Experimenting for self and for others: farmer experimentation as a public good

Michael Ndegwa,[a,*] Hope Michelson,[b] Sarah Kariuki,[a] Mercy Mbugua,[a] Annemie Maertens,[c] Jordan Chamberlin,[a] Jason Donovan[a]

[a] International Maize and Wheat Improvement Center (CIMMYT)
[b] Department of Agricultural and Consumer Economics, University of Illinois
[c] Department of Economics, University of Sussex

11/30/2023

Abstract

Though farmer experimentation is fundamental to technology adoption, innovation, and climate adaptation, such experimentation often takes place without sufficient community awareness. The semi-private nature of this experimentation limits the positive externalities of the experimenter effort, potentially slowing learning by others. To address that problem we implement a village-level randomized controlled trial to evaluate the effects of three scalable strategies to promote farmer experimentation and learning. Across 156 Kenyan villages and 4160 farmers, we vary the proportion of farmers to whom we offer an improved hybrid maize seed; we further split the low concentration treatment arm to an incentive and non-incentive group. In one treatment arm, 10% of villagers receive 0.5 kgs of maize seed. In a second treatment arm, 10% of the farmers receive the maize plus financial compensation for their experimentation. In the third treatment arm 35% of the village farmers receive the maize seed. All participant farmers receive signage to post next to these experimental fields to promote awareness of their work. Our design will therefore test the effects of different levels of saturation within a village against a model where farmers are compensated for their experimental efforts. We study the relative effects of community experimentation and incentives on farmer experimentation, learning, yield expectations, and adoption.

Keywords: risk, agriculture, climate, technology adoption, poverty
JEL codes: D81; D25; O12; O13; Q12; Q54

Introduction

When farmers experiment with new technologies, new crops, or new markets, their experimentation is often private as well as modest in scale (Sumberg & Okali, 1997). Farmers normally bear the costs associated with experimentation, investments in time and effort as well as land and inputs (fertilizer, chemicals, labor) diverted from producing other crops.

While this private small-scale experimentation can generate positive externalities, the scope for social learning from a given farmer's experiments will depend on the amount of local heterogeneity in growing conditions and natural capital (Marenya & Barrett, 2009; McCullough et al., 2022; Richter & Babbar, 1991; Rosenzweig & Udry, 2020; Suri, 2011; Suri & Udry, 2022; Zingore et al., 2007). Social learning will also depend on how much other farmers in the region know and understand these experiments, on how much the farmer shares about what they learn, and on the degree to which other farmers actively seek out information from their peers.[1]

Even so, farmers likely under-experiment or under-provide the insights they derive from their experimentation if they fail to account for the social benefits of their efforts.

This paper evaluates efforts to increase both the scale and the visibility of farmer experimentation. We use a randomized controlled trial to study the effects of compensation for experimentation on social learning and technology adoption. Our design and data allow us to test for effects both on the experimenters and also on non-experimenting but neighboring village farmers. We contrast the effects of these payments with a saturation model, varying the share of farmer experimenters within a village. We compare two scalable models of facilitating within-village learning and adoption: experimental intensity and experimental scope.

Researchers have recently evaluated the effects of rewarding farmers for sharing information. Benyishay & Mobarak (2019) show that peer-to-peer learning increased with incentives paid to the disseminating farmers based on other farmers' gains in knowledge about the technology. Shikuku & Melesse, (2020) and Shikuku et al., (2019) show that social recognition incentives increase knowledge spillovers to other farmers and also changed information networks of both disseminating farmers and their neighbors. In contrast, Okello et al., (2023) find that social incentives reduced the likelihood of the trained progressive farmers reaching out to co-villagers to share information and discuss farming.[2]

Our approach both frames payments to farmers as compensation for the public good that they provide and pays farmers in a lump sum, distinct from any measures of effort or effects. These lump sum payments are easily scaled and do not require monitoring; they may also tap into intrinsic rather than extrinsic motivations of participants (Ryan & Deci, 2000). In a context of poor and time-constrained farmers, these modest financial contributions can effectively recognize and compensate aggregate

[1] A technology does not have to have been successful for the knowledge to benefit others. Establishing that a technology is not suitable for a given climate or market conditions can be valuable if there is sufficient similarity across farmers in terms of growing conditions and market access.

[2] Considerable research has focused on the optimal way to seed information and technologies to promote farmer learning and adoption (Beaman et al., 2021; Beaman & Dillon, 2018; Hinz et al., 2011; Kempe et al., 2005) (Bandiera et al. 2022) and economists have studied a range of frictions that can impede technological adoption and (Bandiera & Rasul, 2006; Conley & Udry, 2010; De Janvry et al., 2017; Magruder, 2018; Munshi, 2004; Suri & Udry, 2022).

effort. Experimenting farmers may be more available to share information, when approached, and may more actively seek others out to share the information.

We conduct the research in two counties in Eastern and two counties of Western Kenya, a country where encouraging and systematizing farmer experimentation is particularly salient. Farmers in Kenya have an enormous amount of choice when it comes to what maize hybrid they plan to grow. More than 350 maize varieties have been released since 1965, the majority of these in the last decade. Maize yields remain relatively stagnant in the country however (De Groote et al., 2005; Mumo et al., 2018) and most farmers report having grown the same maize hybrid for decades, even as growing conditions and better adapted seed choices have evolved significantly (Rutsaert & Donovan, 2020). Farmers have a lot of choices but little guidance. Like most countries in Sub-Saharan Africa, Kenya lacks a national maize variety trial program to provide annual performance testing data. The national varietal release systems require only that any new variety releases must be at least as productive as currently available varieties (while possibly also optimizing on other traits). Farmers report trying new hybrids on small plots but this experimentation is often fragmented and isolated within villages. For example, a woman we interviewed during focus group discussions complained that many farmers in her village were in the habit of trying a new maize seed every year. She couldn't tell what people were growing; she could not make comparisons; she could neither aggregate nor learn from what they were doing.

Across 156 Kenyan villages, we vary the share of farmers to whom we offer a choice between two recently released (less than five years in the market, available in local agri-dealer shops and suited to the local growing conditions but not yet widely adopted in the area) hybrid maize seed: either a relatively low share of farmers – in which 10% of villagers receive 0.5 kg of the improved hybrid maize seed – or a high share – in which 35% of villagers receive the seed.[3] We further split the low concentration treatment arm to incentive and non-incentive arms.

Our sample includes 4160 experimenting and non-experimenting farmers and we randomly select the experimenters so that we can test for heterogenous effects with respect to wealth and other relevant farmer attributes. We collect data on passive, active, observational, and conversational learning styles and on the behavior and adoption of non-experimenting farmers. We also collect beliefs data over the growing season to study how yield expectations evolve in time.

We contribute to literature elucidating the tradeoffs among models of encouraging experimentation and learning. Several papers have contrasted strategic seeding information within villages and social networks with what is known as broadcast (Beaman et al., 2021; Hinz et al., 2011; Matous, 2023) (Kelley et al. 2023; Banerjee et al. 2019); others have compared the effects of demonstration plots worked by community members with field days attended by outside villagers on learning and adoption (Maertens et al., 2021). This work has established the importance of both social and physical proximity to social learning (Kondylis et al., 2017 for example) and has found evidence of spillovers (Hörner et al., 2022) across farmers in communities with demonstration plots and field days.

Our design and data collection are attentive to a possibility that has received less attention in the literature: that increased scale of experimentation (plots and farmers) could have negative effects within a village. It is not a priori clear what the optimal amount of experimentation is, either for an

[3] We capped the high treatment share at 35% because we wanted to implement treatments that were scalable by government extension and/or seed companies in the future. Shares above this were considered implausible due to their cost and

individual farmer or for a village as a whole. There is some sense that more is better when it comes to experimentation; that more plots to observe or more farmers experimenting can speed up learning or increase the quality of learning in a village, contributing to adoption decisions informed by more trials in a given season and more complete assessments of yield and production outcomes. Even so, it may be possible for an individual or a village to have *too much* experimentation. Traditional Bayesian models of learning assume more information is always better, at least if the goal is learning the true signal. Although there are diminishing returns to this learning (so the tenth signal is less informative than the ninth), the marginal contribution is positive. But each experiment comes with a cost; at some stage the additional value of information may fail to exceed the marginal cost of producing it. Those calculations change fundamentally of course if the farmer includes the social benefits their experimentation generates for other farmers.

For example, more observations of a new technology in a given season might not be better if heterogenous growing conditions within a village contribute to different yield outcomes from the same technology. This relates to the model in Kondylis et al. (2023) in which belief dispersion can lead to what they term "noisy adoption". If individuals focus on the negatives – on low yield realizations – more observations could deter adoption but may provide farmers more complete information about heterogeneity in returns. Village farmers may be more likely to observe a negative outcome, especially within a village characterized by significant heterogeneity in soil quality, growing conditions, or farmer ability. If farmers anchor on these lower tail yields or if they compare themselves to farmers with low yields, they may fail to adopt technologies that would have benefitted them. Another possibility: more individuals adopting a technology (or demonstrating it) could create some pressure within a village which could be welfare reducing if there is heterogeneity in returns.

The rest of the paper proceeds as follows: Section two presents research design details including the context in which we conduct this study, our experimental design (main outcomes, hypotheses, intervention and determination of sample size and statistical power). Section three presents details on the data for this study including methods for data collection and processing as well as possible deviation from the original sample size, reasons for this and ways of dealing with it. Section four presents the analysis approaches which will be applied for the full paper including the econometric models, procedures for dealing with attrition, missing values and outliers, multiple outcome and multiple hypothesis testing and methods for assessing heterogenous effects from our treatments.

2. Research Design

A. Context

We work with small maize farmers in two counties in Eastern Kenya and two counties in Western Kenya. The context is well-suited to the research questions related to farmer learning and low technology adoption. Farmers mostly grow hybrid maize in Kenya (De Groote & Omondi, 2023), but with low varietal turnover. This means that farmers have been growing their hybrid of choice for many years; they are not switching despite the plethora of options in the markets years (De Groote & Omondi, 2023; Smale & Olwande, 2014), despite extension efforts and agronomic demonstrations indicating that newer varieties perform better in the context of emerging production challenges related to temperature and the timing and quantity of rainfall during the growing season.

Varietal turnover is important for adaptation to climate change, to sustaining and furthering yield increases, to protecting production from new stresses including pest outbreaks, new pathogens, and new temperature and rainfall patterns (Atlin et al., 2017; Chivasa et al., 2022). Governments and development and research institutions recognize the positive externalities to varietal switching; they are also concerned about stagnation in the seed sector.

Farmers face many choices but little guidance, with limited organized information about how particular varieties have performed compared with other options and in a range of growing conditions. Farmers have a lack of familiarity with new varieties and express low trust in new varieties; in particular, they express concern that the new varieties will increase production costs without increasing yields. Organized, sustained experimentation by farmers to help guide other farmers in this context could reduce information-related constraints about local performance of new seed varieties.

We selected the hybrid seed varieties offered to the treatment farmers through a careful scoping process involving discussions with experienced agro-dealers in the study areas. We visited 11 agro-dealers in Kenya's Eastern (Embu and Kirinyaga counties) and 14 in Western (Busia and Vihiga counties) to gather information on the maize hybrid seed varieties that they stocked, when sellers began stocking these varieties, and sellers' assessments of adoption levels and local performance. Insights were based on the sellers' own experimentation and on feedback sellers have received from farmers. Our hybrid seed selection criteria identified varieties with the following characteristics:

1. Relatively new with at most 5-years since their commercial release in Kenya
2. Preferred varieties in the region, as recommended by breeders and agronomists
3. Varieties are available in local markets but are not as yet widely purchased and by local farmers. Market share remains below 10% of buyers

We used these criteria to choose two varieties per region (two in Eastern and two in Western) to be presented to the treatment farmers: Duma 419 and Pan 3M-05 were selected for Eastern region and Tsavo 411 and Adv 2302W for Western region. We gave treated farmers the choice between these two varieties and we provided each farmer with 0.5 kg of the seed they selected, enough to plant one eighth of an acre.

B. Experimental Design

Primary outcomes of interest

Our baseline survey is conducted at the start of the long rains season, when farmers are making choices about the seeds they will plant and the investments they will make. The treatments are implemented in the weeks just after baseline. We measure post-treatment outcomes at both midline and endline. The midline is November 2023, in advance of the 2023 short rains season. The endline is January 2024, in advance of the 2024 long rains season. We can use the midline and endline to test how farmer decisions about seed choice and experimentation change in both the short rains and long rains season. This is important as farmers may have different predispositions to experiment in long vs shour rains and they also may face different opportunities and pressures across those seasons. We measure outcomes for host farmers (those that receive the seed and the experimental protocols and signs) as well as the non-host farmers living in the treatment villages.

- **Host farmer outcomes measured during the season with phone calls and at midline**
    - Adherence to experimental recommendations by hosts/experimenting farmers
    - Investment (labor, non-labor inputs) in experimental plot relative to non-experimental field
    - Within-farmer yield comparison across experimental seed and their status quo variety
    - Host dissemination of experimental process and results

- **Outcomes measured for host farmers and non-host farmers** residing in treatment villages. These outcomes have to do with varietal experimentation and will be measured both at midline and at endline.
    - Do farmers set up their own experimental plots during the 2023 short rains and 2024 long rains seasons? (after treatment has ended)
    - Do farmers share information about their experiments?
    - Varietal switching to anything new for the farmers. Are farmers adding new varieties to their cultivation? Here we measure any seeds that are new to the farmer, not just the varieties used in treatment.
        - Proportion of the farmer's planted varieties that are new to the farmer
        - Proportion of the farmer's maize area cultivated with new varieties
        - Weighted average age of varieties planted by households
        - Categorical varietal age – Ultra old (over 20 years), Old (11-20 years), new (6-10 years) and ultra new (5 years and below)
    - Perceptions towards new varieties relative to the established and older varieties. This will be an index variable computed from summing up scores from a 5-point scale applied across a set of 14 questions that compares new to old on multiple aspects. The perceptions survey questions are in the appendix.
    - Expected yield distributions for new varieties vis-à-vis established ones (status quo)

- **Outcomes measured for host farmers and non-host farmers** related to the varieties given out to host farmers during treatment, to be measured at midline and endline.
    - Planting the varieties used in the study
        - Proportion of the farmer's planted varieties that are the treatment varieties
        - Proportion of the farmer's maize area cultivated with the treatment varieties
        - Yield expectations for the treatment varieties

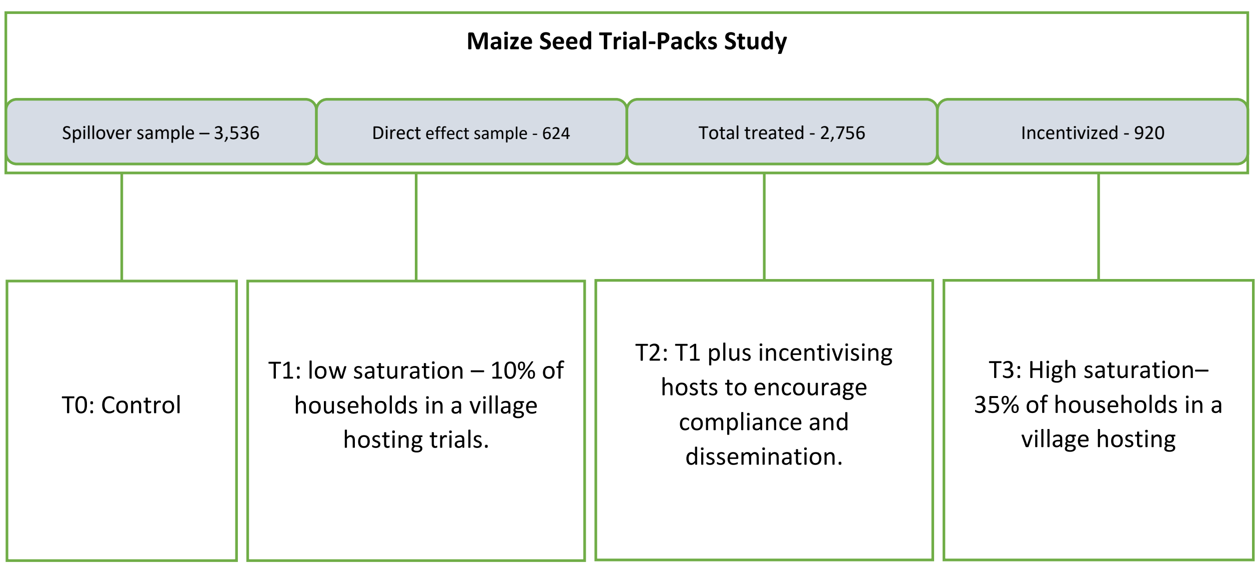
Basic methodological framework / identification strategy

The experimental design for this study is presented in Figure 1. We use a randomized controlled trial to vary the proportion of farmers to whom we offer the improved hybrid maize seed across the village. We further split the low concentration treatment arm to an incentive and non-incentive group. Our experimental design includes a control group and three treatment arms:

- T1: Low concentration in terms of the share of farmers hosting the trials in a village. For this treatment arm, we distributed 0.5 kg of maize seed to 10% of households/farmers in respective villages. All farmers received a sign to post next to their experimental field.
- T2: Low concentration + incentives to host farmers. We maintained a 10% concentration in as in T1. In addition, host farmers received Ksh 1500 (about $10 USD) in three instalments of equal amount (Ksh 500). All farmers received a sign to post next to their experimental field.
- T3: High concentration: For this treatment arm, we distributed 0.5 kg to 35% of households in treatment villages. All farmers received a sign to post next to their experimental field.
- T0: Control – comparison group where no intervention was implemented.

In all treatment villages, randomly selected trial hosts were given 0.5kg of seeds to experiment with. Selected farmers could choose between two hybrids pre-selected by the research team. Follow-up phone calls were made to all trial hosts to collect some data on the progress of their experimental plots as well as their dissemination efforts. The calls were made twice during the season (around topdressing and tasseling stages) and once after they had already harvested and possibly consumed/tasted the grain from the experimental plots. The calls coincided with the three instalments of incentives payment.

Figure 1: The experimental design



Hypotheses

Hypothesis 1: Adherence by trial hosts to experimental recommendations, especially on dissemination of information, will differ across treatments arms.

Hypothesis 2: intervention will lead to improvements across the target outcomes among the treatment groups: T1|T2|T3>T0 where T0=control group

Hypothesis 3: outcomes will differ across the first treatment arm (T1) and the other two treatment arms (T2 and T3): Test T1<T2|T3

Hypothesis 4: outcomes will differ across T2 and T3, but we have no priori expectation of the direction of the differential effects: Test T2=T3

Hypothesis 5: outcomes will vary across the host farmers **(n1)** (those directly treated/received and planted trial-packs) and the non-host farmers **(n2)** (the survey sample who did not receive trial-packs but are used to assess social learning and information spillovers): where n1>n

Intervention
Our treatments vary at the village level. Treatments are defined at the village-level in order to evaluate village-level effects on awareness, perception and demand for experimental varieties.
Randomly selected households from the treatment villages were presented with trial packs of two preselected experimental maize varieties for them to endogenously select one.

Within each treatment village, the 0.5kg seed packs were distributed to households randomly selected from village level sampling frames. Every household in the village had an equal chance of being selected as a trial host. The lottery-based distribution, coupled with random selection of households to the study, should partially alleviate concerns about potential negative impacts of freely distributed goods on subjective valuation (and consequent diminished willingness to pay for seed in post-experimental periods), because while the seed is freely distributed to lottery winners, it is not freely distributed to everyone in the community.

Seed distribution was conducted at the same time as the baseline survey and was accompanied by information about the variety and experimentation protocol. The campaign was the same across all treatments and consisted of a village/cluster level meeting for all the selected trial hosts. Trained enumerators conducted training and sensitization sessions at the meeting describing the purpose of the experiment (i.e., to facilitate village-level learning about new varieties) and norms for experimentation. In particular, farmers learned how to set up experimental plots on farms. They were advised to save a small portion of their land to plant the trial pack, an area not at high risk for animal destruction or theft, and to manage the trial field in the same way that they managed their other maize fields. They knew that the harvested grain was theirs to consume or sell as they wished. We include the protocols communicated to farmers in the appendix. All seed pack distributions were accompanied by signs and farmers were encouraged to display the sign in a visible spot next to the experimental plots they established with the trial packs. The implementation of such experimental norms and signage will not be enforced but will be monitored and recorded as part of the experiment's data collection activities.

Sample and statistical power

We have selected four Kenyan counties for the implementation of this study: Embu, Kirinyaga, Vihiga and Busia counties. We also obtained a Kenya villages map with their names and GIS coordinates and extracted villages for the four counties. After some minor cleaning to remove duplicates, we were left with 211 villages in Embu County, 230 in Kirinyaga, 294 in Vihiga and 256 in Busia. After excluding villages in urban, semi-urban and in forests areas, we have randomly selected 52 villages per county and 208 in total. We hence propose a random administration of 208 villages into 3 treatments and one control group, all of equal size (i.e., 52 villages in each group). Since we do not have data on outcomes of interest in our target population, our power calculations require assumptions. Figures 2 and 3 below show the detectable effect size for pairwise comparisons of different numbers of control and treatment groups. If we want to be able to compare each treatment against the control, as well as compare treatments against one another, then we maximize power by making each group the same size. This means that with a total of 208 candidate villages, we can work with four groups (one control and three treatment arms) of 52 villages each. We use the likelihood of purchasing a new variety in any given year as the basis for power analysis. However, data on this outcome is hard to come by, particularly for smallholders in the SSA region and other low-income countries. As such, we must assume some distributional characteristics.

To begin with, let us assume that 10% of HHs are switchers, or do purchase a new maize variety in any given year. With a fixed number of clusters, we then examine the number of households required for data collection in each village to detect a treatment-induced change from that 10%, assuming that the intra-cluster correlation is 0.2 and standard levels of acceptable type I and II errors (i.e., alpha = 0.05 and power of 80%, using a one-sided test).

We first consider an analysis of spillover effects among non-hosts that compares T1 (the lightest treatment) to the control group. Assuming baseline/control outcome of 10% switchers and MDE of 0.086 or 8.6 percentage points increase in uptake of new varieties among the treatment villages, we need 16 households per village. For the analysis of trial-host outcomes, one can assume a higher MDE since they receive the treatment directly (heavier treatment). With 52 clusters and MDE of 0.112 or 11.2 percentage points increase in uptake of new varieties among the hosts in treatment villages, we need 4 host households interviewed per village. These results are shown in Figure 2 below.

For an analysis that compares treatment groups among each other, we assume T1 as the comparison group for T2 and T3, and assume that, after the intervention, 20% of T1 will be switchers. Assuming equal sample size across experimental groups, we shall be able to detect an MDE of 0.106 (10.6 percentage points) with 16 non-host households surveyed and 0.137 (13.7 percentage points) with 4 hosts households surveyed. These results are shown in Figure 3 below.

Figure 2: detectable effect size from assumed 10% baseline share of farmers buying new varieties in any given year

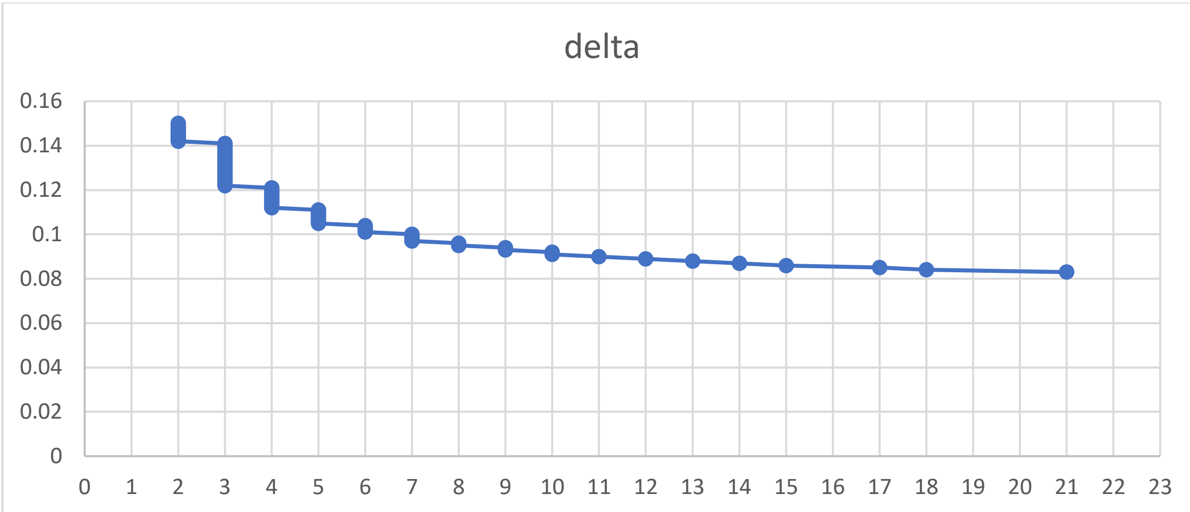
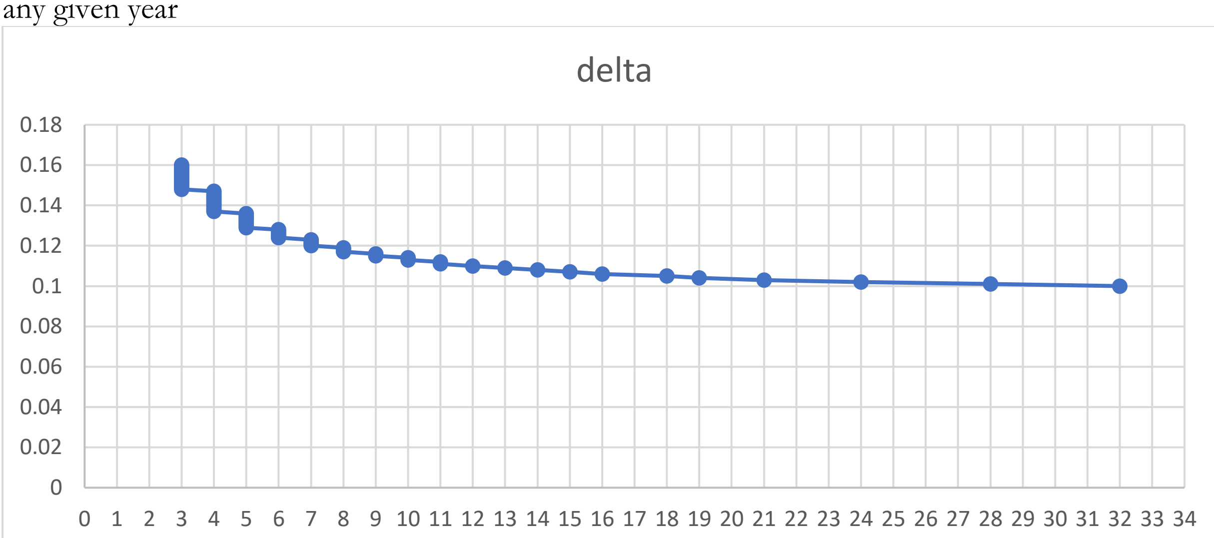


Figure 3: detectable effect size from assumed 20% baseline share of farmers buying new varieties in any given year



3. Data

Data collection and processing

Data for this study come from two main sources:

1. Three period panel of face-to-face household surveys targeting 4,160 households. Out the total sample, 1,040 are from control villages, 2,496 are from treatment villages but did not receive trial-packs (were not directly treated) while 624 are from treatment villages and received trial packs (were directly treated). The three-period panel will consist of:

- Baseline survey which was implemented jointly with trial packs distribution in February-March 2023. The baseline survey collected household demographic and socio-economic information which will be necessary for characterizing our sample, checking for treatments randomization effectiveness through balance checks and controlling for households' heterogeneity at baseline in our regression estimations to increase precision. We also collected information on our key outcomes which will we intend to include in our regression estimations as controls to increase precision.
- Midline follow-up survey which will be implemented in October-November 2023, or right after the seed-purchasing and planting period for the second season is over. All the study areas experience bimodal rainfall season with long rains season coming first between March – July and short rains season coming second between October – December. While the interventions are implemented during the long rains season, we plan to track our outcomes at the two subsequent seasons. Our interventions are more about behavioral outcomes and less about promotion of particular hybrids and we want to assess if this will affect the farmers' behavior and choices during the short rains season.
- Endline follow-up survey which will be implemented in April-May 2024 or one year after the interventions were implemented. This will be the main follow-up survey and unlike the mid-line survey, it will help assess the outcomes of interest for a season similar to one that the project interventions were implemented. The follow-up surveys, both midline and endline, will collect data on outcome measures that we listed above.

2. Three round phone surveys implemented among all the trial hosts during the March-August cropping season which is the intervention season for this study. The calls started at about three weeks after planting period with two consecutive calls scheduled at about mid the season and a few weeks after harvesting period. During the first calls, we collected basic household demographics and socio-economic information which will be used to characterize the all the hosts, check for randomization balance and control for host related outcome estimation. We also collected information regarding trial set up compliance including whether they planted or not, planting date and an assessment of how they treated trial plots in reference to their other maize fields – regarding plot selection, seeding and fertilizer application rate, intercropping, irrigation, etc. During all the three call rounds, we ask them the number of farmers they have reached out to or have shared information about the trials with since the beginning of the season and in the last past one week.

Besides the two main sources of quantitative data above, we shall also be drawing insights from a pre-RCT qualitative scoping work which included 8 FGDs with farmers and several KIIs with agro-dealers. The farmer FGDs were particularly helpful in finalizing the study design and the questionnaires while the KIIs with agro-dealers were particularly helpful in selection of maize varieties to be used for the study.

Variations from the intended sample size

This study has enrolled 20 farmers in 208 clusters (totaling 4,160 farmers - 624 trial hosts and 3,536 non-host)) for three rounds of household surveys timed to coincide with cropping seasons. We surveyed all farmers at baseline, with minimal replacements for households initially listed but were either unavailable or unwilling to participate in the study.

The study has also enlisted a separate random sample of 2,132, trial hosts with varying numbers per cluster depending on the treatment arm for respective clusters, who are taking short surveys over the phone. There are two main threats to maintaining this sample size: (1) attrition and (2) imperfect compliance.

To minimize attrition, we collected detailed contacts for all participants, including alternative contact numbers in case we miss them either at home or on the primary contact numbers. We also work with village elders who act as guides and help with identification and tracking of farmers. We will also make repeated interviews and book appointments with those who may be initially busy. Since the follow-up tool will be substantially short, we shall also conduct phone interviews as a last resort for those who are willing to participate but are either not available during the survey period or may have moved out of the study areas. We are also implementing this study in-house and we shall endeavor to maintain the same survey teams throughout the study as much as possible. Previous experiments in similar contexts have reported attrition rates of less than 10% and we also anticipate low attrition in this study as well.

Imperfect compliance could arise due to either 1) a selected host not receiving the trial packs either due to implementations hurdles or non-willingness to participate or 2) a host farmer receiving the trial pack but not setting up experimental plots. To reduce non-compliance resulting from the former cause, farmers selected as hosts were required to come to a central place to collect their pack and to receive introduction to the protocol. Farmers who could not attend but were still willing to participate were visited at home by the survey team. We, however, had few cases of non-willingness of which were replaced on the spot with randomly selected fellow villages. For the second cause of non-compliance, we made every effort to confirm that the selected farmers were maize farmers and that they were willing and able to set aside a plot enough for 0.5 kg of trial seeds. During the first round of phone surveys, we asked them if they planted the trial packs and only 0.7% indicated that they did not plant. We shall compute a village level compliance rate variable which we shall use to control for this in our main estimation models.

4. Analysis

We shall mainly rely on simple means difference (SMD) estimated with Ordinary Least Squares (OLS) and Linear Probability Models (LPM) to estimate the effect of our treatments on the continuous and binary (respectively) outcomes listed above. For robustness and consistency check of our SMD results, we shall also apply analysis of covariance (ANCOVA), and fixed effects (FE) and present the results side by side.

Hypothesis 1: adherence by trial hosts to experimental recommendations will differ across treatments arms.

Adherence/compliance to the recommendations will be measured in three different ways: 1) Planted the trials which will be a binary variable equal to one if one planted the trial seeds and zero if one did not, 2) protocol followed which will be an index variable computed from a variety of questions which seeks to understand if the farmer treated the trial plots differently from the other maize fields. This includes planting time, fertilizer application rate, seeding rate, intercropping, plot selection, and irrigation, and 3) self-reported dissemination efforts which will be a discrete variable comprising of the number of farmers/other people a host reports to have reached at three different points within the season. Data for this hypothesis is being collected using three-rounds phone surveys among the trial hosts.

We expect that adherence to the recommendations the host farmers were given during seed distribution will differ across the three treatment arms. The first treatment, low saturation group, is the most basic treatment and hence we expect higher compliance among farmers in T2 and T3 in response to the "heavier" nature of the treatments. We believe the incentives in T2 should produce differential effect from T1, especially on dissemination. Incentivizing host farmers could encourage them to share information among their counterparts, which could then improve the effectiveness of trial packs model, even with low concentration. We also anticipate increased peer-to-peer learning and competition with higher saturation in T3 to yield higher adherence and dissemination outcomes in comparison to low saturation in T1. We will use regression equation 1 to compare adherence across treatment arms:

$$\forall_{i,j} = \alpha + \beta_1 T2_{ij} + \beta_2 T3_{ij} + \sum_{i=1}^{K} \beta_k X_{i0} + \pi_t + \varepsilon_j + \varepsilon_{ij} \qquad (1)$$

Where $\forall_{i,j}$ represents adherence for household i in village/cluster j, $T2_j$ and $T3_j$ are village/cluster level treatment indicators equal to one if a household was randomly assigned to the respective treatment group (as defined above) and zero if in T1, X_{i0} represents a vector of farmers baseline socio-economic characteristics, π_t is the panel period indicator equal to zero at baseline, one at midline and two at endline, ε_j is the cluster level error term and $\varepsilon_{i,j}$ is the farmer/individual idiosyncratic error term. The parameters of interest are β_1 and β_2 which capture the treatment effects. These coefficients will reveal whether T2 and T3 effects were substantially and significantly different from T1 effects. We shall then use coefficient comparison approach to explore whether T2 and T3 are significantly different.

Hypothesis 2: intervention will lead to improvements on the target outcomes among the treatment groups: T1|T2|T3>T0 where T0=control group… *add cluster level percentage of compliers in the model*

In reference to the control group, we expect our interventions to yield some benefits among the treated villages and households in regard to awareness and perception of new maize varieties as well as experimentation and uptake of newer varieties. We will use regression equation 2 to estimate the effects of our treatments on a set of target outcomes:

$$Y_{i,j} = \alpha + \beta_1 T1_{ij} + \beta_2 T2_{ij} + \beta_3 T3_{ij} + \beta_4 Y_{ij,0} + \sum_{i=1}^{K} \beta_k X_{i,0}^{k} + \pi_t + \varepsilon_j + \varepsilon_{ij} \qquad (2)$$

Where $Y_{i,j}$ is a vector of the observed outcome variables for household i in village/cluster j post treatment, $T1_j$, $T2_j$ and $T3_j$ are village/cluster level treatment indicators equal to one if a village was randomly assigned to the respective treatment group (as defined above) and zero otherwise, π_t is the panel period indicator equal to zero at baseline, one at midline and two at endline survey, ε_j represents cluster/village fixed effects, and $\varepsilon_{i,j}$ is the idiosyncratic error term for household i in village j. To enhance precision in our estimation, we also include $Y_{ij,0}$ which represents the baseline measure of the outcome variable and $X_{i,0}^{k}$ which represents baseline socioeconomic covariates chosen through double post-lasso procedure as defined by Belloni et al. (2014), The parameters of interest are β_1, β_2 and β_3 which respectively capture the treatment effects for the three treatment arms.

Hypothesis 3: outcomes will differ across the first treatment arm (T1) and the other two treatment arms (T2 and T3): Test T1<T2|T3

We expect that the layered treatments (T2 and T3) will yield higher benefits than the most basic treatment arm namely low saturation (T1). We believe the host farmers in T2 who received a monetary incentive to encourage them to comply to the protocol will put more effort in disseminating information about the new varieties and experimentation in general. This follows findings by (Benyishay & Mobarak, 2019). We will use the coefficients comparison approach to test this hypothesis, where, from Equation 2, we will examine if β_2 is substantially and significantly greater than β_1:

$$\beta_1 T1_{ij} < \beta_2 T2_{ij}$$

Similarly, the higher saturation in terms of proportion of farmers planting trial plots in T3 villages should lead to higher effects compared to low saturation in T1. We expect the differential effect to come from the fact that non-host farmers in T3 have more sources to learn from and be influenced by which should increase learning and awareness of new varieties as well as improve their perception towards them and eventually encourage experimentation with and adoption of the new varieties. With higher saturation, besides an increase in absolute number of learning points, it increases the chances of receiving information from a source (someone) that is trusted by the recipient. This also increases the chances of non-hosts receiving information from peers who are both socially and physically proximate to them. Further, existing literature shows that household and farm heterogeneity affect how farmers learn from each other, where people learn more from peers who they identify with or who they think their farm and economic conditions are similar to theirs (Benyishay & Mobarak, 2019; Berazneva et al., 2023; Kondylis et al., 2017; Matous, 2023) The higher saturation treatment arm increases the probability of such matches. To test this hypothesis, we will examine if β_3 in Equation 2 is substantially and significantly greater than β_1:

$$\beta_1 T1_{ij} \ < \ \beta_3 T3_{ij}$$

Hypothesis 4: outcomes will differ across T2 and T3: Test T2=T3

We hypothesize that the effects of low saturation with incentives treatment arm (T2) will be significantly different from the high saturation treatment arm (T3) but we have no priori expectation of the direction of the differential effects. We hence explore the appropriate strategy to achieve scale: is it low saturation combined with incentives or just high saturation. To test this hypothesis, we will examine if β_3 in Equation 2 is substantially and significantly different than β_2:

$$\beta_2 T2_{ij} = \beta_3 T3_{ij}$$

Hypothesis 5: outcomes will vary across the host farmers **(n1)** and the non-host farmers **(n2)**

The study design consists of farmers who were directly treated within treatment villages and a random sample from the same villages who did not receive the trial packs but are used to assess the spillovers. We anticipate substantially larger effects on those who are directly treated (n1=624) compared to their village mates who did not receive trial packs (n2=2,496). To examine this, we shall run separate regression models for all outcomes (apart from host farmers outcomes) splitting hosts and non-host samples and for the entire (pooled) sample. We shall then compare the magnitude and statistical significance of treatment effects across the three models for each outcome.

Procedures for dealing with attrition, missing values and outliers

To deal with attrition at mid- and end-line surveys, we will: 1) make extra efforts as mentioned above to reach almost all the respondents, 2) test for non-random attrition or check for attrition bias in our sample. We will generate a binary indicator for each follow-up round (midline and endline) equal to one if a household dropped from the study at the respective survey round and 0 otherwise. We will then estimate a series of probit models with this indicator as the dependent variable to analyze association between attrition and (i) random assignment to control and treatment arms, (ii) outcome variables, and (iii) socioeconomic control variables for the baseline sample. This will be done separately for each experimental group and then for the pooled/overall sample, 3) use bounds approach to examine the robustness of our results to attrition, if any (Lee, 2009; Tauchmann, 2014).

We were able to reach our target sample size at baseline, albeit with minimal replacements as explained above. While analyzing the RCT, we shall dummy out any missing baseline data for the selected covariates. We will create dummy variables for all covariates with missing data which will be equal to one if missing and zero if non missing, and then replace the missing values with zero across all the baseline covariates, and include both the original variables and their respective dummies as controls in our estimation models.

To deal with large outliers, we will winsorize our continuous outcome variables at 98th percentile.

Multiple outcome and multiple hypothesis testing

In this study, we estimate the treatment effects of three treatment arms on several outcomes translating to multiple hypotheses tests. This raises the false discovery rate (FDR) concern where significant coefficients may emerge by chance when there are a large number of measured outcomes and tested hypotheses, even when there are no true treatment effects on the outcomes. To check for robustness of our results to this potential biasness, we will conduct multiple hypotheses correction tests using sharpened q-values (Anderson, 2008; Benjamini et al., 2006). Further, following Young (2019), we will conduct an F-test for all our outcome regression estimations to test the null hypothesis that the joint effect of the three treatments is zero.

Heterogeneous Effects

We have collected baseline information on households' demographic and socio-economic characteristics which we shall use to check for heterogeneity in treatment effects. As Chernozhukov et al (2023) advises, we are not restricting ourselves to examine treatment effects heterogeneity to a limited number of pre-determined subgroups as that amounts to throwing away a large amount of potentially valuable information. Instead, we aim to use the data to discover ex-post whether there is any relevant heterogeneity in treatment effect by covariates. However, to avoid overfitting, we shall use their (Chernozhukov et al., 2023) generic machine learning approach for predicting and making inference on heterogeneous treatment effects.

Bibliography

Anderson, M. L. (2008). 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. https://doi.org/10.1198/016214508000000841

Atlin, G. N., Cairns, J. E., & Das, B. (2017). Rapid breeding and varietal replacement are critical to adaptation of cropping systems in the developing world to climate change. *Global Food Security*, *12*, 31–37. https://doi.org/10.1016/J.GFS.2017.01.008

Bandiera, O., & Rasul, I. (2006). Social networks and technology adoption in Northern Mozambique. *Economic Journal*, *116*(514), 869–902. https://doi.org/10.1111/J.1468-0297.2006.01115.X

Beaman, L., BenYishay, A., Magruder, J., & Mobarak, A. M. (2021). Can Network Theory-Based Targeting Increase Technology Adoption? *American Economic Review*, *111*(6), 1918–1943. https://doi.org/10.1257/AER.20200295

Beaman, L., & Dillon, A. (2018). Diffusion of agricultural information within social networks: Evidence on gender inequalities from Mali. *Journal of Development Economics*, *133*, 147–161.

Belloni, A., Chernozhukov, V., & Hansen, C. (2014). High-Dimensional Methods and Inference on Structural and Treatment Effects. *Journal of Economic Perspectives*, *28*(2), 29–50. https://doi.org/10.1257/JEP.28.2.29

Benjamini, Y., Krieger, A. M., & Yekutieli, D. (2006). Adaptive linear step-up procedures that control the false discovery rate. *Biometrika*, *93*(3), 491–507. https://doi.org/10.1093/BIOMET/93.3.491

Benyishay, A., & Mobarak, A. M. (2019). Social learning and incentives for experimentation and communication. *Review of Economic Studies*, *86*(3), 976–1009. https://doi.org/10.1093/RESTUD/RDY039

Berazneva, J., Maertens, A., Mhango, W., & Michelson, H. (2023). Paying for agricultural information in Malawi: The role of soil heterogeneity. *Journal of Development Economics*, *165*, 103144.

Chernozhukov, V., Demirer, M., Duflo, E., & Fernández-Val, I. (2023). *Generic Machine Learning Inference on Heterogeneous Treatment Effects in Randomized Experiments, with an Application to Immunization in India* (No. 24678; NBER Working Papers).

Chivasa, W., Worku, M., Teklewold, A., Setimela, P., Gethi, J., Magorokosho, C., Davis, N. J., & Prasanna, B. M. (2022). Maize varietal replacement in Eastern and Southern Africa: Bottlenecks, drivers and strategies for improvement. *Global Food Security*, *32*, 100589. https://doi.org/10.1016/J.GFS.2021.100589

Conley, T. G., & Udry, C. R. (2010). Learning about a New Technology: Pineapple in Ghana. *American Economic Review*, *100*(1), 35–69. https://doi.org/10.1257/AER.100.1.35

De Groote, H., & Omondi, L. B. (2023). Varietal turn-over and their effect on yield and food security – Evidence from 20 years of household surveys in Kenya. *Global Food Security*, *36*. https://doi.org/10.1016/J.GFS.2023.100676

De Groote, H., Owuor, G., Doss, C. R., Ouma, J. O., Muhammad, L., & Danda, M. K. (2005). The Maize Green Revolution in Kenya Revisited. *EJADE: Electronic Journal of Agricultural and*

Development Economics, *2*(1), 32–49. https://doi.org/10.22004/AG.ECON.110143

De Janvry, A., Macours, K., Sadoulet, E., Alain De Janvry, K., & Macours, E. S. (2017). *Learning for Adopting: Technology Adoption in Developing Country Agriculture* (A. de Janvry, K. Macours, & E. Sadoulet (eds.)). FERDI. https://doi.org/10.13039/501100001665

Hinz, O., Skiera, B., Barrot, C., & Becker, J. U. (2011). Seeding strategies for viral marketing: An empirical comparison. *Journal of Marketing*, *75*(6), 55–71. https://doi.org/10.1509/jm.10.0088

Hörner, D., Bouguen, A., Frölich, M., & Wollni, M. (2022). Knowledge and Adoption of Complex Agricultural Technologies: Evidence from an Extension Experiment. *The World Bank Economic Review*, *36*(1), 68–90. https://doi.org/10.1093/WBER/LHAB025

Kempe, D., Kleinberg, J., & Tardos, É. (2005). Influential nodes in a diffusion model for social networks. *International Colloquium on Automata, Languages, and Programming*, 1127–1138. https://doi.org/10.1007/11523468_91

Kondylis, F., Loeser, J. A., Mobarak, M., Jones, M. R., & Stein, D. (2023). *Learning from Self and Learning from Others: Experimental Evidence from Bangladesh* (No. 10545; Policy Research Working Paper).

Kondylis, F., Mueller, V., & Zhu, J. (2017). Seeing is believing? Evidence from an extension network experiment. *Journal of Development Economics*, *125*, 1–20. https://doi.org/10.1016/J.JDEVECO.2016.10.004

Lee, D. S. (2009). Training, Wages, and Sample Selection: Estimating Sharp Bounds on Treatment Effects. *The Review of Economic Studies*, *76*(3), 1071–1102. https://doi.org/10.1111/J.1467-937X.2009.00536.X

Maertens, A., Michelson, H., & Nourani, V. (2021). How Do Farmers Learn from Extension Services? Evidence from Malawi. *American Journal of Agricultural Economics*, *103*(2), 569–595. https://doi.org/10.1111/AJAE.12135

Magruder, J. R. (2018). An Assessment of Experimental Evidence on Agricultural Technology Adoption in Developing Countries. *Annual Review of Resource Economics*, *10*, 299–316. https://doi.org/10.1146/ANNUREV-RESOURCE-100517-023202

Marenya, P. P., & Barrett, C. B. (2009). Soil quality and fertilizer use rates among smallholder farmers in western Kenya. *Agricultural Economics*, *40*(5), 561–572.

Matous, P. (2023). Male and stale? Questioning the role of "opinion leaders" in agricultural programs. *Agriculture and Human Values*, *1*, 1–16. https://doi.org/10.1007/S10460-023-10415-9/TABLES/3

McCullough, E. B., Quinn, J. D., & Simons, A. M. (2022). Profitability of climate-smart soil fertility investment varies widely across sub-Saharan Africa. *Nature Food*, *3*, pages275–285.

Mumo, L., Yu, J., & Fang, K. (2018). Assessing Impacts of Seasonal Climate Variability on Maize Yield in Kenya. *International Journal of Plant Production*, *12*(4), 297–307. https://doi.org/10.1007/S42106-018-0027-X/METRICS

Munshi, K. (2004). Social learning in a heterogeneous population: technology diffusion in the Indian Green Revolution. *Journal of Development Economics*, *73*(1), 185–213. https://doi.org/10.1016/J.JDEVECO.2003.03.003

Okello, J., Shikuku, K. M., Lagerkvist, C. J., Rommel, J., Jogo, W., Ojwang, S., Namanda, S., & Elungat, J. (2023). Social incentives as nudges for agricultural knowledge diffusion and willingness to pay for certified seeds: Experimental evidence from Uganda. *Food Policy*, *120*, 102506. https://doi.org/10.1016/J.FOODPOL.2023.102506

Richter, D. D., & Babbar, L. I. (1991). Soil Diversity in the Tropics. *Advances in Ecological Research*, *21*, 315–389.

Rosenzweig, M. R., & Udry, C. (2020). External Validity in a Stochastic World: Evidence from Low-Income Countries. *The Review of Economic Studies*, *87*(1), 343–381.

Rutsaert, P., & Donovan, J. (2020). Sticking with the old seed: Input value chains and the challenges to deliver genetic gains to smallholder maize farmers. *Outlook on Agriculture*, *49*(1), 39–49. https://doi.org/10.1177/0030727019900520

Ryan, R. M., & Deci, E. L. (2000). Self-determination theory and the facilitation of intrinsic motivation, social development, and well-being. *American Psychologist*, *55*(1), 68–78. https://doi.org/10.1037/0003-066X.55.1.68

Shikuku, K. M., & Melesse, M. B. (2020). Networks, incentives and technology adoption: evidence from a randomised experiment in Uganda. *European Review of Agricultural Economics*, *47*(5), 1740–1775. https://doi.org/10.1093/ERAE/JBAA009

Shikuku, K. M., Pieters, J., Bulte, E., & Läderach, P. (2019). Incentives and the Diffusion of Agricultural Knowledge: Experimental Evidence from Northern Uganda. *American Journal of Agricultural Economics*, *101*(4), 1164–1180. https://doi.org/10.1093/AJAE/AAZ010

Smale, M., & Olwande, J. (2014). Demand for maize hybrids and hybrid change on smallholder farms in Kenya. *Agricultural Economics*, *45*(4), 409–420. https://doi.org/10.1111/AGEC.12095

Sumberg, J., & Okali, C. (1997). Farmers' Experiments. In *Farmers' Experiments*. Lynne Rienner Publishers. https://doi.org/10.1515/9781685858094/HTML

Suri, T. (2011). Selection and Comparative Advantage in Technology Adoption. *Econometrica*, *79*(1), 159–209. https://doi.org/10.3982/ECTA7749

Suri, T., & Udry, C. (2022). Agricultural Technology in Africa. *Journal of Economic Perspectives*, *36*(1), 33–56. https://doi.org/10.1257/JEP.36.1.33

Tauchmann, H. (2014). Lee (2009) treatment-effect bounds for nonrandom sample selection. *The Stata Journal*, *14*(4), 884–894.

Young, A. (2019). Channeling Fisher: Randomization Tests and the Statistical Insignificance of Seemingly Significant Experimental Results. *The Quarterly Journal of Economics*, *134*(2), 557–598. https://doi.org/10.1093/qje/qjy029

Zingore, S., Murwira, H. K., Delve, R. J., & Giller, K. E. (2007). Soil type, management history and current resource allocation: Three dimensions regulating variability in crop productivity on African smallholder farms. *Field Crops Research*, *101*(3), 296–305. https://doi.org/10.1016/J.FCR.2006.12.006

APPENDIX 1: The experimentation protocol with instructions for host farmers

Protocol – Maize Hybrid Varietal Experimentation

Thank you for agreeing to participate in a maize hybrid varieties experimentation with the International Maize and Wheat Improvement Center (CIMMYT).

In Feb 2023, you will receive:

- A trial pack (500g) of a relatively new hybrid maize variety (Either: ADVANTA 2308W, TSAVO 4141, PAN 3M-05, DUMA 419)
- A signage to label the trial field

Instructions

- Please save a small portion of your land to plant the trial pack (ADVANTA 2308W), during March 2023 rain season.
- The maize trial field should not be at high risk of animal destruction or theft.
- Manage the trial field, the same way you manage your other maize fields.
- Plant the trial pack in a field where other villagers can easily observe.
- You should share information about ADVANTA 2308W and its performance with other farmers within your village.
- The harvested grain is yours to eat or sell after completion of the study.

Appendix 2: Questions used to compute the perception index

No.	Question How much do you disagree/agree with these statements	1.Strongly disagree 2.Disagree 3.Neutral 4.Agree 5.Strongly agree
1	Food made from new varieties are tastier than food made from old varieties	
2	Old varieties yield more than new varieties	
3	Maize grains harvested from new varieties are denser than grains harvested from old varieties	
4	Produce from old varieties experience less damage by pests and molds post-harvest	
5	Produce from new varieties is less infested with aflatoxins than produce from old varieties	
6	Old varieties are more drought resistant than new varieties	
7	New varieties are more drought avoidant (mature early) than old varieties	
8	Old varieties are more disease resistant than new varieties	
9	New varieties are more pest resistant (at farm) than old varieties	
10	Seeds of new varieties are cheaper than seeds of old varieties	
11	New varieties are more labor intensive than old varieties	
12	Old varieties are more input-intensive than new varieties – inputs include fertilizer, chemicals, etc	
13	Produce from old varieties draws better prices than produce from new varieties	
14	If there were some quality concerns on some seeds in the market, the quality issues are more likely to be with old varieties than new varieties	