An Adoption-Aware Crop Recommender from Farmer World Feedback

Vairaaj Bindal [1]

Abstract

We build an adoption-aware crop recommender from offline farmer survey panel data, treating recorded adoption decisions and realized yields as world feedback. The recommender targets Expected Realized Benefit (predicted yield gain weighted by predicted adoption probability) and is fit on 8,744 cultivated plots from the LSMS-ISA Ethiopia ESS Wave 4 panel. On 800 enumeration-area-disjoint test plots, it beats an accuracy-only baseline by 50 kg/ha (bootstrap 95% CI $[24, 74]$), with Rosenbaum sensitivity $\Gamma^\star \geq 5$ and sign-exchangeable permutation $p < 10^{-3}$. The Ethiopia-trained adoption head reproduces the treatment-effect ordering of the Duflo et al. (2011) Kenya SAFI fertilizer RCT across all four arms. Farmer feedback differs from robot or LLM feedback in that the logging policy is the agent, so its propensities cannot be recovered from the data; we describe the architectural and evaluation choices that let direct-method offline policy learning work in this regime.

1. Introduction

Reinforcement learning from world feedback pushes RLHF past explicit preference labels toward outcomes logged in deployment, such as a robot finishing a manipulation, an LLM agent reaching a downstream goal, or a user clicking a search result. The feedback we study comes instead from human strategic agents, namely smallholder farmers in sub-Saharan Africa, whose adoption decisions and realized yields are recorded together in long-running household survey panels.

The payoff is direct. Yield gains pass through to household consumption and food security, and the LSMS-ISA panels record inputs and outcomes at plot resolution. Our working sample is 8,744 cultivated plots from the Ethiopia ESS Wave 4 panel, spanning four staple crops and 23 districts.

Here the logging policy is itself the agent, which is where the difficulty starts. A farmer chooses whether to apply fertilizer, which crop to plant, whether to act on extension advice; those choices condition on risk tolerance, private information about plot quality, and intra-household bargaining, none of which survey covariates fully capture. The propensities are therefore not identifiable from the data alone. Inverse-propensity (Horvitz & Thompson, 1952; Swaminathan & Joachims, 2015) and doubly-robust (Dudík et al., 2011; Jiang & Li, 2016) estimators need a known or estimable behavior policy, and we do not have one. The offline-RL response is to bound identified policy value under sensitivity assumptions (Tan, 2006; Kallus & Zhou, 2018; Yadlowsky et al., 2022; Manski, 2003), or to reach for proxy and conformal alternatives (Tchetgen Tchetgen et al., 2020; Cui et al., 2024; Taufiq et al., 2022). We take a different route: produce a deployable recommender, then ask which of its model-based policy contrasts survive standard sensitivity, permutation, and placebo tests. The lifts we report are contrasts between two policies scored through the same calibrated heads, not estimates of identified policy value, and we hold that line throughout.

Direct-method offline learning here means predicting yield, predicting adoption, and combining the two analytically. It becomes workable once the predictive heads carry two additions they do not get for free. A structural prior anchors the adoption head to experimentally validated frictions, so it cannot absorb arbitrary correlations from the training distribution; an evaluation protocol then reports which claims survive the unidentified behavior policy. We build both into one system.

Contributions. (1) An adoption-aware crop recommender trained and evaluated end-to-end from observational world feedback, with a hybrid yield head (gradient-boosting base plus literature-anchored re-

[1]Lynbrook High School, San Jose, California, USA. Correspondence to: Vairaaj Bindal <vairaajb@gmail.com>.

RLxF: Reinforcement Learning from World Feedback, ICML 2026 Workshop, Seoul, South Korea, 2026. Copyright 2026 by the author(s).

sponse coefficients), a Mondrian-isotonic-calibrated adoption head, and Mondrian-by-crop split-conformal regression intervals. (2) A 50 kg/ha lift in Expected Realized Benefit over an accuracy-only baseline, robust to confounding (Rosenbaum $\Gamma^\star \geq 5$) and to chance (plot-level and EA-cluster sign-exchangeable permutation, both $p < 5 \times 10^{-4}$). The lift is bracketed at 26 to 44 kg/ha under a partial-identification bound that replaces saturated treated-tier adoption probabilities with SAFI-measured uptake rates. (3) Cross-country directional validation against the Kenya SAFI fertilizer RCT, with the bound above providing its quantitative anchor. (4) An evaluation protocol consisting of Rosenbaum sensitivity, sign-exchangeable permutation, and an education-shuffle placebo, with the placebo separating identified equity targeting from descriptive concentration.

2. The Recommender

The recommender optimizes

$$\mathrm{ERB}(a \mid x) = \mathbb{P}(\text{adopt}\, a \mid x) \cdot \big(\mathbb{E}[Y \mid a, x] - \mathbb{E}[Y \mid a_0, x]\big), \tag{1}$$

where x is a farmer-plot covariate vector, a ranges over a discrete action space with five complexity tiers (T_0 status quo, T_1 small fertilizer nudge, T_2 full input package on the current crop, T_3 crop switch only, T_4 full inputs plus crop switch), a_0 is the status quo, and Y is realized yield in kg/ha. Tiers are ordered by distance from the farmer's logged behavior, and the adoption head consumes that ordering explicitly.

Yield head. A histogram gradient-boosting regressor with Gamma loss is the base learner. On enumeration-area-disjoint splits, the pure HistGBM produces near-zero predicted-yield deltas across counterfactual actions, which collapses the policy comparison. We add a literature-anchored linear-response offset: predicted yield under counterfactual action a is $\hat{Y}(a, x) = \hat{Y}_{\mathrm{HistGBM}}(x) + \beta_{c,\mathrm{fert}} \Delta\mathrm{fert}(a) + \mu_{c,\mathrm{seed}} \mathbf{1}[a \in T_3 \cup T_4]$, with per-crop fertilizer-response coefficients $\beta_{c,\mathrm{fert}}$ from Vanlauwe et al. (2014); Mueller et al. (2012) and improved-seed multipliers $\mu_{c,\mathrm{seed}}$ from Khonje et al. (2015). These coefficients are held fixed and not co-trained. Yield predictions are passed through Mondrian-by-crop split-conformal intervals (Vovk et al., 2003) at 90% nominal coverage (empirical 88.7% on the calibration set).

Adoption head. A logistic baseline propensity over farmer-plot covariates (full feature list in Appendix A), plus per-tier log-odds penalties $\{T_0 : 0, T_1 : -0.4, T_2 : -1.5, T_3 : -3.2, T_4 : -2.8\}$ that encode complexity costs relative to status quo. Penalty magnitudes are taken from the Ruzzante et al. (2021) meta-analysis of 367 adoption studies across 41 sub-Saharan African and Asian countries. Raw logits are recalibrated by household-head education quintile via Mondrian-isotonic regression; shift-aware conformal variants would be a natural extension.

Data and splits. Training and evaluation use LSMS-ISA Ethiopia ESS Wave 4 (Central Statistical Agency of Ethiopia & World Bank LSMS-ISA Team, 2019) after restricting to plots with non-missing crop, fertilizer, and yield records (8,744 cultivated plots). The split is enumeration-area-disjoint at 65/15/20 train/calibration/test with seed 42, so the test set is geographically held out from training. The test set contains 800 plots.

Policy comparison. Let $\hat{a}_i^{\mathrm{aware}} = \arg\max_a \widehat{\mathrm{ERB}}(a \mid x_i)$ and $\hat{a}_i^{\mathrm{acc}} = \arg\max_a \widehat{\mathbb{E}}[Y \mid a, x_i]$ be the two policies' picks per plot, and write $\Delta\mathrm{ERB}(a \mid x_i) = \mathrm{ERB}(a \mid x_i) - \mathrm{ERB}(a_0 \mid x_i)$. The matched-pair difference is

$$d_i = \Delta\mathrm{ERB}(\hat{a}_i^{\mathrm{aware}} \mid x_i) - \Delta\mathrm{ERB}(\hat{a}_i^{\mathrm{acc}} \mid x_i). \tag{2}$$

Both terms are evaluated through the same calibrated heads, so head-level miscalibration cancels in the contrast. Plot-level bootstrap 95% CIs on $\bar{d}$ use $B = 2{,}000$ percentile resamples.

The aggregate matched-pair lift is $\bar{d} = 50$ kg/ha, 95% CI $[24, 74]$. The two policies disagree on 368 of the 800 test plots (the lift is realized on roughly 46% of the test population, with status-quo agreement on the rest). The lift is not reproducible by a yield-minus-λ-complexity heuristic at any $\lambda \in \{0, 100, 250, 500\}$ kg/ha per tier-step (Appendix B, B2).

External directional check and partial-identification anchor. We map the action tiers onto the four arms of the Duflo et al. (2011) Kenya SAFI fertilizer experiment (mapping in Appendix C). The model ranks control strictly below all three treated arms, matching the experimental ordering, but it saturates at $\hat{\mathbb{P}} = 1.000$ on the treated arms where the RCT measures uptake at 0.37 to 0.51. Replacing the saturated treated-tier probabilities with SAFI-measured uptake rates as a partial-identification cap yields 26 to 44 kg/ha across three caps. The most adversarial cap, which replaces every treated-tier probability with the SAFI control rate 0.325, still gives 26 kg/ha CI $[13, 39]$ (Appendix E).

3. Reading the World Feedback

Without identified propensities, the standard offline-policy-evaluation toolkit cannot tell a deployer which claims to trust. The recommender's direct-method estimate is fine on its own terms but silent about robustness, so we pair it with three checks, each aimed at a different way the result could be spurious: Rosenbaum sensitivity, a sign-exchangeable permutation, and a stratified label-shuffle placebo.

Rosenbaum signed-rank sensitivity (Rosenbaum, 2002). We compute the signed-rank statistic on $\{d_i\}$ and report the critical $\Gamma^\star$ at which the one-sided p-value of $H_1 : \bar{d} > 0$ rises above 0.05. Agricultural economics treats $\Gamma = 1.5$ as the convention. The bound here has narrower scope than its usual application: matched pairs in our setting are two policies on the same plot, and the d_i are model-based contrasts through identical calibrated heads rather than observed outcomes. We read $\Gamma^\star$ as the strength of hidden confounding in the adoption-training distribution that would be needed to flip the sign of the contrast. The full mapping onto the marginal sensitivity model of Tan (2006) is in Appendix F.

Sign-exchangeable permutation (plot and EA cluster). Independent sign flips of $\{d_i\}$ ($B = 2{,}000$) test the null of sign-exchangeability around zero. We also run EA-level cluster bootstrap and cluster sign-flip permutation (one shared flip per EA across 44 EAs); both reproduce the aggregate result, with cluster CI $[33, 70]$ and cluster permutation $p < 5 \times 10^{-4}$ (Appendix D). The cluster CI does not widen the plot-level interval because the matched-pair design absorbs EA-shared yield potential and adoption baselines into the differencing.

Stratified label-shuffle placebo. Neither Rosenbaum nor permutation distinguishes a per-subgroup gradient driven by the subgroup feature from one that merely reflects where an aggregate effect concentrates. The placebo holds $\{d_i\}$ fixed, permutes household-head education-quintile labels with a stratified shuffle that preserves marginal subgroup sizes, and recomputes per-subgroup means $\bar{d}_s$ ($B = 2{,}000$). Subgroup $p < 0.05$ would indicate that the education feature carries information beyond its structural correlates such as landholding, market access, and geographic clustering. The aggregate lift and the placebo are independent claims: the placebo can falsify a per-subgroup story without touching the aggregate.

Stratum	n^{dis}	$\Gamma^\star$	Placebo p
Aggregate	368	≥ 8.0	—
No education	257	≥ 8.0	0.14
Early primary	17	5.15	0.86
Late primary	47	≥ 8.0	0.87
Secondary	25	7.85	0.50

Table 1. Aggregate and per-subgroup test results. n^{dis} is the number of test plots where the two policies disagree. $\Gamma^\star$ is the Rosenbaum critical odds-ratio at $\alpha = 0.05$ (ag-econ convention is 1.5). Plot-level sign-exchangeable permutation p-values are $< 10^{-3}$ on aggregate and every subgroup (omitted). Placebo p-values are two-sided against stratified label-shuffles preserving marginal subgroup sizes.

4. What the Protocol Supports

Aggregate. The aggregate sample clears Rosenbaum $\Gamma^\star \geq 5$, well above the ag-econ convention of 1.5. An unobserved confounder would have to shift matched-pair-sign probabilities by at least a 5× odds-ratio to reverse the aggregate result, and the adoption literature does not document confounders of that strength for plausible omitted plot or farmer variables. The sign-exchangeable permutation clears at $p < 10^{-3}$.

Per subgroup. Rosenbaum and permutation clear on every education stratum. The placebo p-values across the four subgroups are $[0.14, 0.86, 0.87, 0.50]$, so the realized per-subgroup means fall well inside the label-shuffle distribution. The aggregate lift concentrates in the no-education stratum because that stratum is 68% of the test set ($n = 543/800$), not because the education feature carries targeting information. The placebo rules out a targeting reading of the per-subgroup pattern while leaving the aggregate untouched.

External validation. The Kenya SAFI directional check (§2) is the second positive line of evidence: an adoption head trained on Ethiopian survey data recovers the realized-uptake ordering across the four arms of an independent Kenyan fertilizer RCT. Direct-method offline policy learning at smallholder scale needs no identifiable propensities, only an adoption head that recovers the right ordering on out-of-distribution actions, and this one does.

Targeting versus coverage. Identified equity targeting and descriptive concentration imply different deployments (deploy preferentially to low-education farmers vs deploy where you can), and both are compatible with the same 53 kg/ha subgroup lift. The placebo separates them.

5. Discussion

What sets farmer world feedback apart from robot or LLM feedback is that the logging policy is the agent. A robot's environment does not change its behavior in anticipation of reward; a farmer's does. Because the off-policy-evaluation literature is built around known or estimable behavior policies, its standard guarantees do not carry over, and the burden moves elsewhere: onto a structural prior that limits what the decision model can absorb from the training distribution, and onto an evaluation protocol that reports which qualitative claims survive the unidentified behavior policy.

A few of these points generalize past the case study itself. Missing propensities can be replaced by priors from independent experiments: the Ruzzante et al. (2021) meta-analysis is what lets us anchor the per-tier adoption penalties without estimating them from the same data the policy is evaluated on, and comparable anchors exist for medical adherence, online tutoring, and energy demand. Drift between observational and experimental calibration turns out to be informative rather than a mere nuisance, since the directional ordering can hold even where absolute magnitudes saturate, and saying so is more useful than collapsing the whole comparison into a point estimate. The protocol itself is cheap: Rosenbaum sensitivity, a sign-exchangeable permutation, and an education-shuffle placebo run in under an hour on 800 plots, and for a world-feedback claim they buy more credibility than one more ablation on the predictive head.

Bounded magnitudes and product of errors. ERB is the product of two head-level estimates, so head-level error can amplify in the realized lift. Our bootstrap is on the contrast $\{d_i\}$, which cancels first-order head-level miscalibration but not joint conformal-yield and adoption-probability uncertainty. The SAFI-anchored cap of §2 brackets the absolute magnitude: even at the most adversarial cap (SAFI control rate 0.325), the contrast remains +26 kg/ha with CI excluding zero. Parent-paper ablations (Appendix B) show the lift clears under ±25% penalty perturbation but collapses under halving. A Tanzania replication does not reproduce the per-subgroup pattern, which is consistent with the descriptive reading the placebo selects.

Open follow-ups. A structural discrete-choice baseline (for example, a multinomial logit over (current tier, x) $\rightarrow$ next tier) would test whether the meta-analytic prior is doing load-bearing work or could be replaced by a more flexible head. Multi-country replication beyond Tanzania (Malawi, Uganda, Nigeria all have comparable LSMS-ISA panels) is the natural next-data extension. A controlled deployment study in which farmers receive ERB-optimal recommendations and adoption is logged on the same scale the model predicts would close the loop between predicted and realized adoption.

6. Conclusion

An adoption-aware crop recommender trained from offline world feedback beats an accuracy-only baseline by 50 kg/ha on Ethiopian smallholder data. The lift is robust to confounding, to chance, and to EA clustering, and a partial-identification bound anchored to the Kenya SAFI fertilizer RCT brackets it at 26 to 44 kg/ha. World feedback from strategic agents is workable as a training signal when the predictive heads are anchored by structural priors and the evaluation protocol reports which claims survive an unidentified behavior policy.

References

Central Statistical Agency of Ethiopia and World Bank LSMS-ISA Team. Ethiopia Socioeconomic Survey 2018-2019, Wave 4. World Bank Microdata Library, Catalog 3823, 2019. URL https://microdata.worldbank.org/index.php/catalog/3823.

Cui, Y., Pu, H., Shi, X., Miao, W., and Tchetgen Tchetgen, E. J. Semiparametric proximal causal inference. Journal of the American Statistical Association, 119(546):1348–1359, 2024.

Dudík, M., Langford, J., and Li, L. Doubly robust policy evaluation and learning. In International Conference on Machine Learning, 2011.

Duflo, E., Kremer, M., and Robinson, J. Nudging farmers to use fertilizer: Theory and experimental evidence from Kenya. American Economic Review, 101(6):2350–2390, 2011.

Horvitz, D. G. and Thompson, D. J. A generalization of sampling without replacement from a finite universe. Journal of the American Statistical Association, 47(260):663–685, 1952.

Jiang, N. and Li, L. Doubly robust off-policy value evaluation for reinforcement learning. In International Conference on Machine Learning, 2016.

Kallus, N. and Zhou, A. Confounding-robust policy improvement. Advances in Neural Information Processing Systems, 31, 2018.

Khonje, M., Manda, J., Alene, A. D., and Kassie, M. Analysis of adoption and impacts of improved maize

varieties in eastern Zambia. World Development, 66: 695–706, 2015.

Manski, C. F. Partial identification of probability distributions. 2003.

Mueller, N. D., Gerber, J. S., Johnston, M., Ray, D. K., Ramankutty, N., and Foley, J. A. Closing yield gaps through nutrient and water management. Nature, 490(7419):254–257, 2012.

Rosenbaum, P. R. Observational Studies. Springer, 2nd edition, 2002.

Ruzzante, S., Labarta, R., and Bilton, A. Adoption of agricultural technology in the developing world: A meta-analysis of the empirical literature. World Development, 146:105599, 2021.

Swaminathan, A. and Joachims, T. The self-normalized estimator for counterfactual learning. In Advances in Neural Information Processing Systems, 2015.

Tan, Z. A distributional approach for causal inference using propensity scores. Journal of the American Statistical Association, 101(476):1619–1637, 2006.

Taufiq, M. F., Ton, J.-F., Cornish, R., Teh, Y. W., and Doucet, A. Conformal off-policy prediction in contextual bandits. In Advances in Neural Information Processing Systems, 2022.

Tchetgen Tchetgen, E. J., Ying, A., Cui, Y., Shi, X., and Miao, W. An introduction to proximal causal learning. arXiv preprint arXiv:2009.10982, 2020.

Vanlauwe, B., Coyne, D., Gockowski, J., Hauser, S., Huising, J., Masso, C., Nziguheba, G., Schut, M., and Van Asten, P. Sustainable intensification and the african smallholder farmer. Current Opinion in Environmental Sustainability, 8:15–22, 2014.

Vovk, V., Lindsay, D., Nouretdinov, I., and Gammerman, A. Mondrian confidence machine. In Technical Report, Royal Holloway, University of London, 2003.

Yadlowsky, S., Namkoong, H., Basu, S., Duchi, J., and Tian, L. Bounds on the conditional and average treatment effect with unobserved confounding factors. Annals of Statistics, 50(5):2587–2615, 2022.

A. Adoption-Head Features and Hyperparameters

The adoption-head logistic baseline propensity $\hat{p}_{\text{base}}(x)$ is fit on the following farmer-plot covariates, all plot-fixed (they do not vary across the counterfactual action tiers T_0–T_4):

- Household-level: household-head age (years), education (years of schooling, with the s2q06 non-attender fix applied so non-attenders are 0 rather than NA, lifting non-null rate from 32% to 92.5%), sex, household size, total cultivated area (ha), distance to nearest market (km), prior-wave extension contact (binary), prior-wave credit access (binary).
- Plot-level: plot area (ha), tenure status (owned vs rented), current crop (one-hot over 30 crops), current input bundle — inorganic fertilizer (kg/ha, winsorized at 3,000), organic fertilizer (binary), irrigation (binary), improved seed (binary), $\log(1 + \text{fert kg/ha})$ as a derived feature, prior-wave realized yield (kg/ha, winsorized at 30,000).
- Agro-ecological (EA-centroid join): CHIRPS annual rainfall (mm), WorldClim mean temperature (°C), NDVI growing-season mean, elevation (m), slope (%), FAO soil-quality indices sq_1–sq_7, agro-ecological zone (one-hot). LSMS-ISA's standard 0–5 km rural / 0–2 km urban GPS offset is inherited; loss is negligible at CHIRPS (5 km) and SoilGrids (250 m) resolutions.
- Local social: leave-one-out enumeration-area peer-adoption rate of the target input (computed excluding the focal plot). Plot-fixed across all action tiers — it is a property of the neighborhood, not of the recommendation.
- Survey-design: survey wave (one-hot over Waves 1–4), district fixed effects, region fixed effects.

Action-complexity features that do vary across counterfactual tiers are added on top of $\hat{p}_{\text{base}}(x)$ as the literature-anchored TIER_LOG_ODDS_PENALTY vector $\{T_0 : 0,\ T_1 : -0.4,\ T_2 : -1.5,\ T_3 : -3.2,\ T_4 : -2.8\}$, modulated by EA peer-adoption rate (coefficient +1.0, social-learning credit) and household-head education (coefficient $-0.25 \cdot \log(1 + \text{edu})$, education-as-friction-reducer). Both modulators are anchored to the Ruzzante et al. (2021) meta-analysis.

Hyperparameters.

- HistGBM yield base. max_depth= 6, learning_rate= 0.05, n_estimators= 500, min_samples_leaf= 20, Gamma loss. Counterfactual yields under action a are computed by adding the literature-anchored offset to the base prediction at x; coefficients are not co-trained.
- Adoption-head logistic baseline. L2 regularization $C = 1.0$, no feature interactions, fit on training-fold plots only.
- Mondrian-isotonic recalibration. Per education-quintile isotonic regressor fit on the held-out calibration set; fallback to global isotonic when a stratum has fewer than 80 calibration plots. Reduces Expected Calibration Error from $0.166 \to 0.138$ on fertilizer adoption and $0.131 \to 0.061$ on improved-seed adoption.
- Mondrian-by-crop split conformal. Nominal coverage 90% at $\alpha = 0.10$; empirical coverage 88.7% on calibration.
- Resampling. Bootstrap $B = 2{,}000$; sign-exchangeable permutation $B = 2{,}000$; education-shuffle placebo $B = 2{,}000$. Random seed 42 for the 65/15/20 EA-disjoint train/calibration/test split and for all resampling.

B. Parent-Paper Ablations

This appendix reproduces the two parent-paper ablations that the body references for baseline breadth and penalty-magnitude robustness. Both use the same $n = 800$ EA-disjoint Ethiopia test set as the headline result.

Penalty-grid sensitivity (B1). Multiplicative rescalings of the TIER_LOG_ODDS_PENALTY vector by $\{0.5, 0.75, 1.0, 1.25, 1.5\}$.

Multiplier	Aware lift (kg/ha)	Acc-only lift (kg/ha)	Δ (aware − acc)	n^{dis}
0.50×	+403 [+377, +430]	+387 [+362, +413]	+16 [−22, +51] ns	196
0.75×	+343 [+319, +367]	+308 [+286, +332]	+35 [+3, +64]	296
1.00× (default)	**+263** [**+244**, **+282**]	**+213** [**+196**, **+230**]	**+50** [**+24**, **+74**]	368
1.25×	+213 [+195, +231]	+155 [+139, +172]	+58 [+37, +79]	442
1.50×	+186 [+170, +203]	+127 [+112, +143]	+60 [+42, +77]	524

Table 2. Penalty-grid sensitivity (B1). Bootstrap 95% CIs from $B = 2{,}000$ resamples on the same 800 EA-disjoint test plots. The headline lift clears at every multiplier ≥ 0.75; halving the penalties (0.5×) collapses the gap. This is the reviewer-strict bound: literature-anchored penalty magnitudes must be within roughly a factor of two of their adopted values for the headline claim to hold.

Heuristic baseline (B2). Yield-minus-λ·tier-complexity policy: $\hat{a}_i^{\lambda} = \arg\max_a[\hat{f}(a, x_i) - \lambda \cdot \text{tier}(a)]$ for $\lambda \in \{0, 100, 250, 500\}$ kg/ha per tier-step. The question is whether a uniform tier-cost penalty can substitute for explicit $\hat{\mathbb{P}}(\text{adopt}\, a \mid x)$ modeling.

λ (kg/ha/tier)	Δ (λ vs acc)	Δ (aware vs λ)	Disagree vs aware	Disagree vs acc
0	+0 [−25, +23] ns	**+50** [**+24**, **+74**]	368	0
100	+19 [−7, +42] ns	+31 [+4, +56]	255	125
250	−3 [−30, +23] ns	+53 [+24, +80]	343	452
500	−64 [−92, −38] worse	+114 [+85, +141]	540	638

Table 3. Heuristic yield-minus-λ·complexity baseline (B2). No λ delivers a statistically significant lift over accuracy-only; adoption-aware ERB beats every λ tested by a margin whose CI excludes zero. A uniform tier-cost penalty cannot reproduce ERB's per-(plot, action) reasoning about which specific high-yield-but-high-complexity recommendations a particular farmer is unlikely to adopt.

Other parent-paper sensitivity axes. Three further ablations are reported in full in the parent work and noted here for completeness: (B3) zeroing the literature-anchored response components and running the head-to-head on a pure HistGBM yield base collapses both policies to $\Delta = +0\ [-10, +10]$ (ns; 97 disagreements vs the hybrid's 368) — the hybrid head is a structural prerequisite, not an add-on optimization. (B4) the education-shuffle placebo whose results sit in Table 1 of the main body. (B5) an adoption-on-predicted-yield feedback-loop sweep (per-(plot, action) log-odds nudge $\beta \cdot \Delta\hat{Y}$) shows the aggregate lift survives for $\beta \leq 5 \times 10^{-4}$ kg^{-1}, comfortably wider than the realistic range implied by elasticity estimates in the smallholder adoption literature.

C. Kenya SAFI Mapping

The Duflo–Kremer–Robinson (Duflo et al., 2011) Kenya SAFI fertilizer experiment is the canonical RCT for fertilizer-adoption nudges in smallholder agriculture. Four arms cover the cross-product of two interventions: a behavioral commitment device (SAFI) and free input delivery. We map each arm to an action-complexity vector matching the adoption head's input schema, hold farmer covariates at their Kenyan SAFI values, and score the Ethiopia-trained adoption head on the resulting (farmer, action) inputs.

SAFI arm	Mapped action tier	n	RCT-measured	Model-predicted	|gap|
Control	T_0 status quo	477	0.325	0.004	0.321
SAFI commitment nudge	T_1 small fertilizer nudge	111	0.405	1.000	0.595
Free delivery (no commitment)	T_2 full input package, current crop	196	0.372	1.000	0.628
SAFI + free delivery	T_2 + behavioral reinforcement	93	0.505	1.000	0.495

Table 4. Kenya SAFI external validation. The control arm receives no intervention and maps to T_0. The three treated arms involve some combination of behavioral commitment device and free delivery; both interventions encourage fertilizer adoption without changing crop choice, so they map naturally onto T_1 (small nudge) and T_2 (full input package on the current crop). The model preserves the control-vs-treated ordering but saturates at $\{0, 1\}$ on the OOD treated arms.

What the directional check supports. The Ethiopia-trained adoption head, presented with Kenyan farmer covariates and the four SAFI action vectors, ranks control strictly below every treated arm — matching the

experimentally-confirmed ordering. The within-treated rank ordering is not recovered: predicted probabilities saturate at 1.000 across the three treated arms, so the model cannot reproduce the $\sim$ 13-percentage-point spread between SAFI-only (0.405) and free-delivery-only (0.372), or the $\sim$ 10-percentage-point lift from adding free delivery on top of SAFI (0.405 $\rightarrow$ 0.505). The absolute calibration gap is approximately 50 percentage points on the three treated arms.

Why the saturation. The adoption head is trained on observational LSMS-ISA data in which adoption labels are paired with the action complexity of the package the farmer actually chose. At inference on counterfactual actions (for example, scoring T_2 for a farmer whose observed choice was T_0), the model extrapolates to action-complexity values it was never trained on, and tends toward the decision-boundary extremes. The same pathology appears internally in the Ethiopia panel for some held-out counterfactual tiers, but the Kenya RCT is the cleanest external setting in which to measure its magnitude.

Quantitative alignment. The four-arm mapping by itself does not align predicted and realized uptake. We address quantitative alignment by inverting the question: instead of asking whether predicted magnitudes match SAFI's, we use SAFI's measured uptake rates to bound the within-treated adoption probabilities the recommender uses. The resulting partial-identification analysis is in Appendix E.

D. EA-Cluster-Robust Inference

The body reports plot-level bootstrap and sign-exchangeable permutation. Because plots within the same enumeration area share peer-adoption rates, micro-climate, and local extension capacity, observations are not independent. This appendix re-runs both procedures at the EA-cluster level on the same matched-pair vector $\{d_i\}$, $i = 1, \ldots, 800$, with EA labels recovered by joining test-plot indices back to the LSMS-ISA panel.

Setup. The test set spans 44 unique EAs with plot counts per EA ranging 2–61 (median 16, mean 18). The cluster bootstrap resamples 44 EAs with replacement and takes all plots in each sampled EA; the cluster sign-flip permutation flips one shared sign per EA across all its plots. Both procedures use $B = 2{,}000$ resamples with seed 42.

Statistic	Plot-level	EA-cluster
Bootstrap mean $\bar{d}$ (kg/ha)	+50	+50
Bootstrap 95% CI	[+24, +74]	[+33, +70]
Sign-exchangeable p (two-sided)	$< 5 \times 10^{-4}$	$< 5 \times 10^{-4}$

Table 5. Plot-level versus EA-cluster bootstrap and permutation on the same $\{d_i\}$. Cluster $p < 5 \times 10^{-4}$ is the smallest resolvable value at $B = 2{,}000$ (no permuted mean meets or exceeds $|\bar{d}|$).

Reading. Cluster inference tightens the plot-level interval slightly rather than widening it. The matched-pair design is the reason: both policies are scored on the same plot through identical calibrated heads, so EA-shared yield potential, EA-shared adoption baselines, and EA-shared peer-adoption rates differ out in $d_i = \Delta\mathrm{ERB}_i^{\mathrm{aware}} - \Delta\mathrm{ERB}_i^{\mathrm{acc}}$. Cluster correction widens an interval only when the quantity being aggregated covaries within cluster, which the matched-pair d_i does not.

Education quintile	n	n_{EA}	$\bar{d}$ (kg/ha)	EA-cluster 95% CI
No education ($q = 0$)	543	44	+52.5	[+33, +74]
Early primary ($q = 1$)	66	15	+37.9	[+9, +90]
Late primary ($q = 2$)	104	24	+40.2	[+23, +60]
Secondary ($q = 3$)	47	16	+48.9	[+25, +71]

Table 6. EA-cluster bootstrap 95% CIs by education quintile. Every subgroup CI excludes zero under cluster resampling; the early-primary CI is widest, reflecting the smallest sample ($n = 66$, 15 EAs) rather than instability of the point estimate.

Per-subgroup cluster bootstrap. The placebo p-values reported in Table 1 of the main body are unchanged: the placebo permutes education labels across plots, not signs across EAs, so plot-level permutation is the correct null. Reproducing the placebo at the EA level would require an EA-stratified shuffle, which we have not run.

E. SAFI-Anchored Partial-Identification Bound

The body reports a +50 kg/ha matched-pair contrast computed with the adoption head's saturated $\hat{\mathbb{P}} = 1.000$ on the treated tiers. We bracket the absolute magnitude under the assumption that the true within-treated adoption probability lies in the range observed in the most relevant RCT we have.

Construction. For each plot i, the adoption-aware policy picks tier $\hat{a}_i^{\text{aware}}$ with predicted adoption probability $\hat{p}_i = \hat{\mathbb{P}}(\text{adopt}\, \hat{a}_i^{\text{aware}} \mid x_i)$. The body uses this $\hat{p}_i$ in $\Delta\text{ERB}_i^{\text{aware}} = \hat{p}_i \cdot (\hat{Y}(\hat{a}_i^{\text{aware}}, x_i) - \hat{Y}(a_0, x_i))$. We replace $\hat{p}_i$ with a capped value $\tilde{p}_i = \min(\hat{p}_i, p_{\text{cap}})$ whenever the chosen tier is a treated tier (T_1, T_2, T_3, T_4), and re-derive $\Delta\text{ERB}_i^{\text{aware}}$ as $\tilde{p}_i \cdot (\hat{Y}(\hat{a}_i^{\text{aware}}, x_i) - \hat{Y}(a_0, x_i)) = \Delta\text{ERB}_i^{\text{aware}} \cdot (\tilde{p}_i / \hat{p}_i)$. The same cap is applied to the accuracy-only policy's picks. The cap is held constant across treated tiers because SAFI does not isolate within-treated rankings the model could learn.

Three caps. We report results at three caps drawn from the SAFI-measured uptake rates:

- SAFI maximum ($p_{\text{cap}} = 0.505$): the highest measured arm, "SAFI commitment + free delivery". Generous bound that asks how much of the unanchored +50 kg/ha survives under the most adoption-favorable RCT-anchored ceiling.
- SAFI treated mean ($p_{\text{cap}} = 0.427$): mean of the three treated arms. Conservative middle reading.
- SAFI control ($p_{\text{cap}} = 0.325$): control rate. Most adversarial setting — assumes the within-treated adoption probability for any recommended tier is no higher than the rate at which Kenyan farmers fertilize in the absence of any intervention.

Cap	p_{cap}	$\bar{d}$ (kg/ha)	EA-cluster 95% CI	$n_{\text{aware}}^{\text{cap}}$
SAFI maximum	0.505	+44	[+27, +62]	186
SAFI treated mean	0.427	+38	[+23, +56]	251
SAFI control (adversarial)	0.325	+26	[+13, +39]	349
Unanchored body number	—	+50	[+33, +70]	0

Table 7. SAFI-anchored partial-identification bound. $n_{\text{aware}}^{\text{cap}}$ is the number of adoption-aware policy picks whose predicted probability is capped (out of 682 non-status-quo picks). CIs are EA-cluster bootstrap with $B = 2{,}000$, seed 42.

Interpretation. The within-treated adoption probabilities the recommender saturates at are clearly wrong in absolute terms, but the contrast survives every plausible replacement consistent with the SAFI experimental evidence. The aware-vs-accuracy-only ranking is therefore not driven by the saturated probabilities themselves: even when they are replaced by the SAFI control rate, which is structurally too low for plots receiving recommendations, the matched-pair contrast remains positive. The calibration of within-treated adoption probabilities matters for cost-benefit translation, not for the headline policy comparison.

Caveat. The cap replaces a single scalar $\hat{p}_i$ per plot; it does not replace the per-(plot, action) variation that distinguishes ERB from a uniform tier-complexity penalty. The relevant complement is the B2 ablation in Appendix B, which shows that a uniform tier penalty cannot reproduce the lift even before any saturation correction. Together, the two analyses establish that the gain is neither an artifact of uniform conservatism (B2) nor of OOD adoption-probability saturation (this appendix).

F. Rosenbaum Γ for Model-Based Matched-Pair Contrasts

Rosenbaum's signed-rank sensitivity bound was developed for matched-pair studies of treatment effects on directly observed outcomes. The classical interpretation of the critical $\Gamma^\star$ at which the one-sided p-value crosses

0.05 is: an unobserved confounder would need to shift the matched-pair treatment-assignment odds-ratio by at least $\Gamma^\star$ to overturn the qualitative conclusion. In our setting, both "treatments" (the adoption-aware and accuracy-only policies) are deterministic functions of x, the matched "pairs" are the same plot scored under both policies through identical calibrated heads, and the d_i are model-based contrasts, not directly observed outcomes.

Null distribution in our setting. The signed-rank statistic on $\{d_i\}$ retains a well-defined null distribution. Under $\Gamma = 1$ (no hidden bias) and sign-exchangeability of $\{d_i\}$, the distribution is the standard signed-rank. Under $\Gamma > 1$, the worst-case distribution shifts toward the null. The critical $\Gamma^\star$ at which the one-sided p-value exceeds 0.05 retains the operational meaning of how strong an asymmetry in the matched-pair sign-generating distribution would have to be to overturn the headline.

What asymmetry, exactly. There is no observed-outcome treatment assignment in this setting, so the relevant asymmetry is in the sign of d_i on a given plot, conditional on a hypothetical confounder that perturbs the conditional log-odds of adoption (the input to the adoption head). Under the marginal sensitivity model of Tan (2006) applied to the adoption-head logit, where the true conditional adoption probability $\mathbb{P}^*(\text{adopt} \mid x, a)$ satisfies $\Gamma^{-1} \leq \text{OR}(\mathbb{P}^*, \hat{\mathbb{P}}) \leq \Gamma$, the signed-rank statistic on $\{d_i\}$ has worst-case distribution shifted by a function of Γ.

Scope. The Γ-bound covers hidden confounding in the observational adoption-training data. It does not address miscalibration of the yield head, which is the job of the SAFI-anchored bound (Appendix E) and of the bootstrap on $\{d_i\}$ that cancels first-order head-level miscalibration in the contrast. It also does not address whether the policy comparison would replicate under a different adoption-head architecture; that is the discrete-choice baseline flagged in §5.

Reporting convention. We report Rosenbaum's $\Gamma^\star$ values (Table 1, body) following standard matched-pair sensitivity practice. The narrower scope is: $\Gamma^\star \geq 5$ means the model-based contrast survives a $5\times$ multiplicative perturbation to the adoption-head logits in the worst-case direction, not that the absolute yield effect is identified.