Demand and supply factors constraining the emergence and sustainability of an efficient seed system: Endline report

Bjorn Van Campenhout*, Proscovia Renzaho Ntakyo†, Robert Sparrow‡, David J Spielman§, Caroline Miehe¶

November 3, 2022

Abstract

Agricultural technology remains under-adopted among smallholder farmers in Sub-Saharan Africa. We investigate how the variation in (perceptions of) the quality of an agricultural technology—improved maize seed—affects its adoption. Using a randomized controlled trials among agro-input dealers and smallholder farmers in Uganda, we test three hypotheses. In a first hypothesis, quality issues that constrain uptake are caused by knowledge gaps at the level of agro-input dealers; a training on proper storage and handling of seed at the agro-input dealer level is expected to increase adoption on the farm. A second hypothesis conjectures that information asymmetries similar to the lemons problem constrain technology adoption. We test this hypothesis by setting up an information clearing house that relies on crowd-sourcing to make hidden features of the technology visible to both agro-input dealers and farmers. We also test a behavioral explanation, whereby we assume that farmers attributing poor outcomes to the wrong causes. In this report, we present results obtained from the endline survey that was collected two agricultural seasons after the interventions took place.

keywords: seed systems, information clearinghouse, learning failures, information, input quality, agricultural technology adoption

JEL codes: O13; Q12; Q16; D82; D83

*Development Strategy and Governance Division, International Food Policy Research Institute, Leuven, Belgium - corresponding author: b.vancampenhout@cgiar.org

†National Agricultural Research Organisation (NARO), Uganda

‡Development Economics Group, Wageningen University and Research, Wageningen, The Netherlands

§Development Strategy and Governance Division, International Food Policy Research Institute, Kigali, Rwanda

¶LICOS Centre for Institutions and Economic Performance, KULeuven, Leuven, Belgium

1 Introduction

One of the most effective ways to increase agricultural productivity is through the adoption of improved agricultural technologies and practices. These include mechanization, but the Green Revolution has demonstrated that the largest gains can be expected from improved inputs such as inorganic fertilizers and high yielding cultivars. However, technology adoption remains lower than projected, particularly among the poor in sub-Saharan Africa (Gollin, Morris, and Byerlee, 2005). As a result, differences in yields between sub-Saharan Africa and areas that experienced a green revolution have nearly doubled since 1961 (Magruder, 2018). To reduce this yield gap, it is important to identify the drivers of, and constraints to, technology adoption.

There are various reasons why some people adopt new technologies while other stick to what they have been doing for ages. Key constraints to technology upgrading include poor access to information (Van Campenhout, 2021; Ashraf, Giné, and Karlan, 2009), procrastination and time-inconsistent preferences (Duflo, Kremer, and Robinson, 2011), heterogeneity in the net benefits to the technology due to high transaction costs (Suri, 2011), the lack of access to insurance (Karlan et al., 2014), and learning failures (Hanna, Mullainathan, and Schwartzstein, 2014).

In this study, we investigate quality related considerations about the technology as a particular constraint to adoption, a topic that has received considerable attention recently (Bold et al., 2017; Michelson et al., 2021). Indeed, as the quality of a technology (like improved seed or fertilizer) is not readily observable and can only be (imperfectly) inferred ex-post, inconsistent quality will lead to lower adoption than would be the case if quality would be consistent and/or perfectly observable. Our study tests three hypotheses to assess the relative importance of potential sources of these (perceived) quality issues. We do this using a field experiment that involves agro-input dealers and smallholder farmers in Uganda, where the focus is on improved maize seed.

In a first hypothesis, quality concerns that constrain uptake of the technology are caused by knowledge gaps at the level of the agro-input dealers, who are assumed to lack awareness and skills related to proper storage and handling of improved agro-inputs. To test this hypothesis, we implement and intensive training program that is expected to lead to improved maize seed quality sold by agro-input dealers, in turn increasing adoption by farmers. A second hypothesis conjectures that information asymmetries between seller and buyer with respect to the quality of seed—a classic lemons technology—leads to under-adoption. To test this hypothesis, we implement a crowd-sourced information clearinghouse similar to yelp.com, where we ask farmers about their experience with agro-input dealers, aggregate this data into rankings and feed this data back to both farmer and agro-input dealers. A third hypothesis targets farmers directly, as sub-optimal adoption is assumed to be caused by a particular type of learning failures: As farmers have to pay more for improved seed, this may also lead to altered expectations and behavior. In particular, we argue that farmers may expect too much from seed and as a result exert less effort than they normally would, use less complementary inputs, and plant the improved seed on sub-optimal plots. Disappointing outcomes may then be attributed to poor input quality. An ICT-mediated information campaign that stresses that even when using improved seed is is important to pay attention to all input dimensions is implemented to test this hypothesis.

This report documents impact at endline after two full agricultural seasons had passed since the intervention. A similar report was written at the time of midline data collection, looking at impact after one agricultural season. Both reports closely follow a mock report, that was prepared after baseline data was collected, and well as a pre-analysis plan, that was created at the time the study was conceptualized and designed.

We find that XXX.

The report starts with an outline of the research questions and hypotheses. The interventions that were used to test the hypotheses are explained in the subsequent section. This was done using a field experiment that is presented in Section 4. Next, there is a section on data, that presents characteristics of farmers and agro-input dealers are baseline, demonstrates balance between treatment groups, and looks at attrition. We then present the results from the impact analysis of the endline data.

2 Research Questions and Hypotheses

Seed quality issues, or the perception thereof, may arise at different stages in the seed supply chain. Poor seed quality may occur as a result of input dealer practices. This may be unintentional when the agro-input dealers is unaware of propper handling and storage practices, or intentional if the agro-input dealer tries to reduce cost by for example mixing poor quality seed with good quality seed. The problem may also be situated at the level of the smallholder farmer. For instance, a farmer may lack confidence in the input dealer or his/her products, and the nature of the input may make it difficult for the farmer to objectively assess quality. It may also be that the farmer wrongly attributes poor outcomes caused by factors other than seed quality to seed quality.

We test potential quality related reasons for low uptake of improved maize seed. One has its origins at the agro-input dealer level and is cognitive in nature. A second reason is situated at the interaction between agro-input dealer and farmer and tests a classic economic explanation. A third potential reason is behavioral and happens in the farm. We discuss each research hypothesis in detail below. In the next section we present the interventions that we will use to test these hypotheses.

H1: Seed is of poor quality due to poor handling and storage at the input dealer level

Lack of information is pervasive in developing countries and often leads to sub-optimal outcomes for the rural poor. As a result, a simple piece of information can make a big difference (Duflo and Banerjee, 2011). Also in the context of agricultural technology adoption among smallholders, knowledge gaps have been identified as a key constraint, and governments around the world invest in public agricultural advisory services (Anderson and Feder, 2004). While the need for policies and interventions that strengthen input marketing capacity and infrastructure has been acknowledged decades ago (Tripp and Rohrbach, 2001), we find few examples of studies that look at knowledge gaps at the input dealer level.

The first hypothesis asserts that poor handling and storage at the level of the input dealer may lead to poor seed quality, in turn reducing the profitability of seeds at the farmer level, resulting in low adoption. There is indeed some evidence of input quality reduction at this level. In a comprehensive study of the seed supply chain in Uganda, Barriga and Fiala (2020) document various issues related to handling and storage that may reduce the quality of the input. For example, farmers often need smaller quantities than what is in the standard bags, and input dealers thus often repackage in smaller bags in sub-optimal environments. Poor rotation of seed stock and storage in open bags in moist conditions or in direct sunlight also reduce seed quality.

H2: Mismatch between seed quality delivered by agro-input dealers and seed quality perceived by farmers

The second hypothesis focuses on the information asymmetry between seed sellers and seed buyers. As argued in Bold et al. (2017), the market for seed in Uganda appears similar to the market for used cars as described in Akerlof's classic study (1970). In such a market, the quality of goods can degrade in cases where the quality is known by the seller, but not (yet) by the buyer. This problem can be solved by reducing information asymmetries between the two parties.

Information asymmetry implies that agro-input dealers are better informed about the quality of the products they sell than the farmers they sell to. This means agro-input dealers can compromise on quality to cut costs. In Uganda, there are some indications that adulteration happens at some point in the seed value chain. Bold et al. (2017) find that hybrid maize seed contains less than 50% authentic seed, while Ashour et al. (2019) find that nearly one in three bottles of herbicide contains less than 75% of the labeled concentration of the active ingredient.

H3: Seed is of good quality but have inflated expectations and face learning challenges

In the context of new agricultural technology, production functions are not known. Farmers learn from own experience (Foster and Rosenzweig, 1995) as well as from observing the experience of others (Conley and Udry, 2010). Learning involves an iterative process of forming and updating beliefs about yield or profit distributions. Many researchers have addressed how individuals process information and update beliefs when making repeated decisions (e.g. Camerer and Hua Ho, 1999). Barham et al. (2015) analyze how learning heuristics vary across farmers and how they affect technology adoption decisions. Gars and Ward (2019) test whether farmers' learning heterogeneity is a barrier to adoption. They find that even though Bayesian learning is well suited to learn about hybrid rice, it is also more cognitively demanding, such that only 25 percent of farmers can be characterized as pure Bayesian learners while 40 percent rely on first impressions. Present-biased learning and relying on first impressions is likely to hinder technology adoption.

Erroneous perceptions and false beliefs at the farmer level may complicate learning and affect technology uptake. For instance, high yielding varieties may be less resistant to particular pests and diseases or to droughts than local maize varieties that farmers in a particular area selected themselves over the course of centuries. Therefore, additional inputs such as pesticides, insecticides and irrigation may be needed to bring the seed to its full potential. Worse, farmers they may think that improved seed is a guarantee for higher yield, leading them to reduce management and use of other inputs. This may lead to disappointing yields, and farmers may erroneously attribute these low returns to poor input quality, which may lead to dis-adoption. The problem may thus be rooted in negative experiences which conflate low product quality with incorrect management practices and can be characterized as a learning failure. Consistent with this, Michelson et al. (2021) find that fertilizers in Tanzania meet the requisite quality standards even though Tanzanian farmers persistently believe that the fertilizer they purchase from the market is adulterated.

3 Interventions

To test the three hypotheses outlined in section 2, we designed three corresponding interventions. In this section, we provide more details on these interventions.

I1: Input dealer training on proper seed handling and storage

To test if lack of adoption is due to poor seed quality as a result of poor handling and storage at the input dealer level, an intensive input dealer training to increase input dealer skills regarding seed handling and storage was implemented. This is expected to improve seed quality, in turn reducing risk and increasing profitability at the level of the farmers. This will lead to more farmers adopting improved seed and agro-input dealer to sell more improved seed. It is important to note that this hypothesis implicitly assumes that the dealer is not aware of the fact that he or she sells poor quality seed. In other words, sales of poor quality seed is not intentional.

The trainings were developed based on expert consultations from the Ugandan ministry of agriculture, from the seed sector and from input dealer associations in Uganda prior to the experiment. A consultation workshop was organized in Bugiri, as well as a series of semi structured interviews with experts of different institutions and organizations. During the workshop or the semi-structured interviews, the facilitator kept the focus on "seed storage and handling". First, problems were identified by determining what input dealers typically do wrong in terms of seed storage and handling, leading farmers to end up with sub-optimal seed quality. These problems were then ranked according to how badly they affect seed quality and how many agro-input dealers are affected by these problems. In a next step, solutions were associated to each of the problems. The solutions were also ranked in terms of effectiveness and it was discussed if these solutions would be within reach of the majority of agro-input dealers.[1]

Based on the information collected, we developed detailed training manuals that the trainers were instructed to adhere to. This ensured standardized treatment. We also created visually appealing posters showing the most important best practices that were given to input dealers to mount in their shop. The training was implemented by three trainers, one from ISSD, one from UNADA, and one agronomist that was part of the research team. It was organized at a location that was easily reachable for all sampled agro-input dealers within the treated catchment area. For each treated shop, we invited the shop owner and the shop manager. As an additional incentive, we also provided treatment agro-input dealers with a portable seed moisture meter. The input dealer trainings were organized in May 2021, such that dealers can use the handling and storage practices for seeds they buy in June/July 2021, which are going to be purchased by farmers for the second agricultural season that begins in August 2021.

I2: Information clearinghouse

This classic lemons technology problem can be solved by reducing information asymmetries between the two parties. In Kenya, seed companies have started marketing their seed using novel packaging features to signal product quality and authenticity (Gharib et al., 2021). Uganda does regulate seed quality by means of certifications and standards, but they provide farmers with a relatively weak and unreliable indication of quality. For instance, during baseline data collection, we purchased seed bags from agro-input dealers in our sample, and only 8% of them had a certification sticker from an inspection agency. Alternatives such as electronic verification systems have also been experimented with, but the cost of implementation has proven challenging, and they depend on the reliability of the underlying seed certification system.

In our study, we test an alternative, decentralized clearinghouse that is based on crowd-sourced information and works through reputational mechanisms, much like yelp.com or tripadvisor.com. Studies in other contexts have shown that new crowd-based sources of pre-purchase information can be particularly useful. For instance, Reimers and Waldfogel (2021) compare the impacts of professional critics and Amazon star ratings of books on consumer welfare.

[1] This led to the exclusion of capital intensive investments such as air conditioners or freezers to preserve quality.

The aggregate effect of star ratings on consumer surplus is more than ten times the effect of traditional review outlets.

The information clearinghouse treatment we implement is a fairly complex intervention that required collecting indicators of perceived quality of seed sold by agro-input dealer from farmers. These quality ratings were then used to construct scores to rank agro-input dealers, and these rankings and scores were then provided back to farmers and agro-input dealers in an appealing way. This entire process also happened twice, as we target two full growing cycles. In particular, prior to the clearinghouse intervention, we collected baseline data of farmers in catchment areas of the input dealers that were enrolled in the study. During this baseline interview, we asked farmers to rate input dealers in their catchment area on a number of characteristics. This was done in the ODK app, which had pre-loaded data on the agro-input dealers such as name, location, and a picture of the store front. Depending on the location of the farmer, the ODK app iterated through all agro-input dealers in the vicinity of the farmer. Furthermore, depending on the treatment status of the agro-input dealer, farmers were also asked to rate the seed the agro-input dealer sold.[2]

To construct the rankings of agro-input dealers, we used various questions that farmers were expected to answer on a likert scale. These included whether seed yields are generally as advertised, are drought tolerant as advertised, are pest/disease tolerant as advertised, and if time to maturity is as promised. We also asked to rate seed germination, and the quality of the seed in general. These ratings were translated into words (ok, good, very good, excellent) and stars, so that they are easily understandable for farmers and dealers who are not used to interpreting numbers.

The ratings were then fed back to farmers and agro-input dealers in various forms before farmers start buying seed for the next agricultural season. To farmers, ratings were disseminated both in-person, as well as through a series of sms messages. During in-person dissemination to farmers, enumerators re-visited all farmers in our sample. We designed a visually appealing dissemination app which was shown to farmers on tablet computers which cycled through all dealers in the proximity of each farmer (again showing picture of store front etc.) and stated: "We wanted to let you know that customers from *shop name* rate the quality of maize seed sold there as okay/good/very good/excellent! The quality of the maize seed that this agro-input shop sells got a score of *score* out of 5!" in treated catchment areas. The app also showed the stars associated with the score.[3] For the sms based dissemination, we sent farmers one text message

[2]It may be argued that by asking farmers to rate agro-input dealers in a particular area, you are also making farmers aware of the existence of all agro-input dealers in the area, and this awareness effect may potentially confound the clearinghouse effect. We thus also iterate through the agro-input dealers for control subjects, to make control farmers similarly aware of the existence of agro-input dealers in their vincinity.

[3]As before, to account for the fact that simply visiting farmers and reminding them about the existence of agro-input dealers may confound the clearing house effect, we revisited all farmers in control catchment areas and also cycled through dealers, but simply asked a couple of questions like e.g. "Do you know this *shop name*?".

per dealer in their proximity with similar information.[4]

Agro-input dealers received their ratings in the form of a report on laminated paper which was delivered to their shops. The front shows a visually appealing certificate with a logo and the shops general rating both in word and in stars. Enumerators advised shopkeepers to display the ratings in the shop, similar to a "certificate of excellence" of TripAdvisor. An example of the front of a report can be found in the Appendix Figure 2. The back of the report shows more detailed information. In addition to the dealer's general rating, it shows the dealer's specific ratings (overall quality, yield, drought and disease resistance, time to mature, and germination) and the average ratings of other agro-input dealers in the same catchment area in a table and visualized by stars. This shows dealers their relative position in the area and is expected to motivate agro-input dealers to perform better on areas where they can improve relative to the immediate competition.

Information clearinghouse mechanisms in developing countries have been studied to some extent, but mostly to address market price information asymmetries between smallholder farmers and middlemen. Assuming that middlemen are better informed about prevailing prices in the market than farmers, theory suggests that providing farmers with price information increases their bargaining power. However, evidence on their effectiveness is mixed: while Goyal (2010) finds that internet kiosks that provide wholesale price information significantly increased soy prices in India, Fafchamps and Minten (2012) do not find a statistically significant effect of market information delivered to farmers' mobile phones in a neighboring state. A study by Hasanain, Khan, and Rezaee (2019), who set up a rating system for public veterinary services in Pakistan is probably the closest to ours. They find that farmers who use the clearinghouse enjoy a 25% higher success rate of artificial insemination. Their research suggests that this is mostly due to increased veterinarian effort, as few farmers seem to be switching from veterinaries that receive poor ratings to veterinaries that receive good ratings.

An information clearinghouse could work through different impact pathways. First, farmers who did not buy (or who bought only a little) seed before may start adopting (or start adopting more) improved seed when they realize that an agro-input dealer in their vicinity has a better rating than what these farmers expected. The literature indicates that this pathway of improving the perception of quality without actually improving quality could be very relevant. Michelson et al. (2021) establish that the nutrient content of fertilizer in Tanzania is good and meets industry standards but that farmers believe that it is adulterated. One would simply need to correct farmers' beliefs which are inconsistent with the reality to increase adoption. Also Wossen, Abay, and Abdoulaye (2022) show that farmers in developing countries routinely misperceive input quality and that rectifying this misperception may improve farmers' investment choices and productivity outcomes. Second, farmers could switch from low rated shops to

[4] Also in control catchment areas, farmers received text messages with the names of dealers in their proximity, so that they were aware of the presence of these dealers.

higher rated shops after learning that their preferred agro-input dealer received a poor rating. Thirdly, farmers may pressure their usual dealers to increase efforts. Fourth, dealers could also start improving on their own (i.e. without the pressure of farmers) after they have learned that their seed is rated poorly and want to improve their products. Lastly, dealers could also increase their efforts after seeing that they were rated well and feel motivated and satisfied due to this reward.

I3: Addressing learning failures at the farmer level

The intervention that was used to test the learning failure hypothesis was built around short, visually appealing videos, shown to the farmers on tablet computers. Video's featuring role models have been found effective in changing people's behaviour in a range of applications (Riley, 2019; Van Campenhout, Spielman, and Lecoutere, 2020; Vandevelde, Van Campenhout, and Walukano, 2018; Bernard et al., 2015). Similar to how the content for the input dealer seed storage and handling training was determined, we consulted experts from the Ugandan ministry of agriculture, from the seed sector and from input dealer associations in Uganda prior to the experiment. Also here, a consultative workshop and semi structured interviews were organized with experts of different institutions and organizations.

Based on the information obtained from these consultations, we developed a video that shows what complementary inputs and practices are important to create an enabling environment for improved seed to flourish. A treatment and a control version of the video was produced. The only difference between the treatment and the control video is that in the treatment video, after each complementary input or practice that is shown, the actor stresses that these inputs or practices are particularly important if improved seed varieties are used. The videos were shown to farmers individually on tablet computers.

4 Experimental design

To test the three hypotheses, the three interventions are combined in a field experiment where various treatment and control groups are randomly assigned to either a treatment or control condition. The randomized control trial (RCT) takes the form of a 2^3 factorial design, with each intervention corresponding to one hypothesis. To test the first hypothesis, a random sub-sample of input dealers received training on proper seed handling and storage. To test the second hypothesis, an information clearing house was set up among a random sub-sample catchment areas of input dealers, and farmer and input dealers will receive feedback on the ratings before the start of the planting season. To test the third hypothesis, a video that points out the importance of combining improved seed with other inputs and careful crop management will be shown to a random subset of villages. Impact is judged by looking at outcomes both at the input dealer level (eg. investments in seed storage infrastructure, quantity of seed sold,...) as well as at the farmer level (eg. likelihood that farmer adopted improved seed, maize yields,...).

Factorial designs allow recycling of treated units in the orthogonal factor to be used as controls. As such, to estimate main effects, less observations are needed than would be the case in parallel designs. The factorial design we will use deviates from commonly used factorial designs in that the experimental unit will differ depending on the factor. For the first two factors, corresponding to the input-dealer training and the information clearinghouse, randomization will happen at the level of the catchment area. For the third factor that address learning failures of farmers, randomization will happen at the level of the village.[5]

The resulting layout, with sample sizes indicated in each treatment cell is illustrated in Figure 1. The first two interventions are implemented at the catchment area level. A total of 112 catchment areas are included in the study. Half of these are randomly allocated to the first treatment: all input dealers in 56 catchment areas receive the input dealer training, while input dealers in the remaining 56 catchment areas function as the control for this treatment. Data that was collected in three of the study districts indicates that this corresponds to about 160 input dealers in each treatment arm. Orthogonal to the first factor, the second factor is placed, corresponding to the second treatment that is also implemented at the catchment area level. Also here, in half of the 112 catchment areas an information clearinghouse will be implemented, and half of the catchment areas will function as a control for this treatment. However, this will be done in such a way that balance with respect to the first treatment exists in both treatment and control groups for the second treatment. This means that the treatment group of the second treatment will consist of 28 catchment areas that received the first treatment and 28 catchment areas that function as the control for the first treatment. Similarly, for the control catchment areas for the second treatment, half will consist of catchment areas where input dealers received the input dealer training and half of catchment areas where input dealers did not get trained.

While the third treatment is implemented at the level of the village, it is also important to preserve balance in the orthogonal factors. In other words, we need to make sure that an equal number of villages that are assigned to receive a treatment against learning failures are drawn from catchment areas where input dealers received training as from catchment areas where the input dealer training did not take place. Similarly, orthogonality should also be maintained for the second treatment. Therefore, in each of the four treatment cells formed by interacting the first two treatments, 40 villages (347 villages in 130 catchment areas, ie. 2,67 villages per catchment area, so that 14 areas correspond to 37,37 villages) will be randomly assigned to the third treatment while another 40

[5]The main motivation to randomize at the village level is to eliminate potential spillover effects for the third treatment. However, as we will discuss later, we make sure there is correspondence between villages and input dealers, which would allow us to also look at the impact of the farmer training treatment on outcomes at the input dealer level. However, at that level, spillovers may affect results.

villages will be assigned to the control.

Data

Baseline data was collected among a sample of about 350 agro-input dealers towards the end of 2020 and among a sample of about 3,500 smallholder farmers in the beginning of 2021. This mock report relies extensively on this baseline data. A first part of the report provides descriptive statistics for the baseline data. Tables 1 to 3 report baseline summary statistics at the agro-input dealer level. Tables 4 to 7 does this for the farmers.

A second part checks for balance between treatment and control groups in a set of pre-registered baseline characteristics—typically the first table in RCT studies. Table 8 does this for agro-input dealers, while Table 9 looks at balance between treatment and control farmer. As we test 3 different hypothesis using three difference interventions, we also need 3 balance tests for each variable. We also have a short section on attrition.

We then turn to the analysis of the (pre-registered) outcomes that will be used to judge the effectiveness of the interventions—the center piece of a mock report. In particular, we simulate mock endline data by randomly sampling from the baseline outcome and adding a hypothesized treatment effect. We then report difference between treatment and control groups using ANCOVA models that also control for the baseline value of the outcome. Results of these regressions are reported for a set of primary outcomes at the agro-input dealer level (Table 12) as well as at the farmer level (Table 25). We also provide various tables of secondary outcomes that are grouped by family, again for agro-input dealers (Tables 14 to 24) and for farmers (Tables 27 to 34).

This document was written using the open source lyx/Latex typesetting software. The analysis is contained in an R script called "mock_report.R", which is run from within lyx/latex using the knitr engine. All code, data and documents are also under revision control using git and publicly accessible via github (https://github.com/bjvca/Seed_systems_project). The fact that the entire project is under revision control using git/github provides detailed and time-stamped recording of any changes made over the course of the project, further reducing scope for fishing by the researchers and increasing transparency with respect to the decisions made.

Descriptive statistics

The first part of this report provides descriptive statistics for the baseline data. Tables 1 to 3 report baseline summary statistics at the agro-input dealer level. Tables 4 to 7 does this for the farmers.

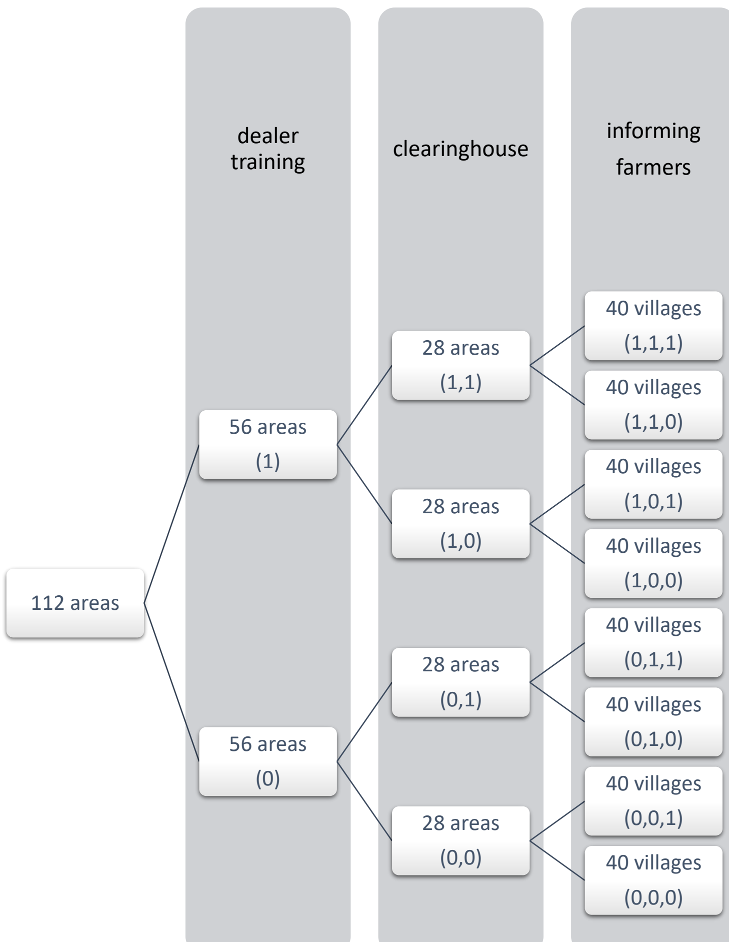

Figure 1: Design

Table 1: Descriptive statistics - Agro-input dealer (baseline)

	mean	min	max	SD	nobs
Respondent's age in years	32.43	15	80	11.49	347
Respondent is male	0.595	0	1	0.492	348
Respondent finished primary education	0.920	0	1	0.271	339
Respondent finished secondary education	0.386	0	1	0.488	339
Respondent owns shop	0.555	0	1	0.498	348
Shop's distance to nearest tarmac road in km	6.556	0	52	10.39	343
Shop's distance to nearest murram road in km	0.190	0	9	0.626	348
Shop only sells farm inputs	0.741	0	1	0.439	348
Number of customers per day	41.49	2	300	46.49	346
Number of customers per day **buying maize seed**	21.27	0	250	26.80	347
Years since shop establishment	5.339	0	33	6.299	348
Shop also sells machinery	0.066	0	1	0.249	348
Shop also sells equipment	0.724	0	1	0.448	348
Shop also sells chemicals	0.945	0	1	0.228	348
Shop also sells fertilizers	0.960	0	1	0.197	348
Respondent received training on maize seed handling/storage	0.526	0	1	0.500	348
Respondent received this training **last year**	0.267	0	1	0.443	348
Respondent received this training last year **by ISSD**	0.106	0	1	0.308	341
Number of maize varieties in stock	2.917	0	10	1.755	348
Number of **hybrid** maize varieties in stock	1.681	0	8	1.330	348
Shop has Longe 10H in stock	0.684	0	1	0.466	348
Shop has Longe 7H in stock	0.161	0	1	0.368	348
Number of **OPV** maize varieties in stock	1.276	0	5	0.686	348
Shop has Longe 5 in stock	0.885	0	1	0.319	348
Shop has Longe 4 in stock	0.264	0	1	0.442	348
Shop stores seed away from other products	0.460	0	1	0.499	348
Shop has problem with pests	0.649	0	1	0.478	348
Shop has leak-proof roof	0.537	0	1	0.499	348
Shop has insulated roof	0.580	0	1	0.494	348
Shop has insulated walls	0.813	0	1	0.390	348
Shop is ventilated	0.793	0	1	0.406	348
Shop has plastered walls	0.920	0	1	0.272	348
Shop's floor is cement/tiles (not mud)	0.974	0	1	0.160	343
Shop's light is ambient (not direct sunlight/dark)	0.825	0	1	0.381	348
Shop stores seed on pallets/shelves (not directly on wood/floor/cardboard)	0.707	0	1	0.456	331
Shop stores maize seed in open containers	0.155	0	1	0.363	348
Shop displays official certificate	0.460	0	1	0.499	348
Shop's cleanness/professionality rating by enumerator	3.451	1	5	1.098	348

Table 2: Descriptive statistics - Agro-input dealer (baseline)

	mean	min	max	SD	nobs
Shop never had expired seed	0.589	0	1	0.493	348
Shop had expired seed but always handled it correctly	0.759	0	1	0.429	145
Shop always explains to customers how seed should be used	0.457	0	1	0.499	348
Shop always recommends complementary inputs to customers	0.529	0	1	0.500	348
Shop offers extension/training to some/ to everyone	0.483	0	1	0.500	348
Shop offers discounts for large quantities	0.750	0	1	0.434	348
Shop's smallest seed bag is 1 kg (not larger)	0.728	0	1	0.445	335
Shop repackages seed if customers want small quantities	0.523	0	1	0.500	348
Shop charges more if customers buy only 1 kg	0.154	0	1	0.362	182
Shop keeps expiry date when repackaging seed	0.588	0	1	0.494	182
Shop provides seed on credit to some	0.595	0	1	0.492	348
Number of customers who received credit (if shop provides credit)	11.02	1	120	13.80	193
Number of **women** who received credit (if shop provides credit)	3.430	0	35	4.692	200
Shop received seed related complaint from customer	0.644	0	1	0.480	348
Shop accepts mobile money as payment	0.348	0	1	0.477	348
Shop sometimes delivers to customers	0.399	0	1	0.490	348
Dealer's self-rating: location	3.876	1	5	0.878	348
Dealer's self-rating: price	3.922	1	5	0.867	348
Dealer's self-rating: product quality	4.046	1	5	0.844	348
Dealer's self-rating: stock, convenient quantities	3.583	1	5	1.002	348
Dealer's self-rating: reputation	4.319	2	5	0.735	348
Shop is registered with UNADA	0.442	0	1	0.497	319
Shop has trading license by local government	0.749	0	1	0.435	338
Shop is member of other professional association	0.345	0	1	0.476	325
Number of inspections by DAO/MAAIF/UNADA last year	1.866	0	43	3.843	335
Shop received warning after inspection	0.317	0	1	0.466	334
Shop's products were confiscated after inspection	0.145	0	1	0.353	337
Shop was closed after inspection	0.009	0	1	0.093	342
Shop has equipment to monitor seed moisture	0.026	0	1	0.159	348
Shop monitors temperature	0.026	0	1	0.159	348
Shop temperature where seed is stored in degrees Celsius	25.31	19.5	52	2.996	345
Moisture in random seed bag in percent	13.576	10.3	17.4	1.522	232
Random seed bag shows expiry date	0.181	0	1	0.386	232
Random seed bag shows expiry date but seed is expired	0.049	0	1	0.218	41
Random seed bag shows packaging date	0.677	0	1	0.469	232
Packaging date is visible but more than 6 months ago	0.039	0	1	0.194	154
Days since packaging date/expiry date minus 6 months	64.96	9	261	47.41	183
Random seed bag is original and undamaged	0.935	0	1	0.246	232
Random seed bag shows certification sticker	0.082	0	1	0.275	232
Random seed bag shows lot number	0.504	0	1	0.501	232
Random seed bag shows e-verification	0.026	0	1	0.159	232

Table 3: Descriptive statistics - Agro-input dealer (baseline)

	mean	min	max	SD	nobs
Shop is known by random farmer in village with most clients	0.557	0	1	0.260	347
Shop sold seed to this farmer	0.247	0	1	0.184	342
Years since shop has this farmer as customer (if shop sold seed)	3.729	0	13	2.310	313
Shop sold seed to someone this farmer knows	0.085	0	1	0.137	341
Shop's maize seed rating on quality by farmers (who (know someone who) bought seed there)	3.772	1.5	5	0.527	176
Shop's maize seed rating on yield by farmers	3.537	1.5	5	0.527	175
Shop's maize seed rating on drought tolerance by farmers	2.938	1	5	0.525	169
Shop's maize seed rating on pest/disease tolerance by farmers	2.445	1	4	0.516	173
Shop's maize seed rating on time of maturity by farmers	3.817	2	5	0.403	172
Shop's maize seed rating on germination by farmers	3.669	2	5	0.540	172
Number of shop's maize seed ratings by farmers	NA	NA	NA	NA	NA
Shop refunds if problem, according to farmers (who (know someone who) bought seed there)	0.331	0	1	0.302	316
Shop gives credit, i.e. inputs one can pay later, according to farmers	0.410	0	1	0.313	314
Shop advises during sales, according to farmers	0.757	0	1	0.263	320
Shop delivers seed to clients, according to farmers	0.235	0	1	0.282	315
Shop provides after-sales service, according to farmers	0.241	0	1	0.288	322
Shop accepts different payment methods, according to farmers	0.420	0	1	0.327	314
Shop sells small quantities if necessary, according to farmers	0.898	0	1	0.188	324

Table 4: Descriptive statistics - Farmer (baseline)

	mean	min	max	SD	nobs
Homestead's distance to nearest tarmac road in km	9.390	0	100	10.81	3302
Homestead's distance to village headquarters in km	0.745	0	15	0.903	3436
Homestead's distance to nearest agro-input shop in km	3.779	0	52	4.789	3339
Number of agro-input shops in farmer's village/neighborhood	2.163	0	25	2.346	3263
Homestead's distance to nearest neighbor in km	0.114	0	2	0.183	3463
Farmer's age in years	48.62	18	97	13.38	3453
Farmer is male	0.777	0	1	0.416	3470
Farmer is married	0.884	0	1	0.320	3470
Farmer had no formal education	0.079	0	1	0.270	3437
Farmer finished primary education	0.507	0	1	0.500	3437
Farmer finished secondary education	0.089	0	1	0.284	3437
Number of people in household (incl. respondent)	8.695	1	25	3.979	3470
Number of rooms in house	3.490	1	10	1.445	3469
Roof is made of iron sheets/tiles (not grass)	0.928	0	1	0.259	3460
Farmer's land for crop production in acres	3.348	0.185	100	4.320	3442
Years since farmer started growing maize	23.09	0	82	13.14	3470
Farmer is member of (maize) farmer group/association/cooperative	0.126	0	1	0.332	3459
Farmer used improved maize seed (OPV/hybrid) for any field last season	0.492	0	1	0.500	3466
Farmer used farmer saved maize seed (if he/she used improved seed)	0.163	0	1	0.370	1664
Farmer bought maize seed at agro-input shop (if he/she used improved seed)	0.668	0	1	0.471	1664
Farmer used maize seed for more than 3rd (not 1st/2nd) time (if he/she used farmer saved seed)	0.529	0	1	0.500	261
Amount of improved maize seed bought from agro-input shop in kg (if farmer bought from shop)	11.07	1	200	13.48	1108
Farmer did not buy seed at agro-input shop because it is too expensive	0.860	0	1	0.347	1813
Farmer did not buy seed at agro-input shop because it is of poor quality	0.088	0	1	0.283	1813
Farmer bought seed at particular agro-input shop because it is of very good quality	0.578	0	1	0.494	948
Farmer says seed at agro-input shop is of poor quality due to disappointing yield	0.252	0	1	0.435	151
Farmer says seed at agro-input shop is of poor quality due to disappointing pest tolerance	0.252	0	1	0.435	151
Farmer says seed at agro-input shop is of poor quality due to disappointing germination	0.318	0	1	0.467	151
Farmers thinks seed at agro-input shop is counterfeit/adulterated	0.685	0	1	0.465	2673
Farmer mentioned Longe 5/Nalongo when asked for improved maize varieties	0.655	0	1	0.475	3470
Farmer mentioned Longe 7R/Kayongo-go when asked for improved maize varieties	0.076	0	1	0.266	3470
Farmer mentioned Wema when asked for improved maize varieties	0.010	0	1	0.099	3470

Table 5: Descriptive statistics - Farmer (baseline)

	mean	min	max	SD	nobs
Farmer grew maize on only 1 field (stand alone/mixed)	0.653	0	1	0.476	3470
Number of fields farmer grew maize on (stand alone/mixed)	1.463	1	5	0.725	3470
Area of randomly selected maize field in acres	1.181	0.075	20	1.001	3465
This maize field was intercropped	0.702	0	1	0.457	3470
This maize field was intercropped with beans	0.419	0	1	0.493	3470
This maize field was intercropped with soybeans	0.108	0	1	0.311	3470
This maize field was intercropped with groundnuts	0.108	0	1	0.311	3470
This maize field was intercropped with cassava	0.267	0	1	0.443	3470
This maize field was intercropped with millet	0.002	0	1	0.045	3470
This maize field was intercropped with sorghum	0.004	0	1	0.061	3470
Percentage allocated to maize (if this field was intercropped)	56.43	5	99	17.84	2414
Farmer planted hybrid maize seed on this field	0.264	0	1	0.441	3124
Farmer planted open-pollinated maize seed on this field	0.260	0	1	0.439	3124
Farmer planted local land race maize seed on this field	0.448	0	1	0.497	3318
Farmer planted hybrid or open-pollinated maize seed on this field	0.552	0	1	0.497	3318
This maize seed was farmer saved	0.579	0	1	0.494	3429
Farmer bought this maize seed at agro-input shop	0.330	0	1	0.470	3429
This maize seed was hybrid but farmer saved	0.050	0	1	0.217	3103
This maize seed was open-pollinated, farmer saved but used 4 or more times	0.031	0	1	0.173	3108
Farmer's rating of this maize seed on general quality	3.385	1	5	1.032	3461
Farmer's rating of this maize seed on yield	3.040	1	5	1.081	3462
Farmer's rating of this maize seed on drought tolerance	2.806	1	5	1.004	3378
Farmer's rating of this maize seed on pest/disease tolerance	2.189	1	5	1.009	3456
Farmer's rating of this maize seed on early maturity	3.416	1	5	1.025	3457
Farmer's rating of this maize seed on demand/market/output price	2.209	1	5	1.096	3299
Farmer's rating of this maize seed on taste	4.031	1	5	0.929	3448
Farmer's rating of this maize seed on price	3.187	1	5	1.223	3163
Farmer's rating of this maize seed on availability	3.362	1	5	1.025	3387
Farmer's rating of this maize seed on germination	3.570	1	5	0.937	3468
Farmer was satisfied with this maize seed	0.678	0	1	0.467	3470
Farmer told supplier that he/she was not satisfied (if not satisfied with maize seed)	0.275	0	1	0.447	509
Farmer would use this maize seed again	0.764	0	1	0.425	3470
Amount of maize seed used on this field in kg	9.616	0.2	200	9.871	3413
Price of this maize seed per kg in UGX	2126	0	14500	3249	3411
Price of this maize seed per kg in dollars	0.597	0	4.071	0.912	3411
Cost of this maize seed in UGX (amount in kg x price per kg)	19289	0	1e+06	47873	3357
Cost of this maize seed in dollars (amount in kg x price per kg)	5.42	0	280.8	13.44	3357

Table 6: Descriptive statistics - Farmer (baseline)

	mean	min	max	SD	nobs
Farmer spaced seeds correctly on this maize field (2.5 feet x 1 foot)	0.109	0	1	0.312	3470
Number of seeds per hill on this maize field	3.350	1	8	1.254	3362
Farmer sowed correct number of seeds per hill on this maize field (1)	0.042	0	1	0.201	3362
Farmer applied organic manure on this maize field	0.074	0	1	0.262	3466
Farmer applied DAP/NPK on this maize field	0.249	0	1	0.432	3465
Amount of DAP in kg on this maize field (if farmer applied DAP/NPK)	16.88	0.1	150	20.68	845
Farmer applied Urea on this maize field	0.075	0	1	0.263	3466
Amount of Urea in kg on this maize field (if farmer applied Urea)	15.81	0.5	150	19.89	251
Number of times farmer weeded this maize field	2.560	0	5	0.650	3466
Farmer weeded this maize field 3 or more times	0.529	0	1	0.499	3466
Days after planting farmer first weeded this maize field	17.53	1	60	6.662	3428
Farmer weeded this maize field first at correct time (18-20 days after planting)	0.063	0	1	0.244	3428
Farmer used pesticides/herbicides/fungicides on this maize field	0.409	0	1	0.492	3463
Farmer planted at correct time on this maize field (1-3 days after 1st rains)	0.699	0	1	0.459	3441
Farmer re-sew where seeds did not germinate on this maize field	0.483	0	1	0.500	3464
Number of maize bags harvested from this field last season (incl. consumed)	5.364	0	250	8.520	3460
Kilograms per bag	100.5	40	149	8.978	3469
Yield in kg (number of harvested maize bags x kg per bag)	544.2	0	25000	858.2	3459
Land productivity in kg/acre (yield/area)	499.5	0	28000	771.2	3454
Market value per bag at harvest in UGX	70259	20000	149999	27821	3400
Market value per bag at harvest in dollars	19.73	5.616	42.12	7.812	3400
Yield in UGX (number of harvested maize bags x market value per bag)	391126	0	36250000	835450	3390
Yield in dollars (number of harvested maize bags x market value per bag)	109.8	0	10178	234.6	3390
Land productivity in UGX/acre (yield/area)	351874	0	1.8e+07	534819	3385
Land productivity in dollars/acre (yield/area)	98.80	0	5054	150.2	3385
Farmer sold maize from this field	0.513	0	1	0.500	3470
Number of maize bags sold from this field (if farmer sold maize)	5.007	0.02	250	9.134	1778
Price farmer charged per bag in UGX	53596	1	750000	38713	1774
Price farmer charged per bag in dollars	15.05	0	211	10.87	1774
Revenue in UGX (number of sold maize bags x price per bag)	313	0.001	70000	1734	1772
Revenue in dollars (number of sold maize bags x price per bag)	0.09	0	20	0.5	1772
Amount kept as seed in kg	17.57	0	400	33.37	1710
Number of maize bags farmer expects to harvest from this field next season (incl. consumed)	8.764	0	280	12.01	2574

Table 7: Descriptive statistics - Farmer (baseline)

	mean	min	max	SD	nobs
Farmer knows particular shop in neighborhood	0.556	0	1	0.380	3449
Farmer bought shop's seed (if he/she knows any shop)	0.258	0	1	0.346	2804
Years since farmer became shop's customer (if he/she bought seed from any shop)	4.104	0	35	3.873	1280
Farmer knows someone who bought shop's seed (if he/she did not buy seed from all shops but knows any)	0.079	0	1	0.235	2462
Farmer bought shop's seed last season	0.448	0	1	0.468	598
Farmer's rating of particular shop on general quality (if he/she (knows someone who) bought seed there)	3.709	1	5	0.946	730
Farmer's rating of particular shop on location	3.874	1	5	1.156	730
Farmer's rating of particular shop on price	3.249	1	5	1.109	730
Farmer's rating of particular shop on product quality	3.780	1	5	1.015	730
Farmer's rating of particular shop on seed stock/availability	3.892	1	5	1.030	730
Farmer's rating of particular shop on reputation/reliability	4.133	1	5	0.910	730
Farmer's rating of particular shop's maize seed on general quality	3.788	1	5	0.841	711
Farmer's rating of particular shop's maize seed on yield	3.542	1	5	0.881	699
Farmer's rating of particular shop's maize seed on drought tolerance	2.997	1	5	0.858	679
Farmer's rating of particular shop's maize seed on pest/disease tolerance	2.446	1	5	0.904	685
Farmer's rating of particular shop's maize seed on timing of maturity	3.845	1	5	0.695	688
Farmer's rating of particular shop's maize seed on germination	3.679	1	5	0.854	699

Orthogonality tests of randomization balance

Testing if treatment and control groups are comparable in terms of a set of pre-registered baseline characteristics are often reported in field experiments. Table 8 does this for agro-input dealers, while Table 9 looks at balance between treatment and control farmer. As we test 3 different hypothesis using three difference interventions, we also need 3 balance tests for each variable.

Orthogonality tests of survey attrition

While random attrition only reduces statistical power, attrition which is correlated with one of our treatments could bias estimates. We focus on limiting attrition during data collection because it is difficult to solve ex post. That is why we recorded the respondents' full names, primary and secondary telephone numbers, and enumerators captured the locations ie. the GPS coordinates of the interviews. We collected this information for every participant during baseline data collection. We additionally asked the agro-input dealers for other (nick)names that they are known by, enumerators took pictures of their shops and wrote down eye-catching features to later identify it. Based on this information, enumerators will trace missing participants for mid- and endline data collection. Due to these measures and because the surveys are conducted over a reasonably short time period, we expect most baseline respondents to also be available for the mid- and endline survey and hence low attrition rates.

Table 10 simulates attrition levels in the treatment and comparison groups. Tables **??** and **??** use this simulated attrition to show how we will test whether attritors and non-attritors differ systematically in our baseline data, at least along observable dimensions. In these tests of survey attrition we include the same variables as in our tests of randomization balance.

right:

If attritors and non-attritors differ significantly and attrition could bias our estimates, we will exploit statistical techniques to identify and adjust for this bias. Manski (1989) and Lee (2002) for example suggest non-parametric bounds on the treatment effect that can be estimated from available data.

Written in May 2022:

Despite our efforts to minimize attrition, we failed to collect midline data from 12.07% of agro-input dealers who were part of the original sample. This reduces our statistical power. Furthermore, the probability that a dealer left the sample is correlated with the clearinghouse treatment status and this might bias the estimates.

Whether our estimates are biased or not depends on the reason why more agro-input dealers in the clearinghouse treatment group were found than in the clearinghouse control group. It is for example plausible that enumerators invested made less of an effort when finding clearinghouse control dealers because they did not have to deliver their SeedAdvisor certificates. Carrying this certificate might have made them more persistent when looking for a shop because

Table 8: Orthogonality tests of randomization balance - Agro-input dealer (baseline)

	mean	dealer training	clearing house	farmer video
Respondent's age in years	32.427	0.555	-2.242^{+}	-1.635
	(11.492)	(1.188)	(1.215)	(1.230)
Respondent is male	0.595	0.015	-0.009	0.043
	(0.492)	(0.057)	(0.057)	(0.053)
Respondent finished primary education	0.920	0.013	-0.013	-0.020
	(0.271)	(0.034)	(0.033)	(0.029)
Shop's distance to nearest tarmac road in km	6.556	-0.919	-1.584	-1.262
	(10.390)	(2.211)	(2.242)	(1.118)
Number of customers per day	41.486	11.355	6.427	3.249
	(46.489)	(7.160)	(6.720)	(4.972)
Years since shop establishment	5.339	-0.092	0.209	0.155
	(6.299)	(0.771)	(0.781)	(0.683)
Respondent received training on maize seed handling/storage	0.526	0.049	0.120^{+}	0.073
	(0.500)	(0.068)	(0.067)	(0.053)
Number of maize varieties in stock	2.917	-0.120	0.326	0.297
	(1.755)	(0.261)	(0.248)	(0.187)
Amount of maize seed sold during last season in kg	910.885	275.416	258.563	117.407
	(2683.235)	(408.648)	(386.410)	(290.018)
Amount of maize seed lost/wasted during last season in kg	3.504	1.995	2.404	1.199
	(18.651)	(2.468)	(2.300)	(2.005)
Shop has problem with pests	0.649	-0.008	-0.028	-0.017
	(0.478)	(0.057)	(0.056)	(0.051)
Shop has leak-proof roof	0.537	-0.014	0.004	-0.047
	(0.499)	(0.057)	(0.057)	(0.054)
Shop received seed related complaint from customer	0.644	-0.106*	0.073	-0.088^{+}
	(0.480)	(0.048)	(0.048)	(0.051)
Respondent knows how seed should be stored after repackaging	0.270	0.067	0.077	-0.094*
	(0.445)	(0.063)	(0.064)	(0.048)
Shop has trading license by local government	0.749	0.013	0.095^{+}	-0.004
	(0.435)	(0.057)	(0.057)	(0.047)
Moisture in random seed bag in percent	13.576	0.122	-0.106	0.164
	(1.522)	(0.257)	(0.273)	(0.203)
Random seed bag shows lot number	0.504	-0.007	-0.011	-0.046
	(0.501)	(0.093)	(0.096)	(0.067)
Shop refunds if problem, according to farmers (who (know someone who) bought seed there)	0.331	0.009	-0.100*	-0.014
	(0.302)	(0.047)	(0.046)	(0.034)
Shop gives credit, i.e. inputs one can pay later, according to farmers	0.410	-0.005	0.005	-0.042
	(0.313)	(0.048)	(0.047)	(0.035)
Shop provides after-sales service, according to farmers	0.241	0.093^{+}	0.055	0.022
	(0.288)	(0.047)	(0.045)	(0.032)
Number of observations	348	348	348	348

Note: First column reports sample means (and standard deviations below); **, * and + denote significance at the 1, 5 and 10 percent levels. Reported standard errors are clustered at the level of randomization (catchment area or village/shop).

Table 9: Orthogonality tests of randomization balance - Farmer (baseline)

	mean	dealer training	clearing house	farmer video
Homestead's distance to nearest tarmac road in km	9.390	0.334	-1.233	-1.275
	(10.810)	(1.826)	(1.844)	(1.048)
Homestead's distance to nearest agro-input shop in km	3.779	-0.109	0.110	-0.039
	(4.789)	(0.405)	(0.403)	(0.353)
Farmer's age in years	48.617	-0.083	-0.239	-0.247
	(13.385)	(0.602)	(0.599)	(0.561)
Farmer is male	0.777	-0.024	0.027	0.022
	(0.416)	(0.028)	(0.028)	(0.021)
Farmer finished primary education	0.507	-0.004	0.040	0.000
	(0.500)	(0.027)	(0.027)	(0.023)
Number of people in household (incl. respondent)	8.695	-0.158	-0.090	0.118
	(3.979)	(0.198)	(0.197)	(0.171)
Number of rooms in house	3.490	-0.006	0.020	0.065
	(1.445)	(0.097)	(0.097)	(0.071)
Farmer's land for crop production in acres	3.348	0.074	-0.002	0.172
	(4.320)	(0.224)	(0.231)	(0.180)
Farmer used improved maize seed (OPV/hybrid) for any field last season	0.492	0.024	0.014	-0.012
	(0.500)	(0.026)	(0.026)	(0.022)
Farmer used improved maize seed bought at agro-input shop	0.325	-0.014	0.011	-0.009
	(0.468)	(0.025)	(0.025)	(0.021)
Amount of improved maize seed bought from agro-input shop in kg (0 if not from shop)	3.533	0.035	-0.193	-0.236
	(9.198)	(0.476)	(0.475)	(0.394)
Farmers thinks seed at agro-input shop is counterfeit/adulterated	0.685	0.012	-0.003	-0.012
	(0.465)	(0.037)	(0.038)	(0.028)
Randomly selected maize field was intercropped with beans	0.419	-0.021	-0.019	-0.011
	(0.493)	(0.036)	(0.035)	(0.024)
Farmer used improved (not too often recyvled) maize seed for randomly selected field last season	0.477	0.023	-0.003	0.009
	(0.500)	(0.030)	(0.029)	(0.023)
Farmer's rating of maize seed planted on randomly selected maize field on general quality	3.385	0.049	0.084	0.083^{+}
	(1.032)	(0.052)	(0.053)	(0.050)
Farmer applied organic manure on randomly selected maize field	0.074	-0.009	0.007	-0.009
	(0.262)	(0.016)	(0.015)	(0.011)
Farmer planted at correct time on randomly selected maize field (1-3 days after 1st rains)	0.699	0.007	-0.028	-0.016
	(0.459)	(0.028)	(0.028)	(0.025)
Yield in kg (number of harvested maize bags x kg per bag)	544.188	6.258	-7.221	2.037
	(858.238)	(40.693)	(40.235)	(34.854)
Land productivity in kg/acre (yield/area)	499.517	43.745	8.614	44.079
	(771.173)	(32.684)	(30.995)	(31.484)
Farmer sold maize from randomly selected maize field	0.513	0.067^{+}	-0.015	0.023
	(0.500)	(0.035)	(0.034)	(0.025)
Number of observations	3470	3470	3470	3470

Note: First column reports sample means (and standard deviations below); **, * and + denote significance at the 1, 5 and 10 percent level. Reported standard errors are clustered at the level of randomization (catchment area or village/shop).

Table 10: Attrition levels in control and treatment groups

	all	control	dealer training	clearing house	farmer video
Number of attritors (dealers)	50	8	25	21	25
Number of dealers	348	42	166	193	175
Percentage of attritors (dealers)	14.37%	19.05%	15.06%	10.88%	14.29%
Number of attritors (farmers)	29	3	11	15	16
Number of farmers	3470	411	1660	1931	1750
Percentage of attritors (farmers)	0.84%	0.73%	0.66%	0.78%	0.91%

Table 11: Survey attrition

	mean	dealer training	clearing house	farmer video	number of observations
Agro-input dealer left the sample	0.144	0.017	-0.079^{+}	-0.002	348
	(0.351)	(0.040)	(0.042)	(0.038)	
Farmer left the sample	0.008	-0.003	-0.001	0.002	3470
	(0.091)	(0.003)	(0.003)	(0.003)	

Note: First column reports sample means (and standard deviations below); **, * and + denote significance at the 1, 5 and 10 percent levels. Reported standard errors are clustered at the level of randomization (catchment area or village/shop).

they did not want to return to their supervisor without having delivered that paper. Moreover, the certificate might have helped the enumerators to find the treated dealers because they were able to show the names to neighbors etc. (instead of just asking) who in turn helped finding them. In that case, a larger number of random dealers left the control sample, meaning that the dealers who were not found are not different from the ones that were found. The subsample of dealers that remained in the control group would then be representative for the entire control group, hence our estimates would be unbiased.

On the other hand, a non-random subset of agro-input dealers might have left our sample. It is for example plausible that the worst performing shops in the clearinghouse control group went bankrupt. Our clearinghouse treatment might have prevented bankruptcy and helped dealers to stay in the market because it served as some kind of advertisement if the rating was good. Hence, as our treatment effects are positive, selection bias is most likely negative. Duflo, Glennerster and Kremer state in their famous toolkit for using randomization in development economics research that in this case, when attrition bias is negative and the treatment effect is positive, the ordinary unadjusted estimates provide a lower bounds for the true effect. In other words, selection-contaminated comparisons provide a lower bound on the impact of the treatment on the outcome variables (Angrist, Bettinger, Kremer, 2006). To construct upper bounds for treatment estimates that are robust to non-random attrition, we can use non-parametric approaches, for example the ones by Manski (1989) or Lee (2002). The upper bound on the treatment effect can be estimated from available data if we make plausible assumptions about the monotonicity of outcomes and attrition, and relative rank restrictions on the distribution of outcomes.

However, we refrain from estimating these upper Manski (1989) or Lee (2002) bounds at this point in time. Instead, we will increase our efforts to find every agro-input dealer at endline data collection. If a dealer is not found despite this, we will include the question why the dealer was not found in the ODK data collection app, and add answer options like shop closed, shop relocated, shop sells different products, dealer does not want to be interviewed, shop could not be located etc. If attrition remains correlated with the clearinghouse treatment, we can use this information to check the assumptions mentioned above, and to eventually estimate the upper bounds. At this point, we do not know which dealers have left our sample, and therefore choose the most conservative approach, namely to look at the ordinary unadjusted estimates, which are lower bounds for the true effect.

Outcome variables

The remaining tables (Table 12 to 34) all test differences between treatment and control groups for the three hypotheses. We have separate sections for outcomes at the agro-input dealer level and the farmer level. We also define a set of primary outcomes to test overall impact and various secondary outcomes to explore impact more in detail and look at mechanisms. We generally provide two tables: and one where p-values are adjusted using the method outlined in Sankoh, Huque, and Dubey (1997).

Agro-input dealer

We start with outcomes at the agro-input dealer level (Tables 12 to 24).

Primary

Primary outcomes at the dealer level are are reported in Table 12 with associated table that corrects p-values for multiple hypothesis testing in Table 13.

Some outcome variables are not included in the overall index: The average sales price of 4 improved maize varieties last season is not included because the impact of the treatments on this outcome is ambiguous (increased adoption could eg. increase demand and hence prices but dealers could also lower prices to be more customer friendly). The index of all seed handling and storage practices is not included because it is a function of the index of capital-intensive practices and the index of labor-intensive practices, which are both included in the index. The index of shop's maize seed ratings by farmers is not included because the questions to compute this index are only collected for all dealers at endline (at midline, these variables are only collected for clearinghouse treated farmers as rating is part of the clearinghouse treatment).

Secondary

Indices As a first set of secondary outcomes, we construct a series of indices to assess things like motivation, knowledge etc. This is reported in table 14. The outcome variables in 14 are not further adjusted for multiple hypothesis testing by means of an overall index or p-values adjusted according to Sankoh, Huque, and Dubey (1997) because they are all indices.

Other secondary A second family of secondary outcomes are in Table 15 with associated Table 16 that adjusts for multiple comparisons.

Some outcome variables are not included in the overall index: The number of maize varieties in stock last season is not included because it is a function of the number of hybrid maize varieties and the number of open-pollinated maize varieties which are both included in the index. Whether the shop has equipment to monitor seed moisture or not is also not included because we provide dealers with this equipment as part of the dealer training. Including this outcome in the overall index would bias our results.

Longe 10H We also look at outcomes for particular seed types. Table 17 looks at farmer level outcomes related to a popular hybrid maize seed, while Table 18 gives the same information but adjusted for multiple comparisons.

Some outcome variables are not included in the overall index: The transformed sales price of Longe 10H per kg at beginning of last season (IHS) is not

Table 12: Differences between treatment and control groups - Agro-input dealer, primary outcome variables

	mean	dealer training	clearing house	farmer video	nobs
Transformed quantity sold of 4 improved maize varieties last season in kg (IHS)†	6.035	-0.633*	0.098	-0.104	249
	(1.989)	(0.257)	(0.249)	(0.225)	
Average sales price of 4 improved maize varieties last season in UGX/kg	4677.213	-23.658	149.095	2.010	227
	(925.835)	(134.322)	(132.987)	(123.177)	
Transformed revenue from 4 improved maize varieties in mln UGX (IHS)†	1.247	-0.236*	0.055	-0.057	249
	(0.977)	(0.107)	(0.106)	(0.098)	
Transformed number of customers who bought maize seed per day last season (IHS)†	2.996	-0.232^{+}	0.226^{+}	-0.035	251
	(0.887)	(0.122)	(0.118)	(0.106)	
Moisture in random seed bag in percent$^{(\dagger)}$	13.354	-0.025	-0.105	-0.169	224
	(1.232)	(0.216)	(0.210)	(0.166)	
Index of **capital-intensive** seed handling and storage practices[1]†	0.007	-0.204*	-0.002	-0.029	232
	(0.480)	(0.084)	(0.071)	(0.065)	
Index of **labor-intensive** seed handling and storage practices[2]†	0.007	0.043	0.025	0.058	239
	(0.445)	(0.071)	(0.070)	(0.057)	
Index of **all** seed handling and storage practices[3]	0.017	-0.030	0.033	0.043	217
	(0.363)	(0.064)	(0.055)	(0.049)	
Index of dealer's efforts and services[4]†	0.000	-0.066	0.054	-0.067	260
	(0.381)	(0.055)	(0.052)	(0.047)	
Index of shop's maize seed ratings by farmers[5]$^{(\dagger)}$	0.005	0.008	0.061	0.008	213
	(0.719)	(0.133)	(0.129)	(0.103)	
Overall index controlling for baseline	0.007	-0.077	0.200*	0.014	199
	(0.613)	(0.090)	(0.085)	(0.074)	
Overall index not controlling for baseline	-0.001	-0.096	0.115	0.026	185
	(0.491)	(0.100)	(0.094)	(0.072)	
Max. number of observations for outcomes from dealer survey					297

Note: 1st column reports sample means and standard deviations below in brackets; 2nd-4th column reports differences between treatment and control groups and standard errors below in brackets; they are clustered at the level of randomization; 5th column reports number of observations; **, * and + denote significance at the 1, 5 and 10 percent levels; † indicates that the variable is included in the overall index; larger indices indicate more desirable outcomes.

[1]The index of capital-intensive seed handling and storage practices contains 6 variables: whether roof is leak-proof, whether roof is insulated, whether walls are insulated, whether shop is ventilated, whether any official certificate is displayed, whether expired seed is handled correctly. This index ranges from -2.54 to 0.56 at endline.

[2]The index of labor-intensive seed handling and storage practices contains 6 variables: whether seed is stored in dedicated area, whether shop has no pest problem, whether seed is stored in correct lighting, whether seed is stored on correct surface, whether seed is not stored in open containers, cleanness and professionality rating by enumerator. This index ranges from -1.21 to 0.81 at endline.

[3]The index of all seed handling and storage practices contains 12 variables: the ones included in the index of capital-intensive practices and the ones included in the index of labor-intensive practices. This index ranges from -1.64 to 0.64 at endline. For only 217 dealers, none of the 12 variables constituting this index is missing at base- or endline.

[4]The index of dealer's efforts and services contains 7 variables: whether shop offers explanations, complementary input recommendations, extension/training, discounts for larger quantities, credit, did not receive seed related customer complaint, accepts mobile money. This index ranges from -0.84 to 0.70 at endline.

[5]The index of shop's maize seed ratings by farmers contains 6 ratings: general quality, yield, drought tolerance, pest/disease tolerance, time of maturity, germination. This index ranges from -2.42 to 2.57 at endline.

Table 13: Differences between treatment and control groups - Agro-input dealer, primary outcome variables: P-values adjusted according to Sankoh, Huque, Dubey (1997)

	mean	dealer training	clearing house	farmer video	nobs
Transformed quantity sold of 4 improved maize varieties last season in kg (IHS)†	6.035	-0.633^{+}	0.098	-0.104	249
	(1.989)	(0.257)	(0.249)	(0.225)	
Average sales price of 4 improved maize varieties last season in UGX/kg	4677.213	-23.658	149.095	2.010	227
	(925.835)	(134.322)	(132.987)	(123.177)	
Transformed revenue from 4 improved maize varieties in mln UGX (IHS)†	1.247	-0.236	0.055	-0.057	249
	(0.977)	(0.107)	(0.106)	(0.098)	
Transformed number of customers who bought maize seed per day last season (IHS)†	2.996	-0.232	0.226	-0.035	251
	(0.887)	(0.122)	(0.118)	(0.106)	
Moisture in random seed bag in percent$^{(\dagger)}$					
Index of **capital-intensive** seed handling and storage practices^{1†}	0.007	-0.204^{+}	-0.002	-0.029	232
	(0.480)	(0.084)	(0.071)	(0.065)	
Index of **labor-intensive** seed handling and storage practices^{2†}	0.007	0.043	0.025	0.058	239
	(0.445)	(0.071)	(0.070)	(0.057)	
Index of **all** seed handling and storage practices^{3}	0.017	-0.030	0.033	0.043	217
	(0.363)	(0.064)	(0.055)	(0.049)	
Index of dealer's efforts and services^{4†}	0.000	-0.066	0.054	-0.067	260
	(0.381)	(0.055)	(0.052)	(0.047)	
Index of shop's maize seed ratings by farmers$^{5(\dagger)}$					
Max. number of observations for outcomes from dealer survey					297

Note: 1st column reports sample means and standard deviations below in brackets; 2nd-4th column reports differences between treatment and control groups and standard errors below in brackets; they are clustered at the level of randomization; 5th column reports number of observations; **, * and + denote significance at the 1, 5 and 10 percent levels; † indicates that the variable is included in the overall index; larger indices indicate more desirable outcomes.

[1] The index of capital-intensive seed handling and storage practices contains 6 variables: whether roof is leak-proof, whether roof is insulated, whether walls are insulated, whether shop is ventilated, whether any official certificate is displayed, whether expired seed is handled correctly. This index ranges from -2.54 to 0.56 at endline.

[2] The index of labor-intensive seed handling and storage practices contains 6 variables: whether seed is stored in dedicated area, whether shop has no pest problem, whether seed is stored in correct lighting, whether seed is stored on correct surface, whether seed is not stored in open containers, cleanness and professionality rating by enumerator. This index ranges from -1.21 to 0.81 at endline.

[3] The index of all seed handling and storage practices contains 12 variables: the ones included in the index of capital-intensive practices and the ones included in the index of labor-intensive practices. This index ranges from -1.64 to 0.64 at endline. For only 217 dealers, none of the 12 variables constituting this index is missing at base- or endline.

[4] The index of dealer's efforts and services contains 7 variables: whether shop offers explanations, complementary input recommendations, extension/training, discounts for larger quantities, credit, did not receive seed related customer complaint, accepts mobile money. This index ranges from -0.84 to 0.70 at endline.

[5] The index of shop's maize seed ratings by farmers contains 6 ratings: general quality, yield, drought tolerance, pest/disease tolerance, time of maturity, germination. This index ranges from -2.42 to 2.57 at endline.

Table 14: Differences between treatment and control groups - Agro-input dealer, secondary outcome variables regarding indices

	mean	dealer training	clearing house	farmer video	nobs
Index of dealer's motivation and satisfaction[1]	0.001	-0.033	-0.014	-0.048	221
	(0.638)	(0.093)	(0.091)	(0.092)	
Index of dealer's self-ratings[2]	0.000	-0.152	0.071	0.083	260
	(0.654)	(0.098)	(0.085)	(0.080)	
Index of dealer's efforts and services according to farmers[3]	0.003	0.012	0.092	-0.022	231
	(0.563)	(0.089)	(0.085)	(0.076)	
Index of dealer's knowledge about seed storage[4]	0.000	0.022	0.092	0.063	260
	(0.490)	(0.057)	(0.055)	(0.060)	
Index of dealer's knowledge about seed[5]	0.000	-0.110	-0.054	-0.018	260
	(0.572)	(0.073)	(0.079)	(0.070)	
Max. number of observations					297

Note: 1st column reports sample means and standard deviations below in brackets; 2nd-4th column reports differences between treatment and control groups and standard errors below in brackets; they are clustered at the level of randomization; 5th column reports number of observations; **, * and + denote significance at the 1, 5 and 10 percent levels; † indicates that the variable is included in the overall index; larger indices indicate more desirable outcomes.

[1]The index of dealer's motivation and satisfaction contains 3 variables: whether dealer sees him-/ herself working as agro-input dealer in future, would recommend working as dealer, how happy dealer feels when he/ she comes to work. This index ranges from -3.33 to 0.61 at endline.

[2]The index of dealer's self-ratings contains 5 ratings: location, price, product quality, stock, reputation. This index ranges from -3.25 to 0.89 at endline.

[3]The index of dealer's efforts and services according to farmers contains 7 variables: whether shop offers refund/insurance, credit, training/advice, delivery, after-sales service, accepts different payment methods, sells small quantities. This index ranges from -2.44 to 1.38 at endline.

[4]The index of dealer's knowledge about seed storage contains 5 variables: whether dealer knows how long seed can be carried over, how seed should be stored after repackaging, what the min. distance between floor and seed is, how seed should be stored in storeroom, whether seed should be repackaged. This index ranges from -1.04 to 1.11 at endline.

[5]The index of dealer's knowledge about seed contains 4 variables: whether dealer knows which seed variety to recommend if farmer complains about poor soil, if farmer complains about little rain, if farmer is late for planting, what to tell clients about yield benefits of hybrid seed. This index ranges from -0.60 to 1.84 at endline.

Table 15: Differences between treatment and control groups - Agro-input dealer, secondary outcome variables

	mean	dealer training	clearing house	farmer video	nobs
Number of maize varieties in stock (incl. hybrids, OPV, landraces)	3.405	-0.357^{+}	-0.041	-0.003	254
	(1.554)	(0.200)	(0.196)	(0.185)	
Number of **hybrid** maize varieties in stock†	2.272	-0.349*	0.086	-0.097	254
	(1.309)	(0.155)	(0.154)	(0.146)	
Number of **open-pollinated** maize varieties in stock†	1.190	-0.107	-0.074	0.093	255
	(0.737)	(0.101)	(0.101)	(0.090)	
Shop has equipment to monitor seed moisture	0.327	0.653**	-0.046	0.035	260
	(0.470)	(0.049)	(0.046)	(0.043)	
Overall index controlling for baseline	-0.012	-0.212*	0.012	0.005	250
	(0.747)	(0.086)	(0.092)	(0.091)	
Overall index not controlling for baseline	0.059	-0.155	-0.042	0.081	224
	(0.704)	(0.094)	(0.100)	(0.098)	
Max. number of observations					297

Note: 1st column reports sample means and standard deviations below in brackets; 2nd-4th column reports differences between treatment and control groups and standard errors below in brackets; they are clustered at the level of randomization; 5th column reports number of observations; **, * and + denote significance at the 1, 5 and 10 percent levels; † indicates that the variable is included in the overall index; larger indices indicate more desirable outcomes.

Table 16: Differences between treatment and control groups - Agro-input dealer, secondary outcome variables: P-values adjusted according to Sankoh, Huque, Dubey (1997)

	mean	dealer training	clearing house	farmer video	nobs
Number of maize varieties in stock (incl. hybrids, OPV, landraces)	3.405	-0.357	-0.041	-0.003	254
	(1.554)	(0.200)	(0.196)	(0.185)	
Number of **hybrid** maize varieties in stock†	2.272	-0.349^{+}	0.086	-0.097	254
	(1.309)	(0.155)	(0.154)	(0.146)	
Number of **open-pollinated** maize varieties in stock†	1.190	-0.107	-0.074	0.093	255
	(0.737)	(0.101)	(0.101)	(0.090)	
Shop has equipment to monitor seed moisture	0.327	0.653**	-0.046	0.035	260
	(0.470)	(0.049)	(0.046)	(0.043)	
Max. number of observations					297

Note: 1st column reports sample means and standard deviations below in brackets; 2nd-4th column reports differences between treatment and control groups and standard errors below in brackets; they are clustered at the level of randomization; 5th column reports number of observations; **, * and + denote significance at the 1, 5 and 10 percent levels; † indicates that the variable is included in the overall index; larger indices indicate more desirable outcomes.

included because the impact of the treatments on this outcome is ambiguous (increased adoption could eg. increase demand and hence prices but dealers could also lower prices to be more customer friendly). The transformed amount of Longe 10H lost/wasted last season (IHS) is not included because many observations are missing at baseline. After midline data collection, we will check whether as many observations are missing and if we have significantly more observations, include the outcome using a regression that does not control for the baseline value. The number of times the shop ran out of Longe 10H last season is not included because the impact of the treatments on this outcome is ambiguous (increased adoption could eg. increase demand and hence this number but the training could improve management and planning, hence decrease it). Furthermore, this number mostly depends on parameters further up the value chain.

Longe 5 Table 19 looks at farmer level outcomes related to a popular OPV maize seed, while Table 20 gives the same information but adjusted for multiple comparisons.

Some outcome variables are not included in the overall index: please see explanations for Longe 10H.

Registration/ trading license/ membership/ inspection We now turn to outcomes that are related to registration and quality control (Tables 21 and 22).

Some outcome variables are not included in the overall index: The transformed number of times shop was inspected by DAO/MAAIF/UNADA last year (IHS) is not included because the impact of the treatments on this outcome is ambiguous (eg. more dealers could ask to be inspected to receive a certificate and signal quality which would increase the number but the treatments could also improve quality which could make inspections less necessary and common).

Seed bag Finally, in each shop, we purchase a random seed bag which is then analyzed. Tables 23 and 24 provides comparisons.

Farmer

We now turn to the impact of the interventions at the farmer level.

Primary

We first define a set of key outcomes. These are in Tables 25 and 26.

Some outcome variables are not included in the overall index: The index of farmer's maize seed ratings for shops nearby and the index of farmer's general ratings of shops nearby are not included because the questions to compute these indeces are only collected from all farmers at endline (at midline, these variables are only collected from clearinghouse treated farmers as rating is part of the clearinghouse treatment). The index of services of shops nearby according

Table 17: Differences between treatment and control groups - Agro-input dealer, secondary outcome variables regarding Longe 10H (most common hybrid variety in area) (selection: if shop had Longe 10H in stock last season)

	mean	dealer training	clearing house	farmer video	nobs
Transformed amount of Longe 10H carried forward in kg (IHS)†	0.308	0.007	-0.129	0.158	214
	(0.965)	(0.134)	(0.126)	(0.132)	
Transformed amount of Longe 10H shop bought from provider in kg (IHS)†	5.709	-0.188	0.024	-0.028	206
	(1.322)	(0.200)	(0.193)	(0.162)	
Transformed cost of Longe 10H in UGX/kg (IHS)[1]	9.264	-0.015	0.033	0.003	150
	(0.174)	(0.035)	(0.036)	(0.027)	
Transformed quantity sold of Longe 10H in kg (IHS)†	5.634	-0.312	0.054	0.053	206
	(1.338)	(0.199)	(0.195)	(0.164)	
Transformed sales price of Longe 10H in UGX/kg (IHS)[2]	9.451	-0.013	0.026	0.004	158
	(0.181)	(0.030)	(0.030)	(0.027)	
Transformed amount of Longe 10H lost/wasted in kg (IHS)$^{(\dagger)}$	0.047	-0.053	-0.047	-0.055	207
	(0.315)	(0.041)	(0.044)	(0.045)	
Transformed number of times per month shop ran out of Longe 10H (IHS)[3]$^{(\dagger)}$	0.530	-0.193	-0.221	0.094	157
	(0.701)	(0.144)	(0.144)	(0.118)	
Overall index controlling for baseline	0.004	-0.125	0.146	-0.041	197
	(0.705)	(0.104)	(0.097)	(0.100)	
Overall index not controlling for baseline	0.011	0.011	0.163*	-0.012	199
	(0.472)	(0.067)	(0.068)	(0.062)	
Max. number of observations					254

Note: 1st column reports sample means and standard deviations below in brackets; 2nd-4th column reports differences between treatment and control groups and standard errors below in brackets; they are clustered at the level of randomization; 5th column reports number of observations; **, * and + denote significance at the 1, 5 and 10 percent levels; † indicates that the variable is included in the overall index; larger indices indicate more desirable outcomes.
[1]Only 150 dealers had Longe 10H in stock in the season before endline and answered this question at base- and endline (182 at baseline and 249 at endline).
[2]Only 158 dealers had Longe 10H in stock in the season before endline and answered this question at base- and endline (189 at baseline and 252 at endline).
[3]Only 157 dealers had Longe 10H in stock in the season before endline and answered this question at base- and endline (188 at baseline and 249 at endline).

Table 18: Differences between treatment and control groups - Agro-input dealer, secondary outcome variables regarding Longe 10H (most common hybrid variety in area): P-values adjusted according to Sankoh, Huque, Dubey (1997) (selection: if shop had Longe 10H in stock last season)

	mean	dealer training	clearing house	farmer video	nobs
Transformed amount of Longe 10H carried forward in kg (IHS)†	0.308	0.007	-0.129	0.158	214
	(0.965)	(0.134)	(0.126)	(0.132)	
Transformed amount of Longe 10H shop bought from provider in kg (IHS)†	5.709	-0.188	0.024	-0.028	206
	(1.322)	(0.200)	(0.193)	(0.162)	
Transformed cost of Longe 10H in UGX/kg (IHS)[1]	9.264	-0.015	0.033	0.003	150
	(0.174)	(0.035)	(0.036)	(0.027)	
Transformed quantity sold of Longe 10H in kg (IHS)†	5.634	-0.312	0.054	0.053	206
	(1.338)	(0.199)	(0.195)	(0.164)	
Transformed sales price of Longe 10H in UGX/kg (IHS)[2]	9.451	-0.013	0.026	0.004	158
	(0.181)	(0.030)	(0.030)	(0.027)	
Transformed amount of Longe 10H lost/wasted in kg (IHS)$^{(\dagger)}$	0.047	-0.053	-0.047	-0.055	207
	(0.315)	(0.041)	(0.044)	(0.045)	
Transformed number of times per month shop ran out of Longe 10H (IHS)[3]$^{(\dagger)}$	0.530	-0.193	-0.221	0.094	157
	(0.701)	(0.144)	(0.144)	(0.118)	
Max. number of observations					254

Note: 1st column reports sample means and standard deviations below in brackets; 2nd-4th column reports differences between treatment and control groups and standard errors below in brackets; they are clustered at the level of randomization; 5th column reports number of observations; **, * and + denote significance at the 1, 5 and 10 percent levels; † indicates that the variable is included in the overall index; larger indices indicate more desirable outcomes.
[1] Only 150 dealers had Longe 10H in stock in the season before endline and answered this question at base- and endline (182 at baseline and 249 at endline).
[2] Only 158 dealers had Longe 10H in stock in the season before endline and answered this question at base- and endline (189 at baseline and 252 at endline).
[3] Only 157 dealers had Longe 10H in stock in the season before endline and answered this question at base- and endline (188 at baseline and 249 at endline).

Table 19: Differences between treatment and control groups - Agro-input dealer, secondary outcome variables regarding Longe 5 (most common open-pollinated variety) (selection: if shop had Longe 5 in stock last season)

	mean	dealer training	clearing house	farmer video	nobs
Transformed amount of Longe 5 carried forward in kg (IHS)†	0.313	-0.052	-0.095	0.222^{+}	228
	(0.948)	(0.145)	(0.148)	(0.127)	
Transformed amount of Longe 5 bought by shop from provider in kg (IHS)†	5.782	-0.311	0.137	-0.074	225
	(1.298)	(0.203)	(0.197)	(0.161)	
Transformed cost of Longe 5 in UGX/kg (IHS)	8.581	0.004	0.024	0.002	197
	(0.148)	(0.019)	(0.019)	(0.019)	
Transformed quantity of Longe 5 sold in kg (IHS)†	5.746	-0.361^{+}	0.149	-0.036	223
	(1.315)	(0.200)	(0.195)	(0.163)	
Transformed sales price of Longe 5 in UGX/kg (IHS)	8.800	0.013	0.028	0.003	206
	(0.165)	(0.023)	(0.023)	(0.019)	
Transformed amount of Longe 5 lost/wasted in kg (IHS)$^{(\dagger)}$	0.089	-0.053	-0.019	-0.029	227
	(0.443)	(0.062)	(0.065)	(0.060)	
Transformed number of times per month shop ran out of Longe 5 (IHS)$^{(\dagger)}$	0.372	0.007	-0.122	0.050	202
	(0.636)	(0.121)	(0.125)	(0.091)	
Overall index controlling for baseline	-0.002	-0.125	0.119	-0.162	217
	(0.732)	(0.110)	(0.107)	(0.107)	
Overall index not controlling for baseline	-0.007	-0.043	0.139^{+}	-0.021	217
	(0.525)	(0.073)	(0.073)	(0.074)	
Max. number of observations					269

Note: 1st column reports sample means and standard deviations below in brackets; 2nd-4th column reports differences between treatment and control groups and standard errors below in brackets; they are clustered at the level of randomization; 5th column reports number of observations; **, * and + denote significance at the 1, 5 and 10 percent levels; † indicates that the variable is included in the overall index; larger indices indicate more desirable outcomes.

Table 20: Differences between treatment and control groups - Agro-input dealer, secondary outcome variables regarding Longe 5 (most common open-pollinated variety): P-values adjusted according to Sankoh, Huque, Dubey (selection: if shop had Longe 5 in stock last season)

	mean	dealer training	clearing house	farmer video	nobs
Transformed amount of Longe 5 carried forward in kg (IHS)†	0.313	-0.052	-0.095	0.222	228
	(0.948)	(0.145)	(0.148)	(0.127)	
Transformed amount of Longe 5 bought by shop from provider in kg (IHS)†	5.782	-0.311	0.137	-0.074	225
	(1.298)	(0.203)	(0.197)	(0.161)	
Transformed cost of Longe 5 in UGX/kg (IHS)†	8.581	0.004	0.024	0.002	197
	(0.148)	(0.019)	(0.019)	(0.019)	
Transformed quantity of Longe 5 sold in kg (IHS)†	5.746	-0.361	0.149	-0.036	223
	(1.315)	(0.200)	(0.195)	(0.163)	
Transformed sales price of Longe 5 in UGX/kg (IHS)	8.800	0.013	0.028	0.003	206
	(0.165)	(0.023)	(0.023)	(0.019)	
Transformed amount of Longe 5 lost/wasted in kg (IHS)$^{(\dagger)}$	0.089	-0.053	-0.019	-0.029	227
	(0.443)	(0.062)	(0.065)	(0.060)	
Transformed number of times per month shop ran out of Longe 5 (IHS)$^{(\dagger)}$	0.372	0.007	-0.122	0.050	202
	(0.636)	(0.121)	(0.125)	(0.091)	
Max. number of observations					269

Note: 1st column reports sample means and standard deviations below in brackets; 2nd-4th column reports differences between treatment and control groups and standard errors below in brackets; they are clustered at the level of randomization; 5th column reports number of observations; **, * and + denote significance at the 1, 5 and 10 percent levels; † indicates that the variable is included in the overall index; larger indices indicate more desirable outcomes.

Table 21: Differences between treatment and control groups - Agro-input dealer, secondary outcome variables regarding registration/ trading license/ membership/ inspection

	mean	dealer training	clearing house	farmer video	nobs
Shop is registered with UNADA†	0.365	-0.121^{+}	0.053	-0.001	223
	(0.482)	(0.065)	(0.065)	(0.060)	
Shop has a trading license issued by local government†	0.773	-0.062	-0.033	-0.032	249
	(0.419)	(0.059)	(0.058)	(0.051)	
Shop is a member of another professional association†	0.194	-0.073	0.053	-0.059	236
	(0.396)	(0.061)	(0.061)	(0.049)	
Transformed number of inspections by DAO/MAAIF/UNADA last year (IHS)†	0.839	-0.049	0.201^{+}	-0.115	238
	(0.676)	(0.104)	(0.106)	(0.083)	
Shop received a warning after inspection†	0.346	0.015	-0.014	0.075	247
	(0.477)	(0.069)	(0.069)	(0.061)	
Shop's products were confiscated after inspection†	0.062	-0.009	-0.035	0.038	248
	(0.242)	(0.036)	(0.037)	(0.030)	
Overall index controlling for baseline	-0.007	-0.057	0.119^{+}	-0.154*	190
	(0.445)	(0.069)	(0.069)	(0.064)	
Overall index not controlling for baseline	0.008	-0.050	0.129	-0.120^{+}	196
	(0.454)	(0.077)	(0.078)	(0.069)	
Max. number of observations					297

Note: 1st column reports sample means and standard deviations below in brackets; 2nd-4th column reports differences between treatment and control groups and standard errors below in brackets; they are clustered at the level of randomization; 5th column reports number of observations; **, * and $+$ denote significance at the 1, 5 and 10 percent levels; † indicates that the variable is included in the overall index; larger indices indicate more desirable outcomes.

Table 22: Differences between treatment and control groups - Agro-input dealer, secondary outcome variables regarding registration/ trading license/ membership/ inspection: P-values adjusted according to Sankoh, Huque, Dubey

	mean	dealer training	clearing house	farmer video	nobs
Shop is registered with UNADA†	0.365	-0.121	0.053	-0.001	223
	(0.482)	(0.065)	(0.065)	(0.060)	
Shop has a trading license issued by local government†	0.773	-0.062	-0.033	-0.032	249
	(0.419)	(0.059)	(0.058)	(0.051)	
Shop is a member of another professional association†	0.194	-0.073	0.053	-0.059	236
	(0.396)	(0.061)	(0.061)	(0.049)	
Transformed number of inspections by DAO/MAAIF/UNADA last year (IHS)†	0.839	-0.049	0.201	-0.115	238
	(0.676)	(0.104)	(0.106)	(0.083)	
Shop received a warning after inspection†	0.346	0.015	-0.014	0.075	247
	(0.477)	(0.069)	(0.069)	(0.061)	
Shop's products were confiscated after inspection†	0.062	-0.009	-0.035	0.038	248
	(0.242)	(0.036)	(0.037)	(0.030)	
Max. number of observations					297

Note: 1st column reports sample means and standard deviations below in brackets; 2nd-4th column reports differences between treatment and control groups and standard errors below in brackets; they are clustered at the level of randomization; 5th column reports number of observations; **, * and + denote significance at the 1, 5 and 10 percent levels; † indicates that the variable is included in the overall index; larger indices indicate more desirable outcomes.

Table 23: Differences between treatment and control groups - Agro-input dealer, secondary outcome variables regarding seed bag (selection: enumerator was able to buy bag of seed)

	mean	dealer training	clearing house	farmer video	nobs
Moisture in random seed bag in percent†	13.354	-0.025	-0.105	-0.169	224
	(1.232)	(0.216)	(0.210)	(0.166)	
Random seed bag shows packaging date†	0.754	-0.080	0.034	0.026	228
	(0.431)	(0.071)	(0.070)	(0.058)	
Days since packaging date/expiry date minus 6 months†	127.537	6.650	4.272	3.924	203
	(58.873)	(9.120)	(9.260)	(8.448)	
Seed is in the original bag without any signs of damage†	0.838	-0.030	0.039	-0.027	228
	(0.370)	(0.058)	(0.058)	(0.049)	
Random seed bag shows lot number†	0.711	-0.153*	-0.002	0.037	228
	(0.455)	(0.069)	(0.068)	(0.060)	
Overall index not controlling for baseline	-0.001	-0.079	0.066	0.018	199
	(0.608)	(0.104)	(0.097)	(0.087)	
Max. number of observations					265

Note: 1st column reports sample means and standard deviations below in brackets; 2nd-4th column reports differences between treatment and control groups and standard errors below in brackets; they are clustered at the level of randomization; 5th column reports number of observations; **, * and + denote significance at the 1, 5 and 10 percent levels; † indicates that the variable is included in the overall index; larger indices indicate more desirable outcomes. We do not control for the baseline values of the outcome variables here because only 183 of the 265 dealers who had seed at endline also had seed at baseline.

Table 24: Differences between treatment and control groups - Agro-input dealer, secondary outcome variables regarding seed bag: P-values adjusted according to Sankoh, Huque, Dubey (selection: enumerator was able to buy bag of seed)

	mean	dealer training	clearing house	farmer video	nobs
Moisture in random seed bag in percent†	13.354	-0.025	-0.105	-0.169	224
	(1.232)	(0.216)	(0.210)	(0.166)	
Random seed bag shows packaging date†	0.754	-0.080	0.034	0.026	228
	(0.431)	(0.071)	(0.070)	(0.058)	
Days since packaging date/expiry date minus 6 months†	127.537	6.650	4.272	3.924	203
	(58.873)	(9.120)	(9.260)	(8.448)	
Seed is in the original bag without any signs of damage†	0.838	-0.030	0.039	-0.027	228
	(0.370)	(0.058)	(0.058)	(0.049)	
Random seed bag shows lot number†	0.711	-0.153	-0.002	0.037	228
	(0.455)	(0.069)	(0.068)	(0.060)	
Max. number of observations					265

Note: 1st column reports sample means and standard deviations below in brackets; 2nd-4th column reports differences between treatment and control groups and standard errors below in brackets; they are clustered at the level of randomization; 5th column reports number of observations; **, * and + denote significance at the 1, 5 and 10 percent levels; † indicates that the variable is included in the overall index; larger indices indicate more desirable outcomes. We do not control for the baseline values of the outcome variables here because only 183 of the 265 dealers who had seed at endline also had seed at baseline.

to farmers is not included because many observations are missing at baseline. After midline data collection, we will check whether as many observations are missing and if we have significantly more observations, include the outcome using a regression that does not control for the baseline value. Whether the farmer switched to a different agro-input shop or not is only included in the regression that does not control for the baseline value because we did not ask this question at baseline.

Secondary

Other secondary We further register a family of secondary outcomes that are somewhat unrelated to each other in Tables 27 and 28. Some outcome variables are not included in the overall index: The index of farmer skills is not included because the questions to compute this index are only collected at endline to avoid priming.

Adoption on randomly selected maize field We also look at plot level outcomes (Tables 29 and 30).

Some outcome variables are not included in the overall index: Whether the farmer planted hybrid not farmer saved seed or an OPV (not used too often) on this field is not included because it is a function of whether the farmer planted hybrid maize seed and whether the farmer planted open-pollinated maize seed, which are both included in the index.

Seed on randomly selected maize field We now go more into detail with respect the the seed that was used on the randomly selected field (Tables 31 and 32). Some outcome variables are not included in the overall index: The transformed cost of seed farmer used on this field last season (amount * price) (IHS) is not included because it is a function of the amount of seed and the price of seed, which are both included in the index.

Yield etc. on randomly selected maize field Finally, we look at production related outcomes and disposal of maize (Tables 33 and 34).

Some outcome variables are not included in the overall index: Production from this field last season is not included because it is included in the yield, which is included in the index. Whether the farmer harvested as much maize as expected from this field last season and whether he/she did not harvest as much maize as expected due to own mismanagement and the transformed amount kept as seed (IHS) are only included in the regressions that do not control for the baseline values because we did not ask these questions at baseline.

5 conclusion

A crowd-sourced information clearinghouse can be an important institutional innovation to solve the problem of asymmetric information in the market for

Table 25: Differences between treatment and control groups - Farmer, primary outcome variables

	mean	dealer training	clearing house	farmer video	nobs
Farmer used quality maize seed for any plot†	0.663	-0.002	0.043^{+}	0.003	2857
	(0.473)	(0.024)	(0.024)	(0.017)	
Farmer bought quality maize seed at agro-input shop for any plot†	0.463	0.009	0.028	-0.013	2811
	(0.499)	(0.024)	(0.023)	(0.019)	
Amount of quality maize seed farmer bought at agro-input shop in kg[1]	6.551	0.348	0.250	-0.682^{+}	531
	(4.666)	(0.428)	(0.434)	(0.383)	
Index of farmer's maize seed ratings for shops nearby (product quality perception)$^{2(\dagger)}$	0.005	0.032	0.086	0.050	1433
	(0.639)	(0.067)	(0.066)	(0.034)	
Index of farmer's general ratings of shops nearby (shop/seller perception)$^{3(\dagger)}$	0.000	-0.001	-0.003	0.000	1472
	(0.641)	(0.054)	(0.051)	(0.033)	
Index of services of shops nearby according to farmers (effort perception)[4]	-0.020	0.044	0.152	0.127^{+}	282
	(0.593)	(0.106)	(0.098)	(0.069)	
Farmer switched to different agro-input shop$^{(\dagger)}$	0.162	-0.021	0.016	-0.009	2997
	(0.369)	(0.017)	(0.016)	(0.014)	
Index of farmer's practices on randomly selected field^{5†}	0.005	0.001	0.020	0.018	2653
	(0.404)	(0.025)	(0.025)	(0.016)	
Farmer thinks that maize seed at agro-input shops is counterfeit/adulterated[6]	0.587	-0.050	0.023	0.045^{*}	1865
	(0.493)	(0.031)	(0.031)	(0.023)	
Farmer used local land race maize seed on randomly selected field†	0.322	-0.002	-0.033	-0.007	2640
	(0.467)	(0.025)	(0.025)	(0.018)	
Overall index controlling for baseline	0.003	0.001	0.071^{+}	0.030	2411
	(0.765)	(0.040)	(0.039)	(0.030)	
Overall index not controlling for baseline	0.122	-0.026	0.072^{*}	0.021	1375
	(0.491)	(0.036)	(0.035)	(0.027)	
Max. number of observations					3441

Note: 1st column reports sample means and standard deviations below in brackets; 2nd-4th column reports differences between treatment and control groups and standard errors below in brackets; they are clustered at the level of randomization; 5th column reports number of observations; **, * and + denote significance at the 1, 5 and 10 percent levels; † indicates that the variable is included in the overall index; larger indices indicate more desirable outcomes.

[1]Only 531 farmers are included in this regression because the variable is defined only for farmers who reported that they bought quality maize seed at an agro-input shop at baseline and at endline (669 farmers). The remaining missing 48 farmers bought quality maize seed at an agro-input shop at base- and endline but did not answer the question about the amount at base- or endline.

[2]The index of farmer's maize seed ratings for shops nearby contains 6 ratings: general quality, yield, drought tolerance, pest/disease tolerance, time of maturity, germination. This index ranges from -3.13 to 1.87 at endline. Only 1433 farmers rated at least one shop nearby on all 6 seed characteristics at endline.

[3]The index of farmer's general ratings of shops nearby contains 6 ratings: general quality, location, price, product quality, stock, reputation. This index ranges from -3.13 to 1.12 at endline. Only 1472 farmers rated at least one shop nearby on all 6 general characteristics at endline.

[4]The index of services of shops nearby according to farmers contains 7 variables: whether shop offers refund/insurance, credit, training/advice, delivery, after-sales service, accepts different payment methods, sells small quantities. This index ranges from -1.84 to 0.83 at endline. Only 282 farmers answered all 7 questions for at least one shop nearby at baseline and at endline.

[5]The index of farmer's practices on randomly selected field contains 10 variables: whether farmer spaced seed correctly, sowed correct number of seeds/hill, applied organic manure, DAP/NPK, Urea, pesticides/herbicides/fungicides, weeded sufficiently, weeded at correct time, planted at correct time, re-sowed. This index ranges from -0.89 to 1.27 at endline.

[6]Only 1865 farmers answered this question at baseline and at endline.

Table 26: Differences between treatment and control groups - Farmer, primary outcome variables: P-values adjusted according to Sankoh, Huque, Dubey (1997)

	mean	dealer training	clearing house	farmer video	nobs
Farmer used quality maize seed for any plot†	0.663	-0.002	0.043	0.003	2857
	(0.473)	(0.024)	(0.024)	(0.017)	
Farmer bought quality maize seed at agro-input shop for any plot†	0.463	0.009	0.028	-0.013	2811
	(0.499)	(0.024)	(0.023)	(0.019)	
Amount of quality maize seed farmer bought at agro-input shop in kg[1]	6.551	0.348	0.250	-0.682	531
	(4.666)	(0.428)	(0.434)	(0.383)	
Index of farmer's maize seed ratings for shops nearby (product quality perception)[2(†)]					
Index of farmer's general ratings of shops nearby (shop/seller perception)[3(†)]					
Index of services of shops nearby according to farmers (effort perception)[4]	-0.020	0.044	0.152	0.127	282
	(0.593)	(0.106)	(0.098)	(0.069)	
Farmer switched to different agro-input shop[(†)]					
Index of farmer's practices on randomly selected field[5†]	0.005	0.001	0.020	0.018	2653
	(0.404)	(0.025)	(0.025)	(0.016)	
Farmer thinks that maize seed at agro-input shops is counterfeit/adulterated[6]	0.587	-0.050	0.023	0.045	1865
	(0.493)	(0.031)	(0.031)	(0.023)	
Farmer used local land race maize seed on randomly selected field†	0.322	-0.002	-0.033	-0.007	2640
	(0.467)	(0.025)	(0.025)	(0.018)	
Max. number of observations					3441

Note: 1st column reports sample means and standard deviations below in brackets; 2nd-4th column reports differences between treatment and control groups and standard errors below in brackets; they are clustered at the level of randomization; 5th column reports number of observations; **, * and + denote significance at the 1, 5 and 10 percent levels; † indicates that the variable is included in the overall index; larger indices indicate more desirable outcomes.
[1]Only 531 farmers are included in this regression because the variable is defined only for farmers who reported that they bought quality maize seed at an agro-input shop at baseline and at endline (669 farmers). The remaining missing 48 farmers bought quality maize seed at an agro-input shop at base- and endline but did not answer the question about the amount at base- or endline.
[2]The index of farmer's maize seed ratings for shops nearby contains 6 ratings: general quality, yield, drought tolerance, pest/disease tolerance, time of maturity, germination. This index ranges from -3.13 to 1.87 at endline. Only 1433 farmers rated at least one shop nearby on all 6 seed characteristics at endline.
[3]The index of farmer's general ratings of shops nearby contains 6 ratings: general quality, location, price, product quality, stock, reputation. This index ranges from -3.13 to 1.12 at endline. Only 1472 farmers rated at least one shop nearby on all 6 general characteristics at endline.
[4]The index of services of shops nearby according to farmers contains 7 variables: whether shop offers refund/insurance, credit, training/advice, delivery, after-sales service, accepts different payment methods, sells small quantities. This index ranges from -1.84 to 0.83 at endline. Only 282 farmers answered all 7 questions for at least one shop nearby at baseline and at endline.
[5]The index of farmer's practices on randomly selected field contains 10 variables: whether farmer spaced seed correctly, sowed correct number of seeds/hill, applied organic manure, DAP/NPK, Urea, pesticides/herbicides/fungicides, weeded sufficiently, weeded at correct time, planted at correct time, re-sowed. This index ranges from -0.89 to 1.27 at endline.
[6]Only 1865 farmers answered this question at baseline and at endline.

Table 27: Differences between treatment and control groups - Farmer, secondary outcome variables

	mean	dealer training	clearing house	farmer video	nobs
Number of improved maize varieties farmer knows†	2.846	0.077	0.031	0.062	2816
	(1.549)	(0.111)	(0.107)	(0.057)	
Farmer knows particular shop in neighborhood†	0.726	0.073*	0.081**	0.014	2857
	(0.358)	(0.029)	(0.030)	(0.013)	
Farmer bought particular shop's seed last season$^{(\dagger)}$	0.246	-0.017	0.047*	0.016	2870
	(0.411)	(0.023)	(0.022)	(0.015)	
Index of farmer's knowledge how to manage improved maize seed$^{1(\dagger)}$	0.000	0.008	0.020	0.028	2870
	(0.582)	(0.042)	(0.042)	(0.021)	
Overall index controlling for baseline	-0.001	0.127*	0.124*	0.041	2803
	(0.789)	(0.061)	(0.059)	(0.028)	
Overall index not controlling for baseline	0.000	0.069^{+}	0.038	0.039^{+}	2849
	(0.566)	(0.039)	(0.038)	(0.021)	
Max. number of observations					3441

Note: 1st column reports sample means and standard deviations below in brackets; 2nd-4th column reports differences between treatment and control groups and standard errors below in brackets; they are clustered at the level of randomization; 5th column reports number of observations; **, * and + denote significance at the 1, 5 and 10 percent levels; † indicates that the variable is included in the overall index; larger indices indicate more desirable outcomes.
[1]The index of farmer's knowledge how to manage improved maize seed contains 5 variables: whether farmers know that when using quality maize seed: 1) good seed handling/management is equally or even more important than when using lower quality seed, that they should 2) weed and remove striga 3) apply the same amount or even more fertilizer, 4) use equally good plots as they would if they'd use lower quality seed, 5) buy both quality seed and fertilizer. This index ranges from -2.30 to 0.46 at endline.

Table 28: Differences between treatment and control groups - Farmer, secondary outcome variables: P-values adjusted according to Sankoh, Huque, Dubey (1997)

	mean	dealer training	clearing house	farmer video	nobs
Number of improved maize varieties farmer knows†	2.846	0.077^{NA}	0.031^{NA}	0.062^{NA}	2816
	(1.549)	(0.111)	(0.107)	(0.057)	
Farmer knows particular shop in neighborhood†	0.726	0.073^{NA}	0.081^{NA}	0.014^{NA}	2857
	(0.358)	(0.029)	(0.030)	(0.013)	
Farmer bought particular shop's seed last season$^{(\dagger)}$					
Index of farmer's knowledge how to manage improved maize seed$^{1(\dagger)}$					
Max. number of observations					3441

Note: 1st column reports sample means and standard deviations below in brackets; 2nd-4th column reports differences between treatment and control groups and standard errors below in brackets; they are clustered at the level of randomization; 5th column reports number of observations; **, * and + denote significance at the 1, 5 and 10 percent levels; † indicates that the variable is included in the overall index; larger indices indicate more desirable outcomes.
[1]The index of farmer's knowledge how to manage improved maize seed contains 5 variables: whether farmers know that when using quality maize seed: 1) good seed handling/management is equally or even more important than when using lower quality seed, that they should 2) weed and remove striga 3) apply the same amount or even more fertilizer, 4) use equally good plots as they would if they'd use lower quality seed, 5) buy both quality seed and fertilizer. This index ranges from -2.30 to 0.46 at endline.

Table 29: Differences between treatment and control groups - Farmer, secondary outcome variables: adoption on randomly selected maize field

	mean	dealer training	clearing house	farmer video	nobs
Farmer planted hybrid seed†	0.351	-0.004	0.039	0.022	2329
	(0.477)	(0.027)	(0.026)	(0.019)	
Farmer planted open-pollinated seed†	0.304	0.000	-0.005	-0.013	2329
	(0.460)	(0.025)	(0.024)	(0.019)	
Farmer planted farmer saved seed†	0.496	-0.011	-0.014	0.005	2820
	(0.500)	(0.024)	(0.023)	(0.018)	
Farmer planted seed bought at agro-input shop†	0.448	0.013	0.032	-0.002	2820
	(0.497)	(0.024)	(0.023)	(0.018)	
Farmer planted hybrid **not farmer saved** seed or an OPV (**not recycled too often**)	0.582	-0.001	0.022	-0.009	2640
	(0.493)	(0.025)	(0.024)	(0.019)	
Overall index controlling for baseline	0.003	-0.009	0.039	0.015	2294
	(0.494)	(0.028)	(0.027)	(0.020)	
Overall index not controlling for baseline	0.003	0.001	0.035	0.019	2572
	(0.494)	(0.030)	(0.029)	(0.020)	
Max. number of observations					3441

Note: 1st column reports sample means and standard deviations below in brackets; 2nd-4th column reports differences between treatment and control groups and standard errors below in brackets; they are clustered at the level of randomization; 5th column reports number of observations; **, * and + denote significance at the 1, 5 and 10 percent levels; † indicates that the variable is included in the overall index; larger indices indicate more desirable outcomes.

Table 30: Differences between treatment and control groups - Farmer, secondary outcome variables: adoption on randomly selected maize field: P-values adjusted according to Sankoh, Huque, Dubey (1997)

	mean	dealer training	clearing house	farmer video	nobs
Farmer planted hybrid seed†	0.351	-0.004	0.039	0.022	2329
	(0.477)	(0.027)	(0.026)	(0.019)	
Farmer planted open-pollinated seed†	0.304	0.000	-0.005	-0.013	2329
	(0.460)	(0.025)	(0.024)	(0.019)	
Farmer planted farmer saved seed†	0.496	-0.011	-0.014	0.005	2820
	(0.500)	(0.024)	(0.023)	(0.018)	
Farmer planted seed bought at agro-input shop†	0.448	0.013	0.032	-0.002	2820
	(0.497)	(0.024)	(0.023)	(0.018)	
Farmer planted hybrid **not farmer saved** seed or an OPV (**not recycled too often**)	0.582	-0.001	0.022	-0.009	2640
	(0.493)	(0.025)	(0.024)	(0.019)	
Max. number of observations					3441

Note: 1st column reports sample means and standard deviations below in brackets; 2nd-4th column reports differences between treatment and control groups and standard errors below in brackets; they are clustered at the level of randomization; 5th column reports number of observations; **, * and + denote significance at the 1, 5 and 10 percent levels; † indicates that the variable is included in the overall index; larger indices indicate more desirable outcomes.

Table 31: Differences between treatment and control groups - Farmer, secondary outcome variables: seed used on randomly selected maize field

	mean	dealer training	clearing house	farmer video	nobs
Index of farmer's seed ratings[1] †	0.001	0.055	0.065	-0.025	2713
	(0.607)	(0.045)	(0.045)	(0.023)	
Farmer was satisfied with seed quality†	0.727	0.022	-0.003	-0.029^{+}	2870
	(0.446)	(0.025)	(0.025)	(0.017)	
Farmer would use seed again†	0.756	0.011	0.015	-0.021	2870
	(0.430)	(0.023)	(0.023)	(0.016)	
Amount of seed farmer used in kg†	6.429	-0.137	-0.138	-0.159	2610
	(4.596)	(0.268)	(0.271)	(0.175)	
Price of seed in UGX/ kg†	2542.690	150.630	171.476	-31.071	2644
	(3220.703)	(170.103)	(168.250)	(120.216)	
Transformed cost of seed in UGX (IHS)	5.048	0.318	0.358	0.099	2564
	(5.351)	(0.259)	(0.252)	(0.207)	
Overall index controlling for baseline	-0.004	0.056	0.055	0.001	2272
	(0.593)	(0.036)	(0.036)	(0.024)	
Overall index not controlling for baseline	-0.004	0.037	0.046	-0.005	2573
	(0.593)	(0.037)	(0.037)	(0.024)	
Max. number of observations					3441

Note: 1st column reports sample means and standard deviations below in brackets; 2nd-4th column reports differences between treatment and control groups and standard errors below in brackets; they are clustered at the level of randomization; 5th column reports number of observations; **, * and + denote significance at the 1, 5 and 10 percent levels; † indicates that the variable is included in the overall index; larger indices indicate more desirable outcomes.

[1]The index of farmer's seed ratings contains 6 ratings: general quality, yield, drought tolerance, pest/disease tolerance, time of maturity, germination. This index ranges from -2.26 to 1.76 at endline.

Table 32: Differences between treatment and control groups - Farmer, secondary outcome variables: seed used on randomly selected maize field: P-values adjusted according to Sankoh, Huque, Dubey (1997)

	mean	dealer training	clearing house	farmer video	nobs
Index of farmer's seed ratings[1]†	0.001	0.055	0.065	-0.025	2713
	(0.607)	(0.045)	(0.045)	(0.023)	
Farmer was satisfied with seed quality†	0.727	0.022	-0.003	-0.029	2870
	(0.446)	(0.025)	(0.025)	(0.017)	
Farmer would use seed again†	0.756	0.011	0.015	-0.021	2870
	(0.430)	(0.023)	(0.023)	(0.016)	
Amount of seed farmer used in kg†	6.429	-0.137	-0.138	-0.159	2610
	(4.596)	(0.268)	(0.271)	(0.175)	
Price of seed in UGX/ kg†	2542.690	150.630	171.476	-31.071	2644
	(3220.703)	(170.103)	(168.250)	(120.216)	
Transformed cost of seed in UGX (IHS)	5.048	0.318	0.358	0.099	2564
	(5.351)	(0.259)	(0.252)	(0.207)	
Max. number of observations					3441

Note: 1st column reports sample means and standard deviations below in brackets; 2nd-4th column reports differences between treatment and control groups and standard errors below in brackets; they are clustered at the level of randomization; 5th column reports number of observations; **, * and + denote significance at the 1, 5 and 10 percent levels; † indicates that the variable is included in the overall index; larger indices indicate more desirable outcomes.
[1]The index of farmer's seed ratings contains 6 ratings: general quality, yield, drought tolerance, pest/disease tolerance, time of maturity, germination. This index ranges from -2.26 to 1.76 at endline.

Table 33: Differences between treatment and control groups - Farmer, secondary outcome variables: yield etc. on randomly selected maize field

	mean	dealer training	clearing house	farmer video	nobs
Production in kg	365.672	17.860	42.442*	2.380	2499
	(320.830)	(20.454)	(19.987)	(12.254)	
Yield in kg/acre†	410.930	2.631	39.211*	-11.361	2506
	(291.318)	(20.016)	(19.379)	(11.529)	
Farmer harvested as much maize as expected(†)	0.190	0.018	0.027	0.000	2832
	(0.392)	(0.022)	(0.021)	(0.015)	
Farmer did not harvest expected amount due to own mismanagement(†)	0.103	-0.006	0.004	0.003	2686
	(0.305)	(0.019)	(0.018)	(0.012)	
Transformed amount of maize sold in kg (IHS)†	1.865	-0.054	0.143	-0.124	2734
	(2.861)	(0.179)	(0.174)	(0.107)	
Transformed revenue in UGX (IHS)†	3.821	-0.129	0.319	-0.327	2702
	(5.985)	(0.381)	(0.371)	(0.226)	
Transformed amount kept as seed in kg (IHS)(†)	1.743	-0.018	0.033	0.037	2508
	(1.655)	(0.110)	(0.107)	(0.067)	
Overall index controlling for baseline	-0.004	-0.002	0.108*	-0.066*	2368
	(0.804)	(0.050)	(0.049)	(0.032)	
Overall index not controlling for baseline	-0.003	0.023	0.036	-0.029	2126
	(0.512)	(0.032)	(0.033)	(0.022)	
Max. number of observations					3441

Note: 1st column reports sample means and standard deviations below in brackets; 2nd-4th column reports differences between treatment and control groups and standard errors below in brackets; they are clustered at the level of randomization; 5th column reports number of observations; **, * and + denote significance at the 1, 5 and 10 percent levels; † indicates that the variable is included in the overall index; larger indices indicate more desirable outcomes.

Table 34: Differences between treatment and control groups - Farmer, secondary outcome variables: yield etc. on randomly selected maize field: P-values adjusted according to Sankoh, Huque, Dubey (1997)

	mean	dealer training	clearing house	farmer video	nobs
Production in kg	365.672	17.860	42.442^{+}	2.380	2499
	(320.830)	(20.454)	(19.987)	(12.254)	
Yield in kg/acre†	410.930	2.631	39.211^{+}	-11.361	2506
	(291.318)	(20.016)	(19.379)	(11.529)	
Farmer harvested as much maize as expected$^{(\dagger)}$					
Farmer did not harvest expected amount due to own mismanagement$^{(\dagger)}$					
Transformed amount of maize sold in kg (IHS)†	1.865	-0.054	0.143	-0.124	2734
	(2.861)	(0.179)	(0.174)	(0.107)	
Transformed revenue in UGX (IHS)†	3.821	-0.129	0.319	-0.327	2702
	(5.985)	(0.381)	(0.371)	(0.226)	
Transformed amount kept as seed in kg (IHS)$^{(\dagger)}$					
Max. number of observations					3441

Note: 1st column reports sample means and standard deviations below in brackets; 2nd-4th column reports differences between treatment and control groups and standard errors below in brackets; they are clustered at the level of randomization; 5th column reports number of observations; **, * and + denote significance at the 1, 5 and 10 percent levels; † indicates that the variable is included in the overall index; larger indices indicate more desirable outcomes.

agricultural inputs. It may be preferable to alternative strategies such as regulating quality due to its likely lower cost, self-sustaining nature and scaleability, and helps to overcome problems such as insufficient public investment in regulatory systems, regulatory enforcement, and market surveillance.

References

Akerlof, G. 1970. "The market for 'Lemons': Quality uncertainty and the market mechanism." *Quarterly Journal of Economics* 84 (3): 488–500.

Anderson, J. R. and G. Feder. 2004. "Agricultural Extension: Good Intentions and Hard Realities." *The World Bank Research Observer* 19 (1): 41–60.

Ashour, M., D. O. Gilligan, J. B. Hoel, and N. I. Karachiwalla. 2019. "Do Beliefs About Herbicide Quality Correspond with Actual Quality in Local Markets? Evidence from Uganda." *The Journal of Development Studies* 55 (6): 1285–1306.

Ashraf, N., X. Giné, and D. Karlan. 2009. "Finding missing markets (and a disturbing epilogue): Evidence from an export crop adoption and marketing intervention in Kenya." *American Journal of Agricultural Economics* 91 (4): 973–990.

Barham, B. L., J.-P. Chavas, D. Fitz, V. Ríos-Salas, and L. Schechter. 2015. "Risk, learning, and technology adoption." *Agricultural Economics* 46 (1): 11–24.

Barriga, A. and N. Fiala. 2020. "The supply chain for seed in Uganda: Where does it go wrong?" *World Development* 130: 104928.

Bernard, T., S. Dercon, K. Orkin, and A. Seyoum Taffesse. 2015. "Will Video Kill the Radio Star? Assessing the Potential of Targeted Exposure to Role Models through Video." *The World Bank Economic Review* 29 (suppl_1): S226–S237.

Bold, T., K. C. Kaizzi, J. Svensson, and D. Yanagizawa-Drott. 2017. "Lemon technologies and adoption: measurement, theory and evidence from agricultural markets in Uganda." *The Quarterly Journal of Economics* 132 (3): 1055–1100.

Camerer, C. and T. Hua Ho. 1999. "Experience-weighted attraction learning in normal form games." *Econometrica* 67 (4): 827–874.

Conley, T. G. and C. R. Udry. 2010. "Learning about a New Technology: Pineapple in Ghana." *American Economic Review* 100 (1): 35–69.

Duflo, E. and A. Banerjee. 2011. *Poor economics.* PublicAffairs.

Duflo, E., M. Kremer, and J. Robinson. 2011. "Nudging farmers to use fertilizer: Theory and experimental evidence from Kenya." *American economic review* 101 (6): 2350–90.

Fafchamps, M. and B. Minten. 2012. "Impact of SMS-based agricultural information on Indian farmers." *The World Bank Economic Review* 26 (3): 383–414.

Foster, A. D. and M. R. Rosenzweig. 1995. "Learning by doing and learning from others: Human capital and technical change in agriculture." *Journal of political Economy* 103 (6): 1176–1209.

Gars, J. and P. S. Ward. 2019. "Can differences in individual learning explain patterns of technology adoption? Evidence on heterogeneous learning patterns and hybrid rice adoption in Bihar, India." *World development* 115: 178–189.

Gharib, M. H., L. H. Palm-Forster, T. J. Lybbert, and K. D. Messer. 2021. "Fear of fraud and willingness to pay for hybrid maize seed in Kenya." *Food Policy* 102: 102040.

Gollin, D., M. Morris, and D. Byerlee. 2005. "Technology adoption in intensive post-green revolution systems." *American Journal of Agricultural Economics* 87 (5): 1310–1316.

Goyal, A. 2010. "Information, Direct Access to Farmers, and Rural Market Performance in Central India." *American Economic Journal: Applied Economics* 2 (3): 22–45.

Hanna, R., S. Mullainathan, and J. Schwartzstein. 2014. "Learning through noticing: Theory and evidence from a field experiment." *The Quarterly Journal of Economics* 129 (3): 1311–1353.

Hasanain, A., M. Y. Khan, and A. Rezaee. 2019. "No bulls: Experimental evidence on the impact of veterinarian ratings in Pakistan." .

Karlan, D., R. Osei, I. Osei-Akoto, and C. Udry. 2014. "Agricultural decisions after relaxing credit and risk constraints." *The Quarterly Journal of Economics* 129 (2): 597–652.

Magruder, J. R. 2018. "An Assessment of Experimental Evidence on Agricultural Technology Adoption in Developing Countries." *Annual Review of Resource Economics* 10 (1): 299–316.

Michelson, H., A. Fairbairn, B. Ellison, A. Maertens, and V. Manyong. 2021. "Misperceived quality: Fertilizer in Tanzania." *Journal of Development Economics* 148: 102579.

Reimers, I. and J. Waldfogel. 2021. "Digitization and pre-purchase information: the causal and welfare impacts of reviews and crowd ratings." *American Economic Review* 111 (6): 1944–71.

Riley, E. 2019. "Role models in movies: the impact of Queen of Katwe on students' educational attainment." *The Review of Economics and Statistics* 1–48.

Sankoh, A. J., M. F. Huque, and S. D. Dubey. 1997. "Some comments on frequently used multiple endpoint adjustment methods in clinical trials." *Statistics in medicine* 16 (22): 2529–2542.

Figure 2: SeedAdvisor certificate

Suri, T. 2011. “Selection and comparative advantage in technology adoption.” *Econometrica* 79 (1): 159–209.

Tripp, R. and D. Rohrbach. 2001. “Policies for African seed enterprise development.” *Food Policy* 26 (2): 147–161.

Van Campenhout, B. 2021. “The role of information in agricultural technology adoption: Experimental evidence from rice farmers in Uganda.” *Economic Development and Cultural Change* 69 (3): 1239–1272.

Van Campenhout, B., D. J. Spielman, and E. Lecoutere. 2020. “Information and Communication Technologies to Provide Agricultural Advice to Smallholder Farmers: Experimental Evidence from Uganda.” *American Journal of Agricultural Economics* n/a (n/a).

Vandevelde, S., B. Van Campenhout, and W. Walukano. 2018. *Spoiler alert! Spillovers in the context of a video intervention to maintain seed quality among Ugandan potato farmers*. Tech. rep., LICOS Discussion Paper.

Wossen, T., K. A. Abay, and T. Abdoulaye. 2022. “Misperceiving and misreporting input quality: Implications for input use and productivity.” *Journal of Development Economics* 157: 102869.