RCTs, Awareness, and Assignment Effects

Florencia M. Hnilo*

January 5, 2026†

Please Click Here for the Latest Version

Abstract

Randomized Controlled Trials (RCTs) are the gold standard for evaluating the effects of interventions because they rely on simple assumptions. Their validity also depends on an implicit assumption: that the research process itself, including how participants are assigned, does not affect outcomes. In this paper, I challenge this assumption by showing that outcomes can depend on the subject's knowledge of the study, their treatment status, and the assignment mechanism. I design a field experiment in India around a soil testing program that exogenously varies how participants are informed of their assignment. Villages are randomized into two main arms: one where treatment status is determined by a public lottery, and another by a private computerized process. My design temporally separates assignment from treatment delivery, allowing me to isolate the causal effect of the assignment process itself. I find that estimated treatment effects differ across assignment methods and that these effects emerge even before the treatment is delivered. The effects are not uniform: the control group responds more strongly to the assignment method than the treated group. These findings suggest that the choice of assignment procedure is consequential and that failing to account for it can threaten the interpretation and generalizability of standard RCT treatment effect estimates.

JEL Classification: C93, D91, O12

Keywords: Assignment, Awareness, Causal Inference, RCTs

*Stanford University, Department of Economics (fhnilo@stanford.edu).

†I am especially grateful to Ran Abramitzky, Arun Chandrasekhar, Marcel Fafchamps, Melanie Morten, and Alessandra Voena for their invaluable support and guidance. I also thank the valuable comments and suggestions of Mohamad Adhami, Daliah Al-Shakhshir, Sebastián Bauer, Gabriele Cristelli, Milena Dotta, Pascaline Dupas, João Francisco Pugliese, Akhila Kovvuri, Eva Lestant, Madeline McKelway, Alexia Olaizola, Dev Patel, Anirudh Sankar, Jason Weitze, and seminar participants at Stanford, J-PAL, and UC Davis. I am grateful to Meenu Mohan and Mithin Nehrubabu for excellent research assistance and indebted to Praveen Banagere, Sharadha Chandankar, and Gayathri Sowrirajan for their continuous support during fieldwork. I gratefully acknowledge financial support from the B. F. Haley and E. S. Shaw Fellowship for Economics, and the Ric Weiland Graduate Fellowship, as well as multiple grants from the George P. Shultz Dissertation Support Fund at SIEPR, the Weiss Fund at the University of Chicago, and the King Center at Stanford University. This paper's protocol has been approved under the Extended Approval Process at Stanford University by the Panel on Non-Medical Human Subjects, IRB protocol number 66074, and by the IFMR Human Subjects Committee under the title of "RCTs & Belief Distortion", and registered on the AEA Social Science Registry on July 08, 2024 (AEARCTR-0013923).

1 Introduction

Nothing is wrong with making assumptions;
causal inference is impossible without making assumptions,
and they are the strands that link statistics to science.
It is the scientific quality of those assumptions,
not their existence, that is critical.

(Rubin, 2005, p. 324)

Governments, organizations, and researchers seek to understand the effects of interventions to make informed decisions, guide policy, and analyze counterfactuals. Through diverse research designs, they test interventions and use the resulting knowledge under the implicit assumption that the research process itself does not affect the potential outcomes (Rubin, 2005). This assumption provides confidence in the accuracy of the study's results and enables the findings to be generalized to other settings.

To illustrate, consider a researcher aiming to estimate the effect of a cash transfer program on employment, Y. The researcher samples subjects from a relevant population and conducts a Randomized Controlled Trial (RCT). Subjects are randomly assigned through a statistical software to either a treatment group that receives the cash transfer ($t = 1$) or a control group that does not ($t = 0$). Following the intervention, the researcher estimates the Average Treatment Effect (ATE), which is the expected difference in effective outcomes: $\tau = \mathbb{E}[Y_i(t_i = 1) - Y_i(t_i = 0)]$.

Imagine those assigned to the control group realize that they have not been selected and feel demoralized, reducing their job search. If that is the case, the interpretation of the ATE is not straightforward. Even setting aside complexities such as general equilibrium effects or spillovers, the research process itself would have altered the subjects' environment. I argue this occurs through a change in what I term the subject's "awareness state," which includes: (i) awareness of participating in a study ($s \in \{0, 1\}$); (ii) awareness of one's own and others' treatment status ($a \in \{0, 1\}$); and (iii) awareness of how assignment into the treatment group or the control group took place ($k \in \{0, 1\}$).[1] A central point of this paper is that the components of the awareness state may not be easily separable from treatment and could therefore introduce bias into standard treatment effect estimates, raising internal and external validity concerns. In the cash transfer program example, the difference in employment rates between the control and treatment groups would capture the effect of the cash transfer together with the demoralization effect. Moreover, it is unclear that the effects of the cash transfer program can be replicated unless the same awareness state is reproduced.

In this paper, I study whether and to what extent subjects' awareness of the assignment method interferes with the treatment effect by varying a and k independently of the treatment status. My

[1] The literature has already identified the threat of some of these effects, as discussed in Duflo et al. (2007). In my framework, variation in s corresponds to the Hawthorne effect (individuals change their behavior when they know they are being observed), and variation in a is linked to the John Henry effect (subjects in the control group try to overcompensate for their exclusion from the intervention).

contribution is fourfold. First, I formalize the notion of an awareness state, developing a formal taxonomy of the potential outcomes that arise from the interaction of treatment and the awareness state. Second, I conduct a proof-of-concept field experiment embedded within a standard agricultural intervention in India. The experimental design exogenously varies two components of the awareness state, a and k, separately from the treatment distribution to test whether their interaction with the treatment status is negligible. I find that awareness of the assignment method cannot be ignored: the treatment effects vary depending on subjects' understanding of the assignment method (k), and the effects appear immediately after subjects know their assignment into the treatment or control group (a), but only for the control group. Third, I develop a set of non-parametric diagnostic tests to determine if the treatment effect can be separated from the assignment method awareness effect. Finally, I explore which underlying behaviors might be driving the results, proposing avenues to model and test the behavioral mechanisms activated with different awareness states.

My field experiment, designed around an existing soil quality testing program in India, allows me to isolate the impact of the research process from the treatment effect. The intervention provides small farmers with information about their soil quality, along with tailored recommendations, aiming to reduce excessive fertilizer use. The first feature of my design is the temporal separation of the assignment stage, where farmers learn their treatment status (so a is activated), from the distribution stage when the intervention is delivered (so t is activated for the treatment group). This separation is crucial because a standard RCT confounds any response to the assignment process (i.e., an assignment effect) with the effects of the intervention itself, varying a, k, and t simultaneously. Yet, the temporal separation between the assignment stage and the distribution stage is not enough to capture the effects coming solely from the assignment stage: if I were to analyze a single assignment method, the effects found after the assignment stage could correspond to subjects responding to knowing the intervention will soon arrive (i.e., anticipation effects).

Therefore, my experimental design has two arms with identical timing and interventions, differing only in the assignment method, which varies subjects' understanding of the assignment method (k). The intuition behind the design is as follows: if the assignment stage has no direct and independent effect on the outcomes, both arms should exhibit identical behaviors before (and after) the distribution of the treatment, even if that means they are reacting in anticipation. If, however, they show different anticipatory effects, that implies that the assignment stage is generating an independent effect.

I randomize clusters of villages into the two most popular random assignment methods used in field experiments: a public lottery arm and the computer-generated random selection arm. The Computerized Randomization method represents the approach most commonly used in RCTs, while the Public Lottery method is usually chosen for its transparency and perceived fairness from the participants' perspective (Pathak and Sethuraman, 2011). While both methods are random, the assignment methods vary in the visual appeal and ease of understanding, yielding $k^{\text{Public Lottery}} > k^{\text{Comp. Rand.}}$. Within villages randomized into one of these two arms, farmers were assigned to treatment using the corresponding assignment method. In both cases, awareness of one's own and others' treatment status (a) is the same: during the assignment step, in both arms, participants were invited to assemble in a public space, and the names of those selected into treatment and into control were read aloud. By comparing the Public Lottery and the Computerized Randomization arms, any outcome differences observed between the assignment and distribution stages can only be attributed to the assignment method rather than effects in anticipation of treatment distribution.

The soil quality testing program provides an ideal setting for my experimental design for three reasons. First, it treats only a subset of farmers within a village, enabling within-village comparisons. Second, a one-month interval separates treatment assignment from the delivery of test results, creating a window to observe anticipatory effects and isolate the impact of assignment itself from information receipt. Third, soil testing represents a standard agricultural intervention, facilitating comparison with existing studies. Moreover, the soil test intervention allows me to measure how farmers interpret and respond to the intervention. Specifically, I can assess whether farmers attribute the test results (which can be either positive or negative depending on their previous expectations) to their actions or external factors, providing further insight into the behavioral impacts of the assignment method.

To know the effects of the soil quality test intervention, I collect a large set of variables related to financial management decisions, on-farm practices, and knowledge of new soil practices. Moreover, I measure several variables related to potential behavioral mechanisms activated through changes in the awareness state: variables on attribution, risk aversion, well-being, and social networks. The variables create a panel across three surveys: the baseline survey (survey 1), a midline survey conducted between the assignment into treatment and the distribution of the treatment (survey 2), and an endline survey done two weeks after the treatment distribution (survey 3). To analyze this large number of variables, I calculate an index for groups of related survey variables. I use both subjective indices, grouping variables according to my survey design, and data-driven indices. This last approach groups the outcomes according to the semantic content of the survey questions, using a natural language processing model to generate embeddings that capture shared aspects of the questions' meaning and context.

Building on this experimental design and focusing on the financial management decisions, on-farm practices, and knowledge of new soil practices indices, I test a series of hypotheses motivated by the conceptual role of the awareness state in shaping causal effects. First, I investigate whether differences in participants' awareness and assignment context can lead to systematically different average treatment effects. Second, I test if these differences appear in anticipation of the treatment distribution.

I find that the assignment method matters and that only the Public Lottery arm yields significant treatment effects. The assignment method not only changes the magnitude of the treatment effects, but it can also reverse the effects' sign: the Public Lottery arm shows negative average treatment effects, while there is no effect in the Computerized Randomization arm. These findings raise concerns about internal and external validity. The internal validity concern appears as the treatment effect is confounded with the effect coming from the assignment process: the same intervention yields different ATEs depending on how the randomization is presented to participants. The external validity concern arises because it is unclear what we can learn from this RCT: unless we reproduce the same awareness state, we may not obtain the same ATE.

I do not find evidence indicating differential average treatment effects in anticipation, but I find differences inside each group: treatment and control. My experimental design allows for analyzing the behavior of the control subjects and of the treated subjects separately, so I include direct comparisons of treated groups across arms and control groups across arms. Analyzing the control and treatment groups separately, I find significant differences in the control group's behavior depending on the assignment method; I find no differences for the treated groups.

To further inspect the differential outcomes of each group between the Public Lottery and the Computerized Randomization arms, I examine their change with respect to a benchmark. In the case of the control group, the cleanest benchmark corresponds to places that are never exposed to the intervention, as spillovers from treated subjects to control subjects are then impossible. Hence, I randomize clusters of villages into a pure control arm, where no farmer knows about the intervention. I also randomize clusters of villages into a pure treatment arm, where every farmer in the village receives the intervention. Then, I compare the treated farmers in the Public Lottery and the Computerized Randomization arm with those in the pure treatment group, and similarly for the control farmers. This approach reveals how the assignment method affects participants differentially depending on the group to which they were assigned, thereby influencing the measured impact of the intervention.

Examining the reactions of the control subjects, I demonstrate that assignment effects are present in both the Public Lottery arm and the Computerized Randomization arm, but not uniformly. Control subjects assigned via a public lottery exhibit distinct behavioral responses relative to the pure control benchmark, changing their farm practices, input use, and financial management practices in advance. The controls in the Computerized Randomization arm are indistinguishable from those in the Pure Control in terms of agricultural outcomes. Still, they do exhibit distinct behavior compared to the Pure Control in some of the other indices. For example, controls in the Computerized Randomization arm attribute more of their success to their personal effort. After the distribution of the treatment, the control group in the Public Lottery continues to diverge from the benchmark, showing more adoption of the information provided, changes in financial management, and higher risk aversion. Controls in the Computerized Randomization arm only exhibit significant differential changes in financial management. The findings in the control group do not replicate when analyzing the treatment groups. For the treated group, the mixed arms do not differ from the Pure Treatment arm at any moment. This suggests an asymmetric impact of the assignment method: in my setting, it only differentially affects the control group, not the treated group.

The asymmetric impact of the assignment between the treated and control groups creates an identification challenge. If the assignment method affects only control subjects, the awareness state may interact with the treatment status $t \in \{0, 1\}$ in ways that confound causal inference. This interaction determines the appropriate empirical strategy. If awareness effects and treatment effects are additively separable, incorporating awareness as a fixed effect suffices to recover unbiased treatment estimates. However, if awareness and treatment status interact non-additively (such that awareness influences *how* farmers respond to being part of the control group) then fixed effects alone cannot identify the treatment effect of interest. In this case, modeling informed by behavioral theory becomes necessary to separately identify the direct effect of treatment from awareness-mediated channels.

I propose a set of logical inequalities for formally testing whether the awareness state operates additively or interacts with treatment status. These tests provide (i) a set of necessary but not sufficient conditions for separability, and (ii) upper bounds on the extent of the bias that the assignment method creates. In my setting, I find that the treatment effect is separable from the assignment effect, but I stress the importance of testing whether separability holds in other settings.

Even when the test suggests that the awareness state is separable from the treatment effects, the heterogeneity in outcomes across arms and groups indicates the existence of group-specific behavioral mechanisms in action. The interaction of treatment status and assignment method may influence how individuals perceive the fairness of the intervention, update beliefs about their own abilities, or attribute their success and failure to luck or effort. In practice, farmers may form new self-assessments or alter social comparisons based on their group assignment and how it was conducted. While my empirical design does not allow for precise pinpointing of the mechanisms, the patterns observed strongly suggest that assignment-induced inferences drive the behavioral changes documented in the analysis.

This paper contributes to the literature highlighting how behavioral responses to the research process can contaminate estimates from RCTs, particularly where blinding is difficult (Barrett and Carter, 2010; List, 2011; Deaton and Cartwright, 2018b). A crucial insight from this literature is that the context in which an experiment is conducted can matter as much as the treatment itself. For instance, Ding and Lehrer (2010) show that treatment effects in an educational intervention were significantly larger when the treatment was more exclusive, suggesting that awareness of treatment rarity induced a behavioral response in teachers and students. Prior work also demonstrates that outcomes are affected by subjects' awareness of being in a study (a Hawthorne effect, corresponding to variation in s in my framework) or their knowledge of treatment assignment (corresponding to variation in a) (Duflo, Glennerster, and Kremer, 2007; Goldberg, 2017; Chassang, Padró i Miquel, and Snowberg, 2012; Bloom, Liang, Roberts, and Ying, 2015). However, the literature does not systematically analyze all these effects, and in particular does not analyze if the method of random assignment itself directly impacts participants' behavior through their understanding of it (k).

I advance this literature in three ways. First, I develop a conceptual framework that formalizes a participant's awareness state to systematically categorize these research-induced effects. Second, I provide the first experimental evidence, to my knowledge, that exogenously varies the assignment method (k) to isolate its causal impact on outcomes. Finally, I offer a set of diagnostic tests for researchers to assess the separability of the assignment effects from the treatment effect which can also be extended to the other components of the awareness state. While discussion of such biases is not new, it is rare in practice in particular in RCTs (Peters, Langbein, and Roberts, 2018). My paper provides a tangible framework and tools to advance the discussion in a structured and testable way.

Within this framework, I empirically test the possibility that control subjects react to being assigned to control. This questions the belief that research participants are used to not getting programs and to arbitrariness, making randomization appear transparent and legitimate (Banerjee and Duflo, 2009; Glennerster, 2014). By experimentally comparing two standard assignment methods and benchmarking against a pure control arm, my results shed light on whether all forms of random assignment are perceived as equally transparent or whether procedural features matter for how participants interpret and respond to their status.

Accordingly, my findings intersect with the procedural utility literature, which emphasizes that subjects value not only outcomes but also fairness and transparency in allocation procedures (Frey, Benz, and Stutzer, 2004). Experimental evidence shows persistent aversion to lotteries and a preference for meritocratic mechanisms, even when alternatives are equally arbitrary in practice (Bouacida and Foucart, 2025). My research extends this literature by empirically linking procedural perceptions directly to experimental design and response, demonstrating that assignment mechanisms are not neutral from the perspective of participants.

By testing the two assignment methods, I also add to a parallel body of research that examines how individuals perceive and respond to randomness and assignment. Kahneman and Tversky (1972) and Benjamin (2019) document widespread biases in reasoning about random processes, including representativeness heuristics and errors in belief updating. While many different behavioral biases have been precisely recorded in laboratory settings (Möbius, Niederle, Niehaus, and Rosenblat, 2022), in this paper I suggest which of the possible mechanisms might be at play behind the assignment bias, testing if the same laboratory behaviors can be generalized to the field.

Rather than engaging with broader debates on whether randomized experiments should be used for causal inference (Heckman, 2020; Deaton and Cartwright, 2018a), I focus on the assumptions inherent in randomization itself. My results show that assignment methods invoke distinct behavioral and psychological responses (which I term assignment effects), highlighting a new layer of complexity that arises after the decision to randomize has been taken. My paper extends beyond critiques of selection or participation bias to demonstrate that the process and context of assignment are themselves essential to understanding and interpreting experimental results.

The implications of my findings extend beyond the scope of the specific intervention tested and RCTs as a method to evaluate policies. My results indicate that subjects' awareness state must be recognized as a significant source of changes that, if ignored, could create biases in field experiments. When assignment awareness is not controlled or standardized, RCTs may yield biased or non-generalizable results, potentially undermining their value for guiding policy decisions. By demonstrating how the assignment method can shape both control and treatment group outcomes, I show the necessity for future RCTs to measure and report on subjects' awareness states. I propose the standardization in reporting assignment and awareness practices in randomized trials, together with careful measurement of behavioral outcomes to capture the impact of interventions across different settings. To better understand the persistence of the effects, I plan to conduct a second endline survey in November 2025.

This paper demonstrates that the process and awareness of random assignment are integral to the estimation and interpretation of causal effects in RCTs. By foregrounding the awareness state, the analysis reveals new threats to external validity and comparability in experimental studies, illustrates the importance of standardizing assignment procedures as a tool for minimizing bias, and highlights the need to identify which behaviors are triggered by different awareness states.

2 Conceptual Framework

In this section, I develop a framework to test whether the research process itself alters potential outcomes. Standard SUTVA requires "no hidden variations of treatment," but I argue that the awareness induced by the research process creates precisely such variations. To discuss this, I introduce the concept of awareness state and incorporate it into the potential outcomes framework.

When evaluating the causal effect of an intervention on outcome Y, researchers design mechanisms to assign subjects to either the treatment group, where they are exposed to the intervention ($t = 1$), or the control group, where they are not ($t = 0$). Researchers aim to design an assignment mechanism that sorts comparable subjects to each of these groups to calculate an unbiased estimator of the average causal treatment effect (ATE). The standard potential outcomes framework (Rubin, 2005) defines the ATE as:

$$\mathbb{E}[Y_i(t_i = 1) - Y_i(t_i = 0)] = \tau$$

For this comparison to be correct, researchers need to make the stable unit treatment value assumption (SUTVA), which assumes: (1) no interference between units, and (2) no hidden versions of treatments (Rubin, 1980; Cox, 1958)[2]. The random assignment mechanism used in RCTs not only complies with SUTVA,[3] but is also assumed to be "ignorable" (Rubin, 1976, 1978).

An assignment mechanism is ignorable or unconfounded if the treatment assignment is independent of the potential outcomes, controlling for covariates. This means that any observed or unobserved factors that influence outcomes do not also influence treatment assignment. Technically, the random assignment in RCTs is ignorable by design. However, behind the technical term of ignorability, there is an assumption on the assignment mechanism not introducing any systematic bias related to the outcomes of interest: "An assumption that is also implicit (...) is that the sci-

[2]This last assumption itself includes the consistency assumption, which ties the potential outcomes to the observed data. Formally, it poses that $Y_i(t) = Y_i$ when $T_i = t$, i.e., the value of Y which would have been observed if T had been set to what in fact was is equal to the value of Y which was in fact observed (VanderWeele and Hernan, 2013).

[3]The literature has already found ways to relax SUTVA even in the special case of random assignment: We allow subjects to interfere with each other and call those interferences "spillovers". Recognizing the existence of spillovers allows researchers to assume structural functions to capture them, and employ empirical saturation strategies to measure the effective treatment effect. For a thorough literature review on spillovers, see Muralidharan and Niehaus (2017) and Han, Basse, and Bojinov (2024).

ence—the covariates and the potential outcomes—is not affected by how or whether we try to learn about it, whether by completely randomized experiments, randomized blocks designs, observational studies, or another method" (Rubin, 2005).

I argue that although the random assignment mechanism technically complies with ignorability, the implementation of the assignment mechanism can introduce new factors that alter the potential outcomes. In particular, I argue that the potential outcomes themselves are not stable between different research processes. Then, I formalize a specific violation of SUTVA that arises from the research process itself, which Rubin (2005) alluded to when he cautioned that the "science" might be affected by "how we learn about it."

Therefore, let's assume that the research process changes potential outcomes, even when the assignment is random. In this case, we need to model that when a subject is exposed to an intervention, it is not only the treatment that affects her outcomes, but also what I term her "awareness state."

Definition 1. The awareness state of subject i is:

$$\boldsymbol{\alpha}_i := (a_i, k_i, s_i),$$

where $a_i \in \{0, 1\}$ denotes if subject i is aware of the existence of different assignment groups and her own status within them; $k_i \in \{0, 1\}$ indicates subject i's awareness or understanding of the assignment mechanism; and $s_i \in \{0, 1\}$ denotes whether subject i is aware she is part of a research study or, more generally, if she is aware she is being watched.

Intuitively, the awareness state of subject i is the subject's awareness of the various aspects of the research intervention she is part of.[4] To illustrate, a can be active for a control subject when she realizes someone else got the treatment and she did not, thus potentially activating the John Henry effect or competition between the control subjects and the treated subjects. By "awareness of the assignment mechanism," k, I mean that subject i understands the rule the experimenter set to exclude individuals from receiving treatment.[5] Finally, variation in s_i captures the Hawthorne effect: subjects being more reactive because they know they are being observed.

Note that varying t together with the three components of $\boldsymbol{\alpha}_i$ yields sixteen possible potential outcomes, listed and explained in Appendix A.[6] Also note that the awareness state does not make any statements or assumptions on the subject's beliefs or understanding of the assignment method.

I can then define the potential outcome of subject i as:

$$Y_i(t_i, \boldsymbol{\alpha}_i) = Y_i(t_i, a_i, k_i, s_i)$$

[4]The awareness state framework can be extended to include, for example, the identity of the intervention implementer (Shenoy and Lybbert, 2024).

[5]It is possible to think of k as a continuous parameter. However, for exposition purposes, I consider it a dichotomous variable: a subject understands the assignment rule, or she does not.

[6]That is, sixty-four possible pairings of treatment and control subjects assuming the experimenter sets different $\boldsymbol{\alpha}$s for the treated and control subjects. For example, the experimenter might inform the treated group about the assignment statuses ($a = 1$), but might not share this information with the control group.

Thus, the unit-level causal treatment effect for subject i is:[7]

$$Y_i(t_i = 1, \boldsymbol{\alpha}_i)) - Y_i(t_i = 0, \boldsymbol{\alpha}_i)) = \tau_i(\boldsymbol{\alpha}_i)$$

RCTs assume the research process to be ignorable, implicitly assuming that the awareness state does not affect potential outcomes. This is equivalent to assuming a strong form of SUTVA where $Y_i(t_i, \boldsymbol{\alpha}) = Y_i(t_i)$ for all possible $\boldsymbol{\alpha}$. Under this assumption, the various awareness-contingent potential outcomes collapse into the standard ones, and the estimand simplifies to the familiar ATE:

$$\mathbb{E}[Y_i(t_i = 1, \boldsymbol{\alpha}_i^{\text{Treated}}) - Y_i(t_i = 0, \boldsymbol{\alpha}_i^{\text{Control}})] = \mathbb{E}[Y_i(t_i = 1) - Y_i(t_i = 0)] = \tau, \ \forall \boldsymbol{\alpha}_i$$

However, if this strong SUTVA assumption is violated (that is, if the "science" is indeed affected by how we learn about it) then the quantity estimated from a simple comparison of means in an RCT is not the standard ATE. Instead, it is an average treatment effect that is specific to the particular awareness states induced by the experimental design:

$$\mathbb{E}[Y_i(t_i = 1, \boldsymbol{\alpha}_i^{\text{Treated}}) - Y_i(t_i = 0, \boldsymbol{\alpha}_i^{\text{Control}})] = \tau(\boldsymbol{\alpha}^{\text{Treated}}, \boldsymbol{\alpha}^{\text{Control}}) \tag{1}$$

The estimand $\tau(\boldsymbol{\alpha}^{\text{Treated}}, \boldsymbol{\alpha}^{\text{Control}})$ is therefore a composite of the direct effect of the treatment and the indirect effect of the treatment operating through the change in awareness. This paper addresses the identification and measurement of this estimand.

3 Hypotheses on Existence and Separability

The conceptual framework above postulates that the ATE is not a fixed quantity but is instead conditional on the participant's awareness state, $\boldsymbol{\alpha}_i$. This central claim leads to a set of directly testable hypotheses. The first hypothesis states that participants' awareness of the research process directly influences outcomes; the second hypothesis refers to the separability of awareness from the treatment effect.

Hypothesis 1. *The subjects' awareness state directly impacts the potential outcomes.*

The first and most fundamental hypothesis directly evaluates the main argument: that a change in the experimental context, as captured by the awareness state, will alter the measured impact of the intervention. To test this, I propose two testable sub-hypotheses.

Hypothesis 1a. *The average treatment effect does not depend on the awareness state.*

$$\tau(\boldsymbol{\alpha_i}) = \tau(\boldsymbol{\alpha_j}) = \tau, \ \textit{with } \boldsymbol{\alpha_i} \neq \boldsymbol{\alpha_j}$$

[7] I ignore spillovers for the sake of simplicity, but the model can easily be extended to allow interference across units, allowing Y_i to depend on other agents' treatment status and outcomes.

Suppose the research process is truly ignorable. In that case, the ATE should remain constant regardless of whether participants are aware of the assignment or how it was conducted. A rejection of the null hypothesis ($\tau(\boldsymbol{\alpha_i}) \neq \tau(\boldsymbol{\alpha_j})$) would provide strong evidence that the context of assignment is not neutral and that failing to account for it confounds the estimation of the true treatment effect.

If Hypothesis 1a is rejected, a natural follow-up question is whether these effects can be predicted in anticipation. That is, if part of the different treatment effects are due to being assigned to a group, conditional on participants knowing or inferring they are in one of the groups. Imagine that the assignment is conducted before and temporally separated from the distribution of the treatment. In that case, it would be possible to directly measure if participants' behavior starts to differ once they know their group assignment and before the treatment effects activate. This difference is what I term assignment effects:

Definition 2. The **assignment effects** are the difference in outcomes born from the assignment into treatment or control, conditional on participants knowing or inferring their group affiliation. The assignment effects are prior to and independent from the treatment effect.

One of the ways to test for the existence of assignment effects is to look at differential anticipation effects. Then, I propose to test Hypothesis 1b.

Hypothesis 1b. *Differences in treatment and control groups after the assignment step of an RCT do not depend on the awareness state.*

Intuitively, consider the moment a subject becomes aware of the assignment rule and her status, triggering behavioral responses. Assignment effects could include changes in motivation, effort, or investment, driven by feelings of fairness, luck, or deservingness associated with the assignment method. Then, by varying how participants are assigned and what they understand about the process (k), I can test whether the assignment method influences participants' behavior. A rejection of Hypothesis 1b would mean that:

$$k_i \neq k_j \Rightarrow \boldsymbol{\alpha_i} \neq \boldsymbol{\alpha_j} \Rightarrow \tau(\boldsymbol{\alpha_i}) \neq \tau(\boldsymbol{\alpha_j})$$

Even if Hypothesis 1 holds, it may be possible to recover the true treatment effect by controlling for the awareness state. This would mean that the assignment effects and the treatment effect are separable.

Definition 3. Suppose the potential outcomes can be decomposed into:

$$Y_i(t, \boldsymbol{\alpha}) = f(t) + g(\boldsymbol{\alpha}) + h(t, \boldsymbol{\alpha})$$

where f depends only on the treatment status, g depends only on the awareness state, and h captures any non-additive interaction between t and $\boldsymbol{\alpha}$. The treatment effect is **separable** from the awareness state if and only if $h(t, \boldsymbol{\alpha}) = 0 \ \forall \ t, \boldsymbol{\alpha}$.

Non-separability would imply that there exists an interaction between the treatment status and the awareness state that cannot be accounted for by controlling for the awareness state. This leads me to my second hypothesis.

Hypothesis 2. *The awareness state is not separable from the treatment status.*

If Hypothesis 2 is rejected, it gives a clue on how to address the spillovers created by different awareness states into the estimation of treatment effects. Namely, controlling for the awareness state could isolate the true treatment effect. However, failure to reject Hypothesis 2 indicates that the participant's outcome is intrinsically context-dependent. In such cases, outcomes are determined not solely by the treatment and the responses to the awareness state, but also by the interaction between the treatment and the awareness state, $h(t, \boldsymbol{\alpha})$. This means that to measure causal effects accurately, researchers must account for how the study's design and the information given to participants influence outcomes.

To that end, it is useful to disentangle the impact of the awareness state on the treatment and control groups separately. A change in the ATE could occur if the awareness state only affects the treated, only affects the controls, or affects both but to different degrees. Hypothesis 2a posits that the effects of awareness are uniform across experimental groups. A rejection of Hypothesis 2a would suggest there is an interaction between treatment status t and the awareness state $\boldsymbol{\alpha}$ that needs to be modeled.

Hypothesis 2a. *Different awareness states affect subjects in the control group and the treatment group uniformly.*

To illustrate what would happen if Hypothesis 2a is rejected, assume that the awareness state influences only the outcomes of the control group, so $g^{\text{Control}}(\boldsymbol{\alpha}) \neq g^{\text{Treatment}}(\boldsymbol{\alpha})$. Then, the control group no longer serves as a pure counterfactual for untreated subjects in an unaware context; instead, it reflects a more complex condition: namely, subjects not receiving the treatment while simultaneously being aware of their assignment status, the procedure, and their participation in the study.

Given that a non-separable scenario is possible, we must be able to identify it. Hence, I develop a set of nonparametric tests that can be applied to cross-sectional data and which rely on simple assumptions about the direction of the treatment effects.[8] The tests are derived directly from Propositions 1 and 2, which provide necessary but not sufficient conditions for separability. However, and most importantly, the tests provide a clue to the extent of the non-separability scenario: they imply upper bounds on the interaction term between the treatment status and the awareness state.

Proposition 1. *Let $Y_i(t, a, k, s)$ denote the potential outcome of subject i under treatment $t \in \{0, 1\}$ and awareness state $\boldsymbol{\alpha} = (a, k, s) \in \{0, 1\}^3$. Suppose we can decompose potential outcomes into:*

$$Y_i(t, \boldsymbol{\alpha}) = f(t) + g(\boldsymbol{\alpha}) + h(t, \boldsymbol{\alpha}),$$

[8] The test is in essence a Bell-type inequality as those proposed to study contextuality in quantum systems (Khrennikov, 2018).

Suppose $a = 1$ and $s = 1$ for all cases such that only k varies inside the awareness state, and define the following differences:

$$
\begin{aligned}
T_1 &= \mathbb{E}[Y_i(t=1, k=0) - Y_i(t=0, k=0)] \\
T_2 &= \mathbb{E}[Y_i(t=1, k=1) - Y_i(t=0, k=1)] \\
T_3 &= \mathbb{E}[Y_i(t=1, k=0) - Y_i(t=0, k=1)]
\end{aligned}
$$

Then, under nonnegative treatment effects and monotonicity in k:

$$
Y_i \textit{ is separable } \Rightarrow T_1 + T_2 - T_3 \geq 0 \tag{2}
$$

If inequality 2 does not hold, it not only implies that separability fails, but also that the interaction term $\mathbb{E}[h(1,1) - h(0,0)]$ is sufficiently negative.[9] Refer to Appendix B for Proposition 1's proof.

Proposition 1 not only provides a simple test of separability, but also an upper bound on the interaction term $\mathbb{E}[h_i(1,1) - h_i(0,0)]$. When Y_i is non-separable, it is impossible to recover $\mathbb{E}[\Delta f_i]$ and $\mathbb{E}[\Delta g_i]$. But the differences T_1, T_2, and T_3 are still fully recoverable from the data. Therefore, it is possible to recover a data-based upper bound under the assumptions that the true treatment effect is non-negative and that awareness k has a non-negative direct effect on the outcome:

$$
\mathbb{E}[h_i(1,1) - h_i(0,0)] \leq T_1 + T_2 - T_3
$$

This means that even if treatment and awareness interact in the most adverse way (so that the control when $k = 0$ reacts in a stronger way than the treated group when $k = 1$), the additional outcome change generated by their coaction cannot exceed $T_1 + T_2 - T_3$.

Similarly, and assuming that the interaction term $\mathbb{E}[h_i(1,0) - h_i(0,1)]$ might be flipping the sign of an intervention (so looking at the other diagonal in the space of t, k), I derive a similar diagnostic inequality, although with different assumptions:

Proposition 2. *Define the following additional difference:*

$$
T_3^* = \mathbb{E}[Y_i(t=1, k=1)] - \mathbb{E}[Y_i(t=0, k=0)]
$$

Then, under the assumption that $\Delta f \geq \Delta g$:

$$
Y_i \textit{ is separable } \Rightarrow T_1 + T_2 - T_3^* \geq 0 \tag{3}
$$

As a corollary, if the inequality is broken, then once again the inequality provides an upper

[9] Note that inequality 2 does not impose supermodularity or submodularity. Instead, it is a necessary condition for separability under weak monotonicity assumptions. A violation implies non-separability; a non-violation does not rule out interaction in other directions.

bound for $\mathbb{E}[h_i(1,0) - h_i(0,1)]$:

$$\mathbb{E}[h_i(1,0) - h_i(0,1)] \leq T_1 + T_2 - T_3^*$$

Empirically, it is also possible to derive a separability test without relying on assumptions regarding Δf and Δg.

Proposition 3. *Define*

$$\psi = (\mathbb{E}[Y_i(1,1)] - \mathbb{E}[Y_i(1,0)]) - (\mathbb{E}[Y_i(0,1)] - \mathbb{E}[Y_i(0,0)])$$

Then:

$$Y_i \text{ is separable } \Rightarrow \psi = 0 \tag{4}$$

If $h(t,k)$ is supermodular (increasing differences), then $\psi \geq 0$; under submodularity of $h(t,k)$, $\psi \leq 0$.

Proposition 3 indicates which group is more affected by changes in awareness: the treated group (when $h(t,k)$ is supermodular) or the control group (when $h(t,k)$ is submodular). Refer to Appendix B for the proofs of Propositions 2 and 3.

4 Experimental Design

To test the hypotheses on existence and separability, I need to independently vary subjects' exposure to treatment (t) and the subjects' awareness state ($\boldsymbol{\alpha}$).

Inside the awareness state, I particularly focus on isolating awareness of the assignment method (k). With this in mind, I vary whether the intervention is randomized at the individual level or at the community level. The logic behind this decision is that when an intervention is randomly assigned between communities, subjects within each community are unaware of the possibility of being excluded from the intervention and are therefore not aware of the assignment process. Then, if the whole community j is randomized into treatment or into control, I consider $k_{ij} = 0 \; \forall i \in j$.[10] In contrast, if randomization into treatment is performed within community j, I consider $k_{ij} > 0$ $\forall i \in j$.

Accordingly, I randomize communities into "pure" arms, where everyone within the community is assigned to the same category (treatment or control), and "mixed" arms, where subjects within the community are randomly assigned to either the treatment group or the control group. Inside the pure arms, I define the "pure treatment" arm, where everyone within the community is assigned to receive the intervention ($t_i = 1 \; \forall i \in j$), and the "pure control" arm, where nobody receives it and subjects are only surveyed ($t_i = 0 \; \forall i \in j$). Figure 1 provides an intuitive representation of the design and the variation in outcomes across survey rounds.

[10]There is still a chance that subjects are aware of their community having been randomly chosen for treatment across many communities and that they react to it, but I assume the effect to be constant across individuals of the same community.

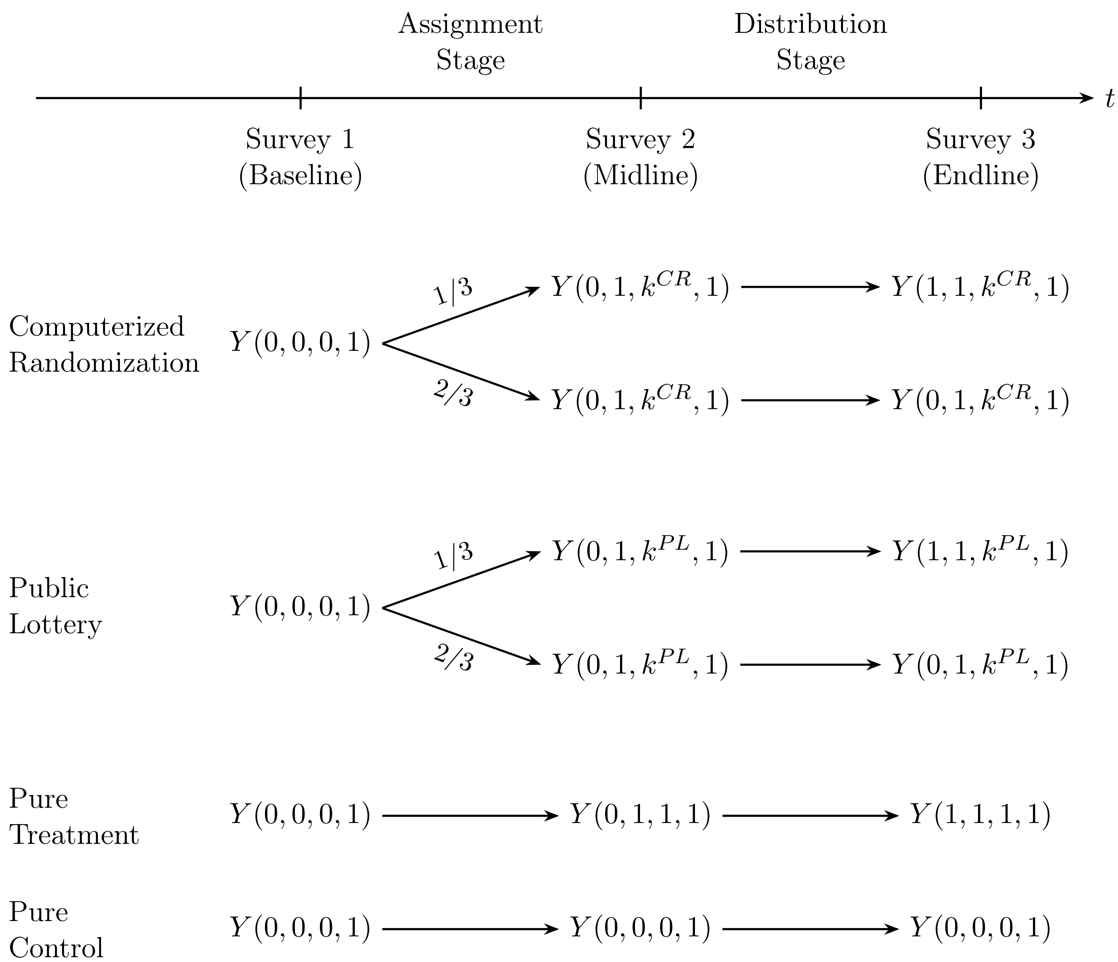


Figure 1: Schematic of the experimental design and outcomes over three surveys. Outcomes are $Y(t_i, a_i, k_i, s_i)$, where a_i denotes if subject i is aware of the existence of different assignment groups and her own status within them; k_i indicates subject i's awareness of the assignment process; and s_i denotes whether subject i is aware that she is part of a research study. The first two arms listed, Computerized Randomization and Public Lottery, are the "mixed" arms that allow for testing the variation in t_i and k_i simultaneously while keeping the same proportion of treated subjects invariant; in those two arms, one-third of the participants are assigned to treatment. The second set of arms, pure treatment and pure control, serves as a benchmark where assignment was done at the community level. The assignment stage refers to the meeting at which farmers were informed about their treatment status in the first three arms, and the distribution stage refers to the day the results of the soil quality tests and the fertilizer subsidies were distributed. By design, $1 \geq k^{\text{PL}} > k^{\text{CR}} > 0$.

The "mixed" arms allow me to vary assignment methods with intrinsically different awareness k. For ethical reasons, it is required to explain the selection method to subjects during field experiments, so it is not possible to keep $k = 0$ in the mixed arms. Then, I selected to test variation in the assignment method using the two most common assignment methods in RCTs: public lottery and computer-generated random selection.[11] The difference between these arms is based on the

[11] Of 5,060 RCTs analyzed from the AEA Registry by May 2025, 81% use computerized randomization and 6% use

explanation given to the subjects: while in the Public Lottery arm (PL) the randomization is done visually, removing balls from a bag to determine who is treated and who is control, in the Computerized Randomization arm (CR) subjects listen to the explanation about randomization, but they can't see it. Hence, the assignment methods vary in their visual appeal and ease of understanding. For this reason, I will assume and later test that $k^{\text{PL}} > k^{\text{CR}}$. Within each community, between one-third and 40% of the subjects are randomly assigned to receive the treatment ($t_i = 1$), while the remaining participants are assigned to the control group ($t_i = 0$). The two different assignment methods directly allow me to test for Hypothesis 1a.

Finally, to disentangle t from a, I separate the assignment stage of the intervention, when subjects were informed about their group assignment (excluding the pure control, where there was no announcement), from the distribution stage of the RCT, when treated subjects received the treatment. I then measure subjects' outcomes and behaviors at three distinct moments in time: before the assignment stage ($a_i = 0$ and t_i is unknown), between the assignment stage and the distribution stage ($a_i = 1$ but treatment has not yet been distributed), and after the distribution stage (t_i is realized). This variation across time directly allows me to test for Hypotheses 1b and 2a.

While anticipation effects can play a role in the period between the assignment and distribution steps, the fact that I have included two differential assignment methods, which are equal except for their explanation to the subjects in each arm, helps me rule them out. If assignment effects do not exist, then there is no reason to expect differential anticipation effects between the Public Lottery and the Computerized Randomization arms. However, if I find a difference, it serves as further proof of concept regarding the assignment effects, which begin to play a role even before the treatment is distributed.

Note that because of ethical concerns, I do not turn off the awareness of being part of a research study (s_i) for most of the outcomes of interest. However, I can observe subjects' farming behaviors without study awareness using satellite imagery (for example, I can analyze changes in crops, planted area, or plowing dates). I plan to run this analysis to track changes during the season, which ends in November 2025.

5 Intervention and Implementation

The intervention to test my hypotheses consists of providing information on soil quality paired with a subsidy to purchase farming inputs. I chose this intervention because it replicates one of India's most widespread agricultural federal programs, the Soil Health Card scheme; knowing whether the effects are well estimated is of particular interest to the country's agricultural development. Additionally, the program is designed to target only a subset of farmers within a village. Finally, this intervention naturally allows me to survey the subjects over time, monitoring behavior, and divide the assignment stage from the distribution stage.

public lottery as the assignment mechanism. The remaining 13% can't be clearly classified.

5.1 Experimental design adapted to a soil quality test program

Given the persistent decline in soil nutrient levels across India over the past three decades, Indian authorities have implemented a suite of agricultural initiatives aimed at enhancing soil fertility and, by extension, agricultural productivity. One such intervention is the Soil Health Card (SHC) program, launched by the Government of India in 2015, which provides farmers with information on critical soil nutrients and crop-specific recommendations for fertilizer use. This initiative seeks to increase yields through more efficient and targeted use of inputs (Ramappa, Manjunatha, Yashashwini, Bangarappa, and Mattihalli, 2024) and has, to date, distributed 250 million SHCs since its inception (Press Information Bureau, 2025).

Specifically, the SHC program provides farmers with a soil quality test: a soil sample is collected from the farmer's plot and analyzed at an accredited laboratory to assess its nutrient profile and physical properties. Based on this analysis, farmers receive a comprehensive report containing tailored recommendations about the optimal types and quantities of fertilizers required to satisfy crop requirements while accounting for existing nutrient availability. The program is relatively well-publicized and is valued by farmers, with empirical evidence showing yield improvements of 3 to 8% and reductions in input costs of 7 to 9% (Abhishek, Deshmanya, Tevari, Lokesh, Ravi, and Suresh, 2020; Ramappa, Kannan, and Lavanya, 2015; Reddy, 2018).

Nonetheless, the program has not reached all farmers due to several barriers, including limited internet access required to access the program, inadequate knowledge of sample collection procedures, insufficient extension support, and the remoteness of testing facilities, all of which contribute to low program uptake (Kaur, Kaur, and Kumar, 2020; Ramappa, Kannan, and Lavanya, 2015). Moreover, there seems to be institutional discrimination against the low caste farmers, as the extension service targets successful or progressive farmers (Sumanth et al., 2020; Kaur et al., 2021).

As the SHC program is a salient and scalable intervention whose expansion is restricted by frictions in transportation, access to inputs, and knowledge dissemination, I replicated the program within the context of my study. Soil testing was conducted in partnership with the University of Agricultural Sciences, Raichur, which houses an official SHC laboratory. J-PAL enumerators were trained in best practices for soil sample collection and handling, and samples were transported under controlled conditions to the laboratory. Following the analyses, each treated farmer received a printed individualized soil quality report accompanied by crop-specific fertilizer recommendations and a subsidy to buy agricultural inputs from a local input dealer. Additionally, all farmers were provided with aggregated information regarding average soil quality within their respective villages in the Computerized Randomization, Public Lottery, and Pure Treatment villages.

The experimental design is readily adapted to the soil quality test intervention, as the government SHC program treats only a subset of farmers within a village and includes a one-month interval between the time farmers learn of their selection into treatment and receive the results. This setup provides an ideal framework for observing behavioral changes during the interim period between assignment to the treatment and control groups and the distribution of the treatment. Moreover, the soil test intervention also offers a unique opportunity to measure how farmers interpret and respond to the intervention. It allows me to determine whether farmers attribute the test results (which can be either positive or negative) to their own actions or to external factors, providing further insight into the behavioral impacts of the assignment process.

5.2 Implementation of the experimental design

The intervention took place in rural villages in the Raichur district in northeast Karnataka, India. Farmers in this region work on small plots, and their agricultural activity depends mainly on the monsoon rains. Knowing whether the effects of a soil quality information provision intervention are accurately measured is of particular interest in Raichur for two reasons. First, the government of India estimates that the district's soil is poor in most of the components necessary for producing crops in an economically viable way.[12] Second, Raichur is one of the districts with higher consumption of fertilizers in Karnataka, which could lead to faster depletion of secondary soil nutrients if misused (Ramappa, Kannan, and Lavanya, 2015; Katsir, Biswas, Urs, Lenka, Jha, and Arora, 2024).

As shown in Figure 2, I randomized clusters of villages into the four different arms, and I invited farmers within each village to take part in a series of surveys.[13] In all cases except for the Pure Control arm, participation in survey 1 made farmers eligible to receive a soil quality test and a subsidy for farming inputs after the second survey. In the assignment stage, eligible farmers from the Computerized Randomization, Public Lottery, and Pure Treatment arms were invited to gather in a caste-neutral area (in front of the government school or next to a tea stall). Appendix E details how the assignment stage was done and explained to the farmers.

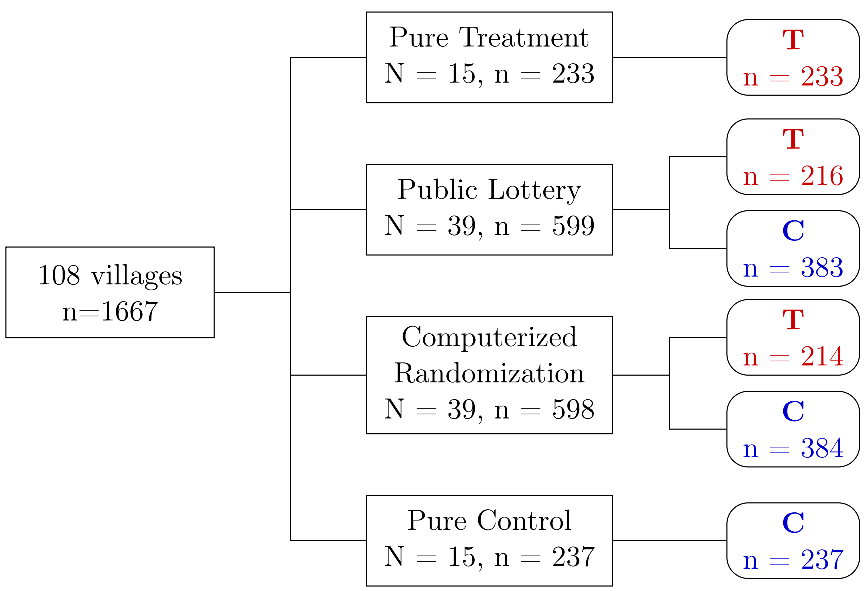


Figure 2: Experimental design flowchart. N indicates the number of villages randomized into each arm, and n indicates the number of farmers in each village based on participation in survey 1.

[12]Data retrieved from https://www.soilhealth.dac.gov.in/piechart, cycle 2023-2024. Moreover, Raichur is one of India's 112 "aspirational districts": the most underdeveloped areas of the country where the Government wants to focus all its efforts to improve the situation of its inhabitants.

[13]Refer to Appendix D to see the differences in consent scripts between the different arms.

Therefore, the trial employs a two-level randomized design: a cluster-level randomization, when villages are randomized into the four arms under a geographic separation constraint that requires villages assigned to different arms to be at least 5 km apart. This constraint minimizes spillovers and interference across arms. The second level is an individual-level randomization within villages. Within each village, eligible individuals are assigned by simple randomization in the Public Lottery and the Computerized Randomization arms. As the goal of the project is to compare alternative assignment mechanisms, maintaining a simple, clearly understandable "equal-chance" process in front of participants is essential for credibility and adherence. Stratified or blocked individual-level randomization (e.g., by age, gender, or other categories) requires complex procedures that are difficult to implement transparently during a Public Lottery, potentially undermining community understanding and trust. This would have added another level of variation to the design that is not intended to be tested, given the hypotheses.

In the Public Lottery arm, farmers were assigned for treatment through a visual lottery. Each farmer present had to select a ball from a closed bag. If the ball was green, they were assigned to the treatment group and accompanied to their plots to gather the soil sample at the end of the Public Lottery.[14]

In the Computerized Randomization arm, an enumerator publicly read the names of the selected farmers and explained that they had been chosen through a computer-generated lottery in the office. Finally, in the Pure Treatment arm, an enumerator publicly read the names of all the eligible farmers and informed them that all of them were receiving a soil quality test.

To maintain a constant number of visits among arms, in the Pure Control arm surveyors asked farmers to gather and provide administrative information about their plot. Enumerators then accompanied each farmer to his plot to get the plot coordinates. As expected, participation in this step was low.[15]

The timing of the surveys allowed me to follow participants across time: survey 1 was conducted in February and March 2025, the assignment phase occurred during April, and survey 2 was conducted between May to mid-June. Because of administrative delays, the laboratory could not share all the soil tests until June. This delay, together with the early start of the rainy season, which made it difficult to reach some villages, meant that the distribution of the soil test results and the fertilizer subsidies was delayed until the second half of June.[16] survey 3 was conducted during July 2025.

[14] Enumerators first contacted non-attending farmers by phone to confirm their continued interest and arrange for a family member to serve as a proxy. If a proxy could not be designated, an enumerator drew from the lottery on the farmer's behalf. This process was conducted in public, with all farmers present to ensure the transparency of the assignment mechanism. To guarantee data integrity, selected absentee farmers were revisited the following day; soil samples were then collected from the correct plot under the farmer's direct supervision.

[15] The only incentive to participate in this stage was to receive information on how to conduct a soil test properly, information that was distributed in all the arms for all eligible farmers.

[16] It is relevant to mention this delay as the Kharif season starts in early June. Therefore, farmers would have found the information more useful two weeks before it was actually distributed. However, the information was still useful as it indicated the amounts of fertilizer to be administered, and the subsidy was designed to cover around one month of fertilizer for a 2 acre plot.

6 Balance and Attrition

6.1 Balance between treatment and control groups, imbalance by arm

Just focusing on the difference between those assigned to treatment and those assigned to control, only the education level and the household income seem to be significantly different according to Table A3. However, what matters the most is that the different arms are balanced in their characteristics, even before sorting farmers into treatment or control. Table A5 shows these differences. In this case, and in particular in the comparison of the Public Lottery and the Computerized Randomization arms, there is imbalance in many covariates. Given that I do not use sample stratification in this study, the imbalance is expected. To account for the imbalance, I all specifications I use an ANCOVA design, including the lagged dependent variable. As a robustness check and to solve the imbalance, I also control for the level of education, household income, and ownership of the plot. Moreover, to limit the discretion in the choice of control variables, I perform post-double selection LASSO (Belloni, Chernozhukov, and Hansen, 2014) and report the main results in Appendix L.

6.2 Differential Attrition: More attrition in the Public Lottery arm

In this section, I analyze attrition from survey 1 to survey 2, and from survey 2 to survey 3. Given my hypotheses, attrition is an interesting outcome in itself, which could provide insights into what determines participation, depending on the awareness and assignment method.

The overall attrition rate for survey 2 is 21.85% but, as expected, it is not homogeneous between those assigned to the treatment group and those assigned to the control group: 25.42% of those assigned to the control group do not participate in survey 2, while only 16.44% of the treated group do not participate.[17] The attrition rate also varies by arm: 19.87% for the Computerized Randomization arm, 25.92% for the Public Lottery arm, 22.46% for the Pure Control, and 15.88% for the Pure Treatment arm.

The picture is different for survey 3, where attrition is higher overall (26.95%). This was expected as, by design, there was almost no incentive to continue participating in the intervention after the soil results and the fertilizer subsidies were distributed. Once again, attrition is higher for those assigned to the control group compared to those assigned to the treatment group (29.31% vs 23.38%).[18] There is almost no variation in attrition rate between arms for survey 3: 26.21% for the Computerized Randomization arm, 27.26% for the Public Lottery arm, 27.54% for the Pure Control, and 27.47% for the Pure Treatment arm.

Based on these attrition rates, the Public Lottery arm appears to create less engagement among participants, with an attrition rate 6 percentage points higher than the Computerized Randomization arm during survey 2. Table 1 confirms that the Public Lottery arm has a higher attrition rate and is statistically different from the Computerized Randomization arm during survey 2, but this

[17]The difference is starker if I focus on those whose soil sample was not taken in the assignment stage and those who had their soil sample taken: 27.16% and 12.18%, respectively.

[18]The difference is higher if I focus on those whose soil sample was not taken in the assignment stage and those who had their soil sample taken: 30.6% and 20.30%, respectively.

is not the case for survey 3. Adding whether the farmer was assigned to treatment in both surveys significantly lowers attrition, but it doesn't have a statistically significant effect when interacting with the Public Lottery arm.

	(1) Left for survey 2	(2) Left for survey 2	(3) Left for survey 3	(4) Left for survey 3
Pure Control	0.0259	0.0012	0.0133	-0.0233
	(0.0412)	(0.0392)	(0.0543)	(0.0466)
Pure Treatment	-0.0399	0.00459	0.0126	0.0784*
	(0.0356)	(0.0407)	(0.0579)	(0.0458)
Public Lottery	0.0605**	0.0803**	0.0105	-0.000272
	(0.0276)	(0.0379)	(0.0304)	(0.0354)
Treated	-	-0.0692**	-	-0.102***
		(0.0275)		(0.0249)
Treated × Public Lottery	-	-0.0539	-	0.0309
		(0.04)		(0.0493)
Constant	0.199***	0.223***	0.262***	0.299***
	(0.0177)	(0.0239)	(0.0196)	(0.0212)
Observations	1,666	1,666	1,666	1,666
p-value: PL = CR	0.0284	0.0343	0.731	0.994

Table 1: The dependent variable is a dummy indicating farmers who did not participate in one of the surveys. The set of dummies, Pure Control, Pure Treatment, and Public Lottery (PL), indicates the arm each farmer belongs to. The constant represents the Computerized Randomization arm (CR), and Treated represents the assigned treatment. At the bottom of the table, I report the p-value of the null hypothesis of equal attrition in the Public Lottery and Computerized Randomization arm.

Tables A6 and A7 show balance tables between farmers who dropped out of the intervention in each survey round and those who did not. In both surveys, younger farmers, those in households that participate less in community roles, those whose land is not registered, and those who don't own livestock, are more likely to abandon the intervention. Farmers who request an agricultural loan and own the selected plot are more likely to remain in the intervention. All this evidence suggests that farmers who are more attached to their land or community are more likely to continue participating in the intervention.

7 Evaluation and Outcomes

In this section, I establish the econometric specifications to evaluate the hypotheses stated in Section 3. Then, I describe the indices used as outcomes: I use both indices defined in advance and indices defined through clustering of vector embeddings.

7.1 Evaluation

First, to test Hypotheses 1a and 1b, I focus on the difference between the two mixed arms to maintain a constant proportion of participants assigned to treatment and avoid differential spillover concerns. I estimate the different intent-to-treat (ITT) effects by arm: I compare farmers *assigned* to the control group to those assigned to the treatment group, irrespective of whether they received the soil quality test or not, and interact the assignment to treatment with an arm dummy. To control for pre-existing group differences or other confounding variables, I use an ANCOVA design:

$$Y_{ir} = \alpha_r + \gamma_r t_i + \delta_r \text{PublicLottery}_i + \beta_r \left(t_i \times \text{PublicLottery}_i\right) + \theta Y_{i1} + \varepsilon_{ir} \tag{5}$$

Then, α_r is the expected outcome value for those assigned to control in the Computerized Randomization arm in survey round r; $\alpha_r + \delta_r$ is the expected outcome value for the control subjects in the Public Lottery arm; $\alpha_r + \gamma_r$ is the expected outcome value for the treated subjects in the Computerized Randomization Arm; and $\alpha_r + \gamma_r + \delta_r + \beta_r$ is the expected outcome for the treated subjects in the Public Lottery arm. The ITT effect on the Computerized Randomization arm is γ_r, while the ITT effect on the Public Lottery arm is $\gamma_r + \beta_r$.

To test Hypothesis 1a, I test the null hypothesis $\beta_3 = 0$, which is the difference in the treatment effects by arm after the distribution of the treatment. If the estimated coefficient is significantly different from zero, it would give evidence in favor of Hypothesis 1a. Moreover, it is the empirical test of Proposition 3.

Focusing on survey 2, the interpretation changes: β_2 is the differential anticipation effects between the Computerized Randomization arm and the Public Lottery arm. If I reject $\beta_2 = 0$, it is evidence in favor of Hypothesis 1b.

Specification 5 also allows me to test Hypothesis 2a: the assignment effect on the control is the difference between the control subjects' expected levels in the Public Lottery arm and the control subjects' expected levels in the Computerized Randomization arm, δ_r. The assignment effect on the treated is captured by the difference between the treated subjects' expected levels in the Public Lottery arm and the treated subjects' expected levels in the Computerized Randomization arm, $\delta_r + \beta_r$. Then, I can test the null hypotheses $\delta_r = 0$ and $\delta_r + \beta_r = 0$.

I perform this exercise and tests for survey rounds $r \in \{2, 3\}$, controlling for the baseline value of the outcome variable Y_{i1}. Although in my main specification I don't control for other covariates, in robustness checks I include a vector of baseline covariates X_i to improve estimation precision: in one specification I control for the farmer's level of education, household income, and ownership status of the relevant plot, while in another I use the Post-Double Selection LASSO (PDL) developed by Belloni, Chernozhukov, and Hansen (2014).

In addition to ITT effects, I also report the local average treatment effect (LATE) as there are non-compliers in my sample. I estimate the effect using the assignment into the treatment as an instrumental variable for receiving the soil test and the fertilizer subsidy, so Z is the excluded instrument for an indicator D of compliance in a two-stage least squares estimation of:

$$Y_{ir} = \alpha + \gamma_r D_i + \delta_r \text{PublicLottery}_i + \beta_r \left(D_i \times \text{PublicLottery}_i\right) + \theta Y_{i1} + \varepsilon_{ir} \tag{6}$$

To tackle Hypothesis 2a more precisely, where I aim to test if changes in the awareness state affect differently how the treated group and the control group behave, I use the four treatment arms. First, focusing on subjects assigned to treatment, I use the Pure Treatment arm as the comparison group and analyze the subjects' behavior in survey 2 and survey 3 in the Computerized Randomization and the Public Lottery arms:

$$Y_{ir} = \alpha + \beta_r^{CR,T} \text{ComputerizedRandomization} + \beta_r^{PL,T} \text{PublicLottery} + \theta Y_{i1} + \varepsilon_{ir} \tag{7}$$

Coefficients $\beta_r^{CR,T}$ and $\beta_r^{PL,T}$ indicate the difference in behavior between the subjects assigned to treatment in the Computerized Randomization arm with respect to subjects in the Pure Treatment arm, and the the difference in behavior between the subjects assigned to treatment in the Public Lottery arm with respect to subjects in the Pure Treatment arm. These coefficients provide insight into how the two different assignment arms behave compared to a context where everyone receives treatment (i.e., with an awareness state of $\boldsymbol{\alpha}^{PT} = (1, 0, 1)$. I also test the null $\beta_r^{CR,T} = \beta_r^{PL,T}$ to evaluate if the difference between the mixed arms is statistically different. As before, in my main specification, I don't control for other covariates, but I include a vector of baseline covariates X_i to improve estimation precision in robustness checks.

For the control group, the specification is the same as in Equation 7, but with the Pure Control arm as the benchmark:

$$Y_{ir} = \alpha + \beta_r^{CR,C} \text{ComputerizedRandomization} + \beta_r^{PL,C} \text{PublicLottery} + \theta Y_{i1} + \varepsilon_{ir} \tag{8}$$

Similarly, coefficients $\beta_r^{CR,C}$ and $\beta_r^{PL,C}$ indicate the difference in behavior between the subjects assigned to control in the Computerized Randomization arm with respect to subjects in the Pure Control arm, and the the difference in behavior between the subjects assigned to control in the Public Lottery arm with respect to subjects in the Pure Control arm. These coefficients provide a clue on how the two different assignment arms are behaving compared to a context where nobody receives treatment (so with an awareness state of $\boldsymbol{\alpha}^{PT} = (0, 0, 1)$. I also test the null $\beta_r^{CR,C} = \beta_r^{PL,C}$ to evaluate if the difference between the mixed arms is statistically significant.

Finally, I conduct the diagnostic tests for the separability assumption stated in Propositions 1 and 2. I empirically calculate the constructs T_1, T_2, T_3, and T_3^*:

$$\begin{aligned} T_1 &= \mathbb{E}[Y(1,1,0,1) - Y(0,1,0,1)] \\ T_2 &= \mathbb{E}[Y(1,1,1,1) - Y(0,1,1,1)] \\ T_3 &= \mathbb{E}[Y(1,1,0,1) - Y(0,1,1,1)] \\ T_3^* &= \mathbb{E}[Y(1,1,1,1) - Y(0,1,0,1)] \end{aligned}$$

As shown in Figure 1, $Y(1,1,0,1)$ is the outcome for subjects assigned to treatment in the Computerized Randomization arm in survey 3. As for $Y(0,1,0,1)$, it refers to the outcomes of the control group in the same arm also during survey 3. So T_1 represents the average difference between the treated and control groups during survey 3 in the Computerized Randomization arm.

Similarly, T_2 represents the average difference between treated and control subjects during survey 3 in the Public Lottery arm. The difference T_3 does not have an intuitive interpretation: it is the average difference between the treatment group in the Computerized Randomization arm and the control group in the Public Lottery during survey 3. The same is true for the difference T_3^*: it is the average difference between the treatment group in the Public Lottery arm and the control group in the Computerized Randomization arm during survey 3.

To test whether inequalities 2 and 4 hold, I define the test statistics $S = T_1 + T_2 - T_3$ and $S^* = T_1 + T_2 - T_3^*$. Formally, the null and alternative hypotheses to test inequality 2 are:

$$H_0 : S \geq 0 \qquad H_1 : S < 0 \tag{9}$$

Likewise, to test whether inequality 4 holds, the null and alternative hypotheses are:

$$H_0 : S^* \geq 0 \qquad H_1 : S^* < 0 \tag{10}$$

To evaluate these hypotheses, I employ a nonparametric bootstrap procedure. Specifically, I repeatedly resample the observed data with replacement within each relevant group to generate bootstrap samples, recalculating T_1, T_2, T_3, T_3^*, S, and S^* for each iteration from the relevant groups. I then approximate the sampling distribution of S under the observed data. The one-sided p-value is given by the proportion of bootstrap replicates in which $S \geq 0$. I also report a bootstrap confidence interval for S, with support for H_1 indicated if the upper bound of this interval is not less than zero. The same procedure is used to approximate S^* and calculate the p-value for $S^* \geq 0$.

Rejection of the null hypotheses $H_0 : S \geq 0$ or $H_0 : S^* \geq 0$ indicates evidence against the additive separability of treatment exposure and the awareness state. In other words, if the observed data lead to $S < 0$ or $S^* < 0$ with statistical significance, this suggests that the causal effect of treatment cannot be decomposed into independent contributions from the treatment assignment and the awareness state.

Finally, in the case of non-separability (rejection of the null hypothesis), S provides an upper bound on how much of the difference between the extreme states of treatment and full-awareness versus control and no-awareness is due purely to the treatment and awareness interaction. Similarly, S^* provides an upper bound on how much of the difference between the opposite extremes: treatment and no-awareness versus control and full-awareness.

As a robustness check, I perform the same calculation using Bayesian bootstrap (Rubin, 1981). This method has two main advantages: first, it delivers estimates that are smoother than the simple bootstrap due to its continuous weighting scheme. Second, the continuous weighting scheme prevents corner cases from arising, as no observation will ever receive zero weight; thus, no collinearity problem will arise.

7.2 Outcomes

In this section, I outline the various outcomes I analyze in both the results and mechanisms sections. To draw general conclusions about the experiment's results, I use summary indices that aggregate information for variables consistently measured across the three survey rounds. Then, I use natural language processing to categorize the questions in a data-driven manner and discuss why this approach proves to be more imprecise. Finally, I present the text analysis of open-ended questions.

7.2.1 Summary indices

Due to the comprehensive nature of the study, I collect a large number of outcome variables listed in Appendix F. Therefore, I expect some of the variables to show significant results due to chance. Following Kling, Liebman, and Katz (2007) and Banerjee, Duflo, Goldberg, Karlan, Osei, Parienté, Shapiro, Thuysbaert, and Udry (2015), I report an index for each group of related survey variables. I define each summary index Y_{iA}^{c} (category c, for subject i within arm A) as the equally weighted average of z-scores of its components, with the sign of each measure oriented so that higher values correspond to better outcomes, or a more proactive behavior. The construction of these composite indices yields greater statistical power to detect effects that go in the same direction within the same conceptual category. Table 2 describes the categories, their interpretations, and the survey variables included in each category.

The first three indices listed in Table 2 correspond to survey questions directly influenced by reports on the state of farmers in India (Niti, 2014; Shekara et al., 2016; Ramappa et al., 2024). I consider these three indices the outcomes of interest for a soil quality information provision intervention. The other four indices are designed to measure the mechanisms behind potential behavioral changes. The Attribution & Locus of Control index is based on Laajaj and Macours (2021), where I replicate the use of visual aids to measure how much farmers believe their actions determine their outcomes, either good or bad. Based on the context, I also added religiosity measures as a second measurement of outcomes determined by an external force. The rationale for adding such measures is based on the literature on attribution bias, which refers to the idea that temporary sensations or situational factors are incorrectly attributed to an underlying, stable characteristic of a person or good (Erkal et al., 2022).

For the Psychological Well-being index, I used similar questions to measure life satisfaction and well-being as Devoto et al. (2012), and included a measure of expected revenue for the season that is based mainly on optimism, as the season is starting by the endline survey. These measures intend to explore the appearance of resentful demoralization or emotional distress as a result of being excluded from the intervention (Cook, Campbell, and Day, 1979).

The Risk Aversion index aims to capture changes in risk behavior born in overconfidence, therefore looking into motivated beliefs as a potential mechanism (Möbius et al., 2022). The questions inside this index include a standard time preferences elicitation task, a variation of the certainty equivalent method, and variations of the Becker-DeGroot-Marschak method.

Finally, for the Social Network index, the questions are similar to Chandrasekhar et al. (2018) but mostly based on interactions around agriculture. The aim of this index is to understand if the assignment method changes interactions between the agents, potentially altering the network and therefore having impacts on information flows.

To standardize each outcome into a z-score to run Equations 5 and 6, I need to compare against the behavior of a common group: the pooled sample of all control group participants from both the Public Lottery and Computerized Randomization arms at baseline (survey 1). Therefore, I subtract the pooled control group mean at survey 1 and divide by the pooled control group standard deviation at survey 1. If an individual has a valid response to at least one survey measure of an index, then any missing values for other survey measures are imputed at the group mean. Then, I average all the z-scores and standardize them to the pooled control group at survey 1. This is a deviation from the most common method, which would involve subtracting the control group mean at the corresponding survey round and dividing by the control group standard deviation at the corresponding survey round, then averaging all the z-scores, and again standardizing to the control group within each round (Banerjee, Duflo, Goldberg, Karlan, Osei, Parienté, Shapiro, Thuysbaert, and Udry, 2015; Dhar, Jain, and Jayachandran, 2022).[19]

My methodological choice ensures that all subsequent measurements are referenced to a fixed, pre-intervention benchmark that reflects the status quo before any possible assignment bias or spillover appears. This approach is critical given my central hypothesis: that participants' awareness state—including awareness of treatment assignment and study participation—may itself influence outcomes, even for those not receiving the direct intervention. If I were to standardize using the control group in later survey rounds, these benchmarks could be confounded by assignment effects that shift behavior within the entire control group, resulting in post-assignment and post-treatment bias. Using the baseline control group preserves the comparability and interpretability of index values as deviations from the initial state (survey 1), with identical awareness states for all arms as shown in Figure 1. Appendix J provides a detailed example of the risk of incorrect standardization.

In practical terms, this anchors all subsequent analyses to a stable reference, allowing treatment and assignment effects to be interpreted as departures from baseline behavior. It also avoids masking awareness-state-induced changes by normalizing them away. This approach thereby provides a transparent and consistent measure of change, directly aligned with the conceptual framework of context-dependent causal inference.

To tackle Hypothesis 2a through Equations 7 and 8, standardizing against each arm's control at baseline would be incorrect. Using three different benchmarks to define the dependent variable introduces unnecessary noise and complicates the interpretation of the coefficients. Then, I construct the outcome indices by standardizing each component variable using the mean and standard

[19] My approach is similar to Dhar, Jain, and Jayachandran (2022) analysis of the gender attitudes index at endline 2: they standardize this index's variables using the mean and the standard deviation of the variables in endline 1 (see page 77 of their online appendix for further detail on their methodology).

deviation calculated from the pooled sample of all control group participants at baseline (survey 1). This approach leverages the fact that all control groups were, in expectation, identical before the assignment, providing a more precise and stable benchmark. It ensures that all estimated coefficients are interpreted in the same units: deviations from the pre-assignment control mean in pre-assignment control standard deviations.

Category (c)	Interpretation	Variables included
Knowledge of New Information	A higher value reflects stronger information-seeking behavior	soil_test_value[a], wtp_max, farming_knowledge
Changes in On-Farm Practices & Inputs	A higher absolute value indicates recent changes in on-farm practices	seed_extension, seed_board, fert_organic, fert_chemical
Change in Financial Management	A higher absolute value indicates recent changes in financial management practices	agri_loan_past, agri_loan_internal, agri_loan_external, purposes_past_loan, agri_loan_future, purposes_future_loan, pur_investment_future_loan, past_agriloan_invest, insurance
Attribution & Locus of Control	A higher value reflects a stronger belief that outcomes are determined by personal effort rather than external factors	generic_poor_farmer, generic_good_farmer, own_poor_farmer, own_good_farmer, hard_work_tokens, yourself_tokens, religiousity_worship, religiousity_pray
Psychological Well-being	A higher value reflects a greater degree of optimism	avg_confidence, life_satisfaction, feeling_cheerful, feeling_worried, feeling_tired, expected_revenuekharif
Risk Aversion	A higher value reflects more risk aversion	beh_measurement_eg, nitro_bet,[b] weather_bet, time_discount, risk_lottery
Social Networks	Indicates a greater number social connections	links_advice, advice_farming, best_farmer1_talk, best_farmer2_talk, best_farmer3_talk

[a] Because of a coding mistake, soil_test_value was not asked in survey 3 for farmers who got a soil test. I removed the variable from the index, but intend to add it back when analyzing the second endline.

[b] The variable nitro_bet is not available for the pure control arm, so it is removed from the index when pure control is used as a benchmark.

Table 2: This table describes the seven conceptual categories used in the analysis, each of which is composed of several survey questions. The variable nitro_bet was not asked for in the Pure Control arm to avoid informing farmers of the intervention taking place in other villages. Each outcome variable is listed and explained in Appendix F.

7.2.2 Data-driven indices

In addition to the previously described summary indices, I implement a data-driven approach to construct outcome groupings based on the semantic content of the survey questions themselves. To achieve this, I use a natural language processing model (OpenAI Text Embedding 3) to generate dense vector embeddings for each outcome description. The embeddings capture subtle aspects of the meaning and context of each question, positioning them in a high-dimensional semantic space.

To better understand the patterns in the semantic relationships between outcomes, I use two techniques. First, I use Principal Component Analysis (PCA) to reduce the complexity of the high-dimensional embedding space, allowing me to visualize each outcome as a point in two dimensions –the first two principal components. These two dimensions are constructed to capture the greatest possible variation in how outcome descriptions differ from one another. While the principal components do not have direct, interpretable meanings (such as “risk” or “well-being”), they summarize the most significant directions in how the survey questions are semantically distributed. As a result, outcomes that are close together on Figure 3 have descriptions that are similar in content and context.

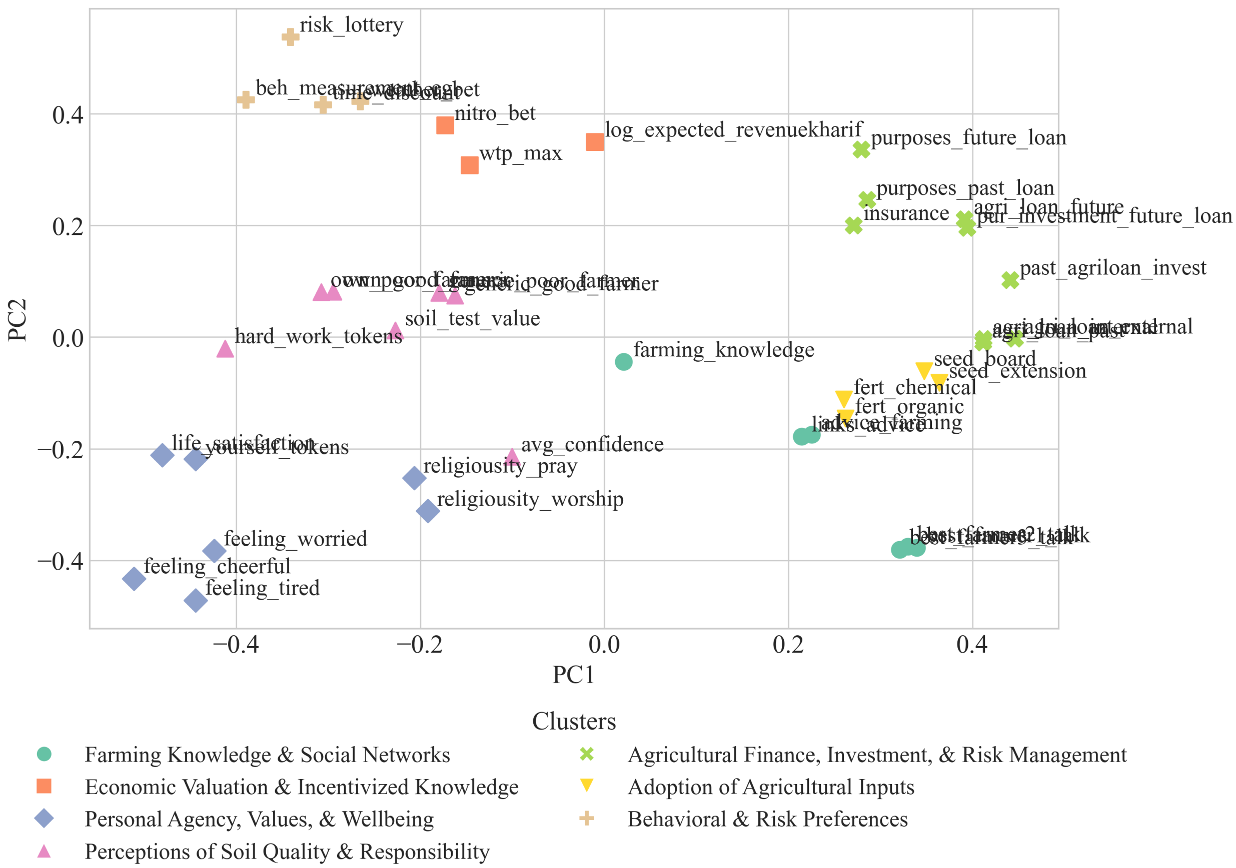


Figure 3: This figure illustrates the results of k-means clustering applied to the variable embeddings, which, for visualization purposes, are compressed to a two-dimensional space using Principal Component Analysis (PCA). Each point represents a different variable, plotted according to its PCA coordinates (PC1 and PC2). The variables are categorized into distinct clusters, as indicated by the various shapes and colors in the legend. The arrangement of points in the plot reveals how these variables group together based on their similarities, highlighting underlying patterns present in the data set.

Next, I apply k-means clustering to group together outcomes that are conceptually related, identifying sets of outcome variables whose textual descriptions are naturally similar. Intuitively, these steps let the similarities in the wording and content of the survey questions guide the formation of groups, revealing clusters of outcomes that may share thematic or conceptual connections—even if they were not grouped in my original design. I set the number of clusters to seven to maintain consistency in the number of groupings and to verify if the data-driven method yields similar groupings as those defined in the previous section.[20] The resulting clusters reflect distinct domains listed in Table 3.

Category (c)	**Interpretation**	**Variables included**
Adoption of Agricultural Inputs	Adoption or use of seeds and fertilizers—measuring input uptake for agricultural productivity.	seed_extension, seed_board, fert_organic, fert_chemical
Personal Agency, Values, & Wellbeing	Combines self-perception (internal locus of control), religious activity, and various facets of life satisfaction and subjective wellbeing.	yourself_tokens, religiousity_worship, religiousity_pray, life_satisfaction, feeling_cheerful, feeling_worried, feeling_tired
Behavioral & Risk Preferences	Includes classic experimental economics measures of risk, ambiguity, time preferences (discounting), and incentivized decisions.	beh_measurement_eg, weather_bet, time_discount, risk_lottery
Agricultural Finance, Investment, & Risk Management	All variables relate to agricultural loans (past/future/sources/uses), investment, and crop insurance—covering the financial domain and risk management in farming.	agri_loan_past, agri_loan_future, purposes_future_loan, pur_investment_future_loan, insurance, past_agriloan_invest, purposes_past_loan, agri_loan_internal, agri_loan_external
Farming Knowledge & Social Networks	This cluster combines objective farming knowledge with social connectedness: giving/receiving advice and interactions with knowledgeable farmers.	farming_knowledge, links_advice, advice_farming, best_farmer1_talk, best_farmer2_talk, best_farmer3_talk
Economic Valuation & Incentivized Knowledge	Includes willingness to pay, betting on nitrogen content (knowledge under incentive), and expected plot revenue—measuring economic judgments or incentives.	wtp_max, nitro_bet, expected_revenuekharif
Perceptions of Soil Quality & Responsibility	Soil test usefulness, responsibility/blame for soil outcomes (generic and own), belief in hard work, and confidence—focusing on attitudes regarding agency and soil management.	generic_poor_farmer, generic_good_farmer, own_poor_farmer, own_good_farmer, hard_work_tokens, avg_confidence, soil_test_value

Table 3: This table describes the seven clusters defined by the data-driven method.

[20] Using the elbow method to define the optimal number of clusters defines five clusters. However, the groupings so formed have poorer conceptual clarity than the data-driven clusters with seven groupings.

To construct quantitative indices for each semantic cluster, I use the same standardization approach as in the construction of the summary indices. The data-driven method provides an alternative lens for interpreting the impacts of the interventions, complementing my constructed categories and ensuring that conclusions are robust to different ways of organizing the outcomes.

Although two of the indices are identical ("Agricultural Finance, Investment, & Risk Management" has the same components as "Change in Financial Management"; and "Adoption of Agricultural Inputs" has the same components as "Changes in On-Farm Practices & Inputs") and others are very similar ("Farming Knowledge & Social Networks" closely resembles "Social Networks", adding the farming_knowledge variable; "Behavioral & Risk Preferences" is the same as the "Risk Aversion" index except for the absence of the nitro_bet variable), my preferred categorization is the summary indices listed in Table 2. These indices offer greater conceptual clarity relative to the data-driven categorization described in Table 3.

The data-driven categorization combines conceptually distinct constructs in a way that risks contaminating the constructs. For example, nitro_bet is an indicator of risk preference, and yet it is placed together with the expected revenue for Kharif and the willingness to pay for a soil test. As another example, the variable farming_knowledge is combined with social networks measurements in Table 3, when it was designed to understand if farmers were incorporating concepts from the soil quality tests.

However, some of the data-driven groupings have a sense I did not foresee. For example, the religiosity measurements may be related to the farmers' well-being, rather than with the attribution measurements as I specify in Table 2. For that reason, I also report the results for the data-driven indices in Appendix Appendix I.

8 Results

In this section, I test each hypothesis using the panel of subjects who participated in all three survey rounds. First, I demonstrate that average treatment effects vary across assignment mechanisms, suggesting that awareness influences the impact of the intervention. Next, I examine whether such differences emerge before treatment delivery and find no evidence of systematic anticipatory effects. I then demonstrate that the assignment mechanism operates asymmetrically across treated and control subjects, generating substantial differences in outcomes among the control group but not among the treated. Finally, I assess separability more formally through a diagnostic test, which evaluates whether treatment effects are independent of the assignment mechanism. The test does not reject separability for any outcome, suggesting that while awareness affects control outcomes, treatment effects themselves remain largely stable across assignment methods.

8.1 Different distribution of treatment effects by awareness state

To examine whether the average treatment effects differ by awareness state, I restrict the sample to survey 3 and test equality of the intent-to-treat (ITT) effects between the Computerized Randomization and the Public Lottery arms. In Equation 5, this corresponds to the null $H_0 : \gamma_3 = \gamma_3 + \beta_3$, which simplifies to $H_0 : \beta_3 = 0$. I test this hypothesis separately for each index outcome and also perform a joint Wald test to determine if (i) the first three indices directly related to the agricultural intervention are jointly equal to zero, and (ii) all seven interaction coefficients are jointly equal to zero, combining the models to account for correlation in the outcome residuals.

Figure 4 displays the estimated coefficients $\widehat{\gamma_3}$ as the ITT for the Computerized Randomization arm, and $\widehat{\gamma_3 + \beta_3}$ as the ITT for the Public Lottery arm, both with 95 percent confidence intervals. It also reports the p-value from the two-sided equality test, which is also presented in Table 4.

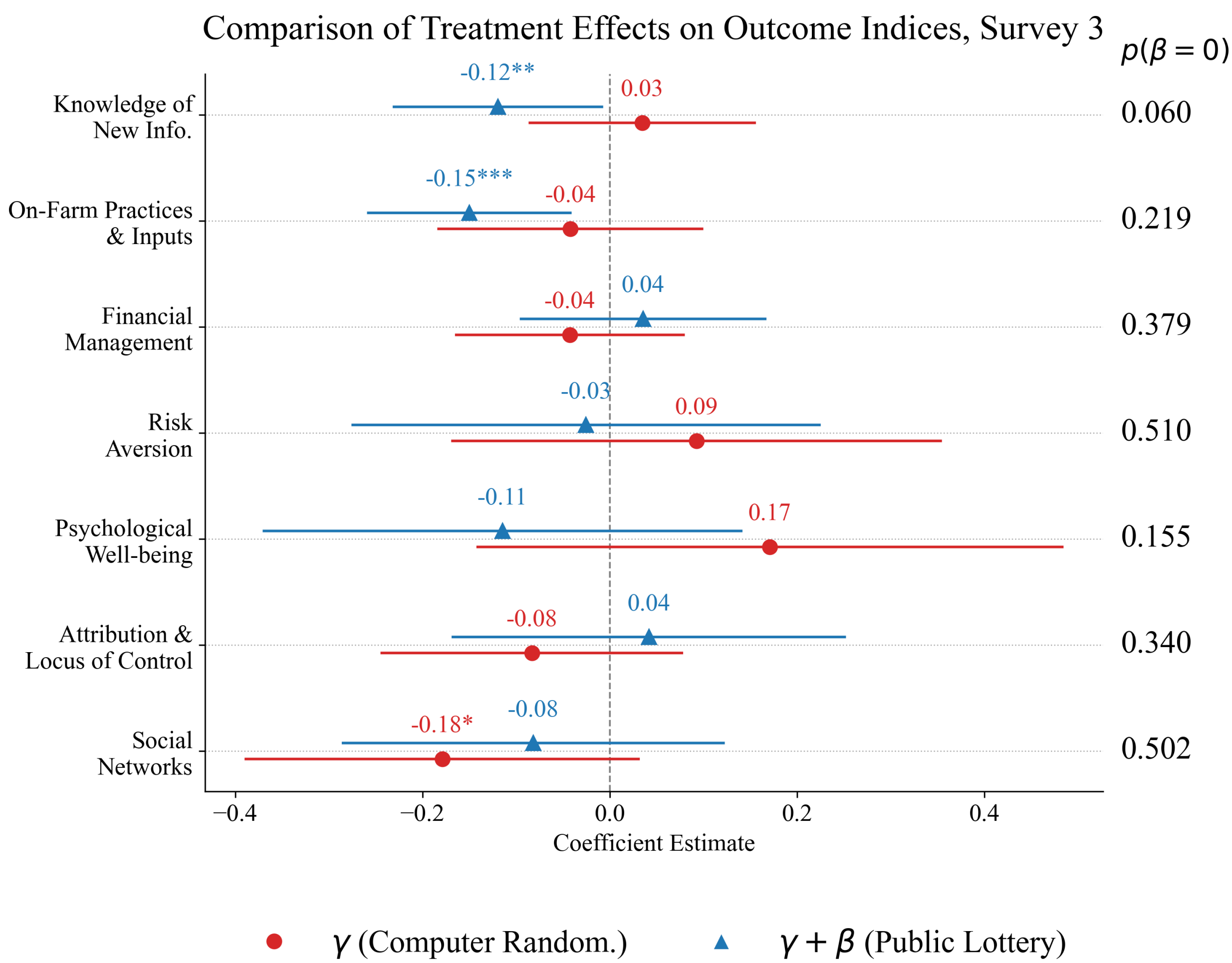


Figure 4: The figure presents coefficient estimates with 95% confidence intervals for the ITT in the Computerized Randomization arm (γ) and for the ITT in the Public Lottery arm ($\gamma + \beta$). The p-values shown above each point represent significance levels for the respective effects. The rightmost values show p-values for the test of $\beta = 0$, indicating whether the Public Lottery effect is statistically distinguishable from the computer randomization effect. The standard errors are clustered at the geographic cluster level, and the estimation was done by adding the baseline level of the corresponding outcome index as stated in Equation 5. All coefficients are expressed in standard deviation (SD) units, standardized relative to the pooled control group. Figure A3 shows the same coefficients for the data-driven indices.

The first striking fact is that the Public Lottery arm shows a negative and significant effect on the index related to the knowledge of new information: 0.12 standard deviations below the pooled control at baseline. It also yields statistically significant results in the changes of on-farm practices and inputs. This means that two weeks after receiving treatment, the effect of the intervention seems to be negative in the adoption of the soil recommendations in the Public Lottery arm. In contrast, there is no treatment effect in the Computerized Randomization arm on these two crucial measures, which are statistically zero. The second striking fact is that for the financial management, risk aversion, psychological well-being, and attribution and locus of control indices, the arms are moving in opposite directions. For the risk aversion and psychological well-being indices in particular, the intervention appears to be having a positive (although not statistically significant) effect in the Computerized Randomization arm. The third interesting result is the adverse effect of the intervention on social networks, particularly for the Computerized Randomization arm.

	Info Adopt.	**Practices Inputs**	**Financial Manag.**	**Risk Aversion**	**Psych. Well-b**	**Attrib. LOC**	**Social Networks**
Treat (γ)	0.035	-0.042	-0.043	0.092	0.171	-0.083	-0.179^{*}
	(0.059)	(0.069)	(0.059)	(0.127)	(0.153)	(0.078)	(0.103)
PL (δ)	0.135^{*}	0.015	0.022	0.171^{**}	0.345^{**}	-0.008	0.098
	(0.069)	(0.099)	(0.071)	(0.079)	(0.158)	(0.073)	(0.073)
Treat $\times$ PL (β)	-0.154^{*}	-0.108	0.078	-0.118	-0.286	0.125	0.097
	(0.079)	(0.086)	(0.088)	(0.177)	(0.196)	(0.129)	(0.143)
Constant (α)	-1.327^{***}	-2.360^{***}	0.866^{***}	-0.244^{***}	-0.358^{**}	0.053	0.243^{***}
	(0.039)	(0.077)	(0.052)	(0.061)	(0.142)	(0.053)	(0.036)
Observations	701	701	701	701	701	701	701

Table 4: Survey 3 regression results as indicated in Equation 5 for summary indices. The baseline corresponds to the Computerized Randomization. PL stands for a dummy variable for the Public Lottery arm, and Treat is a dummy variable indicating whether the subject was assigned to treatment. The coefficients for Treat $\times$ PL allow to test for Hypothesis 1a directly. I control for the lagged variable. The bootstrapped standard errors are clustered at the geographical cluster level. $^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.1$.

Focusing only on the knowledge of new Information, On-Farm Practices & Inputs, and Financial Management indices, the joint Wald test reveals that, when accounting for covariance across these three outcomes, the overall pattern of coefficients departs significantly from equality (p-value = 0.0058). Because the small number of clusters can lead to over-rejection (Cameron, Gelbach, and Miller, 2008), I re-estimate the joint Wald test using two small-sample adjustments. In both cases, the results are robust: A wild cluster bootstrap with 500 replications yields a p-value of 0.022, and a leave-one-cluster-out analysis yields a maximum p-value of 0.036.

For interpretation purposes, the share $\hat{\beta}/(\hat{\gamma}+\hat{\beta})$ can give an approximation of the proportion of the ITT effect found in the Public Lottery arm attributable to using a public lottery rather than a computerized assignment. For the knowledge of new Information, this share is 1.28, meaning that the assignment mechanism's contribution (the β interaction) is 128% of the total Public Lottery treatment effect in magnitude. Intuitively, the public lottery more than fully offsets the positive effect seen under computerized assignment (0.03 SD), pushing the net PL effect to negative (0.12 SD).

When analyzing the seven indices together, while most individual coefficients are statistically indistinguishable across arms, the joint Wald test reveals once again that the overall pattern of coefficients departs significantly from equality (p-value < 0.01). The wild cluster bootstrap yields a p-value of 0.212, but the leave-one-cluster-out analysis also supports that the seven outcomes are simultaneously different from zero (maximum p-value $= 0.047$). Together, the results indicate significant heterogeneity in treatment effects across awareness states, particularly in terms of information adoption and on-farm practices.

	Info Adopt.	**Practices Inputs**	**Financial Manag.**	**Risk Aversion**	**Psych. Well-b**	**Attrib. LOC**	**Social Networks**
D (γ)	0.036	-0.044	-0.045	0.097	0.179	-0.087	-0.188^{*}
	(0.061)	(0.071)	(0.061)	(0.132)	(0.156)	(0.080)	(0.106)
D $\times$ PL (β)	-0.162^{**}	-0.113	0.082	-0.124	-0.300	0.131	0.102
	(0.082)	(0.088)	(0.090)	(0.183)	(0.201)	(0.132)	(0.147)
PL (δ)	0.136^{**}	0.016	0.021	0.171^{**}	0.348^{**}	-0.008	0.098
	(0.068)	(0.097)	(0.070)	(0.078)	(0.155)	(0.072)	(0.072)
Constant (α)	-1.327^{***}	-2.360^{***}	0.867^{***}	-0.245^{***}	-0.359^{***}	0.054	0.243^{***}
	(0.038)	(0.075)	(0.051)	(0.060)	(0.139)	(0.053)	(0.036)
Observations	701	701	701	701	701	701	701

Table 5: Survey 3 regression results as indicated in Equation 6 for summary indices. The baseline corresponds to the Computerized Randomization. PL stands for a dummy variable for the Public Lottery arm, and D is a dummy variable indicating complience (instrumented by assignment into treatment). The coefficients for D $\times$ PL allow to test for Hypothesis 1a directly. I control for the lagged variable. The bootstrapped standard errors are clustered at the geographical cluster level. $^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.1$.

Figure 5 and Table 5 present the treatment effects for compliers (LATE). Compared to the ITT results, the evidence in favor of Hypothesis 1a is similar: the conventional Wald test rejects the null of no effect at the 5% level (p-value $= 0.04$) when looking at the three first indices. The wild cluster bootstrap yields a p-value of 0.08, and the leave-one-cluster-out strategy a maximum p-value of 0.12 and a median p-value of 0.04. Analyzing the seven indices together, the conventional Wald test also rejects the null of no effect at the 5% level (p-value $= 0.016$). Yet, when accounting for the limited number of clusters, the wild cluster bootstrap again does not reject the null hypothesis (p-value $= 0.30$), while the leave-one-cluster-out strategy yields a maximum p-value of 0.10 and a median p-value of 0.017.

These estimates suggest that the assignment mechanism, independent of direct treatment receipt, can shift treatment effects by as much as 0.15 SD in the outcomes of interest. The results are invariant to adding controls, both pre-selected and using Post-Double LASSO, as shown in Appendix K and Appendix L. This provides direct support against Hypothesis 1a: the average treatment effect is not invariant to awareness states induced by the assignment mechanism. More broadly, the results indicate that the experimental design itself can generate effects, implying that measured impacts reflect both the intervention and the context in which it is assigned.

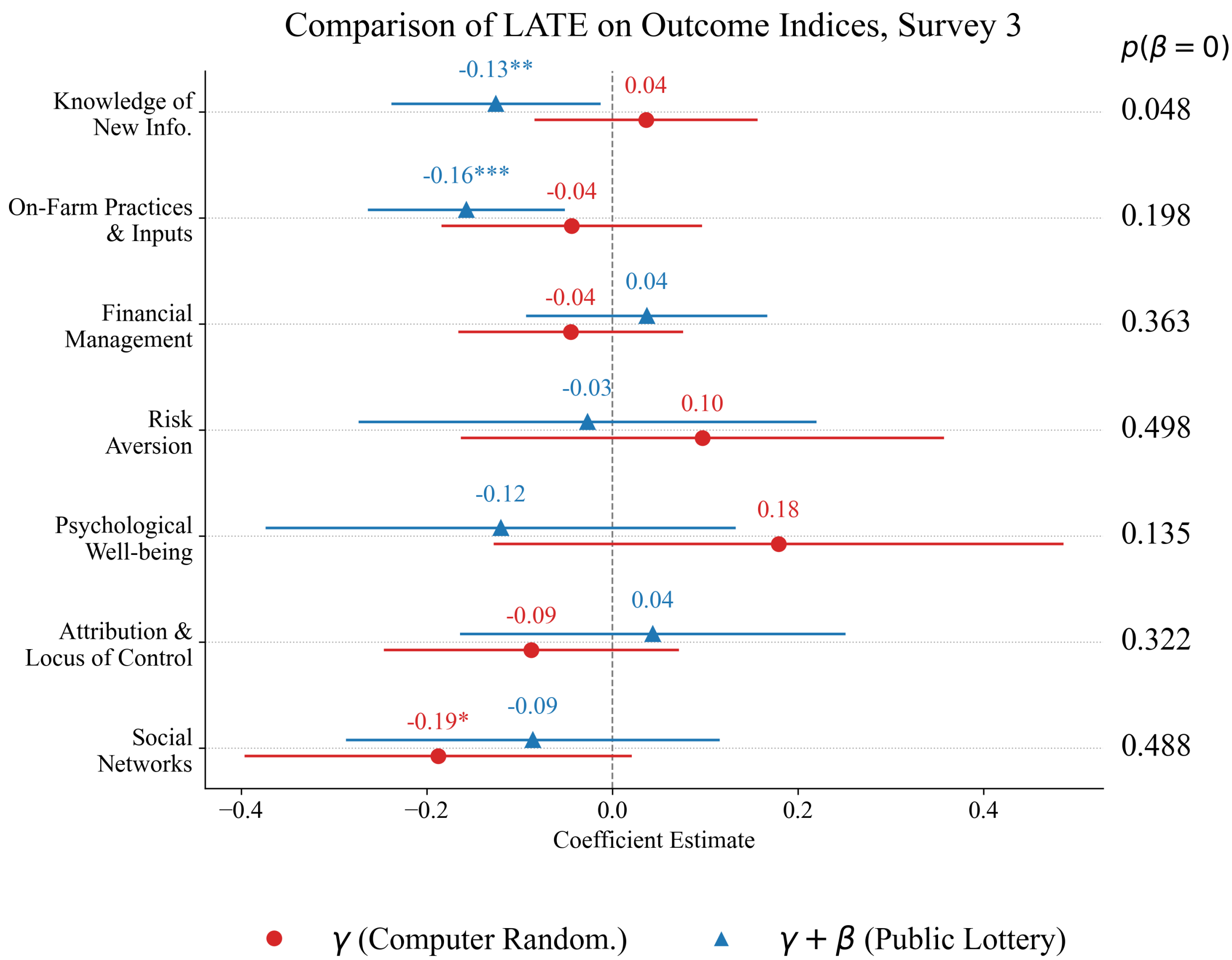


Figure 5: The figure presents coefficient estimates with 95% confidence intervals for the LATE in the Computerized Randomization arm (γ) and for the LATE in the Public Lottery arm ($\gamma + \beta$). The p-values shown above each point represent significance levels for the respective effects. The rightmost values show p-values for the test of $\beta = 0$, indicating whether the Public Lottery effect is statistically distinguishable from the computer randomization effect. The standard errors are clustered at the geographic cluster level, and the estimation was done by adding the baseline level of the corresponding outcome index as stated in Equation 6. All coefficients are expressed in standard deviation (SD) units, standardized relative to the pooled control group. Figure A4 shows the same coefficients for the data-driven indices.

8.2 No evidence of anticipatory assignment effects

Rejecting Hypothesis 1b would imply that effects appear after subjects know their assignment into treatment and control, but before the treated farmers receive the soil quality information and the subsidy. The hypothesis poses whether learning about one's treatment status–before receiving any actual inputs–already shifts behavior or attitudes in anticipation. In other words, whether the assignment step itself is working as a treatment.

To test it, I restrict the sample to survey 2 and test the equality of the intent-to-treat effects between the Computerized Randomization and the Public Lottery arms. In Equation 5, this corresponds to the null $H_0 : \gamma_2 = \gamma_2 + \beta_2$, which simplifies to $H_0 : \beta_2 = 0$. As in the previous subsection, I test this hypothesis separately for each index outcome and also perform a joint Wald test to determine if the first three indices, and then all seven interaction coefficients, are jointly equal to zero, combining the models to account for correlation in the outcome residuals.

Figure 6 displays the estimated coefficients $\widehat{\gamma_2}$ as the ITT for the Computerized Randomization arm, and $\widehat{\gamma_2 + \beta_2}$ as the ITT for the Public Lottery arm, both with 95 percent confidence intervals. It also reports the p-value from the two-sided equality test. Table 6 shows the results of estimating Equation 5 using survey 2 data.

	Info Adopt.	**Practices Inputs**	**Financial Manag.**	**Risk Aversion**	**Psych. Well-b**	**Attrib. LOC**	**Social Networks**
Treat (γ)	-0.060	-0.078	0.002	0.007	0.091	-0.126^{*}	-0.070
	(0.064)	(0.108)	(0.097)	(0.168)	(0.094)	(0.064)	(0.078)
PL (δ)	-0.113	-0.193^{**}	-0.171^{**}	0.157	0.146	-0.240^{***}	0.021
	(0.080)	(0.088)	(0.068)	(0.112)	(0.117)	(0.077)	(0.084)
Treat × PL (β)	0.120	0.172	0.090	-0.085	0.127	0.165	-0.132
	(0.089)	(0.128)	(0.127)	(0.220)	(0.135)	(0.116)	(0.140)
Constant (α)	-0.280^{***}	-2.188^{***}	-0.217^{***}	-0.242^{***}	-0.389^{***}	0.163^{**}	0.369^{***}
	(0.051)	(0.071)	(0.047)	(0.081)	(0.075)	(0.064)	(0.066)
Observations	701	701	701	701	701	701	701

Table 6: Survey 2 regression results as indicated in Equation 5 for summary indices. The baseline corresponds to the Computerized Randomization. PL stands for a dummy variable for the Public Lottery arm, and Treat is a dummy variable indicating whether the subject was assigned to treatment. The coefficients for Treat × PL allow to test for Hypothesis 1a directly. I control for the lagged variable. The bootstrapped standard errors are clustered at the geographical cluster level. $^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.1$.

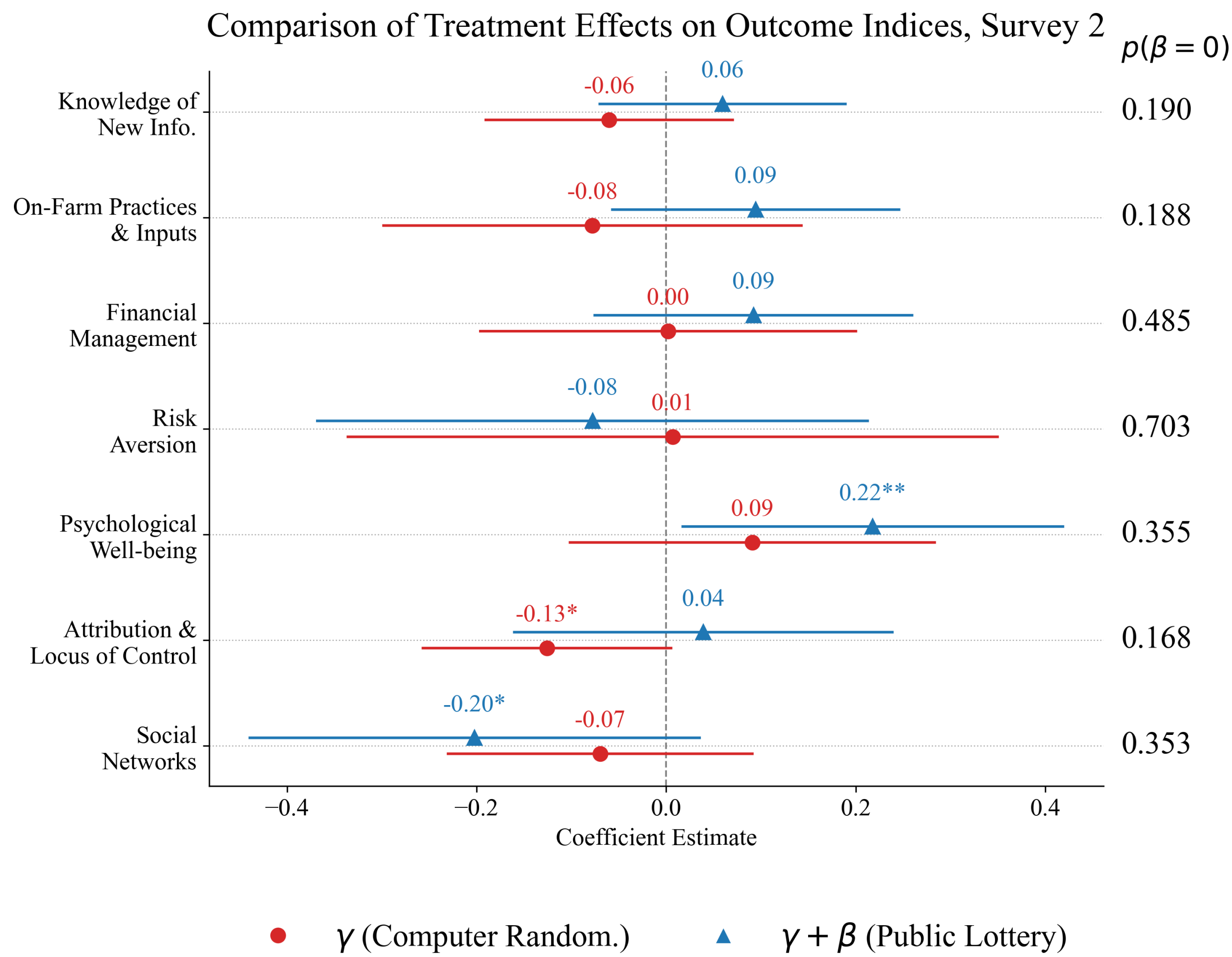


Figure 6: The figure presents coefficient estimates with 95% confidence intervals for the ITT in the Computerized Randomization arm (γ) and for the ITT in the Public Lottery arm ($\gamma + \beta$) after farmers were assigned into groups, but before the treatment was distributed. The p-values shown above each point represent significance levels for the respective effects. The rightmost values show p-values for the test of $\beta = 0$, indicating whether the Public Lottery effect is statistically distinguishable from the computer randomization effect. The standard errors are clustered at the geographic cluster level, and the estimation was done by adding the baseline level of the corresponding outcome index as stated in Equation 5. All coefficients are expressed in standard deviation (SD) units, standardized relative to the pooled control group. Figure A5 shows the same coefficients for the data-driven indices.

Overall, when looking individually at the seven indices, the evidence does not support systematic anticipation effects: The three main indices directly related to the intervention show no meaningful changes, and while some individual coefficients display differences; for example, the Psychological Well-Being index shows a 0.22 SD effect compared to the pooled control baseline under the Public Lottery arm, while the Social Networks index has a weekly significant negative effect in the Public Lottery arm of -0.20 SD (these patterns are not consistent across domains). Looking at the first three indices, joint tests confirm the absence of systematic differences: the standard Wald test yields a p-value of 0.144, the wild cluster bootstrap yields a p-value of 0.246, and the leave-one-cluster-out approach gives a median p-value of 0.156. Likewise, analyzing the seven indices together cannot reject the null of no systematic anticipation effects between arms.

When focusing on compliers in Figure 7 and Table 7, the point estimates are somewhat stronger but follow the same overall pattern. The Wald test of joint equality for the seven indices produces a borderline result (p-value = 0.098), yet the wild cluster bootstrap (p-value = 0.456) and the leave-one-cluster-out method (median p-value = 0.123) again fail to reject the null. Focusing on the three main indices related to agriculture, in all cases (standard Wald test, wild cluster bootstrap, and leave-one-cluster-out method), the null hypothesis of no joint significance cannot be rejected.

Taken together, these results provide evidence in favor of Hypothesis 1b: the differences between treatment and control groups after the assignment step don't seem to depend on the awareness state. Adding controls does not change this conclusion (refer to Appendix K and Appendix L). Therefore, the assignment step alone does not appear to act as an independent treatment. Unlike at endline (survey 3), where assignment mechanisms shaped estimated treatment effects, here the evidence suggests that differences emerge only once the intervention itself is delivered, not at the assignment stage.

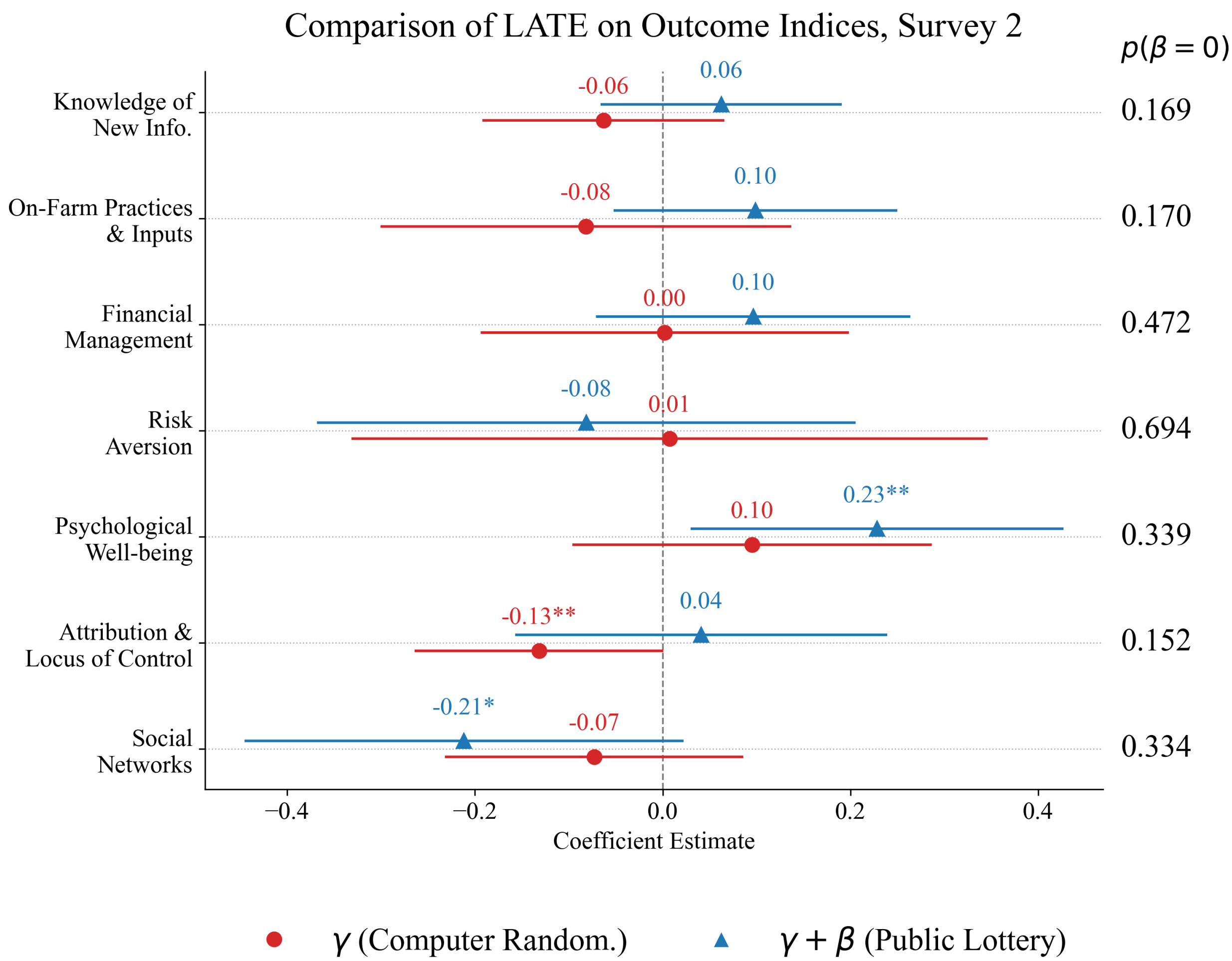


Figure 7: The figure presents coefficient estimates with 95% confidence intervals for the LATE in the Computerized Randomization arm (γ) and for the LATE in the Public Lottery arm ($\gamma + \beta$) after farmers were assigned into groups, but before the treatment was distributed. The p-values shown above each point represent significance levels for the respective effects. The rightmost values show p-values for the test of $\beta = 0$, indicating whether the Public Lottery effect is statistically distinguishable from the computer randomization effect. The standard errors are clustered at the geographic cluster level, and the estimation was done by adding the baseline level of the corresponding outcome index as stated in Equation 6. All coefficients are expressed in standard deviation (SD) units, standardized relative to the pooled control group. Figure A6 shows the same coefficients for the data-driven indices.

	Info Adopt.	Practices Inputs	Financial Manag.	Risk Aversion	Psych. Well-b	Attrib. LOC	Social Networks
D (γ)	-0.063	-0.082	0.002	0.007	0.095	-0.132^{**}	-0.073
	(0.065)	(0.111)	(0.099)	(0.172)	(0.097)	(0.067)	(0.080)
D $\times$ PL (β)	0.126	0.181	0.094	-0.089	0.133	0.173	-0.139
	(0.091)	(0.132)	(0.131)	(0.226)	(0.139)	(0.120)	(0.144)
PL (δ)	-0.114	-0.194^{**}	-0.172^{***}	0.158	0.144	-0.241^{***}	0.023
	(0.079)	(0.087)	(0.067)	(0.111)	(0.114)	(0.076)	(0.083)
Constant (α)	-0.280^{***}	-2.188^{***}	-0.217^{***}	-0.242^{***}	-0.389^{***}	0.164^{***}	0.370^{***}
	(0.050)	(0.070)	(0.046)	(0.080)	(0.074)	(0.063)	(0.065)
Observations	701	701	701	701	701	701	701

Table 7: Survey 2 regression results as indicated in Equation 6 for summary indices. The baseline corresponds to the Computerized Randomization. PL stands for a dummy variable for the Public Lottery arm, and D is a dummy variable indicating complience (instrumented by assignment into treatment). The coefficients for D $\times$ PL allow to test for Hypothesis 1a directly. I control for the lagged variable. The bootstrapped standard errors are clustered at the geographical cluster level. $^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.1$.

8.3 Asymmetric Impact of the Assignment Mechanism: Control group affected by Assignment Mechanism

The central finding of this study emerges from the analysis of Hypothesis 2a, which posits that the awareness state affects control and treatment groups uniformly. Rejecting this hypothesis would imply an asymmetry, which might lead to non-separability as (Hypothesis 2). The evidence strongly suggests that the awareness state induced by the assignment mechanism has a significant impact on the outcomes of the control group, while leaving the treated group largely unaffected.

Figures 8a and 8b visualize the difference in estimated outcomes for control subjects between the PL arm ($\alpha + \delta$) and the CR arm (α). The difference, captured by the PL coefficient (δ), represents the assignment effect on the control group.

In survey 2, before the treatment is distributed, control subjects in the PL arm exhibit significantly different outcomes than their counterparts in the CR arm. Specifically, they report greater changes in On-Farm Practices & Inputs (2.38 SD vs 2.19 SD, the difference being significant at the 5% level) and Financial Management (0.39 SD vs 0.22 SD, the difference being significant at the 5% level). Furthermore, their Attribution & Locus of Control shifts towards external factors, a highly significant difference compared to the CR control group, who are attributing their outcomes to their own actions (-0,08 SD for the PL arm, and 0.16 SD for the CR arm). These results, visualized in Figure 8a, provide clear evidence of anticipatory behavioral changes driven solely by the assignment method.

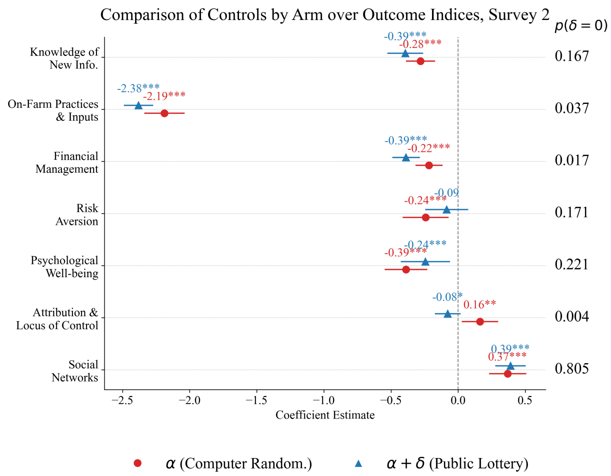
Figure 8b shows that these differences persist and evolve in survey 3 after the treatment period. The PL control group reports significantly higher knowledge of new Information (δ=0.135, p-value $<$ 0.1), higher Risk Aversion (δ=0.171, p-value < 0.05), and greater Psychological Well-being (δ=0.345, p-value < 0.05) compared to the CR control group. Across both survey rounds, the assignment mechanism consistently and significantly alters the outcomes of the control group, demonstrating that the control group's behavior is context-dependent.

The On-Farm Practices and Inputs index, together with the changes in the Financial Management index, shows a significant difference between the Computer Randomization and the Public Lottery arms, as shown in the p-value for $H_0 : \delta = 0$ in Figure 8a. The Wald test for the joint null hypothesis that the assignment effect is zero across the three indices is consistently rejected, with a standard p-value of 0.0015, a wild cluster bootstrap p-value of 0.048, and a maximum p-value of 0.035 for the leave-one-cluster-out analysis. The differences are more nuanced for the controls in survey 3: Figure 8b shows a weak difference between the two arms for the knowledge of new Information index, and this is also reflected in the joint Wald test, where it is not possible to reject the null of joint insignificance.

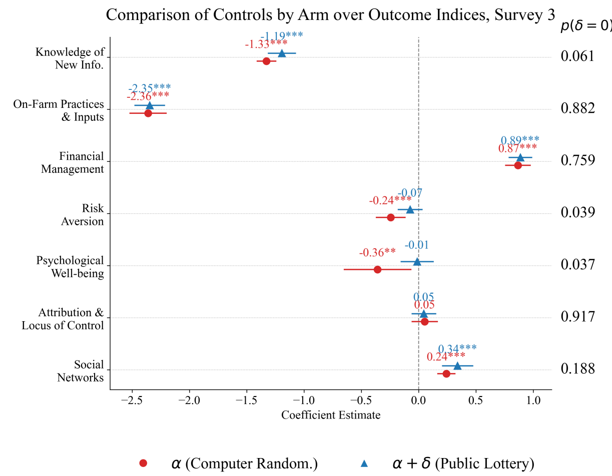
When analyzing the seven indices, the evidence for an assignment effect on the control group remains strong in the pre-treatment period, and it is also significant for survey 3. A standard Wald test for the joint null hypothesis that the assignment effect is zero across all seven indices is rejected at the 1% level in survey 2 (p-value < 0.001) and the 5% level in survey 3 (p-value $= 0.028$).

To assess the robustness of this finding against potential issues related to the number of clusters, I conduct the two aforementioned further analyses: a wild cluster bootstrap, which can be more conservative, yields p-values of 0.112 in survey 2 and 0.426 in survey 3. The leave-one-cluster-out analysis reveals how influential individual clusters are. For survey 2, the joint effect remains highly significant regardless of which cluster is dropped (maximum p-value $= 0.004$), indicating a robust finding. For survey 3, the result is once again fragile; while the median p-value across iterations remains significant at 0.027, the effect loses significance if one specific cluster is removed (maximum p-value $= 0.272$).

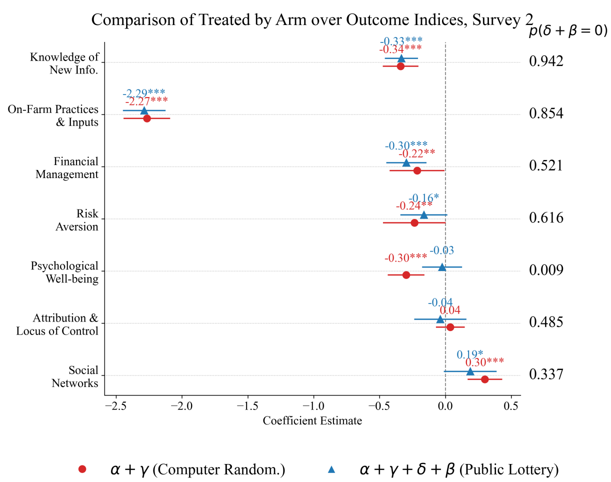
Taken together, these tests indicate a robust and statistically significant assignment effect on the control group in the pre-treatment period (survey 2) for both the agricultural indices alone and the seven indices together. The effect persists post-treatment (survey 3), although it is only significantly different from zero when analyzing the seven indices, and its statistical significance is more sensitive to the choice of inference method.



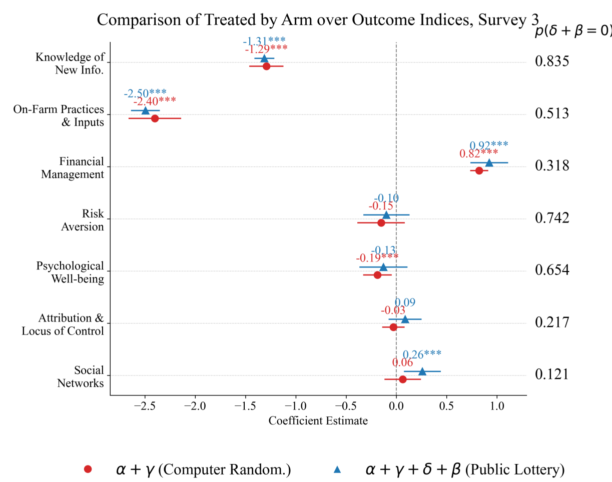
(a) Control, survey 2



(b) Control, survey 3



(c) Treated, survey 2



(d) Treated, survey 3

Figure 8: The figures present coefficient estimates with 95% confidence intervals for the control and treated subjects' outcome values in the Computerized Randomization arm and the Public Lottery arm. Figures (a) and (b) show the coefficients for the control subjects in surveys 2 and 3, respectively. Figures (c) and (d) show the coefficients for the treated subjects in surveys 2 and 3. The p-values shown above each point represent significance levels for the respective effects. The rightmost values show p-values for the test of $\delta = 0$, indicating whether the Public Lottery effect is statistically distinguishable from the computer randomization effect. The standard errors are clustered at the geographic cluster level, and the estimation was done by adding the baseline level of the corresponding outcome index as stated in Equation 6. All coefficients are expressed in standard deviation (SD) units, standardized relative to the pooled control group. Figure A7 shows the equivalent plots for the data-driven indices.

In stark contrast, the assignment method had almost no differential impact on subjects assigned to treatment. Figures 8c and 8d compare the expected outcomes for treated subjects in the CR arm ($\alpha + \gamma$) with those in the PL arm ($\alpha + \gamma + \delta + \beta$). The test for a significant difference (H_0: $\delta + \beta = 0$) cannot be rejected for nearly any outcome in either survey round. The only exception is a significant negative effect on Psychological Well-being in survey 2 for the PL-treated group. A joint test of significance for the difference ($\delta + \beta$) across all seven indices fails to reject the null in both survey waves, and likewise when analyzing the first three agricultural indices.

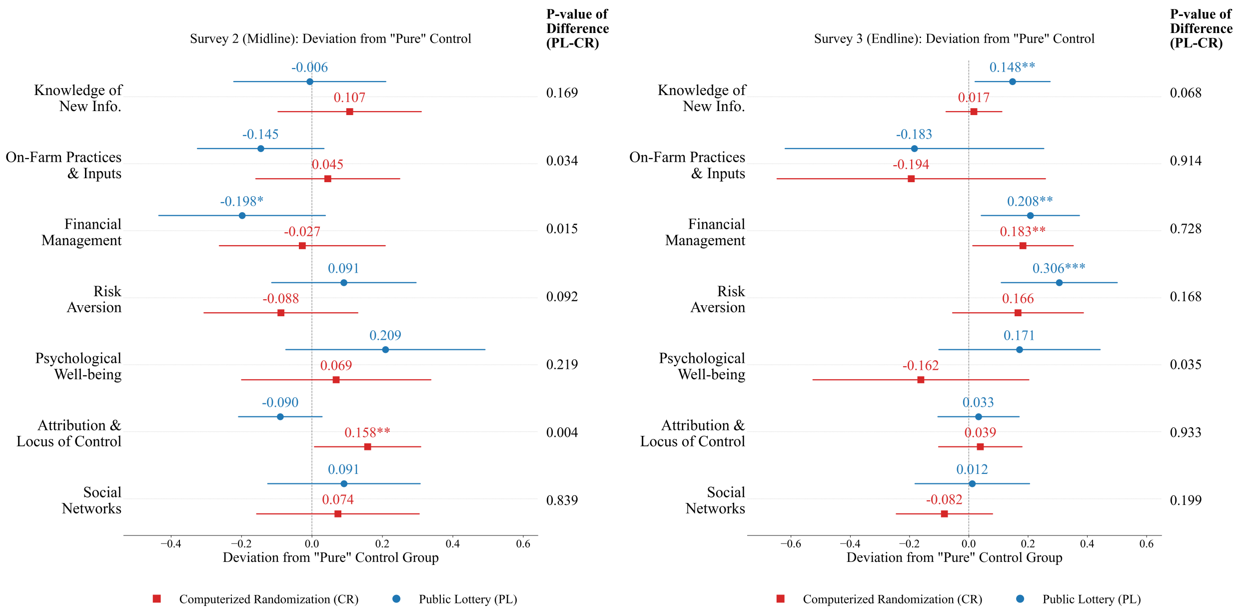
This asymmetry provides strong evidence against Hypothesis 2a and, consequently, in favor of Hypothesis 2. The awareness state is not separable from the treatment status because its influence is almost entirely concentrated on the control group. This implies that the control group in a high-awareness design (like a Public Lottery) does not serve as a pure counterfactual for the treated group; rather, it reflects a complex condition of [Not Treated + Awareness of Not Being Treated], which is demonstrably different from the condition of [Not Treated] in a lower-awareness context.

Moreover, to mitigate the changes caused by seasonal trends, I analyze the behaviors of the control and treated groups in relation to their benchmarks: the pure control arm and the pure treatment arm, respectively. In Figures 9a and 9b I plot coefficients $\beta^{CR,C}$ and $\beta^{PL,C}$ as specified in Equation 8, as well as the p-value of the null test $H = 0$: $\beta^{CR,C} = \beta^{PL,C}$, to evaluate if the difference between the mixed arms is statistically different.

Even before the treatment is administered, the choice of assignment mechanism creates statistically significant and profound differences between the control groups. The midline results from survey 2, shown in Figure 9a, reveal that the PL and CR control groups are significantly different from each other: in the On-Farm Practices & Inputs (p-value = 0.034), Financial Management (p-value = 0.015), Risk Aversion (p-value = 0.092), and Attribution & Locus of Control (p-value = 0.004) indices the coefficients are statistically different and in general move in opposite directions. To illustrate, while control subjects in the CR arm developed a significantly more internal locus of control compared to the Pure Control group (0.158 SD), those in the PL arm trended in the opposite direction. This demonstrates that the two assignment mechanisms induced fundamentally different psychological and behavioral responses among control subjects.

Figure 9b shows these differences persist in survey 3, with the PL control group diverging dramatically from the Pure Control benchmark. Specifically, the PL control group shows significant positive deviations from the Pure Control group in knowledge of new Information (0.148 SD), Financial Management (0.208 SD), and most notably, a large increase in Risk Aversion (0.306 SD). In contrast, the CR control group remains behaviorally much closer to the Pure Control baseline, with only one significant deviation in the Financial Management index (0.183 SD).

The direct comparison between the two mixed arms in survey 3 shows that the PL control group sought significantly more information than the CR control group (p-value = 0.068) and reported significantly higher Psychological Well-being (p-value = 0.035), when compared to the Pure Control group. This analysis provides the strongest evidence in favor of Hypothesis 2a and the non-separability of the awareness state. The behavior of the control group is not a fixed baseline; it is highly contingent on the experimental context. The high-salience Public Lottery created a control group that, over time, became more information-seeking, more risk-averse, and reported higher well-being than an “unaware” control group.

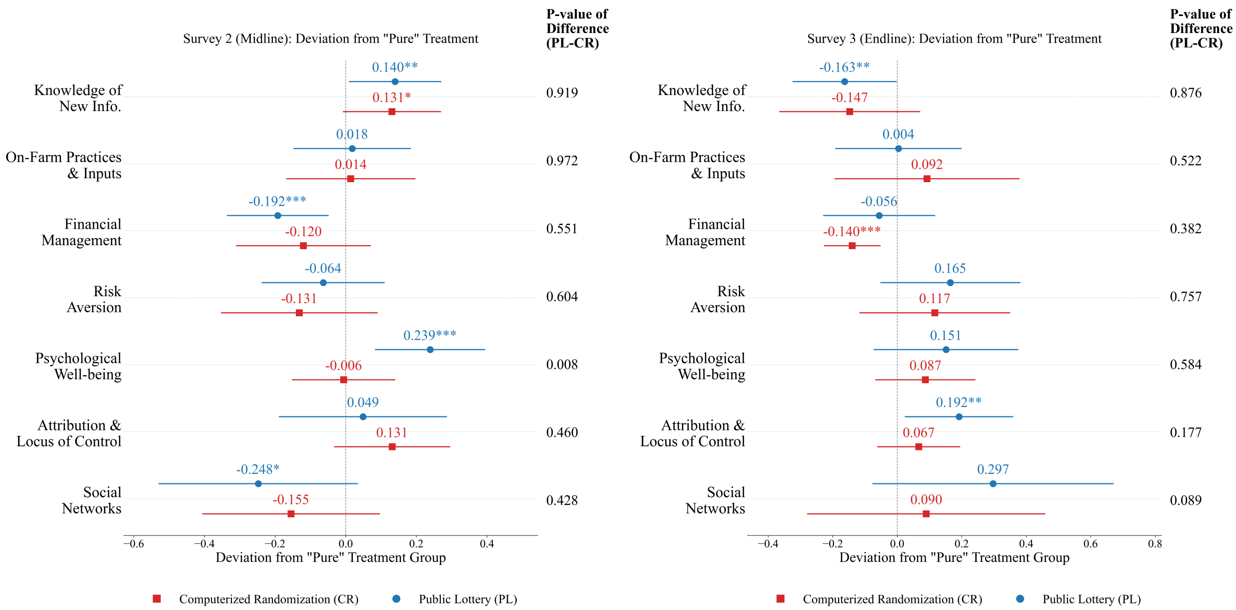


(a) Survey 2

(b) Survey 3

Figure 9: Controls by Arm compared to the Pure Control Benchmark. The figures present coefficient estimates with 95% confidence intervals for the control subjects in the Computerized Randomization arm and in the Public Lottery arm with respect to the Pure Control arm. Figure 9a shows the coefficients in survey 2, while Figure 9b shows the coefficients in survey 3. The rightmost values show the p-values for the test of $\beta^{CR,C} = \beta^{PL,C}$, indicating whether the Public Lottery coefficient is statistically distinguishable from the Computerized Randomization coefficient. The standard errors are clustered at the geographic cluster level, and the estimation is done by adding the baseline level of the corresponding outcome index as stated in Equation 8. All coefficients are expressed in standard deviation (SD) units, standardized relative to the pooled control group. Figure A8 shows the equivalent plots for the data-driven indexes.

Complementing the analysis of the control group, I examine how the awareness state affects subjects assigned to the treatment group in Figures 10a and 10b with respect to the Pure Treatment arm, following Equation 7.



(a) Survey 2 (b) Survey 3

Figure 10: Treated by Arm compared to the Pure Treatment Benchmark. The figures present coefficient estimates with 95% confidence intervals for the treated subjects in the Computerized Randomization arm and in the Public Lottery arm with respect to the Pure Treatment arm. Figure 10a shows the coefficients in survey 2, while Figure 10b shows the coefficients in survey 3. The rightmost values show the p-values for the test of $\beta^{CR,C} = \beta^{PL,C}$, indicating whether the Public Lottery coefficient is statistically distinguishable from the Computerized Randomization coefficient. The standard errors are clustered at the geographic cluster level, and the estimation is done by adding the baseline level of the corresponding outcome index as stated in Equation 7. All coefficients are expressed in standard deviation (SD) units, standardized relative to the pooled treatment group. Figure A9 shows the equivalent plots for the data-driven indexes.

The results from survey 2 reveal significant anticipatory effects, particularly in the Public Lottery arm. Figure 10a shows the estimated deviations from the Pure Treatment group. They report higher knowledge of new Information (0.14 SD), a significant change in Financial Management (-0.19 SD), and a notable decrease in their Social Networks index (-0.248 SD). Most strikingly, they experience a large and highly significant boost in Psychological Well-being (0.239 SD) compared to the Pure Treatment group. In contrast, treated subjects in the Computerized Randomization (CR) arm show much less deviation, with only an increase in the knowledge of new Information index.

Critically, the difference between the two mixed arms is itself significant for the Psychological Well-being index (p-value = 0.008). This indicates that the high-awareness assignment mechanism of the Public Lottery generated a distinct psychological "boost" for the treated even before the intervention began, an effect not present in the lower-awareness CR arm. This provides evidence that the awareness state itself generates anticipatory behavioral and psychological responses among the treated.

In survey 3, after the treatment has been delivered, the pattern of effects shifts. As shown in Figure 10b, the direct difference between the PL and CR arms is no longer statistically significant for any outcome index. However, both groups continue to exhibit significant deviations from the Pure Treatment baseline, albeit on different dimensions. Treated subjects in the Computerized Randomization arm now show a significant deviation in Financial Management (-0.140 SD), an effect that was not significant for this group in survey 2. Meanwhile, the Public Lottery arm also shows persistent deviations. Their initial boost in information adoption reverses into a significant decrease relative to the Pure Treatment group (-0.163 SD). They also report a significant positive shift in their Attribution & Locus of Control index (0.192 SD), suggesting they attribute outcomes more to personal effort post-treatment compared to those in the Pure Treatment arm.

In sum, this analysis provides nuanced support for Hypothesis 2a and the non-separability of the awareness state even in the case of the treated subjects: the results demonstrate that the treated group is not immune to awareness effects. Being in an experimental arm with a control group alters behavior and well-being compared to a setting where everyone receives the intervention. Moreover, the level of awareness matters. The high-awareness Public Lottery mechanism generates significant anticipatory effects (notably on psychological well-being) that are not present in the lower-awareness Computerized Randomization arm. These effects interact with treatment over time. The initial differences between the two mixed arms dissipate post-treatment, but both arms remain behaviorally distinct from the Pure Treatment benchmark.

The empirical evidence reveals an asymmetric impact of the awareness state, conditional on treatment assignment. The comparison between the high-awareness (Public Lottery) and low-awareness (Computerized Randomization) arms reveals that the assignment mechanism significantly alters the outcomes of the control group across multiple indices. This finding is reinforced when benchmarking against a Pure Control group–control subjects in the mixed arms, particularly the Public Lottery arm, exhibit significant deviations in outcomes such as risk aversion, financial management, and information adoption. In contrast, the outcomes for the treated group are largely insensitive to the change in awareness between the mixed arms. This asymmetry demonstrates that the control group is not a passive baseline but an active state whose behavior is contingent on the experimental context.

The observed asymmetry provides direct evidence against the separability of the awareness state from treatment status. The effect of awareness is not an additive constant but functions as an interactive term that is pronounced for control subjects and attenuated for treated subjects. Furthermore, the analysis shows that the outcomes of the treated group are also context-dependent, with subjects in both mixed arms exhibiting significant behavioral deviations from the Pure Treatment benchmark. This specific comparison, however, must be interpreted with caution: the Pure Treatment arm differs from the mixed arms not only in the absence of a control group but also in its 100% treatment density, which may generate distinct information spillovers. The design cannot definitively disentangle the effect of awareness from the mechanical effect of interference spillovers.

8.4 Results of the diagnosis tests for separability

Figure 11 presents the estimation of the null hypothesis 9 (H_0: $S \geq 0$) born from Proposition 1, which tests whether treatment effects are separable from the awareness state. For each summary index, the estimated value of the separability statistic S is displayed along with its bootstrap confidence interval and corresponding one-sided p-value. The p-value represents the probability of observing a non-negative value of S under the null hypothesis $H_0 : S \geq 0$, providing an empirical assessment of whether treatment effects are independent of the assignment mechanism.

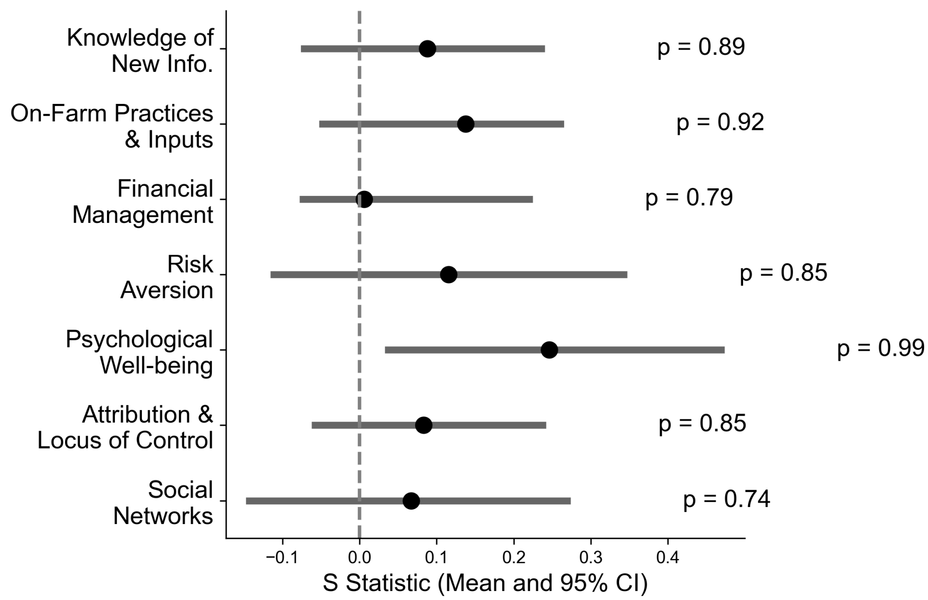


Figure 11: Results from testing the null hypothesis 9 on separability using the summary indices. The plot displays the calculated statistic S by index outcome, along with their 95% confidence intervals, where $S = T_1 + T_2 - T_3$. In the case of the On-Farm Practices & Inputs and the Financial Management indices, $S = |T_1| + |T_2| - |T_3|$, as the direction of treatment effects is uninterpretable. The p-values shown are the one-sided bootstrap p-values for $S \geq 0$ corresponding to each index; intuitively, the p-value provides the empirical probability that the null hypothesis $S \geq 0$ holds. Then, the lower the p-value, the stronger the evidence against the null hypothesis. See Appendix M for the results using Bayesian bootstrap.

This diagnostic test captures a necessary but not sufficient condition for separability. A rejection of the null indicates non-separability–that is, that the assignment mechanism and treatment interact in determining outcomes. However, a failure to reject the null does not imply that separability holds, only that no significant deviation is detected given the sample size and precision. In this setting, I cannot reject separability for any of the outcome indices under this stringent test, suggesting that there is no firm evidence of interaction effects between treatment and awareness state. This result complements the earlier findings: while awareness influences outcomes among controls, the interaction between awareness and actual treatment appears limited once the intervention is implemented.

As for Proposition 2, Figure 12 presents the estimation of the statistic for the null hypothesis 10 ($H_0 : S^* \geq 0$). As expected from the previous results, the null hypothesis is rejected for three of the seven indices: Knowledge of New Information, Psychological Well-being, and Social Networks. These results then provide evidence against separability: the term $[h(1,0) - h(0,1)]$ is negative enough to flip the sign of S^* under minimal assumptions. That is, the control group, under a state of awareness, reacts strongly enough.

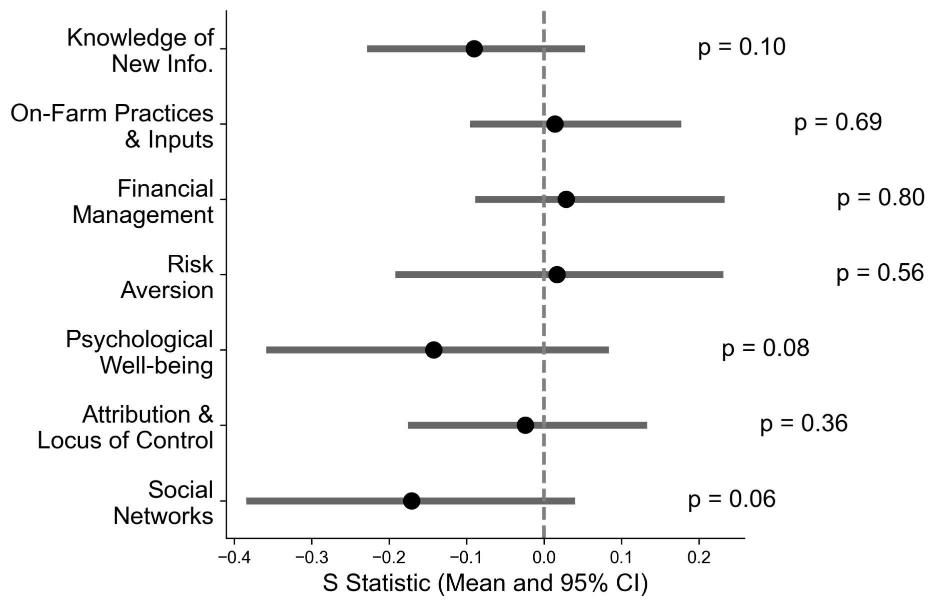


Figure 12: Results from testing the null hypothesis 10 on separability using the summary indices. The plot displays the calculated statistic S^* by index outcome, along with their 95% confidence intervals, where $S^* = T_1 + T_2 - T_3^*$. In the case of the On-Farm Practices & Inputs and the Financial Management indices, $S^* = |T_1| + |T_2| - |T_3^*|$, as the direction of treatment effects is uninterpretable. The p-values shown are the one-sided bootstrap p-values for $S^* \geq 0$ corresponding to each index; intuitively, the p-value provides the empirical probability that the null hypothesis $S^* \geq 0$ holds. Then, the lower the p-value, the stronger the evidence against the null hypothesis. See Appendix M for the results using Bayesian bootstrap.

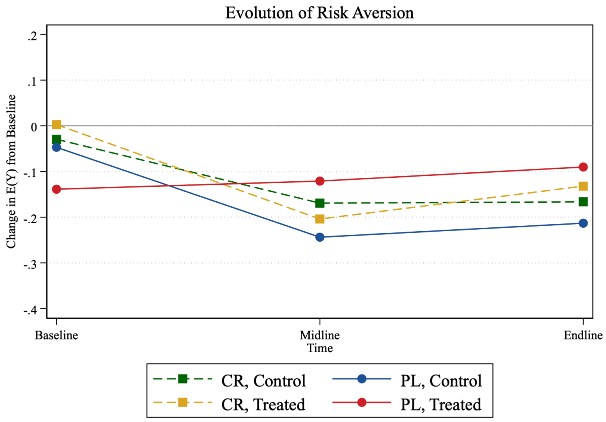
9 Possible Mechanisms

The previous section established that treatment effects differ across assignment mechanisms: the Public Lottery arm shows a reduction in the knowledge of new Information index, and a significant change in the On-Farm Practices and Inputs, while there is no measurable treatment effect when looking at the Computerized Randomization arm. In this section, I explore which underlying behaviors might be driving this result: I discuss behavioral and psychological mechanisms that could explain the heterogeneous impacts. First, I inspect the changes during surveys 2 and 3 for the treated group and the control group in the Public Lottery and the Computerized Randomization arms for the four non-agricultural indices with respect to the benchmarks. Then, I examine further measures taken in surveys 2 and 3 to try to disentangle the behavioral responses even further.

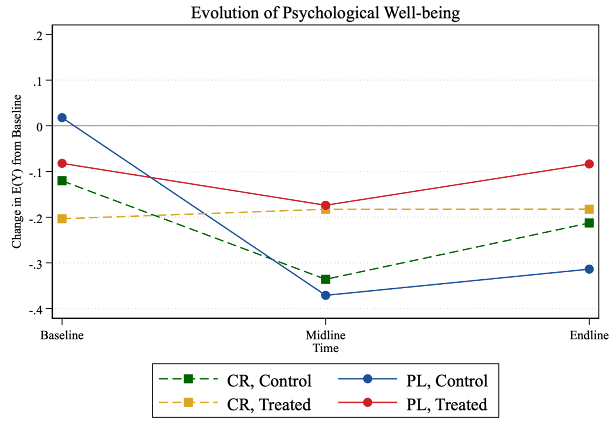
Figure 13 tracks the change in the expected outcome for each one of the four experimental groups, PL-treated, PL-control, CR-treated, and CR-control, across the three surveys. The figure shows the absolute adjusted mean of each group at each survey round, controlling for baseline levels via the ANCOVA regression. The points at baseline (survey 1) show the initial disparities among the groups, while the points at midline (survey 2) and endline (survey 3) show the post-regression-adjusted expected levels. Then, it is easier to compare levels across groups at each survey, interpreting the changes as standard deviations from the pooled control groups at baseline.

The first thing to note is that the PL-treated group is more risk-loving from baseline, but its assignment into treatment made it more risk-averse. This is opposite to what happened in the other groups, where the assignment step made them more risk-loving, particularly so for the PL-control group. At the endline, this behavior creates the gap treatment effect in risk aversion for the PL arm, even when both groups became more risk-loving. There is no differential behavior between groups in the CR arm, but both are also more risk-loving at the endline (which could correspond to a seasonal trend). The combination of assignment through a public lottery and assignment into treatment appears to make farmers more cautious than the general trend. The salience of the assignment may interact with the perceived self-attribution, thereby explaining these trends.

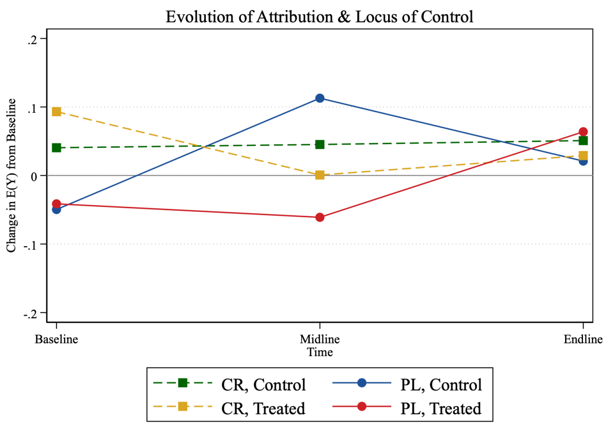
The general literature on motivated reasoning indicates that subjects selected for a treatment or given some degree of advantage tend to interpret their selection as a signal of higher ability or good fortune, leading to increased optimism, self-confidence, and risk-taking (see (Kőszegi, 2006; Brunnermeier and Parker, 2005; Zimmermann, 2020)). Most studies find that control subjects, who lack positive reinforcement or feedback, do not exhibit similar increases in risk-taking behavior. The finding that controls exhibit greater risk-taking behavior in the Public Lottery arm is thus surprising. One possible psychological mechanism for increased risk tolerance among control subjects is the gambler's fallacy: following a perceived negative event (not being selected for treatment), individuals may believe their luck is going to change, prompting greater willingness to take risks (Rabin, 2002). This effect, while not aligned with classic motivated reasoning, reflects a systematic bias in beliefs about random sequences and luck.



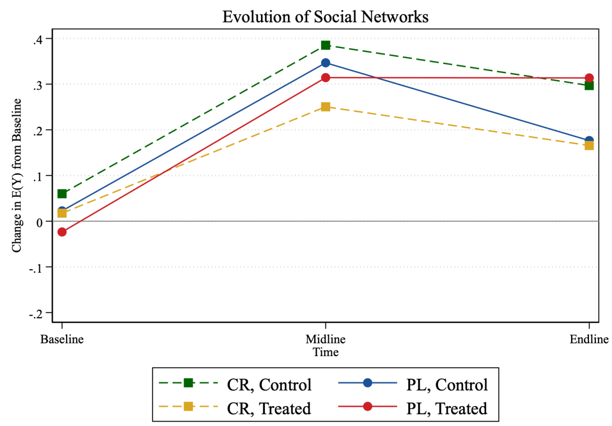
(a)



(b)



(c)



(d)

Figure 13: Evolution of the Risk Aversion, Psychological Well-Being, Attribution and Locus of Control, and Social Networks indices over time. The midline (survey 2) captures the anticipation effects and, therefore, the different assignment effects. The endline (survey 3) captures the treatment effect. The red and blue lines correspond to the Public Lottery arm, while the dashed yellow and dashed green lines correspond to the Computerized Randomization arm. The y-axis is expressed in standard deviations with respect to the pooled control of the two arms (Computerized Randomization and Public Lottery) at baseline.

On the psychological well-being index, the CR-treated group is marginally more optimistic in each survey round. In contrast, the other groups show a decrease in optimism, particularly in the PL-control group. There seems to be an interplay between comparison effects and fairness perceptions. An initial and temporary reduction in self-esteem might be reversed by receiving the treatment: the PL-treated group recovers to its initial level by survey 3 (endline), but this is not true for the controls, who are stuck at a lower value. The PL then creates a difference of more than 0.2 SD between the groups, where the treated end up much better off than the control group. This is not the case for the CR arm, whose adverse effects on the psychological well-being of the control group are short-lived.

These behaviors are contrary to my initial expectation that greater transparency regarding the assignment would mitigate psychological impacts. Specifically, I anticipated that control subjects would experience resentful demoralization or emotional distress as a result of being excluded from the intervention (Cook, Campbell, and Day, 1979), while treated individuals would report higher well-being. Furthermore, I expected these effects to be most pronounced in the Computerized Randomization arm, where comprehension of the assignment process was lower. Instead, it is the arm with higher awareness of the assignment that produces a lasting well-being gap, suggesting assignment transparency may shape psychological reactions in unanticipated ways.

In the attribution and locus of control index, all groups end at a similar level of the index, but with very different paths: the PL-control has a dramatic increase in the index by survey 2, showing that they attribute their outcomes to their actions. This is the opposite of what the treated group experiences in the same arm, maybe internalizing luck from being selected through a public lottery into everyday life. The CR-treated also tends to attribute less to its own actions, while the CR-control doesn't appear to be reading any signal from the intervention. It is interesting that at the endline, farmers in the PL-treated group do change their opinion and now over-attribute their outcomes to their actions.

The social networks index is particularly interesting: in the Public Lottery arm, there is little difference between treated and control farmers at survey 2, but treated farmers become and remain more socially engaged over time, while controls do not. In contrast, in the Computerized Randomization arm, it is the controls who experience a persistent boost in their social networks index. This pattern suggests that assignment mechanisms can alter community interactions and the subsequent flow of information. Having more socially engaged treated farmers is desirable for information diffusion, but it is also essential for control farmers to remain connected and avoid isolation. In line with theoretical perspectives on network interventions (Valente, 2012; VanderWeele and An, 2013; Fafchamps, 2015), these results suggest that assignment mechanisms may serve as a form of "network alteration," actively shaping engagement and information flow. Therefore, collecting detailed network outcome data is valuable–not only because the *treatment* can change the network structure (Chandrasekhar and Jackson, 2025), but also because participants' awareness of the research process and assignment itself may influence the network dynamics.

To sum up, assignment through a public lottery appears to make participants more emotionally reactive and socially engaged, while assignment through private computerized randomization produces more muted and internally consistent responses. The public nature of the lottery seems to amplify both positive and negative reactions–greater shifts in risk attitudes, psychological well-being, and social engagement–consistent with heightened salience and social comparison effects.

Some of these results are consistent with participants' responses to specific questions in each survey. In survey 3, participants reported their perceptions of their soil quality with respect to the rest of the village. While there is no apparent difference in the control group, where 58% of the PL arm reported their soil to be the best vs 53% in the CR arm, the treated in the Public Lottery are significantly more optimistic about their soil quality than the same group in the Computerized

Randomization arm, as shown in Figure 14.

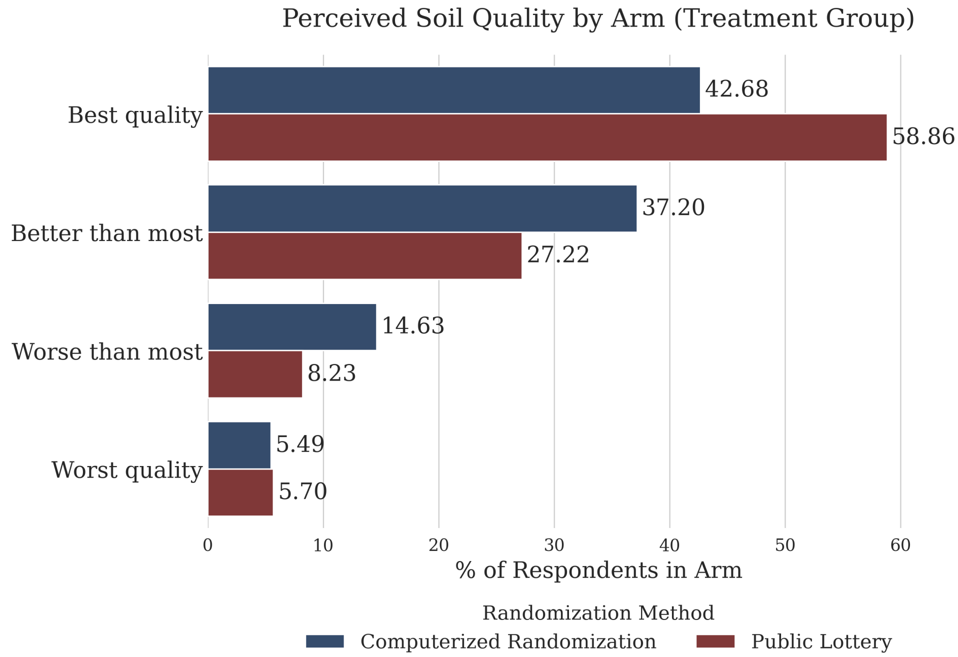
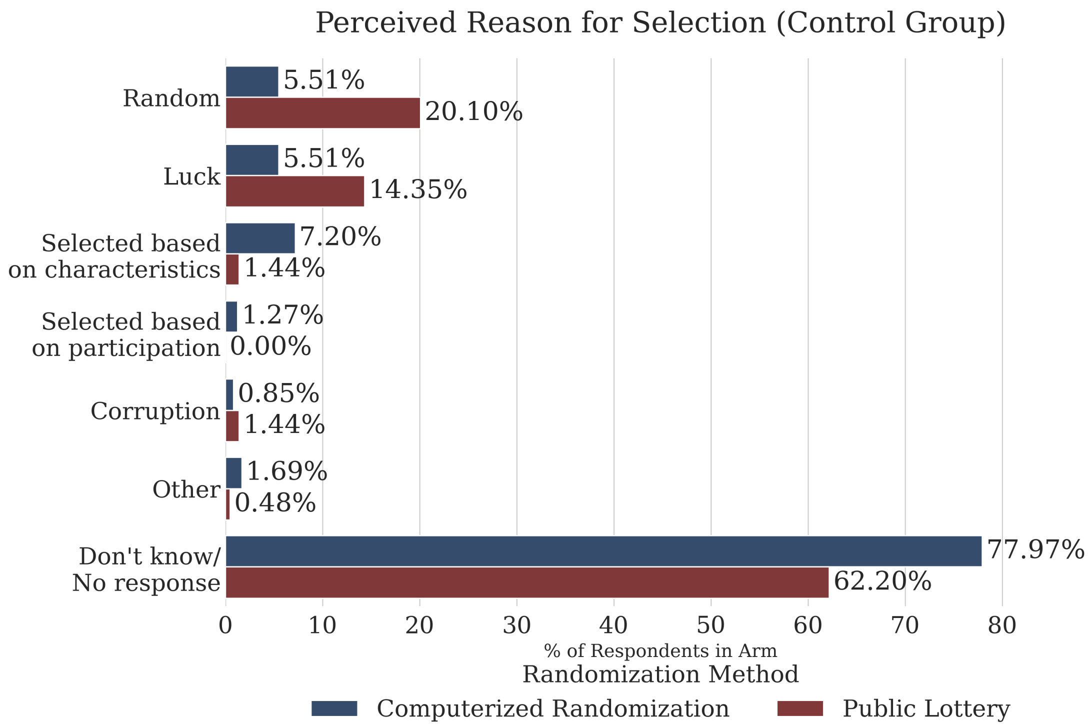
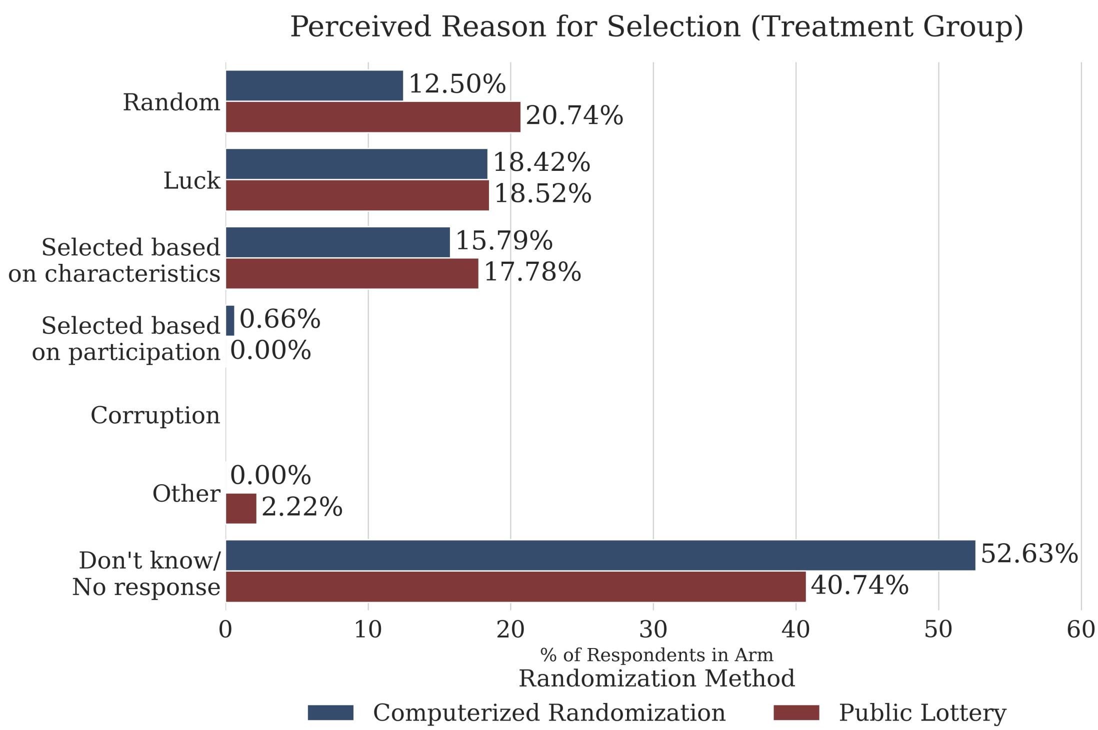


Figure 14: Percentage of farmers in each arm responding to the question "We did multiple soil tests in your village. Given that you know the results now, how do you think your soil is compared to the rest of the village?" during survey 3.

All these differences could correspond to farmers' interpretation of the assignment method. At the very end of survey 3, farmers had to explain the assignment method in general and then explain why they were or weren't selected. These were qualitative questions that surveyors then coded into categories. Figure 15 shows the results: in the PL arm, 20% of both treated and control farmers could perfectly explain the selection method. In the CR, the correct response is highly unbalanced: only 5.5% of the control group responded with something that can clearly be attributed to randomness, while 12.5% did so in the treatment group. For the control group in the CR-arm, 7.20% believe farmers were selected based on their personal characteristics, while this number is higher for the treated. What is surprising is that this is the case even for treated farmers in the Public Lottery arm. These might be farmers who could not attend the public lottery and did not ask their proxy (usually a son or brother) or their neighbors how the process took place.



(a)



(b)

Figure 15: Percentage of farmers responding to their beliefs about the assignment method at the end of Survey 3, coded by surveyors. Figure (a) shows the responses of farmers assigned to control, Figure (b) the responses of farmers assigned to treatment.

10 Discussion

My paper examines the often-overlooked but core assumption of ignorability of the research process underlying diverse research designs used to measure the effect of interventions. Designing an RCT around a standard agricultural information provision intervention in India, I provide evidence that the assignment process can generate behavioral responses that confound causal inference, leading to different treatment effects.

This finding challenges the assumption that the research process is ignorable, even in an RCT where the assignment is statistically exogenous. By introducing the awareness state into the potential outcomes conceptual framework, I discuss how participants' awareness of group assignment, the method by which assignment is conducted, and their knowledge of being part of a research study all influence measured outcomes. The experimental evidence reveals that assignment effects are real and substantial: treatment and control groups respond differently depending on their assignment context, and commonly used methods such as public lotteries and computerized randomization are not neutral in their impact.

Recent discussions in economics have highlighted that the context used in experiments can significantly influence measured treatment effects and their interpretation for policy or scale-up (Deaton and Cartwright, 2018b; List, Suskind, and Al-Ubaydli, 2019; Fischer and Karlan, 2015; Bouguen, Huang, Kremer, and Miguel, 2019). My results demonstrate that different assignment mechanisms not only shape psychological and behavioral responses among treated and control subjects, but can also produce divergent estimates of treatment effects. This underlines the importance of carefully considering the research process and subjects' awareness of it when designing interventions and extrapolating findings. In line with the literature, these findings suggest that treatment effects estimated under one assignment method may not be portable to other settings unless the awareness state is emulated, or the social and psychological dimensions of assignment are explicitly accounted for. Thus, assignment mechanisms are a key determinant of both internal validity and the scalability of experimental results to new populations.

The observed heterogeneity in outcomes may arise from several underlying mechanisms, including shifts in perceived fairness, attribution of success or failure, motivated reasoning, and beliefs about luck or merit. These findings signal a need for greater attention to assignment procedures in experimental design and interpretation. Standardizing assignment practices and carefully measuring intermediate behavioral outcomes will be essential to uphold the validity and policy relevance of RCTs. More broadly, accounting for assignment bias extends the conversation beyond traditional discussion of randomization and blinding, highlighting the importance of contextual and procedural features in shaping both behavior and the credibility of causal claims.

Building on these findings, I propose that researchers must explicitly account for the awareness state of subjects when estimating causal effects: the estimand should be expressed as $\tau(\boldsymbol{\alpha}_i) = E[Y_i(1, \boldsymbol{\alpha}_i) - Y_i(0, \boldsymbol{\alpha}_i)]$. Incorporating the awareness state into the estimand formalizes the idea that participants' awareness conditions–what they know about the research process and how they interpret it–enter directly into the potential outcomes framework. To enable meaningful comparison across studies and contexts, field experiments must either standardize or document the awareness conditions under which estimates are obtained.

Moreover, my paper identifies designable dimensions along which future research should proceed, allowing researchers to characterize both the direction and intensity of assignment and awareness effects. Treatment status alone is insufficient; it is treatment awareness that must become primitive to the analysis of causal interventions. Further tests varying awareness dimensions are needed, both to understand which dimensions might confound the treatment effects, but also to understand if the awareness effects are separable from the treatment effects.

Assuming the research process is ignorable is not always valid. Research design decisions and participants' awareness of the design are, therefore, not mere procedural details. They are core determinants of both internal and external validity of causal estimates.

References

V Abhishek, Jagrati B Deshmanya, Prabhuling Tevari, GB Lokesh, MV Ravi, and K Suresh. Impact of soil health card scheme on paddy farmers' income in north-eastern karnataka. *International Journal of Current Microbiology and Applied Sciences*, 9(9):786–792, 2020.

Michael L Anderson. Multiple inference and gender differences in the effects of early intervention: A reevaluation of the abecedarian, perry preschool, and early training projects. *Journal of the American statistical Association*, 103(484):1481–1495, 2008.

Abhijit Banerjee, Esther Duflo, Nathanael Goldberg, Dean Karlan, Robert Osei, William Parienté, Jeremy Shapiro, Bram Thuysbaert, and Christopher Udry. A multifaceted program causes lasting progress for the very poor: Evidence from six countries. *Science*, 348(6236):1260799, 2015.

Abhijit V Banerjee and Esther Duflo. The experimental approach to development economics. *Annu. Rev. Econ.*, 1(1):151–178, 2009.

Christopher B Barrett and Michael R Carter. The power and pitfalls of experiments in development economics: Some non-random reflections. *Applied economic perspectives and policy*, 32(4):515–548, 2010.

Alexandre Belloni, Victor Chernozhukov, and Christian Hansen. Inference on treatment effects after selection among high-dimensional controls. *Review of Economic Studies*, 81(2):608–650, 2014.

Daniel J Benjamin. Errors in probabilistic reasoning and judgment biases. *Handbook of Behavioral Economics: Applications and Foundations 1*, 2:69–186, 2019.

Nicholas Bloom, James Liang, John Roberts, and Zhichun Jenny Ying. Does working from home work? evidence from a chinese experiment. *The Quarterly journal of economics*, 130(1):165–218, 2015.

Elias Bouacida and Renaud Foucart. Rituals of reason: Experimental evidence on the social acceptability of lotteries in allocation problems. *Games and Economic Behavior*, 152:23–36, 2025.

Adrien Bouguen, Yue Huang, Michael Kremer, and Edward Miguel. Using randomized controlled trials to estimate long-run impacts in development economics. *Annual Review of Economics*, 11 (1):523–561, 2019.

Markus K Brunnermeier and Jonathan A Parker. Optimal expectations. *American Economic Review*, 95(4):1092–1118, 2005.

A Colin Cameron, Jonah B Gelbach, and Douglas L Miller. Bootstrap-based improvements for inference with clustered errors. *The review of economics and statistics*, 90(3):414–427, 2008.

Arun G Chandrasekhar and Matthew O Jackson. Experimenting with networks. *arXiv preprint arXiv:2506.11313*, 2025.

Arun G Chandrasekhar, Cynthia Kinnan, and Horacio Larreguy. Social networks as contract enforcement: Evidence from a lab experiment in the field. *American Economic Journal: Applied Economics*, 10(4):43–78, 2018.

Sylvain Chassang, Gerard Padró i Miquel, and Erik Snowberg. Selective trials: A principal-agent approach to randomized controlled experiments. *American Economic Review*, 102(4):1279–1309, 2012.

Thomas D Cook, Donald Thomas Campbell, and Arles Day. *Quasi-experimentation: Design & analysis issues for field settings*, volume 351. Houghton Mifflin Boston, 1979.

David Roxbee Cox. Planning of experiments. 1958.

Angus Deaton and Nancy Cartwright. Reflections on randomized control trials. *Social science & medicine*, 210(C), 2018a.

Angus Deaton and Nancy Cartwright. Understanding and misunderstanding randomized controlled trials. *Social science & medicine*, 210:2–21, 2018b.

Florencia Devoto, Esther Duflo, Pascaline Dupas, William Parienté, and Vincent Pons. Happiness on tap: Piped water adoption in urban morocco. *American Economic Journal: Economic Policy*, 4(4):68–99, 2012.

Diva Dhar, Tarun Jain, and Seema Jayachandran. Reshaping adolescents' gender attitudes: Evidence from a school-based experiment in india. *American economic review*, 112(3):899–927, 2022.

Weili Ding and Steven F Lehrer. Estimating context-independent treatment effects in education experiments. *Unpublished manuscript*, 2010.

Esther Duflo, Rachel Glennerster, and Michael Kremer. Using randomization in development economics research: A toolkit. *Handbook of development economics*, 4:3895–3962, 2007.

Nisvan Erkal, Lata Gangadharan, and Boon Han Koh. By chance or by choice? biased attribution of others' outcomes when social preferences matter. *Experimental Economics*, 25(2):413–443, 2022.

Marcel Fafchamps. Causal effects in social networks. *Revue économique*, 66(4):657–686, 2015.

Greg Fischer and Dean Karlan. The catch-22 of external validity in the context of constraints to firm growth. *American Economic Review*, 105(5):295–299, 2015.

Bruno S Frey, Matthias Benz, and Alois Stutzer. Introducing procedural utility: Not only what, but also how matters. *Journal of Institutional and Theoretical Economics (JITE)/Zeitschrift für die gesamte Staatswissenschaft*, pages 377–401, 2004.

Rachel Glennerster. The complex ethics of randomized evaluations. *Running Randomized Evaluations (blog), April*, 14, 2014.

Jessica Goldberg. The effect of social pressure on expenditures in malawi. *Journal of Economic Behavior & Organization*, 143:173–185, 2017.

Sander Greenland, James J Schlesselman, and Michael H Criqui. The fallacy of employing standardized regression coefficients and correlations as measures of effect. *American journal of epidemiology*, 123(2):203–208, 1986.

Kevin Han, Guillaume Basse, and Iavor Bojinov. Population interference in panel experiments. *Journal of Econometrics*, 238(1):105565, 2024.

James J Heckman. Randomization and social policy evaluation revisited, 2020.

Daniel Kahneman and Amos Tversky. Subjective probability: A judgment of representativeness. *Cognitive psychology*, 3(3):430–454, 1972.

Stephanie Katsir, AK Biswas, Kshithij Urs, Narendra Kumar Lenka, Pramod Jha, and Kim Arora. Governing soils sustainably in india: Establishing policies and implementing strategies through local governance. *Soil Security*, 14:100132, 2024.

Harpreet Kaur, Arjun Srinivas, and Amir Bazaz. Understanding access to agrarian knowledge systems: Perspectives from rural karnataka. *Climate Services*, 21:100205, 2021.

Sarvjeet Kaur, Prabhjot Kaur, and Pankaj Kumar. Farmers' knowledge of soil health card and constraints in its use. *Indian Journal of Extension Education*, 56(1):28–32, 2020.

Andrei Khrennikov. Bell's inequality for conditional probabilities as a test for quantum-like behaviour of mind. *arXiv preprint quant-ph/0402169v1*, 2018.

Jeffrey R Kling, Jeffrey B Liebman, and Lawrence F Katz. Experimental analysis of neighborhood effects. *Econometrica*, 75(1):83–119, 2007.

Botond Kőszegi. Emotional agency. *The Quarterly Journal of Economics*, 121(1):121–155, 2006.

Rachid Laajaj and Karen Macours. Measuring skills in developing countries. *Journal of Human resources*, 56(4):1254–1295, 2021.

John A List. Why economists should conduct field experiments and 14 tips for pulling one off. *Journal of Economic perspectives*, 25(3):3–16, 2011.

John A List, Dana Suskind, and Omar Al-Ubaydli. The science of using science: Towards an understanding of the threats to scaling experiments. *University of Chicago, Becker Friedman Institute for Economics Working Paper*, (2019-73), 2019.

Markus M Möbius, Muriel Niederle, Paul Niehaus, and Tanya S Rosenblat. Managing self-confidence: Theory and experimental evidence. *Management Science*, 68(11):7793–7817, 2022.

Karthik Muralidharan and Paul Niehaus. Experimentation at scale. *Journal of Economic Perspectives*, 31(4):103–124, 2017.

Lok Niti. State of indian farmers: A report. *Delhi: Center for the Study of Developing Societies*, 2014.

Parag A Pathak and Jay Sethuraman. Lotteries in student assignment: An equivalence result. *Theoretical Economics*, 6(1):1–17, 2011.

Jörg Peters, Jörg Langbein, and Gareth Roberts. Generalization in the tropics–development policy, randomized controlled trials, and external validity. *The World Bank Research Observer*, 33(1): 34–64, 2018.

Government of India Press Information Bureau. Press release: Press information bureau, Aug 2025. URL https://www.pib.gov.in/PressNoteDetails.aspx?NoteId=155036&ModuleId=3&reg=3&lang=2.

Matthew Rabin. Inference by believers in the law of small numbers. *The Quarterly Journal of Economics*, 117(3):775–816, 2002.

K. B. Ramappa, A. V. Manjunatha, M. A. Yashashwini, S. Bangarappa, and S. Mattihalli. Impact of soil health card scheme on production, productivity and soil health in india, 2024.

KB Ramappa, Elumalai Kannan, and BT Lavanya. Adoption of recommended doses of fertilisers on soil test basis by farmers in karnataka. *Agric. Dev. Rural. Transform. Cent., Inst. Soc. Econ. Chang*, 89, 2015.

A Amarender Reddy. Impact study of soil health card scheme. *National Institute of Agricultural Extension Management (MANAGE), Hyderabad-500030*, page 106, 2018.

Donald B Rubin. Inference and missing data. *Biometrika*, 63(3):581–592, 1976.

Donald B Rubin. Bayesian inference for causal effects: The role of randomization. *The Annals of statistics*, pages 34–58, 1978.

Donald B Rubin. Randomization analysis of experimental data: The fisher randomization test comment. *Journal of the American statistical association*, 75(371):591–593, 1980.

Donald B Rubin. The bayesian bootstrap. *The annals of statistics*, pages 130–134, 1981.

Donald B Rubin. Causal inference using potential outcomes: Design, modeling, decisions. *Journal of the American statistical Association*, 100(469):322–331, 2005.

P Chandra Shekara, Ajit Kumar, N Balasubramani, Bakul C Chaudhary, Rajeev Sharma, Chitra Shukla, and Mr Max Baumann. Farmer's handbook on basic agriculture. 2016.

Ashish Shenoy and Travis J. Lybbert. Implementer desirability bias in program evaluation. *Available at SSRN*, https://papers.ssrn.com/sol3/papers.cfm?abstract$_i d = 4764716, 2024$.

N Sumanth, P Sanjay, KR Kaveri, SS Moitreyee, S Sree, and S Raj. Agricultural extension and support systems in india: An agricultural innovation systems (ais) perspective (karnataka, maharashtra and west bengal states of india). Technical report, Discussion Paper 20. MANAGE-Centre for Agricultural Extension Innovations ..., 2020.

Thomas W Valente. Network interventions. *science*, 337(6090):49–53, 2012.

Tyler J VanderWeele and Weihua An. Social networks and causal inference. *Handbook of causal analysis for social research*, pages 353–374, 2013.

Tyler J VanderWeele and Miguel A Hernan. Causal inference under multiple versions of treatment. *Journal of causal inference*, 1(1):1–20, 2013.

Florian Zimmermann. The dynamics of motivated beliefs. *American Economic Review*, 110(2):337–363, 2020.

Appendix A Decomposition of Potential Outcomes

	t	a	k	s	**Interpretation**
Y(0,0,0,0)	✗	✗	✗	✗	Observational data, no intervention; impossible to achieve in an RCT
Y(0,0,1,1)	✗	✗	✓	✓	Control group or treated before distribution, with assignment method clear but unaware of treatment statuses
Y(1,0,0,1)	✓	✗	✗	✓	Directed survey after government intervention; or treated subject does not understand assignment rule, nor is aware of others' assignment status
Y(1,0,1,1)	✓	✗	✓	✓	Treated subject understands assignment rule but is unaware of others' assignment status
Y(0,0,0,1)	✗	✗	✗	✓	Baseline in an RCT, before assignment step
Y(0,1,0,1)	✗	✓	✗	✓	Subjects after assignment but before distribution, aware of assignment statuses but not aware of assignment method
Y(0,1,1,1)	✗	✓	✓	✓	Subjects after assignment but before distribution, aware of assignment method and assignment statuses
Y(1,1,0,1)	✓	✓	✗	✓	Treated after treatment distribution, not aware of assignment method
Y(1,1,1,1)	✓	✓	✓	✓	Treated after treatment distribution, aware of assignment method and assignment statuses
Y(0,0,1,0)	✗	✗	✓	✗	Observational data, after a government explains an assignment rule but has not yet implemented the intervention; impossible and unethical in an RCT
Y(1,0,0,0)	✓	✗	✗	✗	Observational data, generalized intervention; unethical in an RCT
Y(1,0,1,0)	✓	✗	✓	✗	Observational data, transparent rule settled by a government; unethical in an RCT

Table A1: Table 1 of 2. Separation of a and t is sometimes not done in an RCT: the treatment is distributed the moment participants know about their assignment. In red, I signal the treatment and awareness states that are unethical in an RCT. By this, I mean that a researcher should not be allowed to perform surveys around an intervention without giving information to the participants that they are in a research study; IRB requests $s = 1$. In blue, I signal the treatment and awareness states combinations that I include in my experimental design.

	t	a	k	s	**Interpretation**
Y(0,1,0,0)	✗	✓	✗	✗	Unethical in an RCT, but in practice happens: subjects are not informed of their study participation and assignment rule
Y(0,1,1,0)	✗	✓	✓	✗	Possible through an intervention run by a government. Unethical in an RCT, but in practice happens: subjects are not informed of their study participation
Y(1,1,0,0)	✓	✓	✗	✗	Possible through an intervention run by a government. Unethical in an RCT, but in practice happens: subjects are not informed of their study participation and assignment rule
Y(1,1,1,0)	✓	✓	✓	✗	Unethical in an RCT, but in practice happens: subjects are not informed of their study participation

Table A2: Table 2 of 2. Separation of a and t is sometimes not done in an RCT: the treatment is distributed the moment participants know about their assignment. In red, I signal the treatment and awareness states that are unethical in an RCT. By this, I mean that a researcher should not be allowed to perform surveys around an intervention without giving information to the participants that they are in a research study; IRB requests $s = 1$. In teal, I signal the treatment and awareness states combinations that I include in my experimental design.

Appendix B Proof of Propositions 1, 2, and 3

Proposition 1. The three differences can be decomposed into:

$$
\begin{aligned}
T_1 &= \mathbb{E}[Y_i(t=1,k=0) - Y_i(t=0,k=0)] = \mathbb{E}[\Delta f_i + h_i(1,0) - h_i(0,0)] \\
T_2 &= \mathbb{E}[Y_i(t=1,k=1) - Y_i(t=0,k=1)] = \mathbb{E}[\Delta f_i + h_i(1,1) - h_i(0,1)] \\
T_3 &= \mathbb{E}[Y_i(t=1,k=0) - Y_i(t=0,k=1)] = \mathbb{E}[\Delta f_i - \Delta g_i + h_i(1,0) - h_i(0,1)]
\end{aligned}
$$

The difference $T_1 + T_2 - T_3$ reduces to:

$$T_1 + T_2 - T_3 = \mathbb{E}[\Delta f_i + \Delta g_i + h_i(1,1) - h_i(0,0)]$$

Under nonnegative treatment effects ($\Delta f \geq 0$), monotonicity in k ($\Delta g \geq 0$), and additive separability ($h(t, \boldsymbol{\alpha}) = 0 \ \forall t, \boldsymbol{\alpha}$):

$$T_1 + T_2 - T_3 = \mathbb{E}[\Delta f_i + \Delta g_i] \geq 0$$

□

Proposition 2. The three differences can be decomposed into:

$$T_1 = \mathbb{E}[Y_i(t=1,k=0) - Y_i(t=0,k=0)] = \mathbb{E}[\Delta f_i + h_i(1,0) - h_i(0,0)]$$
$$T_2 = \mathbb{E}[Y_i(t=1,k=1) - Y_i(t=0,k=1)] = \mathbb{E}[\Delta f_i + h_i(1,1) - h_i(0,1)]$$
$$T_3^* = \mathbb{E}[Y_i(t=1,k=1) - Y_i(t=0,k=0)] = \mathbb{E}[\Delta f_i + \Delta g_i + h_i(1,1) - h_i(0,0)]$$

The difference $T_1 + T_2 - T_3^*$ reduces to:

$$T_1 + T_2 - T_3^* = \mathbb{E}[\Delta f_i - \Delta g_i + h_i(1,0) - h_i(0,1)]$$

Under $\Delta f \geq \Delta g$ and additive separability $(h(t,\boldsymbol{\alpha}) = 0 \ \forall t, \boldsymbol{\alpha})$:

$$T_1 + T_2 - T_3^* = \mathbb{E}[\Delta f_i - \Delta g_i] \geq 0$$

□

Proposition 3. Decomposing ψ:

$$\psi = (\mathbb{E}[Y_i(1,1)] - \mathbb{E}[Y_i(1,0)]) - (\mathbb{E}[Y_i(0,1)] - \mathbb{E}[Y_i(0,0)])$$
$$\psi = (\mathbb{E}[g_i(1) + h_i(1,1)] - \mathbb{E}[g_i(0) + h_i(1,0)]) - (\mathbb{E}[g_i(1) + h_i(0,1)] - \mathbb{E}[g_i(0) + h_i(0,0)])$$
$$\psi = \mathbb{E}[h_i(1,1)] - \mathbb{E}[h_i(1,0)] - \mathbb{E}[h_i(0,1)] + \mathbb{E}[h_i(0,0)]$$

If Y_i is separable, then $h(t,\boldsymbol{\alpha}) = 0 \ \forall t, \boldsymbol{\alpha}$. Substituting this into the expression for ψ:

$$\psi = \mathbb{E}[0] - \mathbb{E}[0] - \mathbb{E}[0] + \mathbb{E}[0] = 0$$

If $h(t,k)$ is supermodular, then $\psi \geq 0$ by definition of supermodularity. □

Appendix C Balance and Attrition Tables

		(1) Control		(2) Treated	(1)-(2) Pairwise t-test
	N	Mean/(SE)	N	Mean/(SE)	Mean difference
Age	1003	42.422 (0.383)	663	42.220 (0.479)	0.202
Gender	1003	0.031 (0.005)	663	0.038 (0.007)	-0.007
Level of educ.	992	4.379 (0.169)	660	5.055 (0.206)	-0.676**
Number of hh members	1003	5.696 (0.101)	663	5.668 (0.125)	0.028
Number of children in household	522	2.425 (0.081)	314	2.350 (0.090)	0.075
HH community participation	1003	0.276 (0.014)	663	0.291 (0.018)	-0.015
HH income, categories	976	5.597 (0.085)	649	5.963 (0.101)	-0.366***
Land registered (RTC card)	1003	0.790 (0.013)	663	0.778 (0.016)	0.011
Above Poverty Line card	1003	0.095 (0.009)	663	0.081 (0.011)	0.013
Below Poverty Line card	1003	0.806 (0.013)	663	0.819 (0.015)	-0.013
AAY card (subsidy for food)	1003	0.077 (0.008)	663	0.089 (0.011)	-0.012
Owns livestock	1003	0.730 (0.014)	663	0.736 (0.017)	-0.006
Religion, Hindu	1003	0.921 (0.009)	663	0.940 (0.009)	-0.018
Religion, Muslim	1003	0.070 (0.008)	663	0.056 (0.009)	0.014
Farming experience (years)	1003	16.363 (0.167)	662	16.687 (0.203)	-0.324
Age started farming	1003	23.028 (0.211)	662	22.994 (0.285)	0.034
Works as laborer	1003	0.425 (0.016)	662	0.458 (0.019)	-0.033

Table A3: Balance table between farmers assigned to treatment (in the PL, CR, or Pure Treatment arms) and farmers assigned to control (in the PL, CR, or Pure Control arms).

		(1) Control		(2) Treated	(1)-(2) Pairwise t-test
	N	Mean/(SE)	N	Mean/(SE)	Mean difference
Asked for an agri. loan	1003	0.480 (0.016)	663	0.495 (0.019)	-0.015
Number of plots	1003	1.682 (0.039)	662	1.831 (0.083)	-0.149*
Owns the plot	997	0.724 (0.014)	660	0.680 (0.018)	0.044*
Expected input costs	992	82119.677 (5831.071)	660	82165.795 (4801.646)	-46.118
Has done a soil test before	995	0.153 (0.011)	656	0.159 (0.014)	-0.006

Table A4: (Continuation) Balance table between farmers assigned to treatment (in the PL, CR, or Pure Treatment arms) and farmers assigned to control (in the PL, CR, or Pure Control arms).

	(1) PL-CR	(2) PL-PC	(3) PL-PT	(4) CR-PC	(5) CR-PT	(6) PC-PT
Age	1.548**	2.012**	1.560*	0.463	0.012	-0.451
Gender	-0.038***	-0.016	0.006	0.023	0.044***	0.021
Level of educ.	0.751**	1.590***	-0.564	0.838**	-1.315***	-2.154***
Number of HH members	-0.550***	-0.339	-0.198	0.212	0.352	0.141
Number of children in HH	-0.351***	-0.402**	0.156	-0.051	0.507***	0.558**
HH community participation	-0.006	0.017	-0.021	0.023	-0.015	-0.038
HH income, categories	0.347**	0.743***	-0.053	0.397*	-0.400*	-0.796***
Land registered (RTC card)	-0.109***	-0.079**	-0.029	0.030	0.080***	0.050
Above Poverty Line card	-0.025	0.025	0.033	0.050**	0.058**	0.008
Below Poverty Line card	0.073***	0.017	-0.037	-0.057*	-0.110***	-0.053
AAY card (food subsidy)	-0.027*	-0.026	-0.010	0.000	0.016	0.016
Owns livestock	0.021	-0.031	0.002	-0.052	-0.019	0.033
Religion, Hindu	-0.002	0.011	-0.022	0.013	-0.020	-0.033
Religion, Muslim	0.002	-0.021	0.012	-0.023	0.010	0.033
Farming experience (years)	1.007***	0.838**	-0.411	-0.168	-1.418***	-1.250**
Age started farming	1.052**	0.931*	0.485	-0.121	-0.567	-0.446
Works as laborer	-0.010	-0.083**	-0.098**	-0.072*	-0.087**	-0.015
Asked for an agri. loan	0.039	0.016	0.070*	-0.023	0.030	0.054
Number of plots	0.149	0.213	0.141	0.064	-0.008	-0.073
Owns the plot	-0.145***	-0.159***	-0.006	-0.015	0.138***	0.153***
Expected input costs	10460.870	4634.588	-748.396	-5826.281	-1.12e+04	-5382.985
Has done a soil test before	-0.016	-0.021	-0.006	-0.005	0.010	0.015

Table A5: This table shows the pairwise mean comparison between the different arms, where PL stands for Public Lottery; CR for Computerized Randomization; PT for Pure Treatment; and PC for Pure Control. For the number of observations, mean, and standard errors at each group level, refer to A8.

	(1) Responded survey 2	(2) Left for survey 2	(3) Difference
Age	42.764	40.830	-1.935***
	(12.221)	(12.041)	(0.722)
Gender	0.035	0.030	-0.004
	(0.183)	(0.171)	(0.011)
Level of educ.	4.565	4.948	0.383
	(5.320)	(5.280)	(0.316)
Number of hh members	5.762	5.409	-0.353*
	(3.326)	(2.692)	(0.190)
Number of children in household	2.425	2.294	-0.131
	(1.821)	(1.486)	(0.149)
HH community participation	0.296	0.234	-0.062**
	(0.457)	(0.424)	(0.027)
HH income, categories	5.735	5.772	0.036
	(2.637)	(2.603)	(0.158)
Land registered (RTC card)	0.802	0.725	-0.077***
	(0.399)	(0.447)	(0.024)
Above Poverty Line card	0.093	0.077	-0.016
	(0.290)	(0.267)	(0.017)
Below Poverty Line card	0.810	0.813	0.003
	(0.392)	(0.390)	(0.023)
AAY card (subsidy for food)	0.084	0.071	-0.013
	(0.278)	(0.258)	(0.016)
Owns livestock	0.745	0.687	-0.058**
	(0.436)	(0.464)	(0.026)
Religion, Hindu	0.929	0.926	-0.004
	(0.256)	(0.262)	(0.015)
Religion, Muslim	0.065	0.060	-0.005
	(0.247)	(0.239)	(0.015)
Farming experience (years)	16.479	16.538	0.060
	(5.177)	(5.568)	(0.312)
Age started farming	22.992	23.096	0.105
	(7.101)	(6.359)	(0.412)
Works as laborer	0.431	0.462	0.030
	(0.495)	(0.499)	(0.029)
Asked for an agri. loan	0.500	0.434	-0.066**
	(0.500)	(0.496)	(0.030)
Number of plots	1.769	1.643	-0.126
	(1.775)	(1.108)	(0.098)
Owns the plot	0.729	0.627	-0.102***
	(0.445)	(0.484)	(0.027)
Expected input costs	79,994.703	89,913.164	9,918.459
	(167807.312)	(140104.859)	(9,697.726)
Has done a soil test before	0.159	0.140	-0.019
	(0.366)	(0.348)	(0.022)
Observations	1,302	364	1,666

Table A6: Balance table between farmers who continued their participation in the intervention from survey 1 to survey 2 vs those who left the intervention.

	(1)	(2)	(3)
	Responded survey 3	Left for survey 3	Difference
Age	42.659	41.481	-1.178*
	(12.241)	(12.075)	(0.673)
Gender	0.032	0.038	0.006
	(0.176)	(0.191)	(0.010)
Level of educ.	4.685	4.552	-0.133
	(5.275)	(5.415)	(0.294)
Number of hh members	5.752	5.503	-0.249
	(3.167)	(3.288)	(0.177)
Number of children in household	2.372	2.473	0.101
	(1.640)	(2.069)	(0.141)
HH community participation	0.298	0.238	-0.060**
	(0.458)	(0.427)	(0.025)
HH income, categories	5.689	5.890	0.201
	(2.592)	(2.722)	(0.147)
Land registered (RTC card)	0.800	0.744	-0.056**
	(0.400)	(0.437)	(0.023)
Above Poverty Line card	0.089	0.091	0.003
	(0.284)	(0.288)	(0.016)
Below Poverty Line card	0.811	0.811	-0.000
	(0.392)	(0.392)	(0.022)
AAY card (subsidy for food)	0.083	0.078	-0.005
	(0.276)	(0.268)	(0.015)
Owns livestock	0.751	0.682	-0.070***
	(0.433)	(0.466)	(0.024)
Religion, Hindu	0.928	0.931	0.003
	(0.259)	(0.254)	(0.014)
Religion, Muslim	0.066	0.060	-0.006
	(0.248)	(0.238)	(0.014)
Farming experience (years)	16.558	16.312	-0.247
	(5.313)	(5.127)	(0.291)
Age started farming	23.177	22.575	-0.602
	(7.123)	(6.421)	(0.383)
Works as laborer	0.438	0.437	-0.002
	(0.496)	(0.497)	(0.027)
Asked for an agri. loan	0.507	0.428	-0.079***
	(0.500)	(0.495)	(0.028)
Number of plots	1.749	1.719	-0.030
	(1.800)	(1.167)	(0.091)
Owns the plot	0.726	0.654	-0.072***
	(0.446)	(0.476)	(0.025)
Expected input costs	78,552.211	91,924.438	13,372.228
	(168186.594)	(144471.016)	(9,006.822)
Has done a soil test before	0.168	0.119	-0.049**
	(0.374)	(0.324)	(0.020)
Observations	1,217	449	1,666

Table A7: Balance table between farmers who continued their participation in the intervention from survey 1 to survey 3 vs those who left the intervention.

	PL		CR		PControl		PTreat	
	N	Mean/(SE)	N	M./(SE)	N	M./(SE)	N	M./(SE)
Age	598	43.401 (0.490)	599	41.853 (0.503)	236	41.390 (0.819)	233	41.841 (0.785)
Gender	598	0.018 (0.005)	599	0.057 (0.009)	236	0.034 (0.012)	233	0.013 (0.007)
Level of educ.	594	5.064 (0.213)	595	4.313 (0.224)	232	3.474 (0.330)	231	5.628 (0.347)
Number of hh members	598	5.411 (0.119)	599	5.962 (0.144)	236	5.750 (0.193)	233	5.609 (0.212)
Number of children In hh	273	2.220 (0.082)	326	2.571 (0.103)	127	2.622 (0.215)	110	2.064 (0.123)
HH community participation	598	0.279 (0.018)	599	0.285 (0.018)	236	0.263 (0.029)	233	0.300 (0.030)
HH income, categories	587	5.964 (0.105)	575	5.617 (0.112)	231	5.221 (0.175)	232	6.017 (0.171)
Land registered (RTC card)	598	0.731 (0.018)	599	0.840 (0.015)	236	0.809 (0.026)	233	0.760 (0.028)
Above Poverty Line card	598	0.089 (0.012)	599	0.114 (0.013)	236	0.064 (0.016)	233	0.056 (0.015)
Below Poverty Line card	598	0.834 (0.015)	599	0.761 (0.017)	236	0.818 (0.025)	233	0.871 (0.022)
AAY card (subsidy for food)	598	0.067 (0.010)	599	0.093 (0.012)	236	0.093 (0.019)	233	0.077 (0.018)
Owns livestock	598	0.736 (0.018)	599	0.715 (0.018)	236	0.767 (0.028)	233	0.734 (0.029)
Religion, Hindu	598	0.926 (0.011)	599	0.928 (0.011)	236	0.915 (0.018)	233	0.948 (0.015)
Religion, Muslim	598	0.064 (0.010)	599	0.062 (0.010)	236	0.085 (0.018)	233	0.052 (0.015)
Farming experience (years)	598	16.915 (0.211)	598	15.908 (0.217)	236	16.076 (0.356)	233	17.326 (0.331)
Age started farming	598	23.592 (0.265)	598	22.540 (0.315)	236	22.661 (0.452)	233	23.107 (0.386)
Works as laborer	597	0.409 (0.020)	599	0.419 (0.020)	236	0.492 (0.033)	233	0.506 (0.033)
Asked for an agri.loan	598	0.512 (0.020)	599	0.472 (0.020)	236	0.496 (0.033)	233	0.442 (0.033)
Number of plots	598	1.844 (0.098)	598	1.696 (0.040)	236	1.631 (0.080)	233	1.704 (0.062)
Owns the plot	597	0.631 (0.020)	594	0.776 (0.017)	234	0.791 (0.027)	232	0.638 (0.032)
Expected input costs	592	86450.828 (4626.812)	594	75989.958 (6457.563)	234	81816.239 (17878.794)	232	87199.224 (8429.295)
Has done a soil test before	592	0.145 (0.014)	594	0.162 (0.015)	234	0.167 (0.024)	231	0.152 (0.024)

Table A8: This table shows the number of observations, mean, and standard errors at each arm level, where PL stands for Public Lottery; CR for Computerized Randomization; PT for Pure Treatment; and PC for Pure Control.

Appendix D Consent scripts for each arm, survey 1

Consent script for the Pure Control arm, survey 1:

> *My name is [NAME]. I am part of a team from IFMR- Institute of Financial Management & Research. IFMR is a private, non-governmental organization that works to understand and improve people's lives. We will be conducting surveys in this village over the next 3 months.*
>
> *You are invited to participate in a* ***research study*** *about the adoption of different crops and fertilizers and on agricultural knowledge. You will be asked to participate in a short interview about the agricultural methods you use and your beliefs and knowledge of agricultural conditions in your village, including your knowledge and perceptions of agricultural services.*
>
> *With your permission, the interview will be audio-taped. This way we can conduct the survey faster, transcribing later your responses in detail instead of writing them down now. The audio will be destroyed after we transcribe your responses.*
>
> *Your participation will take approximately 40 minutes.*
>
> *There are minimal risks associated with this study. The benefits that may reasonably be expected to result from this study are getting knowledge to make more informed agricultural decisions. We will come back to the village another 2 times to ask for other much shorter surveys (only 20 minutes) during the next 3 months. During each survey, you will get the chance to win between Rs. 50-160 as a UPI reward or phone recharge.*
>
> *If you want to participate, I will start the survey now and we will visit you two more times to ask follow-up questions about the same topics.*
>
> *Please understand your participation is voluntary and you have the right to withdraw your consent or discontinue participation at any time. You have the right to refuse to answer particular questions. Your privacy and confidentiality of the information you provide will be maintained in all published and written data resulting from the study. A summary of all the interviews done, without your name or the name of your village, or the name of your neighbors, will be shared with other researchers. They will not be able to identify anyone that responds the interviews, nor your plot or village.*
>
> *If you have any questions, concerns, or complaints about this research or your rights as a participant, you can reach out to [J-PAL Project Associate].*

Consent script for the Pure Treatment, Public Lottery, and Computerized Randomization arms, survey 1:

> *My name is [NAME]. I am part of a team from IFMR- Institute of Financial Management & Research. IFMR is a private, non-governmental organization that works to understand and improve people's lives. We will be conducting surveys in this village over the next 3 months.*

You are invited to participate in a ***research study*** *about the adoption of different crops and fertilizers and on agricultural knowledge. You will be asked to participate in a short interview about the agricultural methods you use and your beliefs and knowledge of agricultural conditions in your village, including your knowledge and perceptions of agricultural services.*

With your permission, the interview will be audio-taped. This way we can conduct the survey faster, transcribing later your responses in detail instead of writing them down now. The audio will be destroyed after we transcribe your responses.

Your participation will take approximately 50 minutes.

There are minimal risks associated with this study. The benefits that may reasonably be expected to result from this study are getting knowledge to make more informed agricultural decisions. We will come back to the village another 2 times to ask for other much shorter surveys (only 20 minutes) during the next 3 months. Moreover, during each survey, you will get the chance to win between Rs. 50-260 as a UPI reward or phone recharge.

Besides, you will be eligible to be selected to receive a test on the soil quality of one of your plots and a voucher to buy items from a local input dealer.

If you want to participate, I will start the survey now and we will visit you two more times to ask follow-up questions about the same topics.

Please understand your participation is voluntary and you have the right to withdraw your consent or discontinue participation at any time. You have the right to refuse to answer particular questions. Your privacy and confidentiality of the information you provide will be maintained in all published and written data resulting from the study. A summary of all the interviews done, without your name or the name of your village, or the name of your neighbors, will be shared with other researchers. They will not be able to identify anyone that responds the interviews, nor your plot or village.

If you have any questions, concerns, or complaints about this research or your rights as a participant, you can reach out to [J-PAL Project Associate].

There are no differences in the consent scripts for the other two survey rounds.

Appendix E Assignment stage, explanations by arm

Public Lottery arm:

Thank you for coming. My name is [NAME]. I am part of a team from IFMR-Institute of Financial Management & Research. IFMR is a private, non-governmental organization that works to understand and improve people's lives. You participated in an interview some days ago, making you eligible to get a soil test of your plot and a coupon for a local input dealer together with the results of the soil test. However, because of budget restrictions, we will not be able to provide the soil test and the coupon for everyone that is participating in *your village. Keep in mind we are not only reaching this village, but other 110 villages in the area to know the impacts of better information on the soil quality on farming practices.*

We will do a Public Lottery to make the selection of who will receive the soil test and a coupon for a local input dealer as transparent as possible. Of [NUMBER FARMERS] farmers who participated in our first interview, we can only give the benefits to [TREATED FARMERS]. In this bag I will then put [TREATED FARMERS] green balls, and [CONTROL FARMERS] red balls. If the ball you grab is green, you are selected to receive the soil test and a coupon for a local input dealer. If you grab a red ball, then you will not receive the soil test and the coupon. Please, note that who is selected is random, it does not depend on the farmer's characteristics. I will call each of you here and you will choose a ball from the bag. We order your names alphabetically, so we will follow that order for calling you.

Whether you were selected or not, you will still receive information on how and where you can do a soil test.

We will film and/or take pictures of this process to make sure the process is fair. The recording and pictures will be checked by researchers today after the lottery is done and then eliminated. Please, if you have any questions, remember you can contact us.

Computerized Randomization arm:

Thank you for coming. My name is [NAME]. I am part of a team from IFMR-Institute of Financial Management & Research. IFMR is a private, non-governmental organization that works to understand and improve people's lives. You participated in an interview some days ago, making you eligible to get a soil test of your plot and a coupon for a local input dealer together with the results of the soil test. However, because of budget restrictions, we will not be able to provide the soil test and the coupon for everyone who is participating in your village. Keep in mind we are not only reaching this village, but other 110 villages in the area to know the impacts of better information on the soil quality on farming practices.

We held a lottery in our office to make the selection of who will receive the soil test and a coupon for a local input dealer as transparent as possible. The lottery was done using a computer program that randomly selects farmers. Of [NUMBER FARMERS] farmers that participated in our first interview, we can only give the benefits to [TREATED FARMERS], so the software randomly chose [TREATED FARMERS] farmers. I will proceed to read the names of the selected farmers in a moment.

Whether you were selected or not, you will still receive information on how and where you can do a soil test.

We will film the reading of the names of the selected farmers to be sure the process is fair. The recording will be checked by researchers today and then destroyed.

Please, if you have any questions, remember you can contact us using this contact information we will be distributing.

Pure Treatment arm:

Thank you for coming. My name is [NAME]. I am part of a team from IFMR-Institute of Financial Management & Research. IFMR is a private, non-governmental organization that works to understand and improve people's lives. You participated in an interview some days ago, making you eligible to get a soil test of your plot and a coupon for a local input dealer together with the results of the soil test. We will be able to provide the soil test and the coupon for the [NUMBER FARMERS] farmers who participated and completed the survey we did some days ago in your village. We are not only reaching this village, but other 110 villages in the area to know the impacts of better information on the soil quality on farming practices. We will film the reading of the names of the participating farmers to be sure the process is fair. The recording will be checked by researchers today and then destroyed. Please, if you have any questions, remember you can contact us using this contact information we will be distributing.

Pure Control arm (no assignment, but visit to keep the number of visits per village Cons.):

Thank you for coming. My name is [NAME]. I am part of a team from IFMR-Institute of Financial Management & Research. IFMR is a private, non-governmental organization that works to understand and improve people's lives. You participated in an interview some days ago. We are not only reaching this village, but other 110 villages in the area to know the impacts of better information on the soil quality on farming practices. You will receive information on how and where you can do a soil test. Please, if you have any questions, remember you can contact us using this contact information we will be distributing.

Appendix F Description of variables used for the creation of the indices

Variable	Description
soil_test_value	How useful would a soil quality test of your plot be? 1 = Not at all useful, 2 = Probably not useful, 3 = Probably useful, 4 = Definitely useful.
wtp_max	Imagine you could bring a small soil sample to be tested and receive information about soil quality and recommendations to improve it. What is the maximum amount you would be willing to pay for such a test? (Rupees)
farming_ knowledge	Number of farming knowledge questions answered correctly (count).
generic_poor_ farmer	If a generic farming plot has very poor soil quality, how much is the farmer to blame compared to the weather and the soil? 0–10, higher = farmer more to blame.
generic_good_ farmer	If a generic farming plot has very good soil quality, how much is the farmer responsible for the good result compared to the weather and the soil? 0–10, higher = farmer more responsible.
own_poor_ farmer	If your plot has very poor soil quality, how much are you to blame compared to the weather and the soil? 0–10, higher = you more to blame.
own_good_ farmer	If your plot has very good soil quality, how much are you responsible for the good result compared to the weather and the soil? 0–10, higher = you more responsible.
hard_work_ tokens	When you get what you want, is it mostly due to hard work or luck? 0–10, higher = more due to hard work.
yourself_ tokens	On a scale from 0 to 10, where 0 means "completely determined by other people (family/friends/people in power)" and 10 means "completely determined by me," where would you place yourself?
religiousity_ worship	In the past two weeks, how many times did you visit a place of worship for worship only (not for social or other activities)?
religiousity_ pray	On how many of the past 14 days did you offer thanks or prayers to God/your deity?
agri_loan_ past	Have you taken an agricultural loan during the past month?
agri_loan_ internal	Did you get an agricultural loan from someone in your village? (Microfinance group, landlord, village moneylender, friend/relative)
agri_loan_ external	Did you get an agricultural loan from someone outside your village or an institution outside your village? (bank, agricultural society, another source)
agri_loan_ future	Do you plan to take an agricultural loan for Kharif 2025 (June–October)? Yes/No.
purposes_ future_loan	Planned purposes for future agricultural loan (count).

Variable	Description
insurance	Do you plan to buy crop insurance for next season (from June 2025 onwards)?
seed_ extension	Did you obtain new seeds from the agricultural extension service? Yes/No.
seed_board	Did you obtain new seeds from an agricultural board? Yes/No.
fert_organic	Did you use organic fertilizer this season? Yes/No. If yes, list the types.
fert_chemical	Did you use inorganic fertilizer this season? Yes/No. If yes, list the types.
avg_ confidence	Overall, how confident do you feel when answering farming-related questions? 1 = Not at all confident, 5 = Very confident.
life_ satisfaction	How would you classify your satisfaction with your life right now? Select a number from 1 to 10, where 1 is not satisfied at all and 10 is completely satisfied.
feeling_ cheerful	In the last 7 days, how often did you feel happy and optimistic? 1 = Never, 6 = All the time.
feeling_ worried	In the last 7 days, how often did you feel worried about the future? 1 = All the time, 6 = Never.
feeling_ tired	In the last 7 days, how often did you feel low energy or enthusiasm? 1 = All the time, 6 = Never.
beh_measure ment_eg	Risk preference in coin-toss lottery (1–5; higher = more risk-loving). Enumerator: read the five predefined options (e.g., Option 1: Rs 40 if heads, Rs 40 if tails; . . . ; Option 5: Rs 110 if heads, Rs 0 if tails). Record chosen option.
nitro_bet	Think of the soil in your village. First, state how many kilograms of nitrogen there are in one acre of soil. Then choose one payoff option: Option 1: Win Rs 20 if your answer is within 50 units of the correct value, otherwise Rs 10; . . . ; Option 4: Win Rs 50 if within 10 units, otherwise Rs 0.
weather_bet	Predict total rainfall (mm) in Kalaburagi/Gulbarga next week. Choose one payoff option: Option 1: Win Rs 10 if within 10 mm; . . . ; Option 4: Win Rs 70 if within X mm, otherwise Rs 0.
time_ discount	Choose between Rs 100 today and Rs X in one month across 5 rows where X increases each row. Record the row where choice switches
risk_lottery	Choose between a 50–50 lottery (Rs 200 if tails, Rs 0 if heads) and a sure amount. There are 4 rows; the sure amount increases each row. Record switching point.
links_advice	Think about those farmers in the village you will ask for advice on farming (which seeds or fertilizer to use, where to buy inputs, how to use a machine or apply a pesticide). The farmer must be someone you consider knowledgeable, and that you feel comfortable asking for advice. How many farmers in this village did you ask for farming advice in the last two weeks?
advice_ farming	Think about those farmers in the village that ask for YOUR advice on farming (which seeds or fertilizer to use, where to buy inputs, how to use a machine or apply a pesticide). Without naming them, how many farmers asked you for advice on farming in the last 2 weeks?

Variable	Description
best_farmer1_ talk	In the last two weeks, did you speak with the farmer you consider the most knowledgeable in your village about farming? Yes/No.
best_farmer2_ talk	In the last two weeks, did you speak with the farmer you consider the second most knowledgeable in your village about farming? Yes/No.
best_farmer3_ talk	In the last two weeks, did you speak with the farmer you consider the third most knowledgeable in your village about farming? Yes/No.
past_agriloan_ invest	Did you use the past agricultural loan to invest in machinery or to buy/rent land? Yes/No
purposes_past_ loan	Planned purposes for past agricultural loan (count).
expected_revenue kharif	How much do you expect to earn with the produce of your plot in the Kharif season, in rupees?

Appendix G Other weighting methods: first principal component and inverse covariance weighting

Although equal weighting is standard in the literature due to its transparency, it does not account for the potential high correlation between some variables within each index. For example, if an index of three variables contains two that are highly positively correlated, the equal-weighted index will overweight the underlying component that is behind these two variables. Moreover, suppose the treatment or the arms have a substantial effect on this underlying component. In that case, the index will show a significant effect even when the third variable, which measures another underlying component, remains unchanged.

To address this issue, I present two different weighting methods, which result in varying levels of importance for each variable within an index, as shown in Figure A1. First, I run the first principal component on the component variables of each index. The first principal component is the linear combination of all the variables that explains the maximum possible variance in the data. It is, in a nutshell, the most prominent underlying dimension in the set of variables. Then, the weights assigned to each variable using this method are not equal; they are derived from the data to maximize variance. Results from estimating Equations 5 and 6 are presented in Tables A9, A10, A11, and A12.

	Info Adopt.	**Practices Inputs**	**Financial Manag.**	**Risk Aversion**	**Psych. Well-b**	**Attrib. LOC**	**Social Networks**
Treat (γ)	-0.060	-0.111	-0.011	-0.115	0.118	-0.098^{**}	-0.070
	(0.064)	(0.108)	(0.096)	(0.086)	(0.087)	(0.037)	(0.055)
PL (δ)	-0.113	-0.246^{**}	-0.142^{**}	-0.172^{**}	0.143	-0.146^{**}	0.008
	(0.080)	(0.089)	(0.053)	(0.081)	(0.104)	(0.059)	(0.071)
Treat $\times$ PL (β)	0.120	-0.061	0.140	0.140	0.097	0.072	-0.154
	(0.089)	(0.131)	(0.124)	(0.125)	(0.134)	(0.076)	(0.126)
Constant (α)	-0.280^{***}	-1.708^{***}	-0.371^{***}	0.192^{***}	-0.246^{***}	0.143^{***}	0.195^{***}
	(0.051)	(0.086)	(0.036)	(0.066)	(0.069)	(0.041)	(0.055)
Observations	701	701	701	701	701	701	701

Table A9: Survey 2 regression results as indicated in Equation 5 for summary indices, giving more weight to indicators that contribute more to the overall variation in the data. The baseline corresponds to the Computerized Randomization. PL stands for a dummy variable for the Public Lottery arm, and Treat is a dummy variable indicating whether the subject was assigned to treatment. The coefficients for Treat $\times$ PL allow to test for Hypothesis 1a directly. I control for the lagged variable. The bootstrapped standard errors are clustered at the geographical cluster level. $^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.1$.

	Info Adopt.	**Practices Inputs**	**Financial Manag.**	**Risk Aversion**	**Psych. Well-b**	**Attrib. LOC**	**Social Networks**
Treat (γ)	0.035	-0.075	-0.047	-0.021	0.156	-0.079	-0.193^{*}
	(0.059)	(0.053)	(0.065)	(0.110)	(0.150)	(0.067)	(0.097)
PL (δ)	0.135^{*}	-0.083	0.032	-0.021	0.279^{*}	0.000	0.105
	(0.069)	(0.101)	(0.062)	(0.127)	(0.152)	(0.066)	(0.065)
Treat $\times$ PL (β)	-0.154^{*}	-0.049	0.055	-0.090	-0.142	0.034	-0.012
	(0.079)	(0.081)	(0.092)	(0.150)	(0.193)	(0.104)	(0.134)
Constant (α)	-1.327^{***}	-1.745^{***}	-0.288^{***}	0.048	-0.224	0.056^{*}	0.167^{***}
	(0.039)	(0.087)	(0.050)	(0.100)	(0.132)	(0.030)	(0.036)
Observations	701	701	701	701	701	701	701

Table A10: Survey 3 regression results as indicated in Equation 5 for summary indices, giving more weight to indicators that contribute more to the overall variation in the data. The baseline corresponds to the Computerized Randomization. PL stands for a dummy variable for the Public Lottery arm, and Treat is a dummy variable indicating whether the subject was assigned to treatment. The coefficients for Treat $\times$ PL allow to test for Hypothesis 1a directly. I control for the lagged variable. The bootstrapped standard errors are clustered at the geographical cluster level. $^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.1$.

	Info Adopt.	Practices Inputs	Financial Manag.	Risk Aversion	Psych. Well-b	Attrib. LOC	Social Networks
D (γ)	-0.063	-0.116	-0.011	-0.121	0.123	-0.102^{***}	-0.073
	(0.077)	(0.124)	(0.119)	(0.120)	(0.119)	(0.038)	(0.066)
D $\times$ PL (β)	0.126	-0.065	0.146	0.146	0.102	0.075	-0.161
	(0.099)	(0.151)	(0.136)	(0.149)	(0.187)	(0.072)	(0.130)
PL (δ)	-0.114	-0.245^{**}	-0.143^{**}	-0.173^{**}	0.141	-0.146^{**}	0.010
	(0.075)	(0.111)	(0.061)	(0.079)	(0.139)	(0.072)	(0.092)
Constant (α)	-0.280^{***}	-1.708^{***}	-0.371^{***}	0.192^{***}	-0.246^{**}	0.144^{***}	0.195^{***}
	(0.055)	(0.109)	(0.044)	(0.068)	(0.105)	(0.039)	(0.071)
Observations	701	701	701	701	701	701	701

Table A11: Survey 2 regression results as indicated in Equation 6 for summary indices, giving more weight to indicators that contribute more to the overall variation in the data. The baseline corresponds to the Computerized Randomization. PL stands for a dummy variable for the Public Lottery arm, and D is a dummy variable indicating complience (instrumented by assignment into treatment). The coefficients for D $\times$ PL allow to test for Hypothesis 1a directly. I control for the lagged variable. The bootstrapped standard errors are clustered at the geographical cluster level. $^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.1$.

	Info Adopt.	Practices Inputs	Financial Manag.	Risk Aversion	Psych. Well-b	Attrib. LOC	Social Networks
D (γ)	0.036	-0.079	-0.049	-0.022	0.164	-0.083	-0.203^{*}
	(0.069)	(0.067)	(0.089)	(0.128)	(0.153)	(0.062)	(0.116)
D $\times$ PL (β)	-0.162^{**}	-0.051	0.058	-0.095	-0.149	0.036	-0.013
	(0.081)	(0.088)	(0.117)	(0.161)	(0.216)	(0.093)	(0.142)
PL (δ)	0.136^{**}	-0.082	0.032	-0.020	0.279^{*}	0.000	0.106^{*}
	(0.065)	(0.108)	(0.081)	(0.140)	(0.161)	(0.070)	(0.060)
Constant (α)	-1.327^{***}	-1.745^{***}	-0.288^{***}	0.048	-0.225	0.057^{*}	0.168^{***}
	(0.038)	(0.095)	(0.065)	(0.113)	(0.141)	(0.033)	(0.038)
Observations	701	701	701	701	701	701	701

Table A12: Survey 3 regression results as indicated in Equation 6 for summary indices, giving more weight to indicators that contribute more to the overall variation in the data. The baseline corresponds to the Computerized Randomization. PL stands for a dummy variable for the Public Lottery arm, and D is a dummy variable indicating complience (instrumented by assignment into treatment). The coefficients for D $\times$ PL allow to test for Hypothesis 1a directly. I control for the lagged variable. The bootstrapped standard errors are clustered at the geographical cluster level. $^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.1$.

Still, PC weighting does not solve the correlation problem: the principal component's weighting method does not find unique information, but rather aims to find the direction of maximum variance in the data. Principal components identify the shared variance as the most important "story" in the data. If these happen to be variables that are highly correlated, the principal component method will not solve for them. In fact, it is precisely because they are highly correlated that the PC weighting might define them as the primary axis of variation.

Therefore, I use inverse covariance weighting. This method systematically searches for unique information, down-weighting shared variance. To implement this method, I compute the variance-covariance matrix of the component variables for each index in survey 1, thereby avoiding any post-assignment and post-treatment effects that may influence the weights. The weight for each component variable is the sum of the elements in the corresponding row of the inverse of the covariance matrix, normalized so that the weights sum to 1. Intuitively, if a variable moves very similarly to others, its unique contribution is small and its weight is reduced; if it captures something less correlated, it gets more weight.

As mentioned by Anderson (2008), this weighting method is conceptually analogous to the logic of Generalized Least Squares. Just as GLS down-weights noisy or correlated observations to produce an efficient estimate of a regression coefficient, the inverse covariance weighting scheme down-weights noisy or correlated component variables to produce a more informative summary index. Results from estimating Equations 5 and 6 are presented in Tables A13, A14, A15, and A16.

	Info Adopt.	**Practices Inputs**	**Financial Manag.**	**Risk Aversion**	**Psych. Well-b**	**Attrib. LOC**	**Social Networks**
Treat (γ)	-0.060	-0.071	0.009	0.011	0.089	-0.108	-0.065
	(0.064)	(0.104)	(0.098)	(0.165)	(0.093)	(0.066)	(0.113)
PL (δ)	-0.113	-0.189^{**}	-0.206^{*}	0.178	0.137	-0.246^{***}	0.029
	(0.080)	(0.083)	(0.118)	(0.116)	(0.123)	(0.078)	(0.102)
Treat × PL (β)	0.120	0.197	-0.005	-0.085	0.136	0.166	-0.111
	(0.089)	(0.126)	(0.132)	(0.216)	(0.136)	(0.126)	(0.164)
Constant (α)	-0.280^{***}	-2.193^{***}	0.177^{*}	-0.252^{***}	-0.409^{***}	0.167^{**}	0.449^{***}
	(0.051)	(0.066)	(0.090)	(0.088)	(0.075)	(0.065)	(0.080)
Observations	701	701	701	701	701	701	701

Table A13: Survey 2 regression results as indicated in Equation 5 for summary indices, weighting them inversely to their correlation with each other. The baseline corresponds to the Computerized Randomization. PL stands for a dummy variable for the Public Lottery arm, and Treat is a dummy variable indicating whether the subject was assigned to treatment. The coefficients for Treat × PL allow to test for Hypothesis 1a directly. I control for the lagged variable. The bootstrapped standard errors are clustered at the geographical cluster level. $^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.1$.

	Info Adopt.	**Practices Inputs**	**Financial Manag.**	**Risk Aversion**	**Psych. Well-b**	**Attrib. LOC**	**Social Networks**
Treat (γ)	0.035	-0.042	0.008	0.109	0.174	-0.065	-0.157
	(0.059)	(0.064)	(0.135)	(0.126)	(0.150)	(0.072)	(0.104)
PL (δ)	0.135^{*}	0.006	0.016	0.196^{**}	0.358^{**}	-0.006	0.085
	(0.069)	(0.098)	(0.111)	(0.085)	(0.156)	(0.072)	(0.079)
Treat × PL (β)	-0.154^{*}	-0.099	0.065	-0.135	-0.317	0.125	0.155
	(0.079)	(0.081)	(0.173)	(0.173)	(0.193)	(0.133)	(0.151)
Constant (α)	-1.327^{***}	-2.357^{***}	2.328^{***}	-0.274^{***}	-0.375^{**}	0.063	0.272^{***}
	(0.039)	(0.077)	(0.062)	(0.067)	(0.138)	(0.057)	(0.041)
Observations	701	701	701	701	701	701	701

Table A14: Survey 3 regression results as indicated in Equation 5 for summary indices, weighting them inversely to their correlation with each other. The baseline corresponds to the Computerized Randomization. PL stands for a dummy variable for the Public Lottery arm, and Treat is a dummy variable indicating whether the subject was assigned to treatment. The coefficients for Treat × PL allow to test for Hypothesis 1a directly. I control for the lagged variable. The bootstrapped standard errors are clustered at the geographical cluster level. $^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.1$.

	Info Adopt.	**Practices Inputs**	**Financial Manag.**	**Risk Aversion**	**Psych. Well-b**	**Attrib. LOC**	**Social Networks**
D (γ)	-0.063	-0.075	0.009	0.011	0.093	-0.113	-0.068
	(0.064)	(0.105)	(0.121)	(0.203)	(0.118)	(0.078)	(0.130)
D × PL (β)	0.126	0.207	-0.005	-0.089	0.142	0.174	-0.116
	(0.100)	(0.126)	(0.150)	(0.252)	(0.179)	(0.136)	(0.173)
PL (δ)	-0.114	-0.191^{**}	-0.206	0.179	0.135	-0.247^{**}	0.030
	(0.082)	(0.078)	(0.129)	(0.113)	(0.119)	(0.101)	(0.132)
Constant (α)	-0.280^{***}	-2.193^{***}	0.177^{*}	-0.252^{***}	-0.410^{***}	0.167^{**}	0.449^{***}
	(0.048)	(0.064)	(0.093)	(0.085)	(0.084)	(0.074)	(0.101)
Observations	701	701	701	701	701	701	701

Table A15: Survey 2 regression results as indicated in Equation 6 for summary indices, weighting them inversely to their correlation with each other. The baseline corresponds to the Computerized Randomization. PL stands for a dummy variable for the Public Lottery arm, and D is a dummy variable indicating complience (instrumented by assignment into treatment). The coefficients for D × PL allow to test for Hypothesis 1a directly. I control for the lagged variable. The bootstrapped standard errors are clustered at the geographical cluster level. $^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.1$.

	Info Adopt.	**Practices Inputs**	**Financial Manag.**	**Risk Aversion**	**Psych. Well-b**	**Attrib. LOC**	**Social Networks**
D (γ)	0.036	−0.043	0.008	0.114	0.182	−0.068	−0.165
	(0.051)	(0.058)	(0.139)	(0.141)	(0.156)	(0.100)	(0.117)
D × PL (β)	-0.162^{**}	−0.103	0.068	−0.141	-0.333^{**}	0.131	0.163
	(0.077)	(0.071)	(0.166)	(0.177)	(0.169)	(0.158)	(0.174)
PL (δ)	0.136^{**}	0.008	0.015	0.197^{**}	0.361^{***}	−0.007	0.084
	(0.066)	(0.110)	(0.110)	(0.090)	(0.140)	(0.069)	(0.085)
Constant (α)	-1.327^{***}	-2.357^{***}	2.328^{***}	-0.275^{***}	-0.376^{***}	0.063	0.272^{***}
	(0.034)	(0.079)	(0.061)	(0.061)	(0.145)	(0.063)	(0.048)
Observations	701	701	701	701	701	701	701

Table A16: Survey 3 regression results as indicated in Equation 6 for summary indices, weighting them inversely to their correlation with each other. The baseline corresponds to the Computerized Randomization. PL stands for a dummy variable for the Public Lottery arm, and D is a dummy variable indicating complience (instrumented by assignment into treatment). The coefficients for D × PL allow to test for Hypothesis 1a directly. I control for the lagged variable. The bootstrapped standard errors are clustered at the geographical cluster level. $^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.1$.

It is worth noting that both methods can produce negative weights for some variables, as Figure A1 shows. In PCA, negative loadings indicate that a variable moves in the opposite direction along the primary axis of variation, contributing to the definition of that dimension. In inverse covariance weighting, negative weights arise when a variable is strongly correlated with the others: the method assigns it a compensating negative value to down-weight redundant information and highlight independent variation. In both cases, negative weights are not a problem but reflect the geometry of the data.

In summary, while equal weighting implicitly treats all variables as equally informative, both PCA and inverse covariance weighting exploit the correlation structure of the data in different ways—PCA by emphasizing shared variance and inverse covariance weighting by highlighting unique contributions—ensuring that results are robust to alternative, principled ways of constructing indices.

Appendix H The data-driven indices with five clusters

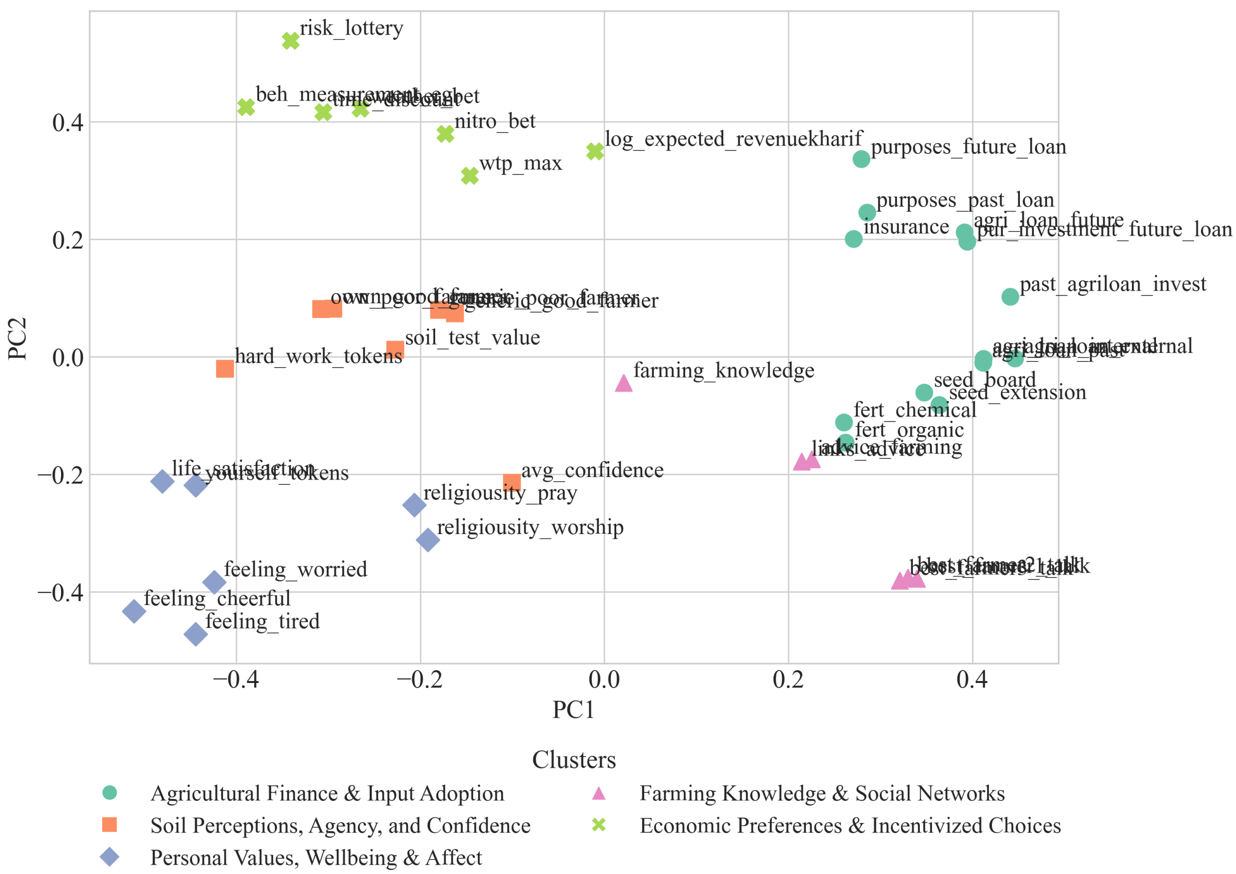


Figure A2: This figure illustrates the results of k-means clustering applied to the embeddings.

Category (k)	Interpretation	Variables included
Agricultural Finance & Input Adoption	This group is unified by financial products (loans, insurance), sources, and also adoption of seeds and fertilizers—capturing both access to resources and uptake of essential agricultural inputs.	agri_loan_past, agri_loan_future, purposes_future_loan, pur_investment_future_loan, insurance, seed_extension, seed_board, fert_organic, fert_chemical, past_agriloan_invest, purposes_past_loan, agri_loan_internal, agri_loan_external
Soil Perceptions, Agency, & Confidence	All variables relate to beliefs about soil quality, responsibility/blame/credit (own and generic), hard work (agency), and confidence in farming—a cognitive/attitudinal grouping.	soil_test_value, generic_poor_farmer, generic_good_farmer, own_poor_farmer, own_good_farmer, hard_work_tokens, avg_confidence
Personal Values, Wellbeing & Affect	Composed of variables on locus of control, religiosity, emotional states, and life satisfaction—focusing on subjective wellbeing and intrinsic values.	yourself_tokens, religiousity_worship, religiousity_pray, life_satisfaction, feeling_cheerful, feeling_worried, feeling_tired
Farming Knowledge & Social Networks	This cluster blends objective farming knowledge with indicators of advice-seeking/giving and local farming expertise—emphasizing knowledge-sharing networks.	farming_knowledge, links_advice, advice_farming, best_farmer1_talk, best_farmer2_talk, best_farmer$_talk$
Economic Preferences & Incentivized Choices	WTP, behavioral measures, incentivized bets, time discounting, risk/reward choices, and revenue expectations—all reflect economic decision-making or behavioral preferences.	wtp_max, beh_measurement_eg, nitro_bet, weather_bet, time_discount, risk_lottery, expected_revenuekharif

Table A17: This table describes the five conceptual categories, each of which is composed of several survey questions. The variable nitro_bet was not asked for in the Pure Control arm to avoid informing farmers of the intervention taking place in other villages.

Appendix I Results for the data-driven indices

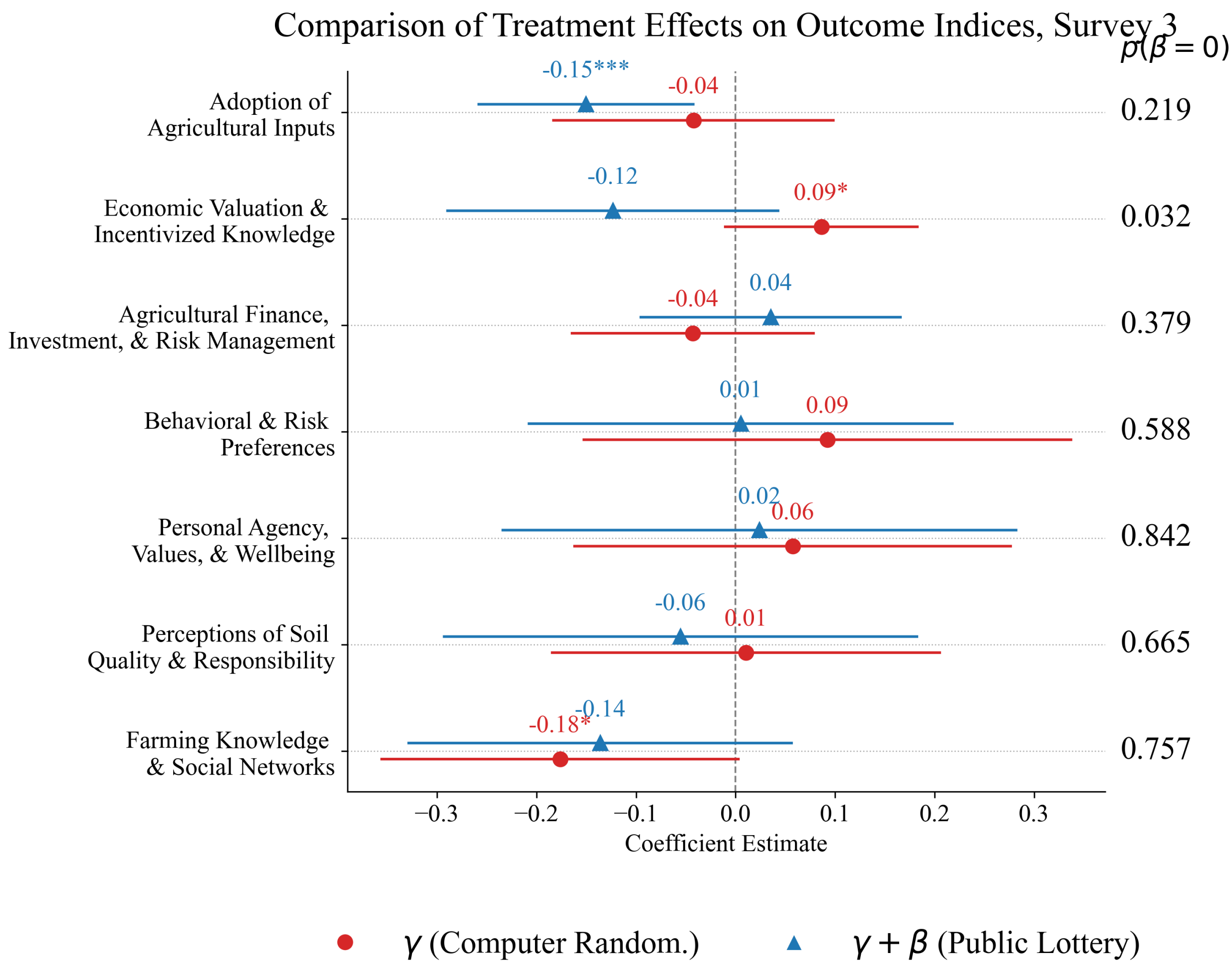


Figure A3: The figure presents coefficient estimates with 95% confidence intervals for the ITT in the Computerized Randomization arm (γ) and for the ITT in the Public Lottery arm ($\gamma + \beta$). The p-values shown above each point represent significance levels for the respective effects. The rightmost values show p-values for the test of $\beta = 0$, indicating whether the Public Lottery effect is statistically distinguishable from the computer randomization effect. The standard errors are clustered at the geographic cluster level, and the estimation was done by adding the baseline level of the corresponding outcome index as stated in Equation 5. All coefficients are expressed in standard deviation (SD) units, standardized relative to the pooled control group. This figure is the equivalent of Figure 4.

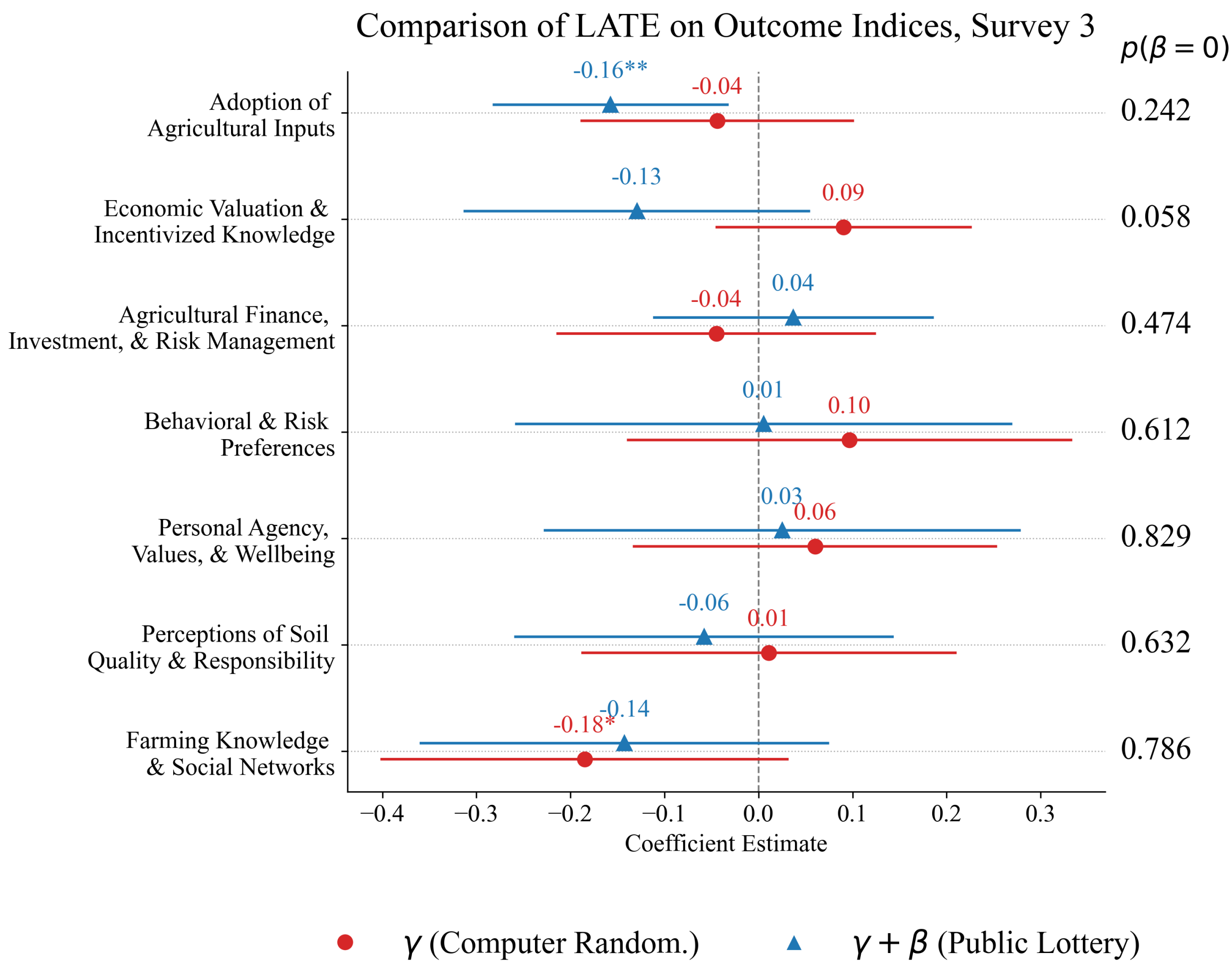


Figure A4: The figure presents coefficient estimates with 95% confidence intervals for the LATE in the Computerized Randomization arm (γ) and for the LATE in the Public Lottery arm ($\gamma + \beta$). The p-values shown above each point represent significance levels for the respective effects. The rightmost values show p-values for the test of $\beta = 0$, indicating whether the Public Lottery effect is statistically distinguishable from the computer randomization effect. The standard errors are clustered at the geographic cluster level, and the estimation was done by adding the baseline level of the corresponding outcome index as stated in Equation 6. All coefficients are expressed in standard deviation (SD) units, standardized relative to the pooled control group. This figure is the equivalent of Figure 5.

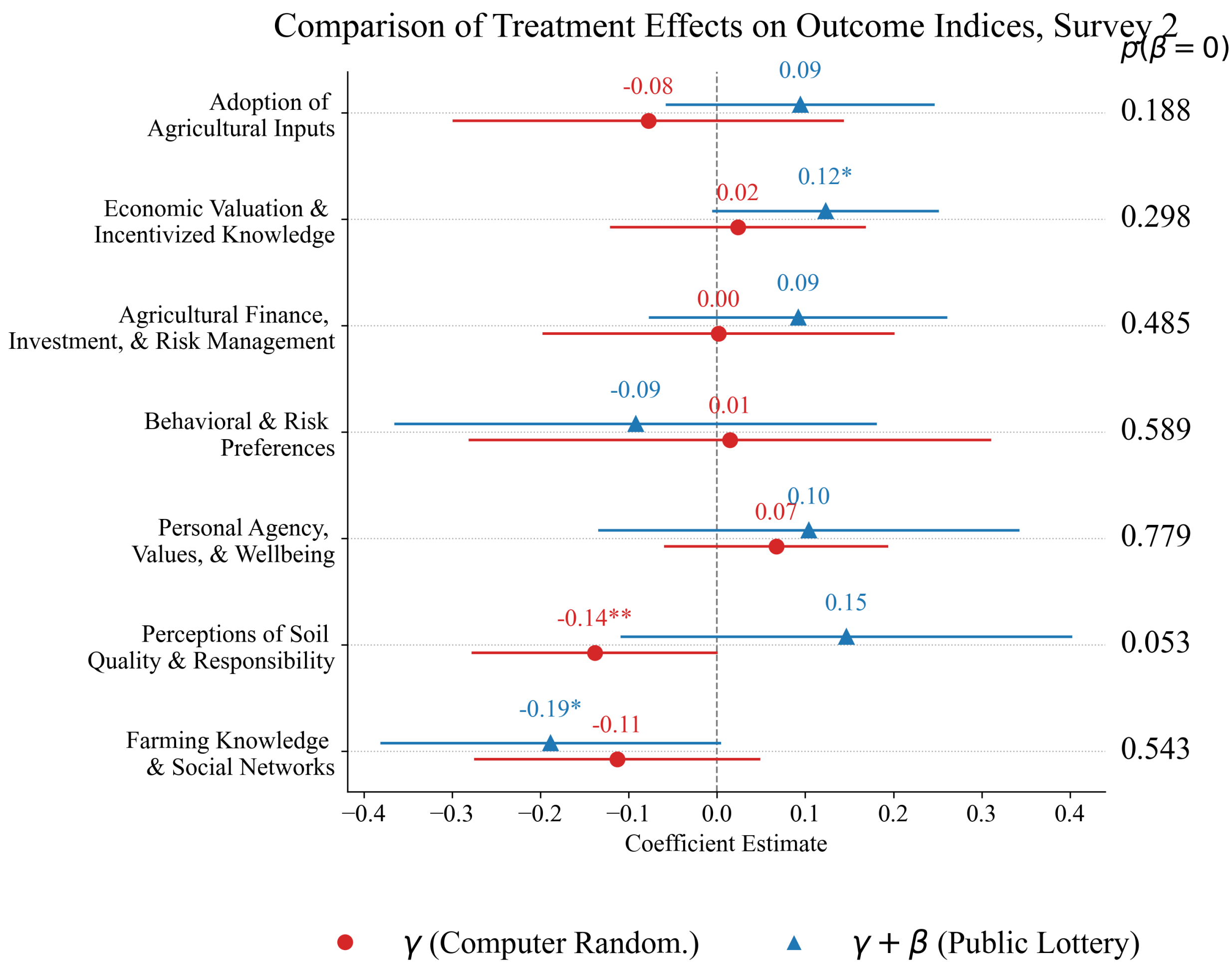


Figure A5: The figure presents coefficient estimates with 95% confidence intervals for the ITT in the Computerized Randomization arm (γ) and for the ITT in the Public Lottery arm ($\gamma + \beta$) after farmers were assigned into groups, but before the treatment was distributed. The p-values shown above each point represent significance levels for the respective effects. The rightmost values show p-values for the test of $\beta = 0$, indicating whether the Public Lottery effect is statistically distinguishable from the computer randomization effect. The standard errors are clustered at the geographic cluster level, and the estimation was done by adding the baseline level of the corresponding outcome index as stated in Equation 5. All coefficients are expressed in standard deviation (SD) units, standardized relative to the pooled control group. This figure is the equivalent of Figure 6.

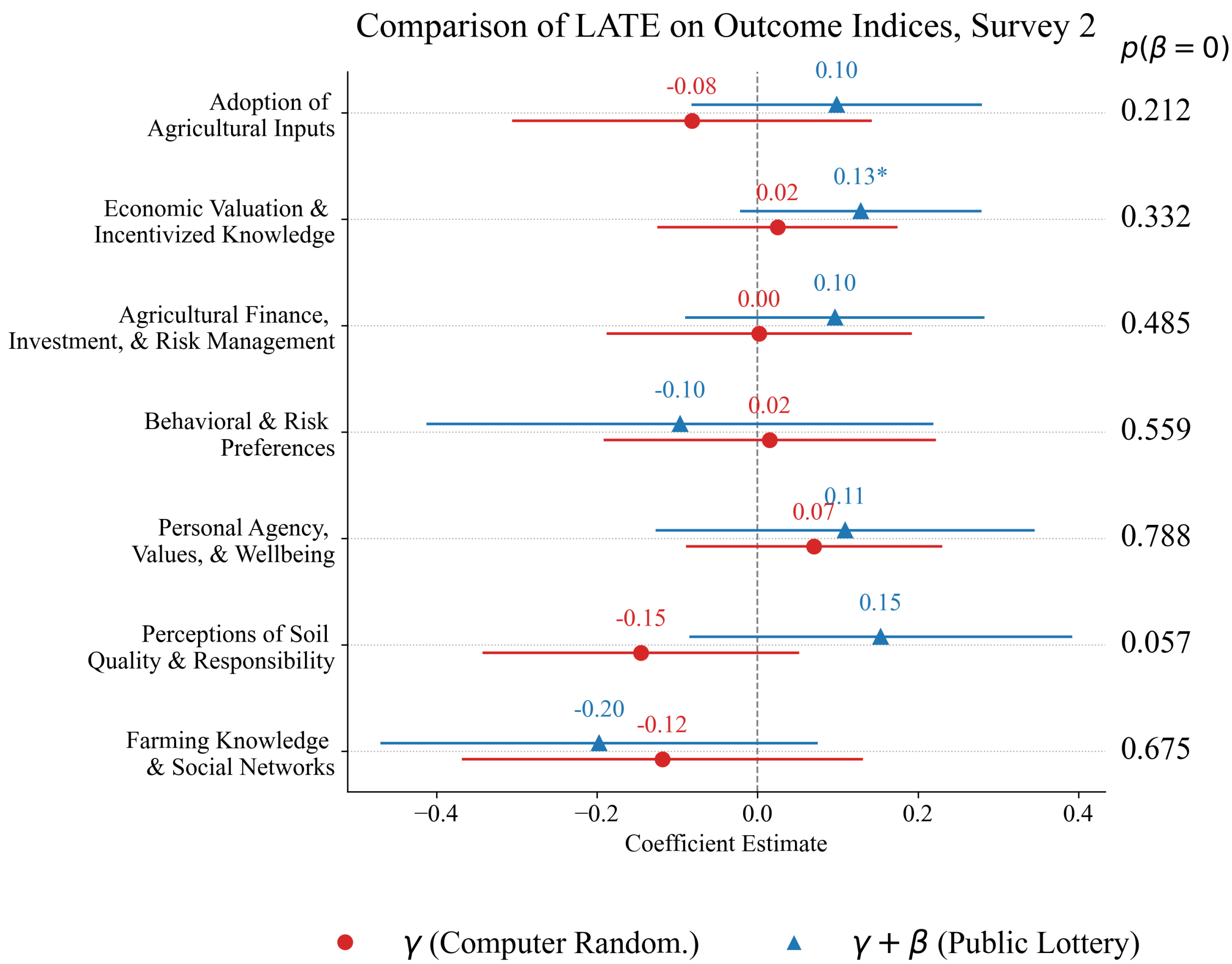
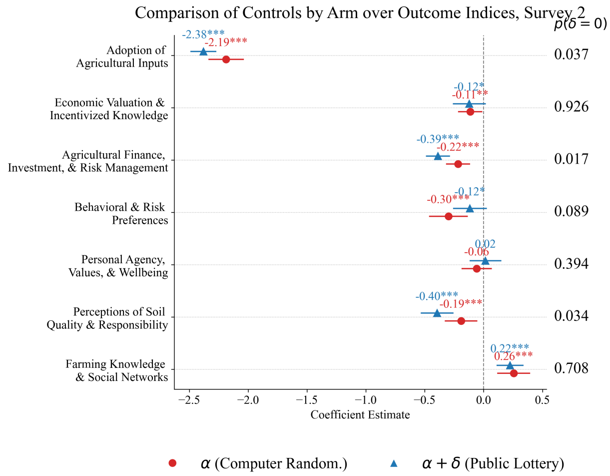
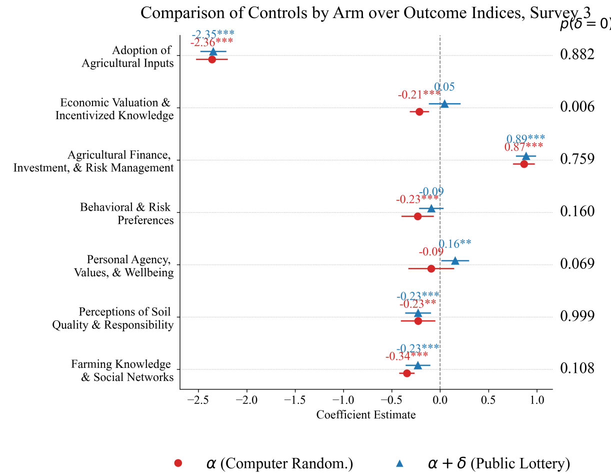


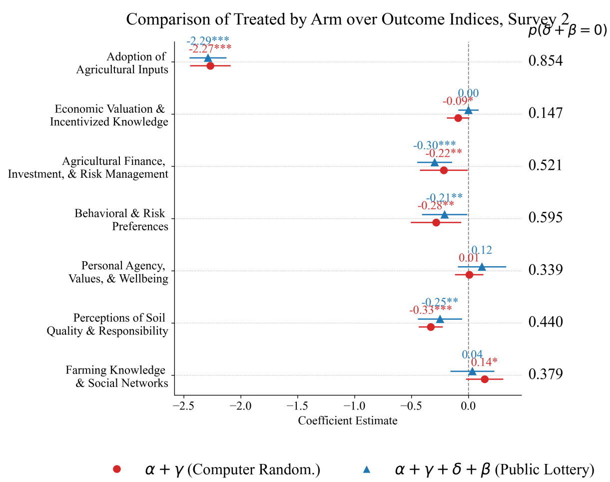
Figure A6: The figure presents coefficient estimates with 95% confidence intervals for the ITT in the Computerized Randomization arm (γ) and for the ITT in the Public Lottery arm ($\gamma + \beta$) after farmers were assigned into groups, but before the treatment was distributed. The p-values shown above each point represent significance levels for the respective effects. The rightmost values show p-values for the test of $\beta = 0$, indicating whether the Public Lottery effect is statistically distinguishable from the computer randomization effect. The standard errors are clustered at the geographic cluster level, and the estimation was done by adding the baseline level of the corresponding outcome index as stated in Equation 5. All coefficients are expressed in standard deviation (SD) units, standardized relative to the pooled control group. This figure is the equivalent of Figure 7.



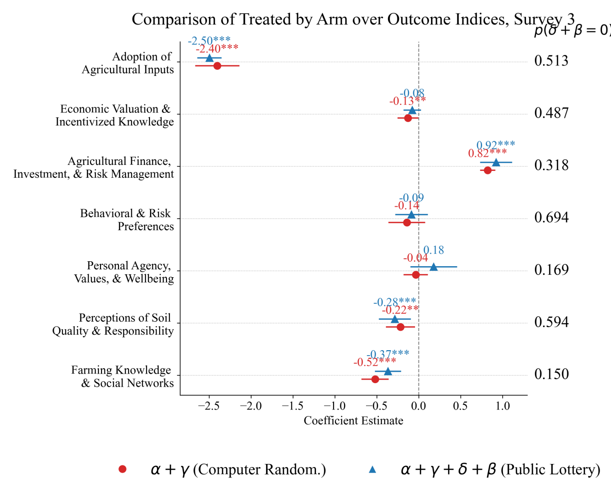
(a) Control, survey 2



(b) Control, survey 3

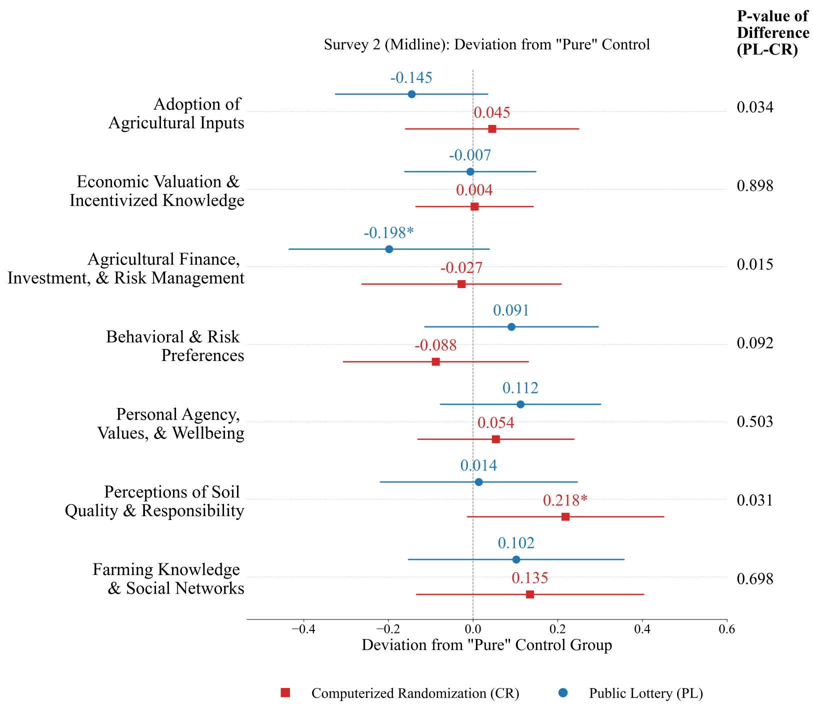


(c) Treated, survey 2

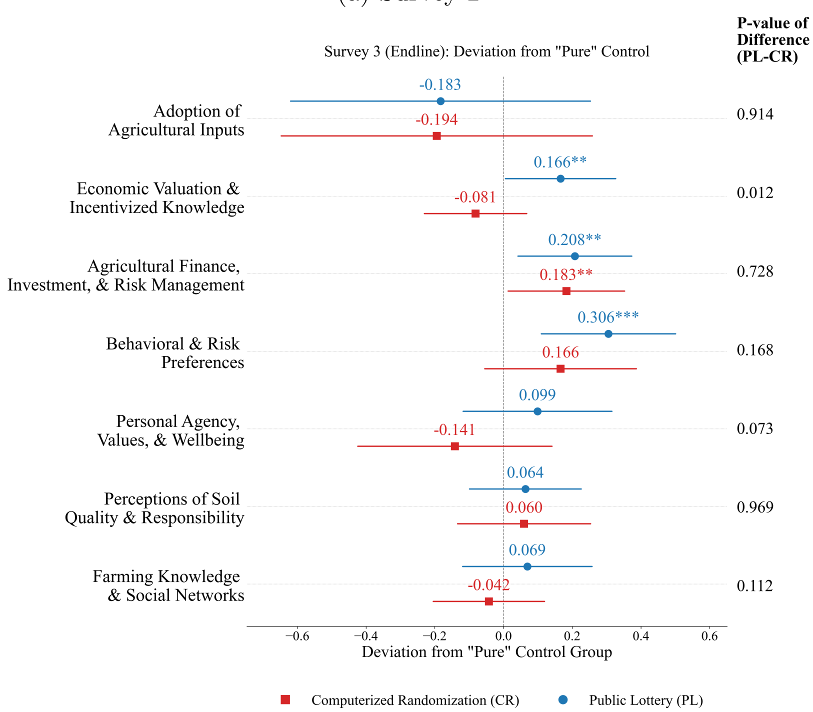


(d) Treated, survey 3

Figure A7: The figures present coefficient estimates with 95% confidence intervals for the control and treated subjects' outcome values in the Computerized Randomization arm and the Public Lottery arm. Figures (a) and (b) show the coefficients for the control subjects in surveys 2 and 3, respectively. Figures (c) and (d) show the coefficients for the treated subjects in surveys 2 and 3. The p-values shown above each point represent significance levels for the respective effects. The rightmost values show p-values for the test of $\delta = 0$, indicating whether the Public Lottery effect is statistically distinguishable from the computer randomization effect. The standard errors are clustered at the geographic cluster level, and the estimation was done by adding the baseline level of the corresponding outcome index as stated in Equation 6. All coefficients are expressed in standard deviation (SD) units, standardized relative to the pooled control group. This figure is equivalent to Figure 8.

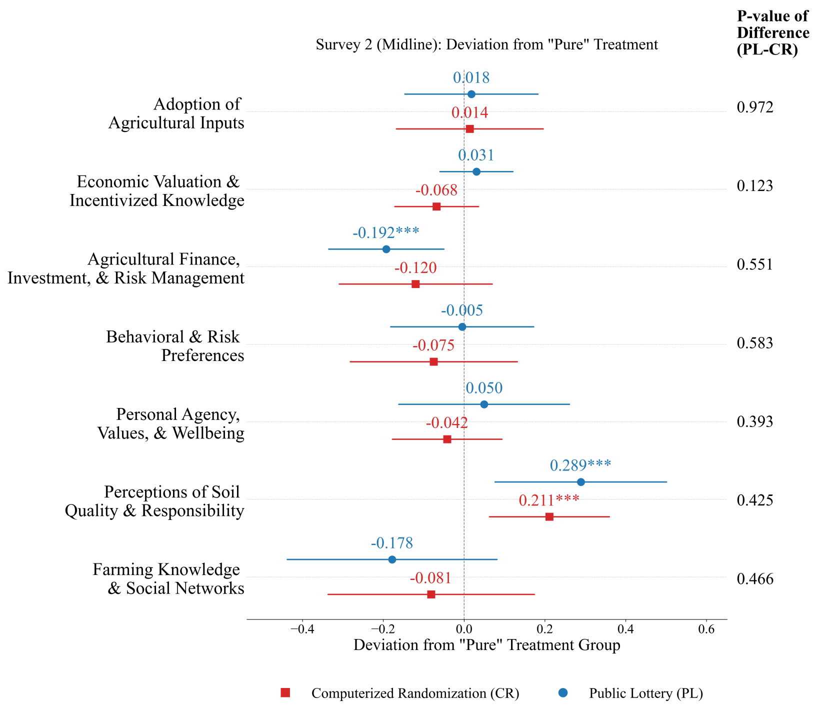


(a) Survey 2

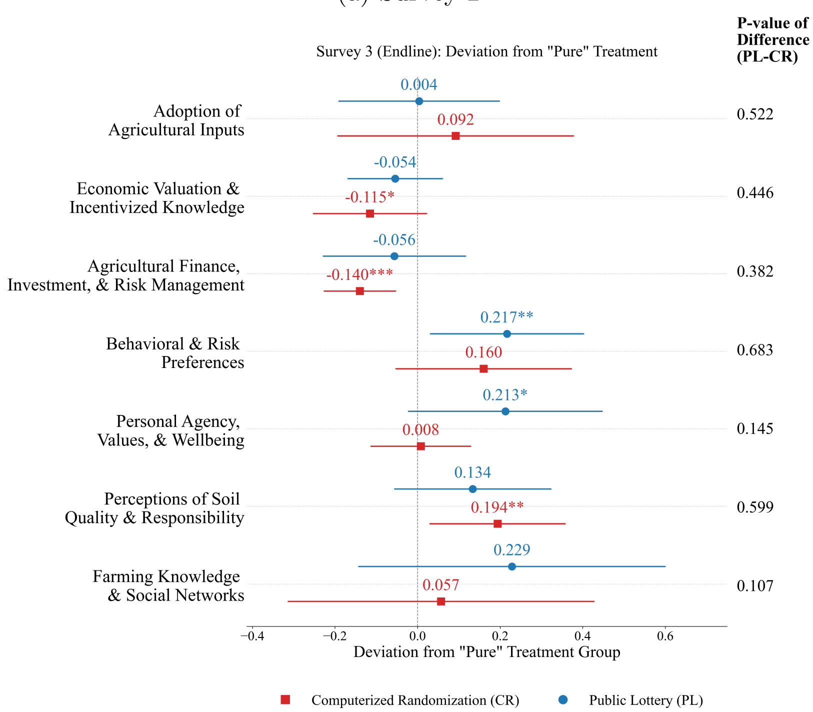


(b) Survey 3

Figure A8: Controls by Arm compared to the Pure Control Benchmark. The figures present coefficient estimates with 95% confidence intervals for the control subjects in the Computerized Randomization arm and in the Public Lottery arm with respect to the Pure Control arm. Figure A8a shows the coefficients in survey 2, while Figure A8b shows the coefficients in survey 3. The rightmost values show the p-values for the test of $\beta^{CR,C} = \beta^{PL,C}$, indicating whether the Public Lottery coefficient is statistically distinguishable from the Computerized Randomization coefficient. The standard errors are clustered at the geographic cluster level, and the estimation is done by adding the baseline level of the corresponding outcome index as stated in Equation 8. All coefficients are expressed in standard deviation (SD) units, standardized relative to the pooled control group. This figure is equivalent to Figure **??**.



(a) Survey 2



(b) Survey 3

Figure A9: Treated by Arm compared to the Pure Treatment Benchmark. The figures present coefficient estimates with 95% confidence intervals for the treated subjects in the Computerized Randomization arm and in the Public Lottery arm with respect to the Pure Treatment arm. Figure A9a shows the coefficients in survey 2, while Figure A9b shows the coefficients in survey 3. The rightmost values show the p-values for the test of $\beta^{CR,C} = \beta^{PL,C}$, indicating whether the Public Lottery coefficient is statistically distinguishable from the Computerized Randomization coefficient. The standard errors are clustered at the geographic cluster level, and the estimation is done by adding the baseline level of the corresponding outcome index as stated in Equation 7. All coefficients are expressed in standard deviation (SD) units, standardized relative to the pooled treatment group. This figure is equivalent to Figure 10.

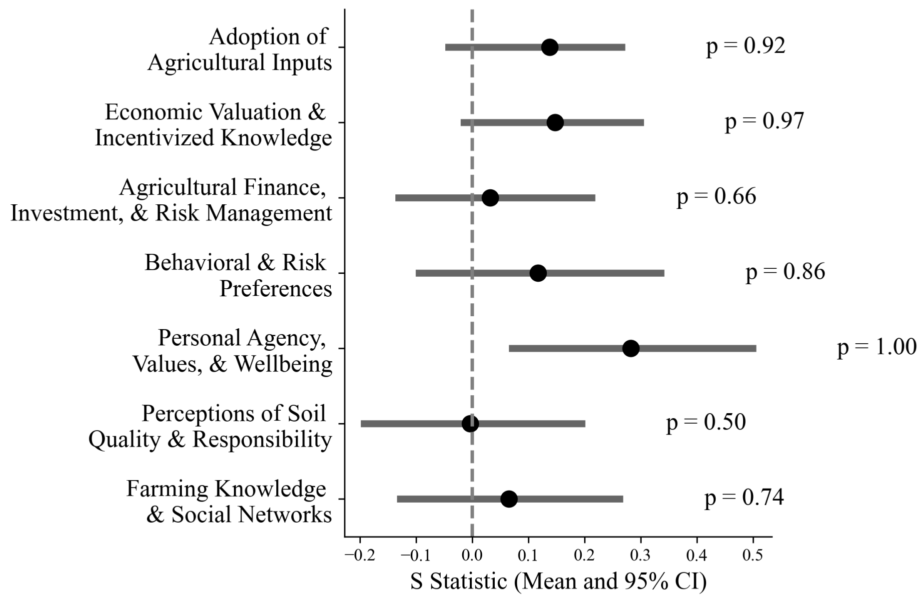


Figure A10: Results from testing the null hypothesis 9 on separability using the data-driven summary indices, standard bootstrap. The plot displays the calculated statistic S by index outcome, along with their 95% confidence intervals, where $S = T_1 + T_2 - T_3$. In the case of the Adoption of Agricultural Inputs index, $S = |T_1| + |T_2| - |T_3|$, as the direction of treatment effects is uninterpretable. The p-values shown are the one-sided bootstrap p-values for $S \geq 0$ corresponding to each index; intuitively, the p-value provides the empirical probability that the null hypothesis $S \geq 0$ holds. Then, the lower the p-value, the stronger the evidence against the null hypothesis. See Figure A13 for the results using Bayesian bootstrap.

Appendix J Methodological Considerations for Index Standardization

In this paper, I construct several outcome indices by aggregating multiple individual variables, following a methodology similar to that of Kling, Liebman, and Katz (2007). A standard procedure for this aggregation involves standardizing each component variable before averaging them into an index. This is typically done by demeaning each variable and dividing by the standard deviation of the control group in the same survey wave (a method I term "Period-Specific Standardization" for exposition purposes). While this approach is appropriate under standard assumptions, it can produce misleading estimates of the treatment effect if the experimental design itself influences the moments of the control group's outcome distribution.

The central argument of this paper is that participants' awareness of the assignment mechanism can generate assignment effects, which may manifest as spillovers onto the control group. These spillovers could affect not only the mean (first moment) but also the variance (second moment) of the outcomes for both the treatment and control groups. If, for example, the variance of the control group at midline or endline is itself affected by the assignment (or even treatment), then using it as a denominator for standardization renders the scale of the outcome variable endogenous to the treatment.

This section formally demonstrates this issue through a simulation exercise. I show that when spillovers (which could be assignment effects) affect the variance of the control group, Period-Specific Standardization leads to an attenuated estimate of the average treatment effect. Consequently, I justify my choice of an alternative method, "Baseline-Only Standardization," which uses the pre-treatment moments of the control group as a fixed benchmark, providing a stable and more interpretable estimate of the average treatment effect under the conditions posited by our theoretical framework.

Let y_{it} be the outcome of interest for individual i at time $t \in \{1, 2\}$, with $t = 1$ the baseline and $t = 2$ the endline. Let C_t denote the set of individuals in the control group at time t. Let $\mu_{c,t}$ and $\sigma_{c,t}$ be the mean and standard deviation, respectively, of the outcome for the control group at time t. We define the standardized outcome, z_{it}, under two alternative methods:

- Method 1: Period-Specific Standardization. This is the standard approach, exemplified by the analysis in Banerjee, Duflo, Goldberg, Karlan, Osei, Parienté, Shapiro, Thuysbaert, and Udry (2015). The endline outcome is standardized using the contemporaneous moments of the endline control group.

$$z_{i,t=1} = (y_{i,t=1} - \mu_{c,t=1}) / \sigma_{c,t=1}$$
$$z_{i,t=2} = (y_{i,t=2} - \mu_{c,t=2}) / \sigma_{c,t=2}$$

- Method 2: Baseline-Only Standardization. In this method, the moments of the baseline control group serve as a fixed benchmark for standardizing outcomes in all periods.

$$z_{i,t=1} = (y_{i,t=1} - \mu_{c,t=1}) / \sigma_{c,t=1}$$
$$z_{i,t=2} = (y_{i,t=2} - \mu_{c,t=1}) / \sigma_{c,t=1}$$

To illustrate the consequences of choosing one or another method, I conduct a simulation exercise. I generate a two-period panel dataset with 100 individuals, where half of whom are assigned to treatment. I simulate a scenario where the treatment has a positive effect on the mean outcome for the treated group, but also a negative spillover effect on the mean outcome for the control group. I then consider two cases.

1. Additive Effects Only. The spillover affects only the mean of the control group's outcomes.

2. Additive and Variance Effects. The spillover affects both the mean and the variance of the control group's outcomes, increasing outcome volatility. This case models the potential real-world consequences of assignment awareness as theorized in this paper.

For each case, I construct outcome indices using both Method 1 and Method 2 and estimate the treatment effect using a standard ANCOVA specification:

$$z_{i,t=2} = \beta_0 + \beta_1 Treatment_i + \beta_2 z_{i,t=1} + \varepsilon_i$$

The results of this exercise are presented in Table A18.

	Method 1 Period-Specific	**Method 2** Baseline-Only
Panel A: Spillover Affects Mean Only		
Treatment	**3.084***	**3.053***
	(0.085)	(0.086)
Index at t=1	0.890***	0.876***
	(0.036)	(0.036)
Intercept	0.000	-1.042***
	(0.060)	(0.060)
Panel B: Spillover Affects Mean and Variance		
Treatment	**1.687***	**2.904***
	(0.085)	(0.150)
Index at t=1	0.649***	1.114***
	(0.035)	(0.063)
Intercept	0.000	-0.793***
	(0.060)	(0.106)
N	100	100

Table A18: The table reports coefficients from an ANCOVA regression of the endline index on treatment status and the baseline index. Standard errors are in parentheses. The data generation process simulates an additive treatment effect of +1.0 baseline standard deviations and a negative spillover effect on the control group mean of -0.5 baseline standard deviations. In Panel B, the spillover also doubles the standard deviation of the control group's outcomes. *** p<0.01, ** p<0.05, * p<0.1.

As shown in Panel A of Table A18, when the spillover affects only the mean, the ANCOVA model is robust. Both standardization methods recover a similar estimate for the treatment coefficient ($\beta_1 \approx 3.0$), correctly identifying the vertical distance between the parallel treatment and control trend lines. The choice of method primarily affects the intercept.

However, Panel B reveals the critical issue. When the spillover also increases the variance of the control group's outcomes, the estimated treatment effect under Method 1 (1.687) is dramatically attenuated compared to the estimate from Method 2 (2.904). This occurs because the denominator of the standardization in Method 1, $\sigma_{c,t=2}$, is now larger due to the spillover. This compresses the scale of the dependent variable, mechanically reducing the estimated coefficient. The resulting β_1

is no longer an interpretable measure of the average treatment effect on the mean, but a complex function of the effects on both the mean and the variance. This is a dynamic variant of the critique of standardized coefficients raised by Greenland, Schlesselman, and Criqui (1986).

Given that the theoretical framework of this paper posits that assignment awareness can affect the entire outcome distribution for all participants, including the control group, the assumption required for Method 1—that the control group's variance is unaffected—is untenable. Therefore, I adopt Method 2 (Baseline-Only Standardization) for the construction of all outcome indices in this paper. This method provides a stable and consistent measure of the average treatment effect on the mean, expressed in fixed units of the baseline control group's standard deviation. While this choice may be less conventional, it is the methodologically appropriate one for providing an interpretable estimate of the treatment effect in the presence of the assignment effects central to my analysis.[21]

Appendix K Tables with pre-selected controls

	Info Adopt.	**Practices Inputs**	**Financial Manag.**	**Risk Aversion**	**Psych. Well-b**	**Attrib. LOC**	**Social Networks**
Treat (γ)	-0.059	-0.080	-0.001	-0.004	0.075	-0.122^{*}	-0.086
	(0.065)	(0.107)	(0.100)	(0.165)	(0.089)	(0.066)	(0.080)
PL (δ)	-0.113	-0.183^{**}	-0.153^{**}	0.132	0.134	-0.225^{***}	0.011
	(0.086)	(0.084)	(0.064)	(0.117)	(0.115)	(0.081)	(0.084)
Treat × PL (β)	0.109	0.181	0.084	-0.076	0.125	0.159	-0.119
	(0.093)	(0.127)	(0.127)	(0.215)	(0.133)	(0.117)	(0.142)
Constant (α)	-0.304^{**}	-2.266^{***}	-0.420^{***}	-0.238	-0.552^{***}	0.094	0.207
	(0.119)	(0.115)	(0.098)	(0.159)	(0.159)	(0.095)	(0.135)
Observations	701	701	701	701	701	701	701

Table A19: Survey 2 regression results as indicated in Equation 5 for summary indices. The baseline corresponds to the Computerized Randomization. PL stands for a dummy variable for the Public Lottery arm, and Treat is a dummy variable indicating whether the subject was assigned to treatment. The coefficients for Treat × PL allow to test for Hypothesis 1a directly. I control for the lagged variable, household income, the farmer's level of education, and whether the farmer owns the selected plot. The bootstrapped standard errors are clustered at the geographical cluster level. $^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.1$.

[21] For the reader curious about the boundary conditions of this finding, my simulation framework confirms two further implications that reinforce my choice of methodology. First, the core conclusion is robust to changes in the treatment group's variance. One might conjecture that if the treatment itself increases outcome variance for the treated it might alter the results. However, the fundamental result will not change: The bias in Method 1 is driven entirely by the endogeneity of the control group's standard deviation, which serves as the standardization denominator. The variance of the treatment group is irrelevant to the construction of this scale.

Second, the bias in Method 1 is symmetric. If a spillover effect were to decrease the control group's variance, Method 1 produces an inflated estimate of the treatment effect, whereas the Method 2 estimate remains stable and correctly centered on the true additive effect.

	Info Adopt.	Practices Inputs	Financial Manag.	Risk Aversion	Psych. Well-b	Attrib. LOC	Social Networks
Treat (γ)	0.020	−0.040	−0.036	0.075	0.150	−0.078	-0.186^{*}
	(0.059)	(0.070)	(0.058)	(0.131)	(0.140)	(0.076)	(0.099)
PL (δ)	0.123^{*}	0.020	0.034	0.163^{*}	0.347^{**}	0.001	0.089
	(0.068)	(0.099)	(0.075)	(0.082)	(0.145)	(0.074)	(0.071)
Treat × PL (β)	-0.153^{*}	−0.109	0.078	−0.104	−0.273	0.119	0.103
	(0.085)	(0.085)	(0.088)	(0.180)	(0.191)	(0.131)	(0.139)
Constant (α)	-1.513^{***}	-2.361^{***}	0.921^{***}	-0.458^{***}	-0.663^{**}	0.061	0.197^{*}
	(0.117)	(0.090)	(0.104)	(0.086)	(0.271)	(0.100)	(0.098)
Observations	701	701	701	701	701	701	701

Table A20: Survey 3 regression results as indicated in Equation 5 for summary indices. The baseline corresponds to the Computerized Randomization. PL stands for a dummy variable for the Public Lottery arm, and Treat is a dummy variable indicating whether the subject was assigned to treatment. The coefficients for Treat × PL allow to test for Hypothesis 1a directly. I control for the lagged variable, household income, the farmer's level of education, and whether the farmer owns the selected plot. The bootstrapped standard errors are clustered at the geographical cluster level. $^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.1$.

	Info Adopt.	Practices Inputs	Financial Manag.	Risk Aversion	Psych. Well-b	Attrib. LOC	Social Networks
D (γ)	−0.062	−0.084	−0.001	−0.004	0.079	-0.129^{*}	−0.090
	(0.067)	(0.110)	(0.102)	(0.170)	(0.092)	(0.069)	(0.082)
D × PL (β)	0.115	0.190	0.088	−0.079	0.131	0.167	−0.124
	(0.095)	(0.131)	(0.130)	(0.220)	(0.136)	(0.120)	(0.146)
PL (δ)	−0.114	-0.184^{**}	-0.154^{**}	0.133	0.132	-0.227^{***}	0.013
	(0.085)	(0.083)	(0.063)	(0.116)	(0.112)	(0.080)	(0.083)
Constant (α)	-0.304^{***}	-2.265^{***}	-0.419^{***}	−0.239	-0.549^{***}	0.094	0.203
	(0.117)	(0.113)	(0.096)	(0.155)	(0.155)	(0.094)	(0.133)
Observations	701	701	701	701	701	701	701

Table A21: Survey 2 regression results as indicated in Equation 6 for summary indices. The baseline corresponds to the Computerized Randomization. PL stands for a dummy variable for the Public Lottery arm, and D is a dummy variable indicating complience (instrumented by assignment into treatment). The coefficients for D × PL allow to test for Hypothesis 1a directly. I control for the lagged variable, household income, the farmer's level of education, and whether the farmer owns the selected plot. The bootstrapped standard errors are clustered at the geographical cluster level. $^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.1$.

	Info Adopt.	Practices Inputs	Financial Manag.	Risk Aversion	Psych. Well-b	Attrib. LOC	Social Networks
D (γ)	0.021	-0.042	-0.037	0.079	0.157	-0.082	-0.196^{*}
	(0.061)	(0.072)	(0.060)	(0.136)	(0.144)	(0.078)	(0.102)
D $\times$ PL (β)	-0.160^{*}	-0.114	0.081	-0.109	-0.287	0.124	0.108
	(0.088)	(0.087)	(0.090)	(0.185)	(0.196)	(0.135)	(0.143)
PL (δ)	0.125^{*}	0.022	0.034	0.164^{**}	0.349^{**}	0.000	0.088
	(0.067)	(0.097)	(0.073)	(0.081)	(0.143)	(0.073)	(0.070)
Constant (α)	-1.515^{***}	-2.363^{***}	0.921^{***}	-0.458^{***}	-0.664^{**}	0.062	0.195^{**}
	(0.113)	(0.088)	(0.102)	(0.084)	(0.264)	(0.097)	(0.096)
Observations	701	701	701	701	701	701	701

Table A22: Survey 3 regression results as indicated in Equation 6 for summary indices. The baseline corresponds to the Computerized Randomization. PL stands for a dummy variable for the Public Lottery arm, and D is a dummy variable indicating complience (instrumented by assignment into treatment). The coefficients for D $\times$ PL allow to test for Hypothesis 1a directly. I control for the lagged variable, household income, the farmer's level of education, and whether the farmer owns the selected plot. The bootstrapped standard errors are clustered at the geographical cluster level. $^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.1$.

Appendix L Post-Double LASSO

To address chance imbalance across a wider set of baseline covariates and to eliminate any discretion in control variable selection, I employ the Post-Double Selection LASSO method (PDL) developed by Belloni, Chernozhukov, and Hansen (2014), ensuring the estimates are robust to any linear or non-linear confounding from the full vector of pre-treatment characteristics. The method has three steps. First, I ran a LASSO regression to select a set of covariates that are good predictors of each index during survey 1. Then I run a second LASSO regression to select a set of covariates that are good predictors of the treatment assignment. To select the optimal penalty parameter for the LASSO regressions, I follow the recommendation of Belloni, Chernozhukov, and Hansen (2014) and employ a theory-driven, "plug-in" method. This approach uses a data-dependent formula to calculate the penalty parameter that is theoretically guaranteed to select all relevant control variables with high probability, which is the primary goal for valid causal inference, rather than optimal prediction.

Finally, I combine these two groups of covariates: I take the union of the variables selected in the previous steps and run Equations 5 and 6 with this combined set of variables as control variables. Note that each index will have a different set of covariates, which is the main advantage of the method: by selecting the most relevant predictors for each index, I maximize the precision of the estimates for that index.

	Info Adopt.	**Practices Inputs**	**Financial Manag.**	**Risk Aversion**	**Psych. Well-b**	**Attrib. LOC**	**Social Networks**
Treat (γ)	-0.062	-0.078	0.003	0.007	0.085	-0.126^{*}	-0.070
	(0.063)	(0.108)	(0.098)	(0.168)	(0.088)	(0.064)	(0.078)
PL (δ)	-0.115	-0.193^{**}	-0.172^{**}	0.157	0.139	-0.240^{***}	0.021
	(0.079)	(0.088)	(0.068)	(0.112)	(0.113)	(0.077)	(0.084)
Treat $\times$ PL (β)	0.113	0.172	0.091	-0.085	0.120	0.165	-0.132
	(0.090)	(0.128)	(0.128)	(0.220)	(0.132)	(0.116)	(0.140)
Constant (α)	-0.336^{***}	-2.188^{***}	-0.241^{***}	-0.242^{***}	-0.470^{***}	0.163^{**}	0.369^{***}
	(0.053)	(0.071)	(0.064)	(0.081)	(0.100)	(0.064)	(0.066)
Observations	701	701	701	701	701	701	701

Table A23: Survey 2 regression results as indicated in Equation 5 for summary indices. The baseline corresponds to the Computerized Randomization. PL stands for a dummy variable for the Public Lottery arm, and Treat is a dummy variable indicating whether the subject was assigned to treatment. The coefficients for Treat $\times$ PL allow to test for Hypothesis 1a directly. I control for the lagged variable and controls selected using the PDL method: the farmer's level of education and whether he works as a laborer for the psychological well-being index; whether the farmer asked for an agricultural loan in the season immediately before the intervention for the financial management index; and for the farmer's education for the adoption of new information index. The bootstrapped standard errors are clustered at the geographical cluster level. $^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.1$.

	Info Adopt.	**Practices Inputs**	**Financial Manag.**	**Risk Aversion**	**Psych. Well-b**	**Attrib. LOC**	**Social Networks**
Treat (γ)	0.031	-0.042	-0.042	0.092	0.166	-0.083	-0.179^{*}
	(0.057)	(0.069)	(0.060)	(0.127)	(0.152)	(0.078)	(0.103)
PL (δ)	0.132^{*}	0.015	0.021	0.171^{**}	0.340^{**}	-0.008	0.098
	(0.069)	(0.099)	(0.071)	(0.079)	(0.155)	(0.073)	(0.073)
Treat $\times$ PL (β)	-0.165^{**}	-0.108	0.079	-0.118	-0.291	0.125	0.097
	(0.079)	(0.086)	(0.087)	(0.177)	(0.194)	(0.129)	(0.143)
Constant (α)	-1.416^{***}	-2.360^{***}	0.835^{***}	-0.244^{***}	-0.425^{**}	0.053	0.243^{***}
	(0.049)	(0.077)	(0.075)	(0.061)	(0.155)	(0.053)	(0.036)
Observations	701	701	701	701	701	701	701

Table A24: Survey 3 regression results as indicated in Equation 5 for summary indices. The baseline corresponds to the Computerized Randomization. PL stands for a dummy variable for the Public Lottery arm, and Treat is a dummy variable indicating whether the subject was assigned to treatment. The coefficients for Treat $\times$ PL allow to test for Hypothesis 1a directly. I control for the lagged variable and controls selected using the PDL method: the farmer's level of education and whether he works as a laborer for the psychological well-being index; whether the farmer asked for an agricultural loan in the season immediately before the intervention for the financial management index; and for the farmer's education for the adoption of new information index. The bootstrapped standard errors are clustered at the geographical cluster level. $^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.1$.

	Info Adopt.	**Practices Inputs**	**Financial Manag.**	**Risk Aversion**	**Psych. Well-b**	**Attrib. LOC**	**Social Networks**
D (γ)	-0.065	-0.082	0.003	0.007	0.089	-0.132^{**}	-0.073
	(0.065)	(0.111)	(0.101)	(0.172)	(0.091)	(0.067)	(0.080)
D $\times$ PL (β)	0.119	0.181	0.095	-0.089	0.126	0.173	-0.139
	(0.092)	(0.132)	(0.132)	(0.226)	(0.136)	(0.120)	(0.144)
PL (δ)	-0.116	-0.194^{**}	-0.173^{***}	0.158	0.138	-0.241^{***}	0.023
	(0.078)	(0.087)	(0.067)	(0.111)	(0.111)	(0.076)	(0.083)
Constant (α)	-0.336^{***}	-2.188^{***}	-0.241^{***}	-0.242^{***}	-0.470^{***}	0.164^{***}	0.370^{***}
	(0.053)	(0.070)	(0.063)	(0.080)	(0.098)	(0.063)	(0.065)
Observations	701	701	701	701	701	701	701

Table A25: Survey 2 regression results as indicated in Equation 6 for summary indices. The baseline corresponds to the Computerized Randomization. PL stands for a dummy variable for the Public Lottery arm, and D is a dummy variable indicating complience (instrumented by assignment into treatment). The coefficients for D $\times$ PL allow to test for Hypothesis 1a directly. I control for the lagged variable and controls selected using the PDL method: the farmer's level of education and whether he works as a laborer for the psychological well-being index; whether the farmer asked for an agricultural loan in the season immediately before the intervention for the financial management index; and for the farmer's education for the adoption of new information index. The bootstrapped standard errors are clustered at the geographical cluster level. $^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.1$.

	Info Adopt.	**Practices Inputs**	**Financial Manag.**	**Risk Aversion**	**Psych. Well-b**	**Attrib. LOC**	**Social Networks**
D (γ)	0.033	-0.044	-0.044	0.097	0.174	-0.087	-0.188^{*}
	(0.059)	(0.071)	(0.061)	(0.132)	(0.156)	(0.080)	(0.106)
D $\times$ PL (β)	-0.173^{**}	-0.113	0.083	-0.124	-0.305	0.131	0.102
	(0.082)	(0.088)	(0.090)	(0.183)	(0.199)	(0.132)	(0.147)
PL (δ)	0.133^{**}	0.016	0.020	0.171^{**}	0.342^{**}	-0.008	0.098
	(0.068)	(0.097)	(0.069)	(0.078)	(0.152)	(0.072)	(0.072)
Constant (α)	-1.417^{***}	-2.360^{***}	0.836^{***}	-0.245^{***}	-0.426^{***}	0.054	0.243^{***}
	(0.047)	(0.075)	(0.074)	(0.060)	(0.151)	(0.053)	(0.036)
Observations	701	701	701	701	701	701	701

Table A26: Survey 3 regression results as indicated in Equation 6 for summary indices. The baseline corresponds to the Computerized Randomization. PL stands for a dummy variable for the Public Lottery arm, and D is a dummy variable indicating complience (instrumented by assignment into treatment). The coefficients for D $\times$ PL allow to test for Hypothesis 1a directly. I control for the lagged variable and controls selected using the PDL method: the farmer's level of education and whether he works as a laborer for the psychological well-being index; whether the farmer asked for an agricultural loan in the season immediately before the intervention for the financial management index; and for the farmer's education for the adoption of new information index. The bootstrapped standard errors are clustered at the geographical cluster level. $^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.1$.

Appendix M Results of the diagnosis test for separability using Bayesian bootstrap

The bootstrap procedure used in the main specification essentially assumes that the sample cumulative distribution function is the population cumulative distribution function. This means that each bootstrap replication is drawn independently from the sample cumulative distribution function. This assumption might be violated when observations are tightly connected with each other. The Bayesian bootstrap has an inherent advantage over the bootstrap with respect to the resulting inferences about parameters: the Bayesian bootstrap generates likelihood statements about parameters rather than frequency statements about statistics under assumed values for parameters (Rubin, 1981).

In A11, I show $\hat{S}$ with their confidence intervals calculated using Bayesian bootstrap for the indexes defined by me; Figure A13 shows the same analysis for the data-driven indices. Bayesian bootstrap has two main advantages: first, it delivers estimates that are smoother than the simple bootstrap due to its continuous weighting scheme. Second, the continuous weighting scheme prevents corner cases from emerging, as no observation will ever receive zero weight; this means that no problem of collinearity will arise.

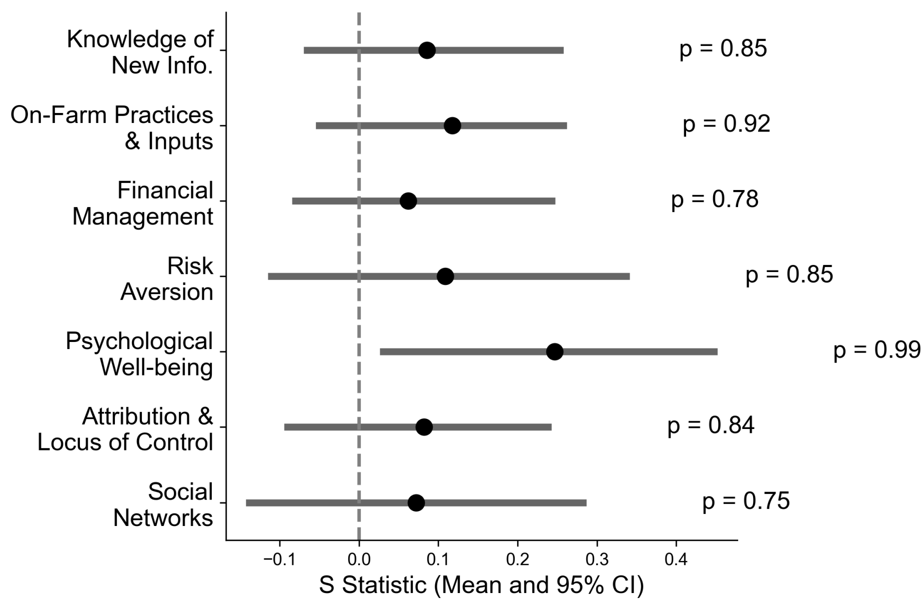


Figure A11: Results from testing the null hypothesis 9 on separability using the summary indices. The plot displays the calculated statistic S by index outcome, along with their 95% confidence intervals, where $S = T_1 + T_2 - T_3$. In the case of the On-Farm Practices & Inputs and the Financial Management indices, $S = |T_1| + |T_2| - |T_3|$, as the direction of treatment effects is uninterpretable. The p-values shown are the one-sided bootstrap p-values for $S \geq 0$ corresponding to each index; intuitively, the p-value provides the empirical probability that the null hypothesis $S \geq 0$ holds. Then, the lower the p-value, the stronger the evidence against the null hypothesis.

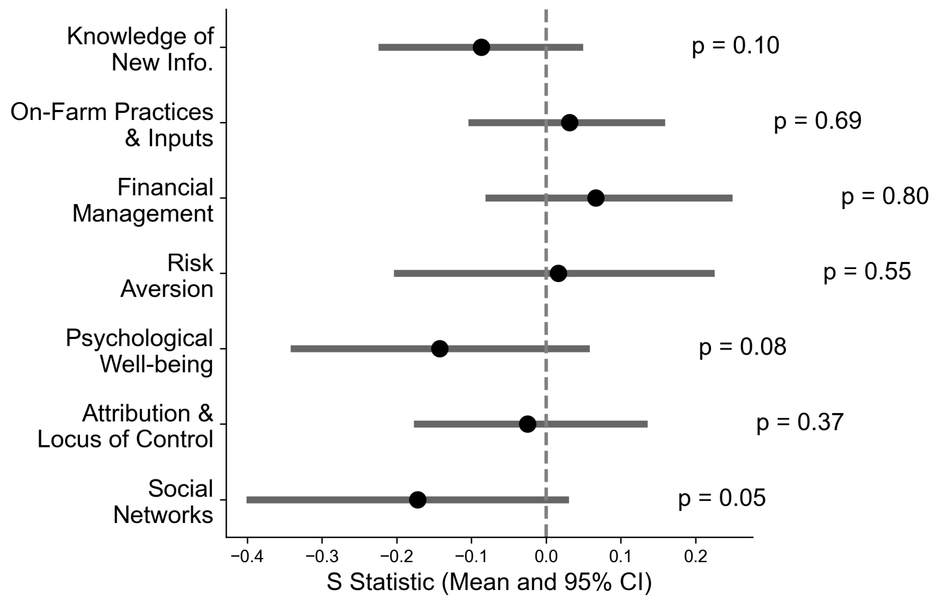


Figure A12: Results from testing the null hypothesis 10 on separability using the summary indices. The plot displays the calculated statistic S^* by index outcome, along with their 95% confidence intervals, where $S^* = T_1 + T_2 - T_3^*$. In the case of the On-Farm Practices & Inputs and the Financial Management indices, $S^* = |T_1| + |T_2| - |T_3^*|$, as the direction of treatment effects is uninterpretable. The p-values shown are the one-sided bootstrap p-values for $S^* \geq 0$ corresponding to each index; intuitively, the p-value provides the empirical probability that the null hypothesis $S^* \geq 0$ holds. Then, the lower the p-value, the stronger the evidence against the null hypothesis.

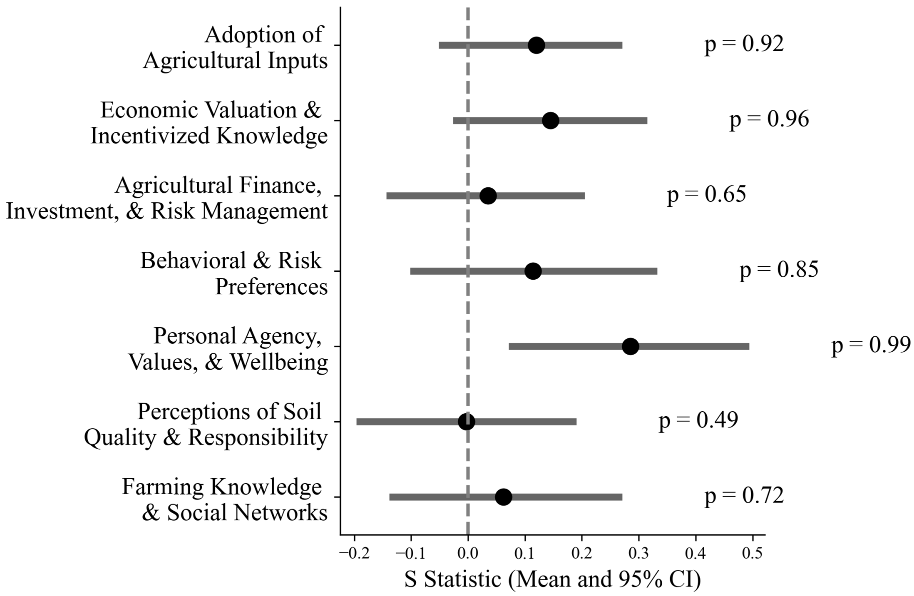
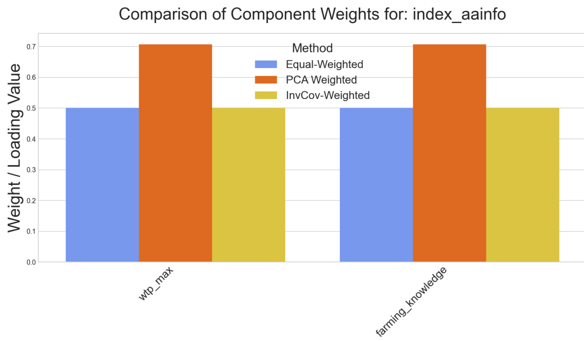
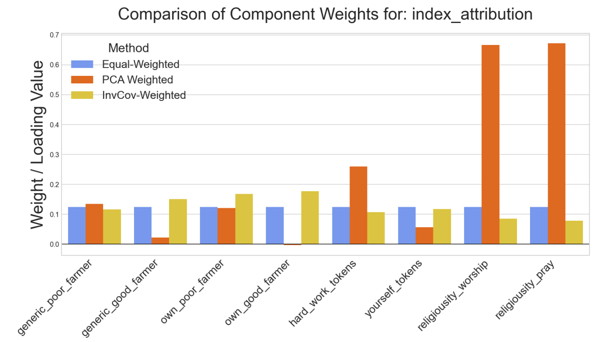


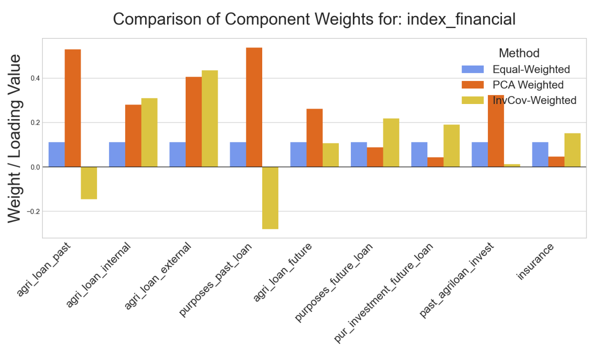
Figure A13: Results from testing the null hypothesis 9 on separability using the data-driven summary indices, Bayesian bootstrap. The plot displays the calculated statistic S by index outcome, along with their 95% confidence intervals, where $S = T_1 + T_2 - T_3$. In the case of the Adoption of Agricultural Inputs index, $S = |T_1| + |T_2| - |T_3|$, as the direction of treatment effects is uninterpretable. The p-values shown are the one-sided bootstrap p-values for $S \geq 0$ corresponding to each index; intuitively, the p-value provides the empirical probability that the null hypothesis $S \geq 0$ holds. Then, the lower the p-value, the stronger the evidence against the null hypothesis.



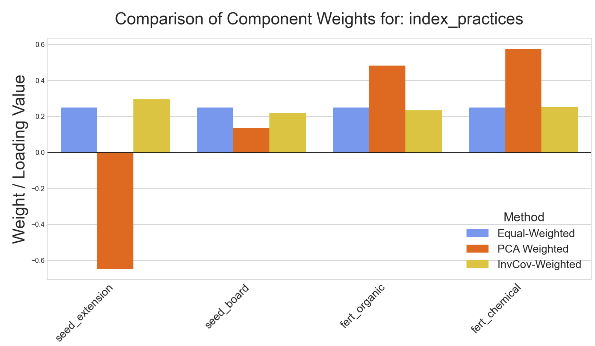
(a)



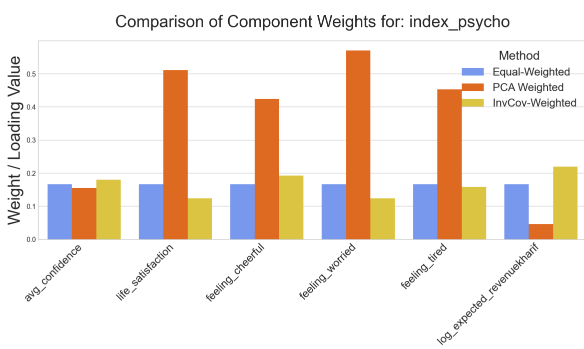
(b)



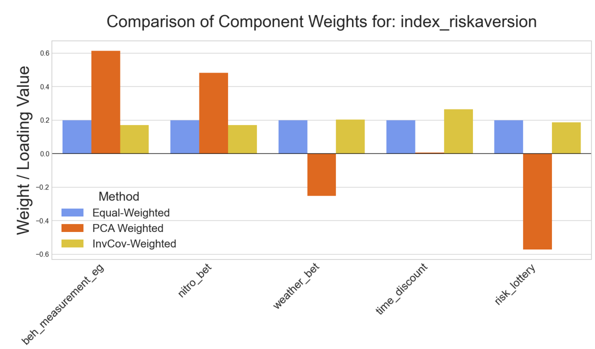
(c)



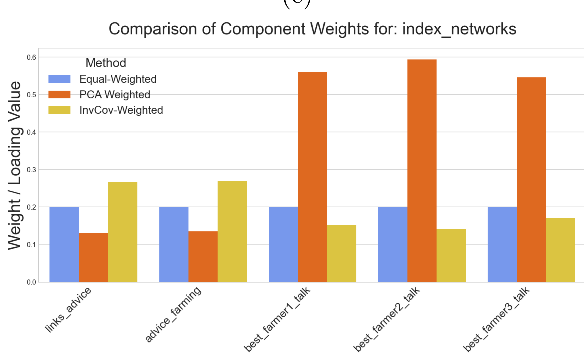
(d)



(e)



(f)



(g)

Figure A1: The plots show the weights for each variable in each index using equal weights, weights given by the first principal component, and using the inverse covariance weighting method. The plots correspond to: (a) knowledge of new Information; (b) Attribution and Locus of Control; (c) Financial Management; (d) On-Farm Practices and Inputs; (e) Psychological Well-being; (f) Risk Aversion; (g) Social Networks.