LEARNING TO SHARE: QUERY-EFFICIENT RECONSTRUCTION OF LLM-SIMULATED POPULATION RESPONSES

Anonymous authors
Paper under double-blind review

ABSTRACT

Simulating how heterogeneous populations respond to alternative policies with large language models (LLMs) requires evaluating a costly Cartesian product of personas and interventions. We introduce the new zero-history task of *population-response reconstruction*: given n personas, m policies, and a budget of $B \ll nm$ teacher queries, the goal is to reconstruct the response distribution for every persona–policy pair. We construct a hierarchical benchmark of 1,024 World Values Survey personas and 600 COFOG-based policies—614,400 pairs per teacher—spanning domains, families, and implementation variants. Diagnostics across four teacher models yield two findings: response sharing is compact but policy-conditioned, and informative queries derive much of their value from improving unobserved pairs. We introduce SCOPE (Shared-response Calibration for Online Population Emulation), a closed-loop framework that calibrates response sharing across personas and policies from queried answers, then uses the calibrated structure to select queries for population-wide reconstruction. Across four teachers and all evaluated query budgets, SCOPE achieves the lowest three-seed mean full-pool error at every reconstruction level. Code, benchmark data, and complete results are available at `https://anonymous.4open.science/r/scope-F646/`.

1 INTRODUCTION

Understanding how heterogeneous people respond to alternative policy interventions is important for comparing designs before implementation. A policymaker may wish to evaluate m variants—for example, housing programs that differ in provider, eligibility, location, funding, or timing—for a population of n people with different socioeconomic circumstances and attitudes. The object of interest is an $n \times m$ response surface that preserves both population heterogeneity and substantive differences among interventions.

The development of large language models (LLMs) has made such simulations increasingly plausible. Prior work conditions LLMs on demographic profiles to emulate human survey samples (Argyle et al., 2023), tests whether LLMs can replicate human-subject experiments (Aher et al., 2023), and compares the opinions expressed by LLMs with those of demographic groups in public-opinion data (Santurkar et al., 2023). This literature primarily asks whether and when an LLM can faithfully represent human answers or opinions. Once a teacher model and prompting protocol have been chosen, however, generating the full population–intervention surface remains expensive: exhaustive simulation requires nm teacher calls.

We study this computational problem as a distinct task. Starting from a zero-history pool of n personas and m natural-language policies, an algorithm receives a budget $B \ll nm$ of LLM calls. Each call returns the teacher's five-way response distribution for one persona–policy pair. Policies are evaluated independently under a fixed persona profile, with the specified teacher's conditional response distribution as the target. The algorithm uses these B observations to train a student that reconstructs all nm teacher distributions. We call this setting *population-response reconstruction*. We shift focus from human-simulation fidelity to recovering a fixed teacher's complete response

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.

surface within a small call budget. For the 1,024 personas and 600 policies in our benchmark, exhaustive simulation would require 614,400 calls per teacher.

Our task is related to pool-based active learning, which selects inputs for labeling under a limited budget (Ash et al., 2020; 2021; Hacohen et al., 2022). It also relates to recent work on reducing the cost of LLM simulation and evaluation. APS queries representative agents and propagates their responses in population-scale agent simulation (Zheng et al., 2026). Its response-sharing mechanism is primarily one-dimensional: conditioned on the current simulation state and scenario, APS shares from queried prototype agents to nearby agents. Our reconstruction pool is intrinsically two-dimensional. Both personas and interventions vary, so one observation may transfer across similar people, across related policies, or through recurring persona–policy interactions. Query selection must therefore identify which cell of the $n \times m$ surface is most informative for the entire surface, extending agent-axis selection to both personas and policies. Active Evaluation Acquisition predicts missing benchmark results from partial evaluations (Li et al., 2024). The latter learns from historical model–benchmark evaluation matrices that are available before acquisition. Our student starts from a zero-history pool for the current teacher, personas, and policies. This assumption reflects a common policy use case: a decision maker has just formulated m new policy variants and wants to estimate the responses of n target personas, creating a fresh task-specific response pool. Observations lie on a two-axis Cartesian pool, and their value depends on transfer across both personas and policies. A more detailed comparison with LLM-based human simulation, efficient model evaluation, active learning, and structured completion is provided in Appendix A.

This setting creates two coupled bottlenecks. The first is *response sharing*: given a sparse set of observed persona–policy responses, how can a student identify which patterns transfer to unobserved pairs? Persona attributes and policy text provide useful side information, but their geometric similarity can differ from the teacher's behavioral similarity. The second is *query allocation*: given the current student, which unobserved pairs should receive the remaining teacher calls so that the resulting answers are maximally useful for learning the full surface? The two problems depend on each other. Query selection requires an estimate of where responses can be shared, while learning that sharing structure requires informative queries. Effective allocation must balance local uncertainty, input coverage, and pool-wide transfer. Figure 1 summarizes the task and these two bottlenecks.

Offline diagnostics yield two findings that map directly to the bottlenecks. **Finding 1: response sharing is compact but policy-conditioned.** Persona and policy main effects, together with concentrated interaction directions, explain most response variation. A small MLP reaches 0.0311 TV under dense supervision, while hidden-state neighborhoods align closely with responses but change across policies. These results motivate multi-scale persona and policy sharing, explicit interactions, and paid calibration of their behavioral relevance. **Finding 2: valuable queries improve unobserved pairs.** Oracle acquisition analyses show that most gains from informative queries occur on other cells in the response surface. Query value should therefore combine exact imputation with pool-wide propagation through the current sharing structure.

These findings directly motivate SCOPE (**S**hared-response **C**alibration for **O**nline **P**opulation **E**mulation), which addresses the two bottlenecks in a closed loop. For response sharing, SCOPE fits candidate kernels containing persona effects, policy effects, and persona–policy interactions, then uses leave-one-out Brier error on the queried responses to calibrate their scales, interaction strength, and regularization. For query allocation, the calibrated mixture defines a response-covariance proxy that scores both the exact imputation value of a pair and its predicted propagation value to all remaining pairs. Each new batch of teacher responses recalibrates the sharing structure, which in turn changes the next acquisition decision.

To evaluate this task, we construct a hierarchical benchmark with 1,024 real WVS respondent profiles (Haerpfer et al., 2022) and 600 policy interventions. We systematically design the policy set around the 10 functional domains in the United Nations COFOG classification (United Nations Statistics Division, 2000), which provides a standardized and reproducible basis for broad policy-domain coverage. Within each domain, we define 10 policy families and six substantive implementation variants per family. This hierarchy yields family-level (1024×6), domain-level ($1024 \times 10 \times 6$), and global ($1024 \times 10 \times 10 \times 6$) reconstruction tasks. The benchmark specifies four teacher checkpoints—Qwen3.5-4B, Qwen2.5-3B-Instruct, Gemma-4-E4B-it, and Qwen3.8-27B—and evaluates exact-imputation full-pool TV under different query budgets;. Comprehensive exper-

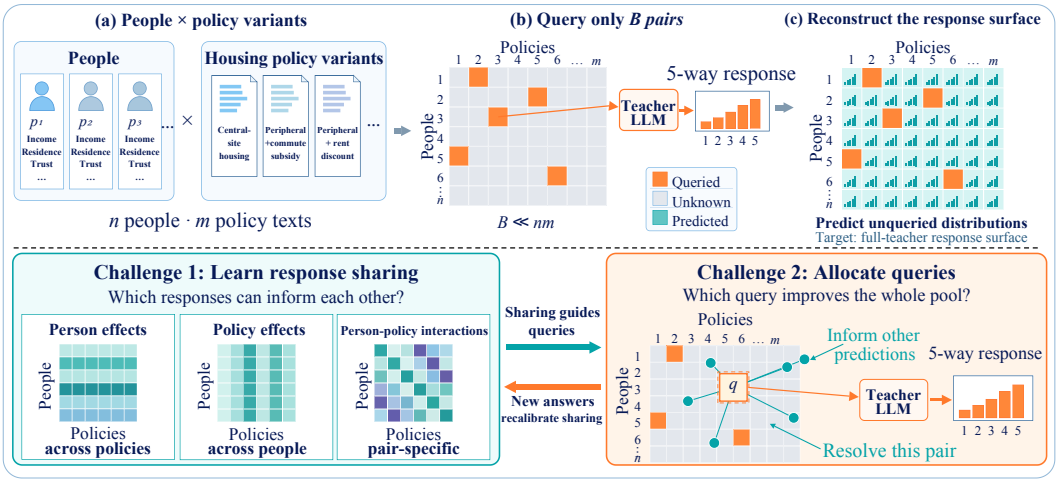

Figure 1: Task and coupled bottlenecks. The system reconstructs nm response distributions using B persona–policy queries. It must identify response sharing supported by persona effects, policy effects, and interactions, then allocate observations according to their value for the full pool. The current sharing structure guides acquisition, and each new teacher answer recalibrates subsequent prediction and querying.

iments demonstrate the effectiveness of SCOPE across teacher models, query budgets, and reconstruction levels.

Our contributions are fourfold. (1) We introduce the zero-history, budget-constrained population-response reconstruction task and a hierarchical persona–policy benchmark. (2) We identify shared main effects, concentrated interaction directions, policy-conditioned response neighborhoods, and cross-pair query value, providing empirical requirements for low-budget reconstruction. (3) We propose SCOPE, a closed-loop method that calibrates response sharing from queried answers and uses the same structure for pool-wide active acquisition. (4) We conduct comprehensive experiments across reconstruction levels, query budgets, and teacher models, demonstrating the effectiveness of the proposed method across all three dimensions.

2 TASK, BENCHMARK, AND PRELIMINARY FINDINGS

2.1 TASK DEFINITION AND INFORMATION BOUNDARY

Let $\mathcal{P} = \{p_i\}_{i=1}^{n}$ be a set of personas and $\mathcal{S} = \{s_j\}_{j=1}^{m}$ a set of policies. The known pool is

$$\mathcal{X} = \mathcal{P} \times \mathcal{S}, \qquad N = nm. \tag{1}$$

For a fixed teacher and prompting protocol, each pair has a response

$$y_{ij} = F_\theta(p_i, s_j) \in \Delta^{C-1}, \qquad C = 5, \tag{2}$$

over acceptance scores 1–5. An algorithm observes at most B responses, $\mathcal{Q}_B \subseteq \mathcal{X}$, and reconstructs the full pool using exact answers at queried cells and student predictions elsewhere:

$$\widetilde{y}_x = \begin{cases} y_x, & x \in \mathcal{Q}_B, \\ \widehat{y}_x, & x \notin \mathcal{Q}_B. \end{cases} \tag{3}$$

Our primary objective is finite-pool individual response error,

$$R_{\text{all}} = \frac{1}{N} \sum_{x \in \mathcal{X}} \text{TV}(y_x, \widetilde{y}_x), \qquad \text{TV}(p, q) = \frac{1}{2} \sum_{c=1}^{C} |p_c - q_c|. \tag{4}$$

We also report unqueried-pair error R_U; for B distinct queries, $R_{\text{all}} = (N - B)R_U/N$.

All persona profiles, policy texts, and public representations are available before acquisition. Every response used for initialization, fitting, calibration, or active selection counts toward the same budget B. Policy-only encoding is accounted for separately. Full response caches support offline diagnostics and evaluation; deployed components use public inputs and budgeted responses.

2.2 HIERARCHICAL BENCHMARK

We sample 256 whole respondent records from each of South Korea, Germany, Canada, and Zimbabwe in WVS Wave 7 (Haerpfer et al., 2022), preserving the joint configuration of 15 fields and representing unavailable values as unknown. The ten official COFOG divisions (United Nations Statistics Division, 2000) provide the coverage scaffold; within each, we define ten study-specific families and six standalone implementation alternatives. The resulting 600 policies yield 100 family tasks (6,144 pairs each), ten domain tasks (61,440), and one global task (614,400), covering progressively broader sharing ranges.

We use Qwen3.5-4B (Qwen Team, 2026a), Qwen2.5-3B-Instruct (Yang et al., 2024), Gemma-4-E4B-it (Gemma Team, 2026), and Qwen3.8-27B (Qwen Team, 2026b). For each independent hypothetical policy, the prompt asks the persona to return an acceptability score from 1 to 5 while treating unspecified attributes as unknown. We normalize the five answer-token logits, producing four complete $1024 \times 600 \times 5$ response tensors and 2,457,600 target distributions. Appendix B documents deterministic persona construction, the complete family taxonomy, policy quality checks, example variants, exact prompts, and response extraction.

2.3 FINDING 1: RESPONSE SHARING IS COMPACT BUT POLICY-CONDITIONED

We apply the centered log-ratio (CLR) transform for compositional data (Aitchison, 1982) and decompose each response as

$$z_{ij} = \mathrm{clr}(y_{ij}) = \mu + a_i + b_j + R_{ij}, \tag{5}$$

where a_i, b_j, and R_{ij} are centered persona, policy, and interaction terms. Across four teachers, the main effects explain 85.79–94.54% of CLR variation. Interactions contribute 5.46–14.21% but are concentrated: their top eight directions retain 75.27–87.02% of interaction energy, and their 95%-energy ranks are 34–183 versus 528–585 after persona permutation. Qwen3.8 is representative: 59.05% persona, 26.74% policy, 14.21% interaction, $r_{95} = 42$, and 81.75% top-eight energy.

With abundant labels, a 26,373-parameter MLP attains 0.0311 held-out TV on about 492,000 Qwen3.5 pairs, while an oracle additive-plus-rank-16 model reaches 0.01790. These results establish sufficient student capacity and locate the low-budget difficulty in identifying sharing. Public similarity is also imperfect: in one housing family, hidden-state neighbors have response TV 0.04461 versus 0.14440 for attribute neighbors, and 48.61% persist across policy variants. Final-layer probes also outperform attribute probes for all three tested teachers. These results motivate multiple sharing scales, explicit interactions, and paid calibration of their behavioral relevance.

2.4 FINDING 2: VALUABLE QUERIES IMPROVE UNOBSERVED PAIRS

For a fixed student on a 512-pair diagnostic panel, oracle query selection reduces TV from 0.10653 to 0.08669 at budget 64. At budgets 64 and 128, 75–80.7% of the gain occurs on pairs other than those queried. Thus, selecting a locally uncertain or atypical cell is insufficient: acquisition should value how an answer propagates through the current sharing structure.

Together, the diagnostics require a learner that infers sharing from paid responses and values a query by its effect beyond the queried cell. The next section instantiates these requirements in a closed loop.

3 METHOD

SCOPE alternates between calibrating how responses are shared and selecting queries using that calibrated structure. Given the current paid set, it fits a mixture of persona, policy, and interaction kernels, converts the mixture into a response-covariance proxy, and selects the next batch by its estimated value for the entire unknown pool. Figure 2 summarizes this loop.

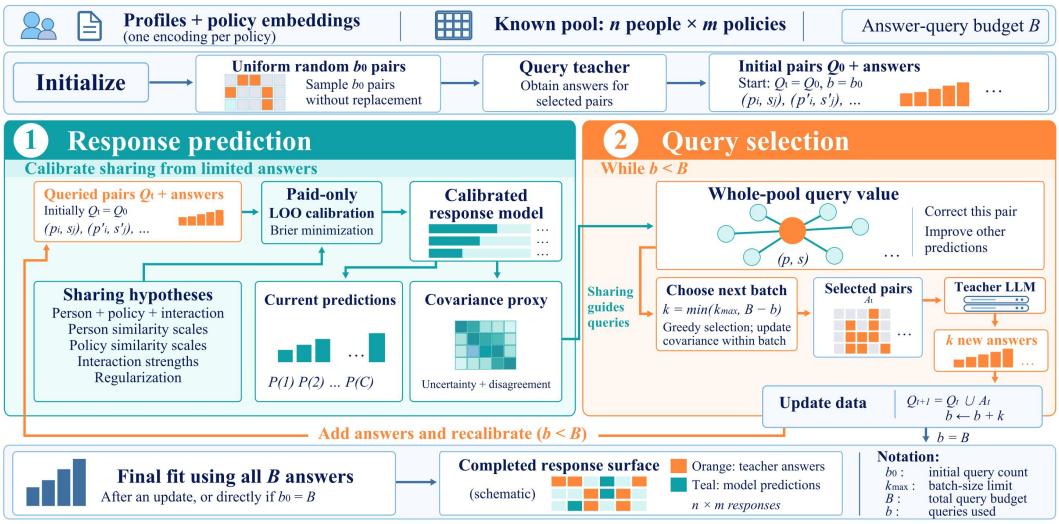

Figure 2: The SCOPE loop. An initial uniform sample provides b_0 paid answers. Leave-one-out Brier calibration produces predictions and a response covariance proxy. The proxy selects a batch of at most $k_{\max}$ queries by pool-wide value while discouraging within-batch redundancy. New answers update the fitted sharing structure. At budget B, queried cells receive exact teacher responses and all remaining cells receive final predictions.

3.1 RESPONSE PREDICTION BY CALIBRATING SHARING

Let v_i be the 74-dimensional persona features and e_j the policy features. Each policy is encoded once by the teacher and reused across all personas:

$$e_j = \mathrm{PCA}_d\big(h_\theta^{\mathrm{last}}[T(p_\varnothing, s_j)]\big)\,, \tag{6}$$

where $p_\varnothing$ is a fixed blank profile and h_θ^{last} is the final-normalized hidden state at the last input position. We retain at most 32 dimensions and zero-pad when needed. PCA (Jolliffe & Cadima, 2016) and block normalization are fitted only from public inputs. This policy-only representation does not reveal a persona–policy response, and its encoding cost is reported separately from the budget B.

For candidate $h \in \{1, \ldots, H\}$, let $K_{P,h}$ and $K_{S,h}$ be centered RBF kernels on persona and policy features, and write $x_{ij} = (p_i, s_j)$. Its joint sharing kernel is

$$K_h(x_{ij}, x_{i'j'}) = \frac{\frac{1}{2}K_{P,h}(i,i') + \frac{1}{2}K_{S,h}(j,j') + \eta_h K_{P,h}(i,i')K_{S,h}(j,j')}{1+\eta_h}. \tag{7}$$

The three terms represent persona main effects, policy main effects, and their interaction. Each public-axis kernel is centered and normalized to unit mean diagonal. The symmetric $1/2$ coefficients are a prior; observed main-effect magnitudes are determined by the data and fitted regression coefficients. Candidates vary the two length scales, the interaction strength, and the ridge regularizer (Appendix C.2).

At round t, the queried set is Q_t and its CLR-transformed teacher responses are Z_t. Each candidate fits kernel ridge regression with an intercept:

$$\widehat{z}_h(x) = \mu_h + k_h(x, Q_t)\alpha_h, qquad\widehat{p}_h(x) = \mathrm{softmax}(\widehat{z}_h(x)). \tag{8}$$

For every queried point q, we compute the analytic prediction $\widehat{p}_h^{(-q)}(q)$ after leaving out its answer. We then learn mixture weights on the probability simplex:

$$\pi_t^* = \arg\min_{\pi \in \Delta^{H-1}} \frac{1}{|Q_t|} \sum_{q \in Q_t} \left\| y_q - \sum_{h=1}^{H} \pi_h \widehat{p}_h^{(-q)}(q) \right\|_2^2. \tag{9}$$

This paid Brier calibration (Brier, 1950) is a convex quadratic problem on the simplex; we solve it with SLSQP (Kraft, 1988). The final probability prediction is

$$\widehat{y}_t(x) = \sum_{h=1}^{H} \pi^*_{t,h} \widehat{p}_h(x). \tag{10}$$

Hence, the answers determine whether broad or local persona sharing, broad or local policy sharing, and stronger or weaker interaction transfer are supported. The squared-probability objective also relates directly to TV because $\mathrm{TV}(p, q) \le \sqrt{C}\|p-q\|_2/2$. On the empirical leave-one-out objective, the exact mixture achieves error at most that of the best single candidate; generalization to unqueried pairs is evaluated empirically.

3.2 Acquisition by pool-wide response value

A teacher query exactly fills its own cell and constrains other cells that share response structure with it. We therefore reuse the calibrated predictor to estimate both effects.

Let $c_h(x, x')$ be the conditional covariance of candidate h with an intercept, and let $S_h \succeq 0$ be an output scale estimated from paid leave-one-out residuals. The softmax Jacobian is

$$J_h(x) = \mathrm{diag}(\widehat{p}_h(x)) - \widehat{p}_h(x)\widehat{p}_h(x)^\top. \tag{11}$$

To capture disagreement across sharing assumptions, define $d_h(x) = \widehat{p}_h(x) - \widehat{y}_t(x)$. Our scalar response-covariance proxy is

$$C_t(x, x') = \sum_h \pi^*_{t,h} \left[\frac{c_h(x, x')}{C} \mathrm{tr}\big(J_h(x) S_h J_h(x')^\top\big) + \frac{\langle d_h(x), d_h(x')\rangle}{C} \right]. \tag{12}$$

The first term measures within-structure uncertainty; the second measures between-structure predictive disagreement. Both use the same paid responses and mixture weights as the predictor. The construction is positive semidefinite, which supports coherent conditional covariance updates (Appendix E).

For unknown set U_t and candidate $q \in U_t$, define

$$V_t(q) = C_t(q, q) + \sum_{u \in U_t \setminus \{q\}} \frac{C_t(u, q)^2}{C_t(q, q) + \tau_t(q)}, \tag{13}$$

where $\tau_t(q) > 0$ is a conservative update noise derived from candidate regularization and output scales. The first term is the value of exactly imputing the queried pair. The second is the predicted reduction for all other unknown pairs. With the current covariance frozen, Eq. equation 13 equals the total scalar Gaussian-proxy risk reduction from removing the queried point and reducing conditional variance elsewhere.

We uniformly sample at most 512 unknown targets to approximate the propagation sum, while still considering every allowable pair as a query candidate. After each selection within a batch, a conditional covariance update discourages redundant selections. Once the teacher returns the batch, all sharing structures and mixture weights are refitted.

The complete loop is

$$(Q_t, Y_{Q_t}) \xrightarrow{\text{calibrate sharing}} (\widehat{y}_t, C_t) \xrightarrow{\text{pool-wide value}} A_t \xrightarrow{\text{query teacher}} (Q_{t+1}, Y_{Q_{t+1}}). \tag{14}$$

All methods share 16 uniformly sampled initial queries and add at most 16 queries per batch, i.e., $b_0 = k_{\max} = 16$. The final batch is truncated to the remaining budget. For example, a family task with 1,024 personas, six policies, and $B = 64$ begins with 16 queries and performs three active batches of 16. The 64 queried cells receive exact teacher responses; the other 6,080 cells use the final student prediction.

Table 1: Family-level full-pool TV (lower is better), aggregated over 100 families. Entries are mean ± sample SD over three paired seeds. Parentheses give the queried fraction; Gain is relative to the best baseline mean.

Teacher	B (pool)	R–MLP	R–XGB	BAIT	TypiClust	SCOPE	Gain
Qwen3.5	16 (0.26%)	0.1241±.0022	0.1538±.0032	0.1241±.0022	0.1241±.0022	**0.1109±.0030**	10.6%
	32 (0.52%)	0.0965±.0064	0.1086±.0050	0.1028±.0024	0.0954±.0030	**0.0809±.0012**	15.2%
	64 (1.04%)	0.0727±.0011	0.0790±.0007	0.0809±.0013	0.0755±.0016	**0.0592±.0006**	18.6%
Qwen2.5	16 (0.26%)	0.2435±.0034	0.2982±.0123	0.2435±.0034	0.2435±.0034	**0.2375±.0044**	2.5%
	32 (0.52%)	0.1989±.0083	0.2212±.0082	0.2085±.0019	0.2028±.0068	**0.1791±.0115**	10.0%
	64 (1.04%)	0.1568±.0063	0.1620±.0083	0.1707±.0004	0.1706±.0037	**0.1289±.0031**	17.8%
Gemma	16 (0.26%)	0.2792±.0196	0.3042±.0232	0.2792±.0196	0.2792±.0196	**0.2791±.0133**	0.04%
	32 (0.52%)	0.2290±.0174	0.2233±.0254	0.2410±.0123	0.2388±.0086	**0.1995±.0016**	10.6%
	64 (1.04%)	0.1753±.0167	0.1564±.0033	0.1917±.0033	0.1782±.0034	**0.1363±.0011**	12.9%
Qwen3.8	16 (0.26%)	0.2297±.0068	0.2582±.0084	0.2297±.0068	0.2297±.0068	**0.2205±.0097**	4.0%
	32 (0.52%)	0.1917±.0052	0.1818±.0103	0.1984±.0048	0.1956±.0020	**0.1778±.0031**	2.2%
	64 (1.04%)	0.1520±.0056	0.1390±.0028	0.1627±.0033	0.1603±.0013	**0.1275±.0004**	8.2%

4 EXPERIMENTS

We evaluate (i) response reconstruction across teachers and sharing ranges, (ii) the separate and joint contributions of prediction and acquisition, and (iii) the relation between calibrated sharing and recovery of the output structure.

4.1 PROTOCOL, BASELINES, AND METRICS

Each family and each domain is an independent task with a separate budget and predictor. Within a teacher and reconstruction level, all tasks together cover the same 614,400 persona–policy pairs. Main results use three paired seeds (719, 720, and 721), with the same 16 initial queries across methods within each task and seed. We first pool all tasks within a teacher, level, method, budget, and seed, then report the mean and sample standard deviation across seeds. The main comparison covers 19,860 task-level budget checkpoints. Thus, methods differ only in prediction and subsequent acquisition. The Qwen3.8 pipeline was first checked on three development domains before expansion; all reported domain aggregates use the complete set of ten domains.

All methods use the same persona and policy features. We compare random-query MLP and XGBoost predictors, Fisher-information acquisition with BAIT–MLP (Ash et al., 2021), coverage-based acquisition with TypiClust–MLP (Hacohen et al., 2022), and SCOPE's jointly calibrated prediction and pool-wide acquisition. Random–MLP serves as the primary reference for separating gains from response sharing and query allocation. Architectures, training settings, feature processing, and adaptations to soft response distributions are detailed in Appendix C.1. Tables abbreviate Random–MLP and Random–XGBoost as R–MLP and R–XGB; "Gain" is $100(R_{\text{best}} - R_{\text{SCOPE}})/R_{\text{best}}$, where R_{best} is the lowest error among the non-SCOPE methods.

The primary metric is exact-imputation full-pool TV, R_{all}, pooled over all tasks of a teacher and reconstruction level. Appendix F also reports unqueried TV, Jensen–Shannon divergence (Lin, 1991), and expected-score MAE. Pointwise TV additionally upper-bounds error in any fixed weighted population average.

4.2 RECONSTRUCTION ACROSS TEACHERS AND HIERARCHY LEVELS

Policy families. Table 1 pools all 100 independent families and all 614,400 pairs for each teacher. SCOPE has the lowest mean TV in all 12 teacher–budget conditions. At the largest budget, its relative reduction against the strongest baseline ranges from 8.2% to 18.6%. The single-seed Qwen3.8 exception in the earlier draft reverses after replication: at $B = 32$, SCOPE improves on Random–XGBoost by 2.2% in mean TV.

Policy domains. Table 2 pools all ten domains. SCOPE achieves the lowest mean TV in all 12 conditions and improves over the strongest baseline mean by 4.3–19.2%. Together with the family results, this yields the lowest mean in all 24 teacher–budget conditions.

Table 2: Domain-level full-pool TV (lower is better), aggregated over ten domains. Entries are mean ± sample SD over three paired seeds. Parentheses give the queried fraction; Gain is relative to the best baseline mean.

Teacher	B (pool)	R–MLP	R–XGB	BAIT	TypiClust	SCOPE	Gain
Qwen3.5	160 (0.26%)	0.0783±.0015	0.0863±.0016	0.0905±.0016	0.0886±.0012	**0.0632±.0006**	19.2%
	320 (0.52%)	0.0626±.0006	0.0714±.0015	0.0750±.0022	0.0762±.0003	**0.0517±.0003**	17.4%
	640 (1.04%)	0.0502±.0003	0.0623±.0009	0.0616±.0011	0.0675±.0005	**0.0436±.0001**	13.2%
Qwen2.5	160 (0.26%)	0.1505±.0015	0.1640±.0025	0.1656±.0009	0.1624±.0007	**0.1249±.0015**	17.0%
	320 (0.52%)	0.1176±.0018	0.1356±.0010	0.1372±.0013	0.1371±.0004	**0.1022±.0005**	13.1%
	640 (1.04%)	0.0991±.0013	0.1175±.0028	0.1160±.0005	0.1245±.0006	**0.0856±.0002**	13.6%
Gemma	160 (0.26%)	0.1749±.0009	0.1733±.0047	0.1955±.0023	0.1983±.0024	**0.1520±.0016**	12.3%
	320 (0.52%)	0.1405±.0035	0.1469±.0035	0.1628±.0034	0.1690±.0031	**0.1207±.0018**	14.1%
	640 (1.04%)	0.1114±.0004	0.1219±.0025	0.1347±.0036	0.1535±.0011	**0.0993±.0006**	10.8%
Qwen3.8	160 (0.26%)	0.1535±.0004	0.1483±.0026	0.1667±.0023	0.1651±.0005	**0.1342±.0009**	9.5%
	320 (0.52%)	0.1229±.0013	0.1291±.0022	0.1389±.0011	0.1432±.0005	**0.1130±.0004**	8.1%
	640 (1.04%)	0.1001±.0014	0.1100±.0017	0.1145±.0009	0.1308±.0022	**0.0959±.0003**	4.3%

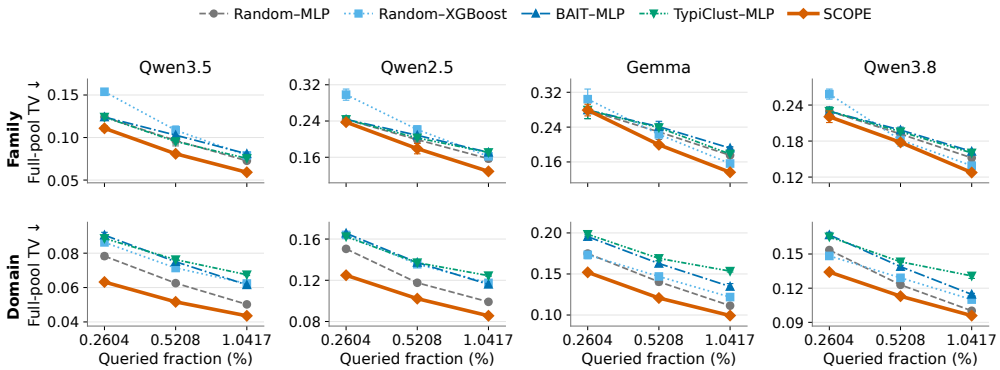

Figure 3: Budget–TV curves across four teachers (columns), with family and domain tasks in the top and bottom rows. Values are means across seeds 719, 720, and 721 after task averaging (100 families or ten domains per seed); bars show one sample SD. Queried fraction uses a log scale, vertical ranges vary by panel, and lower full-pool TV is better.

At matched queried fractions, domain errors are lower than family errors for all four teachers, consistent with the benefit of sharing across a broader set of related policies; domain tasks also receive more queries in absolute terms.

Global pool. The global task places all 1,024 personas and 600 policies into one response surface with 640 queries (0.104%). It jointly stresses cross-domain sharing and low query density. Table 3 shows that SCOPE is best for all four teachers, improving over the strongest baseline by 11.3–18.3% while querying only 0.104% of the pool.

Table 3: Global full-pool TV at $B = 640$ (0.104% of the pool). Gain is SCOPE's relative reduction versus the best baseline in each row.

Teacher	R–MLP	R–XGB	BAIT	TypiClust	SCOPE	Gain
Qwen3.5	0.0849	0.0882	0.0886	0.0952	**0.0694**	18.3%
Qwen2.5	0.1438	0.1587	0.1577	0.1733	**0.1261**	12.3%
Gemma	0.1886	0.1832	0.2091	0.2082	**0.1600**	12.7%
Qwen3.8	0.1580	0.1490	0.1565	0.1648	**0.1322**	11.3%

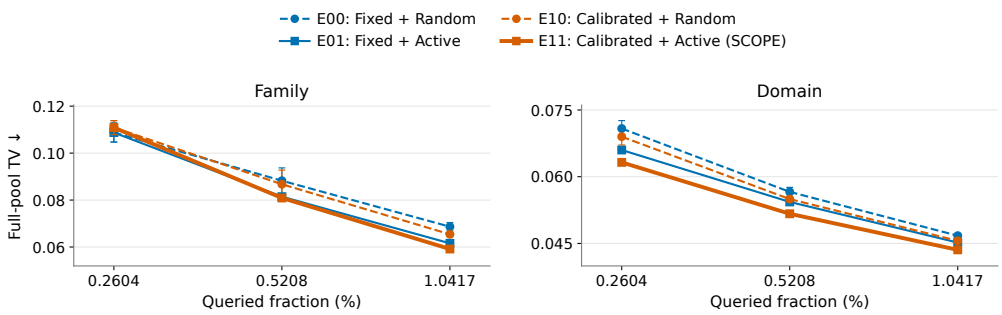

Figure 4: Qwen3.5-4B two-bottleneck factorial ablation at family (left) and domain (right) levels. Blue/orange denote fixed/calibrated sharing and dashed circles/solid squares denote random/active acquisition; E00–E11 map, in order, to fixed–random, fixed–active, calibrated–random, and calibrated–active (full SCOPE). Values are three-seed means after task averaging; bars show one sample SD. Family $B = 16$ is initialization only. Queried fraction uses a log scale; lower full-pool TV is better.

4.3 Separating the two bottlenecks

We cross fixed versus response-calibrated sharing with random versus pool-value acquisition, isolating the prediction and query-allocation bottlenecks. The fixed model uses persona and policy scales of 1, interaction coefficient 0.1, and ridge parameter 0.03; active acquisition uses the value in equation 13 with the covariance supplied by its predictor.

Beyond initialization, calibration alone and pool-value acquisition alone improve over Fixed/R in all five active-budget conditions, and their combination has the lowest mean in every case. At family $B = 64$, the joint system lowers mean TV from 0.0687 to 0.0592, a 13.8% reduction. Family $B = 16$ contains only the shared initial queries and therefore does not test acquisition; the fixed predictor is slightly better there. The interaction contrast is positive for the three domain budgets but negative for the two active family budgets, so the ablation supports complementary contributions without a general claim of superadditivity. Complete values are reported in Appendix F.1.

4.4 Recovery of response structure and mixture calibration

The gains are not explained only by exact answers at queried cells: on a common fully unqueried rectangle, SCOPE recovers interaction variation more accurately than Random–MLP at the largest domain budget for all three teachers with available diagnostics. Random–MLP has higher response-geometry CKA, showing that global geometric alignment need not imply lower pointwise error. On identical Qwen3.5 query prefixes, the Brier-calibrated mixture also consistently improves over uniform averaging and the best single candidate. Appendix F.2 gives definitions and complete results.

5 Conclusion

We introduced zero-history population-response reconstruction from a fixed persona–policy pool under a limited response budget. For the four teachers included in the structural diagnostics, the full response surfaces reveal strong main effects and concentrated interaction structure; hidden-neighborhood and oracle-query diagnostics connect the identification of this sharing structure to the allocation of observations.

SCOPE calibrates response sharing from paid answers and uses the resulting covariance to estimate each new query's value for the entire pool. Prediction and acquisition are therefore updated jointly after every batch. Across four teachers and three paired seeds, SCOPE has the lowest mean error in all 24 family- and domain-level conditions and all four global conditions. The Qwen3.5 factorial study further shows that the joint system has the lowest mean in every evaluated active-budget condition. The central lesson is that, under few queries, each teacher answer should serve both the current prediction and the identification of how future answers can be shared.

Ethics Statement

This work studies the conditional response distributions of specified LLM teachers rather than estimating human beliefs. Persona profiles are derived from de-identified World Values Survey records, and no new human participants were recruited or contacted. The benchmark uses survey attributes and derived artifacts in accordance with the applicable WVS terms. Simulated responses may inherit biases or demographic stereotypes from the teacher models and should not substitute for representative surveys or support high-stakes decisions. We report the persona construction, prompting protocol, and teacher dependence to make these limitations explicit.

Reproducibility Statement

The complete code, benchmark data, experiment configurations, and result artifacts are available in the anonymous repository `https://anonymous.4open.science/r/scope-F646/`. The repository records the persona and policy construction, teacher and inference settings, public representations, random seeds, query traces, predictions, and scripts used to reproduce the analyses, tables, and figures. Appendices B– F provide the corresponding implementation details, candidate grid, analytic calculations, acquisition approximation, and supplementary metrics.

AI Use Statement

Generative AI tools were used solely to polish the manuscript's language and improve readability. The authors performed the research design, method and benchmark development, mathematical derivations, implementation, experiments, and interpretation of results. All AI-assisted edits were reviewed by the authors, who take responsibility for the final content of the paper.

References

Gati V Aher, Rosa I. Arriaga, and Adam Tauman Kalai. Using large language models to simulate multiple humans and replicate human subject studies. In *Proceedings of the 40th International Conference on Machine Learning*, volume 202 of *Proceedings of Machine Learning Research*, pp. 337–371. PMLR, 2023. URL `https://proceedings.mlr.press/v202/aher23a.html`.

John Aitchison. The statistical analysis of compositional data. *Journal of the Royal Statistical Society: Series B (Methodological)*, 44(2):139–177, 1982. doi: 10.1111/j.2517-6161.1982.tb01195.x.

Lisa P. Argyle, Ethan C. Busby, Nancy Fulda, Joshua R. Gubler, Christopher Rytting, and David Wingate. Out of one, many: Using language models to simulate human samples. *Political Analysis*, 31(3):337–351, 2023. doi: 10.1017/pan.2023.2.

Jordan T. Ash, Chicheng Zhang, Akshay Krishnamurthy, John Langford, and Alekh Agarwal. Deep batch active learning by diverse, uncertain gradient lower bounds. In *International Conference on Learning Representations*, 2020. URL `https://openreview.net/forum?id=ryghZJBKPS`.

Jordan T. Ash, Surbhi Goel, Akshay Krishnamurthy, and Sham Kakade. Gone fishing: Neural active learning with fisher embeddings. In *Advances in Neural Information Processing Systems*, volume 34, 2021. URL `https://proceedings.neurips.cc/paper/2021/hash/4afe044911ed2c247005912512ace23b-Abstract.html`.

James Bisbee, Joshua D. Clinton, Cassy Dorff, Brenton Kenkel, and Jennifer M. Larson. Synthetic replacements for human survey data? the perils of large language models. *Political Analysis*, 32 (4):401–416, 2024. doi: 10.1017/pan.2024.5.

Edwin V. Bonilla, Kian Ming A. Chai, and Christopher K. I. Williams. Multi-task gaussian process prediction. In *Advances in Neural Information Processing Systems*, volume 20, pp. 153–160, 2007. URL `https://proceedings.neurips.cc/paper/2007/hash/66368270ffd51418ec58bd793f2d9b1b-Abstract.html`.

Glenn W. Brier. Verification of forecasts expressed in terms of probability. *Monthly Weather Review*, 78(1):1–3, 1950. doi: 10.1175/1520-0493(1950)078⟨0001:VOFEIT⟩2.0.CO;2.

Tianqi Chen and Carlos Guestrin. XGBoost: A scalable tree boosting system. In *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, pp. 785–794, 2016. doi: 10.1145/2939672.2939785.

David A. Cohn, Zoubin Ghahramani, and Michael I. Jordan. Active learning with statistical models. *Journal of Artificial Intelligence Research*, 4:129–145, 1996. doi: 10.1613/jair.295.

Gemma Team. Gemma 4 technical report, 2026. URL https://arxiv.org/abs/2607.02770.

Mehmet Gönen and Ethem Alpaydın. Multiple kernel learning algorithms. *Journal of Machine Learning Research*, 12(64):2211–2268, 2011. URL https://jmlr.org/papers/v12/gonen11a.html.

Haoxiang Guan, Jiyan He, Liyang Fan, Zhenzhen Ren, Shaobin He, Xin Yu, Yuan Chen, Xueyin Xu, Shuxin Zheng, Yan Gao, Enhong Chen, Tie-Yan Liu, and Zhen Liu. Modeling earth-scale human-like societies with one billion agents, 2025. URL https://arxiv.org/abs/2506.12078.

Guy Hacohen, Avihu Dekel, and Daphna Weinshall. Active learning on a budget: Opposite strategies suit high and low budgets. In *Proceedings of the 39th International Conference on Machine Learning*, volume 162 of *Proceedings of Machine Learning Research*, pp. 8175–8195. PMLR, 2022. URL https://proceedings.mlr.press/v162/hacohen22a.html.

Christian Haerpfer, Ronald Inglehart, Alejandro Moreno, Christian Welzel, Kseniya Kizilova, Jaime Diez-Medrano, Marta Lagos, Pippa Norris, Eduard Ponarin, and Bi Puranen. World values survey wave 7 (2017–2022) cross-national data-set. World Values Survey Association, 2022. Version 4.0.0; doi:10.14281/18241.18.

Dan Hendrycks and Kevin Gimpel. Gaussian error linear units (GELUs), 2016. URL https://arxiv.org/abs/1606.08415.

Geoffrey Hinton, Oriol Vinyals, and Jeff Dean. Distilling the knowledge in a neural network, 2015. URL https://arxiv.org/abs/1503.02531.

John J. Horton, Apostolos Filippas, and Benjamin S. Manning. Large language models as simulated economic agents: What can we learn from homo silicus? Working Paper 31122, National Bureau of Economic Research, April 2023. URL https://www.nber.org/papers/w31122. Revised February 2026.

Tiancheng Hu and Nigel Collier. Quantifying the persona effect in LLM simulations. In *Proceedings of the 62nd Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)*, pp. 10289–10307. Association for Computational Linguistics, 2024. doi: 10.18653/v1/2024.acl-long.554. URL https://aclanthology.org/2024.acl-long.554/.

Henrik Imberg, Johan Jonasson, and Marina Axelson-Fisk. Optimal sampling in unbiased active learning. In *Proceedings of the Twenty Third International Conference on Artificial Intelligence and Statistics*, volume 108 of *Proceedings of Machine Learning Research*, pp. 559–569. PMLR, 2020. URL https://proceedings.mlr.press/v108/imberg20a.html.

Prateek Jain and Inderjit S. Dhillon. Provable inductive matrix completion, 2013. URL https://arxiv.org/abs/1306.0626.

Ian T. Jolliffe and Jorge Cadima. Principal component analysis: A review and recent developments. *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences*, 374(2065):20150202, 2016. doi: 10.1098/rsta.2015.0202.

Simon Kornblith, Mohammad Norouzi, Honglak Lee, and Geoffrey Hinton. Similarity of neural network representations revisited. In *Proceedings of the 36th International Conference on Machine Learning*, volume 97 of *Proceedings of Machine Learning Research*, pp. 3519–3529. PMLR, 2019. URL https://proceedings.mlr.press/v97/kornblith19a.html.

Dieter Kraft. A software package for sequential quadratic programming. Technical Report DFVLR-FB 88-28, German Aerospace Center, Institute for Flight Mechanics, 1988.

Andreas Krause, Ajit Singh, and Carlos Guestrin. Near-optimal sensor placements in gaussian processes: Theory, efficient algorithms and empirical studies. *Journal of Machine Learning Research*, 9(8):235–284, 2008. URL https://jmlr.org/papers/v9/krause08a.html.

Akshay Krishnamurthy and Aarti Singh. Low-rank matrix and tensor completion via adaptive sampling. In *Advances in Neural Information Processing Systems*, volume 26, pp. 836–844, 2013. URL https://proceedings.neurips.cc/paper/2013/hash/2050e03ca119580f74cca14cc6e97462-Abstract.html.

Yang Li, Jie Ma, Miguel Ballesteros, Yassine Benajiba, and Graham Horwood. Active evaluation acquisition for efficient LLM benchmarking, 2024. URL https://arxiv.org/abs/2410.05952.

Jianhua Lin. Divergence measures based on the shannon entropy. *IEEE Transactions on Information Theory*, 37(1):145–151, 1991. doi: 10.1109/18.61115.

Ilya Loshchilov and Frank Hutter. Decoupled weight decay regularization. In *International Conference on Learning Representations*, 2019. URL https://arxiv.org/abs/1711.05101.

Felipe Maia Polo, Lucas Weber, Leshem Choshen, Yuekai Sun, Gongjun Xu, and Mikhail Yurochkin. tinyBenchmarks: Evaluating LLMs with fewer examples. In *Proceedings of the 41st International Conference on Machine Learning*, volume 235 of *Proceedings of Machine Learning Research*, pp. 34303–34326. PMLR, 2024a. URL https://proceedings.mlr.press/v235/maia-polo24a.html.

Felipe Maia Polo, Ronald Xu, Lucas Weber, Mírian Silva, Onkar Bhardwaj, Leshem Choshen, Allysson Flavio Melo de Oliveira, Yuekai Sun, and Mikhail Yurochkin. Efficient multi-prompt evaluation of LLMs. In *Advances in Neural Information Processing Systems*, volume 37, 2024b. doi: 10.52202/079017-0707. URL https://proceedings.neurips.cc/paper_files/paper/2024/hash/28236482f64a72eec43706b6f3a6c511-Abstract-Conference.html.

Joon Sung Park, Joseph C. O'Brien, Carrie Jun Cai, Meredith Ringel Morris, Percy Liang, and Michael S. Bernstein. Generative agents: Interactive simulacra of human behavior. In *Proceedings of the 36th Annual ACM Symposium on User Interface Software and Technology*, pp. 1–22. Association for Computing Machinery, 2023. doi: 10.1145/3586183.3606763.

Qwen Team. Qwen3.5-4B model card, 2026a. URL https://huggingface.co/Qwen/Qwen3.5-4B. Hugging Face model repository.

Qwen Team. Qwen3.8-27B model card, 2026b. URL https://huggingface.co/Qwen/Qwen3.8-27B. Hugging Face model repository.

Shibani Santurkar, Esin Durmus, Faisal Ladhak, Cinoo Lee, Percy Liang, and Tatsunori Hashimoto. Whose opinions do language models reflect? In *Proceedings of the 40th International Conference on Machine Learning*, volume 202 of *Proceedings of Machine Learning Research*, pp. 29971–30004. PMLR, 2023. URL https://proceedings.mlr.press/v202/santurkar23a.html.

Gayathri Saranathan, Cong Xu, Mahammad Parwez Alam, Tarun Kumar, Martin Foltin, Soon Yee Wong, and Suparna Bhattacharya. SubLIME: Subset selection via rank correlation prediction for data-efficient LLM evaluation. In *Proceedings of the 63rd Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)*, pp. 30572–30593. Association for Computational Linguistics, 2025. doi: 10.18653/v1/2025.acl-long.1477. URL https://aclanthology.org/2025.acl-long.1477/.

United Nations Statistics Division. *Classifications of Expenditure According to Purpose: COFOG, COICOP, COPNI, COPP*. Number 84 in Statistical Papers, Series M. United Nations, New York, 2000. ISBN 9211614201. ST/ESA/STAT/SER.M/84.

Yuwei Xu, Shulun Zhang, Yingli Zhou, Shipei Zeng, Laks V. S. Lakshmanan, and Chenhao Ma. Topology-aware LLM-driven social simulation: A unified framework for efficient and realistic agent dynamics, 2026. URL https://arxiv.org/abs/2604.18011.

An Yang, Baosong Yang, Beichen Zhang, Binyuan Hui, Bo Zheng, Bowen Yu, Chengyuan Li, Dayiheng Liu, Fei Huang, Haoran Wei, Huan Lin, Jian Yang, Jianhong Tu, Jianwei Zhang, Jianxin Yang, Jiaxi Yang, Jingren Zhou, Junyang Lin, Kai Dang, Keming Lu, Keqin Bao, Kexin Yang, Le Yu, Mei Li, Mingfeng Xue, Pei Zhang, Qin Zhu, Rui Men, Runji Lin, Tianhao Li, Tianyi Tang, Tingyu Xia, Xingzhang Ren, Xuancheng Ren, Yang Fan, Yang Su, Yichang Zhang, Yu Wan, Yuqiong Liu, Zeyu Cui, Zhenru Zhang, and Zihan Qiu. Qwen2.5 technical report, 2024. URL https://arxiv.org/abs/2412.15115.

Yuling Yao, Aki Vehtari, Daniel Simpson, and Andrew Gelman. Using stacking to average bayesian predictive distributions. *Bayesian Analysis*, 13(3):917–1003, 2018. doi: 10.1214/17-BA1091.

Quan Zheng, Yan Gao, Shaobin He, Haoxiang Guan, Yuanhe Tian, Jie Feng, Ming Wang, Shuxin Zheng, and Zhen Liu. APS: Bias-controlled adaptive prototype simulation for population-scale LLM agents, 2026. URL https://arxiv.org/abs/2605.27419.

A Related Work

LLM-based human and social simulation. Persona-conditioned LLMs have been used to emulate survey samples, reproduce human-subject experiments, and instantiate simulated economic decision makers (Argyle et al., 2023; Aher et al., 2023; Horton et al., 2023). Related work compares model opinions with demographic groups or quantifies how persona variables affect subjective-response simulation (Santurkar et al., 2023; Hu & Collier, 2024). This literature primarily studies *fidelity to people*: whether an LLM reproduces human means, subgroup differences, or experimental effects. The evidence is mixed; synthetic survey responses can match some population averages while exhibiting under-dispersion, unstable subgroup relations, and sensitivity to prompt wording or collection time (Bisbee et al., 2024). A complementary line of work studies LLM agents and social dynamics. Generative Agents introduced memory, reflection, and planning for agents in an evolving sandbox (Park et al., 2023), while recent systems scale populations or explicitly model topology and temporal dependencies (Guan et al., 2025; Xu et al., 2026). APS reduces calls by querying representative agents and propagating their responses to nearby agents with bias control (Zheng et al., 2026). These systems generally target trajectories, interactions, or aggregate dynamics under a scenario, and primarily share information along the agent axis. Our task fixes the teacher, prompt, and response-extraction protocol and reconstructs a two-dimensional persona–policy surface. The benchmark crosses common WVS-derived personas with a COFOG-based hierarchy of policy domains, families, and implementation variants (Haerpfer et al., 2022; United Nations Statistics Division, 2000). Each query may transfer across personas, policies, or their interactions, and the target is the complete response distribution at every cell. We therefore study efficient evaluation of a specified simulator over many people and interventions. Human validity, survey replacement, and emergent social trajectories remain distinct evaluation targets.

Active learning and efficient response reconstruction. Classical knowledge distillation trains a compact student from teacher soft targets (Hinton et al., 2015); our problem additionally asks which targets should be purchased when calls are limited. Efficient LLM evaluation usually pursues a different estimand: tinyBenchmarks recovers aggregate scores from small subsets (Maia Polo et al., 2024a), PromptEval estimates prompt-level performance distributions (Maia Polo et al., 2024b), and SubLIME uses anchor models to find subsets that preserve rankings (Saranathan et al., 2025). Active Evaluation Acquisition predicts missing per-example outcomes and adapts the evaluated subset, but learns its dependency model and acquisition policy from evaluation matrices of other models (Li et al., 2024). Our task begins from a zero-history persona–policy pool: the same B observations identify the sharing structure, fit the student, and guide subsequent acquisition. The target is every five-way distribution, complementing aggregate-score and ranking estimands. Pool-based active-learning methods select labels using gradient diversity, Fisher information, or low-budget coverage (Ash et al., 2020; 2021; Hacohen et al., 2022), while variance-reduction and sensor-placement methods value a query through its effect on unobserved locations (Cohn et al., 1996; Krause et al., 2008); local uncertainty and pool-wide error can favor different queries (Imberg et al., 2020). Adaptive

and inductive matrix completion respectively select informative entries or exploit side information (Krishnamurthy & Singh, 2013; Jain & Dhillon, 2013). Our pool is also Cartesian, but each cell is a probability vector and its useful geometry is learned from scarce responses; natural-language policies and persona attributes provide candidate relations. Multitask Gaussian processes, multiple-kernel learning, and predictive stacking provide related tools for shared covariance and model combination (Bonilla et al., 2007; Gönen & Alpaydın, 2011; Yao et al., 2018). SCOPE adapts these ideas by calibrating persona, policy, and interaction kernels through paid leave-one-out Brier error and then using the same mixture to score exact-imputation and cross-pair propagation value. Thus prediction and acquisition form one closed loop coupling completion and query selection.

B BENCHMARK CONSTRUCTION AND TEACHER QUERY PROTOCOL

B.1 PERSONA PANEL AND PUBLIC REPRESENTATION

Sampling. We use the original-coded WVS Wave 7 data, version 6-0-0 (April 15, 2024). Eligible records have a unique interview identifier, a usable country code, and an observed age between 18 and 120. To obtain broad geographic coverage while preserving observed dependencies among attributes, we predefine candidate sets for East Asia, Europe, the Americas, and sub-Saharan Africa. Within each set, among countries with at least 256 eligible records, we select the country with the lowest mean missingness over the retained survey fields. This procedure selects South Korea, Germany, Canada, and Zimbabwe. We then sample 256 whole respondent rows without replacement per country using population seed 20260922. The released population order and teacher tensors are deterministically reproducible from the authorized WVS source data.

Sampling whole rows retains the joint configuration of demographic and attitudinal attributes. Unavailable survey responses are rendered explicitly as unknown in the teacher prompt and receive a separate missingness indicator in the student representation. The panel uses equal country counts and uniform evaluation weights.

Table 4 lists the prompt fields and public student representation. Categorical fields use codebook-defined one-hot encodings; age, income position, and perceived control use observed-only standardization. Missing categorical blocks and missing standardized scalars are set to zero, with field-specific missingness indicators. The resulting 74-dimensional persona vector is RMS normalized over the known population.

Table 4: Persona fields, in prompt order, and their public student representations. Country plus 14 WVS fields produce 74 features after adding field-specific missingness indicators.

Variable	Prompt attribute	Student representation
B_COUNTRY_ALPHA	Country or region	Four-dimensional one-hot
Q260	Sex recorded in the survey	Categorical one-hot + missing indicator
Q262	Age	Standardized scalar + missing indicator
H_URBRURAL	Urban/rural residence	Categorical one-hot + missing indicator
Q273	Marital status	Categorical one-hot + missing indicator
Q274	Number of children	0–6 and 7-or-more categories + missing indicator
Q275	Educational attainment	Codebook-category one-hot + missing indicator
Q279	Employment status	Codebook-category one-hot + missing indicator
Q288	Within-country income ladder	Standardized scalar + missing indicator
Q57	General interpersonal trust	Categorical one-hot + missing indicator
Q71	Confidence in government	Categorical one-hot + missing indicator
Q48	Perceived freedom of choice/control	Standardized scalar + missing indicator
Q1	Importance of family	Categorical one-hot + missing indicator
Q5	Importance of work	Categorical one-hot + missing indicator
Q6	Importance of religion	Categorical one-hot + missing indicator

B.2 POLICY HIERARCHY AND CONSTRUCTION CHECKS

The ten COFOG divisions provide a fixed functional-domain scaffold. Within each division, ten study-defined families pose concrete implementation questions, and each family contains six independently readable alternatives. The alternatives vary consequential choices such as public versus non-public provision, funding and repayment, eligibility and access, geographic allocation, account-

ability, and implementation timing. Related instruments may recur across domains, creating opportunities to test whether paid answers support transfer across policy contexts.

The complete family taxonomy is as follows; each listed family contains six Chinese-language policy descriptions.

General public services. Citizen-facing administrative channels; administrative identity and cross-agency data use; government information disclosure and access; citizen participation in public budgeting; general public procurement; public-sector integrity and conflicts of interest; tax filing and arrears handling; interregional fiscal transfers; public debt and general fiscal constraints; foreign aid and cooperation budgets.

Defence. Defence budgets and funding; adult service obligations and alternative service; reserve training and civilian-employment protection; recruitment and retention benefits; defence procurement and sourcing; foreign military-aid conditions; international defence cooperation; civil-defence alerts and public drills; civil-defence shelters and resource allocation; defence research funding and governance.

Public order and safety. Community policing and patrol resources; police records, complaints, and independent oversight; public-space security monitoring; fraud prevention and reporting; fire prevention in homes and public places; fire-and-rescue coverage and financing; legal aid and public defence; court procedure and small claims; non-custodial and restorative justice; prison conditions and reintegration.

Economic affairs. Minimum-wage adjustment; working hours and scheduling; small-business finance; agricultural risk and income stabilization; strategic-industry and innovation support; import competition and trade facilitation; electricity-demand management and supply investment; public transport and fares; broadband and public communications infrastructure; tourism development and carrying capacity.

Environmental protection. Household-waste sorting and charging; packaging recovery and producer responsibility; reduction of single-use products; wastewater treatment and discharge fees; storm-water pollution; industrial air pollution; noise control; protected areas and use restrictions; ecosystem restoration; and investigation and remediation of contaminated land.

Housing and community amenities. New housing supply and public construction; density and development permits; community land-use and renewal planning; housing repair and accessibility; conversion of vacant space to long-term housing; minimum housing standards and inspection; drinking-water access and provision; residential water pricing and shortage management; water-network renewal and quality disclosure; and street-light coverage.

Health. Primary-care cost sharing; clinic hours and access channels; public-hospital location and waiting-list coordination; essential-medicine procurement; public vaccination delivery; electronic health-record sharing; public mental health entry points; tobacco-harm reduction; regional allocation of health workers; and publicly funded applied medical research.

Recreation, culture, and religion. Public-library provision; museum access and content; heritage protection and public use; public arts funding; sports-facility access and construction; community sports participation; major sporting events and public investment; public broadcasting and content; facilities for religious and other community organizations; and evaluation of cultural and sports services.

Education. Preschool access and fees; compulsory-school funding; public-school admissions and choice; teacher recruitment and disadvantaged-area allocation; school meals; curricula and student assessment; vocational education and apprenticeships; higher-education fees and aid; adult continuing education; and digital teaching and student data.

Social protection. Minimum-income support; old-age income security; unemployment support; sickness-income replacement; disability support; child and family benefits; parental and caregiving leave; survivor benefits; housing-cost assistance; and long-term personal care.

The construction record tracks stable domain, family, and variant identifiers. Structural validation checked the $10 \times 10 \times 6$ hierarchy, identifier uniqueness, and uniqueness of all 600 full texts. Cross-family lexical near-duplicates were screened using Jaccard similarity over contiguous four-character sequences, followed by inspection of the highest-similarity candidates. Content checks covered standalone readability, functional-domain boundaries, policy-logic consistency, and assumptions about unknown persona attributes. Revisions clarified financing constraints, eligibility and appeal procedures, and the allocation of implementation risk.

B.3 A COMPLETE SIX-VARIANT POLICY FAMILY

The following faithful English translations show all six released variants in housing family `cofog06_f01`, which asks how new housing supply should allocate construction responsibility, public land, and long-term constraints. The original Chinese texts are included in the released benchmark.

1. The government directly constructs long-term rental housing on publicly owned land using public investment. Rent primarily funds maintenance and gradual recovery of construction costs. The housing remains publicly owned and cannot be sold as a completed development; eligible applicants enter a transparent allocation process, without a promise to meet all demand immediately.
2. The government provides long-term low-interest construction loans to nonprofit housing organizations. New homes must remain rental properties for 30 years with public operating accounts. Organizations contribute part of the capital and repay the loans; preferential support is recovered after an early sale or change of use, the government does not absorb all vacancy risk, and an independent complaint channel is required.
3. Public housing construction is concentrated in areas with stronger existing public services, despite higher land costs, to reduce residents' distance to services. A fixed budget may therefore produce fewer units. The government publishes site comparisons and provides no direct household cash subsidy for purchase or rent.
4. More housing is constructed in developable peripheral areas. Water and other basic connections are included in the project budget, construction is phased, and occupancy cannot begin before essential services are complete. Travel distances may be longer than for central sites, and no additional transport subsidy is guaranteed.
5. Private developers may lease public land to construct long-term rental housing, with open competition favoring proposals that provide more housing at the specified quality. The land is not sold, rental use must continue for 20 years, the government funds basic site preparation, and the selected developer bears the main construction and operating risks.
6. New public-housing projects are paused for one year while housing need and vacant stock are surveyed and budget comparisons across locations and scales are published. Compliant projects already under construction continue, no new household cash support is introduced during the survey, and construction restarts after one year under public criteria rather than being postponed indefinitely.

B.4 TEACHER PROMPT AND RESPONSE EXTRACTION

The persona block serializes the fields in Table 4 order. Missing fields replace the entire value with "unknown (no usable survey information)." Income is stated as a position on the respondent's national 1–10 survey ladder rather than a cross-country monetary amount, and perceived control retains its 1–10 endpoints.

System message.

> You must simulate the given person's evaluation of a public-policy proposal. Base the answer on the provided persona information and proposal content, and treat experiences

or attributes that are not provided as unknown. The proposal is hypothetical and would be implemented in the person's country; it does not imply that the policy has already been enacted. Each question is independent and does not inherit other proposals or answers. The evaluation may consider the person, other people, and the public interest.

User message.

Persona information:
{profile_text}
Context and proposal:
{policy_text}
From this person's perspective, how acceptable is this proposal?
1: Completely unacceptable
2: Somewhat unacceptable
3: Neither accept nor oppose
4: Somewhat acceptable
5: Completely acceptable
Answer with one number only.

Profile serialization.

Country or region: {country}.
Sex recorded in the survey: {sex}.
Age: {age} years.
Residential area type: {urban_rural}.
Marital status: {marital_status}.
Number of children: {number_children}.
Highest educational attainment: {education}.
Employment status: {employment}.
Household income level: level {income}/10 on the survey ladder in this country (1 is the lowest and 10 the highest; this is not a uniform cross-country monetary amount or an exact income decile).
General trust in other people: {general_trust}.
General confidence in the national government: {government_confidence}.
Perceived freedom of choice and control over life: {control}/10 (1 means none and 10 means very strong).
Importance of family in life: {family_importance}.
Importance of work in life: {work_importance}.
Importance of religion in life: {religion_importance}.

Each pair is evaluated in a fresh context using the teacher's own chat template with an assistant-generation prefix and thinking disabled. Let t_c be the tokenizer-specific single-token encoding of answer $c \in \{1, \dots, 5\}$, and let ℓ_{ij,t_c} be its next-token logit. We verify that the five labels are distinct valid single tokens and define the fixed teacher target as

$$y_{ij,c} = \frac{\exp(\ell_{ij,t_c})}{\sum_{r=1}^{5} \exp(\ell_{ij,t_r})}. \tag{15}$$

Thus each paid query reveals one five-dimensional conditional response vector. The scoring temperature is one. Teacher collection is performed once, and all acquisition seeds access the same fixed response tensor. Public persona processing and policy-only representation extraction occur before acquisition; student fitting, calibration, and query selection access only paid responses from the current task.

C Implementation and Formal Details

C.1 Baseline implementations

All methods use the same 74-dimensional persona features and 32-dimensional teacher-specific policy features. Policy PCA is fitted within the public policy set of each task, and the persona and policy blocks are separately RMS normalized.

Random–MLP uses a 106–128–64–5 GELU network (Hendrycks & Gimpel, 2016), soft-label cross-entropy, and uniform sampling without replacement. It is optimized with AdamW (Loshchilov & Hutter, 2019) for 240 steps using learning rate 0.003 and weight decay 0.01. **Random–XGBoost** fits XGBoost regression (Chen & Guestrin, 2016) in CLR space, maps outputs back through softmax, and uses uniform sampling. It selects depth 2 or 4 and 100 or 300 trees by three-fold TV on queried responses, with learning rate 0.05.

BAIT–MLP uses the same MLP predictor with a Fisher-information batch-design adaptation (Ash et al., 2021). Student features are projected to 32 dimensions and augmented with an intercept; acquisition uses a Gaussian-regression Fisher approximation with forward-$2b$/backward-b selection. **TypiClust–MLP** combines the same MLP with clustering-based coverage and typicality in the public input space (Hacohen et al., 2022). These implementations preserve the original acquisition principles while supporting five-way soft response distributions. SCOPE uses the calibrated prediction and acquisition procedure described in Section 3.

C.2 CANDIDATE STRUCTURES AND PARAMETER RANGES

The main method is a mixture of response kernels over public features. Persona and policy length-scale multipliers are independently selected from $\ell_P, \ell_S \in \{0.25, 1, 4\}$, and interaction strength is $\eta \in \{0, 0.1, 1\}$, giving $3 \times 3 \times 3 = 27$ kernel structures. Each is combined with $\lambda \in \{0.003, 0.03, 0.3\}$, producing $H = 81$ fitted candidates. Both main-effect coefficients in Eq. equation 7 are fixed to $1/2$.

For either public axis, let $d_{ab} = \|x_a - x_b\|_2^2$ and let s be the median of distances strictly greater than 10^{-9}, lower-bounded by 10^{-8}. The base kernel at scale ℓ is

$$R_\ell(a, b) = \exp\left[-\frac{d_{ab}}{2s\ell}\right]. \tag{16}$$

For an axis of size L, we center with $M = I - \mathbf{1}\mathbf{1}^\top/L$ and divide $MR_\ell M$ by its mean diagonal (lower-bounded by 10^{-8}). This makes the two axes comparable in average scale.

Equation equation 9 is optimized by SLSQP from uniform weights, with at most 300 iterations and objective tolerance 10^{-10}. We set weights below 10^{-7} to zero and renormalize, then retain the remaining candidates for prediction and covariance calculations. The exact optimality statement applies to the convex objective before numerical tolerance and sparsification.

C.3 ANALYTIC LEAVE-ONE-OUT PREDICTION AND OUTPUT SCALE

For one candidate, let $b = |Q_t|$, $A = (K_{QQ} + \lambda I)^{-1}$, $v = A\mathbf{1}$, and $d = \mathbf{1}^\top A\mathbf{1}$. Define

$$H_Q = A - vv^\top/d, \qquad \alpha = H_Q Z, \qquad \mu = v^\top Z/d. \tag{17}$$

The exact leave-one-out CLR prediction at queried point q is

$$\widehat{z}_q^{(-q)} = z_q - \alpha_q/(H_Q)_{qq}. \tag{18}$$

We apply softmax before entering the Brier objective. This expression was verified against explicit pointwise refitting.

For arbitrary x, x', the conditional covariance with an estimated intercept is

$$\begin{aligned} c_h(x, x') = K_h(x, x') - k_h(x, Q)Ak_h(Q, x') \\ + \frac{[1 - k_h(x, Q)v][1 - k_h(x', Q)v]}{d}. \end{aligned} \tag{19}$$

Let $E = Z - \mathrm{CLR}(\widehat{P}^{\mathrm{LOO}})$ be the leave-one-out CLR residual matrix. We estimate

$$S_h = \frac{1}{b} E^\top \operatorname{diag}\big(\max\{\operatorname{diag}(H_Q), 10^{-12}\}\big)\, E, \tag{20}$$

which is positive semidefinite and enters Eq. equation 12 after the softmax-Jacobian transformation.

C.4 ALGEBRAIC LIMITS OF RESPONSE FACTORS

CLR coordinates sum to zero. After removing persona and policy main effects, the interaction unfolding satisfies

$$\operatorname{rank}(R_{(1)}) \le \min\{n-1, (m-1)(C-1)\}. \tag{21}$$

For six policies and five responses, at most 20 independent interaction directions exist. A small effective rank must therefore be interpreted relative to this algebraic limit and matched permutation controls. Factor bases are also non-identifiable up to rotation, sign, and scale.

D PUBLIC POLICY ENCODING

We replace the persona profile with the fixed Chinese phrase corresponding to "persona information not provided," retain the original task, policy, and answer format, and extract the final-normalized hidden state at the last input position directly from the transformer. Each policy is encoded once. We compare this *task-last* representation with mean pooling of the bare policy text and mean pooling of the task prompt.

We first fit PCA from public policy representations, then use oracle responses solely to diagnose how much interaction variation lies in the public subspace: In a two-seed housing-domain control with

Table 5: Percentage of interaction energy captured by public policy directions on three 8-persona by 60-policy diagnostic panels.

Domain	Bare top-4	Task-last top-4	Bare top-32	Task-last top-32
General public services	17.04%	**40.41%**	68.91%	**81.85%**
Public order and safety	6.05%	**63.82%**	59.28%	**85.91%**
Housing and community	9.28%	**47.24%**	67.38%	**82.24%**

fixed random queries ($B = 128$), equal input dimensionality, and equal MLP capacity, bare-policy mean pooling, task prompt mean pooling, and task-last obtain TV 0.07282, 0.07118, and 0.05420, respectively. This supports task-last locally, although it can regress on some six-variant family tasks.

If $U_r \in \mathbb{R}^{m\times r}$ contains orthonormal directions from public policy representations and $Y \in \mathbb{R}^{m\times d_Y}$ contains response variation organized by policy, then

$$\min_A \|Y - U_r A\|_F^2 = \|(I - U_r U_r^\top)Y\|_F^2. \tag{22}$$

Thus, capturing more response energy with the same number of public directions lowers the representation-error floor for this linear sharing class. Oracle projection coefficients serve as a diagnostic.

E FORMAL SUPPORT FOR THE JOINT FRAMEWORK

Positive semidefiniteness. Assume the base persona and policy kernels are positive semidefinite, $\eta_h \ge 0$, $\pi_h \ge 0$, and $S_h \succeq 0$. Centering, positive scaling, and nonnegative sums preserve positive semidefiniteness, and the elementwise product of the persona and policy kernels is positive semidefinite by the Schur product theorem. Equation equation 7 is therefore a valid kernel. Moreover, $\operatorname{tr}(J_h(x)S_hJ_h(x')^\top)$ is a Gram kernel of vectorized $J_h(x)S_h^{1/2}$ features. Its product with conditional covariance c_h is positive semidefinite; the predictive-disagreement inner product is also a Gram kernel. Their nonnegative mixture in Eq. equation 12 is consequently positive semidefinite.

Empirical optimality of Brier mixing. Equation equation 9 is a convex quadratic objective on a simplex that contains every single-candidate vertex e_h. Its exact solution satisfies

$$\widehat{R}_{\mathrm{LOO,Brier}}(\pi^*) \le \min_h \widehat{R}_{\mathrm{LOO,Brier}}(e_h). \tag{23}$$

For any fixed evaluation set $\mathcal{E}$, Cauchy–Schwarz further gives

$$\frac{1}{|\mathcal{E}|}\sum_{x\in\mathcal{E}}\mathrm{TV}(p_x,q_x)\le\frac{\sqrt{C}}{2}\sqrt{\frac{1}{|\mathcal{E}|}\sum_{x\in\mathcal{E}}\|p_x-q_x\|_2^2}. \tag{24}$$

Transfer from paid leave-one-out examples to the unknown pool still depends on coverage and the sharing prior and is evaluated empirically.

Pool-wide query value. Freeze the sharing structure, weights, Jacobians, and output scales, and treat C_t as a scalar Gaussian-proxy covariance with update noise $\tau_t(q)$. For $u\neq q$, conditioning on q reduces variance by

$$C_t(u,u)-C_{t+1}(u,u)=\frac{C_t(u,q)^2}{C_t(q,q)+\tau_t(q)}. \tag{25}$$

Once the teacher response at q is observed, q leaves the prediction set and its risk term $C_t(q,q)$ is fully removed. Summing these contributions gives Eq. equation 13. The actual closed loop refits the structure after every batch and is evaluated by TV on teacher probabilities. The proposition characterizes the frozen proxy update; experiments evaluate the full nonlinear process.

E.1 Computational implementation

Define

$$a_h(x)=\mathrm{vec}\Big(J_h(x)S_h^{1/2}\Big)/\sqrt{C}. \tag{26}$$

The implemented update noise is

$$\tau_t(q)=\max\left\{\sum_h\pi_{t,h}\lambda_h\|a_h(q)\|_2^2,10^{-12}\right\}. \tag{27}$$

From U_t, sample T_t uniformly without replacement with $s=\min(512,|U_t|)$. We estimate the query value as

$$\widehat{V}_t(q)=C_t(q,q)+\frac{|U_t|}{s}\sum_{u\in T_t\setminus\{q\}}\frac{C_t(u,q)^2}{C_t(q,q)+\tau_t(q)}. \tag{28}$$

For frozen covariance and fixed q, the first-round target sample gives an unbiased estimate of the propagation sum; maximizing the estimate and making subsequent within-batch choices introduce adaptivity. After each within-batch selection, a Schur update reduces conditional covariance and removes the selected target's prediction risk.

The runtime decomposes as

$$T_{\textsc{Scope}}=T_{\mathrm{public}}(m)+T_{\mathrm{teacher}}(B)+T_{\mathrm{fit}}+T_{\mathrm{acquire}}+T_{\mathrm{predict}}. \tag{29}$$

Exact candidate fitting requires $O(H|Q_t|^3)$ linear algebra per round with $H=81$; full-pool prediction uses cross-kernels between the queried set and the pool.

F Supplementary Results

F.1 Complete two-bottleneck ablation

Table 6 gives the values summarized by Figure 4.

F.2 Response-structure recovery and calibration controls

We evaluate output-structure recovery using saved predictions from SCOPE and Random–MLP for all ten domains. Full 60-policy rows are retained for personas that remain unqueried under both methods, removing exact-imputation effects. After double-centering teacher and student CLR responses, we compute

$$R^2_{\mathrm{int}}=1-\frac{\sum_d\|\widehat{R}_d-R_d\|_F^2}{\sum_d\|R_d\|_F^2} \tag{30}$$

Table 6: Qwen3.5 two-bottleneck factorial ablation (lower is better). Fixed and Calib. denote fixed and calibrated sharing; R/A denote random and pool-value acquisition. Parentheses give the queried fraction.

Level	B (pool)	Fixed/R	Fixed/A	Calib./R	Calib./A
Family	16 (0.26%)	**0.1104**	**0.1104**	0.1124	0.1124
Family	32 (0.52%)	0.0886	0.0829	0.0839	**0.0817**
Family	64 (1.04%)	0.0707	0.0614	0.0659	**0.0587**
Domain	160 (0.26%)	0.0726	0.0656	0.0711	**0.0634**
Domain	320 (0.52%)	0.0573	0.0541	0.0557	**0.0520**
Domain	640 (1.04%)	0.0465	0.0452	0.0455	**0.0437**

Table 7: Interaction recovery and response-geometry CKA on a common unqueried rectangle. Entries are mean ± sample SD over three paired seeds; parentheses give the queried fraction. Higher is better.

Teacher	B (pool)	SCOPE R^2_{int}	MLP R^2_{int}	SCOPE CKA	MLP CKA
Qwen3.5	160 (0.26%)	**0.091±.016**	-0.161±.062	0.508±.028	**0.596±.040**
	320 (0.52%)	**0.191±.007**	-0.065±.045	0.612±.004	**0.657±.043**
	640 (1.04%)	**0.333±.002**	0.092±.040	0.742±.010	**0.748±.023**
Qwen2.5	160 (0.26%)	**0.060±.010**	-0.042±.033	0.544±.015	**0.727±.014**
	320 (0.52%)	**0.116±.011**	-0.006±.034	0.587±.045	**0.789±.016**
	640 (1.04%)	**0.201±.006**	0.048±.034	0.715±.013	**0.843±.014**
Gemma	160 (0.26%)	0.091±.012	**0.113±.021**	0.533±.026	**0.614±.018**
	320 (0.52%)	0.197±.016	**0.219±.023**	0.682±.026	**0.743±.027**
	640 (1.04%)	**0.327±.012**	0.305±.012	0.808±.008	**0.820±.013**
Qwen3.8	160 (0.26%)	**0.089±.006**	0.071±.016	0.537±.030	**0.600±.007**
	320 (0.52%)	**0.177±.006**	0.172±.010	0.650±.022	**0.721±.004**
	640 (1.04%)	**0.304±.004**	0.289±.014	0.762±.010	**0.820±.004**

and linear centered kernel alignment (CKA) of the response geometry (Kornblith et al., 2019).

For all four teachers, SCOPE interaction R^2 and CKA improve with budget. Persona-permutation controls have CKA around 0.010–0.012 and negative interaction R^2. The MLP has higher CKA throughout and higher Gemma interaction R^2 at B160 and B320, whereas SCOPE has higher B640 interaction R^2 for every teacher and lower primary TV.

On all 100 Qwen3.5 families for seed 719, we also compare uniform candidate averaging, selecting one candidate by paid leave-one-out Brier score, and the proposed mixture on identical random query prefixes:

Table 8: Seed-719 calibration controls on identical Qwen3.5 family queries. Gain is the Brier mixture's relative TV reduction versus the best single LOO candidate; boldface follows unrounded values.

B (pool)	Uniform avg.	Best LOO	Brier mixture	Gain
16 (0.26%)	0.1136	0.1127	**0.1124**	0.3%
32 (0.52%)	0.0948	0.0841	**0.0839**	0.3%
64 (1.04%)	0.0764	0.0667	**0.0659**	1.3%

F.3 FIXED AUDIT-SET VALIDATION

An earlier matched experiment used three Qwen3.5 domains, seeds 719 and 920, and 64 personas by 60 policies per condition. For every persona, 12 policies formed a fixed audit set excluded from every acquisition pool. Table 9 averages the six domain–seed conditions. On this discrete budget grid, SCOPE reaches audit TV below 0.05 at budget 256, whereas Random–MLP first reaches it at 512. This complements full-pool evaluation by showing improvement on a common set that remains unqueried for every method.

Table 9: TV on a common held-out audit set. Parentheses give the fraction of the 64×60 pool; Gain is full SCOPE relative to the strongest alternative, SCOPE prediction with random acquisition.

Budget (pool)	Random–MLP	SCOPE + random	Full SCOPE	Gain
32 (0.83%)	0.1194	0.1044	**0.1021**	2.3%
64 (1.67%)	0.0965	0.0841	**0.0820**	2.6%
128 (3.33%)	0.0768	0.0657	**0.0598**	8.9%
256 (6.67%)	0.0564	0.0483	**0.0450**	6.8%
512 (13.33%)	0.0418	0.0386	**0.0366**	5.4%

F.4 AUXILIARY METRICS

Table 10 reports distributional metrics computed from saved predictions for all four teachers.

Table 10: Auxiliary metrics at the largest family and domain budgets. JSD is reported in nats; entries are mean ± sample SD over three paired seeds.

Teacher	Setting (pool)	Full TV	Unqueried TV	JSD	Score MAE
Qwen3.5	Family 64 (1.04%)	0.0592±.0006	0.0598±.0006	0.00461±.00010	0.0985±.0011
Qwen2.5	Family 64 (1.04%)	0.1289±.0031	0.1302±.0031	0.02511±.00120	0.1701±.0051
Gemma	Family 64 (1.04%)	0.1363±.0011	0.1377±.0012	0.03431±.00052	0.1591±.0014
Qwen3.8	Family 64 (1.04%)	0.1275±.0004	0.1289±.0004	0.02357±.00011	0.1840±.0002
Qwen3.5	Domain 640 (1.04%)	0.0436±.0001	0.0440±.0001	0.00241±.00001	0.0706±.0002
Qwen2.5	Domain 640 (1.04%)	0.0856±.0003	0.0865±.0003	0.01148±.00003	0.1073±.0003
Gemma	Domain 640 (1.04%)	0.0993±.0006	0.1004±.0006	0.01997±.00022	0.1133±.0006
Qwen3.8	Domain 640 (1.04%)	0.0959±.0003	0.0969±.0003	0.01363±.00009	0.1335±.0008

F.5 AREA UNDER THE LOW-BUDGET CURVE

For queried fractions $\rho \in \{0.00260417, 0.00520833, 0.01041667\}$, define the normalized area under the log-budget curve

$$\mathrm{AULC} = \frac{1}{\log(\rho_3/\rho_1)} \sum_{k=1}^{2} \frac{R(\rho_k) + R(\rho_{k+1})}{2} \log \frac{\rho_{k+1}}{\rho_k}. \tag{31}$$

Lower is better. Table 11 summarizes the three paired runs.

Table 11: Normalized AULC over the three main queried fractions (lower is better). Entries are mean ± sample SD over three paired seeds; Gain is SCOPE's relative reduction versus the best baseline mean.

Teacher	Level	R–MLP	R–XGB	BAIT	TypiClust	SCOPE	Gain
Qwen3.5	Family	0.0975±.0039	0.1125±.0029	0.1026±.0020	0.0976±.0024	**0.0829±.0012**	14.9%
Qwen2.5	Family	0.1995±.0028	0.2256±.0029	0.2078±.0016	0.2049±.0048	**0.1811±.0067**	9.2%
Gemma	Family	0.2281±.0174	0.2268±.0177	0.2383±.0119	0.2338±.0096	**0.2036±.0037**	10.2%
Qwen3.8	Family	0.1913±.0056	0.1902±.0038	0.1973±.0049	0.1953±.0027	**0.1759±.0038**	7.5%
Qwen3.5	Domain	0.0634±.0006	0.0728±.0012	0.0755±.0017	0.0771±.0003	**0.0525±.0002**	17.2%
Qwen2.5	Domain	0.1212±.0015	0.1382±.0008	0.1390±.0007	0.1403±.0001	**0.1037±.0007**	14.4%
Gemma	Domain	0.1418±.0020	0.1473±.0030	0.1639±.0027	0.1725±.0020	**0.1232±.0014**	13.1%
Qwen3.8	Domain	0.1249±.0010	0.1291±.0019	0.1398±.0011	0.1455±.0006	**0.1140±.0001**	8.7%

G ARTIFACT AND RESULT PROVENANCE

The working artifact records: (i) WVS persona sampling and field processing; (ii) the complete policy text, family, and domain indices; (iii) teacher checkpoint, template, valid-token, and inference-engine configurations; (iv) the complete response-cache hashes and public policy encodings; (v)

frozen random seeds, initial query sets, and per-round acquisition traces; and (vi) saved predictions and scripts that recompute decomposition, TV, JSD, score MAE, CKA, and all table aggregates. Pending results appear as dashes and are omitted from completed-result claims.