Can informed buyers improve goods quality? Experimental evidence from crop seeds

Eric Hsu[a] Anne Wambugu[b]

This version: January 20, 2022
Latest version available **here**.

Abstract

We study whether adding informed buyers to a market can improve the quality of goods supplied by sellers, in an environment where goods quality is difficult to observe. To do so, we implement a market-level intervention, randomizing rural markets in Kenya into a community-wide information campaign. Small-scale maize (corn) farmers in treated market areas were trained to identify hybrid maize seed that is quality-verified under national seed regulations. We do so in a region where there are widespread concerns about deceptive counterfeits and other uncertified seeds of lower quality. In a first main result, we find that observable markers predict seed quality, and treatment increased knowledge of these markers and affected seed purchase decisions. We show that substantial gains in yields were experienced among subgroups of informed buyers that had larger gaps between baseline seed quality and national standards. We do not see effects on seed quality offered to uninformed buyers in the same communities, as revealed by data from secret shoppers. These patterns can be explained by a model in which sellers' ability to discriminate allows lower-quality products to continue to be offered to uninformed buyers, even while informed buyers experience benefits and their increasing numbers hurt firm profits. Consistent with the model predictions, we find that treatment caused some seed sellers to exit the market, and we do not detect changes in seed prices. Taken together, the findings document new stylized facts and provide evidence relevant for boosting yields of a staple crop crucial for food security. More generally, they provide lessons concerning the role of improved consumer information in disciplining firms in low information environments.

[a] Department of Economics, University of California, Berkeley [b] Strathmore Energy Research Centre, Strathmore University. We thank Susanna Berkouwer, Stephanie Bonds, Ben Faber, Frederico Finan, Benjamin Handel, Supreet Kaur, Kei Kawai, Samuel Leone, Jeremy Magruder, Aprajit Mahajan, Edward Miguel, Maximillian Mueller, Duncan Mwangi, Andrés Rodriguez-Clare, Joseph Wairiuko, Guo Xu, and participants at UC Berkeley Development Lunch, UC Berkeley Development Lunch, and IZA/FCDO G2LM-LIC Research Conference for feedback and helpful conversations. We thank Esther Kimani and Simon Maina at the Kenya Health Inspectorate Service, as well as Bright Simons, Carolyne Kimuhu, Timothy Maina, and Charles Ochieng from mPedigree for helpful conversations and their support for the project. We thank Noah Forougi and Cashua Wang for excellent research assistance. We are also grateful to REMIT Kenya staff – and particularly Carol Nekesa, Nettah Isavwa, Erick Emilio, Stephen Omare, and Christine Soita for their work on data collection and other field work for this project. The Agricultural Technology Adoption Initiative, IZA, FCDO, the Center for Effective Global Action, and the Private Enterprise for Development in Lower-Income Counties initiative provided financial support for the project, and we acknowledge that the views expressed are ours and not necessarily those of the above mentioned organizations and initiatives. Data collection was carried out under UCB CPHS protocol 2019-10-12666.

1 Introduction

When buyers cannot perfectly observe product quality, a market breakdown can occur in which offering high-quality products may not be profitable for sellers (Akerlof, 1970). A variety of mechanisms may help counteract such a market failure by providing a credible signal of product quality to consumers (Dranove and Jin, 2010). For example, firms may use branding or warranties to show they are incentivized to provide high quality. Government agencies can set quality standards and provide certifications or licenses to assure consumers of a certain level of quality in a market.

However, mechanisms to counteract the effects of limited consumer information may function imperfectly or be non-existent in many settings. Markets in low and middle income countries (LMICs), for instance, often have features that make them particularly vulnerable to problems with imperfect information. Lower access to information technology or lower levels of education may slow the diffusion of consumer information and hinder a firm's efforts to build a reputation for providing high quality (Aker and Blumenstock, 2015). Policy solutions may be undermined by lower state capacity (Dal Bó and Finan, 2020). This could lead to less effective enforcement of quality certifications. Branding and reputation building mechanisms may also function less well if entry of deceptive counterfeits cannot be deterred through legal institutions. These factors may help explain LMIC settings where reliably obtaining high quality goods can be challenging.

In this paper, we study the effects of improving consumers' ability to detect quality in the hybrid maize (corn) seed market in rural Kenya. Several feature of this setting may hinder the ability of firms to signal that their product is high quality (Bold et al., 2017; Bai, 2021). By nature, the quality of hybrid maize seeds is difficult to observe prior to purchasing and planting them. Even after planting, the performance of maize seeds depends on numerous factors, each of which can have a sizable impact on plant growth and quantities of harvested maize. As a result, maize yields tend to be noisy and give buyers a poor signal of seed quality. Typical crop planting timelines limit farmers to only observing one or two harvests per year. Therefore, infrequency of purchase may also limit consumer learning about seed quality. While regulations exist to establish minimum quality standards for seeds, enforcement is limited and there have been widespread concerns about the prevalence of counterfeit and other uncertified seeds in recent years (Okinda, 2019).

Resolving information problems and improving market-supplied quality is likely to be particularly consequential in this setting. High quality maize seeds are a productive input that is critical for maintaining food security in low-income areas of the world. Maize is a staple food in our study area in Western Kenya but also more broadly in Southern, Central, and Eastern Africa and in Central and South America. In these regions, over 20% total calories are supplied by maize (Shiferaw et al., 2011). Our study area is in the East Africa region, where food security is a widespread and chronic issue and a major focus for policymakers. Over 30% of the population in East Africa experience severe food insecurity, lacking adequate access to enough safe and nutritious food for healthy development (FAO, 2020, 2021). This contributes to malnutrition as a major public health problem. For example, in 2019 an estimated 34.5% of under-age-five children in the region were stunted and 5.3% were moderately or severely wasted, with height-for-age or weight-for-age at least two standard deviations below the median in World Health Organization standards (FAO, 2020).

To evaluate the effects of improved consumer information, we conduct a field experiment in a sample of 386 rural market areas in Kenya. In randomly selected markets, we provided community-wide trainings to enable farmers to better detect quality-verified hybrid maize seeds. This market-level intervention allows us to identify effects of improved information on local market participants, including consumers who became informed due to treatment, uninformed consumers in the same communities, and seed sellers. We examine whether and how informed buyers are able to achieve improved agricultural outcomes. As well we examine whether uninformed buyers in treated communities may also experience benefits through improved seed quality. Using a series of surveys and secret shopper activities, we track outcomes for buyers and sellers (including knowledge, seed sales and purchases, and agricultural outcomes) for over one year after treatment. The inclusion of 104 pure control sites allow us to test whether baseline activities may have influenced outcomes. The extended follow-up period of over one year allows us to track possible convergence to a new equilibrium over the course of three planting seasons following treatment.

The findings first show that observable quality markers correlate with lab tested quality, with packets missing one or more quality marker having lower germination rates.[1] This confirms that agricultural gains should in theory be possible if farmers follow a simple strategy to avoid buying seeds that are not quality-verified. Yet, farmer knowledge about observable quality markers is low at baseline.

The information treatment caused buyers to have better ability to detect seed quality. Treated buyers had greater knowledge of observable quality markers and were more likely to successfully use detection techniques in subsequent planting seasons. Treated buyers frequently reported that the information that was provided affected their purchasing decisions—both which seed packets to buy and which seed seller to buy from. Consistent with these accounts, we confirm that treated buyers were more likely to leave the local market to purchase seeds. Examining agricultural outcomes, we find that treated farmers experienced about 3% higher maize yields overall, though not statistically significant. Farmers in more remote market areas, which had lower initial seed quality, experienced a statistically significant 7.7% increase in yields, with female-headed households especially benefiting. Meanwhile, we do not detect effects on quality offered to uninformed buyers in the same communities, as revealed by data from secret shoppers.

These patterns can be explained by a model in which farmers who are informed are able to conduct a search for high quality seeds–possibly leaving the local market to do so–and their increasing numbers hurt the profits of local firms. Yet, sellers' ability to discriminate between informed and uninformed buyers can allow lower-quality products to continue to be sold to uninformed buyers. Consistent with the model's predictions, we also find that treatment caused some seed sellers to exit the market, and we do not detect effects on prices.

Taken together, the results document new stylized facts and provide important evidence relevant for boosting yields of a staple crop. There is much work on barriers to adopting and interventions

[1] Among packets sampled by secret shoppers, 38% of them lack one or more quality marker.

to encourage adoption of agricultural technologies that can enhance productivity. Jack (2013) surveys this literature, which spans studies on many types of constraints—including credit and liquidity constraints, risk aversion, lack of knowledge on costs and benefits, and heterogeneous costs, among other explanations for under-adoption—and has motivated policy responses to address market failures in the adoption of agricultural inputs (Holden, 2019). Some recent work has explored the role of low input quality as a barrier to improving agricultural productivity. Quantitative research (Bold et al., 2017; Ashour et al., 2019) and news reports (Muchiri, 2019; Okinda, 2019) suggest that the quality of agricultural inputs—such as seeds, fertilizer, and pesticides—is often low in rural African markets. For example, Bold et al. (2017) study retail quality maize seeds and fertilizer from rural Ugandan markets. They find that switching to wholesale quality of the same products causes a 40% increase in maize yields. These reports suggest that low and high-quality products often appear similar along observables (e.g. price and package characteristics), which is consistent with accounts of widespread counterfeiting (Kenya Association of Manufacturers, 2012). Our paper contributes by evaluating an intervention to improve consumer information as a way to boost usage of high quality inputs and improve agricultural productivity. The findings are directly relevant to policymakers working to promote food security. This paper also contributes important descriptive data on the quality of agricultural inputs (Bold et al., 2017; Gharib et al., 2021; Kenya Association of Manufacturers, 2012; Ashour et al., 2019; Michelson et al., 2021)), particularly on the relationship between observables and seed quality.

We also contribute to work related to consumer mistakes, consumer learning, and quality provision. There is a deep theoretical literature on market dynamics when agents have imperfect information (e.g. Akerlof, 1970; Wolinsky, 1983; Shapiro, 1983). However, relatively few empirical studies have tried to test theoretical predictions in a real-world setting (Bai, 2021; Jin and Leslie, 2009; Bjorkman-Nyqvist et al., 2012). We do so in a high-stakes and policy-relevant context, using an experimental evaluation closely tied to a common government policy (certification of minimum quality standards) intended to overcome a market failure arising from a difficult learning environment for consumers. This paper also relates to work on consumer misinformation and costly consumer mistakes (Bronnenberg et al., 2015; Grubb and Osborne, 2015; Handel and Schwartzstein, 2018). We document that many consumers overlook easy-to-use strategies to obtain higher quality products, behavior that is inconsistent with a full-information model with rational consumers.

Lastly, we contribute to work on regulation, monitoring, and enforcement. Several papers examine how consumer information or monitoring by consumers can complement or replace direct enforcement efforts. Different from previous work (Gonzalez-Lira and Mobarak, 2021; Annan, 2021; Naritomi, 2019), we find evidence in our setting that consumers may not be very effective at disciplining sellers, and we discuss the conditions from this settings that likely contributes to this finding.

2 Background

Maize (often called corn in North America) is a staple crop in many areas of the world. In Eastern and Southern Africa, maize is estimated to account for 22% of all calories consumed and is an important crop for subsistence farmers in the region (Shiferaw et al., 2011). A central challenge for policymakers concerned about food security is how to boost agricultural productivity in lower-income countries, where maize yields are a small fraction of the average in OECD countries (OECD/FAO, 2021). Increasing access to improved agricultural inputs such as high yielding seed varieties and fertilizers have been one focus among policymakers in recent decade (Evenson and Gollin, 2003).

However, recently scholarly work and local media coverage has suggested that retail maize seeds in East Africa and other agricultural inputs are variable and often sub-standard (Bold et al., 2017; Ashour et al., 2019; Muchiri, 2019; Okinda, 2019). In Kenya, the regulation of agricultural inputs falls under the responsibility of the Kenya Plant Health Inspectorate Service (KEPHIS). KEPHIS was established in 2012 by the Seed and Plant Varieties Act, and as part of its responsibilities, it tests and certifies crop seeds for sale. All seeds must be certified by KEPHIS before sale, and all sellers must register with KEPHIS as well. Importantly, seeds must test with over 90% germination rate to be certified, meaning that under ideal temperature and moisture conditions randomly sampled seeds must develop normally into a plant. Concerns about counterfeit and other uncertified seeds led to a new initiative in 2018 to mandate e-verification for certified seeds. This requirement for all seeds sold in Kenya was layered on top of pre-existing requirements, such as a printed lot number and packaging or expiration date listed for each packet. The new e-verification allowed for an additional method to obtain and verify printed information that was visible via SMS and could not be physically tampered with. Each packet is assigned a unique secret code that is revealed upon scratching off the sticker, and the code is available for us one time only – subsequent uses would indicate that the code is valid but would alert the user that it had been used before.[2]

Nevertheless, concerns about seed quality remain. As KEPHIS Managing Director Esther Kimani said, tying this issue to food security: "These fake seed sellers have...been the cause of food shortages that make Kenya spend billions of shillings on imports annually." Counterfeits may not be the only possible cause of low quality seeds – other reasons may include poor storage or selling seeds after expiration or too long after testing to ensure high quality. This is consistent with academic research on maize seed performance after storage under less-than-ideal conditions (Ghassemi-Golezani and Mamnabi, 2019). Although seed testing for each seed lot that is manufactured is subject to testing requirements at KEPHIS, there are several points within the supply chain between the seed manufacturer and the small retail outlets that we study, which allows for possible lower-quality seed packets to enter the supply chain. First, the seed manufacturer relies on a network of agents and sub-agents to distribute the seeds, ending with the smallest retail shops and the farmers they sell to. Intermediaries and retail shops may store seed packets improperly, sell expired seed, or source seeds

[2]We thank staff at KEPHIS and mPedigree for helpful conversations about recent changes in seed regulations.

from unauthorized distributors, allowing for counterfeits or other non-quality-verified products to enter the supply chain.

Our study area is in Western Kenya in four counties - Bungoma, Busia, Kakamega and Transnzoia Counties. According to the 2014 Kenya DHS survey, 47% of households in the region experience food insecurity. In 2019, 74.5% of households in this area participated in farming activities, 92.4% of which farmed maize (Kenya National Bureau of Statistics, 2019).

3 A Simple Model With Informed and Uninformed Consumers

To structure our thinking about consumers and seller behavior, we consider a simple one-period model. This is model sellers set price and quality levels in the presence of both informed and uninformed consumers who also have the option of purchasing from a seller outside the local market.

3.1 Demand side

Assume there are N consumers, all of whom seek exactly 1 unit of the good. The good comes in one of two quality levels: high or low. A consumer derives utility $1 - p$ if they obtain a high quality good, and $-p$ if they obtain low quality. Let θ be the proportion of consumers who are informed. We treat θ as exogenous; one could think of the share of informed consumers as coming from a person-specific cost of acquiring information, which leads some to acquire this information and others to not acquire it.

Assume there is either 0 or 1 firm locally.[3] Consumers can buy from the local firm (if there is one), or alternatively, they can buy from a firm outside the local market, paying search cost S and price P_0. As assume that the local market is small enough that P_0 would be negligibly affected by choices of local market participants. An informed consumer, upon taking the outside option, will always get high quality. An uninformed consumer believes they can get high quality with probability $\bar{q}$ in the outside option.

Informed consumers can observe quality markers and upon visiting the local shop will know the quality of the product (either high or low) with certainty. Uninformed consumers, on the other hand, cannot observe quality. They have beliefs about the average quality in the local market ($\hat{q}$). Note that we allow $\hat{q}$ to match or not match the average quality offered by the seller.[4]

We focus on the equilibrium in which the uninformed buy locally, and the informed buy locally *only* if they receive a high quality good. Since the informed expect to get high quality from the outside option at cost $P_0 + S$, and the uninformed expect to get quality $\bar{q}$ in expectation at cost

[3]Here we assume only up to one seller, which is clearly a simplification. However, the median market in our sample has only one maize seed seller. Evidence from some recent research has also suggested that sellers in similar rural settings in Kenya may collude and behave similarly to a local monopolist (Bergquist and Dinerstein, 2020)

[4]Bold et al. (2017), Ashour et al. (2019), and Michelson et al. (2021) suggest that buyers of agricultural inputs do not hold accurate beliefs about product quality and that learning about quality can be very difficult. Moreover, as suggested in (Bold et al., 2017), learning by Bayesian farmers may especially be hindered when the difference in quality between low and high quality is not very large. Thus we think of this as a model of short-run buyer and seller responses when beliefs about average quality do not change.

$P_0 + S$, so the following must be true in this equilibrium:

$$1 - p \geq 1 - P_0 - S \implies p \leq P_0 + S$$
$$\hat{q} - p \geq \bar{q} - P_0 - S \implies p \leq \hat{q} - \bar{q} + P_0 + S$$

We might default to thinking of the second expression as binding, in which buyers believe they get better expected quality outside the local market. Evidence from Gharib et al. (2021) would support this assumption, finding that farmers are willing to pay a premium for a seed packet directly from the seed company, which normally can be obtained outside the local market in town). Either way, in this equilibrium, the seller can raise prices up to $p^* = \min\{P_0 + S, \hat{q} - \bar{q} + P_0 + S\}$ without losing any of the uninformed consumers, or any informed consumers who get high quality. They lose informed customers who are offered low quality, who opt for the outside option. In Appendix B we discuss the conditions needed to ensure this equilibrium, which we think is the most interesting case.

3.2 Supply side

Let x be the number of units that the firm sells. Let $q \in [0, 1]$ be the quality mix chosen by the seller. That is, proportion q of the seller's stock is high quality, while proportion $1-q$ is low quality. For each potential customer, the seller randomizes the quality of the product according to the choice of q, and they make a take-it-or-leave-it offer.

The seller faces variable costs $(c + d\frac{q}{1-\theta+q\theta})x$, where c is the (assumed constant) per-unit cost for a unit of low-quality good, and d is the cost of upgrading a unit's quality from low quality to high quality. $\frac{q}{1-\theta+q\theta}$ is the average quality among units that are sold if the seller chooses quality mix q, but informed buyers refuse if offered low quality (which happens $q\theta$ of the time).

In addition, assume the firm faces possible punishment from regulators. For simplicity, our baseline model sets expected punishment as: $b\theta(1-q)^2N$.[5] One can think of this functional form as being the product of two terms – the first being the probability of detection, and the severity of punishment conditional on detection – where the probability of being detected by authorities is proportional to the number of unhappy informed customers, and the severity of punishment conditional on detection is proportional to how much cheating the firm engages in. Finally let F be the firm's fixed cost.

Thus, the firm's problem is as follows, where we write x as a function of the firm's choice of price p and quality mix q:

$$Max_{p,q}(p - c - d\frac{q}{1-\theta+q\theta}) * x(p,q) - \frac{1}{2}b\theta(1-q)^2N - F$$

[5] In practice, our main takeaways from the model will not be very dependent on the exact functional form.

As shown in Appendix B, when an interior solution exists, optimal quality is:

$$q^* = \frac{(p-c)\theta - d + b\theta}{b\theta} = \frac{(p-c)\theta - d}{b\theta} + 1$$

Quality is increasing in enforcement strength b and decreasing in upgrade cost d. Importantly, quality is increasing in the proportion of buyers who are informed θ.

We examine the effect of more informed buyers on firm profits. From the firm's objective function, by the envelope theorem we get:

$$\frac{\partial \pi^*}{\partial \theta} = -(p-c)(1-q^*)N - \frac{1}{2}b(1-q^*)^2 N \leq 0$$

So having more informed buyers is bad for profits and may cause sellers to exit the market.

3.3 Summary of model predictions

1. Buyers that get informed obtain higher quality goods. Being able to observe quality, they are able to search until they find high quality, leaving the local market if necessary to find another seller.
2. Having more informed buyers does not affect market prices ($\frac{\partial p}{\partial \theta} = 0$). Sellers have local market power, and prices are bounded from above by uninformed consumers' value of buying from the outside option.
3. Having more informed consumers has a negative effect on local sellers' expected profit ($\frac{\partial \pi}{\partial \theta} \leq 0$). More informed consumers reduce the local seller's market share and may also increase the expected punishment from regulators.
4. Uninformed consumers may also benefit from having more informed consumers in the local market. In the interior solution they benefit, but less than informed consumers do ($\frac{\partial q}{\partial \theta} \geq 0$). They do not benefit in an edge case. This is because uninformed consumers only benefit to the extent that more informed consumers affect the local seller's choice of quality mix.

4 Study Design

4.1 Sample

We randomly sampled rural markets that satisfy the following conditions: (1) it must have fewer than 100 shops, (2) it must be more than 2km from a market that has more than 100 shops, (3) it must have at baseline at least one seed seller. We consider markets that satisfy these conditions eligible to be sampled. We sample markets from four counties: Bungoma, Busia, Kakamega, and Transnzoia using a 2-stage sampling strategy–first randomly selecting sublocations (a list of which we obtained from the county commissioner offices), next tabulating all eligible markets in selected sublocations, and lastly randomly selecting one market (or two in randomly selected sublocations in Transnzoia, due to logistical reasons). This sampling strategy helps minimize clusters of markets that are very close to one another (possibly risking spillovers between treatment and control sites), and it helps the sample of markets cover both less remote and more remote areas of each county. The main study sample in treatment and control groups consists of 302 markets in 282 sublocations.

At each market, all seed sellers and eight randomly selected households within 1km from the market center were sampled to be surveyed. The sampled farmers are overwhelmingly small-holder farmers, with the median farmer planting on 1.5 acres.

In addition to markets in the treatment and control groups, 104 pure control markets were also selected using the same methodology, except that we do not observe the number of seed sellers in Feb-Mar 2020. We instead use the number of seed sellers in March 2021. For these sites we collected only data from market audit activities during the 2021 main planting season. Due to funding limitations, we did not conduct seller surveys or household surveys in pure control markets.

4.2 Randomized experiment

Sublocations were randomized into the following treatment groups, and we describe the treatment arms in greater detail below.

Treatment #1	: 68 sublocations
Treatment #1 & Treatment #2	: 68 sublocations
Control	: 146 sublocations
Pure Control	: 104 sublocations

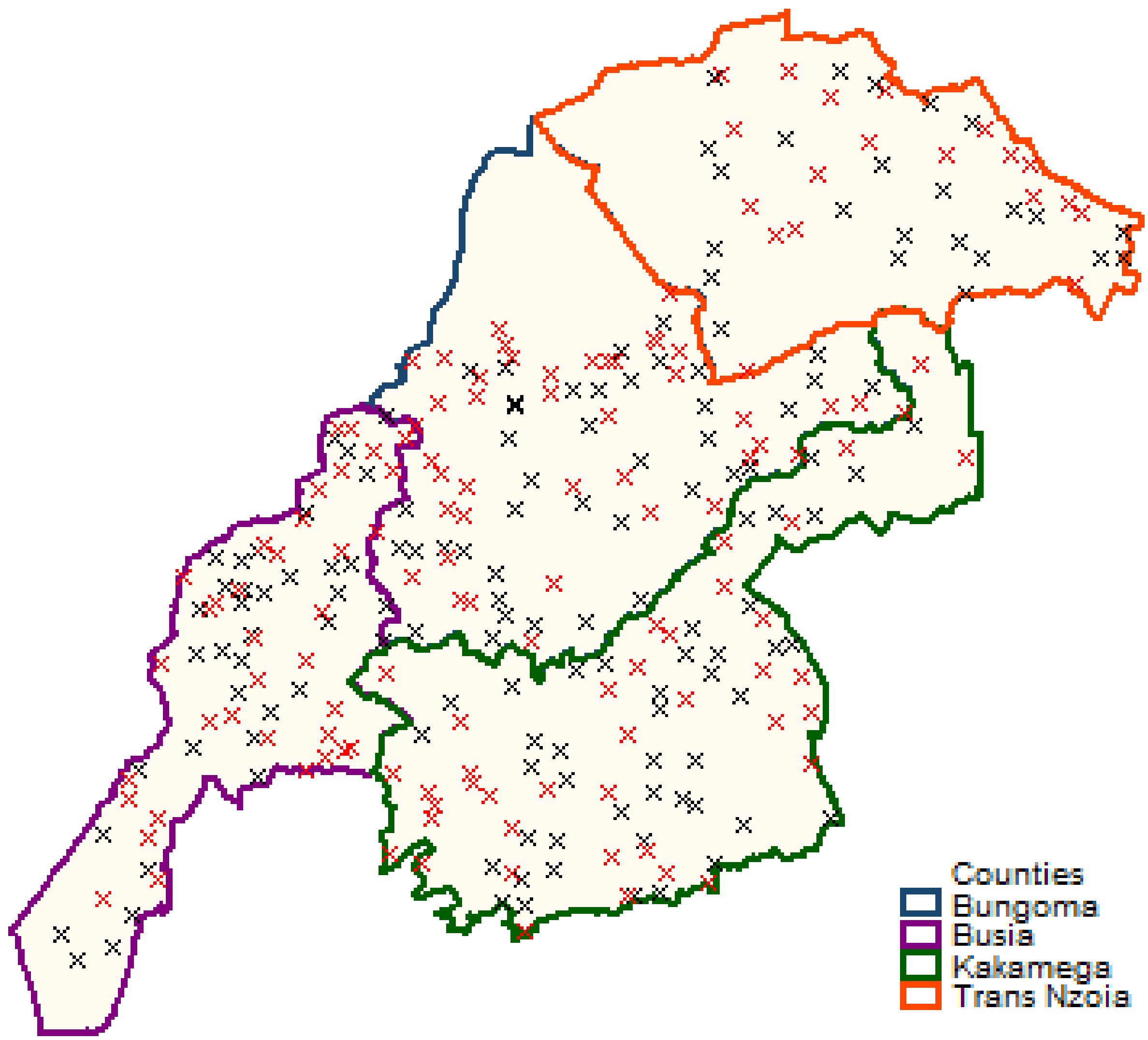
Figure 1 shows the sampled markets within the four counties, with treated markets in red.

Treatment #1 provides information to help consumers distinguish between low and high-quality products. Through pilot activities and discussion with project partners, we identified a set of quality markers that are required for all certified seeds in Kenya and have variation in rural markets. These include the lot number and a sticker containing an SMS code that allows farmers to verify authenticity (see Appendix A). In response to widespread reports of deceptive counterfeiting, starting in 2018 the Kenya Plant Health Inspectorate Service (KEPHIS) required all certified packets of seed to carry one of these scratch-off codes. Using SMS verification allows farmers to detect two types of lower-quality seeds: uncertified seeds and old seeds that were unsold from previous seasons.

Treatment #2 encourages consumers to report sub-standard seeds when they encounter them. Consumers were (1) encouraged to ask for a receipt and keep the packaging for documentation, (2) told how to report the incident anonymously to the Anti-Counterfeit Authority, a corporate representative, an agricultural officer, KEPHIS, or the local chief (who were informed by our team how to escalate reports), and (3) encouraged to discuss seed quality among friends, family, and neighbors. They also were told of documented instances where complaints led a company to give compensation or led to legal action.

Both treatments were administered in the month immediately before the main planting season in 2020, which occurs around the start of rains in March. Our team carried out the campaign by working with the local assistant chief to deliver flyers to and speak to locally influential residents at gatherings, including village elder meetings, local farmer group meetings, and barazas and chamas (local community meetings). We also spoke with individual farming households, who were recruited as survey respondents, and after being surveyed were given the information treatment. We sent a team member to deliver information at each site, taking into account gender dynamics in the local context.

Figure 1: Study Area



The study area in Western Kenya includes Bungoma, Busia, Kakamega, and Transnzoia Counties. Red X's represent treatment sites, where community-wide training to help consumers detect quality-verified seed was carried out. Black X's represent control sites. In total 302 market areas are included in the sample.

All treated sites receive Treatment #1, which was designed to increase the probability of detecting non-quality-verified seeds, corresponding to an increase of θ in the model. As illustrated in the model, this can directly affect purchasing decisions (e.g. refusing a packet or switching sellers) and lead to the adoption of higher quality seeds. As well, treatment may affect sellers' decisions (e.g. selling more high-quality seeds or adjusting the price) and lead to a greater prevalence of high-quality seed in the market for even uninformed buyers. This may happen through economic channels as buyers refuse to buy when observing a poor signal (as in Wolinsky 1983). We look to predictions #1 through 4 to guide our thinking on the expected effects the information treatment. A secondary focus of our project is certain channels such as social pressure or threat of legal action that may also reduce the sellers' payoff to cheating and additionally affect the seller's decisions about firm entry and quality. This is explored by the addition of Treatment #2 which is layered on top of Treatment # 1 for about half the treated markets and which corresponds to increasing b in the model. While this does not affect model predictions #1 and #2, it could affect the seller's entry/exit decision by affected expected profits, and it could affect the seller's choice of quality.

Appendix A shows the flyers which summarize the information that was communicated at treated market areas as well as typical packets of seeds and screenshots from using the e-verification system. (Those who received flyers were encouraged to pass extras along to family, neighbors, and friends.) Focus group discussions prior to project launch as well as treatment meetings suggest that farmers perceived the provided information to be important for their livelihoods and that the information is accessible, particularly with the help of family, friends, and neighbors. We note that one of the quality markers–the KEPHIS sticker–is more technically challenging to use properly. While the application of SMS codes to quality verification may be new to many farmers, we note that the technology is often somewhat familiar from its widespread use for pre-paid mobile phone service, which may explain the ready acceptance of this aspect of the treatment.

We also varied treatment intensity such that two-thirds of treated markets received a more intensive multi-day treatment (two or three days of information dissemination); this allows us to explore differential effects from greater information saturation. We implemented this design out of concern that treatment effects may only be seen for a large enough change in θ. For example, consider the model from section 3 in which some markets begin in an edge case with $q = 0$. In such cases, only a large-enough change in θ that causes a threshold to be crossed will induce a change in the quality provided by sellers. In practice, though, the data shows little evidence of such threshold effects from the variation in our sample in treatment intensity, and section 5 below primarily focuses on simple treatment-control comparisons in order to maximize statistical power.

Treated market areas received on average 173.5 flyers, which we estimate as being approximately 19.7% of local seed customers (see Table D1).[6] On average in treated markets the team trained 81.7 adults face-to-face, or about 9.1% of local seed customers. Considering village elders' estimates for the local population, about 49% of the local (within 1km) adult population received a flyer, and 22% were treated directly face-to-face.

4.3 Data Collection and Timeline

Project activities proceeded as illustrated in Figure 2.

In February 2020 into early March 2020 baseline surveying and treatment were carried out. After a site was initially sampled and our team made contact with village elders, baseline household and

[6]We use several pieces of data to estimate this figure, including: number of sellers in a market, number of customers per seller, and percentage of local customers who shop locally for seeds.

sellers surveys were carried out. Treatment meetings normally were later in the day during the first day of surveying, with participants convening in the mid-afternoon in a public location. Market areas that are assigned to receive additional visits for information dissemination would then have a team member return for additional activities on subsequent days.

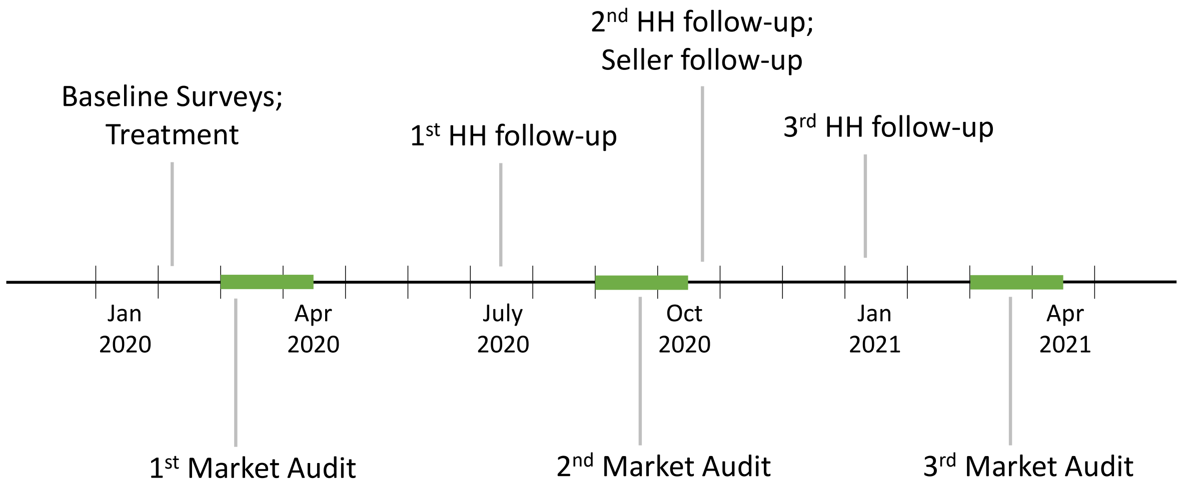
Over the following one year we conducted 3 rounds of household follow-up surveys via phone to document household seed choices and agricultural outcomes. Households self-reported harvested amounts for each maize variety they planted, which we use to compute maize yields for each household. We chose survey times strategically to be spread out around key dates in the usual crop timeline in order to minimize issues with imperfect recall of key variables (e.g. seed purchases and harvested amounts) for both the main season (with planting starting around March 2020) and the short season (with planting starting around September 2020).

To measure seed quality and possible adjustments by sellers, we conducted three rounds of market audit activities. In March 2020, August to September 2020, and March to April 2021, covert shoppers (enumerators on our research team) posed as farmers to visit markets and view up to two packets and purchase up to one packet. They used realistic scripts to purchase "seeds that are popular" or "seeds that are cheap", without naming the exact brand or variety that is desired, leaving the seller to decide which seeds to offer. In the local context, requesting seeds in this way without specifying the exact variety is perceived as normal. When outside of the market area, the secret shopper was instructed to document visible quality markers as well as data about the sellers in the market and the shopping experience. Conversations with field staff lead us to expect that secret shoppers were not recognized or treated differently from other shoppers during the relatively quick transaction.

All seed packets were repackaged into plain paper bags, labeled, and sent to KEPHIS facilities in Nakuru for purity and germination testing, to provide objective measures of two aspects of seed quality. Purity tests measure the percentage of material (by weight) in the packet that are whole seeds. Germination tests measure the percentage of whole seeds that become normally emerging plants under ideal temperature and moisture conditions. While these tests provide an objective test of seed quality, with KEPHIS staff blinded to most seed that are not visible on the seed itself, it does not capture all relevant aspects of seed quality that a farmer may care about. Notably, the lab tests we conducted does not provide information about seed performance when conditions are less-than-ideal. As well, it does not capture aspects of seed quality that affect yields beyond the germination rate. As demonstrated by Ghassemi-Golezani & Mamnabi (2019), it is expected that older or poorly-stored maize seeds will have lower yields due to deterioration of seed quality beyond the effects on germination rate alone.

The team documented the following markers during the market audit: (1) presence of valid SMS verification code, (2) presence of lot number, (3) testing date within 1 year of purchase, (4) damaged packaging (which reports have linked to tampering). A packet can lack a valid SMS verification code several ways. First, a KEPHIS sticker could be absent. Second, the code could be invalid. Third, the code could be valid but already used prior to purchase. Lastly, the code

Figure 2: Timeline



The project activities included: (1) baseline surveys and information treatment conducted starting in February 2020; (2) three market audit activities conducted starting in March 2020, September 2020, and March 2021; (3) a seller follow-up survey to capture prices, quantities, and profits of sellers; (4) 3 rounds of household follow-up surveys, designed to capture key information about seed purchases and agricultural outcomes in both the main growing season and the short growing season.

could be successfully applied for the first time, but the registered information may not match the information on the packet. Similarly, a packet could lack a valid expiration date either because it lacks any date whatsoever, or because it has a date but the seeds are expired.

During follow-up surveys, enumerators were not made aware of the treatment status for the sites and respondents that were being followed up on. Lab testers were not made aware of the seed sources, and all seed samples were repackaged into plain paper bags labeled with unique alpha-numerical identifiers used internally by the research team.

One concern with the survey data may be that measurements on agricultural outcomes are self-reported and could suffer from inaccurate reporting. To guard against these concerns, we used self-reported yields to inform our sample size, anticipating possible noise from mismeasurement.[7] Acreage, kilograms of maize harvested, and kilograms of seed purchased are also checked through repetitive questions in different parts of the survey and cross-validated with each other. For example, the total amount of land dedicated to maize should be consistent with the sum of land dedicated to each variety of maize. The surveys also phrase questions and enumerators use probing in ways that help respondents recall the answers we are seeking. For instance, we ask "how many 90kg bags of maize did you harvest during the long growing season?" as a question respondents often are able to readily recall.

4.4 Baseline Balance, Spillovers, and Other Threats To Identification

Table D2 shows baseline balance for market areas in our sample. Households in treated sublocations had households heads that were slightly less likely to have completed primary or secondary school, and they had slightly smaller plots. However, when we jointly test for differences between treatment and control on all of these outcomes, we cannot reject the null hypothesis that all characteristics are the same in both treatment and control groups (p=0.22). All results below showing treatment effects will control for these baseline characteristics.

In the Appendix C, we explore the possibility that the experimental design may have been contaminated in one of two ways. First, did baseline activities influence sellers by altering their beliefs about the likelihood their products will be scrutinized? Second, did information spillover from treated sites into neighboring market areas and affect buyer and seller knowledge and behavior there? We do not find evidence consistent with these two hypotheses and so in section 5 below we primarily focus on comparisons between treatment and control sites.

5 Results

In this section we present the main results as follows. In subsection 5.1, we confirm that observable markers correlate with lab-tested germination rates. This suggests that farmers can indeed use simple purchasing strategies based on observables to help them acquire better quality and in theory achieve sizeable agricultural gains. In subsection 5.2, we describe treatment effects on buyer knowledge and usage of quality markers, and seed purchases. The findings suggest that informed farmers were able to adjust their purchasing behavior, consistent with the channels in the model. In subsection 5.3 we examine agricultural outcomes. We find gains in yield among more remote farmers, where quality markers were associated with greater gains in seed quality. This finding corresponds to model prediction #1. The higher gains obtained by more remote farmers can be interpreted through the lens of the model, where lower initial knowledge, lower enforcement, and lower initial quality (all features of more remote market areas) lead to higher treatment effects. We also show that prices paid by locally-buying farmers are not affected, corresponding to model prediction #2. In subsection 5.4, we examine effects on sellers' decision to enter or exit the hybrid seeds market.

[7] We used survey data from the Tegemeo Institute.

We document a sizeable effect of treatment on seller exit, consistent with the negative effects of informed buyers on profit from model prediction #3. Lastly in subsection 5.5, we examine effects of information on prices and quality for uninformed buyers in the market. We do not find effects on seed quality for the uninformed, contrary to model prediction #4. We discuss mechanisms within the model that would make the treatment effect on quality for the uninformed to be too small to detect. As well, we explore the consequences if sellers are able to discriminate between informed and uninformed buyers, deviating from the model assumption that sellers make a take-it-or-leave it offer to informed buyers. This results in even smaller effects on uninformed buyers, or no effects at all in the case where sellers can perfectly discriminate (we describe the consequences in the model in Appendix B).

5.1 Summary statistics: quality markers and their correlates

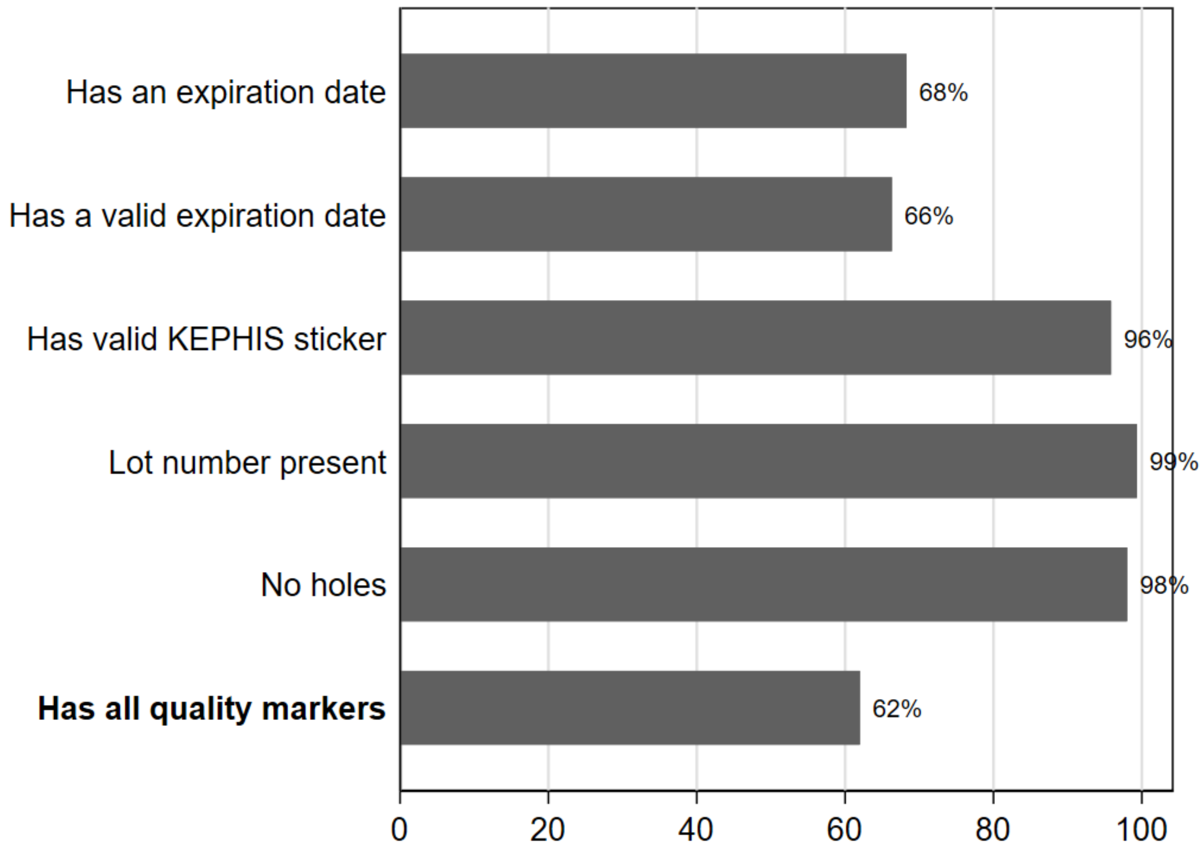
We first describe summary statistics for the data on quality markers that was collected. These descriptive statistics provide new insight into the prevalence of various observable quality markers and their correlates. Positive correlation between observable markers and objective lab-tested quality measures suggest that due to KEPHIS regulations that were implemented with some success, farmers stand to benefit if they use these markers to decide which seed packets to purchase.

Figure 3 presents summary statistics for each quality marker. Overall, 62% of packets that our team observed had all quality markers, while 38% were confirmed to be missing one or more. Most of the missing quality markers were due to not having an expiration date, or not having a valid expiration date on the packet, though missing or invalid KEPHIS stickers and the other observable markers also contributed somewhat to missing quality markers.

We next ask if observable quality markers correlate with aspects of seed quality from lab tests. Table 1 shows that observable quality markers are strongly correlated with germination rate but not purity. The relationship is driven primarily by two of the quality markers: (1) presence of a KEPHIS sticker, and (2) presence of a valid expiration date. Other aspects of seed quality are likely positively correlated with germination rate (Ghassemi-Golezani and Mamnabi, 2019), and so we view this as a lower bound for the yield increase that is possible if non-quality-certified seeds could be eliminated from the marketplace.

We view the observed 4.7% overall difference in lab tested germination rates as a lower bound for the changes in yield that could be achieved should farmers exert perfect discretion in choosing seed. It does not account for other characteristics of seeds that affect yield, which is expected to correlated positively with germination rate (Ghassemi-Golezani and Mamnabi, 2019) and whose effects on yield may far exceed the differences in germination rate. As well, this obviously would not account for any possible general equilibrium effects as the market adjusts. In light of findings from (Bold et al., 2017), this relatively small magnitude for differences in germination rate (which is more easily observable to farmers than other seed characteristics), may help explain the persistence of baseline market equilibrium – if low-quality seeds were even more inferior, farmers would be able to learn more quickly and possibly unravel the pooling equilibrium that we observe where both

Figure 3: Quality marker frequency



This figure shows the percentage of seed packets that feature each quality marker, calculated using data from market audit surveys which were submitted by secret shoppers on our research team. The sample includes 1495 seed packets. This includes observations from all three planting seasons in which market audit activities were conducted (long rains 2020, short rains 2020, and long rains 2021). It also includes both packets that were purchased as well as packets that were closely observed but not purchased. In the latter case, enumerators were not able to scratch the KEPHIS sticker and fully complete the e-verification steps. We categorize those packets as having a valid KEPHIS sticker so long as all other quality markers are present, and the sticker was present with all visible features of the sticker appearing to be valid. Overall, we find that 62% of packets had all quality markers present.

lower and higher quality seeds are sold by the same sellers at the same price.

The relationship between quality markers and lab-tested germination rate differs in more remote versus less remote markets. Table 2 shows this relationship for market areas that are farther than the median distance from the county capital, as opposed to markets that are closer than the median distance. More remote market areas have lower germination rates overall, showing the seed quality on average is farther from national standards. As well the quality markers are more informative in the sense that having all quality markers is associated with a greater gain in germination rate in more remote markets than in less remote markets.

These correlations between observable markers and lab-tested germination rates do not capture

Table 1: Quality markers, seed purity, and germination rate

	Purity	Germination Rate
Has all quality markers	0.01	5.21***
	(0.02)	(1.14)
Constant	99.84***	88.08***
	(0.01)	(0.96)
Observations	467	467

The sample includes 968 seed packets that were purchased by secret shoppers and tested in a lab by KEPHIS staff. The sample includes packets purchased during the long growing season in 2020, the short growing season in 2020, and the long growing season in 2021. The dependent variable in column 1 is the estimated percentage of material that are whole maize seeds (in percentage points). The dependent variable in column 2 is percentage of seeds that germinated in the lab (in percentage points). Table D3 and Table D4 show similar results for the complete list of quality markers.

all relevant aspects of seed quality. However, we expect germination rates to correlate with yields, and the tests provide high-quality objective evidence that farmers – particularly in more remote areas – stand to gain in agricultural output if they are able to successfully use these observables to choose higher quality seeds. Lab tests are not subject to self-reporting biases and testing staff are blinded to the seed source, variety, and any characteristics of the packaging. This also reflects a sort of limited success in the KEPHIS certification process – implementation failures and other mistakes, or widespread counterfeited marks have not prevented observables from being unrelated to seed quality.

5.2 Effects on household knowledge and seed purchases

In the following results, we will focus on overall treatment effects, as well as effects on less-remote and more-remote areas, with special emphasis on more-remote areas where baseline seed quality was worse while quality markers are associated with greater quality gains. Households in more remote areas appear similar socioeconomically (e.g. education levels, housing quality, plot size), though we note that more remote farmers tend to have higher yields and lower population density. We also expect more remote areas to be less intensely monitored by regulators.

We examine treatment effects on several measures of household knowledge, as shown in Table 3. Treatment substantially increased knowledge about visible quality markers, being associated with increased recall of specific elements by 35% to 120%. We note that household-level effects appear to be heterogeneous. For households with more highly educated households heads with primary school or secondary schooling completed, knowledge (by all measures) both started at a higher level as well as increased substantially more due to treatment. Generally we see that households in more remote areas start with lower knowledge and generally have slightly higher gains in knowledge, though not statistically significantly so. Knowledge levels and treatment effects among treated households are similar in more and less remote areas.

Table 2: Remote markets have more to gain

	Germination Rate
Has All Quality Markers * Not Remote	3.30**
	(1.48)
Has All Quality Markers * Remote	7.50***
	(1.77)
Remote	-5.50***
	(1.93)
Constant	90.47***
	(1.27)
Observations	467

The sample includes 467 seed packets that were purchased by secret shoppers and tested in a lab by KEPHIS staff, and were collected from control markets. The dependent variable is the percentage of maize seeds that germinated in the lab (in percentage points). Remote markets are defined as markets with above-median distance to the county capital. "Has all QMs" equals 1 if a seed packet has all quality markers, and equals 0 otherwise.

We also examine the possibility that households at control sites near treated sites may have experienced gains in knowledge due to spillovers. However, the data suggests that any spillover effects that may have occurred are too small to be detectable. As shown in Appendix C, we do not find evidence of spillover effects within 2km, 4km, or 6km. Partly by design, due to our multi-stage sampling strategy, the sample does not have many markets that are clustered very closely together and so we do not have statistical power to detect spillovers across very short distances. At the same time, few markets are at closer distances where we might bee concerned that spillovers could greatly affect estimates of treatment effects. In any case, one could think of the following treatment effects as being lower bounds, to the extent that undetected spillovers affect nearby control sites similarly but to a smaller degree than the treatment sites.

A priori, one concern that could limit the effectiveness of treatment is that many farmers could already possess the information we seek to disseminate. These results show this is not the case and confirm that treatment was able to increase knowledge of quality markers, from a relatively low baseline. While discouraging from the point of view of KEPHIS, it is encouraging for our research design that we find low awareness for the SMS-based verification system, suggesting that the intervention was well-timed to publicize it. As our contacts at KEPHIS explained, they have yet to run a major campaign to inform buyers of the verification system, while the stickers themselves say little about their purpose (see Appendix A). As shown in Table 4, households had little baseline knowledge about quality markers, on average naming only 0.55 markers, and with less than 7% of farmers citing the seeds' expiration date.

Given the timing of data collection, we think it is likely that the increases in knowledge represent lower bounds on the true increase in knowledge due to treatment that is relevant at the time of seed purchases. These data were collected in July 2020 at the earliest, and respondents did not know in advance when they would be called by the research team. While originally planned for immediately after the planting season ended in late April, field operations were disrupted by the onset of the COVID-19 pandemic, and so our first post-treatment measurements of knowledge come from a phone survey with the main respondent approximately five months after treatment.

Table 3: Household knowledge and usage of markers

Panel A: Household Knowledge

	(1) no. quality marks known	(2) e-verif known	(3) expiration known
treated	0.207***	0.104***	0.025***
	(0.035)	(0.011)	(0.008)
Observations	4645	4645	4645
Control Mean	0.56	0.13	0.07
Treatment Effect (%)	37.11	79.95	37.13

Panel B: Usage of markers

	(1) Able to verify	(2) Used e-verif successfully
treated	0.070***	0.075***
	(0.011)	(0.009)
Observations	4645	4645
Control Mean	0.14	0.07
Treatment Effect (%)	51.49	106.68

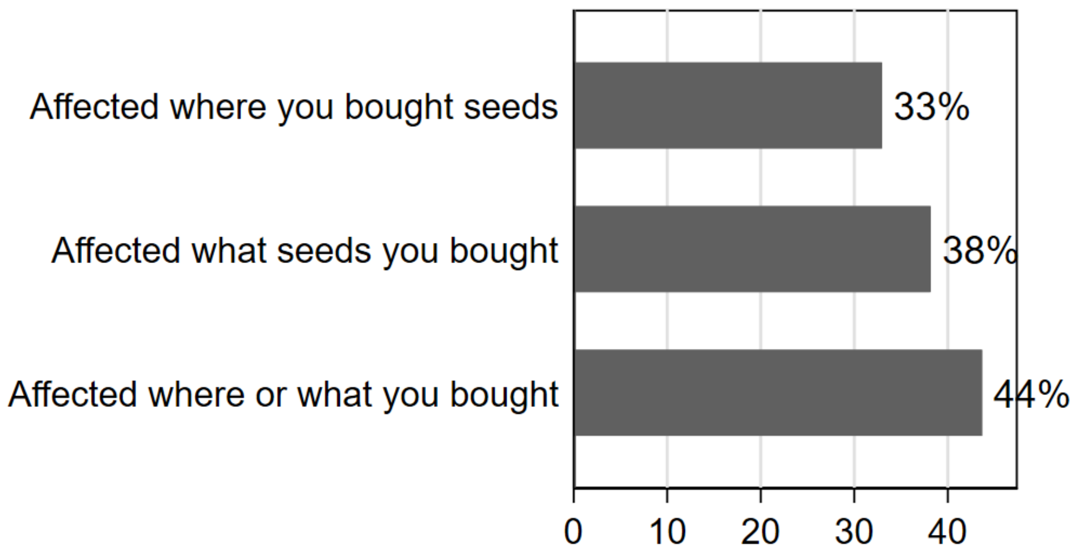
The table shows the effects of information treatment on household knowledge and household usage of markers. The sample includes observations from both the main planting season in 2020 and the short season in 2020, with knowledge measured approximately in late July 2020 and in November 2020. Standard errors are clustered by sublocation. In Panel A, the dependent variables are (1) the number of quality markers named by the respondent, (2) whether e-verification was named as a way to verify seed quality (=1 if yes), (3) whether expiration date was named (=1 if yes), and (4) whether e-verification was reported to be understood at least somewhat well (=1 if yes). In Panel B, the dependent variables are (1) whether the respondent was able to verify that seeds from that season were of high quality, and (2) whether e-verification was used successfully.

After the planting season, we asked treated households whether the treatment affected their seed purchase decisions. Respondents were asked the following questions: (1) "Did the information we provided about quality markers help?", (2) "Did it influence your decision of where to buy seeds? Please explain", (3) "Did it influence your decision of which seeds to buy? Please explain". Figure 4 summarizes their responses, in which 44% of respondents said that the information affected what seeds they purchased or where they purchased them.

We examine in Table 4 whether treatment affected whether and where households bought hybrid maize seed. Columns 1 and 2 show that households shifted their source for seeds, with 5.8% fewer households (though this is 17.7% of households that are expected to purchase locally) sourcing seeds from the local market, and a similar amount (19.2% of the control mean) obtaining seeds elsewhere. Consistent with these findings, evidence from Gharib et al. (2021) suggests that farmers have higher willingness to pay for seed packets that have certain quality markers present.[8] These results are consistent with the mechanisms in the model, in which informed buyers that cannot obtain high

[8] Gharib et al. (2021) estimates the effect of training (on two of our seven quality markers) on willingness-to-pay for maize seed packets.

Figure 4: Do treated households think the information helped?



The sample includes 742 respondents in treated market areas. These respondents were asked the following questions: (1) "Did the information we provided about quality markers help?", (2) "Did it influence your decision of where to buy seeds? Please explain", (3) "Did it influence your decision of which seeds to buy? Please explain". 44% of respondents indicated that the information that was provided affected either where or what seeds were purchased.

quality seeds locally find it worthwhile to continue their search by buying outside the local market.

Overall, however, households were not more likely to buy hybrid seeds. It is possible that even if farmers updated expectations about the quality of seeds that they can expect to get (when using the quality markers), the shift is not near the threshold of pushing expected economic returns to be positive. Another consideration is that additional market imperfections barriers (such as credit constraints or incomplete insurance markets) may be important in limiting movement on the adoption margin (Karlan et al., 2014).

5.3 Effects on household agricultural outcomes

We next examine how the information campaign affected household agricultural outcomes. Germination rate, harvest amount, and yield (kilograms harvested per acre) are the main outcomes we will examine in this section. While inherently noisy and affected by other farmer inputs, these agricultural outcomes also capture elements of seed quality that are not captured in lab tests.

Table 4: Household Seed Choice

Panel A: Full sample

	(1) Bought hybrid at local market	(2) Bought hybrid elsewhere	(3) Bought hybrid
treated	-0.058***	0.053***	-0.001
	(0.019)	(0.018)	(0.013)
Observations	3717	3771	3771
County FE	Yes	Yes	Yes
HH Baseline Controls	Yes	Yes	Yes
Control Mean	0.33	0.28	0.61
Treatment Effect (%)	-17.74	19.27	-0.12

Panel B: Less remote

	(1) Bought hybrid at local market	(2) Bought hybrid elsewhere	(3) Bought hybrid
treated	-0.059**	0.043*	-0.015
	(0.029)	(0.026)	(0.019)
Observations	1852	1876	1876
County FE	Yes	Yes	Yes
HH Baseline Controls	Yes	Yes	Yes
Control Mean	0.35	0.29	0.64
Treatment Effect (%)	-16.85	14.94	-2.33

Panel C: More remote

	(1) Bought hybrid at local market	(2) Bought hybrid elsewhere	(3) Bought hybrid
treated	-0.051**	0.054**	0.013
	(0.023)	(0.023)	(0.017)
Observations	1865	1895	1895
County FE	Yes	Yes	Yes
HH Baseline Controls	Yes	Yes	Yes
Control Mean	0.29	0.26	0.56
Treatment Effect (%)	-17.47	20.96	2.33

This table shows treatment effects on household seed purchase decisions. Results in Panel A include households in both the main season in 2020 and the short season in 2020. All standard errors are clustered by sublocation. The dependent variable in column 1 equals 1 if the respondent bought hybrid seeds at the local market in that planting season (and 0 otherwise). The dependent variable in column 2 equals 1 if the respondent bought hybrid seeds from a location other than the local market in that planting season (and 0 otherwise). The dependent variable in column 3 equals 1 if the respondent bought hybrid seeds from any source in that planting season. In panels B and C, we split the sample by the median distance of a market to the county capital.

Table 5 shows that treatment caused a noisy increase in yields among households, with the more substantial gains among households in smaller and more remote markets. Effects on self-reported and lab-measured germination rates are weakly positive at best and are insufficient to explain measured increases in yields. There are a few possibilities to explain this despite positive effects on yields. Expired seeds do not germinate as well, while non-hybrid seed may germinate as well as hybrid seed but will have less yield. We cannot rule out the presence of both of these types of sub-standard seed. Research shows that other aspects of poor quality seed can affect yields in ways other than those mediated by germination rates (Ghassemi-Golezani and Mamnabi, 2019). Of course, it is also possible that self-reported germination rates are too noisy to detect the relatively modest differences that we expect between high and low quality, being a fact from months before harvest and being possibly a less salient fact for farmers to remember.

To check the plausibility of these yield gains, let's assume that 38% of packets do not have a quality marker (average in control group for secret shoppers), and remote markets see a 7.7% increase in yield. Assuming the households follow the strategy from the model in section 3, then this implies that informed buyers that got a different quality level than in the counterfactual got a 20.3% gain in yield, while the remaining buyers would have gotten high quality even in the counterfactual and had a gain of 0%.

For reference figures to help us judge the plausibility of these gains we can look toward related research. Fabregas et al. (2019) do a meta-analysis of 7 digital agricultural extension interventions and find an average of 4% increase in output. Bold et al. (2017) find a 13-18% increase in yield when switching from retail to wholesale quality seeds; they find a 28-38% increase in yield when switching from local to wholesale quality seeds.[9] Ghassemi-Golezani and Mamnabi (2019) used artificial aging to lower seed quality seeds that had 2% or 7% lower germination. Compared to these comparison groups, the higher quality seeds had 23.5% or 64.5% higher yield. Thus, our results compare favorably to other agricultural informational interventions, and the magnitudes are comparable with what might be expected (given evidence from poor quality seeds in Uganda, or evidence on the performance of old seeds from controlled trials).

We do not see any effects on the prices that households paid (Table D21), nor do we see evidence of greater price dispersion among treated households (Table D22). This suggests that the estimated agricultural gains led to welfare gains for the households.

Gains in remote areas are driven by female-headed households (Table D9). This suggests that resolving information frictions may help close the gender gap in agricultural productivity in LMICs (Diiro et al 2018; Wambua et al 2018). This finding also mirrors a result from Annan et al (2021) in which female customers in Ghana benefit more than male customers from an information treatment that helps them avoid being defrauded by mobile money agents.

For completeness we show results splitting the sample by remoteness as well as the other three pre-specified subsets of farmers to examine heterogeneous effects. As discussed, we see larger gains (and significantly so) for more remote farmers (Table D12), and we see somewhat larger gains (though not statistically significantly so) for female-headed households (Table D13) and lower education households (Table D14). Table D15 Together with the evidence we have seen from more remote farmers as compared to less remote farmers, this is suggestive of a pattern in which larger

[9] Even larger increases were seen when fertilizer quality was also increased.

gains are experienced by sub-groups with lower seed quality, since female-headed households and lower-education households self-report lower germination rates and have less knowledge abut quality markers (see Table D10). These sub-groups also tend to be less likely to use hybrid seeds.

One caveat to our results is that we cannot rule out changes in other complementary inputs (e.g. labor input, quality of fertilizer and pesticide, etc). We are able however to check whether the information treatment affected usage of fertilizer, and we do not see effects (see Table D16)

5.4 Effects on seller entry/exit

We next examine the effect of information treatment on seller entry and exit. The dependent variable in this analysis is the number of sellers present at the market. [10] As shown in Table 6, the number of sellers per market decreased by about 0.27 seller in the SGS2020 and LGS2021 seasons due to treatment. This is about 17% percent decrease in the number of sellers. Panels B and C show that these effects are driven especially by less-remote markets with decrease of 0.42 sellers (24% decrease).

The pattern of sellers exiting due to treatment is consistent with model prediction #3, where informed buyers cause sellers to expect increased costs, fewer customers, and greater risk of punishment, all driving down expected profits. The pattern of sellers exiting due to treatment is also confirmed in our separate dataset of recruited sellers, from whom we collected baseline and endline data. We speculate that the pattern of greater treatment effects on seller exit in less-remote markets may be due to their having better alternative business opportunities. We observe that sellers in less remote markets tend to offer a greater variety of agricultural inputs, being both more likely to sell other types of vegetable seed as well as being more likely to sell other agricultural inputs.

Other possible stories do not appear to find support in the data. For example, based on our household surveys of local residents (see Table 4), sellers in less remote markets do not appear to have lost more market share due to treatment when compared to less remote areas. As well, we don't see evidence consistent with lower profit margin per unit in less remote areas (no differences in effects on prices, and no differences in measured profit-per-unit among stayers.) While firms with lower profits were somewhat more likely to exit (Table D18), we do not see differences in baseline profit between more and less remote sellers.

5.5 Effects on seed quality and prices for uninformed buyers

The data show no detectable effects on visible quality markers or on purity and germination rates of sampled seeds obtained by covert shoppers working for the research team.

[10] Due to logistical difficulties during the main planting season in 2020, we did not collect the total number of sellers present at markets if there were more than 2 seed sellers, and we did not visit all markets. Due to the limit time in March 2020 available before field operations were forced to stop, the field team visited a subset of about 75% of markets in the sample, with this subset driven by logistical considerations and time constraints rather than treatment status or other market characteristics. For the 2020 main planting season, we know for this subset of markets how many sellers were present, or if there are more than 2 sellers, then how many are present among up to two sellers that were randomly selected during baseline activities.

Table 5: Household agricultural outcomes (All seasons)

Panel A: Full sample

	(1) Germination rate	(2) Kgs harvested	(3) Yield
treated	-0.22	32.86	24.55
	(0.72)	(39.96)	(30.21)
Observations	2274	2328	2111
County FE	Yes	Yes	Yes
HH Baseline Controls	Yes	Yes	Yes
Control Mean	85.01	850.07	890.77
Treatment Effect (%)	-0.26	3.87	2.76

Panel B: Less remote markets

	(1) Germination rate	(2) Kgs harvested	(3) Yield
treated	-0.95	-40.84	-41.36
	(1.08)	(56.17)	(46.33)
Observations	1182	1241	1110
County FE	Yes	Yes	Yes
HH Baseline Controls	Yes	Yes	Yes
Control Mean	85.10	745.45	866.14
Treatment Effect (%)	-1.11	-5.48	-4.78

Panel C: More remote markets

	(1) Germination rate	(2) Kgs harvested	(3) Yield
treated	-0.13	104.61*	77.79*
	(0.96)	(61.21)	(42.52)
Observations	1092	1087	1001
County FE	Yes	Yes	Yes
HH Baseline Controls	Yes	Yes	Yes
Control Mean	84.88	1013.02	928.37
Treatment Effect (%)	-0.15	10.33	8.38

This table shows treatment effects on household agricultural outcomes. Results in Panel A include households in both the main season in 2020 and the short season in 2020. All standard errors are clustered by sublocation. The dependent variable in column 1 is the self-reported germination rate of seeds planted in that season (in percentage points). The dependent variable in column 2 is the number of kilograms harvested in that planting season. The dependent variable in column 3 is the nubmer of kilograms per acre achieved in that planting season. In panels B and C, we split the sample by the median distance of a market to the county capital.

As Table 7 shows, overall and in more remote markets, we see a 0-2% increase in packets with all quality markers present. However, these estimates are not statistically significantly different from zero and are far too small to credibly be a driving reason for the increases in yields of the magnitudes that we see. We do not detect effects on purity or germination rate from lab tests.

Table 6: Seller Entry/Exit

Panel A: Full sample

	Number of sellers at market
treated	-0.273**
	(0.133)
Observations	495
County FE	Yes
Control Mean	1.58
Treatment Effect (%)	-17.26

Panel B: Less Remote

	Number of sellers at market
treated	-0.417**
	(0.186)
Observations	256
County FE	Yes
Control Mean	1.71
Treatment Effect (%)	-24.38

Panel C: More Remote

	Number of sellers at market
treated	-0.114
	(0.193)
Observations	239
County FE	Yes
Control Mean	1.39
Treatment Effect (%)	-8.18

This table shows treatment effects on seller entry and exit in the local market. The sample in Panel A includes markets in the main season in 2020 and the short season in 2020. All standard errors are clustered by sublocation. The dependent variable equals the number of sellers present at that market, as reported by secret shoppers. In panels B and C, we split the sample by the median distance of a market to the county capital.
market.

We interpret this finding to mean that uninformed buyers in treated markets did not obtain higher quality. One reason effects on the uninformed would be even smaller would be if sellers can discriminate between informed and uninformed buyers. In the extreme case in which sellers can perfectly discriminate between informed and uninformed buyers, the uninformed never benefit at all. We formalize this case in Appendix B.

We also do not see evidence that treatment affected prices for uninformed buyers as revealed by the secret shopper data (Table D23).

Table 7: Quality Markers (all seasons)

Panel A: Full sample

	(1)	(2)	(3)
	Has all quality markers	purity	germ_rate
treated	0.012	-0.277	-0.539
	(0.028)	(0.274)	(0.753)
Observations	1212	878	878
County FE	Yes	Yes	Yes
Control Mean	0.57	99.85	91.84
Treatment Effect (%)	2.16	-0.28	-0.59

Panel B: Less Remote

	(1)	(2)	(3)
	Has all quality markers	purity	germ_rate
treated	0.005	-0.007	-1.325
	(0.041)	(0.017)	(0.853)
Observations	654	462	462
County FE	Yes	Yes	Yes
Control Mean	0.59	99.85	92.98
Treatment Effect (%)	0.82	-0.01	-1.42

Panel C: More Remote

	(1)	(2)	(3)
	Has all quality markers	purity	germ_rate
treated	0.038	-0.548	1.033
	(0.039)	(0.553)	(1.405)
Observations	558	416	416
County FE	Yes	Yes	Yes
Control Mean	0.53	99.84	90.03
Treatment Effect (%)	7.16	-0.55	1.15

This table shows treatment effects on seed quality offered to secret shoppers posing as uninformed buyers. The sample in Panel A includes markets in the main season in 2020, the short season in 2020, and the main season in 2021. All standard errors are clustered by sublocation. The dependent variable in column 1 equals 1 if the seed packet has all quality markers present. The dependent variable in column 2 is the estimated percentage of material that are whole maize seeds (in percentage points). The dependent variable in column 3 is percentage of seeds that germinated in the lab (in percentage points). the number of sellers present at that market, as reported by secret shoppers. In panels B and C, we split the sample by the median distance of a market to the county capital. In panels B and C, we split the sample by the median distance between a market and the county capital.

6 Discussion

Widespread and effective information campaigns can be difficult and costly to implement. While we pursue an in-person community approach to disseminating information, this design was informed at least partly by research needs – limiting spillovers, simplicity, highly targeted exposure to information, bundling treatment with data collection activities to save on costs. Using plausible assumptions to estimate the cost of only providing the information campaign (with no data collection costs), we estimate that 22,000 USD was spent to train 12255 residents via face-to-face conversations. This is equivalent to 1.77 USD per person. Valuing harvested maize at 30 USD per 90 kilograms, the benefit to more remote farmers is about 34:87 USD.

Of course, we take such cost-effectiveness figures with a grain of salt. Point estimates come with confidence intervals, and extrapolating will depend on being able to effectively target, given the heterogeneous effects in our study sample. One may think of the intervention as lowering the cost of acquiring information. We would expect that other related information interventions may have similar effects and may scale well. For example, reforming the messaging in existing e-verification system to improve accessibility to the information, updating standards for product labeling including standardizing formatting for expiration dates, and a mass-media campaign via newspaper or radio may cost-effectively disseminate information to consumers.

7 Conclusion

In this paper, we study empirically the effects of reducing information frictions in the hybrid maize seed market in rural Kenya. We evaluate a randomized market-level information campaign to examine quantify effects for both informed and uninformed buyers in treated market areas. We monitored prices and seed quality that sellers offer for over one year to allow for sellers to respond to the increase in the number of informed buyers.

We show first that farmers stand to gain from receiving information about observable quality markers. The information campaign affected farmers' purchasing decisions and led to gains in maize yields, particularly in more remote areas where baseline seed quality is lowest and the observable quality markers are associated with the greatest gain in seed quality. Second, while improved information caused sellers to exit the market, we do not observe effects on prices or quality among the stayers. This is consistent with a model in which sellers are able to discriminate between informed and uninformed buyers. Patterns in the data for other outcomes also tell a story consistent with a simple model in which treatment induces households to adjust purchasing behavior, and dampen expected seller profit, causing sellers to exit.

Our findings suggest that that policies that help improve consumer information may be beneficial to supporting the productivity of small-holder farmers. Reforming existing requirements for information displayed on seed packets for consumers to promote greater understanding about seed certification may be helpful. Yet, the absence of benefits in this study for uninformed consumers points to remaining challenges to overcome the difficult learning and reputation-building environment and their impacts on the market equilibrium.

References

Aker, Jenny and Joshua Blumenstock, "The Economic Impacts of New Technologies in Africa," in Celestin Monga and Justin Lin, eds., *Handbook of Africa and Economics*, Oxford: Oxford University Press, 2015.

Akerlof, George A., "The Market for "Lemons": Quality Uncertainty and the Market Mechanism," *The Quarterly Journal of Economics*, 1970, *84* (3), 488–500.

Annan, Francis, "Misconduct and Reputation under Imperfect Information," Working Paper, Georgia State University July 2021.

Ashour, Maha, Daniel Orth Gilligan, Jessica Blumer Hoel, and Naureen Iqbal Karachiwalla, "Do Beliefs About Herbicide Quality Correspond with Actual Quality in Local Markets? Evidence from Uganda," *Journal of Development Studies*, June 2019, *55* (6), 1285–1306.

Bai, Jie, "Melons as Lemons: Asymmetric Information, Consumer Learning and Seller Reputation," Technical Report 2021.

Bergquist, Lauren Falcao and Michael Dinerstein, "Competition and Entry in Agricultural Markets: Experimental Evidence from Kenya," *American Economic Review*, December 2020, *110* (12), 3705–3747.

Bjorkman-Nyqvist, Martina, Jakob Savensson, and David Yanagizawa-Drott, "Can Good Products Drive Out Bad? Evidence from Local Markets for (Fake?) Antimalarial Medicine in Uganda," CID Working Papers 242, Center for International Development at Harvard University August 2012.

Bold, Tessa, Kayuki C. Kaizzi, Jakob Svensson, and David Yanagizawa-Drott, "Lemon Technologies and Adoption: Measurement, Theory and Evidence from Agricultural Markets in Uganda," *The Quarterly Journal of Economics*, 2017, *132* (3), 1055–1100.

Bronnenberg, Bart J., Jean-Pierre Dubé, Matthew Gentzkow, and Jesse M. Shapiro, "Do Pharmacists Buy Bayer? Informed Shoppers and the Brand Premium," *The Quarterly Journal of Economics*, 2015, *130* (4), 1669–1726.

Dal Bó, Ernesto and Frederico Finan, "At the Intersection: A Review of Institutions in Economic Development," *The Handbook of Economic Development and Institutions*, 2020.

Dranove, David and Ginger Zhe Jin, "Quality Disclosure and Certification: Theory and Practice," *Journal of Economic Literature*, December 2010, *48* (4), 935–963.

Evenson, R E and D Gollin, "Assessing the impact of the green revolution, 1960 to 2000," *Science*, May 2003, *300* (5620), 758–762.

Fabregas, R., M. Kremer, and F. Schilbach, "Realizing the potential of digital development: The case of agricultural advice," *Science*, 12 2019, *366* (6471).

FAO, *2019 Africa Regional Overview of Food Security and Nutrition*, FAO, 2020.

FAO, *2020 Africa regional overview of food security and nutrition*, FAO, 2021.

Gharib, Mariam H., Leah H. Palm-Forster, Travis J. Lybbert, and Kent D. Messer, "Fear of fraud and willingness to pay for hybrid maize seed in Kenya," *Food Policy*, 2021, *102* (C).

Ghassemi-Golezani, Kazem and Sohrab Mamnabi, "SOME PHYSIOLOGICAL RESPONSES AND YIELD OF MAIZE AFFECTED BY SEED AGING AND PRIMING DURATION," *Plant Breeding and Seed Science*, 12 2019, *79*, 63–70.

Gonzalez-Lira, Andres and Ahmed Mushfiq Mobarak, "Slippery Fish: Enforcing Regulation when Agents Learn and Adapt," Working Paper 28610, National Bureau of Economic Research March 2021.

Grubb, Michael D. and Matthew Osborne, "Cellular Service Demand: Biased Beliefs, Learning, and Bill Shock," *American Economic Review*, January 2015, *105* (1), 234–271.

Handel, Benjamin and Joshua Schwartzstein, "Frictions or Mental Gaps: What's Behind the Information We (Don't) Use and When Do We Care?," *Journal of Economic Perspectives*, Winter 2018, *32* (1), 155–178.

Holden, Stein T., "Economics of Farm Input Subsidies in Africa," *Annual Review of Resource Economics*, October 2019, *11* (1), 501–522.

Jack, B Kelsey, "Market Inefficiencies and the Adoption of Agricultural Technologies in Developing Countries," Technical Report 2013.

Jin, Ginger Zhe and Phillip Leslie, "Reputational Incentives for Restaurant Hygiene," *American Economic Journal: Microeconomics*, February 2009, *1* (1), 237–267.

Karlan, Dean, Robert Osei, Isaac Osei-Akoto, and Christopher Udry, " Agricultural Decisions after Relaxing Credit and Risk Constraints *," *The Quarterly Journal of Economics*, 02 2014, *129* (2), 597–652.

Kenya Association of Manufacturers, "The Study to determine the severity of the counterfeit problem in Kenya," Technical Report 2012.

Kenya National Bureau of Statistics, "Kenya Population and Housing Census Results," Technical Report 2019.

Michelson, Hope, Anna Fairbairn, Brenna Ellison, Annemie Maertens, and Victor Manyong, "Misperceived quality: Fertilizer in Tanzania," *Journal of Development Economics*, 2021, *148* (C).

Muchiri, Joseph, "Fight against fake seeds bearing fruit, says Kephis official," *The Standard*, Jul 2019.

Naritomi, Joana, "Consumers as Tax Auditors," *American Economic Review*, September 2019, *109* (9), 3031–3072.

OECD/FAO, "OECD-FAO Agricultural Outlook," July 2021.

Okinda, Brian, "Alarm as fake seed merchants come up with new ways of selling products," *The Daily Nation*, June 2019.

Shapiro, Carl, "Premiums for High Quality Products as Returns to Reputations," *The Quarterly Journal of Economics*, 1983, *98* (4), 659–679.

Shiferaw, Bekele, Boddupalli M. Prasanna, Jonathan Hellin, and Marianne Bänziger, "Crops that feed the world 6. Past successes and future challenges to the role played by maize in global food security," August 2011, *3* (3), 307–327.

Wolinsky, Asher, "Prices as Signals of Product Quality," *Review of Economic Studies*, 1983, *50* (4), 647–658.

A Appendix A: Information campaign details

Figure A1: Treatment #1 Flyer

How to check maize seed quality

Quality markers to look for:

1) Clearly displayed company logo together with the seed variety
2) Clearly displayed weight of the seed packet. For example: 2kg, 10kg.
3) Packaging with no holes. The seed packet should not have broken or taped seals that could allow seeds to be removed or tampered with.
4) A printed lot number—this code allows seed packets to be easily traceable
5) A recent packaging or testing date. These dates will tell you if the seeds are old. Older seeds may germinate poorly.
6) Seeds should not be split or broken, and the coloring on the seeds should not come off easily
7) A KEPHIS sticker:

- Check that the SMS code is valid
- Check that the variety, lot number, and testing date match the packaging

Valid	Not valid
OK Monsanto DK8031 Species: Zea mays Variety:DK8031 Lot No: 18-23463HP Class: C1G Testing Date: Jan/2019 More: ghub.ai/zH9d	No 219823200694 IS NOT A VALID CODE. Check and send correct code. The seed may not be genuine. Call 0709891000 or ke@mpedigree.net. More: ghub.ai/X2XD
	No: 261710114026 was a Valid Code BUT was used on 2019-03-14 14:23:56 by 7151****8. MPedigree service. hub.goldkeys.net More: ghub.ai/zzfc

Figure A2: Treatment #2 Flyer

If you are concerned about the quality of your seeds, you may contact the following:

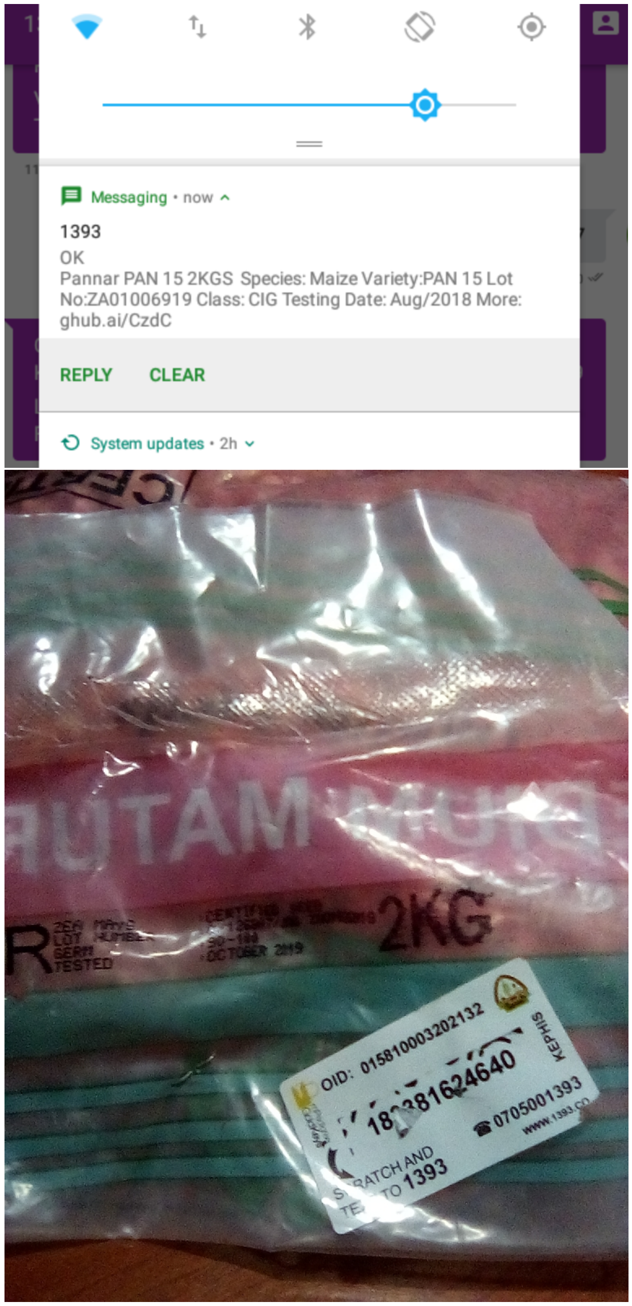
- Kenya Plant Health Inspectorate Service (KEPHIS)
  Phone Number: 020-3597209
- Anti-Counterfeit Authority (ACA)
  Phone Number: 020 2280111
- Your assistant chief

Other Information:

- Not all seeds are of the same quality. It is important to discuss with friends, family, and neighbors the quality of seeds that you purchase. This can help them avoid seeds that may be lower quality.
- When buying and planting maize seeds, it can be helpful to ask the seller for a receipt and keep the empty packet after planting. In case there is any issue with the quality of the seeds, you can refer to these items to confirm the type of seed, as well as the time and place where you purchased the seeds.

Figure A3: Maize variety with and without quality marks

Figure A4: E-verification



B Appendix B: Model

B.1 Model details

In section 3, the firm's problem is as follows, where we write x as a function of the firm's choice of price p and quality mix q:

$$Max_{p,q}(p - c - d\frac{q}{1-\theta+q\theta}) * x(p,q) - \frac{1}{2}b\theta(1-q)^2N - F$$

Accounting for how consumers respond to the firm's choice of p and q, the firm's problem becomes:

$$Max_q(p - c - d\frac{q}{1-\theta+q\theta})(1-\theta+\theta q)N - \frac{1}{2}b\theta(1-q)^2N - F$$

This simplifies to:

$$Max_q(p-c)(1-\theta+\theta q)N - dqN - \frac{1}{2}b\theta(1-q)^2N - F$$

Taking derivatives with respect to q, we get:

$$\frac{\partial}{\partial q}(\cdot) = (p-c)\theta N - dN + b\theta(1-q)N$$
$$\frac{\partial^2}{\partial q^2}(\cdot) = -b\theta N \leq 0$$

So the second derivative of the firm's objective function with respect to q is everywhere weakly negative. Thus, we can define conditions under which the firm chooses $q = 1$ or $q = 0$:

$$q^* = 1 \text{ if } \frac{\partial}{\partial q}(\cdot)\bigg|_{q=1} = (p-c)\theta N - dN \geq 0$$
$$q^* = 0 \text{ if } \frac{\partial}{\partial q}(\cdot)\bigg|_{q=0} = (p-c)\theta N - dN + b\theta N \leq 0$$

An interior solution exists when $(p-c)\theta - d \in (-b\theta, 0)$. In this case, we can solve for the optimal quality:

$$q^* = \frac{(p-c)\theta - d + b\theta}{b\theta} = \frac{(p-c)\theta - d}{b\theta} + 1$$

This relationship is increasing in θ, but this relationship weakens when enforcement is higher.

$$\frac{\partial q^*}{\partial \theta} = \frac{d}{b\theta^2}$$
$$\frac{\partial^2 q^*}{\partial \theta \partial b} = -\frac{d}{b^2\theta^2}$$

B.2 Model with perfect discrimination

We present in this sub-section a variant of the base case model. What if sellers can discriminate between informed and uninformed consumers? For example, sellers may be able to sell lower quality, unless informed consumers reveal themselves by asking about a missing quality marker, in which case the seller is able to offer a different quality mix.

Below, we consider perfect discrimination, where the seller can sell different quality mixes q_1 to informed buyers and q_2 to uninformed buyers. The firm's costs now are:

- $(c + d)\theta N q_1$ for the informed
- $(c + dq_2)(1 - \theta)N$ for the uninformed
- $\frac{1}{2}b\theta(1 - q_1)^2 N$ for expected punishment

The firm's problem is:

$$Max_{p,q_1,q_2}(p - c - d)\theta q_1 N + (p - c - dq_2)(1 - \theta)N - \frac{1}{2}b\theta(1 - q_1)^2 N - F$$

Since only the informed will leave the local market or report, the firm's optimal choice for q_2 is zero. Taking the derivative with respect to q_1, we get:

$$\begin{aligned}\frac{\partial}{\partial q_1}(\cdot) &= (p - c - d)\theta N + b\theta(1 - q_1)N \\ &\geq 0\end{aligned}$$

Optimal quality for the informed is $q_1^* = 1$. And so the quality that informed consumers are offered doesn't depend on θ.

Also as with the model without seller discrimination, more informed customers is bad for profits:

$$\begin{aligned}\frac{\partial \pi^*}{\partial \theta} &= (p - c - d)N - (p - c)N \\ &= -dN \\ &\leq 0\end{aligned}$$

Summary of model predictions:

1. Buyers that get informed obtain higher quality
2. Having more informed buyers does not affect market prices ($\frac{\partial p}{\partial \theta} = 0$)
3. Having more informed consumers has a negative effect on seller profit ($\frac{\partial \pi}{\partial \theta} \leq 0$)
4. Uninformed consumers never benefit. They always receive low quality. ($\frac{\partial q_2}{\partial \theta} = 0$)

B.3 Conditions for equilibrium

In section 3 we consider the equilibrium in which sellers set price low enough to entice uninformed buyers and informed buyers who are offered high quality to buy locally. Below, we discuss the alternative cases where price is set higher or lower.

Alternative case #1: the firm could choose a lower price to keep even informed consumers who get low quality. Those consumers get $1 - P_0 - S$ in the outside option and $-p$ at the local firm. Therefore, this informed consumer who gets low quality will stay if: $-p \geq 1 - P_0 - S$, or $p < P_0 + S - 1$.

The optimal quality in this case will be q=0. The firm's profits are:

$$(p - c)N - \frac{1}{2}b\theta - F$$

Price cannot drop below c or else the firm is sure to exit the market. Thus we can rule out this scenario if buyer's valuation of a high quality product is large enough. We also note that we see a good number of packets with all quality markers, suggesting that we generally do not observe an equilibrium with $q = 0$.

Alternative case #2: the firm chooses a higher price $P_0 + S$ to keep only the informed buyers who get high quality (uninformed buyers feel they are better off getting the higher quality in the outside option). Note that as $\hat{q} - \bar{q}$ goes to zero, this becomes a sub-optimal choice for any values for the parameters, since the firm loses a large segment of customers (the uninformed customers) for a vanishingly small gain in price per unit, and must stock exclusively high quality for all of the informed customers to keep them buying locally and prevent the outside option from being the more attractive option. We note also that the share of uninformed is generally large in our sample. Among control group households, and using either knowledge of expiration date or knowledge of the KEPHIS sticker to mean that the consumers is "informed", we estimate that 84% are uninformed.

C Appendix C: Spillovers Effects And Effects of Baseline Surveys

C.1 Spillover Effects

In this subsection, we examine whether nearby treatment sites had effects on the major outcomes. To do so, in the main estimating equation to recover treatment effects, we control for the number of study sites within 2, 4, or 6 kilometers, and we also include the number of treated sites within that distance.

In Table C1 we examine possible spillovers on household knowledge.

Table C1: Households knowledge (spillovers)

Panel A: Spillovers within 2km

	(1) no. quality marks known	(2) e-verif known	(3) expiration known
Number of treatment sites within 2km	-0.035	0.002	0.003
	(0.057)	(0.018)	(0.013)
Number of control sites within 2km	0.024	0.027	-0.004
	(0.061)	(0.019)	(0.014)
Observations	2305	2305	2305
County FE	Yes	Yes	Yes
HH Baseline Controls	Yes	Yes	Yes

Panel B: Spillovers within 4km

	(1) no. quality marks known	(2) e-verif known	(3) expiration known
Number of treatment sites within 4km	-0.019	0.006	-0.004
	(0.031)	(0.010)	(0.007)
Number of control sites within 4km	0.018	0.011	-0.001
	(0.025)	(0.008)	(0.006)
Observations	2305	2305	2305
County FE	Yes	Yes	Yes
HH Baseline Controls	Yes	Yes	Yes

Panel C: Spillovers within 6km

	(1) no. quality marks known	(2) e-verif known	(3) expiration known
Number of treatment sites within 6km	0.014	0.008	0.004
	(0.020)	(0.006)	(0.005)
Number of control sites within 6km	0.033**	0.007	0.005
	(0.017)	(0.005)	(0.004)
Observations	2305	2305	2305
County FE	Yes	Yes	Yes
HH Baseline Controls	Yes	Yes	Yes

We examine in Table C2 whether treatment may have had other spillover effects. First, we examine effects on firm entry and exit.

As well we examine in Table possible spillover effects on quality markers and lab tested quality.

C.2 Effects of Baseline Surveys

We compare control sites (which received baseline surveys but no information campaign) with pure control sites (which received no baseline activities of any kind and were visited for the first time in March to April 2021).

For data from the market audit survey, we have data for the LGS2021 planting season for both control markets and pure control markets. Table C4 and **??** show no statistically significant differences in the numbers of sellers or in the prevalence of quality markers in control markets as compared to pure control markets.

Table C2: Seller Entry/Exit Spillovers (spillovers)

Panel A: Spillovers within 2km

	(1) Number of sellers at market
Number of treatment sites within 2km	-0.184
	(0.156)
Number of control sites within 2km	0.053
	(0.206)
Observations	371
County FE	Yes
HH Baseline Controls	

Panel B: Spillovers within 4km

	(1) Number of sellers at market
Number of treatment sites within 4km	-0.101
	(0.126)
Number of control sites within 4km	-0.072
	(0.077)
Observations	371
County FE	Yes
HH Baseline Controls	

Panel C: Spillovers within 6km

	(1) Number of sellers at market
Number of treatment sites within 6km	-0.057
	(0.067)
Number of control sites within 6km	-0.038
	(0.061)
Observations	371
County FE	Yes
HH Baseline Controls	

Table C3: Quality Markers and Lab Tests Spillovers (spillovers)

Panel A: Spillovers within 1km

	(1) purity	(2) germ_rate	(3) Has all quality markers	(4) purity	(5) germ_rate
Number of treatment sites within 2km	0.014	-1.372	-0.036	0.014	-1.372
	(0.019)	(1.301)	(0.050)	(0.019)	(1.301)
Number of control sites within 2km	-0.008	0.019	-0.016	-0.008	0.019
	(0.017)	(1.171)	(0.044)	(0.017)	(1.171)
Observations	484	484	662	484	484
County FE	Yes	Yes	Yes	Yes	Yes
HH Baseline Controls					

Panel B: Spillovers within 4km

	(1) purity	(2) germ_rate	(3) Has all quality markers	(4) purity	(5) germ_rate
Number of treatment sites within 4km	-0.009	-0.060	-0.017	-0.009	-0.060
	(0.009)	(0.645)	(0.024)	(0.009)	(0.645)
Number of control sites within 4km	0.003	0.080	0.005	0.003	0.080
	(0.008)	(0.557)	(0.021)	(0.008)	(0.557)
Observations	484	484	662	484	484
County FE	Yes	Yes	Yes	Yes	Yes
HH Baseline Controls					

Panel C: Spillovers within 6km

	(1) purity	(2) germ_rate	(3) Has all quality markers	(4) purity	(5) germ_rate
Number of treatment sites within 6km	-0.000	0.584	-0.021	-0.000	0.584
	(0.006)	(0.399)	(0.015)	(0.006)	(0.399)
Number of control sites within 6km	-0.001	0.472	0.006	-0.001	0.472
	(0.005)	(0.360)	(0.014)	(0.005)	(0.360)
Observations	484	484	662	484	484
County FE	Yes	Yes	Yes	Yes	Yes
HH Baseline Controls					

Table C4: Control vs Pure Control: LGS2021 Firm Entry/Exit

Panel A: Full Sample

	(1) Number of sellers at market
Baseline Surveys Administered	0.098
	(0.157)
Observations	195
County FE	Yes
Control Mean	1.771

Table C5: Control vs Pure Control: LGS2021 Quality Markers

	(1) Has all quality markers	(2) Has valid KEPHIS sticker	(3) Lot number present
Baseline Surveys Administered	-0.014	-0.025	-0.008
	(0.058)	(0.021)	(0.008)
Observations	258	254	265
County FE	Yes	Yes	Yes
Control Mean	0.616	0.966	0.994

	(1) purity	(2) germ_rate
Baseline Surveys Administered	0.054	3.154*
	(0.039)	(1.818)
Observations	165	165
County FE	Yes	Yes
Control Mean	99.86	91.90

Table C6: Control vs Pure Control: LGS2021 Prices

	(1) Price paid for 2kg hybrid
gotbaseline	-2.635
	(7.227)
Observations	266
County FE	Yes
Control Mean	483.2

Appendix D: Extra Tables

Table D1: Treatment Delivery

	Flyers distributed	Residents trained directly
Avg number per site	173.5	81.7
% of local customers	19.7%	9.1%

Table D2: Baseline Balance

Variable	Control Mean	Treatment Mean	Difference
HH Head is Male	0.700	0.710	0.0100
			'(0.02)
HH Head Age	49.97	50.64	0.670
			'(0.64)
Completed Primary School	0.620	0.580	-0.04*
			'(0.02)
Completed Secondary School	0.270	0.230	-0.04**
			'(0.02)
Home quality index	1.440	1.440	0
			'(0.05)
Acres Planted (2019 main season)	1.220	1.180	-0.0400
			'(0.07)
Germination rate (2019 main season)	82.03	82.29	0.260
			'(0.94)
Hybrid maize yield	930.5	910.0	-20.44
			'(69.73)

Sample includes 2431 households surveyed at baseline. P-value for joint test for differences between treatment groups is 0.229. For all tests for differences between the treatment group and the control group, we control for county dummies (site selection was stratified by county) and calculate standard errors clustered at the sub-location level, which is the unit of randomization)

Table D3: Quality markers and seed purity

	Purity	Purity	Purity	Purity	Purity	Purity
Has all quality markers	0.01					
	(0.02)					
Has valid KEPHIS sticker		-0.04				
		(0.03)				
Lot number present			-0.03			
			(0.06)			
Has expiration date				0.04**		
				(0.02)		
Has valid expiration date					0.03*	
					(0.02)	
No holes						-0.05
						(0.06)
Constant	99.84***	99.88***	99.87***	99.81***	99.82***	99.89***
	(0.01)	(0.03)	(0.06)	(0.02)	(0.02)	(0.06)
Observations	467	464	467	468	468	468

Table D4: Quality markers and germination rate

	Germination Rate	Germination Rate	Germination Rate	Germination Rate	Germination Rate	Germination Rate
Has all quality markers	5.21***					
	(1.14)					
Has valid KEPHIS sticker		0.92				
		(2.11)				
Lot number present			-2.58			
			(4.37)			
Has expiration date				3.24**		
				(1.45)		
Has valid expiration date					6.33***	
					(1.20)	
No holes						-4.31
						(4.36)
Constant	88.08***	90.91***	94.29***	89.03***	86.96***	96.00***
	(0.96)	(2.03)	(4.34)	(1.33)	(1.05)	(4.33)
Observations	467	464	467	468	468	468

Table D5: Households knowledge (by gender of household head)

Panel A: Male-headed

	(1) no. quality marks known	(2) e-verif known	(3) expiration known
treated	0.241***	0.119***	0.029***
	(0.044)	(0.014)	(0.010)
Observations	3141	3141	3141
Control Mean	0.58	0.14	0.07
Treatment Effect (%)	41.33	85.96	41.57

Panel B: Female-headed

	(1) no. quality marks known	(2) e-verif known	(3) expiration known
treated	0.132**	0.070***	0.016
	(0.056)	(0.018)	(0.013)
Observations	1502	1502	1502
Control Mean	0.51	0.11	0.06
Treatment Effect (%)	25.90	62.88	26.23

Table D6: Households knowledge (by education)

Panel A: At Least Primary School Completed

	(1) no. quality marks known	(2) e-verif known	(3) expiration known
treated	0.290***	0.116***	0.033***
	(0.050)	(0.016)	(0.011)
Observations	2720	2720	2720
Control Mean	0.64	0.16	0.08
Treatment Effect (%)	45.35	72.02	39.74

Panel B: Primary School Not Completed

	(1) no. quality marks known	(2) e-verif known	(3) expiration known
treated	0.122***	0.097***	0.020*
	(0.045)	(0.015)	(0.010)
Observations	1923	1923	1923
Control Mean	0.43	0.08	0.05
Treatment Effect (%)	28.23	118.87	43.26

This table shows treatment effects on household knowledge about quality markers. The sample is split according to whether the household head has completed primary school.

Table D7: Household Seed Choice (by season)

Panel A: Long Growing Season 2020

	(1) Bought hybrid at local market	(2) Bought hybrid elsewhere	(3) Bought hybrid
treated	-0.064**	0.052*	-0.001
	(0.026)	(0.027)	(0.012)
Observations	1833	1882	1882
County FE	Yes	Yes	Yes
HH Baseline Controls	Yes	Yes	Yes
Control Mean	0.49	0.45	0.93
Treatment Effect (%)	-13.13	11.58	-0.15

Panel B: Short Growing Season 2020

	(1) Bought hybrid at local market	(2) Bought hybrid elsewhere	(3) Bought hybrid
treated	-0.053**	0.055***	0.000
	(0.021)	(0.017)	(0.023)
Observations	1884	1889	1889
County FE	Yes	Yes	Yes
HH Baseline Controls	Yes	Yes	Yes
Control Mean	0.17	0.10	0.29
Treatment Effect (%)	-30.60	52.41	0.01

Table D8: Household Seed Choice (by education)

Panel A: At Least Primary School Completed

	(1) Bought hybrid at local market	(2) Bought hybrid elsewhere	(3) Bought hybrid
treated	-0.040*	0.052***	0.016
	(0.022)	(0.020)	(0.015)
Observations	2247	2275	2275
County FE	Yes	Yes	Yes
HH Baseline Controls	Yes	Yes	Yes
Control Mean	0.31	0.29	0.60
Treatment Effect (%)	-13.15	17.95	2.70

Panel B: Primary School Not Completed

	(1) Bought hybrid at local market	(2) Bought hybrid elsewhere	(3) Bought hybrid
treated	-0.085***	0.054**	-0.027
	(0.028)	(0.027)	(0.022)
Observations	1470	1496	1496
County FE	Yes	Yes	Yes
HH Baseline Controls	Yes	Yes	Yes
Control Mean	0.36	0.25	0.62
Treatment Effect (%)	-23.61	21.60	-4.42

This table shows treatment effects on household seed choice. The sample is split according to whether the household head has completed primary school.

Table D9: Household agricultural outcomes (Remote Markets; By Gender)

Panel A: Male-headed Households

	(1) Germination rate	(2) Kgs harvested	(3) Yield
treated	-1.51	90.30	18.48
	(1.20)	(72.06)	(52.28)
Observations	757	748	696
County FE	Yes	Yes	Yes
HH Baseline Controls	Yes	Yes	Yes
Control Mean	84.92	1058.80	987.12
Treatment Effect (%)	-1.77	8.53	1.87

Panel B: Female-headed Households

	(1) Germination rate	(2) Kgs harvested	(3) Yield
treated	3.30	126.88	202.59**
	(2.02)	(117.79)	(83.50)
Observations	335	339	305
County FE	Yes	Yes	Yes
HH Baseline Controls	Yes	Yes	Yes
Control Mean	84.77	910.46	792.15
Treatment Effect (%)	3.89	13.94	25.57

Table D10: Seed quality vs gender and education

	(1) Germination rate
HH head is male	1.45*
	(0.82)
Constant	81.09***
	(0.67)
Observations	2317

	(1) Germination rate
Completed primary	2.40***
	(0.78)
Constant	80.65***
	(0.60)
Observations	2317

	(1) Germination rate
Above median plot size	-0.47
	(0.76)
Constant	82.31***
	(0.54)
Observations	2317

	(1) Can identify low quality
HH head is male	0.03*
	(0.02)
Constant	0.12***
	(0.01)
Observations	2415

	(1) Can identify low quality
Completed primary	0.05***
	(0.01)
Constant	0.11***
	(0.01)
Observations	2415

	(1) Can identify low quality
Above median plot size	0.01
	(0.01)
Constant	0.13***
	(0.01)
Observations	2415

Table D11: Household agricultural outcomes (By Season)

Panel A: Long Growing Season 2020

	(1) Germination rate	(2) Kgs harvested	(3) Yield
treated	0.02	21.06	18.57
	(0.76)	(50.03)	(32.68)
Observations	1748	1709	1631
County FE	Yes	Yes	Yes
HH Baseline Controls	Yes	Yes	Yes
Control Mean	84.91	1058.79	933.97
Treatment Effect (%)	0.03	1.99	1.99

Panel B: Short Growing Season 2020

	(1) Germination rate	(2) Kgs harvested	(3) Yield
treated	-0.99	21.79	41.99
	(1.35)	(24.11)	(72.56)
Observations	526	619	480
County FE	Yes	Yes	Yes
HH Baseline Controls	Yes	Yes	Yes
Control Mean	85.34	284.34	747.31
Treatment Effect (%)	-1.16	7.66	5.62

Table D12: Household agricultural outcomes (By Remoteness)

Panel A: Less Remote

	(1) Germination rate	(2) Kgs harvested	(3) Yield
treated	-0.95	-40.84	-41.36
	(1.08)	(56.17)	(46.33)
Observations	1182	1241	1110
County FE	Yes	Yes	Yes
HH Baseline Controls	Yes	Yes	Yes
Control Mean	85.10	745.45	866.14
Treatment Effect (%)	-1.11	-5.48	-4.78

Panel B: More Remote

	(1) Germination rate	(2) Kgs harvested	(3) Yield
treated	-0.13	104.61*	77.79*
	(0.96)	(61.21)	(42.52)
Observations	1092	1087	1001
County FE	Yes	Yes	Yes
HH Baseline Controls	Yes	Yes	Yes
Control Mean	84.88	1013.02	928.37
Treatment Effect (%)	-0.15	10.33	8.38

Table D13: Household agricultural outcomes (By Gender)

Panel A: Male-headed Households

	(1) Germination rate	(2) Kgs harvested	(3) Yield
treated	-0.92	36.15	-3.05
	(0.85)	(46.70)	(36.12)
Observations	1591	1617	1473
County FE	Yes	Yes	Yes
HH Baseline Controls	Yes	Yes	Yes
Control Mean	85.09	894.98	918.20
Treatment Effect (%)	-1.08	4.04	-0.33

Panel B: Female-headed Households

	(1) Germination rate	(2) Kgs harvested	(3) Yield
treated	1.31	41.30	88.61
	(1.34)	(68.18)	(57.52)
Observations	683	711	638
County FE	Yes	Yes	Yes
HH Baseline Controls	Yes	Yes	Yes
Control Mean	84.83	747.69	826.87
Treatment Effect (%)	1.54	5.52	10.72

Table D14: Household agricultural outcomes (By Education)

Panel A: At Least Primary School Completed

	(1) Germination rate	(2) Kgs harvested	(3) Yield
treated	0.22	53.46	13.86
	(0.79)	(58.50)	(39.46)
Observations	1380	1395	1276
County FE	Yes	Yes	Yes
HH Baseline Controls	Yes	Yes	Yes
Control Mean	84.66	1010.47	969.79
Treatment Effect (%)	0.27	5.29	1.43

Panel B: Primary School Not Completed

	(1) Germination rate	(2) Kgs harvested	(3) Yield
treated	-1.04	-0.09	36.02
	(1.26)	(51.81)	(43.15)
Observations	894	933	835
County FE	Yes	Yes	Yes
HH Baseline Controls	Yes	Yes	Yes
Control Mean	85.57	592.48	761.76
Treatment Effect (%)	-1.21	-0.02	4.73

Table D15: Household agricultural outcomes (By Plot Size)

Panel A: Above median plot size

	(1) Germination rate	(2) Kgs harvested	(3) Yield
treated	0.36	62.43	52.33
	(0.98)	(68.86)	(40.57)
Observations	1166	1179	1067
County FE	Yes	Yes	Yes
HH Baseline Controls	Yes	Yes	Yes
Control Mean	84.64	1085.80	866.99
Treatment Effect (%)	0.42	5.75	6.04

Panel B: Below median plot size

	(1) Germination rate	(2) Kgs harvested	(3) Yield
treated	-0.80	-10.44	-7.61
	(0.94)	(38.55)	(43.09)
Observations	1108	1149	1044
County FE	Yes	Yes	Yes
HH Baseline Controls	Yes	Yes	Yes
Control Mean	85.41	603.07	916.01
Treatment Effect (%)	-0.94	-1.73	-0.83

Table D16: Household complementary inputs (LGS2020/SGS2020)

Panel A: Full sample

	(1)
	Used Fertilizer
treated	-0.01
	(0.01)
Observations	2433
County FE	Yes
HH Baseline Controls	Yes
Control Mean	0.94
Treatment Effect (%)	-0.94

Panel B: Less Remote

	(1)
	Used Fertilizer
treated	-0.02
	(0.02)
Observations	1285
County FE	Yes
HH Baseline Controls	Yes
Control Mean	0.95
Treatment Effect (%)	-2.19

Panel C: More Remote

	(1)
	Used Fertilizer
treated	0.00
	(0.02)
Observations	1148
County FE	Yes
HH Baseline Controls	Yes
Control Mean	0.93
Treatment Effect (%)	0.47

We split the sample by median distance from the county capital

Table D17: Seller Entry/Exit (LGS2020/SGS2020/LGS2021)

Panel A: Full Sample

	Number of sellers at market
treated	-0.179*
	(0.103)
Observations	719
County FE	Yes
Control Mean	1.43
Treatment Effect (%)	-12.52

Panel B: Less Remote

	Number of sellers at market
treated	-0.291*
	(0.149)
Observations	365
County FE	Yes
Control Mean	1.56
Treatment Effect (%)	-18.68

Panel C: More Remote

	Number of sellers at market
treated	-0.050
	(0.145)
Observations	354
County FE	Yes
Control Mean	1.25
Treatment Effect (%)	-4.00

All standard errors are clustered by sublocation

Table D18: Seller Entry/Exit (Surveyed Sellers Only; By Baseline Profit)

Panel A: Above median baseline profit

	(1) lgs_sold_seeds	(2) sgs_sold_seeds
treated	0.003	-0.038
	(0.034)	(0.047)
Observations	236	236
County FE	Yes	Yes
Control Mean	0.92	0.18
Treatment Effect (%)	0.31	-21.51

Panel B: Below median baseline profit

	(1) lgs_sold_seeds	(2) sgs_sold_seeds
treated	-0.041	-0.099*
	(0.038)	(0.056)
Observations	178	178
County FE	Yes	Yes
Control Mean	0.95	0.27
Treatment Effect (%)	-4.37	-37.24

Table D19: Household reporting

Panel A: Full Sample

	(1) report_yesno_lgssgs	(2) n_report_recip_lgssgs	(3) report_to_auth_lgssgs	(4) report_to_nei_lgssgs
treated	0.012**	0.033**	0.004	0.012***
	(0.006)	(0.014)	(0.003)	(0.004)
Observations	4493	4493	4493	4493
Control Mean	0.04	0.07	0.01	0.01
Treatment Effect (%)	34.59	50.59	37.79	83.72

Panel B: Less Remote

	(1) report_yesno_lgssgs	(2) n_report_recip_lgssgs	(3) report_to_auth_lgssgs	(4) report_to_nei_lgssgs
treated	0.011	0.019	0.013**	0.003
	(0.008)	(0.019)	(0.005)	(0.006)
Observations	2248	2248	2248	2248
Control Mean	0.03	0.07	0.01	0.02
Treatment Effect (%)	31.01	27.10	124.58	17.91

Panel C: More Remote

	(1) report_yesno_lgssgs	(2) n_report_recip_lgssgs	(3) report_to_auth_lgssgs	(4) report_to_nei_lgssgs
treated	0.013	0.046**	-0.003	0.021***
	(0.009)	(0.020)	(0.004)	(0.006)
Observations	2245	2245	2245	2245
Control Mean	0.04	0.06	0.01	0.01
Treatment Effect (%)	36.29	73.17	-28.57	182.76

Table D20: Seller Entry/Exit (By Type; SGS2020/LGS2021)

Panel A: Full Sample

	(1) agrovet_or_stockist	(2) duka	(3) offer_otherseeds	(4) offer_otherinputs	(5) offer_hhgoods
treated	-0.047	0.100**	-0.034	-0.009	-0.024
	(0.052)	(0.050)	(0.056)	(0.043)	(0.054)
Observations	447	447	447	447	447
County FE	Yes	Yes	Yes	Yes	Yes
Control Mean	0.60	0.29	0.55	0.78	0.38
Treatment Effect (%)	-7.83	33.88	-6.07	-1.08	-6.23

Panel B: Less Remote

	(1) agrovet_or_stockist	(2) duka	(3) offer_otherseeds	(4) offer_otherinputs	(5) offer_hhgoods
treated	-0.102	0.145*	-0.073	-0.016	0.013
	(0.075)	(0.075)	(0.078)	(0.059)	(0.075)
Observations	243	243	243	243	243
County FE	Yes	Yes	Yes	Yes	Yes
Control Mean	0.65	0.25	0.58	0.82	0.34
Treatment Effect (%)	-15.81	57.16	-12.50	-1.96	3.74

Panel C: More Remote

	(1) agrovet_or_stockist	(2) duka	(3) offer_otherseeds	(4) offer_otherinputs	(5) offer_hhgoods
treated	0.044	0.032	0.018	0.023	-0.094
	(0.073)	(0.070)	(0.086)	(0.073)	(0.082)
Observations	204	204	204	204	204
County FE	Yes	Yes	Yes	Yes	Yes
Control Mean	0.53	0.37	0.51	0.73	0.45
Treatment Effect (%)	8.40	8.70	3.52	3.17	-20.86

Table D21: Prices for hybrid maize seed (household self-reported)

Panel A: Full sample

	(1) Avg price per kg	(2) Avg price per kg from local market	(3) Avg price per kg from elsewhere
treated	1.782	0.265	2.816
	(1.917)	(2.128)	(2.454)
Observations	2110	1105	974
County FE	Yes	Yes	Yes
HH Baseline Controls	Yes	Yes	Yes
Control Mean	220.22	224.77	219.80
Treatment Effect (%)	0.81	0.12	1.28

Panel B: Less remote

	(1) Avg price per kg	(2) Avg price per kg from local market	(3) Avg price per kg from elsewhere
treated	2.666	3.569	1.390
	(2.840)	(3.167)	(3.793)
Observations	1102	601	498
County FE	Yes	Yes	Yes
HH Baseline Controls	Yes	Yes	Yes
Control Mean	223.89	228.20	222.59
Treatment Effect (%)	1.19	1.56	0.62

Panel C: More remote

	(1) Avg price per kg	(2) Avg price per kg from local market	(3) Avg price per kg from elsewhere
treated	0.273	-3.319	2.605
	(2.358)	(2.921)	(3.414)
Observations	1008	504	476
County FE	Yes	Yes	Yes
HH Baseline Controls	Yes	Yes	Yes
Control Mean	214.73	219.39	215.68
Treatment Effect (%)	0.13	-1.51	1.21

Table D22: Price dispersion for hybrid maize seed (household self-reported)

Panel A: Full sample

	(1) purity	(2) germ_rate	(3) sd_price_hybridS_h	(4) rng_price_hybridS_h
treated			0.299	-1.900
			(2.075)	(4.498)
Observations	165	165	198	222
County FE	Yes	Yes	Yes	Yes
HH Baseline Controls			Yes	Yes
Control Mean	99.86	91.90	15.15	31.55
Treatment Effect (%)			1.98	-6.02

Panel B: Less remote

	(1) sd_price_hybridS_h	(2) rng_price_hybridS_h
treated	2.369	9.814
	(2.753)	(6.554)
Observations	93	105
County FE	Yes	Yes
HH Baseline Controls	Yes	Yes
Control Mean	15.07	31.04
Treatment Effect (%)	15.72	31.61

Panel C: More remote

	(1) sd_price_hybridS_h	(2) rng_price_hybridS_h
treated	-0.109	-9.101
	(2.940)	(6.224)
Observations	105	117
County FE	Yes	Yes
HH Baseline Controls	Yes	Yes
Control Mean	15.23	32.15
Treatment Effect (%)	-0.72	-28.30

Table D23: Prices for hybrid maize seed (from market audit; All Seasons)

Panel A: Full sample

	(1)
	Price paid for 2kg hybrid
treated	0.091
	(5.102)
Observations	693
County FE	Yes
HH Baseline Controls	Yes
Control Mean	480.30
Treatment Effect (%)	0.02

Panel B: Less Remote

	(1)
	Price paid for 2kg hybrid
treated	2.115
	(6.867)
Observations	390
County FE	Yes
HH Baseline Controls	Yes
Control Mean	479.15
Treatment Effect (%)	0.44

Panel C: More Remote

	(1)
	Price paid for 2kg hybrid
treated	-2.009
	(7.118)
Observations	303
County FE	Yes
HH Baseline Controls	Yes
Control Mean	482.34
Treatment Effect (%)	-0.42

All standard errors are clustered by sublocation. We examine effects on market areas separately in the LGS2020, SGS2020 and LGS2021 planting seasons. We examine effects on above and below median baseline population; below and above median distance to the nearest major market. Keeping in mind that lower-population areas had larger shifts in the market equilibrium, and that sellers may have taken at least one season to shift their entry/exit decisions, we also examine the subset of markets with below median population during LGS2020 and during SGS2020 and LGS2021