigma_r = F_r F_r^\top + \mathrm{diag}(\mathrm{diag}(S) - \mathrm{diag}(F_r F_r^\top)) + 10^{-8} I$ uses the leading eigenvectors of S scaled by $\sqrt{0.8\lambda}$, preserving all marginal variances. Full covariance was the initial specification; rank eight was recorded as a follow-up hypothesis before confirmation queries were selected. Both enter the confirmation family.

After paying for provider i, let $a = \Sigma_{:,i}$, $u = \max(\Sigma_{ii}, 10^{-10})$ and $r = \mathrm{clip}(v_i - \mu_i, -4\sqrt{u}, 4\sqrt{u})$. Update

$$\mu \leftarrow \mu + ar/u, \qquad \Sigma \leftarrow \Sigma - aa^\top/u. \tag{4}$$

The observed mean is set to v_i and its covariance row/column to zero. A 0.002 standard-deviation floor applies only to acquisition and stop features. Clipping makes this a heuristic residual update rather than an exact Gaussian posterior.

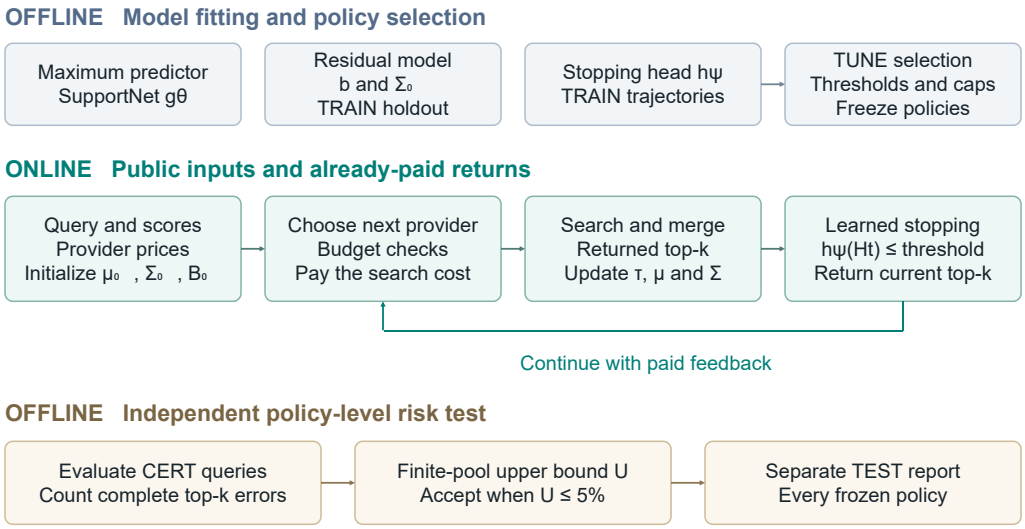

Figure 1: Offline fitting supplies the predictor, residual model and stopping head. Online decisions use only public information and paid returns. Independent testing assesses the complete frozen policy, separately from its acquisition and stop scores.

Algorithm 1 Executable paid-feedback policy (frozen parameters)

Require: Query q, prices c, public scores s, $B_0(q)$, C, $g_\theta, b, \Sigma_0, h_\psi, \theta_\pi$

$V \leftarrow \varnothing$, $R \leftarrow \varnothing$, $W \leftarrow 0$, $\mu \leftarrow g_\theta(q) + b$, $\Sigma \leftarrow \Sigma_0$
while an unvisited provider remains **do**
  $\mathcal{A} \leftarrow \{i \notin V : c_i \le B_0 - W\}$; **break** if $\mathcal{A} = \varnothing$
  if $V = \varnothing$ **then**
    $i \leftarrow \arg\max_{i \in \mathcal{A}} s_i$ ▷ common public seed
  else if fixed actor **then**
    $i \leftarrow$ first unvisited provider in the frozen support order
  else
    $i \leftarrow \arg\max_{i \in \mathcal{A}} A_{i,t}$
  end if
  if $i \notin \mathcal{A}$ or $W + c_i > C$ **then**
    break ▷ no cheaper substitution
  end if
  Pay c_i; obtain local top-k; set $W \leftarrow W + c_i$, $V \leftarrow V \cup \{i\}$
  Canonically merge into R; update τ_t, μ, Σ using the paid maximum
  if $h_\psi(H_t) \le \theta_\pi$ **then**
    break
  end if
end while
return R, including any incomplete result at budget exhaustion

3.2 Acquisition and Adaptive Stopping

After a common public-score seed, let τ_t be the returned kth score. The dynamic actor ranks unvisited providers by

$$A_{i,t} = c_i^{-1} \Phi(\mathrm{clip}((\mu_{i,t} - \tau_t)/\sigma_{i,t}, -12, 12)) \,. \tag{5}$$

It selects the highest priority affordable under $B_0 - W_t$, then checks C. If the selected action exceeds C, it stops without substituting a cheaper one. The fixed actor follows the strict initial $g_\theta + b$ order after the same seed, stopping at an unaffordable next provider. Algorithm 1 specifies both admission rules.

A histogram-boosted classifier h_ψ estimates whether the current result is incomplete from 31 public, state and return features (Appendix D). Each residual model receives a head fitted on TRAIN trajectories and shared by its fixed/dynamic pair. For each actor, TUNE selects a threshold minimizing mean work plus 0.25 P99 under a 3.5% empirical omission guard. If the guard is infeasible, the lowest-risk point is retained and marked. Because score calibration depends on the visited states, each actor's threshold is part of its policy even when the head is shared.

Equation 5 scores direct competitiveness rather than expected information gain. It can therefore overlook providers whose returns are informative but absent from the final answer; Appendix K gives an explicit construction. The controller is a tractable search heuristic, without an optimal-exploration guarantee.

3.3 Independent testing of the complete policy

Models, thresholds, budgets and numerical semantics are frozen before CERT labels are opened. CERT draws $n = 8{,}192$ queries uniformly without replacement from an eligible population of size N_{pool}. For e complete-policy errors, compute

$$K^+ = \max\{K : \Pr[\text{Hypergeom}(N_{\text{pool}}, K, n) \le e] > \delta/J\}, \qquad U = \frac{K^+ - e}{N_{\text{pool}} - n}. \tag{6}$$

The maximization ranges over feasible counts $e, \ldots, N_{\text{pool}} - n + e$. Exact inversion and a union bound provide simultaneous coverage at least $1 - \delta$ across $J = 10$ frozen policy–condition slots; acceptance requires $U \le \alpha$. We set $\delta = 0.05$. This follows the learn-then-test perspective (Angelopoulos et al., 2022), with a hypergeometric test for our without-replacement sampling design. The bound covers the eligible pool remaining after CERT, including TEST; it provides neither per-query correctness nor protection against distribution shift. The work cap is a separate pathwise property: only actions satisfying both budget checks are charged.

4 Experiments

4.1 Data, controls and evaluation protocol

The semantic condition assigns all 8,841,823 MS MARCO passages (Bajaj et al., 2016) to 100 frozen semantic providers. The source condition retains 992,502 documents from the 100 largest accepted hostnames, one source per provider. Both use 384-dimensional `all-MiniLM-L6-v2` sentence embeddings based on MiniLM (Wang et al., 2020). They differ in corpus composition as well as ownership. FiQA from BEIR (Thakur et al., 2021) supplies a development diagnostic on 57,638 documents with 1,024-dimensional BGE-M3 embeddings (Chen et al., 2024). Appendix B records pinned model cards, checkpoints and partition provenance.

Each MS MARCO condition has 4,096 TRAIN, 1,024 TUNE and 2,048 development EVAL queries; FiQA has 2,048/512/1,024. The semantic SupportNet has 33.97M parameters and eight layers, using the authors' score/gradient-target losses. Training uses 100,000 steps, batch 32 and four normalized augmentations per prior-fit query. TUNE maximum-score RMSE selects among 30k/60k/100k checkpoints. This matches the XS parameter fraction but uses 3.2M sample presentations, compared with 1B in the authors' routing example (Olausson et al., 2026). The comparison measures adaptation at a shared training budget, rather than AMIPS's fully trained performance. All actors use the same RAGRoute-style public scorer, a matched reimplementation whose earlier training queries are included in confirmation exclusions.

Independent confirmation uses four disjoint new query roles: CERT and TEST for each MS MARCO condition, each with 8,192 queries. We exclude recorded IDs, source rows and normalized-vector fingerprints from both development roots before uniform sampling. The remaining 174,115 representatives form populations of 87,057 and 87,058. Full and rank-eight fixed/dynamic pairs, plus a same-predictor static prefix, make ten frozen slots; TEST selects none of them. Acquisition, training and price interventions use the existing TRAIN/TUNE/EVAL roles and carry no new independent certificates. TEST replay and threshold exchange diagnose already-confirmed results. Appendix A maps these distinct roles to models and code.

Table 1: Neural confirmation: all ten frozen slots. Work is in M documents on TEST. Both CERT and TEST contain 8,192 queries per condition. The finite pool sizes are $N_{\text{pool}} = 87{,}057$ (semantic) and 87,058 (source). e is the CERT error count; K^+ is the upper population error count; $U = (K^+ - e)/(N_{\text{pool}} - 8{,}192)$. All slots satisfy $U \leq 5\%$ with a joint failure budget of 0.05.

Condition	Policy	Mean	P99	e	K^+	U (%)	TEST errors
Semantic	Static prefix	3.571	3.621	254	3136	3.65	287
	Full / fixed	2.661	3.758	237	2942	3.43	274
	Full / dynamic	2.821	3.758	233	2896	3.38	247
	Rank-8 / fixed	2.465	3.758	251	3102	3.62	289
	Rank-8 / dynamic	2.303	3.757	268	3295	3.84	291
Source	Static prefix	0.816	0.820	286	3500	4.08	214
	Full / fixed	0.776	0.866	285	3488	4.06	234
	Full / dynamic	0.765	0.866	276	3386	3.94	267
	Rank-8 / fixed	0.777	0.866	294	3591	4.18	243
	Rank-8 / dynamic	0.777	0.866	250	3090	3.60	234

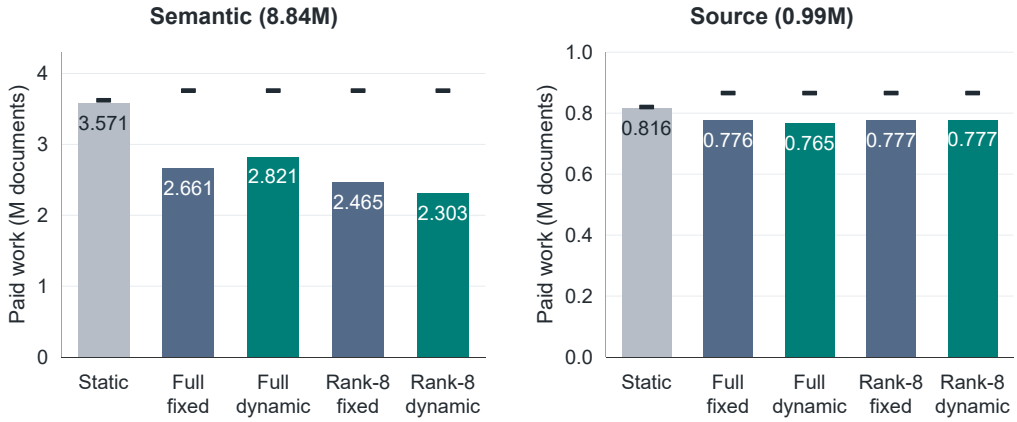

Figure 2: All frozen neural policies on 8,192 TEST queries per condition. Bars show mean paid work; black ticks show P99. Rank-eight dynamic order reduces semantic mean work, with essentially unchanged common-cap tails, but has no mean advantage on source providers.

Canonical targets use source float32 vectors normalized and scored in float64, with global-ID tie breaking. Policy evaluation reveals cached exact responses only after charging. The live benchmark instead recomputes public predictions and searches provider indexes, checking every timed work/error outcome and canonical return against the frozen policy.

4.2 Independent policy confirmation

All ten slots pass their finite-pool risk tests (Table 1). On semantic providers, the rank-eight dynamic policy uses 2.303M mean work versus 2.465M for the fixed policy: a 6.59% reduction with a descriptive paired query-bootstrap interval of [6.27%,6.88%]. TEST errors are 291 versus 289; independent CERT upper bounds are 3.84% versus 3.62%. The saving relative to the same predictor's static prefix is 35.52%. The bootstrap interval measures work variation across queries, separately from risk certification and training randomness.

The outcome depends on the residual model and provider condition. With full covariance, the dynamic policy uses 6.00% more work than the fixed policy, 2.821M versus 2.661M, while reducing omissions. On source providers, full covariance saves 1.31%, whereas rank eight is effectively tied: dynamic spends 23 additional document units per query. Static prefixes have lower P99 in both conditions. Table 1 reports every frozen slot under its own model identity.

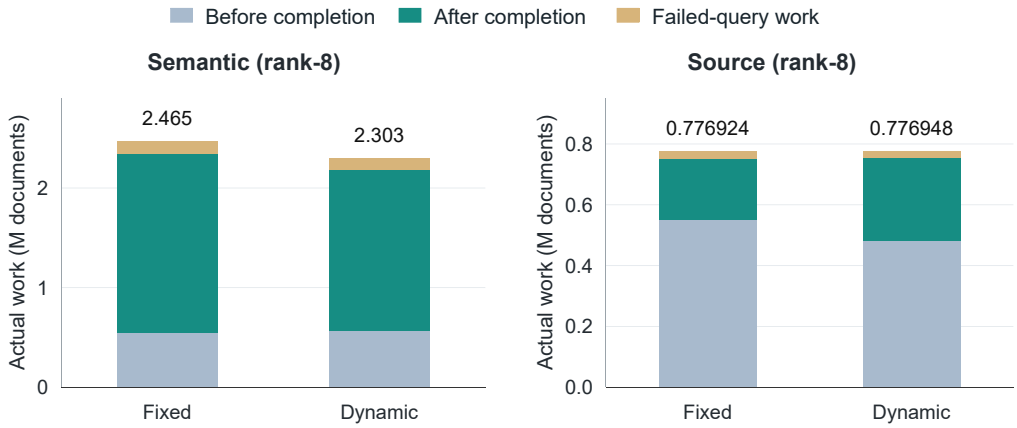

Figure 3: Replay of the same frozen rank-eight policies and TEST queries as Table 1. The components sum exactly to actual mean work. Completion is identified in hindsight; later work may still be needed for an online stopping decision.

4.3 Work before and after result completion

Instrumented replay reproduces all frozen work, error, search-count and forced-cap outcomes. For successful queries, we split actual cost at the first complete global top-10; all failed-query work forms the third component. Unexecuted continuations contribute no actual cost.

On semantic rank eight, dynamic order spends 16,585 more document units before completion, 171,291 fewer afterwards, and 7,637 fewer on failures, with each component averaged over all 8,192 queries. These differences sum to the 162,342 saving. Among the 7,819 queries where both methods succeed, fixed-minus-dynamic differences are $-13{,}722$ before completion and $+180{,}506$ afterwards. Thus the saving's location persists when the success set is held fixed. The corresponding all-query after-completion share is 105.51% ([101.85%,109.20%], paired bootstrap): the signed components offset one another.

Source rank eight shows the opposite balance. Among 7,857 common successes, dynamic saves 75,235 before completion but spends 75,266 more afterwards. Its near-zero total makes a percentage decomposition uninformative. Failures also differ: semantic fixed/dynamic have 264/238 path failures and 25/53 premature stops; source has 101/91 and 142/143. None is forced solely by the cost of necessary providers. Appendix F reports all success strata, numerical safeguards and paid-state traces.

4.4 Ordering and stopping act together

Two acquisition controls isolate the information used to choose the next provider. After the common seed, *one-shot cost* computes Equation 5 once using initial corrected means/scales and the seed's boundary. *Boundary only* updates τ_t while keeping those action means/scales fixed. Neither conditions its action prior on the seed. Both choose among B_0-affordable providers, as dynamic does; fixed support instead follows a strict prefix. All four retain full paid conditioning in their stop features. We compare them with the frozen paired head and with a common head refitted on all four actors' TRAIN trajectories.

At the 3.5% TUNE guard with the frozen head, one-shot cost uses 2.276M work and makes 65/2,048 EVAL errors; full feedback uses 2.273M with 50 errors. With the four-actor head, one-shot is cheaper (2.298M versus 2.383M), but omits more queries (65 versus 48). At the declared 4.5% guard, full feedback improves both metrics: 1.997M/67 versus 2.018M/79. Across the complete sweep in Figure 4, cost prioritization accounts for much of the work reduction over fixed support order. Repeated feedback changes the omission/work tradeoff, but does not dominate at every operating point. Appendix G includes both secondary conditions and infeasible FiQA guards.

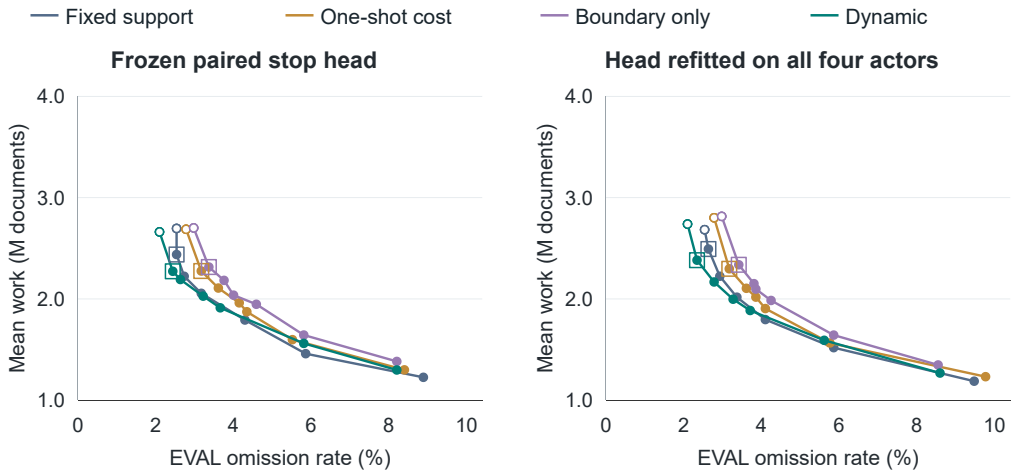

Figure 4: Semantic development EVAL across all ten declared TUNE guards (1–10%). Open circles mark infeasible guards; outlined squares mark 3.5%. Each threshold is selected on TUNE. The curves describe empirical tradeoffs, not certified frontiers.

Table 2: Threshold exchange on unchanged final rank-8 legal paths. Cells show mean work (M documents) and TEST omissions. The original frozen policies are the fixed-path/fixed-threshold and dynamic-path/dynamic-threshold diagonal cells. Swapped cells are post-confirmation diagnostics without new certificates.

Condition	Path	Work at θ_F	Errors	Work at θ_D	Errors
Semantic	Fixed	2.464975	289	2.256568	308
Semantic	Dynamic	2.517896	268	2.302632	291
Source	Fixed	0.776924	243	0.786234	201
Source	Dynamic	0.765206	281	0.776948	234

Exchanging the two frozen thresholds tests how that operating point interacts with the route. On semantic rank-eight TEST paths, θ_F gives 2.465M/289 errors for fixed and 2.518M/268 for dynamic; θ_D gives 2.257M/308 and 2.303M/291. At either common numerical threshold, dynamic spends more but misses fewer queries. The confirmed 6.59% gain therefore belongs to the route *together with its tuned stopping point*. Swapped cells are counterfactual policies without certificates, and equal numerical thresholds need not imply equal risk. Source paths show the complementary effect: the more conservative dynamic threshold nearly cancels a same-threshold work advantage.

4.5 Sensitivity to training, prediction error and prices

We cross full/rank-eight residuals with four stop-training regimes: all 4,096 TRAIN queries; 1,024 mixed queries preserving the 3:1 prior-fit/holdout ratio; 1,024 prior-fit queries; and 1,024 prior-unfitted queries. Each comparison fixes the corrected prior, marginal variances, head capacity, cap and tuning rule, then refits one head on both actors' trajectories. Full-covariance dynamic increases work and fails the 3.5% TUNE guard in all four anchors; rank eight saves 0.99–15.46% on development EVAL. Two bootstrap TRAIN draws per quarter-sized regime add variability, including one positive full-covariance result (Appendix G). Thus covariance rank and the distribution used to train stopping must be assessed jointly.

Two complete neural initialization/minibatch repeats yield rank-eight savings of 8.94% and 5.80%, compared with 6.79% for the confirmed model on the same development EVAL. Their fixed/dynamic error counts are 49/54 and 53/58. Checkpoints and thresholds are selected on TUNE in every run. These repeats test training randomness with fixed queries, partitions, encoder and augmented targets; they do not replace the independently confirmed model.

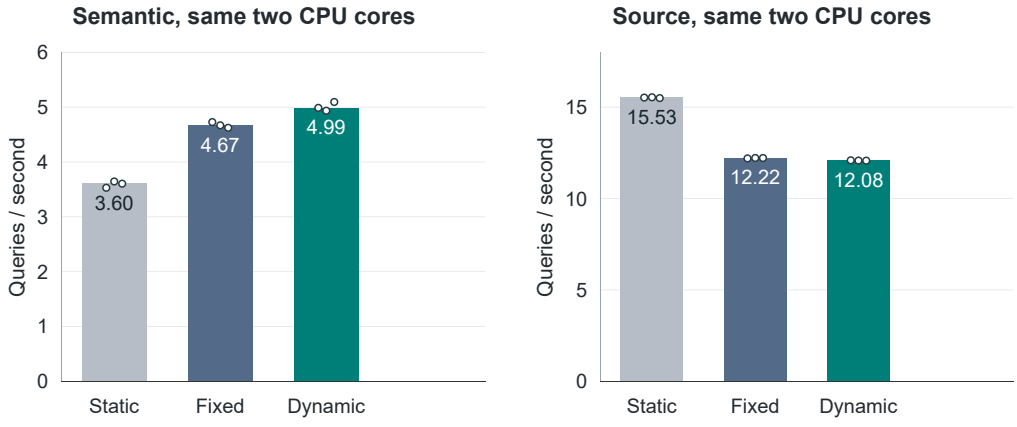

Figure 5: Frozen neural policies with two concurrent queries under the same two-core allowance. Bars show median throughput over three randomized blocks; dots show individual blocks. Every block executes the first 128 TEST rows, including public scoring, GPU prior inference, controller work, actual scans and canonical replay.

Prediction error and decision cost can move in opposite directions. On identical 16-provider semantic prefixes, residual conditioning changes unvisited-maximum RMSE from 0.05680 to 0.05946 while lowering label-visible sufficient remaining-prefix work from 211,560 to 186,350. Each comparison holds the paid prefix and returned boundary fixed. This offline ranking diagnostic (Appendix I) shows why maximum-score RMSE alone does not measure routing utility.

Finally, keeping the semantic corpus, ownership, queries and prior fixed, we test near-uniform virtual fees and two declared permutations of the original fees. Each condition refits a shared four-actor head. Dynamic saves 2.48%, 10.10% and 6.70% over fixed, with fixed/dynamic errors of 52/50, 54/60 and 51/57. Savings persist with nearly uniform prices and vary with their assignment. These development comparisons measure TUNE-calibrated virtual charges, not physical scan or monetary savings (Appendix H).

4.6 Actual neural execution

On semantic providers, dynamic reaches 4.99 queries/s versus 4.67 for fixed and 3.60 for static, with median process CPU times of 401, 428 and 553 ms/query. The shared warm public stage takes roughly 0.15–0.18 seconds per 128-query batch and is included in throughput. Under the same two-core allowance, parallel static search has lower single-query retrieval service time: 284 ms versus dynamic's 398 ms. Source static is also faster and more CPU-efficient than dynamic.

Timing starts from embeddings and excludes text encoding, network, cold loading and arrival queueing. All timed canonical returns match cached execution. Appendix J reports treatments, numerical expansions and validation; its rank-eight factor-state microbenchmark measures covariance scaling, not 1,000-provider retrieval. Tree-predictor and native ConANN comparisons retain their separate model and query identities in Appendix L.

5 Related work and implications

Correlated Pandora's Box and active feature acquisition distinguish useful outcomes from information about unopened alternatives (Chawla et al., 2020; 2023; Li & Oliva, 2021; Valancius et al., 2024). We study this distinction in complete top-k retrieval, using competitiveness per cost to compare dynamic and fixed ordering under feedback-based stopping.

RAGRoute predicts source relevance (Guerraoui et al., 2025), Optimist uses score moments for optimistic routing (Bruch et al., 2025), and AMIPS learns support functions (Olausson et al., 2026).

ConANN adapts search effort using intermediate results (Horchidan et al., 2025). These methods motivate our controls and native ConANN adapter. The latter retains its own query cohort and changes ownership, caps and loss to match complete top-10 retrieval. Learn-then-test, conformal risk control and utility-aware prediction sets provide established risk–utility tools (Angelopoulos et al., 2022; 2024; Lin et al., 2023; Cortes-Gomez et al., 2025). Here, independent testing assesses the executable policy rather than its internal residual model.

Ordering and stopping jointly determine work: earlier completion can be offset by later searches. Maximum-score error alone also misses changes in acquisition quality. Actual execution adds another distinction, because lower document work need not reduce latency.

The quality experiments fix provider count and k. Semantic and source conditions differ in corpus and ownership; virtual fees are not measured service prices. Cross-domain confirmation and fixed-corpus partition interventions remain needed to assess transfer. Additional acquisition, retraining and price controls have no new independent certificates.

6 Conclusion

The value of paid feedback depends on the complete route and stopping rule it induces. On the large semantic collection, the frozen neural rank-eight policy saves 6.59% mean work, with replay locating the reduction after result completion. Matched controls connect the gain to cost prioritization and the stopping point; source providers show how earlier completion can be offset by later searches. Joint work–risk evaluation and actual execution measurements reveal when these policy changes yield an efficiency benefit.

Reproducibility statement

Appendices A–J specify evidence identities, data provenance, model fitting, all 31 stop features, numerical semantics and execution. The supplement provides editable figure sources and result arrays, with reproduction code and checks. Its entry index separates result reproduction from experiments requiring large embeddings, provider indexes or neural checkpoints, which are not bundled.

AI use statement

Generative AI assisted literature review, hypothesis development and mathematical reasoning, writing, and experimental figure preparation. AI assistance also supported synthetic diagnostic experiments, which are separate from evaluations on public retrieval datasets.

Ethics statement

Retrieval omissions measure complete embedding top-k recovery, not factual correctness of generated answers. An average-risk target does not ensure equal performance across query groups. Consequential deployment requires subgroup, distribution-shift, privacy and resource evaluation.

References

Anastasios N. Angelopoulos, Stephen Bates, Emmanuel J. Candès, Michael I. Jordan, and Lihua Lei. Learn then test: Calibrating predictive algorithms to achieve risk control. *arXiv preprint arXiv:2110.01052v5*, 2022. URL https://arxiv.org/abs/2110.01052v5.

Anastasios N. Angelopoulos, Stephen Bates, Adam Fisch, Lihua Lei, and Tal Schuster. Conformal risk control. In *International Conference on Learning Representations*, 2024. URL https://openreview.net/forum?id=33XGfHLtZg.

Payal Bajaj, Daniel Campos, Nick Craswell, Li Deng, Jianfeng Gao, Xiaodong Liu, Rangan Majumder, Andrew McNamara, Bhaskar Mitra, Tri Nguyen, et al. MS MARCO: A human generated machine reading comprehension dataset. *arXiv preprint arXiv:1611.09268*, 2016. URL https://arxiv.org/abs/1611.09268.

Sebastian Bruch, Aditya Krishnan, and Franco Maria Nardini. Optimistic query routing in clustering-based approximate maximum inner product search. In *Advances in Neural Information Processing Systems*, volume 38, pp. 8370–8402, 2025. doi: 10.52202/085713-0256.

Shuchi Chawla, Evangelia Gergatsouli, Yifeng Teng, Christos Tzamos, and Ruimin Zhang. Pandora's box with correlations: Learning and approximation. In *IEEE Symposium on Foundations of Computer Science*, pp. 1214–1225, 2020. doi: 10.1109/FOCS46700.2020.00116.

Shuchi Chawla, Evangelia Gergatsouli, Jeremy McMahan, and Christos Tzamos. Approximating pandora's box with correlations. In *Approximation, Randomization, and Combinatorial Optimization. Algorithms and Techniques*, volume 275 of *Leibniz International Proceedings in Informatics*, pp. 26:1–26:24, 2023. doi: 10.4230/LIPIcs.APPROX/RANDOM.2023.26.

Jianlyu Chen, Shitao Xiao, Peitian Zhang, Kun Luo, Defu Lian, and Zheng Liu. M3-embedding: Multi-linguality, multi-functionality, multi-granularity text embeddings through self-knowledge distillation. In *Findings of the Association for Computational Linguistics: ACL 2024*, pp. 2318–2335, 2024. doi: 10.18653/v1/2024.findings-acl.137.

Santiago Cortes-Gomez, Carlos Patiño, Yewon Byun, Zhiwei Steven Wu, Eric Horvitz, and Bryan Wilder. Utility-directed conformal prediction: A decision-aware framework for actionable uncertainty quantification. In *International Conference on Learning Representations*, 2025. URL https://proceedings.iclr.cc/paper_files/paper/2025/hash/0c6b452f1bbfb6905f6bac957d73b321-Abstract-Conference.html.

Rachid Guerraoui, Anne-Marie Kermarrec, Diana Petrescu, Rafael Pires, Mathis Randl, and Martijn de Vos. Efficient federated search for retrieval-augmented generation. In *Proceedings of the 5th Workshop on Machine Learning and Systems (EuroMLSys)*, pp. 74–81, 2025. doi: 10.1145/3721146.3721942. URL https://euromlsys.eu/pdf/euromlsys25-14.pdf.

Sonia Horchidan, Fabian Zeiher, Henrik Boström, and Paris Carbone. ConANN: Conformal approximate nearest neighbor search. *Proceedings of the VLDB Endowment*, 19(1):29–42, 2025. doi: 10.14778/3772181.3772184.

Yang Li and Junier Oliva. Active feature acquisition with generative surrogate models. In *International Conference on Machine Learning*, volume 139 of *Proceedings of Machine Learning Research*, pp. 6450–6459, 2021. URL https://proceedings.mlr.press/v139/li21p.html.

Zhen Lin, Shubhendu Trivedi, Cao Xiao, and Jimeng Sun. Fast online value-maximizing prediction sets with conformal cost control. In *Proceedings of the 40th International Conference on Machine Learning*, volume 202 of *Proceedings of Machine Learning Research*, pp. 21182–21203, 2023. URL https://proceedings.mlr.press/v202/lin23j.html.

Theo X. Olausson, João Monteiro, Michal Klein, and Marco Cuturi. Amortized maximum inner product search with learned support functions. *arXiv preprint arXiv:2603.08001v3*, 2026. URL https://arxiv.org/abs/2603.08001v3. Version 3, July 1, 2026.

Nandan Thakur, Nils Reimers, Andreas Rücklé, Abhishek Srivastava, and Iryna Gurevych. BEIR: A heterogeneous benchmark for zero-shot evaluation of information retrieval models. In *NeurIPS Datasets and Benchmarks*, 2021. URL https://datasets-benchmarks-proceedings.neurips.cc/paper/2021/hash/65b9eea6e1cc6bb9f0cd2a47751a186f-Abstract-round2.html.

Michael Valancius, Maxwell Lennon, and Junier Oliva. Acquisition conditioned oracle for nongreedy active feature acquisition. In *International Conference on Machine Learning*, volume 235 of *Proceedings of Machine Learning Research*, pp. 48957–48975, 2024. URL https://proceedings.mlr.press/v235/valancius24a.html.

Wenhui Wang, Furu Wei, Li Dong, Hangbo Bao, Nan Yang, and Ming Zhou. MiniLM: Deep self-attention distillation for task-agnostic compression of pre-trained transformers. In *Advances in Neural Information Processing Systems*, volume 33, 2020. URL https://proceedings.neurips.cc/paper/2020/hash/3f5ee243547dee91fbd053c1c4a845aa-Abstract.html.

A EVIDENCE IDENTITIES AND REPRODUCTION ENTRY POINTS

Table 3 specifies which experiments support which claims. "Historical" means query roles already available during development; it does not mean that EVAL labels were used to fit these new models. The new controls select thresholds on TUNE, evaluate all declared variants on EVAL, and are not replacements for the independently confirmed policies. Post-confirmation TEST diagnostics reuse labels to understand a frozen result and make no new independence claim.

Table 3: Evidence map. F/D denotes fixed/dynamic actors. All substantive neural mechanisms in the main text use the final corrected initial mean.

ID	Model / variant	Queries	Purpose and executable entry point
E1	Frozen neural full and rank-8 F/D; static	New CERT/TEST: 8,192 each per condition	Confirmation; `run_mechanism_confirmation.py`
E2	Exactly E1 F/D	All 8,192 TEST rows per condition	Actual-cost decomposition; `replay_final_policy_mechanism.py`
E3	Final prior, rank-8, four acquisition rules	Historical TRAIN/TUNE/EVAL	Acquisition controls; `study_final_action_attribution.py`
E4	Final prior, full/rank-8, refitted paired heads	Historical roles; four TRAIN regimes plus six resamples	Stop-training factorial; same study entry as E3
E5	E1 rank-8 F/D and static	First 128 TEST rows; three blocks	Real searches and neural inference; `benchmark_final_neural.py`
E6	Two additional complete neural fits	Identical fit/held queries and augmented targets; historical TUNE/EVAL	Initialization/minibatch sensitivity; `repeat_final_neural_training.py`
E7	Synthetic rank-4/8 covariance	One query, eight updates, seven repeats per case	State representation only; `benchmark_final_rank_scaling.py`
E8	E1 legal paths, exchanged frozen thresholds	All TEST rows, post-confirmation	Threshold sensitivity; `diagnose_final_threshold_exchange.py`
E9	Final neural rank-8, three virtual-price assignments	Same historical semantic TRAIN/TUNE/EVAL	Price sensitivity; `study_final_virtual_prices.py`

The supplementary entry index separates three reproduction levels. Saved JSON and per-query arrays suffice to recompute tables, work differences and finite-pool bounds; editable figure sources accompany the result arrays. Figures in this manuscript are vector exports of the editable PPTX master. Re-executing policies additionally requires frozen response caches, public predictions and model files. Backend execution requires document/query embeddings, provider assignments and indexes, while training also requires the authors' AMIPS dependency and augmentation targets. Large vectors, indexes and neural checkpoints are not bundled.

B DATASETS, ENCODERS AND PARTITIONS

Semantic collection. The BEIR MS MARCO passage corpus contains 8,841,823 indexed rows. The encoder is `sentence-transformers/all-MiniLM-L6-v2`, revision

`c9745ed1d9f207416be6d2e6f8de32d1f16199bf`.

The pinned model card identifies the sentence-embedding variant and its pooling conventions. The full-corpus builder reads JSONL in file order, strips title/text, joins a nonempty title and text with one space, and uses `SentenceTransformer.encode` with normalized embeddings saved as float32. The snapshot's sentence configuration sets a maximum sequence length of 256; no local override is made. The progress record confirms corpus completion, encoder revision and normalization. Query embeddings use the same snapshot and all query JSONL rows in file order, with stripped text, normalized float32 output and batch 512. The query pool contains 509,962 rows, distinct from the number of indexed passages.

The 100 semantic centers are fitted on the first one million passage embeddings with spherical Faiss k-means, seed 42, 15 iterations and at most 512 training points per centroid. All 8.84M rows are then

assigned by maximum inner product to those frozen normalized centers. Thus these are extrapolated frozen centers, not a claim to refit 100 clusters on the entire corpus. Provider sizes range from 42,309 to 208,973. Canonical IDs are source-row integers. The supplement records hashes for the corpus vectors, query vectors, centers, assignments and sizes.

Source-pure collection. This condition derives from the MS MARCO *document-ranking* TSV, not a relabeling of the preceding passage corpus. Its 3,213,835 rows contain document ID, URL, title and body. Embeddings use stripped title plus body, excluding URL, with the same pinned MiniLM snapshot and normalization. Eight rows lacking an acceptable explicit URL source are removed, leaving 3,213,827 documents and 445,772 canonical hostnames. No unknown-source bucket is created. Sources are ranked by document count, with hostname as deterministic tie breaker, and the first 100 are retained whole. This yields 992,502 documents and one hostname per provider, with sizes 2,657–165,854. It is a head-source stress test: the long tail is excluded. Hostnames do not establish independently administered organizations, trust domains or policy autonomy. Query outcomes do not determine this selection.

FiQA. The historical cross-domain collection has 57,638 documents and uses `BAAI/bge-m3`, revision

`5617a9f61b028005a4858fdac845db406aefb181`.

The pinned model card documents the dense-vector interface. The pipeline loads the tokenizer and model directly with Transformers. Document text is title plus text, and query text is stripped. Tokenization uses padding and truncation to 256 tokens, batch 32; the embedding is the final hidden layer's CLS vector, L2-normalized in Torch, converted to float32 on CPU and normalized again. Document-only spherical k-means uses 100 centers, seed 42, 20 iterations, one restart and at most 256 training points per centroid. These are simulated providers, not institutions. Current FiQA controls reuse historical query roles; their results are not a fresh cross-domain confirmation.

Scope of provenance. Checkpoint identifiers describe the executed encoders, not which contemporary model would be best. The benchmark evaluates exact embedding top-10 recovery, not relevance judgments or downstream generated-answer quality. Main and source differ in corpus, source selection and ownership together, so their contrast cannot identify partition purity as a causal factor. Row-order and hash records define the reproducibility boundary; no dataset size is padded to ten million.

C Model fitting, budgets and thresholds

The common public scorer is a RAGRoute-style reimplementation with a 32-dimensional TRAIN-fitted query PCA, 16-dimensional owner profile derived from training centroid scores, and public score/rank/cost features. Its two hidden layers have widths 256/128, LayerNorm, ReLU and dropout 0.4. Its earlier 8,192-query training role includes a 20% internal validation split (6,553/1,639); the main and source scorers both select epoch 26 from at most 80. This role is distinct from the later 4,096-query TRAIN for the prior and stopping head and is included in confirmation exclusions.

The new maximum predictor fits 3,072 of those 4,096 queries, with a fixed split seed 2026092309. Each produces its original normalized vector plus four Gaussian perturbations of standard deviation 0.02, normalized afterwards, for 15,360 training examples. Targets are the strict maximizing document within each provider; its normalized vector supplies the gradient-matching target. The remaining 1,024 unaugmented queries estimate residual moments. FiQA analogously uses 1,536 fit queries and 512 held-out queries.

SupportNet uses eight layers, principled initialization, bias, input injection every four layers, identity rectifier, homogenization, and soft-leaky-ReLU with slope 0.1 and beta 20. The parameter target is $\max(4{,}000{,}000, 0.01 N_{\mathrm{docs}} d)$; actual counts are 33,970,340 (semantic), 4,039,108 (source), and 4,086,788 (FiQA). Score and gradient losses have weights 0.01 and one. Adam uses gradient clipping at one, 2,500 warmup steps to 5×10^{-5}, cosine decay to 5×10^{-6}, and EMA decay $0.999^{32/128}$. There are 100,000 batches of 32. Minimum TUNE maximum-score RMSE among 30k/60k/100k

selects the checkpoint, not EVAL risk or work. Generic optimizer/loss defaults in the architecture configuration are unused; the executed training specification defines these values.

The residual model uses the same held-out TRAIN residuals for full and rank-eight variants. Fixed order and dynamic belief both receive the same mean correction. The conditional engine is passed zero additional residual bias to avoid correcting twice. In the final pair, observed means are reset to actual maxima, so the older common-residual-shift feature remains zero; it is not a second update. Residual covariance is allowed to change the stop features, which is why each residual variant refits its paired head.

The stop classifier has 100 histogram-boosting iterations, maximum depth four, minimum leaf size 150, learning rate 0.08, L_2 regularization 10, no early stopping, and seed 2026092310. Query-balanced sampling uses up to the first 12 states and up to 12 evenly spaced later states per trajectory. The final paired head sees fixed and dynamic TRAIN trajectories. The four-actor control instead uses all four actors with the same state-sampling and capacity rules.

Head fitting uses trajectories admitted by B_0 before imposing the terminal absolute cap C, so some offline TRAIN states extend beyond the serving cap. The TUNE score-quantile grid likewise uses the recorded eligible B_0 states, while each candidate's evaluated outcome obeys both limits. This convention is held fixed across the matched controls; no beyond-cap return is revealed to an online policy. It is part of the training distribution whose sensitivity the factorial study examines, not a claim that all offline states are visited at serving time.

For centroid rank $r(q)$ of the last necessary provider, TUNE sets r_0 to the $\lceil (n_{\text{tune}} + 1)0.95 \rceil$th order statistic, or exhaustive rank when that index exceeds the sample. The recorded reference ranks are 49/94/55 for semantic/source/FiQA. Absolute caps are 3,760,104 / 866,297 / 30,451 document units. These are shared within each comparison. Dynamic admission selects among providers affordable under B_0 before applying C; the chosen provider is not silently replaced if it fails the latter check.

Threshold candidates are 129 TUNE score quantiles plus endpoints. The objective is mean plus 0.25 P99 work, under a primary 3.5% empirical guard. The historical acquisition study declares the guard set $\{1, 2, 2.5, 3, 3.5, 4, 4.5, 5, 7, 10\}\%$ and retains infeasible points. These guards are development choices, not independently certified operating points. The static neural prefix selects its work cap and common-seed convention on TUNE, with no feedback-based stopping.

The final semantic rank-eight thresholds are $\theta_F = 0.0032108694670632877$ and $\theta_D = 0.004999333651734538$; source thresholds are 0.02640186497074884 and 0.021273587942013807. Swapping them changes the complete policy and does not preserve certification.

D THE COMPLETE STOPPING-FEATURE DEFINITION

Let s_i be the public source score, z_i the query–centroid score, V the searched set, U its complement, $C_{\text{tot}} = \sum_i c_i$, $d_s = \max(\sum_i s_i, 10^{-9})$, and $r_1 \geq \cdots \geq r_k$ the current returned scores with $\tau = r_k$. Write $\max_U x = \max_i\{x_i \text{ if } i \in U, -1 \text{ otherwise}\}$. Features are converted to float32 in the stated order. Despite the legacy name "public features" for indices 1–17, several depend on paid state, and feature 15 uses the current conditioned mean. They are not all available before the first search.

Table 4: All 31 features of the final conditional stopping head. Indexing is one-based.

Indices	Group	Definition in order
1–3	Public source	$\max_i s_i$, $\text{mean}_i\, s_i$, $\text{std}_i\, s_i$
4–6	Centroid	$\max_i z_i$, $\text{mean}_i\, z_i$, $\text{std}_i\, z_i$
7–11	Budget/state	B_0/C_{tot}, $\lvert V\rvert/M$, probed fraction (zero), $W/\max(B_0, 1)$, $(B_0 - W)/C_{\text{tot}}$
12–17	Remaining state	$\sum_{i\in V} s_i/d_s$, $\max_U s$, $\max_U z$, $\max_U \mu$, $\sum_{i\in U} s_i/d_s$, $\sum_{i\in U} c_i/C_{\text{tot}}$
18–22	Paid returned scores	τ, r_1, $k^{-1}\sum_{j=1}^{k} r_j$, $r_1 - \tau$, $r_5 - \tau$ $(k = 10)$
23–26	Gain/support	Latest change in τ (zero at seed), max observed probe score (-1), $\max_U \mu - \tau$, $\max_U \mu - \tau$
27–31	Uncertainty/coverage	$\max_U \sigma$, $\max_i p_i$, $M^{-1}\sum_i p_i$, residual shift (zero), number of providers represented in returned top-k divided by k

Here $p_i = \mathbf{1}\{i \in U\}\Phi(\mathrm{clip}((\mu_i - \tau)/\sigma_i, -12, 12))$, and $\sigma_i = \sqrt{\max(\Sigma_{ii}, 0.002^2)}$. The two support-gap coordinates coincide in the primary conditional engine because observed coordinates already equal their paid values and the inherited extra shift is zero. Probe fields are compatibility coordinates: the evaluated primary policies buy no probes. The provider-count feature uses owners of returned document IDs, not hidden necessary-provider labels. No stop feature reads an unvisited local answer. Training labels and offline diagnostics are attached only after trajectory generation.

E FINITE-POPULATION RISK CALCULATION

Eligible query construction first excludes recorded identifiers, source rows and normalized-vector fingerprints from two historical run roots, including the inherited scorer's training roles. The lowest-index representative of each remaining fingerprint is fixed before randomization. The 174,115 representatives are randomly partitioned into disjoint condition populations; each supplies CERT and TEST by a fixed uniform permutation. CERT is therefore a uniform sample without replacement conditional on the frozen eligibility and partition. A missing/unreadable audit artifact causes failure, not silent exclusion from the audit.

Let the frozen policy make K errors in its condition population. The CERT count E has distribution $\mathrm{Hypergeom}(N_{\mathrm{pool}}, K, n)$. Inverting its lower tail gives $\Pr(K^+ < K) \le \delta/J$. A union bound across ten slots gives joint coverage at least $1 - \delta$, without requiring independent errors across methods. On that coverage event, removal of the observed e CERT errors leaves at most $K^+ - e$ errors among $N_{\mathrm{pool}} - n$ queries. This justifies Equation 6. The finite-population assumption is about the executed sampling design, not a claim of IID production traffic.

For semantic rank-eight dynamic, $N_{\mathrm{pool}} = 87{,}057$, $n = 8{,}192$, $e = 268$ and $K^+ = 3{,}295$, hence $U = 3{,}027/78{,}865 = 3.8382045\%$. Source's remaining population has 78,866 queries. Neither denominator removes TEST: TEST belongs to the remaining population. The table reports all ten e, K^+ values. The critical integer boundaries were independently checked using log-combination tail summation: the CDF at K^+ exceeds 0.005 and at $K^+ + 1$ does not. This is a numerical cross-check, not formal verification of the full software stack.

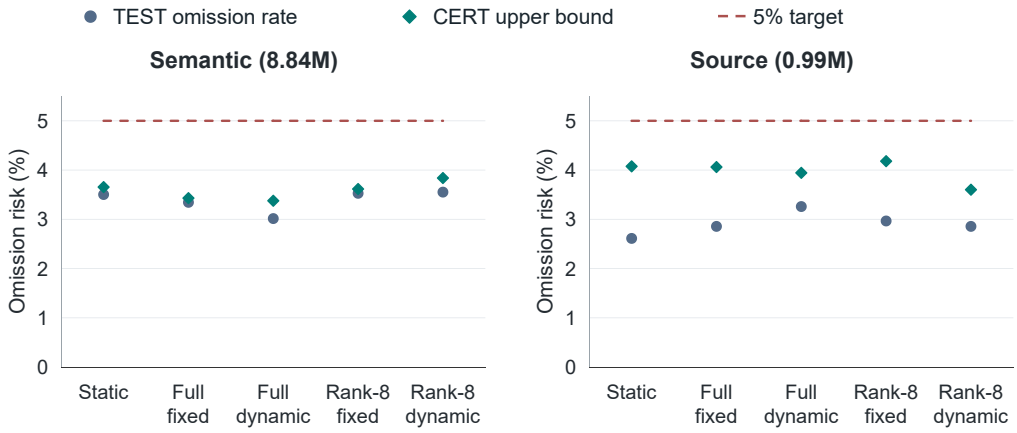

Figure 6: Separate markers for TEST empirical risk and the independent CERT upper bound on the remaining finite pool. A bound is not an error bar centered on the adjacent TEST point. Both use 8,192 queries per condition; no point is a per-query correctness guarantee.

The deterministic work bound is separate. Each admitted search satisfies $W + c_i \le \min(B_0, C)$ and adds exactly c_i, so induction proves the cap. Under disjoint ownership, exact local top-k plus canonical merging recovers the global top-k if all necessary providers were searched. An omitted necessary provider necessarily withholds one of those documents. This proves the loss equivalence used by cache replay. It does not apply without change to approximate local search or to native baseline heaps with differing arithmetic; those returns are evaluated directly by canonical document IDs.

Table 5: Actual work decomposition for final rank-8 on all 8,192 TEST queries per condition. Each cost column is averaged over all queries. Failed-query work is not an oracle continuation. Budget-infeasible necessary-provider sets occur zero times in these two pairs.

Condition	Policy	Before	After	Failed work	Errors	Path	Premature
Semantic	Fixed	547,013	1,794,249	123,712	289	264	25
Semantic	Dynamic	563,598	1,622,958	116,075	291	238	53
Source	Fixed	551,850	201,902	23,172	243	101	142
Source	Dynamic	480,274	274,603	22,071	234	91	143

Table 6: Paired success strata for final rank-8. Costs are conditional means within each stratum, in document units, and report fixed minus dynamic. The three components sum to the total; a positive value favors dynamic.

Condition	Stratum	n	Before	After	Failed work	Total
Semantic	Both succeed	7819	-13,722	180,506	0	166,783
Semantic	Fixed only succeeds	84	2,412,372	654,526	-2,846,874	220,024
Semantic	Dynamic only succeeds	82	-2,819,637	-769,976	3,588,825	-788
Semantic	Both fail	207	0	0	35,812	35,812
Source	Both succeed	7857	75,235	-75,266	0	-32
Source	Fixed only succeeds	92	723,540	48,937	-759,088	13,389
Source	Dynamic only succeeds	101	-706,249	-86,136	782,917	-9,468
Source	Both fail	142	0	0	-1,535	-1,535

F FINAL-POLICY TRAJECTORY DIAGNOSTICS

For each frozen policy, we record its prescribed legal trajectory without learned early stopping. Labels identify the first complete result afterwards. Actual stopping restricts which recorded steps were executed; the rest are explicitly counterfactual and are never charged as actual work. For successful queries define W_{before} as work through first completion and $W_{\text{after}} = W - W_{\text{before}}$. For failed queries, both are zero and $W_{\text{fail}} = W$. Thus

$$W = W_{\text{before}} + W_{\text{after}} + W_{\text{fail}}$$

query by query, not only on average. A label-visible same-path oracle on a failed actual query may stop later than the real policy; its cost is a different counterfactual quantity.

The failure hierarchy is: necessary-provider cost exceeds the cap; otherwise the prescribed legal path never contains all necessary providers; otherwise learned stopping precedes a later complete prefix. This decomposition is conditional on the action model and path, not a proof that some other policy could safely recover every feasible-budget query. For semantic rank-eight, both-success/fixed-only/dynamic-only/both-failure counts are 7,819/84/82/207. For source they are 7,857/92/101/142. The four groups are reported rather than conditioning the main efficiency claim only on successful cases.

Paired query bootstrap uses 2,000 replicates and seed 2026092525 for descriptive 95% decomposition intervals. The semantic mean-work difference interval is $[154{,}795, 169{,}974]$ document units. For source, it is $[-516, 454]$ units and the after/total ratio is suppressed. These are repeated-query intervals, not training uncertainty, post-selection-adjusted mechanism tests, or new certificates.

Innovation clipping occurs in 0.936%/0.998% of actual semantic fixed/dynamic rank-eight updates and 0.444%/0.436% of source updates. The unvisited standard-deviation floor activates zero times in these actual replay states. These observations quantify the numerical safeguards' use in the evaluated runs; they do not justify removing them from the frozen policy.

Threshold exchange uses only the two already-frozen thresholds, not a TEST-tuned grid. Writing $W_{a,b}$ for order a and threshold from b, the symmetric path contrast is $\frac{1}{2}[(W_{F,F}-W_{D,F})+(W_{F,D}-W_{D,D})]$; the threshold contrast is $\frac{1}{2}[(W_{F,F} - W_{F,D}) + (W_{D,F} - W_{D,D})]$. They sum algebraically to the observed diagonal difference. For semantic rank eight they are $-49{,}493$ and $+211{,}835$; source gives $+10{,}502$ and $-10{,}526$. These contrasts describe numerical-score operating points with different

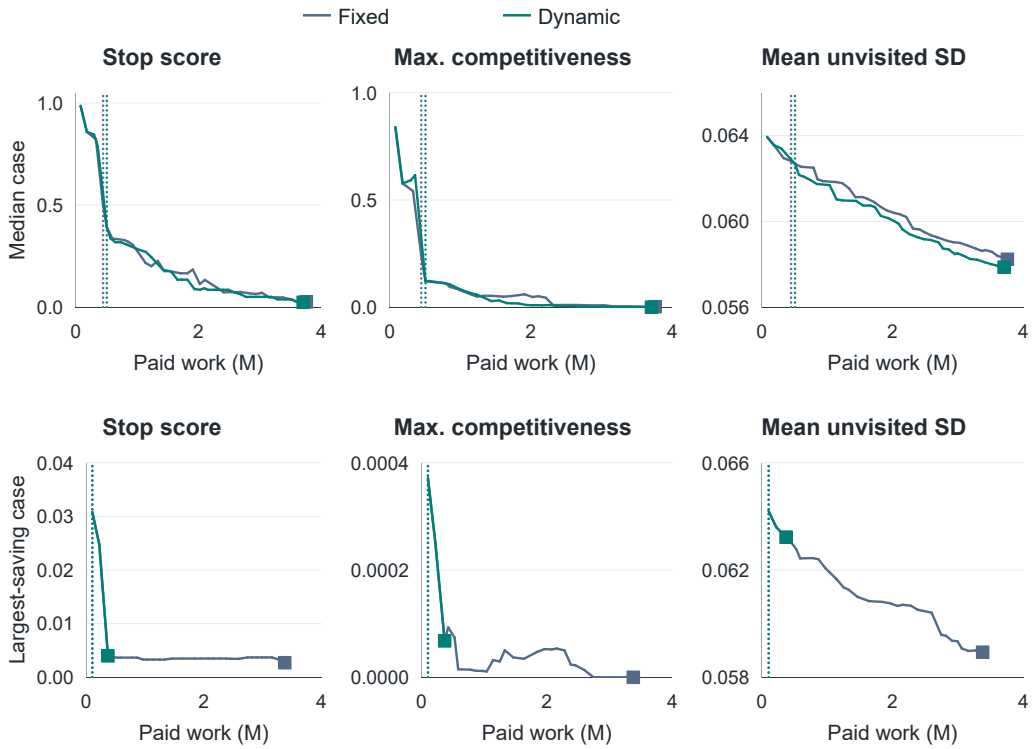

Figure 7: Actual paid-state traces of final semantic rank-eight policies. Dotted vertical lines mark first completion and squares mark actual stopping. Selection is explicit: the median and largest paired saving among common successes, with cache order breaking ties. The largest-saving row is an illustrative extreme, not a prevalence estimate or evidence used to tune a policy.

realized risks, not uniquely identified causal shares of information value. The full-covariance counterparts are $-5{,}590/-154{,}064$ (semantic) and $+9{,}885/+306$ (source).

G ACQUISITION CONTROLS AND MATCHED TRAINING

The one-shot control freezes μ_0, σ_0 and the seed-return boundary τ_1 in its action score. Boundary-only freezes μ_0, σ_0 but uses current τ_t. Both preserve full paid conditioning for stop features. One-shot uses a fixed *priority among affordable actions*, not the strict-prefix affordability rule of fixed support; hence fixed-support to one-shot changes both the priority and that admission behavior. One-shot to boundary-only isolates the boundary in acquisition, and boundary-only to full feedback adds the conditioned action means/scales under the same affordability rule. All use the common seed and check C after selecting within B_0.

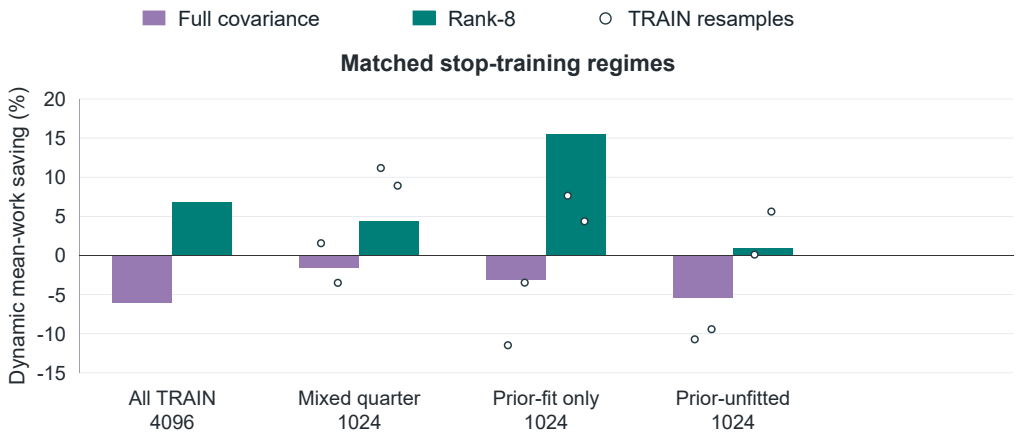

Figure 8: Historical semantic EVAL: anchor stop-training regimes (bars) and two additional query-bootstrap draws per quarter-sized regime (open dots). Positive means lower dynamic work. Every comparison fixes the corrected prior and refits a common paired head. Dots are training-query perturbations, not confidence intervals or new independent test sets.

The mixed quarter contains 768 prior-fit and 256 prior-unfitted queries; prior-fit-only and prior-unfitted-only contain 1,024 each. Anchor subset seed is 2026092530. Resampling seeds 2026092531/32 draw within those same groups with replacement, preserving the declared composition. The resulting shared head and each actor's threshold are refitted, so variability includes stop-model fitting and its TUNE operating point. "Prior-unfitted" means the maximum predictor did not fit those query targets. Those queries still estimate residual moments, and the inherited source scorer has its own earlier training history; this is not full cross-fitting of every component.

The complete neural repeats change initialization and minibatch sampling while holding the exact augmented target file fixed. They are more extensive than changing the seed of a deterministic tree head, but do not resample the corpus, encoder, partition, augmentation noise or query split. Each run gets its own residuals, head and TUNE thresholds. All runs, including infeasible and negative results, are retained. They provide no replacement certificate and are not used to pick a new winner for the previously opened TEST pool.

Table 7: Acquisition controls on historical semantic EVAL. All use the same neural rank-8 prior and full paid conditioning for stop features. Each threshold is chosen on TUNE at a 3.5% guard. The refitted head sees all four actors during training. None of these additional policies inherits a confirmation certificate.

Stop head	Action rule	Mean (M)	P99 (M)	EVAL errors	TUNE feasible
Frozen	Fixed support	2.439	3.758	52/2048	yes
Frozen	One-shot cost	2.276	3.757	65/2048	yes
Frozen	Boundary only	2.315	3.757	69/2048	yes
Frozen	Dynamic	2.273	3.757	50/2048	yes
Refitted	Fixed support	2.493	3.758	54/2048	yes
Refitted	One-shot cost	2.298	3.757	65/2048	yes
Refitted	Boundary only	2.337	3.757	70/2048	yes
Refitted	Dynamic	2.383	3.757	48/2048	yes

Table 8: The same acquisition controls on source and FiQA historical EVAL. Infeasible TUNE guards remain visible. FiQA is a historical cross-domain diagnostic, not an untouched confirmation.

Condition	Head	Action rule	Mean work	Errors	TUNE feasible
Source	Frozen	Fixed support	777,339	63/2048	yes
Source	Frozen	One-shot cost	841,221	142/2048	no
Source	Frozen	Boundary only	809,913	63/2048	yes
Source	Frozen	Dynamic	777,387	51/2048	yes
Source	Refitted	Fixed support	772,259	67/2048	yes
Source	Refitted	One-shot cost	844,277	142/2048	no
Source	Refitted	Boundary only	808,149	62/2048	yes
Source	Refitted	Dynamic	769,329	60/2048	yes
FiQA	Frozen	Fixed support	19,505	38/1024	yes
FiQA	Frozen	One-shot cost	21,159	47/1024	yes
FiQA	Frozen	Boundary only	28,394	53/1024	no
FiQA	Frozen	Dynamic	28,161	30/1024	no
FiQA	Refitted	Fixed support	19,689	36/1024	yes
FiQA	Refitted	One-shot cost	21,562	47/1024	yes
FiQA	Refitted	Boundary only	28,377	53/1024	no
FiQA	Refitted	Dynamic	28,134	30/1024	no

Table 9: Matched stopping-training factorial on historical semantic EVAL (2,048 queries). Initial mean correction, predictor, caps and head capacity are held fixed. Thresholds are retuned only on TUNE. F/D errors are reported separately; these are not equal-risk or newly certified comparisons.

Residual	Stop TRAIN	n	F work (M)	D work (M)	Saving (%)	F/D errors	D feasible
Full	all	4096	2.633	2.793	-6.06	52/53	no
Full	mixed quarter	1024	2.729	2.774	-1.64	52/54	no
Full	fit quarter	1024	2.961	3.055	-3.15	49/52	no
Full	unfitted quarter	1024	2.600	2.743	-5.49	52/55	no
Rank-8	all	4096	2.439	2.273	+6.79	52/50	yes
Rank-8	mixed quarter	1024	2.551	2.437	+4.44	54/47	yes
Rank-8	fit quarter	1024	2.798	2.365	+15.46	51/54	yes
Rank-8	unfitted quarter	1024	2.688	2.661	+0.99	50/45	yes

Table 10: Two additional query-bootstrap TRAIN draws per 1,024-query regime; same prior and residual covariance. This resamples training queries, including duplicates, rather than changing a deterministic classifier seed. It does not measure neural-training or partition uncertainty. Seeds end in 31 and 32.

Residual	Regime	Seed	F work (M)	D work (M)	Saving (%)	F/D errors	D feasible
full	mixed	31	2.955	2.909	+1.57	48/53	no
full	mixed	32	2.780	2.877	-3.49	49/53	no
full	fit	31	2.948	3.286	-11.47	48/52	no
full	fit	32	2.750	2.845	-3.47	48/53	no
full	unfitted	31	2.701	2.990	-10.71	50/53	no
full	unfitted	32	2.626	2.874	-9.42	53/54	no
rank8	mixed	31	2.812	2.498	+11.18	48/46	yes
rank8	mixed	32	2.794	2.544	+8.92	51/47	yes
rank8	fit	31	2.818	2.602	+7.64	49/47	yes
rank8	fit	32	2.649	2.533	+4.37	53/47	yes
rank8	unfitted	31	2.692	2.689	+0.12	52/46	yes
rank8	unfitted	32	2.667	2.518	+5.61	51/50	yes

Table 11: Complete neural-training repeats on historical semantic EVAL. Two additional initialization/minibatch seeds reuse the identical prior-fit queries and augmented targets. Each run performs 100,000 steps, chooses its checkpoint on TUNE RMSE, refits rank-8 residuals and a shared paired stop head, and tunes both thresholds. All runs are reported; none replaces the frozen confirmed model.

Training seed	Fixed work (M)	Dynamic work (M)	Saving (%)	F/D errors	D feasible
2026092501 (original)	2.439	2.273	+6.79	52/50	yes
2026092531	2.357	2.146	+8.94	49/54	yes
2026092532	2.257	2.126	+5.80	53/58	yes

H SAME-CORPUS VIRTUAL-PRICE INTERVENTIONS

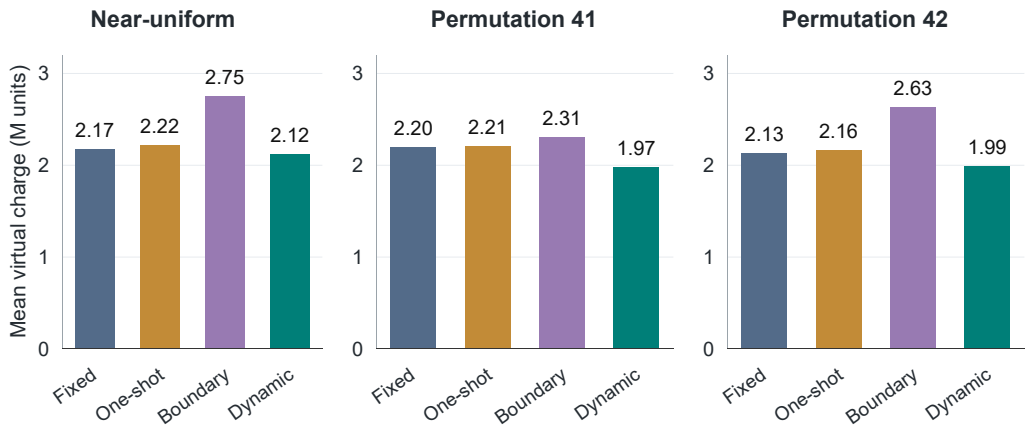

Figure 9: The corpus and prior are fixed; virtual prices change. Bars show historical EVAL charges at the declared 3.5% TUNE guard, with all four actors sharing a newly fitted head per condition. Corresponding omissions and feasibility are in Table 12; lower bars alone do not imply dominance.

The near-uniform vector divides the original total price as evenly as integer units permit, so prices differ by at most one unit. The other two vectors permute the original fees using seeds 2026092541/42. The price multiset is therefore identical under permutation, but its assignment to providers is randomized. All prices remain positive and every vector sums to 8,841,823; the absolute cap stays at 3,760,104, while B_0 is recomputed from the unchanged public centroid order and new fees.

Corpus vectors, partition, queries, encoder, upstream scorer, neural means, bias correction and rank-eight covariance are fixed. Each condition refits one same-capacity stop head on all four TRAIN actors and tunes thresholds under the same rule. Thus it tests the price response of a retuned pipeline, not an isolated coefficient change with a frozen classifier. No CERT or final TEST queries enter the study. These are virtual charge units, not observed currency prices or counts of physically scanned documents.

Near-uniform pricing gives dynamic mean charge 2.119M with 50 errors, versus fixed 2.173M/52 and one-shot 2.217M/62. The two permutations give dynamic 1.974M/60 and 1.988M/57, versus fixed 2.195M/54 and 2.131M/51. Dynamic is cheaper in all three, but the permuted cases involve more errors. This excludes dependence on the original size-based price assignment as the sole explanation, not all possible pricing or stopping-calibration effects. Full declared guard curves remain in the accompanying results.

Table 12: Same semantic corpus and queries, changed virtual prices. Near-uniform charges differ by at most one unit; permutations preserve the exact original price multiset. The total charge and absolute cap are unchanged. A common four-actor head is refitted per price condition, and thresholds use only TUNE at 3.5%. These are virtual charges, not measured document scans or currency.

Virtual prices	Action rule	Mean (M)	P99 (M)	EVAL errors	TUNE feasible
Near-uniform	Fixed support	2.173	3.714	52/2048	yes
Near-uniform	One-shot cost	2.217	3.714	62/2048	yes
Near-uniform	Boundary only	2.750	3.714	58/2048	yes
Near-uniform	Dynamic	2.119	3.714	50/2048	yes
Permutation 41	Fixed support	2.195	3.757	54/2048	yes
Permutation 41	One-shot cost	2.208	3.758	56/2048	yes
Permutation 41	Boundary only	2.307	3.758	57/2048	yes
Permutation 41	Dynamic	1.974	3.757	60/2048	yes
Permutation 42	Fixed support	2.131	3.757	51/2048	yes
Permutation 42	One-shot cost	2.164	3.756	55/2048	yes
Permutation 42	Boundary only	2.626	3.758	54/2048	yes
Permutation 42	Dynamic	1.988	3.757	57/2048	yes

I SAME-PREFIX PREDICTION AND DECISION MEASUREMENTS

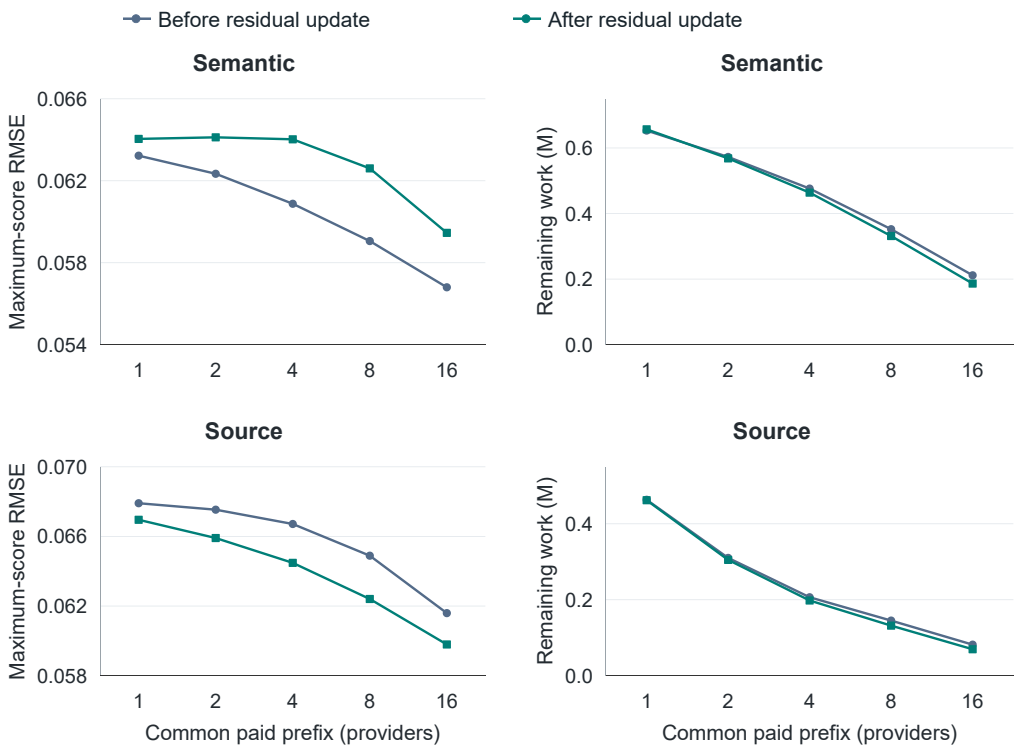

Figure 10: Final neural rank-eight model on identical paid prefixes, all 8,192 TEST queries per condition. Left: unvisited-maximum RMSE. Right: offline sufficient remaining-prefix work under the acquisition ranking. The returned boundary is held fixed within every before/after comparison. The two metrics can move in opposite directions.

At prefix lengths 1, 2, 4, 8 and 16, both rankings see the same paid providers and canonical returned boundary. The "prior" ranking uses the initial corrected means/scales; "updated" uses the residual-conditioned ones. A label-visible calculation ranks the remaining providers and sums costs through

Table 13: Final rank-8 same-prefix diagnostics on TEST. The paid prefix and returned boundary are identical before/after conditioning. Labels measure prediction error and a sufficient remaining ranked-prefix cost offline; the latter is not an executable zero-risk stopping rule.

Condition	Prefix	Eligible n	Prior RMSE	Updated RMSE	Prior work	Updated work
Semantic	1	8192	0.06322	0.06404	653,073	656,872
Semantic	2	8192	0.06234	0.06412	572,627	568,189
Semantic	4	8192	0.06088	0.06402	476,120	463,715
Semantic	8	8192	0.05906	0.06260	352,190	331,390
Semantic	16	8192	0.05680	0.05946	211,560	186,350
Source	1	8192	0.06791	0.06696	463,165	462,015
Source	2	8192	0.06754	0.06591	310,114	304,864
Source	4	8192	0.06671	0.06448	206,513	197,983
Source	8	8192	0.06489	0.06241	145,058	131,757
Source	16	8192	0.06159	0.05979	81,697	69,612

the last necessary provider. It does not apply learned early stopping, need not fit in the remaining budget, and is not an executable oracle. Eligibility uses only whether the common paid prefix fits the prescribed budget. It holds for all 8,192 queries at these prefix lengths in both conditions.

Source rank eight improves both RMSE and this remaining-prefix metric, yet has no net actual-work gain. Semantic rank eight worsens RMSE at all five common prefixes, while reducing sufficient remaining-prefix work at four of them. Combined with the actual-cost decomposition, these results caution against treating maximum-score RMSE, offline rank utility and stopped-policy work as interchangeable objectives.

J Actual neural execution and numerical semantics

The timing host has Xeon Silver 4210R CPUs and an RTX 3090 (24 GiB). The process is pinned to two distinct physical cores. At concurrency two, each query uses one native search thread; the single-query parallel static treatment uses two threads inside selected-provider search. The read-only index is shared. Treatments are randomized within three blocks, and all use the same first 128 TEST rows without outcome filtering. Other users' processes are not stopped. GPU-intensive jobs are serialized; measured timings precede the additional neural retraining jobs.

The public stage recomputes centroid scores, the required common source scorer and the restored SupportNet prior in GPU batches of 64. It is charged once per treatment batch in elapsed throughput and process CPU time. Text encoding, network communication, cold loading and arrival queueing are excluded. Median block statistics are reported; medians of separate components need not add exactly. The 128-query subset's mean work (semantic fixed/dynamic 2.609M/2.429M) differs from the all-8,192 means and is not substituted for them.

Before timing, the recomputed public predictions, source scores and centroid orders exactly match the saved values on the measured inputs. On 32 historical TRAIN queries per actor, the compiled state engine matches reference actions, all 31 features, stopping decisions and outcomes; the maximum observed feature difference is zero. Every timed canonical returned list and work/error outcome is additionally compared with frozen execution. These tests establish observed equivalence on the checked inputs, not a proof for every floating-point input.

Numerical candidate validation. The native Faiss backend scans normalized float32 vectors, then recomputes candidate scores from the source vectors normalized in float64. With $u = 2^{-24}$ and embedding dimension d, set

$$\gamma_d = \frac{(4d+32)u}{1-(4d+32)u}, \qquad \epsilon = \gamma_d(\|q_{32}\| + \max_i \|d_{i,32}\|)^2 + \Delta(4+\Delta) + 2\times 10^{-10},$$

where $\Delta = \|q_{32} - q_{64}\| + \max_i \|d_{i,32} - d_{i,64}\|$ and maxima range over the selected providers. This implements a conservative arithmetic/normalization screen, with actual candidate discrepancies checked against the envelope. Search initially asks for 32 candidates plus an omitted-candidate boundary. The canonical tenth score must strictly exceed the upper bound for omitted candidates

Table 14: Actual neural-policy backend timing: 128 fixed TEST rows, three randomized blocks, median block statistics, two physical Xeon Silver 4210R cores and an RTX 3090 for the shared prior. Q is concurrent queries. Throughput and CPU time include batched public scoring and the neural prior. Retrieval service excludes the preceding public batch and queueing; public time is the amortized cost per query, not individual arrival latency.

Condition	Method	Q	QPS	CPU ms/q	Service ms	Public ms/q	Control ms/q
Semantic	Static	1	1.81	555.3	551.6	1.34	0.09
Semantic	Static parallel	1	3.50	566.7	284.4	1.35	0.10
Semantic	Fixed	1	2.29	438.3	434.6	1.38	4.96
Semantic	Dynamic	1	2.50	401.7	398.1	1.34	4.66
Semantic	Static	2	3.60	552.7	549.4	1.35	0.11
Semantic	Fixed	2	4.67	427.8	424.9	1.26	5.17
Semantic	Dynamic	2	4.99	400.9	397.9	1.33	5.02
Source	Static	1	8.33	122.1	118.8	1.19	0.08
Source	Static parallel	1	15.48	130.0	63.5	1.12	0.09
Source	Fixed	1	6.68	151.8	148.6	1.17	8.58
Source	Dynamic	1	6.66	152.2	148.9	1.17	8.95
Source	Static	2	15.53	129.0	125.8	1.22	0.12
Source	Fixed	2	12.22	160.9	160.8	1.16	10.29
Source	Dynamic	2	12.08	161.9	162.2	1.17	11.01

(rounded outwards); otherwise candidates double. If separation is unresolved at width 4,096, a bounded-memory exact float64 scan returns the canonical result. Ties use global IDs. A boundary that remains unresolved is not accepted as exact merely because it is numerically close.

On the measured semantic sample no expansion occurs. In source, each fixed/dynamic treatment has one expansion per 128-query block; static has none. Across all blocks and settings, the maximum expansion count per query is one, so no width-4,096 exact fallback is reached. Source adaptive native distance counts exceed paid document work by 22.39 evaluations/query due to that repeated scan. Canonical replay counts and native distance counts are separate from charged unique document work: the cap is not a total-FLOP bound. A static plan merges selected providers within one native call; adaptive search returns between providers, which incurs both extra control and more canonical replay.

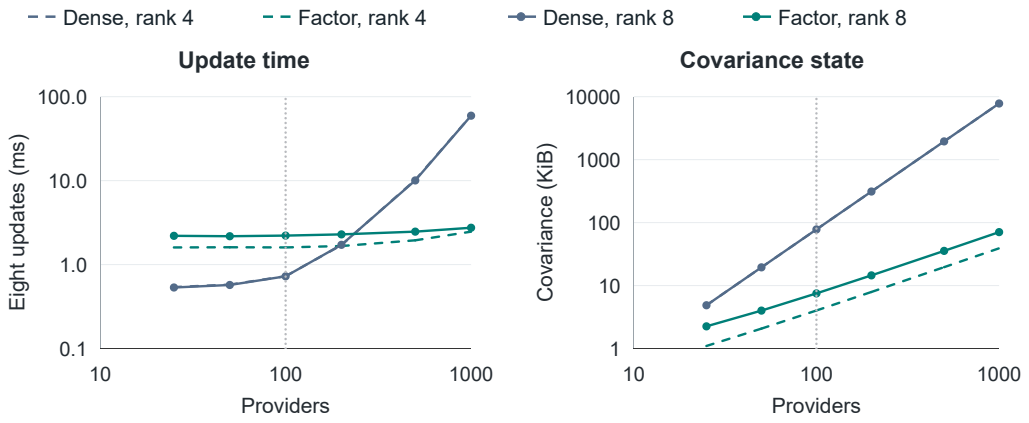

Figure 11: Reference dense/factor covariance microbenchmark, ranks four and eight, eight updates with seven repetitions. The dotted line marks the actual retrieval setting $M = 100$. All means and standard deviations are checked after every update. This is covariance-state scaling, not end-to-end retrieval quality or latency at 1,000 providers.

The rank-eight factor representation takes 2.22 ms for eight updates at $M = 100$, versus 0.73 ms for dense; at $M = 1{,}000$ the figures are 2.75 versus 59.47 ms. Covariance storage is 7,712 versus 80,000 bytes at $M = 100$ and 72,512 versus 8,000,000 at $M = 1{,}000$. The live serving experiment uses its

verified native dense state at $M = 100$, not this Python factor implementation. Numerical mean/scale agreement in the microbenchmark is within 2.3×10^{-16}.

K AN INFORMATION-ONLY ACTION THE ACQUISITION RULE CAN MISS

This response-level construction illustrates the established distinction between direct retrieval value and information value (Chawla et al., 2020; 2023). It is not an additional embedding benchmark or a new theorem. A common seed costs 10 and returns ten scores at most 0.4. One of providers B/C, each costing 100, holds all ten globally best documents with scores above 0.89; the other returns scores below 0.31. Their roles are equiprobable. Provider A costs $a \geq 10$, contributes no final top-10 document, and returns a signal about the winner with accuracy $(1 + \rho)/2$. Varying ρ keeps each provider's marginal score distribution fixed.

With the same known public moments, competitiveness per cost searches B/C before A. A diagnostic information-first policy instead pays A after the seed and then resumes that same rule. Both stop using the public score threshold valid in this construction, not the hidden winner label. Exact enumeration gives expected work 160 for the direct rule and $160 + a - 50\rho$ for information-first, both with zero omissions under the full common budget. Buying the signal helps iff $a < 50\rho$. Conditioning on past paid observations therefore does not itself make an acquisition rule optimal for information gathering.

L HISTORICAL BASELINE AND BACKBONE CONTEXT

The comparisons in this appendix predate the final neural confirmation. They retain their own model identities, query cohorts and calibration assumptions. They are useful baseline and sensitivity evidence, but their efficiency percentages, mechanism shares and runtime measurements are not assigned to the final neural policy.

L.1 ORIGINAL CONFIRMATION AND NATIVE CONANN

The original tree-based family has 24 policy–condition slots and separate 8,192-query CERT/TEST roles for each condition. It uses the one-sided binomial upper limit

$$U_j = \mathrm{Beta}^{-1}(1 - 0.05/24; e_j + 1, n_j - e_j),$$

with $U_j = 1$ if all queries fail. Its confidence interpretation assumes IID queries independent of the policy. The final neural confirmation instead matches its explicit without-replacement sampling to the finite-pool calculation in Equation 6. Neither version licenses a bound on arbitrary future traffic.

The original semantic tree policy uses 30.99% less mean work than support-order native ConANN and 12.57% less than its same-model fixed order; source reductions are 8.33% and 5.90%. These are not the neural policy's percentages. Original fixed order ranks raw maximum predictions while its dynamic prior includes the residual mean. New calibrated controls correct both equally, as does the final neural confirmation.

The ConANN adapter executes the authors' Faiss-based C++ online kernel. The bridge supplies existing provider lists, external centroid/scorer/support orders, a hard work cap and the canonical any-top-10 loss. The original algorithm's squared-L2 search uses normalized float32 vectors, $k = 10$ and maximum distance four. Its unaffordable order suffix is replaced by -1. Scanned-list and distance counters determine charged work, including the paid look-ahead when the previous heap is returned. Native returned IDs are compared directly to canonical targets rather than assuming numerical agreement.

Historical TUNE evaluates $k_{\mathrm{reg}} = 1$, $\gamma \in \{0, 0.001, 0.01, 0.1\}$ and 129 score-quantile thresholds cast to float32 before selection/execution. It uses the same mean-plus-quarter-P99 objective and 3.5% guard. A scanner tuning trace is checked against the actual online routine before its frozen confirmation. The changed ownership, order, cap and loss are adapter choices; this is not a reproduction of the authors' entire original benchmark.

Retaining the paid look-ahead repairs zero canonical omissions on semantic providers and one on source. Canonical reconstruction on the same original retained prefix removes zero and 22 numerical

Table 15: Original confirmation family, including rejected native centroid-order variants. Mean/P99 use the original TEST; upper bounds use its CERT.

Method	Mean M	P99 M	TEST misses	CERT upper %	Pass
Semantic owners					
Support-Cost	3.709	3.759	260/8192	3.86	Yes
Cost	3.708	3.759	282/8192	4.11	Yes
Public-stop	3.426	3.759	295/8192	4.24	Yes
Feedback-stop	3.030	3.759	286/8192	4.14	Yes
Support-feedback	2.965	3.758	263/8192	3.91	Yes
Conditional-fixed	2.829	3.758	265/8192	3.94	Yes
Adaptive (primary)	2.473	3.757	275/8192	3.82	Yes
Conditional-EI	2.922	3.758	248/8192	3.58	Yes
Independent-residual	3.237	3.759	305/8192	4.41	Yes
ConANN, centroid	3.608	3.759	459/8192	6.08	No
ConANN, upstream	3.585	3.758	286/8192	4.19	Yes
ConANN, support	3.584	3.758	260/8192	3.90	Yes
Source owners					
Support-Cost	0.860	0.866	249/8192	3.48	Yes
Cost	0.860	0.866	261/8192	3.77	Yes
Public-stop	0.855	0.866	266/8192	3.89	Yes
Feedback-stop	0.857	0.866	264/8192	3.82	Yes
Support-feedback	0.848	0.866	267/8192	3.69	Yes
Conditional-fixed	0.837	0.866	282/8192	3.77	Yes
Adaptive (primary)	0.788	0.866	290/8192	4.24	Yes
Conditional-EI	0.789	0.866	288/8192	4.23	Yes
Independent-residual	0.858	0.866	349/8192	4.64	Yes
ConANN, centroid	0.844	0.866	1448/8192	18.80	No
ConANN, upstream	0.860	0.866	287/8192	4.24	Yes
ConANN, support	0.860	0.866	276/8192	3.90	Yes

Table 16: ConANN support-order diagnostics on its original 8,192 TEST queries. Native/canonical columns change only return arithmetic on the same retained prefix. Keep retains the already-paid look-ahead. Retuned keep is a separate non-author variant, not newly certified.

Condition	Native	Canonical	Keep native	Keep canon.	Retuned (M)	Errors
Semantic owners	260	260	260	260	3.508	260
Source owners	276	254	275	253	0.860	275

return errors. The diagnostic variant that keeps look-ahead and retunes on historical TUNE reduces semantic mean work from 3.584M to 3.508M, still above the original tree controller's 2.473M. It is explicitly not the authors' method and does not inherit an old certificate. A same-path reconstruction also cannot measure every effect of retraining a canonical implementation.

L.2 Prior calibration and residual scans

After both fixed and dynamic actors receive the same initial residual mean correction, the full-covariance semantic tree pair uses 2.864M/2.540M work with 64 errors each on historical EVAL. Its actual-work difference splits into 707 before completion, 318,306 after completion and 4,423 on failures. The often-large after-completion saving is thus also present with this weaker prior. The final neural replay in the main text now tests the phenomenon directly instead of relying on transfer from the tree model.

The earlier diagonal/rank-1/rank-4/rank-8/full sweeps refit a same-capacity stop head for each residual model but retain the original raw fixed ranking. Their complete outcomes are preserved in the supplementary historical result files, not combined with the corrected final-pair definition. Diagonal variances are matched across residual models; the rank-one case tests one shared factor with learned loadings. The sweep motivated the declared, corrected rank-eight follow-up before its independent confirmation. It does not establish monotonic gains with larger rank.

Table 17: Historical EVAL, not new confirmation. Both orders share a bias-corrected prior, covariance and stop head. Work is in thousands of documents; feasible means TUNE risk at most 3.5%.

Condition	Prior	Order	Mean	P99	EVAL errors	Feasible
Semantic owners	Trees	Fixed	2863.6	3758.2	64/2048	yes
		Dynamic	2540.2	3757.6	64/2048	yes
	SupportNet	Fixed	2632.9	3758.3	52/2048	yes
		Dynamic	2792.6	3757.3	53/2048	no
Source owners	Trees	Fixed	854.5	866.2	51/2048	no
		Dynamic	790.0	865.7	69/2048	yes
	SupportNet	Fixed	775.0	866.1	63/2048	yes
		Dynamic	765.6	865.8	56/2048	yes
FiQA / BGE-M3	Trees	Fixed	28.4	30.4	42/1024	no
		Dynamic	28.9	30.4	39/1024	no
	SupportNet	Fixed	20.1	30.4	34/1024	yes
		Dynamic	29.2	30.4	34/1024	no

An earlier full-model holdout-stop comparison changed both stop-training population and, relative to another table, the fixed ranking's mean correction. It cannot isolate either change. The new equal-size, identically corrected factorial in Table 9 supersedes that interpretation, while preserving the old files and outcomes. This is a correction to attribution, not an alteration of a previously computed result.

L.3 Other attempted explanations

Fixed support prefixes followed by cheap-provider tails test a restricted class of cost-aware schedules, not all static strategies. In the expanded tree-model family, prefixes up to 32 fail the 3.5% TUNE guard; 48/64 pass but do not beat the fixed-order fallback chosen on TUNE. The new neural one-shot cost control in Table 7 is a separate and stronger check of the cost-prioritization explanation.

Earlier development also selected directly among providers affordable under $\min(B_0, C) - W$, preserving B_0 in the feature definition. Without retraining/retuning, FiQA dynamic work changed from 26,320 to 26,482, and errors from 42 to 41; fixed remained 25,714 with 40 errors. This admission repair did not eliminate the efficiency reversal. A separate refitted acquisition family selected fixed order for FiQA on TUNE. Those are tree-model diagnostics, not silent changes to the certified final neural policy.

Table 18: Executed neural training. TUNE RMSE chooses the checkpoint; all runs perform 100k steps. Fit min is recorded training-loop/selection time, not total preprocessing, augmentation-oracle or deployment cost.

Condition	Params M	Fit queries	30k	60k	100k	Selected	Fit min
Semantic owners	33.97	3072	0.0776	0.0692	0.0691	100000	53.9
Source owners	4.04	3072	0.0824	0.0722	0.0702	100000	15.2
FiQA / BGE-M3	4.09	1536	0.0311	0.0249	0.0236	100000	19.7

The three original final-backbone fits selected 100k steps by TUNE RMSE. Training time excludes embedding, partition construction, public-scorer fitting, augmentation-target computation and index construction. These common costs also serve static controls, so they cannot all be charged only to adaptation. GPU seconds, CPU seconds and document counts are different units; a monetary or amortized break-even claim requires an explicit price model, which these experiments do not provide.