NBER WORKING PAPER SERIES

TARGETING, PERSONALIZATION, AND ENGAGEMENT IN AN AGRICULTURAL ADVISORY SERVICE

Susan Athey
Shawn Cole
Shanjukta Nath
Jessica Zhu

Working Paper 34951
http://www.nber.org/papers/w34951

NATIONAL BUREAU OF ECONOMIC RESEARCH
1050 Massachusetts Avenue
Cambridge, MA 02138
March 2026

The study in this paper was conducted in cooperation with Precision Development, a non-profit that Cole co-founded, and which employed Zhu. Athey and Cole serve as non-compensated advisors to the non-profit. The Golub Capital Social Impact Lab at Stanford GSB and the DFRD at Harvard Business School funded this research. The views expressed herein are those of the authors and do not necessarily reflect the views of the National Bureau of Economic Research.

Targeting, Personalization, and Engagement in an Agricultural Advisory Service
Susan Athey, Shawn Cole, Shanjukta Nath, and Jessica Zhu
NBER Working Paper No. 34951
March 2026
JEL No. C93, O13, Q16

ABSTRACT

We conduct a series of iterative experiments to evaluate approaches for optimally targeting calls for an agricultural advisory service serving over one million farmers in rural India. We estimate the value of alternative targeted policies using "off-policy" evaluation on data from randomized call assignments. When we evaluate off-policy using held-out data from the same time periods used to design the targeted policies, we find that targeted policies increase engagement by up to 8%. However, when we design a policy using current data and then implement the policy for randomly selected users in subsequent weeks, our "on-policy" evaluation estimates show realized gains that are substantially smaller. We develop tools to diagnose the causes of this underperformance and adopt a transfer-learning approach to policy learning and off-policy evaluation that accounts for temporal changes in farmer behavior, substantially improving off-policy estimates of performance in subsequent weeks. We further develop novel approaches to targeted policy design that respond to organizational objectives and resource constraints, including distributional goals (e.g., reaching women farmers) and prioritizing farmers likely to benefit most from the information on downstream outcomes (e.g., yields). We propose a method for organizations to quantify trade-offs between objectives, and we demonstrate the value of targeting for alleviating these trade-offs.

Susan Athey
Stanford University
Graduate School of Business
Department of Economics
and NBER
athey@stanford.edu

Shawn Cole
Harvard University
Harvard Business School
and NBER
scole@hbs.edu

Shanjukta Nath
University of Georgia
shanjukta.nath@uga.edu

Jessica Zhu
jzhu@precisiondev.org

A randomized controlled trials registry entry is available at
https://www.socialscienceregistry.org/trials/13161

1 Introduction

The combination of large data sets, theoretical advances in machine learning, and experimentation has dramatically improved organizations' ability to personalize marketing efforts. However, best practices for developing and evaluating personalized policies are still evolving. One way to assess the value of a particular policy (that is, a mapping from a user's characteristics and history to treatments) is to run a randomized experiment where one set of users receives the policy and another set of users does not. In this paper, we refer to such an evaluation as an "on-policy" evaluation (in industry, it may also be referred to as an "A/B Test" or a "champion vs. challenger" evaluation). When there are many alternative policies to consider, it may be prohibitively expensive to evaluate all policies of interest in this way.

Simester et al. (2020a) propose a powerful and attractive approach to evaluate multiple targeting policies with just a single, large experiment. If there is data available where treatments are randomly assigned, organizations can evaluate ex post a wide variety of targeted policy rules without running additional experiments. The researcher splits the data, using part of the data as "training data" to design targeted policies (see, e.g. Athey and Wager (2021); Kitagawa and Tetenov (2018)) and part as "evaluation data" where the value of the targeted policy can be estimated in an approach we refer to as "contemporaneous off-policy" evaluation (that is, estimation of the value of a counterfactual treatment assignment policy).[1] This approach has been used in a growing variety of contexts, especially in the marketing literature (Rafieian and Yoganarasimhan, 2021; Yoganarasimhan et al., 2023). There is, however, little work evaluating the efficacy of these "off-policy" methods. A natural approach is to compare policies using off-policy evaluation, select one, and then conduct an on-policy evaluation in subsequent periods using a randomized controlled trial where users are randomly assigned to either the policy or a baseline. However, the literature provides little guidance on how to diagnose or improve performance when there is a discrepancy between the performance anticipated by off-policy evaluation and the results obtained by on-policy evaluation.

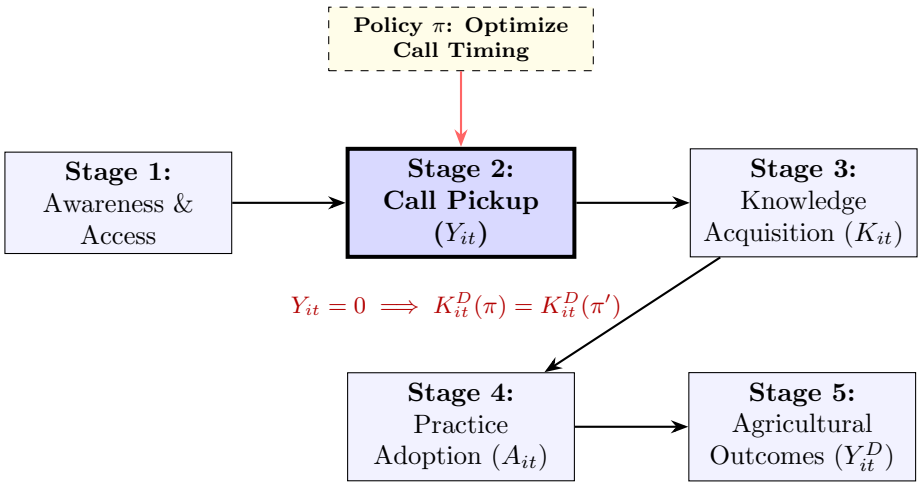
This paper develops, implements, and evaluates targeted policies that customize an information marketing intervention (a weekly phone call with timely agricultural advice) for a population of over one million farmers in an agrarian state in India. Our primary outcome is farmer engagement with the service, a binary indicator of whether a farmer responds to a call (call pickup) and receives agricultural advice. We iteratively design multiple targeted policies to maximize farmer engagement with the service by calling at a time the farmer is most likely to pick up, and evaluate each policy with three complementary techniques: contemporaneous off-policy evaluation, on-policy evaluation in subsequent weeks, and off-policy evaluation using the purely randomized data in subsequent weeks ("K-step ahead off-policy," where the targeted policy is designed on data from week t and evaluated on data from week $t + k$ where $k > 0$).

[1] With J treatment arms under uniform randomization, overlap with a targeted policy is approximately $\frac{1}{J}$ of the sample, though all data can be used to estimate the baseline uniform policy value. In our setting with $J = 91$, this yields roughly $\frac{N}{91}$ farmers contributing to the targeted policy estimate in off-policy evaluation. In contrast, on-policy evaluation testing a targeted policy against a baseline in a 33-67 randomization allows one-third of the data (approximately $\frac{N}{3}$) to directly inform the targeted policy estimate, providing substantially higher statistical power.

Our findings suggest that the policy evaluation approach matters a great deal. Our contemporaneous off-policy evaluation indicated that our final targeted policy would increase call pickup by 2.6 percentage points (an 8% increase), equivalent to reaching 34,000 additional farmers at no additional cost. Yet, the actual gains estimated using on-policy evaluation were much more modest. When the NGO deployed the targeted policy in the final week of our program, the actual gains were 0.4 percentage points.

Earlier work has indeed recognized the risk that implemented policies may not perform as well as expected when the relationship between covariates and outcome changes over time (temporal drift), but it has not proposed tools to diagnose and mitigate these risks.[2] We propose and implement a "transfer learning" approach, adapting a solution designed for prediction problems to the problem of designing and evaluating targeted policies. Using K-step-ahead off-policy evaluations and simulations, we show that this transfer learning (TL) approach does not degrade performance in settings with no temporal drift but can dramatically improve performance when there is substantial temporal drift. When we design targeted policies using the TL approach and evaluate them using K-step ahead off-policy methods, we achieve 7-8% pickup gains using the targeted policy (relative to the uniform policy). We further evaluate potential gains from alternative approaches to designing the targeted policies. In the on-policy experiments, we design and implement targeted policies, treating each hour block in the week as a discrete treatment arm. We compare this targeted policy with an alternative targeted policy that uses sieve estimators to flexibly model treatment effects and their interactions with covariates in the outcome model estimation. Using K-step ahead off-policy evaluation, the targeted policy designed using outcome models with sieve estimators achieves 5.18 pp improvement in pickup rates (relative to uniform policy), compared to 2.29 pp when using outcome models with discrete treatment indicators.

Figure 1: Marketing Funnel for Agricultural Advisory Service.



[2] Simester et al. (2020b) calls this instability in the relationship between outcome and covariates "concept shift". We will use the term "temporal drift" to denote this time-varying instability in the relationship between outcome and covariates.

The literature has recognized tradeoffs in selecting outcome variables in an environment of iterative experimentation (Athey, 2025; Athey et al., 2025a; Yang et al., 2024). Iterative experimentation requires the selection of outcome variables that can be measured in a short time frame. In addition, an outcome that is directly impacted by the intervention often has a better signal to noise ratio, which reduces variance when estimating treatment effect heterogeneity, improving the performance of policies. In our context, Figure 1 illustrates the causal chain between our intervention and long-term agricultural outcomes. Call pickup is necessary for our intervention to have downstream effects, and the primary mechanism through which our intervention affects long-term outcomes (yields, profits) is through improving call pickup rates. Long-term outcomes are expensive to measure and require months or years to observe, making them unsuitable for such quick iterative policy design and evaluation. Thus, we optimize short-term engagement (call pickup rates) rather than long-term agricultural outcomes. We benefit from several rigorous impact evaluations of this same type of advisory service, which have demonstrated that access significantly improves downstream agricultural outcomes. Studies in this same Indian eastern state show that access to the advisory service increases crop yields by 2-7%, total harvest by 4-9%, and reduces severe crop loss by 10-21% (Cole et al., 2025; Harigaya and Zhu, 2025). Harigaya and Zhu (2025) documents that even disadvantaged farmers with low education, low wealth, and poor infrastructure who engage less with the service still benefit substantially, experiencing a 5% increase in harvest and a 14% reduction in crop loss.

Although these papers show positive average treatment effects, our targeted policies prioritize some types of farmers over others in order to improve overall pickup rates. We examine whether targeting can effectively prioritize farmers most likely to benefit in terms of downstream outcomes such as agricultural yields. Using data from Cole and Fernando (2021), where the authors evaluated this same type of service in Gujarat, we estimate heterogeneous treatment effects (of access to agricultural advisory) on yield growth using causal forests. Large landholders experience estimated yield gains 5 times higher than small landholders (12.6pp vs 2.5pp), although this difference is imprecisely estimated due to small sample size and high variance in agricultural outcomes. We then bring this result into our setting and compare two targeting approaches: (1) equal-weight targeted policy that maximizes overall pickup, and (2) prioritized-weight targeted policy that assigns large landholders 20-fold higher weight in the constrained optimization objective. Using contemporaneous off-policy evaluation, the prioritized policy increases large landholder pickup by 1.2 percentage points (from 29.0% to 30.2%) while reducing small landholder pickup by 0.6 percentage points (from 32.2% to 31.6%). Combining these pickup changes with yield gains and weighting by the proportion of each group, the prioritization contributes +0.032pp to expected aggregate yields from increased large landholder engagement but reduces expected yields by -0.011pp due to decreased small landholder engagement. The limited improvement in high-benefit farmer engagement reveals practical constraints: large landholders exhibit systematically lower baseline pickup rates, possibly reflecting higher opportunity costs of time, suggesting that call timing optimization alone provides limited scope for benefit-based prioritization. Alternative targeting strategies beyond timing (e.g.,

targeted message content) may be needed to more effectively engage high-benefit farmer segments.

We make several substantive contributions by designing and evaluating these targeted policies in a novel context. First, while targeted marketing has been developed almost exclusively in sales contexts in developed countries, we show that these tools can be used productively in low-resource environments, a setting where there is growing recognition that digital information delivery can be highly valuable. Second, the non-profit with which we work has a novel objective function that may prioritize outreach for specific groups, such as women. We devise methods for accommodating such "distributional aware preferences" into targeted policy design, and, perhaps even more importantly, demonstrate that we can calculate the marginal "cost" of pursuing such goals in terms of expected reduction in average outcomes (e.g., how many fewer male farmers will be reached, and the associated reduction in overall pickup rates, in order to increase engagement of women farmers), depicting these trade-offs in a decision frontier. We argue that this frontier may be an important object of interest for organizational decision-making.

Second, operating in resource-limited environments requires solving the challenge of targeted policy under constraints. We therefore integrate a mixed-integer programming problem that optimizes targeting in an operationally feasible manner. This method also allows us to estimate the shadow value of relaxing the bandwidth constraint, thereby enabling the non-profit to reach more women and men.

Third, we extend constrained optimization to incorporate prediction uncertainty and supply-side shocks. We use conformal prediction to quantify group-specific uncertainty and penalize assignments where predictions are unreliable. In contemporaneous off-policy evaluation, the uncertainty penalty reduces predicted pickup rates from 32% to 29.2%, suggesting limited overfitting. Our framework also accommodates bandwidth shocks (server maintenance, holidays, operational disruptions) by reallocating farmers to available call times when capacity constraints change unexpectedly for certain call times.

The paper proceeds as follows. In the next section, we identify gaps in the literature and explain how our approach addresses them. Section 3 introduces the framework's context and setup and defines key terminology. Section 4 shows there is variation in pick-up rates over time, and provides descriptive statistics. Section 5 presents our main contemporaneous off-policy and on-policy results. Section 6 demonstrates how our framework can incorporate transfer learning to account for temporal drift and can be used flexibly to evaluate the efficacy and trade-offs associated with prioritizing specific groups, such as women or farmers for whom estimated yield gains are greatest. Section 7 concludes.

2 Related Work

Our work contributes to a large, increasingly sophisticated body of literature on designing targeted policies to maximize the delivery of customized marketing messages or offers. Building on the off-policy evaluation framework of Simester et al. (2020a) (detailed in Section 1), this approach is

applied across diverse marketing contexts: Ascarza (2018) compares predictive targeting to targeting based on estimated treatment effects using data from an experiment designed to improve customer retention; Athey et al. (2025b) compares alternative approaches to targeting using off-policy evaluation in an application to nudging students to fill out financial aid forms with text message nudges; Rafieian and Yoganarasimhan (2021) evaluates mobile ad targeting policies counterfactually; Ellickson et al. (2023) estimates heterogeneous effects of targeted email promotion, and estimates gains from a targeted policy using an off-policy evaluation. Yoganarasimhan et al. (2023) uses off-policy evaluation to compare personalized free trial policies and finds that targeted policies designed using outcome models estimated via LASSO perform best relative to other methods on short-term conversion outcomes. Moreover, policies optimized for short-term conversions perform as well as, or better than, policies directly optimized for long-term outcomes when evaluated on long-term retention and revenue.

A smaller literature explores the robustness of policy design using both off-policy and on-policy methods. Simester et al. (2020b) estimates outcome models using seven ML methods for customer prospecting at a US retailer, designs targeted policies based on these models, and evaluates the resulting policies in a prospective on-policy experiment; they find that targeted policies designed using outcome models estimated via LASSO (which we also use for outcome model estimation) perform best out of sample when the training data has substantial information. Interestingly, only one of the seven targeted policies outperforms a "simple" policy of a constant discount. Simester et al. (2020b) describes how covariate shift (changes in the distribution of covariates over time) and concept shift (which we refer to as temporal drift[3]) affect performance. Yang et al. (2024) and Dubé and Misra (2023) each use two experiments; the first includes random assignment of offers and is used to design a targeted policy, which they evaluate on data in a subsequent period. Dubé and Misra (2023) shows substantial gains from targeted pricing. Yang et al. (2024) reports gains in Boston Globe subscriber revenue and addresses potential ongoing temporal variation through continuous randomized assignment, enabling re-optimization with bootstrapped Thompson sampling, though it does not report how much this approach affects policy design or performance.

We advance the policy design literature by taking seriously the possibility that "one-shot" off-policy design or two-period experiments with off-policy followed by on-policy may not produce durable, robust policies. Our six-period experimental design enables diagnosis of temporal instability that shorter designs cannot detect (see Section 6.2 for empirical findings). We demonstrate diagnostic tools for identifying temporal drift through iterative experimentation and develop data-driven methods based on transfer learning to address it.

Our work extends the unconstrained recommendation systems literature (Agrawal et al., 2025; Athey et al., 2024; Rafieian and Yoganarasimhan, 2021; Zhou and Zou, 2023) to settings with operational capacity constraints. Our partner organization's resource constraints limit the number of farmers that can be contacted per hour, creating a constrained optimization problem with 91 call time capacity constraints and hundreds of thousands of individual farmer constraints (each

[3]See footnote 2 in Section 1.

farmer may be assigned only one call time). We develop a two-step solution: predict individual-level treatment effects, then solve the constrained assignment problem using mixed integer programming. Lu et al. (2025) contemporaneously develops optimization methods for targeted policies under constraints and demonstrates off-policy evaluation of different constraint specifications. Our work complements theirs by demonstrating large-scale deployment (over one million farmers) in a real-world setting.

Second, distributional goals may also require accommodation in policy design. Methodologically, we develop a framework for incorporating distributional preferences into constrained optimization, enabling organizations to prioritize specific subgroups while quantifying tradeoffs in terms of expected outcomes. We construct decision frontiers providing transparent decision support for stakeholders evaluating preferential policies (Athey, 2025; Athey et al., 2024).[4]

Empirically, we apply this framework to our setting where the NGO receives funding from organizations seeking to improve outcomes for priority groups such as women farmers. We show how these tradeoff frontiers shift when capacity constraints are relaxed, revealing that expanding bandwidth creates more room to pursue distributional goals without sacrificing overall performance.

Our approach of optimizing short-term engagement (pickup rates) rather than long-term outcomes (yields, profits) builds on marketing literature demonstrating that early-stage surrogate optimization can improve downstream outcomes when the causal path runs through initial engagement (Hu et al., 2014; Li and Kannan, 2014; Sahni et al., 2018; Wei et al., 2024; Yoganarasimhan et al., 2023). Athey et al. (2025a) and Yang et al. (2024) formalize the surrogacy assumptions required for short-term optimization to translate into long-term improvements. We extend this literature by testing whether targeting based on estimated downstream heterogeneity can overcome engagement barriers (see Section 6.4 for empirical findings).

Finally, we contribute to a growing literature that emphasizes the social value of delivery of information in emerging markets, in sectors such as public health, education, labor, and agriculture.[5] This paper demonstrates the feasibility of implementing sophisticated targeting policies at a substantial scale in a low-income population: over 1 million farmers across 91 treatment arms.

3 Setup and Experimental Design

To investigate the impact of targeted policy on call timing, we conducted a multistage experiment with a phone-based agricultural extension service in India. This digital extension service, launched in 2018, was developed and implemented by an NGO in collaboration with an Indian state government.[6] By the end of 2021, it served 1.3 million smallholder farmers throughout the state with a

[4]This contribution is related to, but distinct from, Ascarza and Israeli (2022) BEAT (Bias Elimination via Adaptive Trees), which seeks to create "fair" policies to eliminate heterogeneity correlated with protected attributes by changing the causal tree splitting criteria.

[5]The benefits of automated delivery are demonstrated in public health (Araya et al., 2021; Lee et al., 2021), education (Agrawal et al., 2025; Rodriguez-Segura, 2022), labor (Dammert et al., 2015), and agriculture (Cole and Fernando, 2021; Fabregas et al., 2019).

[6]Specifically, this digital extension service was developed and implemented by Precision Development, with support from the Abdul Latif Jameel Poverty Action Lab, in partnership with an Indian state government. The Bill & Melinda

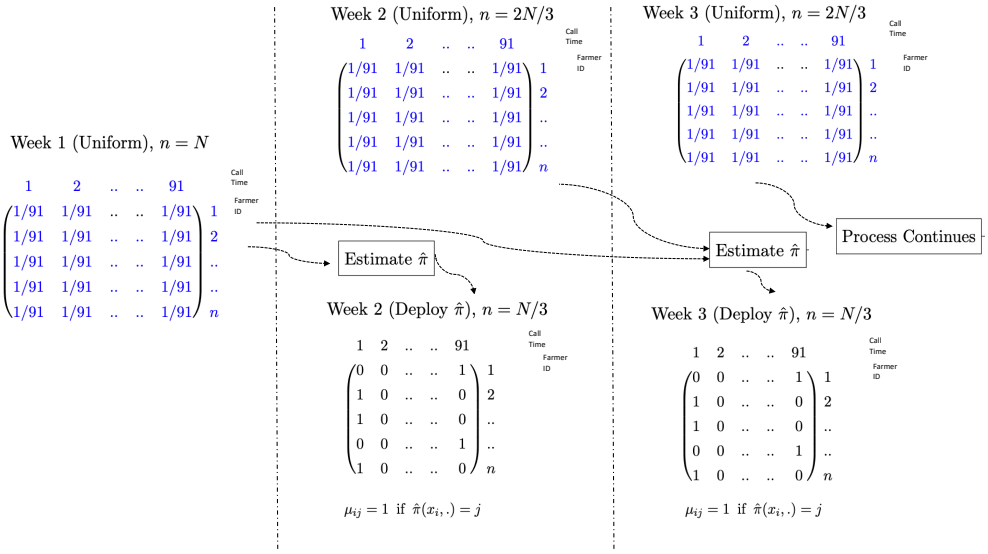
two-way, mobile-phone-based platform and a live call center. This service was transferred to the State Government at the end of 2021. It expanded rapidly; by 2024, it had grown to serve 6.8 million enrolled farmers.[7]

3.1 Setup

The agricultural extension service provided customized advisory on crops, livestock, and fisheries. The customization was done using farmer covariate information (e.g., language, location, crop, water management) and agricultural data (e.g., weather forecasts, market information, pest/disease outbreaks). Users of this service received agricultural information through three channels: 1) weekly interactive voice response (IVR) calls that provided farmers with customized farming advisories timed to the crop calendar (outbound calls); 2) an IVR platform that farmers could call in to listen to content from an advisory library and record their questions; and 3) a call center where farmers could call in and ask agriculture-related questions (inbound calls). Local agronomists answered inbound calls by sending recorded responses within 48 hours.

Our primary outcome variable is binary call pickup (outbound calls), depicted as Stage 2 in Figure 1. As discussed in Section 6.4, we optimize pickup rather than downstream outcomes because the causal path runs through initial engagement: without pickup ($Y_{it} = 0$), the policy cannot affect downstream outcomes. Rigorous impact evaluations have demonstrated that access to this digital service significantly improves farmers' downstream outcomes (Cole and Fernando, 2021; Cole et al., 2025; Harigaya and Zhu, 2025).

Figure 2: Two Data Collection Methods for 6 Experiments



We conducted six experiments over six weeks in October and November 2021 ($t \in \mathcal{T} \equiv$

Gates Foundation provided financial support.

[7] Appendix A1 provides additional information on the agricultural advisory service and the messages sent out through the call center.

$\{1, 2, 3, .., 6\}$). Figure 2 shows the data collection process for every week of the experiment. The intervention in this experiment consisted of deploying two data-collection methods. The data collection method determined the probability (μ_{ij}) that farmer $i \in \{1, 2, .., N_t\}$ was called in call time $j \in \mathcal{J} = 1, .., 91$. The first data collection method, which we refer to as "uniform randomization/uniform policy," assigned each farmer i to each of 91 call times j with equal probability. The second data collection method, which we call a "targeted policy," was deterministic and depended on the observable characteristics of a farmer x_i. The targeted policy π maps farmer covariates x_i to call times j. Thus, the probability that farmer i was allocated to call time j was 1 if $\pi(x_i, \cdot) = j$.

In selected weeks (weeks 2, 4, 5, and 6), we conducted an on-policy evaluation comparing the targeted policy with uniform randomization. In these weeks, farmers were randomized (with constant probability) to either uniform randomization or a targeted policy. Different targeted policies were used in different weeks. In each iteration, data collected under the uniform policy from previous weeks were used to estimate an outcome model and to design a targeted policy for the subsequent week.

Table 1: Experiment Weeks in October and November 2021

Week No.	Date	Uniform Randomization		Targeted Policy	
		Sample Name	Sample Size	Sample Name	Sample Size
1	Oct 5-10*	$N_{1,\mathcal{U}}$	881,891	–	–
2	Oct 18-24	$N_{2,\mathcal{U}}$	616,656	$N_{2,\hat{\pi}_A}$	265,188
3	Oct 25-31	$N_{3,\mathcal{U}}$	879,109	–	–
4	Nov 1,Nov 4-7*	$N_{4,\mathcal{U}}$	707,644	$N_{4,\hat{\pi}_B}$	167,995
5	Nov 8-14	$N_{5,\mathcal{U}}$	624,801	$N_{5,\hat{\pi}_C}$	234,276
6	Nov 17-23	$N_{6,\mathcal{U}}$	587,608	$N_{6,\hat{\pi}_D}$	227,386

Notes: *In these weeks, data was collected for only the indicated days of the week and not the entire week. $N_{t,\mathcal{U}}$ is used to denote the sample that is allocated to uniform randomization in week t. $N_{t,\hat{\pi}_e}$ is the sample that receives targeted policy $\hat{\pi}_e$ where $e \in \{A, B, C, D\}$ and t denotes the week. Our sample includes 880,000 farmers for whom we observe a complete set of covariates. The sample varies from around 800,000-880,000, depending on which farmers are designated to receive messages each week.

The weeks of this multistage experiment are given in Table 1. Every week, we sent an agricultural advisory push call to nearly 1 million farmers, prioritizing advisories on rice, one of the most important staple crops in this Indian state. Appendix A1 shows example scripts of agricultural advisory messages sent to farmers.

3.2 Targeted and Uniform Policies

This section defines the notation and framework for the analysis. $x_i \in \mathcal{X}$ is the vector of fixed covariates for individual i, such as gender, land size, access to irrigation, smartphone ownership, district of residence, and historical engagement with the service. We let F_X denote the distribution

of X with support in $\mathcal{X}$.

The outcome is the binary call pickup dummy Y_{it}^{obs}. Let $Y_{it}(j)$ be the potential outcome for farmer i in week t and call time j. The observed outcome can therefore be written in terms of potential outcomes as $Y_{it}^{\text{obs}} = \sum_{j=1}^{|\mathcal{J}|} \mathbb{1}(J_{it} = j) Y_{it}(j)$. The potential outcome $Y_{it}(j)$ follows a Bernoulli distribution with probability of pickup given by $\mu(x, j, t)$. Given a machine learning model $m \in \mathcal{M}$ and a training dataset $\mathcal{S}$, $\hat{\mu}(x, j, t; m, \mathcal{S})$ is the estimate of $\mu(x, j, t)$. Moreover, the population value for targeted policy π is defined as:

$$V(\pi) = \mathbb{E}_{X \sim F_X}[\mathbb{E}_Y[Y_t(\pi(X))|X]] \tag{1}$$

where t indexes evaluation week.

The optimal policy π^* is $\pi^*(\Pi) = \arg\max\{V(\pi) : \pi \in \Pi\}$ where Π is the set of policies that respect the budget constraint. The targeted policy designed using the estimated outcome model is denoted as $\hat{\pi}(X, t, \Pi; m, \mathcal{S})$. $\hat{\pi}$ returns the best call time based on $\hat{\mu}$:

$$\hat{\pi}(x; t, \Pi; m, \mathcal{S}) = \arg\max_j \hat{\mu}(x; j, t; m, \mathcal{S}), \quad \text{s.t. } \hat{\pi} \in \Pi \tag{2}$$

The uniform policy can be expressed as a function of 91 "fixed time policies" π^j where every farmer is called during call time j, $\pi^j(x) = j \ \forall x$ and the value of fixed time policies can be expressed as $V(\pi^j) = \mathbb{E}[Y_t(\pi^j(X))]$ where t indexes the evaluation week. Consequently, the population value for the uniform policy is $\bar{V}(\mathcal{U}) = \frac{1}{91} \sum_{j=1}^{91} V(\pi^j)$.

We use the uniform policy as our baseline (comparison) policy because it was collected simultaneously with the outcome data for the targeted policy. We do not use historical data as a baseline because it was collected during a different period, and, importantly, historically, the non-profit made very few calls during high-engagement weekend hours (see Figure A1 in Appendix), meaning we would not have a representative set of farmer characteristics across all call times with historical data.

4 Average Treatment Effect of Call Time on Pickup

Before analyzing the results of targeted policies, we present summary statistics on data collected using the uniform policy. Figure A2 shows farmers' pickup rates. The average pickup rate is around 31%, but there is substantial variation in pickup rates at different points of time during the week. The evening and morning hours have relatively higher pickup than the afternoon hours. Table A1 shows the average pickup in the afternoon is 1.3 (1.8) percentage points lower than the morning (evening). The uniform policy data helped us identify several hours on weekends as high-pickup periods, particularly the evening hours. While the NGO transmitted calls on all days of the week, it inadvertently failed to schedule many calls during these higher-pickup hours.

We present summary statistics on farmer covariates used in designing the targeted policy and examine variation in call pickup across farmer subgroups. Table A2 provides summary statistics

on the primary farmer characteristics: about 18% of the sample are female farmers, approximately 37.4% use smartphones, and slightly over 44% use irrigation systems. Information on residential districts is also included in the covariate data. In addition to these variables, we designed and evaluated alternative targeted policies using additional covariates, including marketing variables capturing recency and pickup frequency constructed from historical pickup data (collected prior to our experiment), and weather variables such as daily rainfall data (see Table A3).

Call pickup varies across subgroups. Figure A3 shows a persistent but time-varying gap in pickup rates for female farmers (0.29) relative to male farmers (0.32). The graph also suggests that the high pickup times often overlap for male and female farmers. Moreover, the average pickup rate among smartphone users is 4.1 percentage points lower than that among non-smartphone users; we also find that farmers at the higher end of the land-size distribution (i.e., wealthier) have lower pickup rates than those with less land (Figure A4).

5 Method for Designing and Evaluating a Targeted Policy

As outlined in Figure 2, we conducted a six-week iterative experiment (detailed in Section 3.1). Below, we describe the methods used to design and evaluate the targeted policies.

Design Targeted Policy Using Data from Uniform Policy This section describes how targeted policies are designed. First, we select an estimator for the outcome model. Among many reasonable machine learning models (e.g., LASSO, random forest) or matrix factorization methods (e.g., recommendation systems), we choose LASSO regularized regression as our primary method to estimate the outcome model due to a large number of potential treatments and covariates.[8]

Because the outcome model will be used to design the targeted policy, it is important that our estimator $\hat{\mu}(X, J, t; m, \mathcal{S})$ yields useful estimates of the difference in expected outcomes across treatment arms (call times). Since our training data $\mathcal{S}$ has uniform randomization, we are confident that farmer characteristics do not systematically vary with treatment assignment; the randomization ensures independence ($\{Y_i(j)\}_{j=1}^{J} \perp\!\!\!\perp J$) and thus unconfoundedness. The probability of pickup is modeled as:

$$\text{logit}(\mu_{ij}) = X_i\beta + \sum_{j'=1}^{91} \delta_{j'} D_{ij'} + \sum_{j'=1}^{91} \gamma_{j'} X_i D_{ij'} \tag{3}$$

where $D_{ij'} = \mathbb{1}(J_i = j')$ is an indicator for whether farmer i is assigned to treatment $j' \in \{1, \ldots, 91\}$. The coefficients $\delta_{j'}$ capture treatment main effects, while $\gamma_{j'}$ capture treatment-effect heterogeneity (interactions with farmer characteristics). The selection of the regularization parameter λ for LASSO is done using cross-validation. We do not penalize the coefficients on the treatment indicators $\{D_{ij'}\}$.

[8] Table A4 in the Appendix shows that targeted policies designed using outcome models estimated via random forest produce similar results to those using outcome models estimated via LASSO.

Below, we evaluate the expected benefit from assigning farmers using a targeted policy derived from $\hat{\mu}$. If the same data is used to design the targeted policy and evaluate its effectiveness, the result will likely overstate the benefits of the policy. We use cross-fitting to address this concern, dividing the data into K equal groups (folds). Denoting $k(i)$ the fold which contains farmer i, for each fold k, we estimate the model on all folds except k, and then use the model parameters to predict pickup for farmers in the k^{th} fold. We repeat this k times to generate $\hat{\mu}_{-k(i)}$ for all farmers in the sample.

The estimate of $\hat{\pi}_{-k}$ also depends on the constraints related to technological limits on the number of farmers that can be called in a call time. We account for this by using a two-step process. Step 1 uses LASSO and data collected using uniform randomization to estimate $\hat{\mu}$ with cross-fitting. Step 2 uses $\hat{\mu}$ and the budget constraints to allocate farmers to their best or near-best call times as shown below:

$$\begin{aligned} &\max_{z} \sum_{i=1}^{n} \sum_{j=1}^{J} z_{ij} \hat{\mu}_{-k(i)}(x_i, j, \cdot) \\ &s.t. \sum_{j} z_{ij} = 1 \ \ \forall i, \ \ \sum_{i} z_{ij} \leq b_j \ \ \forall j \ \text{ and } \ z_{ij} \in \{0, 1\} \end{aligned} \tag{4}$$

where b_j is the capacity limit on call time j. The decision variables are z_{ij}, which takes a value of 1 if farmer i is allocated to call time j and 0 otherwise. The decision variables can be represented in the form of a $N \times J$ matrix z, where each row corresponds to a farmer and each column represents a treatment arm.

The set of constraints is combined to generate a matrix A that has dimension $(N+J) \times (N*J)$. The first N rows of A ensure that every farmer i can be allocated to only one call time j. The remaining J rows ensure that the sum of allocation in each call time cannot exceed the total capacity, and, therefore, they should add to be less than or equal to b_j. The resulting set of equations for constraints is $A \times \text{vec}(z) \leq [1, \ldots, 1, b_1, \ldots, b_J]'$.[9] Since the decision variables are discrete, this can be set up as an integer programming problem. We use the Gurobi Parallel Mixed Integer Programming solver. This solver uses a linear-programming-based branch-and-bound algorithm for mixed integer programming problems. All the steps are summarized in the targeting algorithm 1 in the online appendix.[10]

Off-Policy Evaluation Off-policy evaluation (both contemporaneous and K-step ahead) is used to compute the value of targeted policies from a dataset that is collected using a different policy (Athey and Wager, 2021; Zhou et al., 2022). For instance, for this experiment, for weeks $t \in \{1, 2, .., 6\}$, a subset of users was selected to receive uniform randomization, which assigned call times uniformly at random. The data from this subset of users can be used to design and evaluate

[9] In our application, total capacity exceeds the number of farmers, so the inequality $\sum_j z_{ij} \leq 1$ binds at optimality, equivalent to the equality constraint in equation 4.

[10] More precisely, we estimate $\hat{\mu}_{-k}(x_i, j, \cdot)$ for all $j \in \{1, \ldots, 91\}$, generating a complete $n_k \times 91$ prediction matrix for use in the optimization problem.

the targeted policy ($\hat{\pi}$). The key idea is that under uniform randomization, $\frac{N}{91}$ farmers receive a treatment under $\mathcal{U}$ that matches their assignment under $\hat{\pi}$. The outcome of this subset of farmers can be used to estimate the value of targeted policy counterfactually (see Appendix A4 for details).

The estimators for the population objects of interest (defined in Section 3.1) are obtained using data collected from the uniform policy. All targeted policies are based on an underlying outcome model estimated via cross-fitting. The estimator for the value of $\hat{\pi}$ under the off-policy evaluation is denoted by $\hat{V}^{\text{Off}}(\hat{\pi}, \mathcal{S}^{\text{eval}})$.

$$\hat{V}^{\text{Off}}(\hat{\pi}, \mathcal{S}^{\text{eval}}) = \frac{\sum_{i \in \mathcal{S}^{\text{eval}}} Y_{it}^{\text{obs}} \mathbb{1}(\hat{\pi}(x_i, \cdot) = J_{it})}{\sum_{i \in \mathcal{S}^{\text{eval}}} \mathbb{1}(\hat{\pi}(x_i, \cdot) = J_{it})} \tag{5}$$

This is an unbiased estimator of the value of the targeted policy given randomized data and the use of cross-fitting in designing the policy. Note that after we design $\hat{\pi}$ using the uniform randomization data, there is a further important step of conducting the off-policy evaluation, as we want to counterfactually estimate the value of $\hat{\pi}$ prior to deployment. Without gains in the counterfactual evaluation relative to the baseline uniform policy, it would not make sense to deploy the policy $\hat{\pi}$ in a production setting.

As outlined in the Introduction, our experimental approach enables two types of off-policy evaluation. First, because policy design and evaluation are implemented using k-fold cross-fitting, the same week of data can be used to evaluate $\hat{\pi}$. We call this contemporaneous off-policy evaluation. Second, $\hat{\pi}$ can be designed using training data from week t and evaluated on week $t + k$, where $k > 0$. This second off-policy evaluation is referred to as K-step ahead off-policy evaluation in the paper.

On-Policy Evaluation Once we verified the off-policy evaluation showed significant gains from deploying $\hat{\pi}$ over the uniform policy, the next step was to deploy the targeted policy $\hat{\pi}$. In the subsequent week, we randomly divided the farmers into two groups. Group 1 received calls according to $\hat{\pi}$. Group 2 received calls according to the uniform policy. At the end of the week, pickup data was collected for both groups, as illustrated in Figure 2. This data enabled both on-policy and off-policy evaluation using a single dataset.

The population object of interest for the on-policy evaluation is defined as $\delta(\hat{\pi}, \mathcal{U}) = V(\hat{\pi}) - \bar{V}(\mathcal{U})$. Note that on-policy estimation is straightforward: it simply entails taking a sample mean on a dataset where the relevant policy was applied. For reference, we define $\hat{V}^{\text{On}}(\pi, \mathcal{S}) = \frac{\sum_{i \in \mathcal{S}} Y_{it}^{\text{obs}}}{|\mathcal{S}|}$. We note here that while off-policy evaluation enables comparison of multiple candidate policies without deployment, on-policy evaluation provides substantially higher statistical power due to the larger effective sample size. In off-policy evaluation, only the subset of farmers whose uniform assignment matches the targeted policy ($\hat{\pi}(x_i) = J_{it}$) contributes to the estimate, whereas on-policy evaluation uses the entire treatment group.

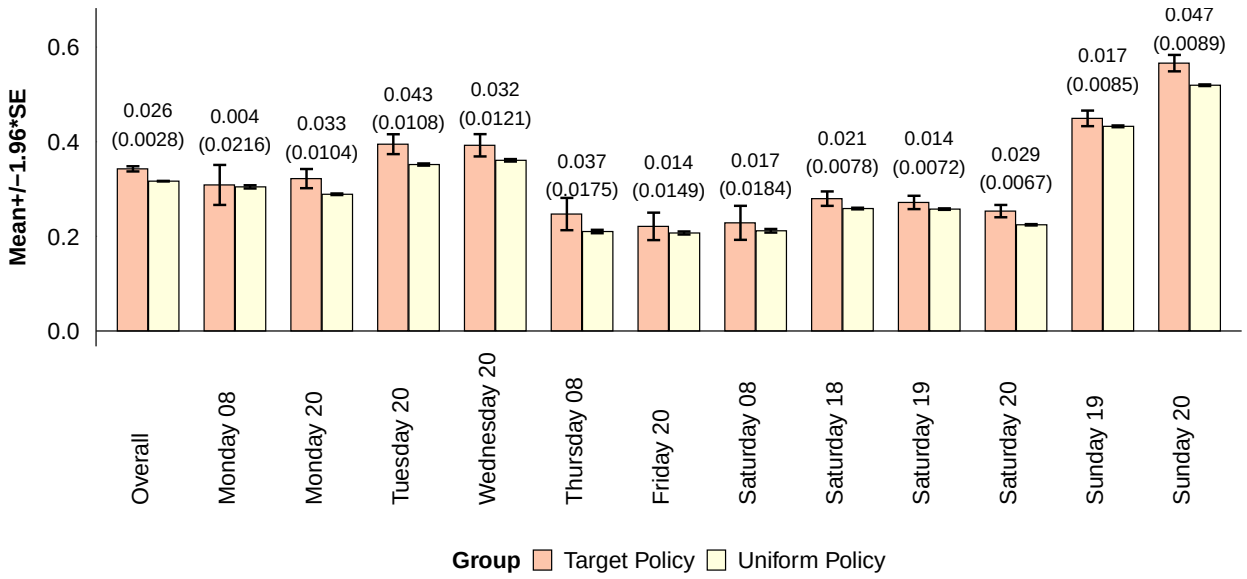
6 Results

6.1 Evaluation of Targeted Policies

Off-Policy Evaluation: We are interested in the average pickup rate achieved by the targeted policies we design. Recall our approach is to implement uniform policies and use the data generated by these calls to design targeted policies. We present results for the final targeted policy $\hat{\pi}_D$, which was designed using uniform policy data from weeks 1, 2, and 3.[11]

We first assess the out-of-sample prediction performance of the LASSO method used to estimate the outcome model, using uniform policy data from the first three weeks.[12] Panel (b) in Figure A5 shows the Receiver Operating Characteristics (ROC) plot for the binary pickup. The Area Under the Curve (AUC) is 0.683. An AUC of 0.683 indicates moderate predictive performance. This suggests that the model outperforms a random classifier by 36.6% in distinguishing between farmers who will and won't pick up the advisory's calls. Panel (c) shows the distribution of predicted responses by the class of the binary pickup variable. The predicted response for farmers with actual pickup as 1 is higher than those who did not pick up the calls. Lastly, Panel (d) in Figure A5 shows the calibration plot between the predicted and true pickup. We also provide calibration plots separately by farmer covariates in Figure A6.

Figure 3: Off-Policy Evaluation: $\hat{\pi}_D$ and Uniform



Notes: Sample includes data from uniform randomization in weeks 1, 2, and 3. Targeted policy estimates are computed using $\hat{V}(\hat{\pi}_D, \mathcal{S}^{\text{eval}})$ (Section 5). Uniform policy estimates are computed using the sample mean (Section 5). The estimate for the differences between $\hat{V}(\hat{\pi}_D, \mathcal{S}^{\text{eval}})$ and $\hat{V}(\mathcal{U}, \cdot)$ is displayed at the top of the bars and the standard error of this difference is in parentheses.

[11] Three other targeted policies ($\hat{\pi}_A$, $\hat{\pi}_B$, $\hat{\pi}_C$) were also designed and deployed in weeks 2, 4, and 5 as we iteratively refined our approach. Details on these intermediate policies are provided in Appendix A5.

[12] Panel (a) in Figure A5 shows the cross-validation error corresponding to different values of λ.

Following Algorithm 1, we design targeted policy $\hat{\pi}_D$ using uniform policy data from weeks 1, 2, and 3. We estimate its value using contemporaneous off-policy evaluation via the estimator $\hat{V}^{\text{Off}}(\hat{\pi}_D, \mathcal{S}^{\text{eval}})$ (defined in Section 5). Figure 3 shows that there are substantial gains (2.6 pp or 8%) to calling farmers according to $\hat{\pi}_D$ over the baseline of uniform policy. To provide a sense of the gains for different targeted call times, Figure 3 also reports the off-policy gains for the 12 most popular call times (which comprise 91% of the targeted policy sample).

Table 2: On-Policy Evaluation in Week 6

Data Collection-policy	$\hat{\pi}_{\text{D}}$	Uniform	Difference
Outcome Variable			
	All Farmers		
First Call	0.313	0.309	0.004
	[0.001]	[0.001]	[0.001]
	Female Farmers		
First Call	0.2971	0.2901	0.0069
	[0.0023]	[0.0014]	[0.0027]
	Male Farmers		
First Call	0.3167	0.3135	0.0032
	[0.0011]	[0.0007]	[0.0013]
N	227,386	587,608	

Notes: On-policy evaluation estimates use sample means for each of the two data collection mechanisms from Week 6.

Next, we explore whether estimating the outcome model with additional covariates improves targeted policy performance. We incorporate the additional marketing and weather variables shown in Table A3. To assess the model's predictive accuracy, we repeat the steps above and compute the ROC curve and AUC. The AUC is 0.712 when we incorporate these additional measures from the historical data. The targeted policy designed using this outcome model produces gains close to 8%, which is comparable to the gains observed using $\hat{\pi}_D$.[13]

Additionally, we estimate an alternative outcome model using random forest and design a targeted policy based on this model. Table A4 in the Appendix compares gains from targeted policies designed using random forest-estimated outcome models versus LASSO-estimated outcome models. Using contemporaneous off-policy evaluation, the random forest-based targeted policy achieves 2.0 pp gains relative to the uniform policy, compared to 2.6 pp for the LASSO-based targeted policy. However, this difference is not statistically significant at conventional levels.[14]

As a final robustness check, we vary the number of folds used in cross-validation for both outcome model estimation and policy evaluation. Our primary analysis uses three-fold cross-fitting; we additionally test two-fold and four-fold cross-fitting. Table A5 shows that the value of the targeted

[13]The additional marketing variables (frequency and recency) and the rainfall variable were added as a response to an anonymous reviewer as an additional robustness check.

[14]The AUC for the random forest outcome model is comparable to that of the LASSO outcome model, indicating similar predictive accuracy.

policy remains stable across different fold specifications, demonstrating that our results are robust to the choice of cross-fitting folds.

On Policy Evaluation of Targeted Policy: As part of the experiment, the targeted policy was designed, deployed, and evaluated; we iteratively updated the targeted policy as we collected additional uniform policy data. We also refined some details of our approach. For example, in week 2, we used 21 call times (morning, afternoon, and evening) over seven days, subsequently shifting to 91 call times (13 hour blocks per day); in other weeks, holidays or other real-world constraints affected the set of days in which we could implement policies. These details are described in Appendix A5.

We now present results from the final week of our study, in which we conduct an on-policy evaluation of $\hat{\pi}_D$.[15] In the final week, we randomly assigned farmers into two groups. The first group was assigned targeted policy $\hat{\pi}_D$, and the second group was called according to the uniform policy. We compare the average pickup for the two groups in Table 2. Table 2 shows the difference between the sample mean of farmers who received calls according to $\hat{\pi}_D$ and those that got called according to the uniform policy. The differences are shown for the overall sample as well as for the sample of female and male farmers. Surprisingly, there is only a 0.4 percentage point increase in overall pickup relative to the control group (30.9 pp); this gain is much smaller than the 2.6 percentage point gains predicted by contemporaneous off-policy evaluation (Figure 3). In the next subsection, we investigate the factors that cause this discrepancy and then propose approaches to improve the robustness of designed policies.[16]

6.2 Robust Targeted Policy

6.2.1 Transfer Learning and Sieve Estimator

A robust policy should perform well not only out-of-sample in data from the period it is designed (contemporaneous off-policy evaluation), but also out-of-sample in subsequent weeks when it is prospectively deployed (on-policy and K-step ahead off-policy). However, if the farmer's behavior changes over time, a policy designer may face a trade-off: placing greater weight on more recent data may yield more accurate predictions, at the cost of using less information from earlier weeks.

In this section, we leverage the availability of six weeks of data generated under the uniform policy. This allows us to detect temporal drift in farmer response patterns, which is difficult or impossible to detect with single- or two-period designs. This setting is conceptually similar to

[15] Appendix A5 provides details on the intermediate policies deployed during the experiment in weeks 2, 4, and 5, when we were developing and iterating the targeted policy to test multiple components such as granularity of treatment arms, bandwidth constraints, and the weeks of uniform policy data used for designing the targeted policy.

[16] We used uniform policy data to evaluate the targeted policy in week 6 of the experiment and test if the on-policy underperformance also shows up when we use the K-step ahead off-policy estimator. Table A6 in the Appendix shows that the gains obtained with K-step-ahead off-policy evaluation in week 6 are comparable to those obtained with on-policy evaluation. Hence, the underperformance cannot be attributed to estimator variance between off-policy and on-policy. An approach sometimes useful in these contexts is doubly robust augmented inverse propensity weighting. We do not use this here because with uniform randomization and known propensity scores, the standard motivation for AIPW does not apply.

transfer learning problems, in which a researcher has access to multiple datasets (auxiliary datasets) that may or may not be informative for estimating an outcome model on the target dataset. We build on recent work by Bastani (2021); Li et al. (2022), which demonstrated how transfer learning can be used for prediction problems. We formally adapt the ideas from Li et al. (2022), which is concerned with prediction accuracy, to the context of policy learning and evaluation. Our objective is to determine which treatment to assign to each farmer to maximize call pickup rates. This difference necessitates several modifications to Li et al. (2022) method. First, the objective changes from minimizing prediction error to maximizing the number of farmers reached. Second, we modify the model to account for treatment-effect heterogeneity and to correct for temporal drifts. Third, we handle 91 high-dimensional treatment arms using sieve basis functions to flexibly model treatment-effect heterogeneity.

A useful first step to assess the utility of this approach is to conduct out-of-sample diagnostics to measure the ranking stability of the best call times for farmers over the six-week period. Table A7 shows targeted treatment assignment stability, while Table A8 shows top-K treatment overlap under two distinct scenarios.

First, we estimate outcome models using weeks 1 and 2 and identify each farmer's best call time based on predicted pickup rates using data from week 3. We then estimate a second outcome model using weeks 1, 2, and 3 and again identify best call times using week 5 data. Only 51.1% of farmers receive the same recommended call time across the two models.[17] This demonstrates substantial temporal drifts as engagement patterns evolve. When we examine the top-5 recommended call times (rather than only the single best), overlap increases to 63.1%, rising to 79.2% for top-20, indicating that while the single best recommendation shifts dramatically, the broader set of good options remains more stable.

We conduct a similar diagnostic in a second scenario. We estimate outcome models for week 6 using two different training sets: weeks 3-5 versus week 5 only. The two models recommend the same best call time for only 20.0% of farmers. This striking instability demonstrates that predicted best call times are highly sensitive to which historical weeks are included when estimating the outcome model. The top-K overlap shows less disagreement (47.6% for K=5, jumping to 65.1% for K=20), again suggesting the models agree on which call times are generally poor but disagree substantially on which are best.

These diagnostics reveal two sources of instability that justify our transfer learning approach: (1) treatment effects genuinely change over time (temporal drifts), and (2) naive pooling of historical data can produce unreliable predictions.

For our transfer learning analysis, we note that the six weeks of uniformly randomized data are structured as a temporal sequence. We treat weeks 1-4 as auxiliary samples, week 5 as the target training data, and week 6 for policy evaluation (K-step ahead off-policy evaluation).[18] Let $D_{ij}^{(t)}$

[17] We compare models on week 5 predictions rather than week 4, as in week 4, the server did not send calls for all 91 call times due to routine maintenance.

[18] We use this setup with the full six weeks as it provides multiple auxiliary data sets, as well as a separate week for target data and another for evaluation. Such a complete sequence is not possible if fewer than six weeks are used in

indicate whether the farmer i is assigned to the call time j in week t. Formally, $D_{ij}^{(t)} = \mathbb{1}(J_{it} = j)$. The outcome model for each week is $\text{logit}(\mu_{ij}^{(t)}) = X_i\beta^{(t)} + \sum_{j=1}^{91} \delta_j^{(t)} D_{ij}^{(t)} + \sum_{j=1}^{91} \gamma_j^{(t)} X_i D_{ij}^{(t)}$ where superscript t indexes the week of the experiment and the target week parameters are $(\beta^{(5)}, \delta^{(5)}, \gamma^{(5)})$. The contrast vectors capture the temporal variation and are denoted by $\Delta\beta^{(t)} = \beta^{(5)} - \beta^{(t)}, \Delta\delta^{(t)} = \delta^{(5)} - \delta^{(t)}, \Delta\gamma^{(t)} = \gamma^{(5)} - \gamma^{(t)}$. Unlike in a prediction setup, this paper's setting incorporates drift in treatment-effect heterogeneity.

To understand how temporal drift affects policy performance, we establish the following decomposition:

Lemma 1 (Policy Value Decomposition Under Temporal Drift) *Consider a policy $\hat{\pi}$ estimated using data from auxiliary weeks $t \in \mathcal{T}_{aux}$ with weights w_t satisfying $\sum_{t\in\mathcal{T}_{aux}} w_t = 1$. The policy value in target week s can be decomposed as:*

$$V_s(\hat{\pi}) = \sum_{t\in\mathcal{T}_{aux}} w_t V_t(\hat{\pi}) + \sum_{t\in\mathcal{T}_{aux}} w_t \mathbb{E}_X[\mu(X, \hat{\pi}(X), s) - \mu(X, \hat{\pi}(X), t)] \tag{6}$$

where the first term represents the weighted average performance of policy $\hat{\pi}$ in auxiliary weeks and the second term captures temporal drift in farmer response functions.

Proof. See Appendix A7.

Lemma 1 reveals the fundamental tradeoff in policy learning under temporal drift. When there is no drift ($\mu(x, j, s) = \mu(x, j, t)$ for all x, j, t), the second term vanishes and equal weighting ($w_t = 1/|\mathcal{T}_{\text{aux}}|$) maximizes statistical efficiency by pooling all available data. However, when drift is present, the optimal weighting strategy must balance minimizing the temporal drift term against maintaining estimation precision. We adapt the transfer learning approach of Li et al. (2022) to select weights that address this tradeoff. In our application, $\mathcal{T}_{\text{aux}} = \{1, 2, 3, 4\}$ and $s = 5$. Evaluation is performed on week 6 data (K-step ahead off-policy). Below, we describe the algorithm for data-driven weight selection.

To identify which auxiliary datasets are informative about target week 5, a measure of similarity between weeks is needed. A marginal statistic is used to measure this similarity, capturing the association between the outcome and the covariates, treatments, and their interactions. Let $Z_i^{(t)}$ denote the full vector stacking covariates X_i, treatment indicators $D_{ij}^{(t)}$ for all j, and their interactions $X_i D_{ij}^{(t)}$ for farmer i in week t. The contrast in marginal statistics between week t and the target week 5 is:

$$\hat{\Delta}^{(t)} = \frac{1}{n_t} \sum_{i=1}^{N_t} Z_i^{(t)} y_i^{(t)} - \frac{1}{|I|} \sum_{i\in I} Z_i^{(5)} y_i^{(5)} \tag{7}$$

where $I \subset \{1, \ldots, N\}$ is a holdout subset from week 5 (after sample splitting), and $y_i^{(t)} \in \{0, 1\}$ is the pickup indicator. This vector $\hat{\Delta}^{(t)}$ has dimension equal to the total number of parameters in

our setting, because the experiment included two weeks during which uniform randomization was not applied across the 91 call times due to operational constraints. See discussion in Appendix A5.

$\theta^{(t)} = (\beta^{(t)}, \delta^{(t)}, \gamma^{(t)})$. Small values of $\|\hat{\Delta}^{(t)}\|$ indicate week t is similar to week 5 and is therefore informative for policy learning on target week 5, while large values indicate temporal drift.

We rank the four auxiliary weeks from most similar to least similar based on the magnitude of $\|\hat{\Delta}^{(t)}\|$. Weeks with smaller values exhibit pickup patterns more similar to week 5. This ranking allows us to construct five candidate sets based on cumulative similarity: $\hat{G}_0 = \emptyset$ (use no auxiliary data, estimate outcome model using LASSO only on target week 5), $\hat{G}_1$ (use only the most similar week), $\hat{G}_2$ (use the two most similar weeks), and so on up to $\hat{G}_4$ (use all four auxiliary weeks). For each candidate set $\hat{G}_\ell$, we estimate outcome model parameters $\hat{\theta}(\hat{G}_\ell)$ using a two-step procedure that leverages both auxiliary and target week data.

Step 1 (Initial estimation): Pool data from auxiliary weeks in $\hat{G}_\ell$ and estimate initial parameters $\tilde{\theta}^{(\ell)}$ using LASSO regression. This provides a starting point informed by the selected auxiliary weeks.

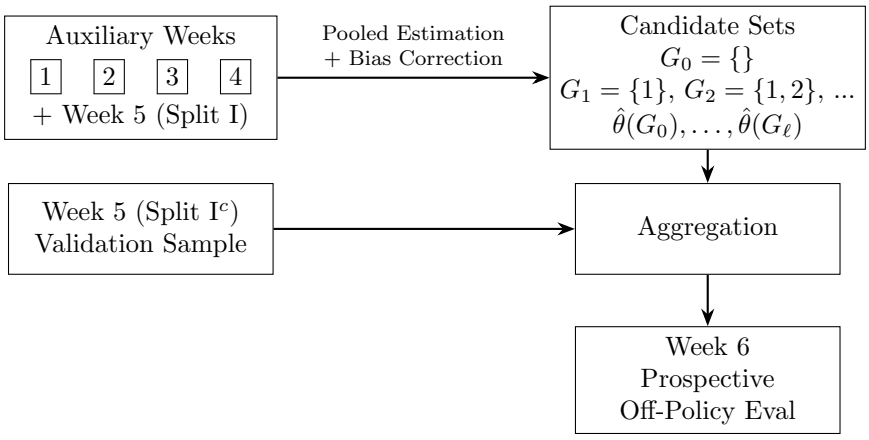
Step 2 (Bias correction): Correct for temporal drift by fitting LASSO on residuals from target week 5 data, obtaining bias correction $\hat{\Upsilon}^{(\ell)}$. The final estimator is $\hat{\theta}(\hat{G}_\ell) = \tilde{\theta}^{(\ell)} + \hat{\Upsilon}^{(\ell)}$.

Rather than selecting a single best candidate, we combine all candidates using data-driven weights. We evaluate each candidate's prediction accuracy on held-out Week 5 data and then compute a weighted average, with better-performing candidates receiving greater weight. This approach automatically adapts to temporal drift: when historical data are informative, multiple candidates receive weights; when drift is severe, the weight concentrates on the candidate using only recent data.

This two-step approach addresses the tradeoff identified in Lemma 1: auxiliary data improves precision (reduces variance) while target week data corrects for temporal drift (reduces bias). Importantly, the LASSO penalty applies to covariate main effects (β) and interactions (γ) but not to treatment indicators (δ_j). This complete transfer learning framework for policy learning is illustrated in Figure 4. This shows that the split I from week 5, combined with auxiliary weeks, yields candidate estimators via pooled estimation and bias correction. Split I^c, held out from all candidate construction, evaluates each candidate's prediction accuracy and determines aggregation weights. The final targeted policy is then designed for farmers in Split I^c using the weighted combination of candidates. The targeted policy is evaluated using week 6 data that was not used in any of the transfer-learning steps.

Simulation Validation. To validate our transfer learning approach, we conduct a Monte Carlo simulation study with 100 replications. Each simulation generates data for 50,000 farmers observed across six time periods with five available treatments (reduced the number of treatments for faster computation). Farmer covariates (gender and continuous characteristics) are fixed, whereas treatment assignments and binary outcomes (analogous to call pickup in our empirical application) vary across periods. We consider two scenarios: (1) a no-drift scenario where outcome model parameters remain stable across all periods, meaning there is no change in the best treatment across the multiple time periods, and (2) a strong-drift scenario where the best two (of five) treatments

Figure 4: Transfer Learning for Targeted Policy Design



shift from $\{2, 4\}$ to treatments $\{1, 3\}$ between periods 4 and 5, mimicking a regime shift. Both methods, i.e., equal weighting and transfer learning, are trained on data from periods 1–4 plus 40% of period 5 (matching our sample-splitting procedure). We evaluate policy performance in period 6 by computing the expected pickup rate under each designed policy, using the true outcome model parameters, thereby allowing comparison with the oracle policy that knows the true parameters.

Table 3: Simulation Results on Policy Performance Relative to Oracle

Method	Metric	No Drift	Strong Drift
Equal Weighting Policy			
	Policy Value	0.6069	0.5616
	Gap from Oracle	0.0016	0.1308
		[0.0012, 0.0019]	[0.1298, 0.1318]
TL Policy			
	Policy Value	0.6065	0.6921
	Gap from Oracle	0.0020	0.0003
		[0.0017, 0.0023]	[0.0002, 0.0004]
Oracle			
	Policy Value	0.6085	0.6924

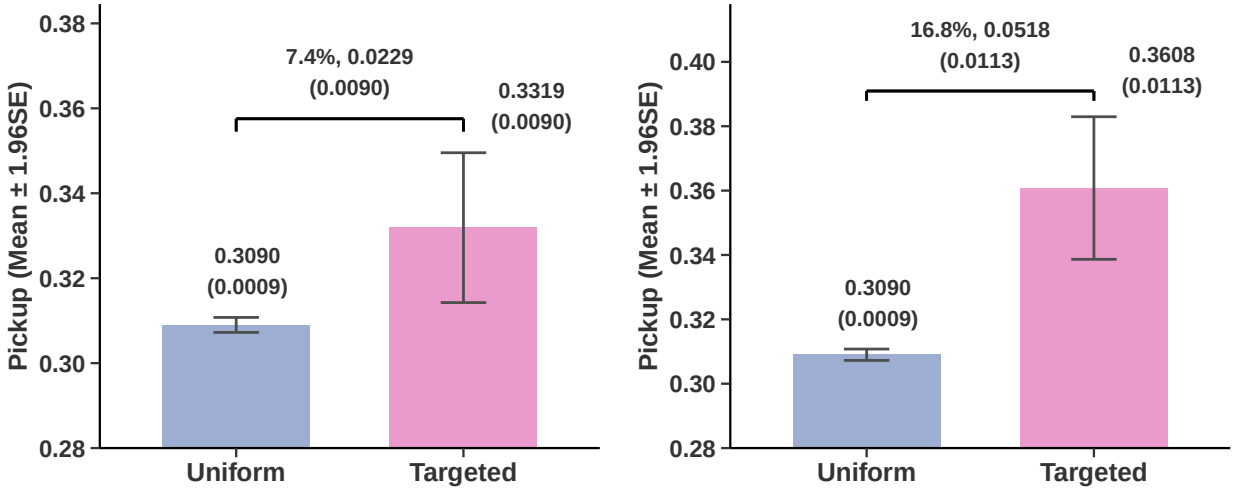
Notes: This table reports policy performance across 100 Monte Carlo simulations with 50,000 farmers each. Policy value represents the expected pickup rate under each method. Gap from oracle is the difference between oracle and estimated policy values (in probability units), with 95% confidence intervals shown in brackets.

Table 3 presents the results. In the no-drift scenario, both methods perform essentially identically, with gaps from the oracle of 0.16 and 0.20 percentage points (pp), respectively. This

demonstrates that transfer learning does not harm performance when parameters are stable. In the strong-drift scenario, however, the methods diverge dramatically. Equal weighting, which pools historical data without adaptation, yields a targeted policy that is 13.08 pp worse than optimal [95% CI: 12.98, 13.18]. In contrast, transfer learning achieves near-oracle performance with only a 0.03 pp gap [95% CI: 0.02, 0.04]. These results demonstrate that transfer learning provides substantial value in the face of temporal drift, while remaining robust when parameters are stable.

Results on Data. Figure 5(a) reports results for the TL-based targeted policy evaluated on real data. The targeted policy is designed using target week 5, with the weights for auxiliary weeks (1-4) determined using the sparsity indices described above. When evaluated in week 6, the transfer learning policy achieves a 7.4% improvement (2.29 percentage points) over the uniform policy, with the targeted policy achieving a pickup rate of 33.19% compared to 30.90% under uniform randomization (K-step ahead off-policy evaluation). The transfer-learning targeted policy correctly captures the temporal drift between the auxiliary weeks and the target week and corrects for the bias introduced by the regime shift.

Figure 5: Targeted Policy Gains over Uniform Policy



a. TL with 91 treatments

b. TL with Sieve Estimator for Treatments

Notes: This figure shows off-policy evaluation results on week 6 data using transfer learning methods, where weeks 1–4 are auxiliary weeks and week 5 is the target week. Panel (a) shows pickup rates under the targeted policy using 91 discrete treatment indicators. Panel (b) shows pickup rates using the sieve estimator approach with B-spline basis functions for hours and day-of-week indicators.

To further improve robustness, we modify the treatment indicators and instead use a sieve estimator for modeling the treatment heterogeneity. Precisely, let $h \in \{8, 9, \ldots, 21\}$ denote the hour of day and $d \in \{\text{Mon}, \ldots, \text{Sun}\}$ denote the day of week. The treatment effect function is modeled using a sieve basis. A cubic B-spline basis function is used for the continuous hour variable $\mathcal{B}_h = \{B_1(h), B_2(h), \ldots, B_{K_h}(h)\}$, where K_h denotes the number of cubic B-spline basis

functions. The basis functions provide a smooth flexible way to approximate the hour-of-day effect. Dummy indicators are used for the days of the week $\mathcal{D} = \{\mathbb{1}_{\text{Mon}}(d), \mathbb{1}_{\text{Tue}}(d), \ldots, \mathbb{1}_{\text{Sun}}(d)\}$. To allow the treatment effect of hour to vary by day of the week, we include the full tensor product of hour-by-day combinations.

$$\mathcal{B}_h \otimes \mathcal{D} = \{B_k(h) \cdot \mathbb{1}_{d'}(d) : k = 1, \ldots, K_h,\ d' \in \{\text{Mon}, \ldots, \text{Sun}\}\} \tag{8}$$

In addition, we include discrete indicators for structural breaks: Weekend $= \mathbb{1}\{d \in \{\text{Sat}, \text{Sun}\}\}$ and Evening $= \mathbb{1}\{h \geq 18\}$. These indicators capture discontinuities in farmer availability identified in the uniform-randomization data, where weekend and evening calls had higher pickup rates. The complete treatment basis is $\mathcal{T} = \mathcal{B}_h \cup \mathcal{D} \cup (\mathcal{B}_h \otimes \mathcal{D}) \cup \{\text{Weekend}, \text{Evening}\}$. With the modified treatment definition, the predicted outcome under call time (h, d) is:

$$\hat{\mu}_i(h, d) = \hat{\delta}' T(h, d) + \hat{\beta}' X_i + \hat{\gamma}'[X_i \otimes T(h, d)] \tag{9}$$

where $T(h, d)$ denotes the vector representation of the treatment basis $\mathcal{T}$ evaluated at call time (h, d).

The sieve approach reduces dimensionality for the interactions from $91 \times p$ to approximately $(8K_h + 9) \times p$, where p is the number of covariates. By exploiting smoothness in the hour effect, the estimator leverages similar patterns across neighboring time points, reducing estimation variance while allowing treatment effects to vary smoothly over hours rather than exhibiting changes between adjacent hours. When evaluated on week 6 (K-step ahead off-policy), the sieve estimator achieves 5.18 percentage points improvement (16.8%) over uniform policy, approximately double the 2.29 percentage point gain (7.4%) obtained using 91 treatment dummies (see Panel (b) in Figure 5).[19]

6.2.2 Follow-up Calls

The overall objective of this project is to maximize farmer engagement with the service. The policy was designed and deployed for the timing of the first call attempt each week to reach a farmer. In fact, if a farmer does not answer the phone, the dialing software automatically attempts a second call 24 hours later, and if that is not answered, a third and final attempt is made 24 hours later (Figure A7). While, in principle, it is possible to design and implement a targeted policy for the second and third calls, the nonprofit did not want to take on such operational complexity.

We can, however, use these follow-up calls to illustrate how they can mitigate some of the shocks from the first call. This section explores the efficacy of the rule of thumb of calling farmers 24 hours later and proposes a method for counterfactual evaluation of all three calls. We provide the first evidence for this mechanism using the overall pickup over the three calls for our on-policy evaluation of $\hat{\pi}_D$ (Table A9). We find the impact of targeted policy is higher when examining overall engagement by adding the pickup over call attempts 1, 2, and 3 than only examining the

[19] For the implementation, we select K_h using cross-validation, evaluating candidate values $K_h \in \{3, 5, 7, 9\}$ based on out-of-sample prediction error.

engagement over the first call. If there are shocks to the first call, follow-up calls can mitigate some of the cost.

Next, we provide evidence of the impact of the first-call targeted policy on overall engagement using contemporaneous off-policy evaluations. Figure A8 provides evidence that the benefit of targeting the first calls has benefits for overall farmer engagement. For this exercise, the targeted policy is designed using uniformly randomized data from the first calls in weeks 1 and 2. We evaluate the first-call policy (contemporaneous off-policy evaluation) not only on the initial call but also on subsequent follow-up calls. The sample used to evaluate the targeted policy comprises a subset of individuals in the evaluation set whose actual assignment matched the targeted policy according to the first call. The first two bars in Figure A8 show the gains of targeting the first call on the first call. Next, we add the pickup over the first and second calls for the evaluation sample. We continue to see that the pickup over the first two calls for the targeted policy group is higher than the pickup for the first and second calls for the uniform group (2 pp (SE=0.36)). Next, we add pickup over calls 1, 2, and 3. We see that the targeted policy engagement is higher than the uniform policy engagement (1.6 pp (SE=0.35)).

6.3 Distribution and Risk Aware Targeting System

Decision tool to Assess Higher Preference for Subgroups Organizations often prioritize specific groups in their optimization efforts for several reasons, such as donor preferences. Our system provides the organization with a decision tool to incorporate group-specific weights into the optimization and assess the trade-offs of these preferences.

As discussed in Section 4, there is a substantial overlap between the top pickup call times for male and female farmers, as seen in Panel (b) and (c) of Figure A3. This suggests that a targeted policy designed to maximize overall pickup rates may assign high-value slots to men; there may be a trade-off between maximizing overall pickup rates and prioritizing female farmers.

We demonstrate how weighting can be easily implemented in practice in our setting. To do so, we first draw a random sample of 600,000 farmers from the datasets collected using uniform randomization in weeks 1, 2, and 3 of the experiment.[20] We estimate $\hat{\mu}$ as before, modifying the constrained optimization step of Algorithm 1 to add weights as follows. Let G be the group indicator random variable with support in $\mathcal{G}$ and g be the realization of this random variable. X follows a mixture distribution $X|G = g \sim F_g$. The population objective function in this case is

$$V(\pi, w) = \sum_{g \in \mathcal{G}} w(g) \mathbb{E}_{X \sim F_g}[\mu(X, \pi(X), \cdot)], \tag{10}$$

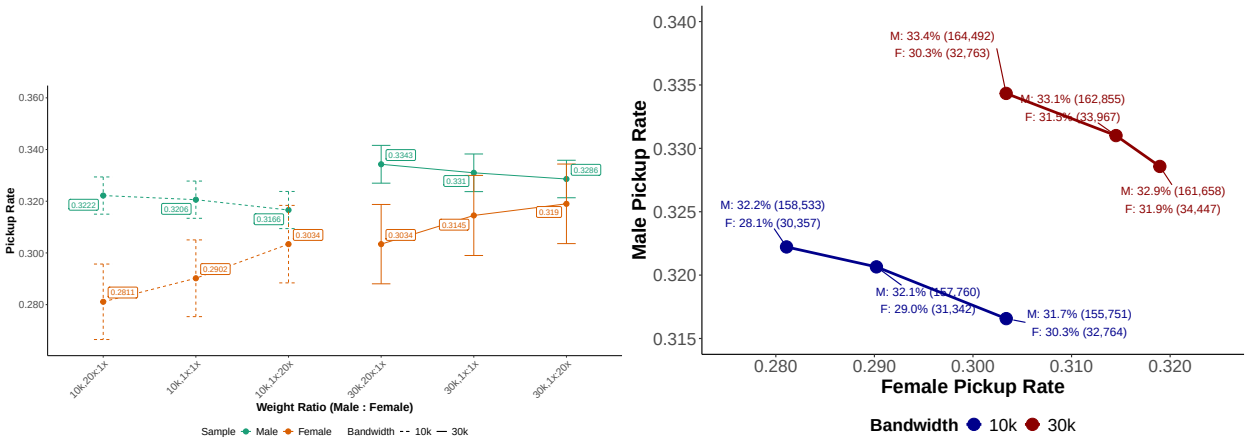
where w is the weight function $w : \mathcal{G} \to [0, 1]$ with $\sum_{g \in \mathcal{G}} w(g) = 1$.

For any given set of weights, we design a targeted policy for the corresponding weighted objective function. These alternative targeted policies are then evaluated using contemporaneous off-policy

[20] We identify unique farmers across weeks 1-3, randomly sample 600,000 of them, and retain all their uniform randomization observations from these weeks.

evaluation. We characterize the trade-offs by constructing a frontier that shows the allocations across different weighting schemes. Each point on this frontier represents an optimal allocation for a given weight ratio between male and female farmers; improving pickup rates for one group requires reducing them for the other when bandwidth remains constant.[21]

Figure 6: Constrained Optimization with Varying Weights (Male and Female Farmers)



a. Pickup Rates by Weight Scenario

b. Decision Frontiers

Notes: Panel (a) shows pickup rates for male and female farmers across three weight scenarios (20x:1x male: female, 1x:1x, and 1x:20x male:female) at two bandwidth levels (10,000 and 30,000 farmers per hour). Panel (b) presents the frontier showing the pickup rate and the number of male and female farmers reached by multiplying the pickup rate by the proportion of male and female farmers and the sample size.

Panel (a) in Figure 6 illustrates the pickup rates across three weighting scenarios for male and female farmers at two bandwidth levels. We examine three weight ratios: 20x:1x (heavily favoring males), 1x:1x (equal weights), and 1x:20x (heavily favoring females). The figure demonstrates two key findings. First, at a fixed bandwidth (shown separately for 10,000 and 30,000 farmers per hour), there exists a clear trade-off between reaching male and female farmers. Moving from equal weights (1x:1x) to female-favored weights (1x:20x) at 10,000 bandwidth increases female pickup rates while decreasing male pickup rates, as the constraints are tight and binding in the optimization. Second, expanding the bandwidth from 10,000 to 30,000 shifts the entire frontier outward, enabling Pareto improvements in which both male and female pickup rates increase simultaneously. We note that we assess these trade-offs counterfactually: in our 6-week experiment, the non-profit did not modify the bandwidths, as they are determined by long-term contracts and are therefore fixed in the short term.

Panel (b) in Figure 6 translates these pickup rates into the number of farmers reached, making the trade-offs more tangible for policy decisions. The frontier indicates that at a bandwidth of 10,000 with equal weights (1x:1x), the organization reaches approximately 31,342 female farmers and 157,760 male farmers. Moving to female-favored weights (1x:20x) enables reaching 1422 additional

[21] The bandwidth constraints have been scaled down because the analysis includes a smaller sample than 1.3 million farmers, and we assume scheduling only one message per farmer, which also excludes follow-up calls.

female farmers, but at the cost of reaching 2009 fewer male farmers. However, this difference between the two scenarios changes substantially when the sample size is expanded to 30,000 farmers. The frontier under relaxed constraints shifts out, and the trade-offs associated with favoring women farmers also change. We note that constructing the decision frontier requires evaluating multiple alternative targeted policies; due to experimental constraints, we evaluate these alternatives using contemporaneous off-policy evaluation rather than prospective on-policy deployment.

Risk Aware Targeting System The targeted policies designed and evaluated in this paper are based on point estimates of predicted pickups for each farmer-call-time combination. The transfer learning methods for policy design discussed in Section 6.2.1 address temporal drifts in outcome. However, prediction uncertainty can arise even in cross-sectional settings or in environments without temporal drifts. Some farmer groups may have high predicted engagement but also high variance (measured by the dispersion of residuals within group call time slots), which makes them risky targets for the limited number of high-engagement slots. In this section, we extend the constrained optimization framework to account for uncertainty in predicted pickup rates by incorporating conformal prediction methods.

We propose a conservative targeted policy that penalizes assignments to high-uncertainty farmer groups. The modified objective function is:

$$\max_{z} \sum_{i=1}^{n} \sum_{j=1}^{J} z_{ij} \left[\hat{\mu}^{(-k(i))}(x_i, j, \cdot) - \lambda \cdot \varepsilon_{g(i),j} \right] \tag{11}$$

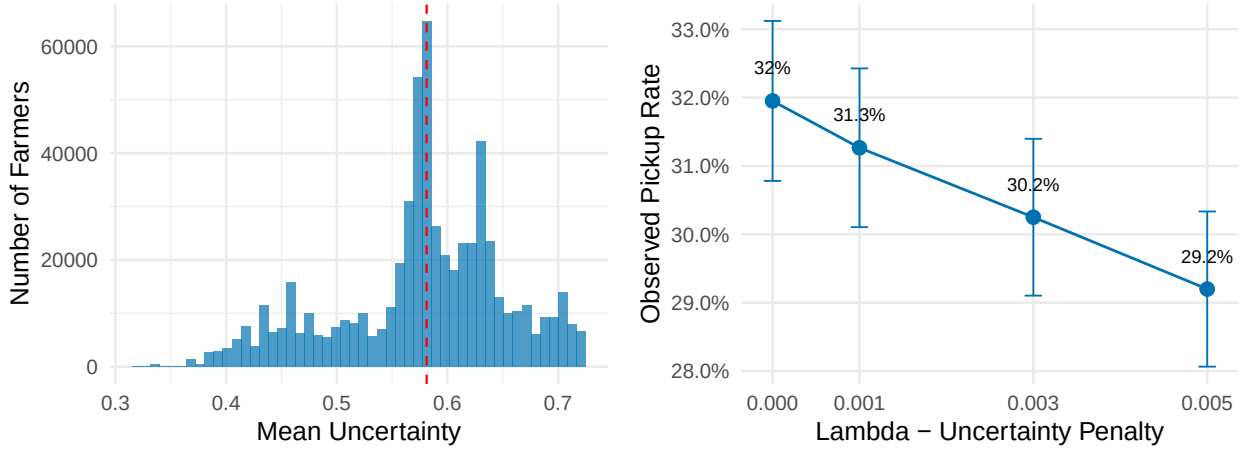
subject to the same constraints as in Equation 4, where $\varepsilon_{g(i),j}$ is the uncertainty measure for farmer i's group at call time j and $\lambda \in [0, 1]$ is a penalty parameter ($\lambda = 0$ recovers the original approach).

To estimate the uncertainty measure using conformal prediction, we compute absolute residuals $r_i = |Y_i - \hat{\mu}_{(-k(i))}(x_i, J_i, \cdot)|$, where Y_i is the binary pickup indicator and J_i is the assigned time slot for farmer i. Farmers are grouped by characteristics g (e.g., gender, land size). For each group-time combination, we construct the uncertainty measure as the $(1 - \alpha)$-quantile of absolute residuals: $\varepsilon_{g,j} = \text{Quantile}_{1-\alpha}(\{|r_i| : i \in \text{group } g, J_i = j\})$. Farmers in volatile groups (high residual variance) receive larger penalties for time slots where their group exhibits high prediction uncertainty.

We evaluate the uncertainty-penalized policy using contemporaneous off-policy evaluation on weeks 1-3. Figure A10 presents the results. Panel (a) shows substantial heterogeneity in pickup-prediction uncertainty across farmers, with a mean uncertainty ranging from 0.3 to 0.7. Panel (b) shows the tradeoff between the uncertainty penalty λ and pickup rates. At $\lambda = 0$ (no penalty), the policy achieves 32% pickup. As λ increases, pickup rates decline monotonically, reaching 29.2% at $\lambda = 0.005$. The uncertainty penalty can improve out-of-sample performance when the prediction model overfits. Our finding that the penalty reduces performance in contemporaneous off-policy evaluation suggests that there is no substantial overfitting in our predictions during this period.

Next, we extend the framework to accommodate bandwidth shocks from server maintenance, festivals, or holidays. The call center may need to shut down operations during specific call times,

Figure 7: Prediction Uncertainty and Conformal Prediction



a. Distribution of Prediction Uncertainty

b. Pickup Rate vs Uncertainty Penalty

Notes: Panel (a) shows the distribution of prediction uncertainty across farmers. Panel (b) shows pickup rates under different values of the uncertainty penalty parameter λ in off-policy evaluation.

requiring the optimization to redistribute calls accordingly. We modify the capacity constraints by introducing shutdown indicators $s_j \in \{0, 1\}$, where $s_j = 1$ if call time j is unavailable. The modified optimization problem becomes:

$$\begin{aligned} &\max_{z} \sum_{i=1}^{n} \sum_{j=1}^{J} z_{ij} \left[\hat{\mu}_{(-k(i))}(x_i, j, \cdot)\right] \\ &\text{subject to} \quad \sum_{j=1}^{J} z_{ij} = 1 \quad \forall i, \quad \sum_{i=1}^{n} z_{ij} \leq b_j \cdot (1 - s_j) \quad \forall j, \quad z_{ij} \in \{0, 1\} \quad \forall i, j \end{aligned} \tag{12}$$

When $s_j = 1$, the capacity constraint becomes $\sum_i z_{ij} \leq 0$, effectively blocking all assignments to that call time. Figure A9 illustrates this reallocation.

In summary, this subsection demonstrates that the targeted policy design methods developed and deployed in this paper can accommodate both prediction uncertainty (via conformal prediction) and bandwidth shocks (supply-side uncertainty).

6.4 Targeting Based on Heterogeneous Impacts on Agricultural Yields

Our optimization framework can incorporate weights to prioritize farmers most likely to benefit from downstream outcomes, such as agricultural yields, when provided the agricultural advisory service. To implement this, we first estimate heterogeneous treatment effects on yields using data from a previous impact evaluation. We use data from Cole and Fernando (2021), who conducted an impact evaluation study of the same type of advisory service in Gujarat, another state in India. Cole and Fernando (2021) reports only average treatment effects of advisory access on downstream agricultural outcomes such as yields and profit; we use their data and extend the analysis to identify

systematic heterogeneity in treatment effects on crop yield.

The conditional average treatment effect (CATEs) are estimated using causal forests (Wager and Athey, 2018). The outcome is the growth rate of crop yield. The treatment is a binary indicator for receiving access to the advisory service. Covariates include land size, farmer age and baseline input costs for fertilizer, irrigation, labor, and pesticides. Summary statistics for the outcome, treatment indicator, and covariates are provided in Table A10 (Appendix).

Figure A10 Panel (a) presents the Targeting Operator Characteristic (TOC) curve, which plots the gain from prioritizing farmers based on predicted treatment effects relative to random allocation. The curve lies substantially above the diagonal (random targeting), indicating that farmers predicted to benefit most from the advisory service indeed experience larger yield growth gains. The area under the TOC curve, the Rank-Weighted Average Treatment Effect (RATE), provides visual evidence of treatment-effect heterogeneity. The RATE for agricultural yield growth is 0.194 (SE = 0.053), indicating that prioritization based on CATEs would yield better outcomes than random allocation on yields.

To obtain valid out-of-sample CATE predictions and avoid overfitting, we implement 3-fold cross-fitting. We partition the sample into three equal folds, train separate causal forests on two folds (excluding the third), and predict CATEs for the held-out fold. This process is repeated three times, such that each farmer receives an out-of-fold CATE prediction. Each causal forest is trained with 10,000 trees, with hyperparameters (sample fraction, minimum node size, honesty fraction, and regularization parameters) tuned via cross-validation on the full sample and then applied to each fold.

Table A11 Panel (a) decomposes the heterogeneity by dividing farmers into quintiles based on their predicted CATEs and estimating treatment effects within each quintile using linear regression with heteroskedasticity-robust standard errors. Treatment effects for higher quintiles exceed those for the lowest quintile, with only differences between the highest and lowest quintiles being statistically significant at conventional levels. Figure A10 Panel (b) examines which farmer characteristics drive treatment effect heterogeneity by displaying average covariate values across CATE quintiles. The heatmap reveals that land size farmers in higher CATE quintiles operate with substantially larger farms (land size), and baseline costs for fertilizer, pesticides, and labor costs. This suggests that larger landholders with high baseline costs benefit more from agricultural advisory services in terms of yield growth. To quantify this relationship, we compare average treatment effects across land size groups using a median split.

Farmers with above-median land size (n = 1,013) experience cotton yield growth gains of 12.6pp from advisory access (SE = 6.8, p = 0.063). In contrast, farmers with below-median land size (n = 1,056) experience statistically insignificant yield growth gains of 2.5pp(SE = 9.4, p = 0.787). While the point estimates suggest large landholders benefit approximately 5 times more than small landholders, this contrast is imprecisely estimated at the conventional levels due to high variance in agricultural outcomes. Nevertheless, this motivates an important question: can our optimization framework be adapted to prioritize reaching these high-benefit farmers?

As an illustrative exercise, we test this by implementing a weighted two-step optimization for targeted policy design in the dataset used in the project of this paper. Formally, we solve the constrained optimization problem using modified objective weights, where the weights are determined by land size.[22] Table A12 reports the gains from targeted policy under the two scenarios using contemporaneous off-policy evaluation. Under equal weighting, large landholders have a baseline pickup rate of 29.0%, approximately 3 percentage points lower than the 32.2% rate for small landholders. When we weight large landholders 20-fold more heavily in the optimization objective, their pickup rate increases to 30.2% (a 1.2 percentage-point gain). However, this comes at a cost: the pickup rate for the small landholder farmers declines by 0.6 percentage points to 31.6%. This translates into reaching an additional 1474 large landowners at the cost of not reaching 2607 small landowners.[23]

To assess whether this tradeoff is welfare-improving, we calculate the expected aggregate yield impact by combining pickup changes with heterogeneous treatment effects. For large landholders, the 1.2pp pickup gain combined with their 12.6pp yield impact contributes +0.032pp to aggregate yields. For small landholders, the 0.6pp pickup loss combined with their 2.5pp yield impact reduces aggregate yields by -0.011pp. The net expected aggregate yield gain from prioritization is +0.021pp. This is a modest improvement driven by the 5-fold difference in yield impacts between farmer groups.

These results highlight a fundamental tension between maximizing downstream agricultural impacts and ensuring service coverage across all farmer groups. Even when prioritizing the farmers most likely to generate substantial yield gains, the optimization increases their engagement by only 1.2 percentage points, a modest gain that yields only a 0.021pp improvement in aggregate yields. This small improvement likely arises from two factors. First, large landholders exhibit systematically lower baseline pickup rates, likely reflecting higher opportunity costs of time that make them harder to reach through call timing optimization alone. Second, call timing represents only one dimension of targeting; alternative strategies, such as personalized message content, may be needed to effectively reach high-benefit farmers who are less responsive to timing adjustments. Organizations implementing such services must therefore explicitly weigh the tradeoff between concentrating resources on high-impact farmers versus ensuring broad reach across their beneficiary population.

7 Discussion

This project demonstrates that advanced targeted policies that leverage machine learning methods with RCTs can meaningfully improve engagement in resource-constrained development settings using technology as simple as automated voice calls. Our iterative experimental design, comprising

[22] We define large landholders as the top quartile in our setting by land size because this threshold provided sufficient variation in pickup rates to detect differences between equal-weight and prioritized targeting in our setting.

[23] We note here that this is just an illustrative exercise since the land size distribution can vary for the farmers between the two states, as well as their heterogeneity in treatment effects of access to the agricultural advisory service.

six field experiments with over 1 million farmers, reveals a critical challenge that static one-period or two-period designs would miss: temporal instability in treatment effects degrades policy performance. We develop a unified framework to address this instability through transfer learning adapted for policy learning (providing data-driven auxiliary week selection), sieve estimators that exploit temporal smoothness in treatment effects, and risk-aware optimization that incorporates prediction uncertainty. These methods recover and exceed the initial off-policy gains when evaluated using K-step ahead off-policy evaluation. Beyond efficiency, our distribution-aware targeting framework provides transparent decision tools for organizations facing priorities for certain groups, quantifying how capacity constraints and subgroup weighting affect coverage allocation across farmer populations.

Our findings provide actionable guidance for organizations deploying digital marketing interventions. Off-policy evaluation offers flexibility for exploring distributional preferences, whereas on-policy evaluation reveals temporal drift that degrades performance. We demonstrate that temporal drift is solvable: transfer learning adapted to targeted policy design provides robust policies when agricultural seasonality creates instability. We optimize short-term engagement (pickup rates) rather than long-term outcomes (yields, profits) because downstream agricultural outcomes require months to observe and are confounded by weather, pests, and markets. Our approach is supported by rigorous impact evaluations of this service, which demonstrate it increases yields by 2-7% and reduces crop loss by 10-21%. Critically, even disadvantaged farmers benefit substantially (5% harvest gains, 14% crop loss reduction), validating our focus on optimizing the critical first stage of the marketing funnel as pickup is necessary for all downstream impacts.

While our study demonstrates the feasibility of sophisticated targeting at scale in development contexts, several limitations suggest important directions for future work. First, our experimental design randomized treatments independently each week, precluding identification of dynamic treatment sequence effects. Future research can randomize farmers to predetermined multi-week policy sequences to understand how repeated exposure patterns affect long-term engagement and learning. Second, we optimize for call pickup as a surrogate outcome; future work can directly evaluate how different targeting strategies affect downstream agricultural outcomes (yields, profits, practice adoption) and examine whether policies optimized for engagement also maximize long-term welfare gains. Third, while our distribution-aware framework quantifies coverage trade-offs across subgroups, we cannot assess the long-term impacts of such prioritization on sustained engagement across farmer communities. Fourth, our outcome model is estimated using LASSO rather than deep learning architectures. While this choice prioritizes interpretability and computational efficiency in resource-constrained settings, future work could explore whether neural network models capture additional complex interaction patterns that improve the performance of a targeted policy over a uniform policy. Finally, future research could explore how targeted follow-up timing, combined with message content customization, further enhances the effectiveness of digital advisory services in resource-constrained settings.

References

Agrawal, K., Carleton Athey, S., Kanodia, A., Nath, S., and Palikot, E. (2025). The economics of algorithmic personalization: Evidence from an educational technology platform. *Available at SSRN 5996014*.

Araya, R., Menezes, P. R., Claro, H. G., Brandt, L. R., Daley, K. L., Quayle, J., Diez-Canseco, F., Peters, T. J., Cruz, D. V., Toyama, M., et al. (2021). Effect of a digital intervention on depressive symptoms in patients with comorbid hypertension or diabetes in Brazil and Peru: Two randomized clinical trials. *JAMA*, 325(18):1852–1862.

Ascarza, E. (2018). Retention futility: Targeting high-risk customers might be ineffective. *Journal of Marketing Research*, 55(1):80–98.

Ascarza, E. and Israeli, A. (2022). Eliminating unintended bias in personalized policies using bias-eliminating adapted trees (beat). *Proceedings of the National Academy of Sciences*, 119(11):e2115293119.

Athey, S. (2025). Presidential address: The economist as designer in the innovation process for socially impactful digital products. *American Economic Review*, 115(4):1059–1099.

Athey, S., Byambadalai, U., Cersosimo, M., Koutout, K., and Nath, S. (2024). The heterogeneous impact of changes in default gift amounts on fundraising. *Available at SSRN 4785704*.

Athey, S., Chetty, R., Imbens, G. W., and Kang, H. (2025a). The surrogate index: Combining short-term proxies to estimate long-term treatment effects more rapidly and precisely. *Review of Economic Studies*, page rdaf087.

Athey, S., Keleher, N., and Spiess, J. (2025b). Machine learning who to nudge: causal vs predictive targeting in a field experiment on student financial aid renewal. *Journal of Econometrics*, 249:105945.

Athey, S. and Wager, S. (2021). Policy learning with observational data. *Econometrica*, 89(1):133–161.

Bastani, H. (2021). Predicting with proxies: Transfer learning in high dimension. *Management Science*, 67(5):2964–2984.

Cole, S. and Fernando, A. N. (2021). 'Mobile'izing agricultural advice: Technology adoption, diffusion, and sustainability. *Economic Journal*, 131(633):192–219.

Cole, S., Goldberg, J., Harigaya, T., and Zhu, J. (2025). The impact of digital agricultural extension service: Experimental evidence from rice farmers in india. Working paper.

Dammert, A. C., Galdo, J., and Galdo, V. (2015). Integrating mobile phone technologies into labor-market intermediation: A multi-treatment experimental design. *IZA Journal of Labor & Development*, 4(1):1–27.

Dubé, J.-P. and Misra, S. (2023). Personalized pricing and consumer welfare. *Journal of Political Economy*, 131(1):131–189.

Ellickson, P. B., Kar, W., and Reeder III, J. C. (2023). Estimating marketing component effects: Double machine learning from targeted digital promotions. *Marketing Science*, 42(4):704–728.

Fabregas, R., Kremer, M., and Schilbach, F. (2019). Realizing the potential of digital development: The case of agricultural advice. *Science*, 366(6471):eaay3038.

Harigaya, T. and Zhu, J. (2025). Customized agricultural advisory services for smallholder farmers in india. Working paper.

Hu, Y., Du, R. Y., and Damangir, S. (2014). Decomposing the impact of advertising: Augmenting sales with online search data. *Journal of marketing research*, 51(3):300–319.

Kitagawa, T. and Tetenov, A. (2018). Who should be treated? empirical welfare maximization methods for treatment choice. *Econometrica*, 86(2):591–616.

Lee, J., Foschini, L., Kumar, S., Juusola, J., Liska, J., Mercer, M., Tai, C., Buzzetti, R., Clement, M., Cos, X., et al. (2021). Digital intervention increases influenza vaccination rates for people with diabetes in a decentralized randomized trial. *NPJ digital medicine*, 4(1):1–8.

Li, H. and Kannan, P. (2014). Attributing conversions in a multichannel online marketing environment: An empirical model and a field experiment. *Journal of marketing research*, 51(1):40–56.

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

Lu, H., Simester, D., and Zhu, Y. (2025). Optimizing scalable targeted marketing policies with constraints. *Marketing Science*.

Mazumder, R., Hastie, T., and Tibshirani, R. (2010). Spectral regularization algorithms for learning large incomplete matrices. *The Journal of Machine Learning Research*, 11:2287–2322.

Rafieian, O. and Yoganarasimhan, H. (2021). Targeting and privacy in mobile advertising. *Marketing Science*, 40(2):193–218.

Rodriguez-Segura, D. (2022). Edtech in developing countries: A review of the evidence. *The World Bank Research Observer*, 37(2):171–203.

Sahni, N. S., Wheeler, S. C., and Chintagunta, P. (2018). Personalization in email marketing: The role of noninformative advertising content. *Marketing Science*, 37(2):236–258.

Simester, D., Timoshenko, A., and Zoumpoulis, S. I. (2020a). Efficiently evaluating targeting policies: Improving on champion vs. challenger experiments. *Management Science*, 66(8):3412–3424.

Simester, D., Timoshenko, A., and Zoumpoulis, S. I. (2020b). Targeting prospective customers: Robustness of machine-learning methods to typical data challenges. *Management Science*, 66(6):2495–2522.

Wager, S. and Athey, S. (2018). Estimation and inference of heterogeneous treatment effects using random forests. *Journal of the American Statistical Association*, 113(523):1228–1242.

Wei, Q., Mu, Y., Guo, X., Jiang, W., and Chen, G. (2024). Dynamic bayesian network–based product recommendation considering consumers' multistage shopping journeys: a marketing funnel perspective. *Information Systems Research*, 35(3):1382–1402.

Yang, J., Eckles, D., Dhillon, P., and Aral, S. (2024). Targeting for long-term outcomes. *Management Science*, 70(6):3841–3855.

Yoganarasimhan, H., Barzegary, E., and Pani, A. (2023). Design and evaluation of optimal free trials. *Management Science*, 69(6):3220–3240.

Zhou, B. and Zou, T. (2023). Competing for recommendations: The strategic impact of personalized product recommendations in online marketplaces. *Marketing Science*, 42(2):360–376.

Zhou, Z., Athey, S., and Wager, S. (2022). Offline multi-action policy learning: Generalization and optimization. *Operations Research*.

Appendix

A1 Details on the Agricultural Advisory Service

The agricultural extension service utilized in this study is a digital advisory service developed and implemented jointly by a non-governmental organization (Precision Development) and a state government in India. The service provides agricultural advice covering 21 crops, as well as livestock and fisheries, and serves millions of farmers across the state.

During the agricultural season, advisory messages are delivered to farmers' mobile phones through interactive voice response (IVR) calls. Messages are customized based on farmers' preferred language, crop profile, geographic location, land characteristics, sowing method, and irrigation type, and are scheduled according to the crop calendar on a weekly basis. The messages cover major agricultural decisions throughout the production cycle, including land preparation, seed varietal selection, sowing practices, nutrient management, pest and disease management, harvesting and post-harvest practices, and market price information. The service also provides preventive and mitigation guidance in response to weather events and pest outbreaks. Below, we provide two examples of advisory messages delivered during the experimental period.

If an IVR call is not answered, the service system automatically reschedules the call up to twice, with a 24-hour interval between attempts. Precision Development records all user interactions with the service. Engagement with the service, defined as picking up the advisory calls, is considered the first step in the program's theory of change and is hypothesized to translate into downstream benefits for farmers (see Figure 1).

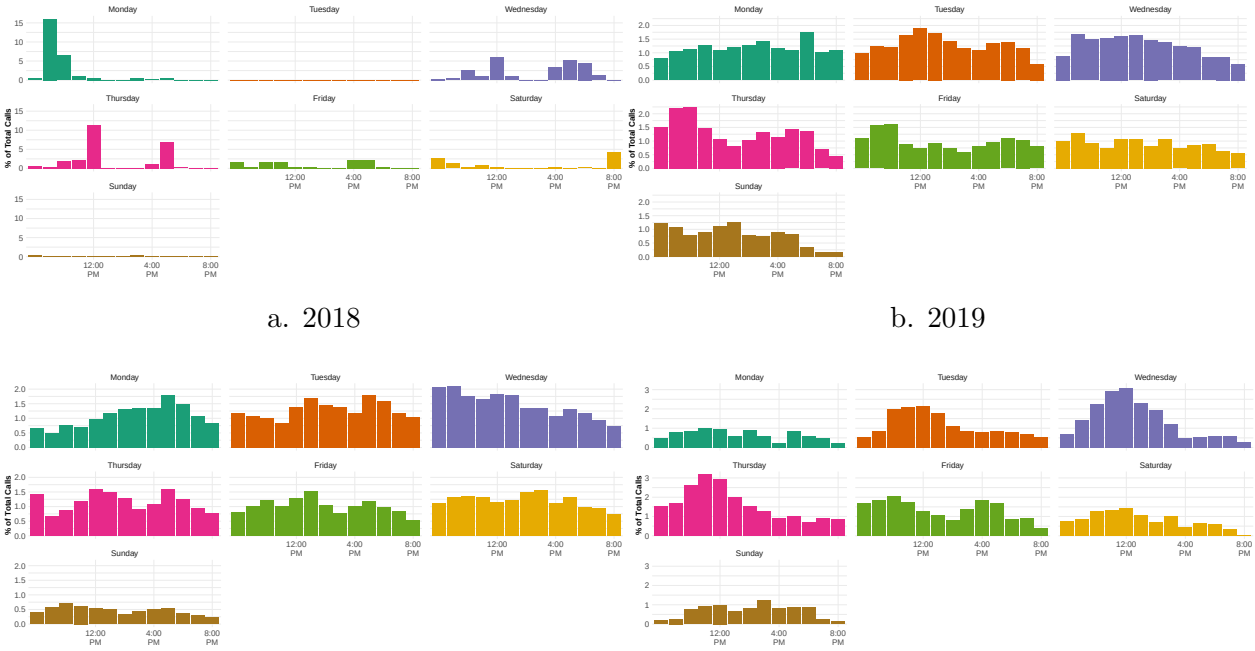
Message example 1. Advisory on pest management: *Namaskar. Today we will discuss about neck blast disease and its management in paddy crop. Due to high relative humidity and differential day and night temperature Neck Blast disease incidence can be seen in paddy crops. To manage these diseases, first drain out excess water from the paddy field. Spray Hexaconazole (Contaf Plus/Hexadhan Plus/Trigger Pro) @ 400-ml/acre or Azoxystrobi+ Difenoconazole (Amistar Top/ Chemistar /Karishma) @ 200-ml/acre or Tebuconazole + Trifloxystrobin (Nativo) @ 80-gram/acre. Thank you and remember that if you have questions about this advisory or need more information, you can call the hotline number on 155333.*

Message example 2. Advisory on basal fertilizer application for transplanting: *Namaskar. Today we will share a few tips for applying basal fertilizers correctly for farmers who are growing HYV paddy. If you have not yet done so, we advise you to complete your transplanting by August 15th. Before applying fertilizers at the time of sowing, you should determine what kind of soil you have. This is because the fertilizers recommended are different for different soil. You should apply 35 kg DAP, 30 kg Potash, and 8 kg Urea per acre during the last puddling. Remember again that you should apply 35 kg DAP, 30 kg Potash, and 8 kg Urea per acre at the time of last puddling. Please*

remember 1acre=25 guntha. However, you should apply Potash in two equal splits at the basal and PI stages if you have sandy soil. Also, do not forget that zinc deficiency is the most widespread micronutrient disorder in paddy, affecting plant growth. For this reason, we suggest that you can apply 10 kg of Zinc Sulphate micronutrient per acre based on a soil test report or if your soil is deficient in zinc. Thank you and remember that if you have questions about this advisory or need more information, you can call the hotline number on 155333, and we will provide this message to your mobile via SMS.

A2 Additional Tables and Figures

Figure A1: Number of push calls sent during each call hour in the past



a. 2018

b. 2019

c. 2020

d. 2021

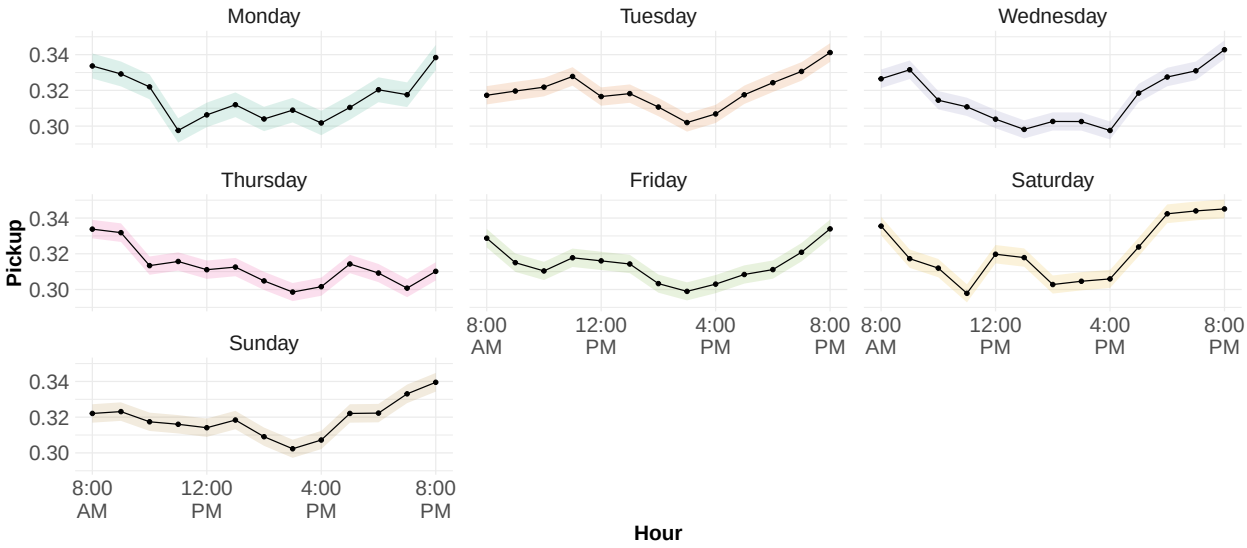
Notes: These figures show the number of push calls sent between July 2018 and September 30, 2021. It illustrates that push calls used to be sent in an ad hoc manner in the past.

Table A1: Variation in Engagement by Call Times

	Morning	Afternoon	Difference
Mean	0.320	0.307	0.013
Std. Error	[0.00051]	[0.00046]	[0.00069]
	Evening	Afternoon	Difference
Mean	0.325	0.307	0.018
Std. Error	[0.00052]	[0.00046]	[0.00069]

Notes: We pool the data collected using uniform randomization in weeks 1, 2, and 3 of the experiment (see Table 1 for sample size). Morning hours are between 8 AM to 12 PM, afternoon hours are between 12 PM to 5 PM, and evening hours are between 5 PM to 9 PM.

Figure A2: Pickup by 91 Treatment Arms



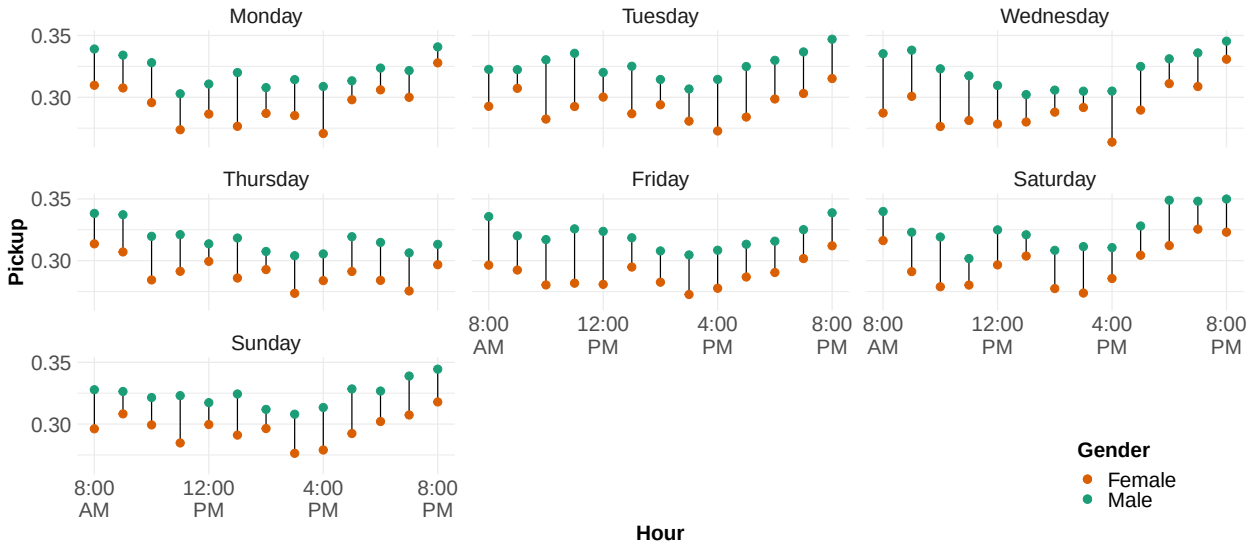
Notes: We pool the data collected using uniform randomization in weeks 1, 2, and 3 of the experiment (see Table 1 for sample size). It illustrates the variation in engagement between the 91 call times, which is a combination of the day of the week and the hour of the day (Mean +/- 1.96SE).

Table A2: Summary Statistics

Variable	Mean	Std.	Median
	A. Outcome Variable		
Pick-up	0.318	0.283	0.250
	B. Covariates		
Female	0.181	0.385	0.000
Smartphone	0.374	0.484	0.000
Irrigation	0.444	0.497	0.000
Landsize	1.286	1.881	0.809
N	884,194		

Notes: We also include the district of residence and historical engagement as covariates for estimating $\hat{\pi}(x_i, \cdot)$.

Figure A3: (a) Pickup by time and gender, (b) Highest Pickup arms for female farmers and (c) male farmers in the dataset.



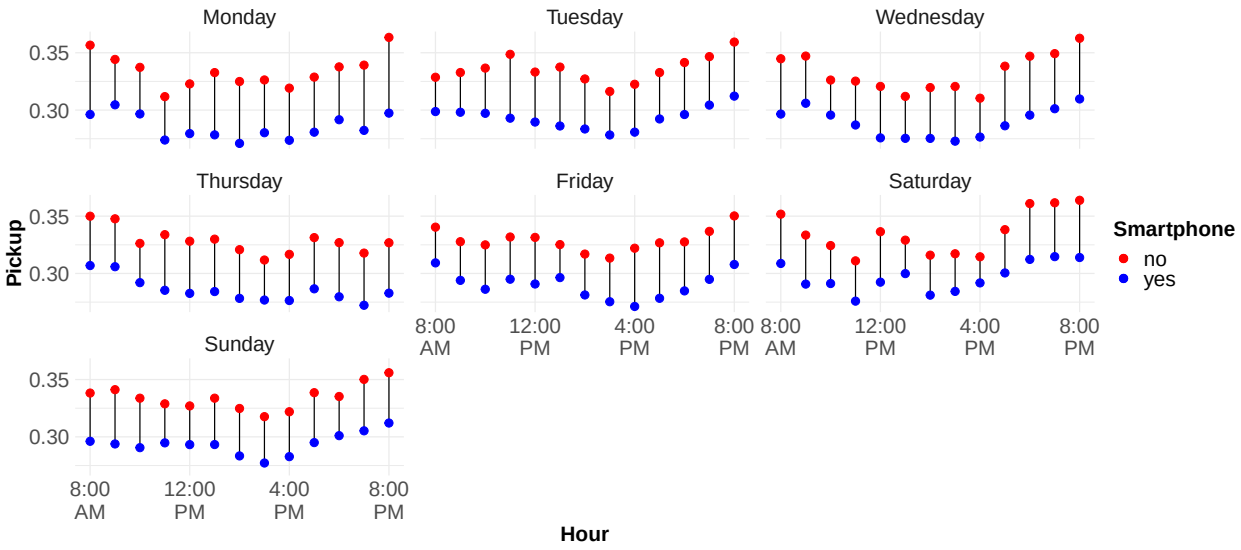
a. Variation in Pickup by Gender for 91 Treatment Arms

b. Top 10 Buckets for Females

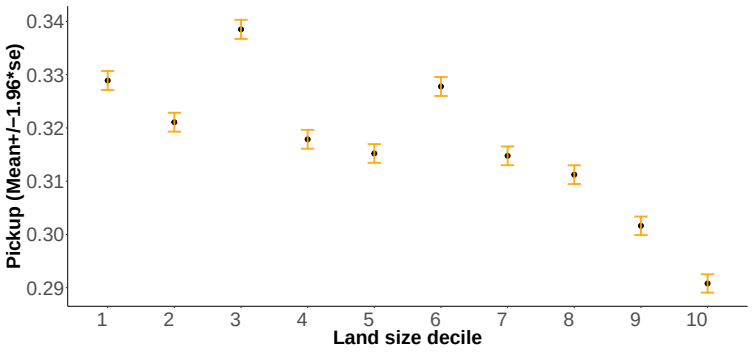
c. Top 10 Buckets for Males

Notes: We pool the data collected using $\mathcal{U}$ in weeks 1, 2, and 3 of the experiment. Panel (a) illustrates the gender gap in engagement over the 91 call times. Panel (b) shows the most popular call times for female farmers arranged in increasing order of popularity, with the pickup for male farmers for comparison. Panel (c) shows the most popular call times for male farmers arranged in increasing order of popularity with the pickup for female farmers.

Figure A4: (a) Pickup by time and smartphone, (b) Pickup by the distribution of land size.



a. Variation in Pickup by Smartphone Dummy for 91 Treatment Arms



b. Pickup by Land size

Notes: For this figure, we pool the data collected using uniform randomization in weeks 1, 2, and 3 of the experiment. Panel (a) illustrates the smartphone users' and nonusers' gap in engagement with the service over the 91 call times. Panel (b) shows the variation in pickup by land size (Mean +/- 1.96se). We divide the sample of farmers into deciles based on their land size. For each decile, we compute the mean and standard error for pickup.

Table A3: Summary Statistics: Additional Covariates

Variable	Mean	Std.	Median
Recency, Frequency and Weather Variables			
Total Calls	77.270	49.168	64.000
Total Pickups	23.835	18.691	20.000
Recency Last Call	3.610	4.074	2.000
Overall Pickup Rate	0.327	0.174	0.328
Max Pickup Streak	4.524	2.817	4.000
Weekday Pickup Rate	0.325	0.178	0.324
Evening Pickup Rate	0.334	0.219	0.333
Pickup Trend	-0.041	0.204	-0.037
Recent Pickup Rate	0.292	0.221	0.280
Rainfall (mm)	2.196	6.826	0.000
N	884,194		

Notes: Recency, Frequency variables are computed from historical engagement data prior to the experiment. Rainfall is daily rainfall in mm for 2021.

Algorithm 1 Algorithm for Estimating $\hat{\pi}$

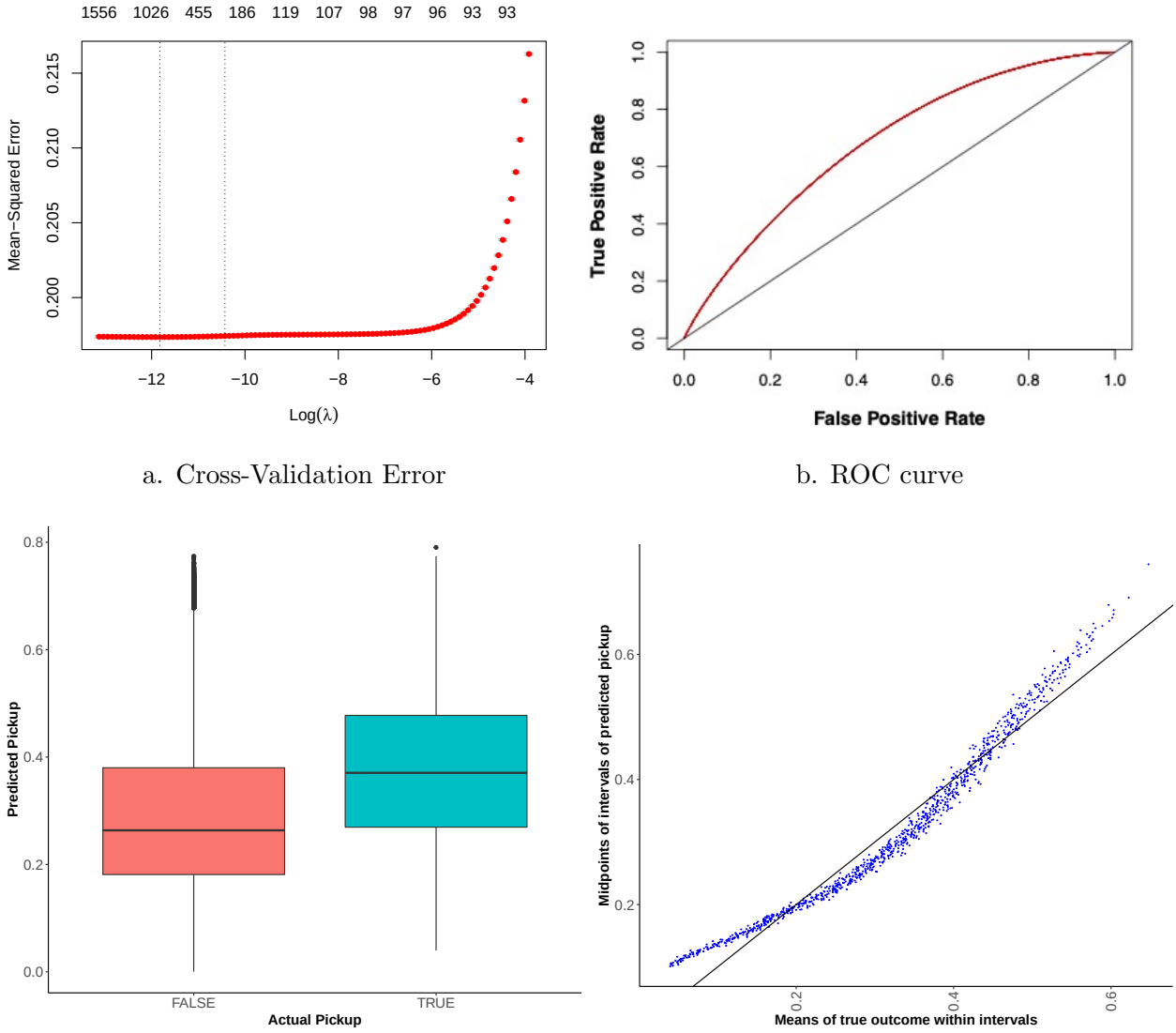
Input: Capacity limits b; Training data $\mathcal{S}$
Result: Model parameters $(\hat{\delta}, \hat{\beta}, \hat{\gamma})$, Allocation z

1: Partition data into K mutually exclusive folds
2: **for** $k = 1$ to K **do**
3: Estimate LASSO using all folds except k (Equation 3): Obtain $\hat{\delta}, \hat{\beta}, \hat{\gamma}$
4: Predict for each i in fold k: $\hat{\mu}_{-k}(x_i, J_i, \cdot)$
5: Append $\hat{\mu}_{(-k)}$ $\forall$ $k \in K$
6: Constrained Optimization: $\max_z \sum_{i=1}^{n} \sum_{j=1}^{J} z_{ij} \hat{\mu}_{-k(i)}(x_i, j, \cdot)$ s.t. $\sum_j z_{ij} = 1 \forall i$ and $\sum_i z_{ij} \leq b_j \forall j$ (Equation 4)
7: Use $\hat{\mu}$, capacity limit b and solve for z
8: **return** $[\hat{\delta}, \hat{\beta}, \hat{\gamma}, z]$

Table A4: Targeted Policy Design with Random Forest for Outcome Model Estimation

	Targeted Policy	Uniform Policy	Difference
Mean	0.3367	0.3168	0.0200
	[0.0028]	[0.0003]	[0.0028]

Figure A5: Out-of-Sample Predictive Performance



a. Cross-Validation Error

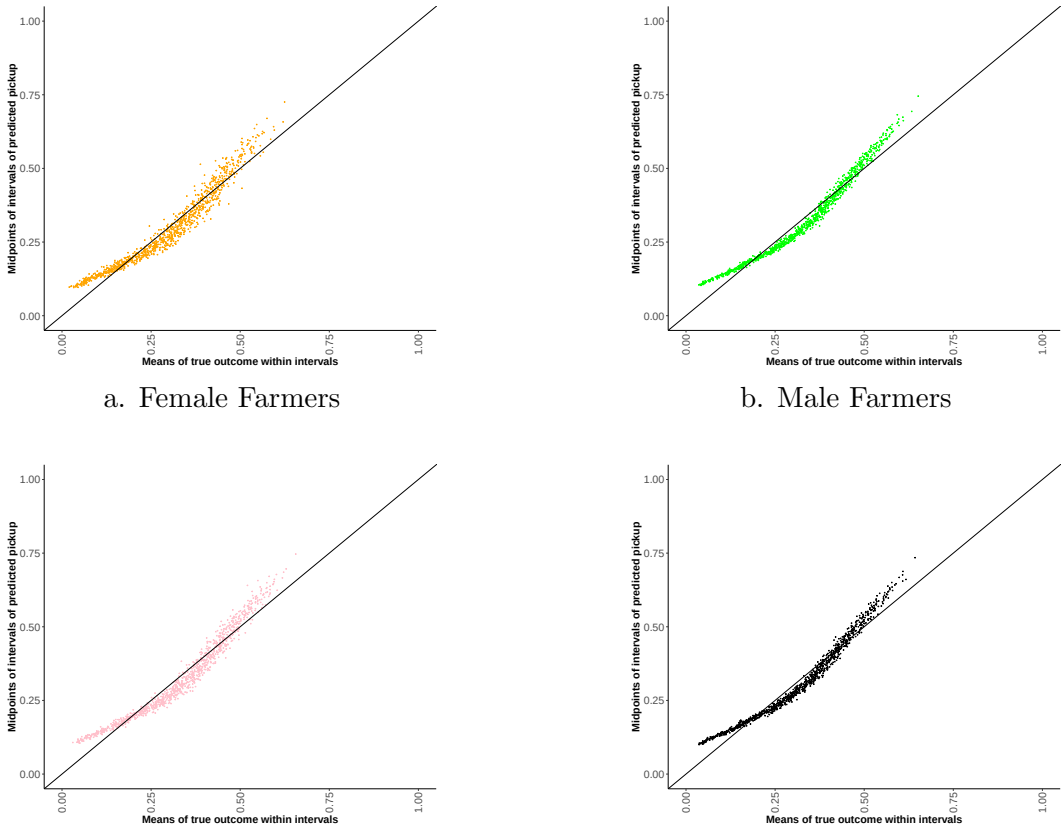
b. ROC curve

c. Predicted Response

d. Calibration Plot

Notes: (a) shows the MSE corresponding to the different regularization parameters (λ) used in our analysis. (b) shows the ROC curve for the out-of-sample prediction using cross-fitted data. The AUC is 0.683. (c) shows the distribution of predicted response by the binary true pickup in the data. (d) shows the calibration plot for true and predicted pickup using cross-fitted data.

Figure A6: Calibration Plots for LASSO by Farmer Covariates



a. Female Farmers

b. Male Farmers

c. Small Landowners

d. Large Landowners

Notes: Panel (a) shows the calibration plots between the true and predicted outcome for the female farmers. Panel (b) shows the calibration plots for the male farmers. Panel (c) shows the calibration plot for the small landowners whose land size is below or the same as the 25^{th} percentile of the land-size distribution. Panel (d) shows the calibration plot for large landowners. Large landowners are defined as those farmers whose land holding is above 25^{th} percentile of the land-size distribution.

Table A5: Robustness to Cross-Fitting Specification: Policy Value Estimates

Cross-Fitting Folds	Uniform Policy	Targeted Policy	Improvement
2-Fold	0.3167	0.3418	0.0251
	[0.0003]	[0.0028]	[0.0028]
3-Fold	0.3167	0.3427	0.0260
	[0.0003]	[0.0028]	[0.0028]
4-Fold	0.3167	0.3431	0.0264
	[0.0003]	[0.0028]	[0.0028]

Notes: This table reports off-policy evaluation results for the targeted calling policy under different cross-fitting specifications. Improvement is the difference between targeted and uniform policies. Standard errors are shown in square brackets.

Table A6: K-step ahead Off-Policy Evaluation Using Week 6 Uniform Randomization Data

	Targeted Policy	Uniform Policy	Difference
mean	0.3122	0.3093	0.0028
se	[0.0059]	[0.0006]	[0.006]

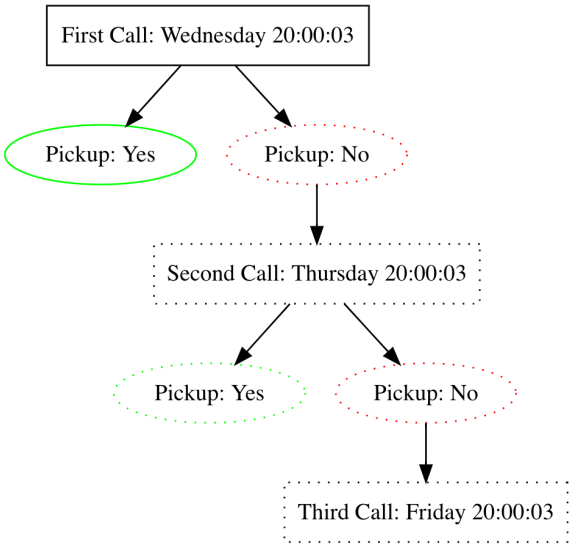
Table A7: Treatment Assignment Stability based on Best Predicted Call time

Subgroup	Train 1,2 vs 1,2,3 (Eval: 3 vs 5)	Train 3,4,5 vs 5 (Eval: 6)
Overall	51.1	20.0
Female	50.2	21.2
Male	51.3	19.7
Small Land	47.2	23.0
Large Land	55.6	16.5
Smartphone	54.3	10.3
Non-Smartphone	49.2	26.1

Table A8: Top-K Treatment Overlap: Percentage of top-K treatments that remain in top-K under different training scenarios

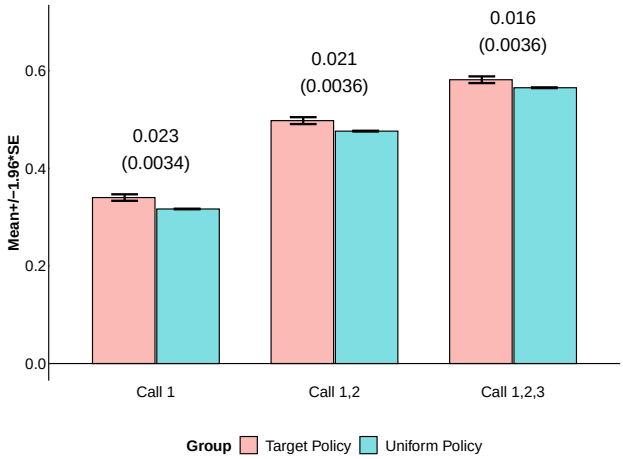
K	Subgroup	Train 1,2 vs 1,2,3 (Eval: 3 vs 5)	Train 3,4,5 vs 5 (Eval: 6)
5	Overall	63.1	47.6
	Female	59.2	47.7
	Male	63.9	47.6
	Small Land	62.5	48.3
	Large Land	63.7	46.9
	Smartphone	64.6	47.3
	Non-Smartphone	62.1	47.8
10	Overall	72.0	47.6
	Female	71.5	47.6
	Male	72.1	47.5
	Small Land	71.4	47.7
	Large Land	72.7	47.4
	Smartphone	72.3	47.4
	Non-Smartphone	71.8	47.7
20	Overall	79.2	65.1
	Female	80.0	66.3
	Male	79.0	64.9
	Small Land	78.8	65.1
	Large Land	79.5	65.2
	Smartphone	80.2	65.2
	Non-Smartphone	78.5	65.1

Figure A7: Repeat Calls from the Call Center



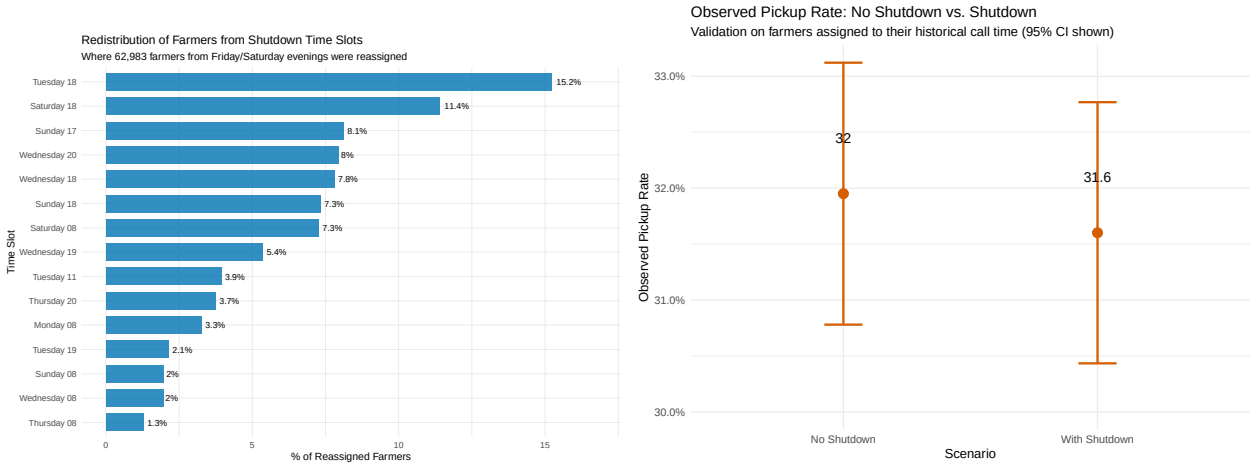
Notes: This figure illustrates the process of follow-up calls from the call center. If the farmer does not pick up the call on the first attempt, the call center makes a call exactly 24 hours after the first call. Moreover, if the farmer does not pick up the call on the second attempt, the third call is made 24 hours after the second call.

Figure A8: Impact of Targeting First Call on Overall Farmer Engagement: Off-Policy Evaluation



Notes: The training and evaluation data for estimating the value of $\hat{\pi}$ counterfactually and the uniform policy uses the uniform randomization data for week 1 and week 2. We estimate $\hat{\pi}$ on the first calls for the two weeks of data using cross-fitting (following Algorithm 1) and counterfactually evaluate $\hat{\pi}$. We estimate $\hat{V}^{\text{Off}}(\hat{\pi}, \mathcal{S}_{\mathcal{U},1,2})$, where the outcome changes to the total pickup for call 1, call 2 for second pair of bars. The evaluation dataset for the $\hat{\pi}$ group corresponds to the subset of people whose actual assignment matched the first call $\hat{\pi}$. We estimate $\hat{V}^{\text{Off}}(\hat{\pi}, \mathcal{S}_{\mathcal{U},1,2,3})$, where the outcome changes to the overall pickup for call 1, 2, 3 for right most pair of bars.

Figure A9: Bandwidth Shocks



a. b.

Notes: Panel (a) shows the redistribution of calls across time slots when specific slots are unavailable due to bandwidth shocks (e.g., server maintenance, festivals, or holidays). Panel (b) compares observed pickup rates between the baseline allocation and the shock-adjusted allocation using contemporaneous off-policy evaluation.

Table A9: Incorporating Follow-Up Calls to Mitigate Shocks: On-Policy Evaluation

Data Collection-policy	$\hat{\pi}_{\text{D}}$	Uniform	Difference
Outcome Variable			
	All Farmers		
Call 1,2,3	0.552	0.544	0.008
	[0.001]	[0.001]	[0.001]
	Female Farmers		
Call 1,2,3	0.5249	0.5133	0.0116
	[0.0025]	[0.0016]	[0.0029]
	Male Farmers		
Call 1,2,3	0.5577	0.5500	0.0077
	[0.0011]	[0.0007]	[0.0014]
N	227,386	587,608	

Notes: This table shows the on-policy evaluation results for targeted policy $\hat{\pi}_D$ in week 6. The farmers are randomized between two groups. Group A gets called according to $\hat{\pi}_D$, and Group B gets called according to uniform randomization. Sample means are used to estimate the value of $V(\hat{\pi}_D, \mathcal{S}^{\text{eval}})$ and $V(\bar{\mathcal{U}})$.

Table A10: Summary Statistics: Treatment, Outcomes, and Covariates from Cole and Fernando (2021)

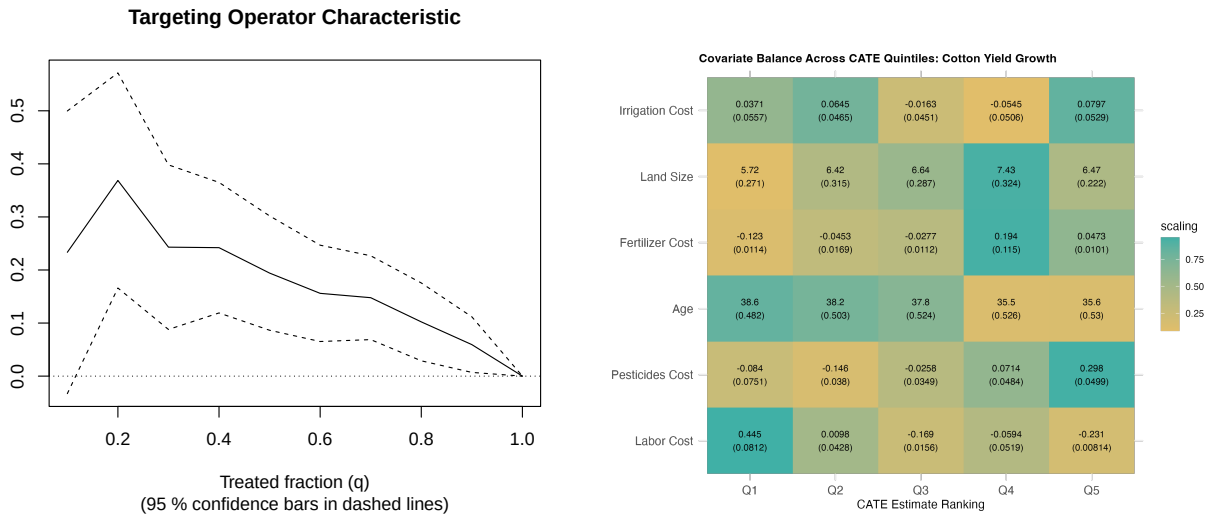
Variable	Mean	SD	Median	N
Treatment				
Ever received advisory services	0.666	0.472	1.000	2069
Outcomes				
Cotton yield (% growth rate)	0.046	1.313	-0.229	2069
Covariates				
Land size (acres)	6.536	5.838	4.800	2069
Age (years)	37.120	10.518	35.000	2069
Baseline fertilizer cost (logs)	0.009	1.073	-0.089	2069
Baseline irrigation cost (logs)	0.022	1.024	-0.506	2069
Baseline labor cost (logs)	-0.001	1.001	-0.275	2069
Baseline pesticide cost (logs)	0.023	1.054	-0.213	2069

Notes: Summary statistics for the clean sample used in the downstream outcome analysis.

Table A11: Heterogeneous Treatment Effects: ATE Differences Relative to Quintile 1 (Cotton Yield Growth)

comparison	estimate	std.err	CI$_{l}ower$	CI$_{u}pper$	t.value	p.value	stars
Q 2 - Q1	0.2508	0.1938	-0.1291	0.6308	1.29	0.1958	
Q 3 - Q1	0.3038	0.2270	-0.1411	0.7486	1.34	0.1809	
Q 4 - Q1	0.2818	0.2053	-0.1206	0.6842	1.37	0.1700	
Q 5 - Q1	0.4557	0.1963	0.0709	0.8405	2.32	0.0204	**

Figure A10: Downstream Outcome: Growth in Cotton Yield



a. TOC

b. Covariate Ranking by CATE

Table A12: Land Size Prioritization: Pickup Rates by Weighting Scenario

Land Size Group	**Equal Weights** (1x:1x)	**20x Priority** (20x:1x)
Small Landholders	0.3219	0.3164
	(0.0038)	(0.0037)
Large Landholders	0.2901	0.3018
	(0.0070)	(0.0070)

Notes: Pickup rates under two optimization scenarios. Equal weights (1x:1x) assigns all farmers identical priority. The 20x priority scenario (20x:1x) assigns large farmers 20 times the weight of small landholders. Standard errors in parentheses. Bandwidth = 10,000 calls per time slot.

A3 Additional Historical Variables

In this section, the additional historical variables constructed using the data from October 2018 to September 30th, 2021, are described. This historical engagement data is matched at the farmer level with the farmer IDs in the experiment sample. The goal of this exercise is to construct variables from historical observational data that capture farmers' engagement patterns with the service before the start of the experiment. We construct the following variables.

Call Volume Variables

1. Total Calls: Total sum of the calls made to the farmer since they registered with the advisory service.

2. Total Pickups: Sum of the total calls that each farmer picked up successfully.

Recency Variable

1. Recency Last Call: This variable calculates the total number of days elapsed between the last call attempt and the reference date of September 30, 2021. Note that the experiment started in October, hence the historical data is considered till the end of September 2021.

2. Recent Pickup Rate: In the 90 days prior to September 30, 2021, this variable computes the recent pickup rate as the ratio of recent pickups divided by the number of recent calls.

Engagement Variables

1. Overall Pickup Rate: Ratio of total Successful Pickups to the total call attempts. It captures the engagement of the farmer over their entire tenure with the service.

2. Max Pickup Streak: This variable calculates the longest consecutive sequence of calls for which the farmer did pickup up the call.

3. Weekday Pickup Rate: Calculates the pickup rate for the weekday calls.

4. Evening Pickup Rate: Calculates the pickup rate for the evening calls.

5. Pickup Trend: This captures the variation in engagement between the first and the latter half of their tenure with the advisory service.

A4 Simplified Setting

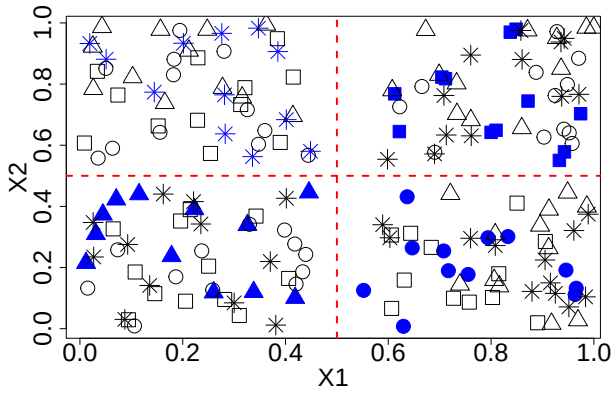
A4.1 Off-Policy Evaluation

The concept of off-policy evaluation can be explained using a simplified setting with four treatment arms and two covariates. Assume two uniformly distributed covariates $X = [X_1, X_2]$ and the outcome variable is Y. Let there be only four treatments (W_1, W_2, W_3, W_4). The two policies $\mathcal{U}$ and $\hat{\pi}$ are defined below.

$$\mathcal{U} = \begin{cases} W_1, & p = 0.25 \\ W_2, & p = 0.25 \\ W_3, & p = 0.25 \\ W_4, & p = 0.25 \end{cases} \qquad \hat{\pi} = \begin{cases} W_1, & x_1 > 0.5, x_2 > 0.5 \\ W_2, & x_1 > 0.5, x_2 < 0.5 \\ W_3, & x_1 < 0.5, x_2 < 0.5 \\ W_4, & \text{otherwise} \end{cases}$$

Under the uniform randomization, every individual could be allocated to any of those four treatments with a 0.25 probability. Under the targeted policy, individuals in the top left quadrant, top right quadrant, bottom left quadrant, and bottom right quadrant would be assigned to treatment W_4, W_1, W_3, and W_2, respectively (Figure A11). Because of uniform randomness, $\frac{N}{\text{No. of arms}}$ individuals under the uniform randomization are actually assigned to the treatment arm they would have received if the targeted policy $\hat{\pi}$ was deployed. Those individuals are highlighted using blue points in the figure, and they would be the treated individuals in the off-policy evaluation.

Figure A11: Off-Policy Evaluation: Simplified Setting



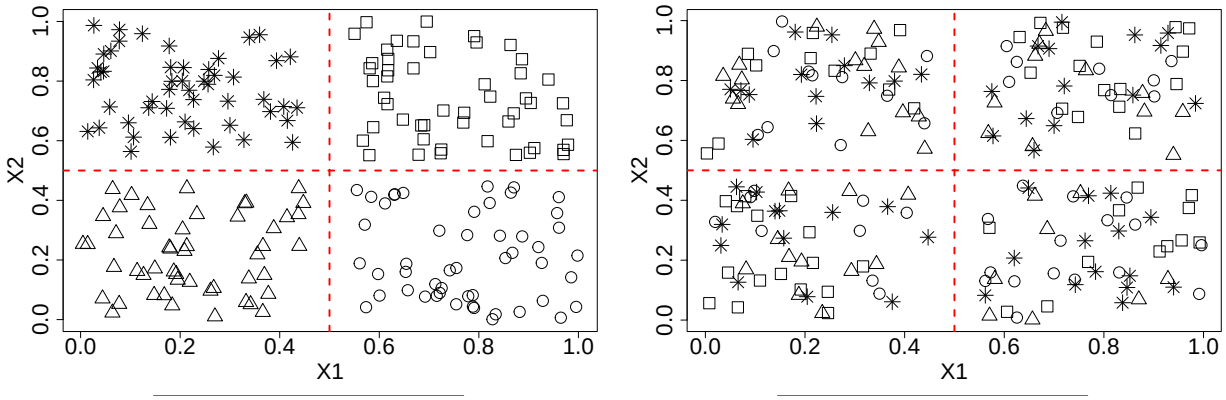
Notes: Data is collected using the uniform randomization, where probability of allocation to each arm is 0.25. 25% of farmers get the same assignment under $\mathcal{U}$ as they would have received under $\hat{\pi}$ (blue points).

A4.2 On-Policy Evaluation

Using the same simplified setting with two covariates, we illustrate graphically the concepts of on-policy evaluation used in this study. Figure A12 illustrates the policy for the above setting. The difference in expectation for the subset of individuals who should get W_1 based on $\hat{\pi}$ is obtained

by comparing the outcomes for the upper right quadrant in the uniform policy group and the targeted policy group. In other words, the covariate space is kept the same and only the policy is varied between the two groups. Similarly, the differences can be computed for the other arms. Additionally, the overall benefit of deploying $\hat{\pi}$ over uniform policy can be computed by comparing outcomes in Panel (a) with the outcomes in Panel (b).

Figure A12: On-Policy Evaluation: Simplified Setting



a. Deploy $\hat{\pi}$ b. Deploy $\mathcal{U}$

Notes: Here, individuals are randomly allocated between the group that receives the treatment according to $\hat{\pi}$ (left) and uniform randomization (right).

A5 Estimated Policies: Implementation Phase

In this section, we provide details on the implementation phase of the targeted policies. While collecting the data using the uniform randomization, we simultaneously developed the technology for deployment of targeted policies. The goal of this exercise was to test the technological constraints with the advisory system and update the targeted policies according to our learning. For the second week, we estimated and evaluated a crude policy ($\hat{\pi}_A$), where the treatment arms were morning, afternoon, and evening trained on uniform randomization data in week 1 (refer to Figure 2).

Next, as we collected more data using the randomization between 91 call times for additional weeks, we updated the policy ($\hat{\pi}_B$) to take into account the hour of the call along with the day of the week. We could only test the updated policy for a limited number of days for the week of November 1st to 7th (week 4) because, during this week, the call center only operated for five days instead of seven (this week had two major festivals, Diwali and Bhai Dooj). Next, we tested another intermediate policy ($\hat{\pi}_C$) in week 5. This was the first week that we could implement the targeted policy with 91 treatment arms for all seven days. However, during this time, we were also trying to learn about the capacity constraints and bandwidth limits for the call center. We deployed and tested the intermediate policy for week 5 but with tighter constraints. For the last week, week 6, we deployed the policy for all seven days and relaxed the constraint further to accommodate additional farmers in their best-predicted call times ($\hat{\pi}_D$). Ideally, we would want to implement a few more targeted policies taking into account the possibility of shocks, but we used off-policy evaluation to estimate and evaluate those additional policies. We only had time to deploy a few limited, targeted policies during the six weeks.

Table A13: Implementation Phase (Week 2)

Data Collection Method	$\hat{\pi}_A$	Uniform Policy	Difference
All	0.3030	0.3178	-0.0148
	[0.0009]	[0.0006]	[0.0011]
N	265,188	616,656	

Notes: This table shows the on-policy evaluation for $\hat{\pi}_A$ in week 2. The farmers are randomized between two groups. Group A gets called according to $\hat{\pi}_A$ and group B gets called according to uniform randomization. Sample means are used to estimate the value of $V(\hat{\pi_A}, \mathcal{S}^{\text{eval}})$ and $\bar{V}(\mathcal{U}, \mathcal{S}^{\mathcal{U}})$.

Table A14: Implementation Phase (Week 4)

Data Collection Method	$\hat{\pi}_B$	Uniform Policy	Difference
All	0.3125	0.3172	-0.0047
	[0.0011]	[0.0006]	[0.0013]
N	167,995	707,644	

Notes: This table shows the on-policy evaluation for $\hat{\pi}_B$ in week 4. The farmers are randomized between two groups. Group A gets called according to $\hat{\pi}_B$ and Group B gets called according to uniform randomization. Sample means are used to estimate the value of $V(\hat{\pi_B}, \mathcal{S}^{\text{eval}})$ and $\bar{V}(\mathcal{U}, \mathcal{S}^{\mathcal{U}})$.

Table A15: Implementation Phase (Week 5)

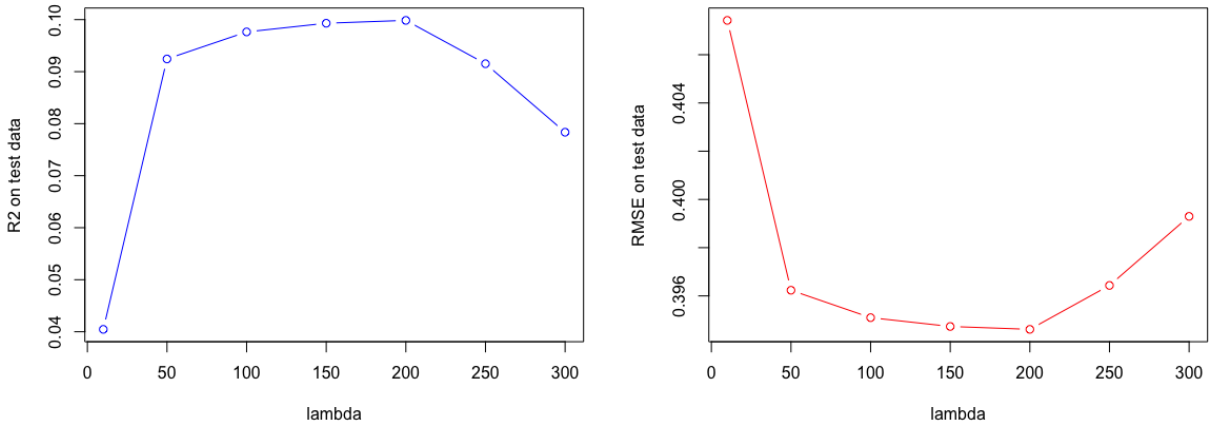
Data Collection Method	$\hat{\pi}_C$	Uniform Policy	Difference
All	0.3133	0.3195	-0.0063
	[0.0011]	[0.0006]	[0.0011]
N	234,276	624,801	

Notes: This table shows the on-policy evaluation for $\hat{\pi}_3$ in week 5. The farmers are randomized between two groups. Group A gets called according to $\hat{\pi}_C$ and Group B gets called according to uniform randomization. Sample means are used to estimate the value of $V(\hat{\pi_C}, \mathcal{S}^{\text{eval}})$ and $\bar{V}(\mathcal{U}, \mathcal{S}^{\mathcal{U}})$.

A6 Matrix Completion on Historical Data

We describe the matrix completion exercise of the historical data in this section. Note that the historical data for the experimental sample is available from July 2018 to September 2021. However, we do not have the historical engagement for every farmer for each of the 91 call times, as calls were scheduled in an ad hoc manner in the past. The non-missing entries constitute about 49% of the total number of entries. Hence, a matrix completion algorithm is needed to complete the

Figure A13: Nuclear norm penalty for the Matrix Completion on Past Engagement



a. Out of Sample R^2 b. RMSE

Notes: (a) shows the R^2 on the test data. (b) shows the RMSE corresponding to different values of the nuclear norm penalty.

past engagement matrix. The Soft-Impute algorithm described in Mazumder et al. (2010) is used for the matrix completion exercise. This algorithm uses nuclear norm regularization for matrix completion.

As described in Mazumder et al. (2010), the optimization problem for completing the historical engagement matrix Y_{history} is provided below. Note that the matrix Y_{history} has dimension $N \times J$ where each (i, j) element corresponds to the historical engagement of farmer i in call time j.

$$\min_{M} \frac{1}{2}||P_{\Omega}(Y_{\text{history}} - M)||_F^2 + \lambda||M||_*$$

, where $||M||_*$ is the sum of singular values of M and $P_{\Omega}(Y_{\text{history}})$ is the projection matrix with observed elements in Y_{history}.

The regularization parameter is chosen using the following steps. The matrix elements are split 80:20 into training and test entries. The non-missing entries in the training matrix that are part of the test IDs are made NAs for this exercise. We choose several values of λ and do the matrix completion using Soft-Impute. Both the R-MSE and out-of-sample R^2 were computed for the test

entries. The λ corresponding to the lowest R^2 was chosen as the optimal λ. Panels (a) and (b) of Figure A13 display these measures for the different values of λ. The matrix completion exercise provides a 10% improvement over imputing the missing entries with the mean of the matrix. The predicted past engagement is used as a covariate for estimating and evaluating targeted policies on the uniform randomized data.

A7 Proofs

Proof of Lemma 1. Starting from the definition of policy value in week s:

$$V_s(\hat{\pi}) = \mathbb{E}_{X \sim F_X}[\mu(X, \hat{\pi}(X), s)]$$

For any auxiliary week t, we add and subtract $\mu(X, \hat{\pi}(X), t)$:

$$V_s(\hat{\pi}) = \mathbb{E}_X[\mu(X, \hat{\pi}(X), t)] + \mathbb{E}_X[\mu(X, \hat{\pi}(X), s) - \mu(X, \hat{\pi}(X), t)]$$

Since the first term equals $V_t(\hat{\pi})$ by definition:

$$V_s(\hat{\pi}) = V_t(\hat{\pi}) + \mathbb{E}_X[\mu(X, \hat{\pi}(X), s) - \mu(X, \hat{\pi}(X), t)]$$

Multiplying both sides by w_t and summing over all auxiliary weeks with $\sum_t w_t = 1$ yields:

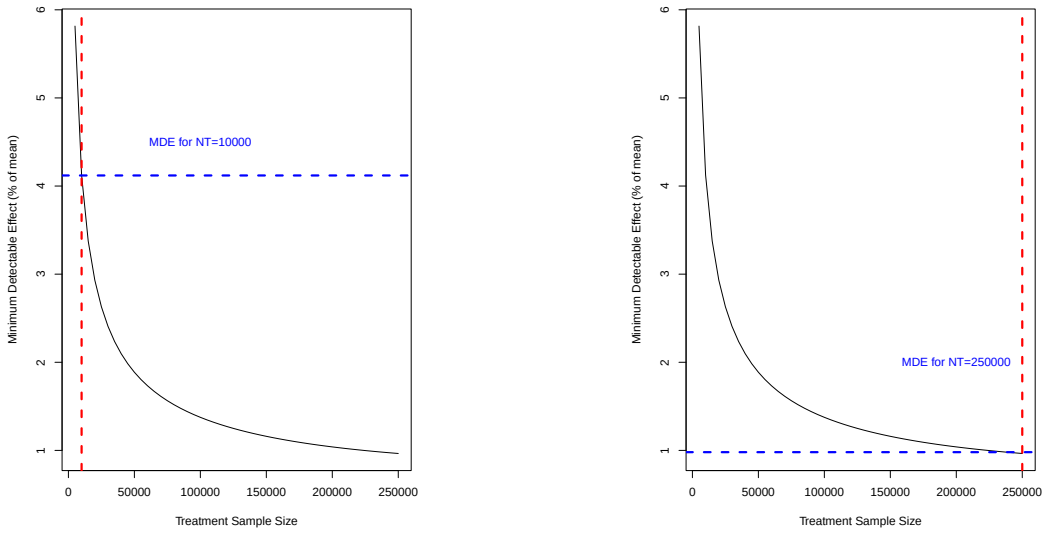
$$V_s(\hat{\pi}) = \sum_{t \in \mathcal{T}_{\text{aux}}} w_t V_t(\hat{\pi}) + \sum_{t \in \mathcal{T}_{\text{aux}}} w_t \mathbb{E}_X[\mu(X, \hat{\pi}(X), s) - \mu(X, \hat{\pi}(X), t)]$$

■

A8 Power Calculations

This section provides an outline of the power calculations that we did before starting the multistage experiment. The goal of this exercise is to assess the minimum detectable effect for the off-policy and on-policy evaluations that we intended to do for each week. For each N, we estimate the minimum detectable effect for the differences in the estimate of the value of targeted policy and uniform policy under off-policy and on-policy evaluations. For the off-policy evaluation, let us assume that 900,000 farmers are allocated to uniform randomization in a week. This implies about $1/91 \times 900,000$ would be receiving calls as per the targeted policy under the uniform randomization data. Varying the sample size for those getting called according to $\hat{\pi}$ and those according to the uniform policy, we illustrate the relationship between the minimum detectable effect and the sample size. Figure A14 shows that under the sample sizes close to the ones we see in our project, the MDE for off-policy evaluation is much higher than the MDE for on-policy evaluation. This is happening as we can allocate a higher sample size (250,000) to receive calls according to the targeted policy under on-policy evaluation.

Figure A14: Power Calculations



a. Off-Policy Evaluation

b. On-Policy Evaluation

Notes: Panel (a) shows the relationship between the MDE and sample size for the off-policy evaluation. Panel (b) shows the relationship between MDE and sample size for the on-policy evaluations.