Market Access and Quality Upgrading: Evidence from Three Field Experiments*

January 2021

Tessa Bold,[1] Selene Ghisolfi,[2] Frances Nsonzi,[3] Jakob Svensson[4]

Abstract

Smallholder farming in many developing countries is characterized by low productivity and low quality produce. Low quality limits the price farmers can command and thus their potential income from farming. We conduct a series of measurement and field experiments among smallholder maize farmers in western Uganda to shed light on the barriers to quality upgrading at the farm level and to study its potential in raising productivity and rural incomes. First, we measure maize quality at the farm gate and in the lab and show that quality is low and at least partly observable at the farm gate. Second, we generate exogenous variation in the quality of the maize farmers sell to local markets. The causal return to quality is zero; that is, the market for quality maize is effectively missing. Third, we generate experimental variation in access to a market for premium quality maize. Over time, the majority of treatment farmers sold maize of high quality. Profit from maize farming in the treatment group increased by 40-80%; an effect driven both by increased productivity and higher prices for both premium and lower quality maize in treatment villages. Our findings reveal the importance of demand-side constraints in limiting rural income and productivity growth.

* Financial support from the Swedish Research Council (2017-05665) and Riksbankens Jubileumsfond is greatly appreciated.

[1] IIES, Stockholm University, Stockholm, Sweden; CEPR (tessa.bold@iies.su.se).
[2] Leap, Bocconi University, Italy (ghisolfi@unibocconi.it).
[3] Department of Food Technology and Nutrition, Makerere University, Kampala, Uganda; PNDK Solutions, Kampala, Uganda (nsonzif@gmail.com).
[4] IIES, Stockholm University, Stockholm, Sweden; BREAD; CEPR (jakob.svensson@iies.su.se).

1. Introduction

Smallholder farmers in low income countries produce and sell output of low quality. Low quality limits the price farmers can command and can help explain why the returns to smallholder farming are low. At the same time many experts and policy makers argue quality-upgrading is key to raising income and productivity and the World Bank promises `double dividends' to poorer countries that participate in global value chains (World Bank, 2020). Yet, few farmers upgrade the quality of their produce. This, in turn, suggests that either the link between quality upgrading and higher income is not as strong as hypothesized or strong constraints – possibly both on the demand and the supply side – trap farmers in a low quality-low productivity equilibrium.

This paper conducts three experiments among smallholder maize farmers in western Uganda to shed light on the impediments to quality upgrading at the farm level and to study its potential. We proceed in three steps. In the first experiment, we measure the quality of maize sold at the farm gate and confirm that it is poor. The low quality of the final product sold to consumers can thus be traced back all the way to the farmer. We also show that maize quality measured through simple and quick tests at the farm gate strongly predicts maize quality measured by more elaborate laboratory tests. Maize quality is thus, at least partly, observable. In the second experiment, we randomly assign a sample of farmers into two groups, and offer treatment farmers a service package to improve the quality of the maize they sell. We find no return to selling high quality maize: traders pay the same price for high and low quality maize.

Being unable to access a market that rewards high quality may thus explain why farmers are loath to incur the extra cost and effort to increase quality. To test this hypothesis, we conduct a third experiment to investigate if and how farmers respond when offered access to a market where quality maize is paid a (market) premium. Following farmers over four years, including for four post-treatment seasons, we find that a majority of farmers, when offered the opportunity, sell maize of high quality. Treatment farmers' profit increases substantially – an effect driven both by increased farm productivity and higher prices for both premium and lower quality maize in treatment villages.

The market access intervention was designed in close collaboration with a Ugandan vertically integrated agro trading company to emulate a situation where treated households gain access to an output market for quality maize. The company committed to buy quality maize at a premium throughout the main buying seasons in treatment villages, with the

premium determined by the difference in the amount of waste and defected grains in high versus low quality maize, valued at high quality prices. Because the market access intervention was randomized at the village level, and because not all farmers sold maize to the agro trading company, we can in addition to the intervention's direct effect also study its effect on the local market equilibrium. We find that the entry of the high-quality buyer resulted in an increase in prices received also for farmers who continued to sell to incumbent local traders. This effect raised income for farmers continuing to sell average quality maize and thus reduced the relative price of higher quality maize, weakening the incentives for quality upgrading. Adjusting for selection, the local market equilibrium effects can account for one-third of the increase in average prices in the treatment relative to the control group.

To ensure that farmers had up-to-date knowledge about pre- and post-harvest practices necessary for producing maize of sufficiently high quality, the market access intervention was combined with a learning-by-doing extension service component. In a parallel trial (Bold et al., 2020), we investigate the impact of this supply-side intervention alone. We find little evidence that farmers changed their farm practices as a result of the training program: revenue, expenses, and yield remain essentially the same in the treatment and the control group. The results from the two trials thus suggest that market access is necessary and critical for agricultural transformation, but also that it is, in some respects, not sufficient: while the use of modern technologies (hybrid seeds and inorganic fertilizer) increased significantly in the market access intervention group, modern input use across all groups in the two trials remained low. Therefore, we conclude more broadly that transformative change in small scale agriculture likely requires relaxing both binding supply and demand/market constraints.

Our results relate to a number of recent papers on the implications (for farmers or firms) of market (buyer) driven quality upgrading in a developing country setting. As in Atkin et al. (2017), we exploit experimental variation in access to a market/buyer of quality products. Their intervention, which connects rugs producers in Egypt to foreign buyers paying a premium for higher quality rugs, led to large improvements in both quality and productivity. We focus on poor smallholders working with little or no physical capital. Further, we worked closely with an agro trading company that makes direct outreach at the village level, and constrain the firm's actions in several ways so as to mimic a situation in which farmers have access to a local market for quality. Our village level intervention allows us not only to assess the direct effect of the intervention, but also to study how local traders that are the dominating buyers at the village level responded to the entry of a high-quality

buyer. Finally, we assess the implications of having access to a market for quality over a longer period.

Our paper is also related to, and complements, Macchiavello and Miquel-Florensa (2019). They employ a difference-in-difference strategy to estimate the impact of a quality upgrading program in Colombia and find that eligible farmers increased quality and received higher farm gate prices. While the intervention we exploit here also involves a vertically integrated domestic buyer – in our case a firm who buys quality maize at the farm gate and sells high quality maize flour in Kampala – the intervention, at the farm gate, was designed for research purposes. This enables us to directly measure and identify smallholders' choices and outcomes using experimental variation, including the costs of quality upgrading, in response to getting access to an output market for quality maize, and to decompose the treatment effect on output into components attributable to changes in the inputs that we can measure and to increases in total factor productivity. By comparing outcomes from the market intervention with the extension-service only intervention, we can also, at least partly, unbundle the impact of demand and supply (extension service) factors.[5]

Knowledge about the pre-conditions and determinants for agricultural technology adoption has grown vastly over the last decade (for a review, see de Janvry, et al., 2017). The evidence – drawn primarily from randomized controlled trials in Sub-Saharan Africa and South Asia – makes clear that there are productivity-enhancing supply side innovations available today that can increase technology adoption and productivity among smallholders. Measured effects on farmer income, however, have been much more limited.[6] We add to this literature by studying the impact of lifting both a demand (inclusion in a value chain) and a supply (knowledge) constraint. De Janvry and Sadoulet (2019) discuss the complexity of rigorously evaluating value chain inclusion, noting the double challenge of having to implement treatment at the community level – making the intervention costly – and the difficulty in finding implementing partners who are willing to expand their business in a way

[5] Our findings also relate, albeit less directly, to a large literature on the effects of quality upgrading on productivity and growth – often through exporting higher quality goods – although the buyer in our experiment is not an exporter. Important theoretical contributions in this literature are Hausman et al. (2007), who emphasize the link between specialization in exporting high quality and subsequent higher growth. Empirical results (largely non-experimental) are reviewed, for example, in De Loecker and Goldberg (2014). Ashraf, Gine and Karlan (2009) experimentally evaluated an intervention in Kenya that helped farmers to adopt and market export crops. The authors find small effects on adoption and income, but a 32% income gain for adopters. More broadly, our results also relate to a large (mainly non-experimental) literature on the effects of market access and market integration reviewed in Donaldson (2015).

[6] One exemption is the evaluation of the One Acre Fund's small farmer program – a bundled program where participating farmers receive training on improved farming practices, input loans, and crop insurance. Deutschmann, et al. (2019) document an increase in profit of 8%-16% after one season.

that is amenable to rigorous evaluation. We overcome these problems here, and document widespread increases in adoption (here in terms of quality upgrading) over time and a significant increase in farm profit.

Our findings should be interpreted within the context of the study – a poor area of western Uganda. All farmers participating in the experiments are smallholders and farm largely with traditional methods. This context can help explain why farmers raised yields predominantly by working more and producing a higher output per hour worked, rather than applying modern inputs, such as improved seeds and fertilizer to their fields. The poor rural setting may also have contributed to the large increase in farmers' income: by targeting the poorest farmers who are currently entirely excluded from global value chains, the agro-trading company provided market access to precisely those farmers who had the largest potential to benefit.

To implement the intervention and to circumvent potential agency- and information problems, the collaborating company ran a vertically integrated operation. As quality upgrading is potentially a core motive for why firms change their organizational structure, the intervention provides a case study of the constraints of buying, processing, and selling quality maize for a vertically integrated domestic buyer.[7] It does not, however, allow us to study the behavior of the layers of intermediaries that dominate the low-quality segment of the market.[8]

2. Context

2.1. Introduction

Uganda remains highly dependent on agriculture. It is estimated that the sector contributes over 70 percent of export income and 65 percent of the population is active in the sector. As

[7] The decision to work with a vertically integrated buyer to purchase high quality maize grain is reminiscent of the work by Hansman et al. (2020), who study vertical integration as a response to a higher quality premium in the context of the Peruvian fishmeal industry, as well as a theoretical literature that emphasizes the relationship between the nature of output and the boundary of the firm (Baker et al, 2001, Gibbons 2005a,b).

[8] There is a complementary literature examining the structure and competitiveness of intermediaries both as buyers and sellers. Regarding the latter, Bergquist and Dinerstein (2020) show in an experimental study in Kenya that cost-pass through from traders to consumers is very limited and conclude that trader behavior is indistinguishable from a fully collusive model. This finding is echoed by Atkin and Donaldson (2015) who structurally estimate the rate of cost-pass through from world prices in Ethiopia and Nigeria and conclude that the majority of surplus increases due to falling world prices are captured by intermediaries and not passed on to consumers. Regarding intermediaries as buyers, Casaburi and Reed (2015) estimate in the context of Sierra Leonean cocoa producers that the local markets in which traders operate to buy farmer cocoa are highly competitive with a low differentiation parameter. Dillon and Dambro (2017) come to a similar conclusion for a wide variety of agricultural markets in their review.

in most countries in the region, the agricultural sector is dominated by smallholder farmers, a majority of whom cultivate less than two hectares.

Maize is the most important cereal crop and grown primarily as a cash crop.[9] Smallholder farmers account for roughly 75 percent of maize production and 70 percent of marketable surplus.

Maize has different end uses depending on the geographic regions of the producers (Daly, et al. 2016). In the US, for example, only 12% of maize produced is used for human consumption, with the remainder split between animal feed and ethanol fuel production (Ranum, 2014). In Africa, and especially in East-Africa, maize is a staple food crop, accounting for nearly half of the calories and protein consumed (Macauley, 2015). While maize grain of the lowest quality is also used for animal feed, the main feed ingredient sold in markets is maize bran – a byproduct of flour production or grits manufactured from maize grain.

Maize, as most other crops in Uganda, is produced using mainly traditional techniques and few farmers use modern inputs such as hybrid seeds and fertilizer. Yields tend to be low. For example, Bold et al (2017), using four waves of LSMS data for Uganda, report an average yield (metric tons per hectare) for smallholders of 1.4. As a comparison, average maize yield based on data from farm demonstrations in Uganda (with recommended crop management and modern inputs) is over 4 tons per hectare (World Bank 2007) and average corn yield in the U.S. was close to 12 tons per hectare in 2017 (USDA, 2019). There is also concern about the quality of the maize (and other crops) bought and sold in Uganda, though there is little systematic measurement of quality.[10]

The research program is set in an area of western Uganda (Kakumiro and Kibaale districts), where smallholder maize farming is common. Rural Kibaale is poor, with an average consumption expenditure at 0.80 USD per day (UBOS, 2019).[11]

2.2. Local markets

[9] According to official statistics, maize exports accounted for about 2% of the country's total exports (Uganda Bureau of Statistics, 2015). Based on interviews with stakeholders in the sectors, Daly et al. (2016) estimate that 70-80% of maize that is bought and sold in Uganda is channeled through informal channels. To account for the size of the informal market, previous surveys have used multipliers of between 3-3.5 for formal trade data (Gates Foundation, 2014).

[10] For the low quality of maize and coffee in Uganda, see Daly et al. (2016), Gates Foundation (2014), and Morjaria and Sprott (2018), respectively.

[11] The district of Kakumiro was created in 2016 from the split of Kibaale district and separate statistics are not available.

The local, or village, output market for smallholders can be described as a spot market.[12] The farmer and the buyer agree right before the sale, usually after a short visual inspection of the bags by the buyer, about the amount and the price. The farmer is paid directly and the transaction takes place at the farm gate.

There are two types of buyers active in these local markets: (i) local traders or aggregators who often buy from a smaller set of farmers and resell to commercial traders/aggregators that are either passing through the village or located in a nearby trading center, and (ii) commercial buyers, who pass through the village with a truck, and buy directly from individual farmers (and local traders).

Over the five season for which we collected detailed sales data, 80% of the sales and 78% of the sales volume went to local traders (see Table A1 in supplementary appendix). Still, more than half the smallholders in the sample have sold to a commercial trader at least once during the last five seasons. A sale to a commercial trader fetches, on average, an 8% higher prices than a sale to a local trader.

Farmers tend to know the local traders they sell to and repeated transactions across seasons are common. 98% of the smallholders sold to the same buyer in at least two out of the last five seasons and 31% sold to the same buyer in at least four out of five seasons. Most households sell once per season (79%) and accounting for multiple sales to the same buyer in a given season, 90% sell to only one trader per season.

We collected data on market prices from the nearest trading center for each village in the sample (altogether five trading centers). Figure 1 depicts the results. Prices tend to increase (slowly) between harvest seasons, while there is substantial idiosyncratic variation in market prices within harvest seasons. The trading center price follows a similar pattern as the prices in the main wholesale markets in Kampala and Nairobi, with the average trading center price 16% lower than the price in Kampala, which in turn is 30% below the price in Nairobi.

3. Sample Frames and Research Design: Overview

We combine field experiments, and maize quality measurement using laboratory techniques and visual inspections to answer three questions. What maize quality is produced and sold at the farm gate? Is quality rewarded in local markets? Will farmers respond by producing higher quality if offered access to a market where quality maize is paid a (market) premium?

[12] We use data collected from the control group in the experiments discussed below to describe the local market context.

Figure A1 in the appendix illustrates the design of the study, including an overview of the sample frames, and the timing of the surveys and interventions. We draw on data from three sample frames. For Sample Frame 1, we first selected 20 communities (villages), each at least 5 kilometers apart, from digital maps of Kakumiro district in western Uganda. For each of the selected villages, we completed a census and identified smallholder farmer households (with maize gardens of no more than 5 acres of land) who cultivated maize in the previous season. We used the same approach to identify villages and households in the neighboring district Kibaale (Sample Frame 2). Finally, we identified 20 of the closest nearby villages to a subset of villages in Sample Frame 1 and identified smallholder maize farmers in all these villages. This set of villages constitutes Sample Frame 3.

We use data collected from households selected from Sample Frame 3 to measure the quality of maize in local markets. We also use this sample to estimate the causal return to quality-upgrading in local markets. We use Sample Frame 1 to study our main intervention, which combined access to an output market where quality maize is paid a (market) premium with a learning-by-doing extension service component. Finally, to learn about complementarities between demand and supply side constraints, we compare the market-cum-extension experiment with an additional trial focused solely on the impact of the extension service component. The extension service intervention was offered to a randomly selected set of villages from Sample Frame 2. We refer to some of the findings from this experiment here. The full results of the extension service experiment are reported in Bold et al. (2020).

4. Maize Quality and Verifiability of Quality

4.1. Introduction

Maize is sold and handled in large quantities, with the smallest unit typically a 100 kilogram bag. A bag of maize is considered high quality if it contains sufficiently large and dry maize kernels of the right color and neither non-grain substances (e.g. stones, dirt, and insects) nor defective (e.g. broken, immature, damaged, rotten, or moldy) grain. More formally, maize quality in East Africa is classified according to the East African Grading Standard (EAS) (East African Community, 2011), which divides maize into three broad quality categories based on moisture level and amount of non-grain substances and defective grain: graded maize, under-grade maize and reject maize. Graded maize (quality maize) is further categorized into three grades: grades 1, 2 and 3, with grade 1 having the most stringent thresholds for defects.

The quality of maize determines its potential economic and nutritional value and whether it is safe for human consumption. The presence of non-grain substances and defective grain adds to the weight of the bag without adding value and increases processing costs. Non-grain substances and defective grain are also indicators that the maize has not been properly handled. For example, stones and dirt in the bags indicate that the farmer has stored or dried the maize directly on the ground, raising the risk that grains are contaminated by microorganisms such as bacteria and fungi. Insect parts or insect waste, pest damaged, rotten, diseased, and discolored grains, are direct indicators of (acute) infestation. A particular concern is contamination with aflatoxin, poisonous carcinogens that are produced by certain mold species which inhabit the soil.[13] Contamination and infestation can spread quickly through the bag and while waste and defective maize kernels can be sorted and cleaned at a later stage in the value chain, there is an elevated risk that the remaining grain is (already) contaminated.

The moisture content influences expected maize quality through the same two channels: by increasing gross (but not dried) maize weight and by raising the risk of infestation. Dry grains keep longer, are attacked by insects less often, and usually do not support mold growth. In wet grains, on the other hand, fungal growth and release of mycotoxins can occur quickly, especially during storage. Aflatoxin contamination can increase ten-fold in just a few days if maize grain is not dried properly (Hell, at al., 2008).

An effective quality control strategy thus requires the prevention of defects and moisture in the maize as early as possibly in the value chain. Farmers play an important role in quality control and can influence maize quality by harvesting at the right time, shelling the cob without breaking or cracking the grains, not drying or storing cobs on the bare ground, and cleaning and storing the grain correctly.

The different components of maize quality can, in principle, be tested. As the full EAS testing protocol requires maize to be tested in the lab, however, it is rarely applied in the informal market. Instead, quality testing, when performed, relies on inspections of the bags of maize grains for the presence of various defects and measuring of moisture using either subjective tests (e.g. the crush test) or portable grain moisture meters.

[13] There is a large literature testing for the presence of aflatoxin in crops like maize. For a recent discussion of the literature on aflatoxin and health, see de Almeida et al. (2019). For a recent summary of evidence of aflatoxin measurement in Uganda, see Sserumaga et al. (2020). For research on the consequence of the unobservability of aflatoxin, see Hoffmann, et al. (2013).

4.2. Quality Measurement

To measure the quality of maize sold by farmers, and the extent to which quality can be verified at the farm gate, we performed four tests: visual inspection and moisture measurement at the farm gate, and laboratory testing and aflatoxin measurement in Kampala.[14] For the measurement experiment, we selected 10 villages from Sample Frame 3 and between 5-15 households in each village. We enrolled 100 farmers that were about to harvest their maize and assigned half of them to a treatment group and half of them to a control group. This sample forms the basis for the returns to quality experiment discussed in the next section. Here, we focus on the subset of control households.

To measure farm gate quality, trained enumerators visited each farmer at the time of sale and recorded, based on visual inspections, the presence (or absence) of 10 types of defects in each bag put for sale. We denote the mean number of those defects (in a given bag) as "Visually verifiable defects".[15]

After inspection, one bag per farmer was randomly selected and bought from the household and transported to a laboratory in Kampala for testing.[16] Samples of 200g were drawn from each bag, and the weight of all non-grain substances and defective grain recorded. The total weight of the defects, expressed as percent of the sample weight, is denoted "Lab verified defects".

To test for aflatoxins, we randomly sampled an additional 30 households from six Sample Frame 3 villages over two consecutive seasons. In this sample, at the time of sale, one bag per farmer was randomly selected and bought from the household. The purchased bags were brought to Kampala for lab measurement of defects, as described above, and also tested for aflatoxin.[17]

In both samples, the field enumerators measured moisture levels in the bags destined for the lab using a mobile moisture meter. We generate a binary indicator labeled "wet maize", taking the value 1 for maize bags with a moisture content in excess of 13% and zero otherwise.

[14] Details on the test protocol are in supplementary appendix S3.
[15] Enumerators verified and recorded whether the maize in the bag was dirty, included cobs, stones, dust, insects (live or dead), and broken, immature, damaged, rotten, and mold-infested grain.
[16] The testing protocol followed the EAS approved objective test methods for defects. 44 of the bags were tested in the lab (see appendix S3).
[17] AflaCheck test kit (VICAM) was used to detect the presence of aflatoxin. The test strips can detect aflatoxin at two different cutoff levels depending on the protocol followed. We used the 10 ppb (parts per billion) cutoff level, which is the limit imposed by the Uganda National Bureau of Standards (UNBS). As a reference, the European Union standard is 4 ppb (or ng/g) and the US standard is 20 ppb (Sserumaga et al, 2020).

Finally, we combined the visually verified defects and the moisture measure to classify all samples tested in the lab using the East African Quality Standard.

4.3. Results: Quality of maize at the farm gate

The quality of the maize farmers sell is low (see Table 1). The average bag inspected at the farm gate contained 2.5 defects (out of 10 possible) and the maize samples tested in the lab contained on average 26% defects, i.e., a quarter of the weight of maize sold consists of defective grains and waste. The distributions are depicted in Figure 2, panels A and B. For lab verified defects, the distribution is skewed with about 1 in 10 bags tested containing only defective grains (primarily immature or discolored). 28% of the households sold maize with a moisture content higher than 13%, with an average moisture content of wet maize of 14.7%.

The results for grading the lab samples according to the EAS classification are reported in the last four rows of Table 1. None of the bags contained grade 1 grain (the highest grade), 4% of the bags contained maize of grade 2 quality and 11% contained maize of grade 3 quality. The remaining bags, 85%, contained ungraded or rejected maize.

Quality measured at the farm gate predicts quality measured in the lab, as shown in Figure 2, panel C (the corresponding regression is reported in Table A2 in appendix), especially at higher levels of defects. When the number of defects found in the bag increases from 0 to 2, the percentage of waste in the lab sample increases from 10 to 15%. As the number of defects doubles from 2 to 4, the percentage of waste in the lab sample also doubles.

Lab measured defects, in turn, predict whether the sample contains dangerous levels of aflatoxin. Figure 2, panel D (specifications (3) in Table A2) plots the predicted probability of aflatoxin levels exceeding the limit imposed by the Uganda National Bureau of Standards. The relationship is roughly log-linear: as the share of the sample that is lost to waste and defects doubles, so does the predicted probability that aflatoxin levels are too high (> 10 ppb).

In sum, smallholders tend to sell maize of low and possibly unsafe quality. While testing for quality in the lab is costly, visual assessments that are easy and quick to execute at the farm gate can provide a proxy measure.

5. Returns to quality experiment

5.1. Introduction

If the economic value of maize depends on its quality, why is the quality of maize sold by farmers so low? A starting point to answer this question is the neoclassical agriculture household model. In this model, with complete markets, the production decision is separable from the consumption decision. Thus, a utility maximizing farmer chooses a vector of inputs to maximize profit. Consider a version with two inputs, x and z, with unit costs c_x and c_z, where input (or input vector) x primarily affects the quantity of output while input (or input vector) z primarily affects the quality (and potentially the price). The farmer's problem can be stated as:

$$\max_{x,z} \Pi = p(q(z))F(x) - c_x x - c_z z \ , \tag{1}$$

where $p(q)$ is the price as a function of quality q, $F(x)$ is output (assuming that land is a fixed factor). The farmer's choice of inputs is given by two first-order conditions:

$$p(.)F'(x) - c_x \leq 0 \ \text{ and } \ p'(.)q'(z)F(x) - c_z \leq 0 \ . \tag{2}$$

That is, the farmer will set the intensity of use of any particular input until its marginal value product equals its marginal cost. Thus price, or more precisely, the responsiveness of price to quality, is a key driver of the decision to produce high (or low) quality maize. But does the (local) market reward quality? To answer this question, we designed an experiment to measure the returns to quality.

5.2. Intervention

Low quality at the farm gate is determined by a number of factors, several of which the farmer can influence through good agronomic practices in harvesting, decobbing, drying, cleaning and storing the grain. To create random variation in the quality of maize farmers sell, we therefore developed a service package, which included assistance with several key harvest and post-harvest (drying, winnowing, and sorting) activities. The services offered were implemented by agricultural workers with access to portable agricultural machinery (tarpaulin for drying and a sheller/decobber), and were managed by staff from the research team.

5.3. Experimental design and data

We attempted to enroll 100 maize farming household from Sample Frame 3, 99 of whom gave consent. After stratifying by village, the households were randomly assigned into treatment (49 households) and control (50 households) groups. At enrollment, a short survey was administered. Table A3 in the Supplementary appendix compares pre-harvest outcomes

between treatment and control groups. None of the collected covariates show statistically significant differences across assignment groups and a joint balance test fails to reject the null hypothesis that the pre-harvest outcomes do not predict the assignment to treatment.

Before harvest, farmers in both groups were visited by staff from the research team. In the treatment group, households were offered the free service package. The offers were presented as a service from the research team. Compliance was 100%; i.e., all treatment households accepted the offer. The households were also asked to contact the research team at the time of bagging the maize but before selling it, and were promised a reward of UGX 10,000 (approximately USD 3) if they did so. Farmers in the comparison group were also visited and offered a (larger) monetary reward (UGX 30,000; approximately USD 9), if they contacted the research team before selling their maize.

When the farmer was ready to sell, they were visited again, this time by trained enumerators who visually verified the presence (or absence) of defects in all bags the farmer was planning to sell (see section 4.2. and section S3 in Supplementary appendix for a more detailed description of the measurement). The enumerators also weighed the bags and tested the moisture level using a mobile moisture meter. Altogether, 622 bags were visually inspected. In addition, one bag, drawn at random, was bought from each farmer for further quality analysis in the lab.[18] After selling their maize, all farmers were visited for a third and final time and asked about sales volume and prices. Four of the 98 farmers selling maize could not be reached and one farmer decided not to sell. In total, we collected data on 116 sales from 94 households.

5.4. Results: returns to quality

The free service package successfully raised quality in the treatment group. Figure 3 illustrates the results and Table 2 provides summary statistics. The average sale in the control group contained 2.5 defects per bag (maximum 10 defects), ranging from 0 to 7 defects. 86% of the bags contained at least one defect. In the treatment group, the average sale contained less than 0.05 defects per bag and only 4% of bags contained one or more defects. Observable quality differences between treatment and control are equally stark when averaging at the household level or focusing only on the randomly selected bag sent for lab testing.

The observed differences in defects between treatment and control were confirmed in the lab. One third of the content in the average bag in the control group consisted of defective

[18] As discussed in Supplementary appendix S5, one farmer did not sell any maize. In addition, 16 bags could not be tested due to administrative constraints.

maize grain and waste that should be removed before processing the maize. In the treatment group, the corresponding number was 6.5%. None of the bags purchased in the treatment group contained excessive moisture (above 13%), while 14% of the bags in the control group did.

In Panel D, Figure 3 and in the bottom rows of Table 2 we report the results of classifying the bags bought according to the EAS standard-maize grain classification system (see section 4.2). All bags tested in the treatment group were graded maize (13% grade 1, 72% grade 2, and 15% grade 3). In the control group, 56% were classified as ungraded/reject maize, and the remainder as either grade 2 (30%) or grade 3 quality (14%).

Despite the large differences in both visually verified and lab verified quality, however, buyers did not pay higher prices to farmers who had received the service package. Figure 4, Panel A, plots the CDFs of price in the two assignment groups. The two CDFs lie effectively on top of each other and the Kolmogorov-Smirnov test fails to reject the null that the two distributions are equal (D statistic $= 0.11$, $p = .84$).

These non-parametric results are confirmed by regression analysis. Table 3, specification (1), regresses price on the treatment indicator, controlling for week of sale fixed effects and randomization strata (villages). The unit of observation here is a sale (13% of households sold more than once during the season). Specification (2) reports the average treatment effect of quality, measured as the number of defects observed in the bag, on prices, using assignment to treatment as instrument. In both the OLS and the IV specification, the treatment effect is essentially zero, i.e., there is no evidence that higher maize quality – equivalent to fewer defects counted in the bag – systematically yields a higher price. The coefficient is also tightly estimated (95% CI: -11.5, 8.51). We can reject (at the 5%-level) that one additional defect reduces price per kilogram by more than UGX 11.5 shillings (Table 3, column 2). The extent of visually verifiable defects was, on average, 2.3 defects lower in the treatment group than in the control group. Such a reduction, with a point estimate of -11.5, would imply an increase in the price of maize (per kilogram) of UGX 23. This is a small increase for a large quality improvement – equivalent to less than a 3% price increase relative to the average price of maize in the control group.

Why do traders not pay higher prices for better quality maize? A first possible explanation for the absence of a relationship between farm gate quality and price is that traders simply cannot infer the true quality of maize from observable defects and hence do not adjust prices. In Section 4.3, however, we showed that there is a strong relationship

between visually and lab verified maize quality. Thus "true" quality is – at least partly – observable at the farm gate.

A second possible explanation for the results is that adjustments for quality differences are not done through prices, but through deductions in weight. To test for this possibility, Figure 4 plots the CDFs of deductions in the two assignment groups, with deduction defined as $(y - z)/y$, where y is the weight of maize sold as measured by enumerators and z is the agreed upon (or buyer stated) weight. While deductions are common (the mean is 4%, and in one out of four sales more than 5% of the weight is deducted), the extent of deductions is similar across groups.[19] Table 3, specification (3), regresses deductions on the treatment indicator and specification (4) shows the average treatment effect on the net sales price, pz/y, which is the per kilo price scaled by the ratio of the buyer stated weight to the enumerator measured weight. We find no evidence of a systematic relationship between quality and the net price and conclude that quality is not rewarded by lower weight deductions.[20]

There are other possible explanations for the absence of a quality-price relationship that our experiment cannot speak to. A recent literature has emphasized limited contract enforcement and informational asymmetries in both input supply chains and output value chains (Antras, 2015; Blouin and Macchiavello, 2019; Bold et al., 2017), which could severely limit traders', and, in turn, farmers, incentives for quality upgrading. A more direct, but also complementary reason is that the local traders are active in a segment of the value chain where the final product is low quality, and possibly even hazardous, maize flour, and therefore they do not place any additional value on premium quality.

Our experimental results do not rule out that buyers would reward quality over time if the seller sold higher quality maize repeatedly. That is, buyers may offer a price based on expected quality not actual quality. Over time, a seller may be able to acquire a reputation for high quality maize and buyers may be willing to pay for it. Even if this mechanism is at play, our results still show that the farmer would not be rewarded in the first season they upgrade quality, which lowers the return to upgrading.

Our findings also do not tell us about optimal quality at the farm gate, even in a high quality value chain. In the extreme case in which quality is only about the amount of

[19] The Kolmogorov-Smirnov D statistic for the test of equality of the treatment and control distributions is 0.17 [p = .39].

[20] Casaburi and Reed (2020) find, in the context of cocoa production in Sierra Leone, that traders may extend trade credit rather than adjust prices. We do not observe such interlinked contracts in the village markets we study.

defective grains and waste in the maize, the optimal quality at the farm gate, focusing on post-harvest practices, depends on the marginal cost of drying, cleaning, and sorting the maize. If these costs are low at the farm, compared to a processing plant, then it would make sense from an economic point of view to reward farmers to incur them. In the opposite case, drying, cleaning and sorting should be done at a later stage in the production chain. Neither of this implies, however, that traders should not reward maize quality at the farm gate: after all, the net weight of maize kernels per unit of high quality maize is higher than in a unit of low quality maize (which contains waste and defects). More importantly, quality does not just depend on the amount of waste in a bag of maize, but also, as discussed, on the safety of the final product. Quality assurance at an early stage of the value chain is especially important for the latter, implying that it is optimal for quality control to happen already at the farm.

6. Market for quality experiment

6.1. Introduction

The results from the returns to quality experiment in section 5 show that farmers face weak incentives to invest in high quality. As a consequence, we would expect them to invest little in enhancing maize quality and the market to be dominated by low-quality maize, which is what we observe (section 4). An important question then becomes: will farmers produce higher quality if the market values it and what are the implications for farmer profit and productivity of quality upgrading?

To answer this question, we randomly offer households (or rather villages) access to a market for quality maize. That is, maize that exceeds a certain quality threshold $\bar{q}$ is paid a premium ω over and above the prevailing market price $\tilde{p}$, resulting in the inverse demand schedule $p(q) = \tilde{p}(1 + I_{q \geq \bar{q}}\omega)$, where $I_{q \geq \bar{q}}$ is an indicator function equal to 1 if quality exceeds the minimum threshold and zero otherwise.

The intervention also aimed to improve farmers' knowledge on how to produce higher quality maize (as well as increasing their general knowledge of best-practice pre- and post-harvest agronomic activities) through extension services. Below we describe both interventions in detail. We also discuss the trial design and the data we collected, before presenting the main findings.

6.2. Intervention

6.2.1. Market intervention

The intervention was designed to emulate a market for high quality maize. To this end, we collaborated with a Ugandan vertically integrated agro trading company which committed to buy quality maize in treatment villages. The company used agents, overseen by a manager, to run their buying operation. Agents contacted all predetermined households by phone before buying commenced and were present in the villages throughout the buying seasons.[21,22] When a household was ready to sell quality maize, the household and the agent agreed on a time and buying took place at the farm gate. Agents visually inspected the maize, weighed it, and measured moisture with mobile moisture meters. Agents were not allowed to make deductions for lower quality or bargain about the price. Instead, they were instructed to reject bags that included waste (cobs, stones, dust, and insects) or defective maize (dirty, broken, immature, damaged, rotten, and moldy grain), as well as maize with a moisture level above 13% and only buy maize bags that were of sufficient quality at a predetermined price for quality maize. If the farmer was selling several maize bags of different (observed) quality, the agent bought only bags of sufficiently high quality. The households were informed why a maize bag had been rejected. The company then processed the maize bought in the treatment villages, and sold quality flour to customers in Kampala willing to pay a premium.

The research team randomly selected which villages the company should be active in and which (randomly selected) households in the villages should be invited to participate in the trial.

The research team also determined the premium for quality maize with the aim of reproducing a market equilibrium. Since such a market does not exist in the villages (see Section 5), we used a simple model predicting the premium as a function of observable outcomes – the extent of defects and prices for (average quality) maize in nearby trading centers – as a guide.

In the framework we use (see supplementary appendix S5) farmers can produce and sell either low (denoted by subscript L) or high (denoted by subscript H) quality maize, with low quality maize containing waste and defective kernels. Only high quality maize can be processed into quality flour. We further assume that a farmer can turn low quality maize into high quality maize by sorting away defects and waste and, in the benchmark model, that

[21] Most farmers (70%) sell their maize during a two-month period, between mid-January and mid-March in the spring season and mid-July and mid-September in the fall season. Some farmers, typically for financial reasons, sell their maize early, while some sell outside the main selling seasons, when prices usually are higher. In the first follow-up season, the company was active buying in the treatment villages for one month. In the remaining three follow-up seasons, the company was active for the full season (8-10 weeks).

[22] The firm could also buy from other farmers in the treatment villages, conditional on the household selling quality maize, although that seldom happened in practice.

doing so is costless. We then solve for the minimum premium, $(p_H - p_L)$, which the buyer needs to pay for high quality maize, which is simply the share of defective kernels and waste in low quality maize, valued at the premium price. Based on pre-treatment data, we predicted that maize with no visually verifiable defects, and a moisture level below 13%, would constitute maize of essentially grade 3 quality (using the EAS grading system); i.e., that the company would be able to buy grain containing 15 p.p. less waste and defects compared with the average quality on the market (see section 4.3). We further assumed, again based on pre-treatment data, that local prices, on average, are 10% lower than prices in the trading centers. These assumptions yielded a premium relative to trading center prices of 5% and an estimated premium relative to local (or village) prices of approximately 15%. As the price at which the firm bought quality maize changed with a lag, the de facto premium varied somewhat throughout the season.

The quality premium we chose should be viewed as a lower bound of a "market-based" quality premium for several reasons. First, in the model, the premium was set such that the farmers are indifferent between producing and selling high and low quality; i.e. the farmer's participation constraint binds. Second, the premium increases if we relax the assumption that the cost of sorting and cleaning away waste and defective kernels is zero. Third, we did not factor in that more waste and defective kernels, and higher moisture levels, increase the risk that the maize will become unsuitable for consumption (see section 4). Fourth, our estimate of the share of defects in the maize produced by treatment farmers in the intervention turned out to be too high. Specifically, and as reported in section 6.4.2., the company, on average, bought maize of grade 2 rather than grade 3 quality. Finally, and as discussed in detail in section 6.4.6., local traders in the treatment villages offered higher prices in response to the entry of the quality buyer, resulting in a de-facto lower premium relative to local prices.

6.2.2. Extension service intervention

To ensure that farmers had up-to-date knowledge about the pre- and post-harvest practices considered necessary for producing maize of sufficiently high quality, the agro trading company also organized an extension service program in all treatment villages. A smaller plot was selected in each village and with the help of an extension service agent, a demonstration garden was set up. Throughout the first two seasons, five meetings were held at the demo garden, during which the extension service agent provided hands-on training on best agronomic practices for plot preparation, planting, weed and pest management, and harvest

and post-harvest tasks. All treatment households were invited to the demonstrations, and close to 70% of the invitees attended the meetings. Other households in treatment villages could participate in the training as well, but few did.

6.3. Experimental design, data, and power

6.3.1. Trial design

We chose a clustered repeated measurement design for the experiment. Specifically, we restricted the number of clusters (20) but expanded on the number of waves, or seasons (7). The 20 clusters were randomly assigned to two groups: 12 to the buying group and 8 to the control group.

This design was motivated by several features of the local market and market access intervention. First, we chose a clustered rather than an individual design because we anticipated that the buying intervention could also impact households in the treatment clusters who chose not to upgrade quality. Second, we chose to expand on the number of waves rather than the number of clusters for three reasons: (i) the intervention, essentially the creation of an integrated value-chain, was complex, and costly; (ii) it may take time for farmers to decide to upgrade and/or build up a relationship with the new buyer; (iii) because of large aggregate variations, impacts may vary dramatically from season to season (see Rosenzweig and Udry, 2000). Subject to these considerations, the final combination of clusters and waves was then chosen to have sufficient power to detect moderate treatment effects.

The trial design is illustrated in Figure 5. The first three seasons serve as baseline. The intervention(s) began at the end of the third season and ran for four consecutive seasons.

6.3.2. Data

The overall objective of the data collection was to measure the various components of the farmer's profit function (equation (1)).

The household surveys were implemented at the end of the selling season when farmers had either planted or prepared the plot(s) for planting for the following season. The size of all maize plots that households had prepared for maize planting, or had already planted maize on, were collected using GPS trackers. To improve recall of revenues and expenses, households were provided with a form from the second season onwards, listing all maize plots in the current season, to be filled in with inputs and labor use and sales data. In order to ensure data quality, GPS data from the previous season was pre-loaded in the survey form, and farmers were shown satellite photos of their measured plots to confirm the plot

sizes. All calculations were checked by the survey form and any discrepancy was immediately checked and corrected.

Data were collected on the amount harvested, amount sold, and the price and revenue received. For farmers that sold multiple times, sales data were collected for each sale. The survey also collected detailed expense data, including on chemical use, seed varieties, and various pre-harvest and post-harvest practices, referring to the most recent season. Labor expenses and hours were collected for hired and family labor, respectively.

The data collected by the survey firm contained several observations with large positive values. We cannot rule out that these observations are correct (the outliers were rechecked for coding errors), and they therefore remain in our core sample. As these outliers may have an undue influence on the results, however, we also estimate treatment effects on the core components of the profit function in (1) with outliers removed, trimming the top (and in the case of profits, which could take on both large positive and large negative values, also the bottom) 1% observations.

6.3.3. Estimator and power calculations

Our benchmark ANCOVA specification uses only follow-up data and regresses outcome Y_{ijt} , where sub-script i denotes individual, j denotes cluster, and t wave or season, on a treatment indicator, D_{jt}, which takes on the value 1 in treatment clusters and zero in control clusters, a full set of season dummies and a lag-dependent variable, i.e., the value of the outcome pre-treatment $\bar{Y}_{ij,PRE}$:

$$Y_{ijt} = \gamma D_{jt} + \sum_4^7 \delta_t + \theta \bar{Y}_{ij,PRE} + \varepsilon_{ijt} \ . \tag{3}$$

Here, ε_{ijt} is an idiosyncratic error. The coefficient of interest, γ, gives the average treatment effect over the four follow-up rounds. We report point estimates and *p*-values, with the latter estimated both based on clustered-by-village standard errors and computed using randomization inference.

We estimated the power of the design before the trial began using data from a small pilot. Table 4 updates these power calculations for the ANCOVA estimator using data from the control group.[23] As evident, the design is powered to detect small effects for some outcomes (like price), but medium effects for some others (like expenses, harvest, and acreage). Expressed as share of the standard deviation, the MDEs vary from 14% (price) to

[23] For general structures of auto- and intraclass correlation, there are no analytic results for the MDE for the panel design we employ. We use the Stata package in Burlig et al. (2020), and control group data, to estimate MDEs, allowing for time varying serial- and intraclass correlation.

35% (expenses and harvest). Expressed as a share of the mean, the MDEs vary from 6% (price) to 48% (expenses).

6.3.3. Assignment, attrition, and baseline balance

The sample population for the market access experiment consists of smallholder maize farmers in 20 maize farming communities (villages/clusters) in Kakumiro. In each cluster, we randomly selected 10 households which had planted maize in the previous season; i.e., the season before the first baseline season. In addition, we randomly selected up to 5 replacement households in each cluster.

The first three seasons serve as a baseline panel. After the first season, households that did not give consent to continue to participate, or that we could determine had moved, or that were involved in commercial maize trading, were replaced by households from the replacement list. No replacements were added after the first season. At the end of the last pre-intervention season, the sample included 544 household-by-season observations from 189 households in 20 clusters (see Table 5).

Follow-up lasted for four seasons. As reported in Figure 5 and Table 5, less than 5% (9 households) of the 189 households in the final baseline sample attritted. The attrition rates were similar across assignment groups (see Table 6). Of the non-attritters (180 households), 86% were re-surveyed in each follow-up season, and the remainder were surveyed in some but not all seasons, yielding a follow-up sample of 677 household-by-season observations. The re-survey rates; i.e., the share of the 180 households that were surveyed per season, were similar across assignment groups (see Table 6). Combining the baseline and follow-up data, we have a panel of households with both baseline and follow-up data, with 1,198 household-by-season observations for 180 households in 20 clusters over seven seasons.

Table 7 reports summary statistics and mean comparisons between the treatment and control group across a broad set of outcomes. Panel A focuses on household characteristics: the main decision maker (when it comes to agricultural decisions) is male in most households (18% of the decision makers are female) and about 40% of them had completed primary school. Average household size is 6.2.

Panel B and Table A5 in the supplementary appendix present baseline farm enterprise outcomes. There are large variations across seasons (Table A5). For example, farmers invested on average 30% more in season three than in season one; the price of maize was more than 50% higher in the first season than the third; and profit in season two was roughly 50% higher than in season three.

Summary statistics for household characteristics and farm enterprise outcomes (the latter pooled across the three baseline seasons) by assignment group are reported in columns (4)-(5) in Table 7 and a test of baseline balance is reported in columns (6) and (7). We find no evidence of differences in means among the household characteristics or the farm enterprise variables: all p-values for the core farm enterprise outcomes in column (7) are greater than .6. The last four rows of Table 7 test whether the variables listed within each grouping jointly predict treatment assignment. Our joint balance tests fail to reject the null hypothesis that the household characteristics ($p = .964$), farm enterprise outcomes ($p = .612$; $p = .580$), and all baseline variables together ($p = .794$) do not predict assignment to treatment.

6.4. Results: Market access

6.4.1. Summary

We begin by summarizing how market access affects the main outcome of interest: profits. Figure 6, Panel A, shows that the cumulative distribution function (CDF) for profits is strongly shifted to the right for farmers who gained market access and we reject the hypothesis that the two distributions are equal (Kolmogorov-Smirnov D statistic is 0.17, p = .000). Several factors contributed to this profit increase: first, farmers who produced higher quality maize and sold to the agro-trading company received higher prices. Second, and to a lesser degree, farmers who continued to sell to local traders in treatment villages also earned higher prices. Third, farmers in treatment villages grew more maize on a given plot of land. Together this produced a large increase in revenue. At the same time, farmers spent more on cultivating maize. The combined increase in revenue and expenses raised mean profits by \$63-\$98 or 36%-80% (depending on how own and family labor is priced).

6.4.2. Quality upgrading and price

In each post-treatment season, the agro-trading company offered to buy maize of sufficiently high quality from pre-selected households in the treatment villages. Averaging across the four post-treatment seasons, 40% of farmers sold at least some bags of maize of sufficiently high quality to the company in a given season. The share of farmers who sold to the trading company increased with each additional season. In the first season, about one in five households sold quality maize, in the fourth (and last) season, that ratio had tripled (see Figure 7, Panel A). This upward trend suggests that it takes time for many households to make the necessary adjustment in their agricultural practices to produce maize of sufficient quality, but also that the switch to producing high quality maize is a permanent one. For

example, the probability of selling to the high quality buyer for a farmer who has not interacted with the high quality buyer is 24% across all four seasons compared with 75% for a farmer who has sold to the high quality buyer at least once.

Figure 7, Panel B, provides more detailed information on how the farmers and the company interacted, using data from the last two seasons.[24] Approximately four out of 10 of the households did not attempt to sell (quality) maize to the company. One-third of the households sold all they wanted to sell. For 15% of the farmers, the company first refused to buy (and required the maize to be sorted, cleaned and/or dried further), but then bought at least a subset of bags once the farmer had upgraded the quality. For one in ten households, the company refused to buy because the quality was low.

Consistent with the farm gate quality inspection used by the company, lab verified quality of the maize was high. In the last season, quality of the maize bought by the company was measured in the lab (using the method described in Section 4.2). The mean share of defects was 8.1% (std. 2.6%), with the maximum of 16%. To put this in context, recall (see Section 4.2) that the average share of defects in maize sold in nearby villages was 26% (std. 34%). Using the EAS quality classification as a yardstick, 84% of the maize bags bought by the company were graded maize, with 6 of ten bags classified as grade 1 or 2 maize. Of the maize sold in nearby villages, more than half (56%) was ungraded/reject maize, with only 3 in ten bags classified as grade 2 maize (see section 4.2).

Farmers with access to a high quality market received significantly higher prices: the CDF of prices in treatment villages is strongly shifted to the right compared to the control group (Figure 6, Panel B). The Kolmogorov-Smirnov D statistic is 0.34 (p = .000). The regression equivalent is presented in Table 10, column 1: on average, farmers earned $2.40 or 11% more per bag of maize (140 kg) they sold (p = .001, control mean $21).

6.4.3. Investments and productivity

Market access may encourage farmers to invest more via two channels: (i) the intervention offered farmers in the treatment group higher prices conditional on producing high quality maize. It thus incentivized farmers to invest in upgrading quality. (ii) As farmers obtained higher prices for their crop, profit-maximization predicts that they would use more inputs to increase the amount of (high-quality) output to be produced. These predictions are borne out

[24] The agro trading company did not collect information on reasons for not buying in the first two buying seasons.

in the data: treatment farmers increased investments across a wide range of cultivation inputs and activities that improve both quality and productivity.

Treatment farmers bought more inputs and hired more labor for pre-harvest activities (see Table 8), investments that primarily – though not exclusively – affect how much maize is produced. Specifically, farmers spent an additional $2.3 or 60% ($p$ =.048, control mean $3.8) on hybrid and open pollinated seeds as well as inorganic fertilizer. The value of all agricultural input purchases, which also includes plant growth booster, animal manure, pesticides and herbicides, increased by $4.3 or 30% ($p$ =.06, control mean $13.3). Although these treatment effects represent large relative increases, in absolute terms, modern input use is low: 3% of control farmers used inorganic fertilizer, 13% used improved seeds, and input expenses amounted to 13% of all expenses. Second, farmers spent $16 or 30% ($p$ = .275, control mean $54) more on hiring agricultural workers to prepare the land, plant maize seed, weed and spray the crop, although the effect is not precisely estimated.

Farmers also invested more post-harvest, which is viewed as crucial for maize quality (see discussion in Section 4). At baseline and in control villages, few farmers properly processed their crop: two thirds of farmers dried their maize in direct contact with soil (i.e., only one third dried their maize on a tarpaulin or in other ways that avoid contact with the soil), 13% sorted their maize and one fifth winnowed it (see Table 9). With access to a market for high quality maize the share of farmers who engaged in these practices nearly doubled: 60% properly dried their maize (a difference of 24 percentage points, p = .000), 27% sorted the maize (a difference of 14pp, p = .002) and 34% winnowed it (a difference of 15pp, p = .033). Consistent with this, spending on harvest and post-harvest activities rose by 20% (p = .255, control mean $31), an increase mainly driven by higher expenses on hired labor (a difference of 40%, p = .144, control mean $15.6).

Summing across all items of cultivation expenditure, farmers in treatment villages invested $18 more than those in control villages (Table 10, column (6)), an increase of 17% (p = .305, control mean $106). The increase is somewhat smaller (15%) when trimming the expenses data. Farmers did not change the area under cultivation (Table 10, column (2)).

Farmers in treatment villages increased their total maize harvest as well as their yields. Figure 6, Panel C, shows that yield is higher in treatment than in control villages across the entire distribution. On average, yield (measured in kilogram per acre) rose by 112 kg or 15% (p = .036, control mean 793 kg), and total harvest by 239 kg or 13% (p = .308, control mean 1887 kg) as seen in Table 10, columns (3) and (4). This quantity increase together with the price increase translates into a significant and economically important

increase of the value of farmers' harvest, column (5), which rose by $78.7 or almost 30% per season (p = .079, control mean $286.7). In the trimmed data, the percentage increases for yield, harvest and harvest value are the same, but more precisely estimated.

How much do the price effect and the productivity effect each contribute to the increase in revenue? Denoting mean harvest and mean price in assignment group $d = \{1,0\}$ (treatment, control) by $\bar{Y}^d$ and $\bar{p}^d$, and Δx the treatment effect on outcome x, the treatment effect on harvest value (pY) can be decomposed into a pure price/quality-effect ($\bar{Y}^0 \Delta p$), a pure quantity effect ($\bar{p}^0 \Delta Y$) and an interaction:

$$\Delta pY = \bar{Y}^0 \Delta p + \bar{p}^1 \Delta Y + \Delta p \Delta Y \ . \tag{4}$$

Given the treatment effects on price and harvested amount, the quantity effect accounts for 50% of the increase in harvest value, the quality effect accounts for 45%, and the remainder is explained by the interaction. Hence, the quality and quantity channels contribute in (almost) equal measure to the increase in harvest value.

6.4.4. Income

The ultimate aim of linking farmers to a buyer of high quality maize is to increase farmer income and reduce rural poverty. To calculate farmer income, we need to value the farmers' own and family labor, which amount to an average of 408 hours per season (in the control group).

Comparing the treatment effects on family and hired labor, we find that farmers in treatment villages reduce family labor hours by 75 hours per season or 16% (p =.106, control mean 449 hours). On the other hand, and in line with the overall increase in expenses, treatment farmers hire substantially more labor.[25] Farmers in treatment villages increase their spending on hired labor by $26 or 36% ($p$ = .322, control mean $71), equivalent to an additional 121 hours per season at the hourly wage. Summing the two sources of labor, these effects amounted to a net increase in labor of 46 hours per season in the treatment group compared with the control group.

The total effect of these changes in labor composition on investment and profits depends on the relative productivity of family and hired labor and by implication, the relative value of the two types of labor. Valuing family labor is challenging: one possible approach is to value family labor at the market wage, another possible approach is to put zero value on it,

[25] There is an active market for hired labor in the study villages, with farmers in control villages spending on average $71 on hired labor equivalent to 70% of all monetary expenses and roughly 315 hours per season at the hourly wage.

as no monetary costs are incurred. In reality, however, family labor clearly has an opportunity cost. At the same time, it is most likely not a perfect substitute for hired labor: farmers typically hire labor for more difficult and physically demanding tasks, and even for the same task, hired labor tends to be adult labor while own/family labor is a mix of child and adult labor.[26]

In the end, we remain agnostic and let the value of own and family labor vary between zero and the market wage for hired labor. That is, we specify a profit function

$$\Pi = pY - cx - \varphi w L_F \ , \tag{5}$$

where $pY - cx$ is harvest value minus monetary expenses (including hired labor), $\varphi \in [0,1]$, w is the hourly market wage for hired labor, and L_F measures hours of own and family labor. In Table 10, columns (7) and (8), we present the regression results for the two polar cases: family labor valued at zero ($\varphi = 0$) and family labor valued at the market wage ($\varphi = 1$). The treatment effects (and their 95% confidence intervals) for intermediate values of α are plotted in Figure 8.

Linking farmers to a buyer of high quality maize increased their income substantially.[27] Including only monetary expenses in the profit calculation, farmer profit increased by \$63, or 36%, in treatment villages (p = .051, control mean \$178); Table 10, column (7). The treatment effect is larger (43% increase in treatment as compared to control) and more precisely estimated (p = .024) when trimming the top and bottom 1% observations (Table 10, Panel B). Valuing family labor at the market wage, the increase is even starker (since farmers swapped family for hired labor): farmer profits were \$97 higher in treatment than in control villages (p=0.027, control mean \$122) column (8). Trimming reduces the effect size (70% increase in treatment compared with control; p = .020). As shown in Figure 8, profits increase smoothly (given the reduction in own hours in treatment villages) as α increases from zero to one and p-values decrease. For $\varphi = 2/3$, which corresponds to our

[26] We can compare the productivity of the two types of labor in the control group by relating the total amount of hours of hired labor per acre for a specific task (e.g. plot preparation, planting, weeding, harvesting) to the hours the average household member would take to perform the same task. These calculations suggest that family labor is about two thirds as productive as hired labor.

[27] Profits are low and for several farmers even negative in our data set. Specifically, profit turns positive at the 10th percentile in the control group when family labor is valued at zero. As a reference, Karlan et al. (2014) report profits turning positive at the 15th percentile when family labor is valued at zero. They further report that the market for hired labor is thin – the experiment is set among maize farmers in northern Ghana – and that family labor, when valued at market wages for hired labor, is the most important component of total costs. In their data, profits turn positive at the 60th percentile of realized profits when family labor is valued at market wages. In the context of our study, the market for hired labor is active. So while we also find that labor is the most important component of total costs, roughly half of the costs constitute costs of hired labor. When family labor is valued at market wages, profit turns positive at the 25th percentile.

guesstimate of the relative productivity of family and hired labor (see footnote 26), the treatment effect on profits is $88, or 55% ($p$ = .040, control mean $165).

These effects represent large absolute increases in the context of our study, where most people live on less than 1 dollar a day. They also represent large increases relative to average annual income from all sources in the region: additional income from maize farming in the market access group represents a 16-24% increase in average annual income relative to a typical family in the region (UBOS 2019).

Market access did not just increase farmer income, but, at least tentatively, put farmers in treatment villages onto a different income growth trajectory. Specifically, in the first treatment season, income growth, measured as the percentage change in average profit across seasons, was similar in the two assignment groups. But already at the second season, treatment villages started pulling away. By the final season, income growth in treatment villages was 27% higher than in control villages.

6.4.5. Mechanisms and productivity

So far we have shown that access to a market for premium quality resulted in an increase in measured inputs and an increase in output per acre of land. It is also possible that other inputs, that we have not measured, increased and that market access affected how well a given bundle of inputs were used; i.e., total factor productivity (TFP). For example, market access may have increased farmers incentive to perform various agricultural tasks emphasized in the extension service program in a way closer in line with best practice. In this section we go beyond the estimation of treatment effects and attempt to examine the relative importance of both measured and unmeasured inputs in explaining the increase in output. To do so, we face two challenges. First, we need comparable measures of output across farms in a setting were farmers sell different quality produce at different prices. Second, we need to add more structure to the production process.

We choose to measure output with harvest volume. This quantity-based outcome solves problems with measuring TFP related to differential prices and markups. Moreover, although harvested maize may be of different qualities, activities and investments to improve quality during pre-harvest, for instance through improved planting or weeding techniques, will also likely increase output, thus mitigating concerns that measured improvements in quantity-based productivity will be (downward) biased. Activities and investment at the post-harvest stage, on the other hand, will likely result in higher quality but lower volumes of (quality) maize to sell. Thus basing the TFP calculation on volumes sold is more problematic.

To assess relative importance of factors of production, we also need to specify the relationship between inputs and output; i.e., a production function. In the context we are considering, it is reasonable to assume that farmers use no physical capital. The main inputs into the production function are therefore land, A, and labor, L. We relax the assumption of perfect substitutability between different types of labor and thus treat hired, L_H, and family (or own) labor, L_F as separate inputs. We further assume that land quality can be enhanced by investment (e.g. fertilizer). Specifically, we postulate that output is a function of the stock of fertile land E, with $E = Ae^{\omega\mu}$, where μ is the amount of soil and crop enhancing investment and ω is the return to land quality of such investment. Finally, we assume that harvest volume, Y, is well-described by the following Cobb-Douglas production function:

$$Y = \theta L_F^{\alpha_1} L_H^{\alpha_2} E^{\alpha_3} \ , \tag{6}$$

where Y is output (harvest) and θ is the farm's TFP.

We estimate two versions of the production function in (6). In the log-linear version we write:

$$\ln Y^d = \alpha^d + \alpha_1^d \ln L_F^d + \alpha_2^d \ln L_H^d + \alpha_3^d \ln A^d + \alpha_4^d \ln \mu^d + \theta^d + \tilde{\varepsilon}^d \ , \tag{7}$$

where subscript $d \in \{0,1\}$ denotes assignment groups (0 control, 1 treatment), ω is normalized to 1 for notational simplicity, and $\tilde{\varepsilon}^d$ is a zero-mean error term assumed to be independent of the regressors x_j^d and θ^d. The linear version simply replaces the logarithms with the levels.

To consistently estimate the parameters, $\boldsymbol{\alpha^d} = [\alpha_1^d, \alpha_2^d, \alpha_3^d, \alpha_4^d]$, of the measured inputs, $\boldsymbol{x^d} = \left[\ln L_F^d, \ln L_H^d, \ln A^d, \ln \mu^d\right]$, in (7) we need to assume that these inputs are independent of TFP (θ^d), given the treatment status. This is the key condition, when using experimental data, in the sequential ignorability assumption of Imai et al. (2010, 2011). As noted in Heckman and Pinto (2015), data from an experiment can be used to test (a portion) of this assumption. Specifically, if we assume that observed and unobserved inputs are independent in the control group and that the parameters of the production function are the same in the control and the treatment group (i.e., assume autonomy), then we can test whether the experimentally induced changes in unmeasured inputs are independent of experimentally induced changes in measured inputs. The intuition for this test is as follows. The inputs for treated households are the sum of the inputs they would choose if they were assigned to the control group plus the increment due to treatment. Assuming independence of observed and unobserved inputs in the control group plus autonomy implies that a test of H_0: $\widehat{\boldsymbol{\alpha}}^{\mathbf{1}} = \widehat{\boldsymbol{\alpha}}^{\mathbf{0}}$ is equivalent to a test that $Cov(\boldsymbol{\Delta x}, \Delta\theta) = 0$, which is sufficient to obtain an

unbiased estimate of $\boldsymbol{\alpha}$. For both specifications of the production function we cannot reject the null hypothesis of independence of the increments (see Table 11, Panel B).[28] We can thus obtain consistent estimates of the impact of observed inputs on output by regressing log output, $y^d = \ln Y^d$, on the vector of measured inputs, $\boldsymbol{x^d}$, and a dummy for treatment assignment, δ^d,

$$y^d = \delta^d + \boldsymbol{\alpha x^d} + \varepsilon^d \ , \tag{8}$$

where $\delta^d = \alpha^d + E[\theta^d]$ and $\varepsilon^d = \tilde{\varepsilon}^d + [\theta^d - E[\theta^d]]$.

We can now decompose the treatment effect on harvest volume, $E(y^1 - y^0)$, into components attributable to changes in the inputs that we can measure; i.e., land, labor, and crop and land enhancing inputs, and the unmeasured component (TFP):

$$E(y^1 - y^0) = (\delta^1 - \delta^0) + \sum_k \alpha_k \, E[x_k^1 - x_k^0] \ , \tag{9}$$

where $(\delta^1 - \delta^0)$ is the contribution of TFP (or unobserved inputs) to the mean treatment effect and$\sum_k \alpha_k \, E[x_k^1 - x_k^0]$ is the contribution of measured inputs to the mean treatment effect.[29] The vector of observed inputs, $\boldsymbol{x}$, can explain the treatment effect on harvest volumes only if they affect harvest ($\gamma \neq 0$) and, on average, are affected by the experiment, so that $E[x_k^1 - x_k^0] \neq 0$. Both these conditions hold in our experimental data in both specifications (see bottom of Table 11).

Table 11, Panel A reports the estimated treatment effects, using the linear and log-linear production function specifications, and the contributions of measured and unmeasured inputs (TFP). For both specifications, the overall treatment effect is significant. Output is 328 kilograms higher in the linear specification (a 21% increase relative to the control group) and 0.12 log points higher in the log-specification (thus a 12% increase relative to the control group. For both specifications, we conclude that more than half the increase in output works through TFP (56% in the linear specification and 87% in the log-specification). Of the measured inputs, the largest contribution comes from hired labor (which contributes 17% of the increase in harvest in the log-specification and 15% in the log-specification). The contributing effect of family labor and land is positive in the linear specification but negative in the log-specification, while the inputs fertilizer and hybrid seeds have a positive albeit

[28] Alternatively, if we are willing to assume independence of the observed and unobserved inputs (in both treatment and control), the test of H_0: $\hat{\boldsymbol{\alpha}}^{\boldsymbol{1}} = \hat{\boldsymbol{\alpha}}^{\boldsymbol{0}}$ is equivalent to testing autonomy; i.e., the parameters of the production function are the same in the control and the treatment group (see Heckman and Pinto, 2015).

[29] A complementary approach to measure TFP is to estimate the production function (equation (4)), using control group data, and then back out TFP as the residual. The treatment effect on productivity can then be estimated in a second stage (see for example Atkin et al., 2017).

small effect in both specifications. In sum, increases in measured inputs account for 13-44% of the treatment effect on harvest, with the switch from own to hired labor the most important factor, while improvements in TFP account for the lion's share of the increase in output.

6.4.6. Local market equilibrium effects

The intervention involved an offer to buy quality maize at a premium. As documented in section 6.4.3, across the four seasons in which the buying operation was active, on average 40% of farmers sold at least some high quality maize to the premium buyer in a given season and average prices in the treatment group increased by 11% relative to the control group.

The entry of a new buyer in local (village) markets could affect (local) prices in two ways. First, and directly, households who successfully produced higher quality maize could sell it to the new buyer at a premium. Second, the entry of the new buyer could affect other traders' behavior in village markets even in the case of differentiated products (higher or lower quality maize), and therefore also the maize price for households who did not upgrade quality or did not sell to the high quality buyer for other reasons. Our trial, which induced variation in exposure to the new buyer across clusters, was designed to (partly) capture such local market equilibrium effects.

For the local market equilibrium analysis, we use sale level data, which records for each sale the price received, the type of buyer, and for local traders also the name of the buyer. These data were collected from season 3, the last baseline season, onward. Our analysis here focuses on the three types of buyers: local traders or aggregators (LT), commercial traders (CT), and the high quality buyer (HT). Let D denote treatment assignment ($D = 1$, if a household reside in a village being treated and $D = 0$ otherwise) and superscript $d = \{1,0\}$ to represent variables when treatment is fixed at d. Further, let superscript j denote trader type. Dropping time subscripts, we can then estimate type and assignment-specific average prices $\bar{p}^{d,j}$ and market shares $\overline{ms}^{d,j}$ using household-sale data by estimating the following system of regressions

$$p_i = \sum_{d,j} \beta^{d,j} z_i^{d,j} \tag{10}$$

$$z_i^j = \alpha_0^j + \omega^j D_i \text{ for} = \{LT, CT, HT\}\,. \tag{11}$$

p_i is price paid at sale i, $z^{d,j}$ and z^j are dummies for assignment group d and trader type j, and trader type j, respectively. The type and assignment-specific average price is then $\bar{p}^{d,j} = \beta^{d,j}$, while the type and assignment-specific market share is $\overline{ms}^{d,j} = \hat{\alpha}_0^j + \hat{\omega}^j D$. Note that average price in the control group is $\bar{p}^0 = \sum_j \overline{ms}^{0,j} \bar{p}^{0,j}$; i.e., the (average) price paid by

trader type j, weighted by the market share of j in the control group, and that the average treatment effect on price, Δp, can be written as

$$\Delta p = \sum_j \overline{ms}^{1,j}\left(\bar{p}^{1,j} - \bar{p}^{0}\right) . \tag{12}$$

In appendix, Table A6, we show that the market shares ($ms^{d,j}$) and prices ($p^{d,j}$) of local and commercial traders, respectively, are balanced at baseline across treatment and control. Table 12, panel A, shows average market shares and prices paid by commercial and local traders, relative to the control group mean over the four intervention seasons.[30] Entry of the high quality buyer decreased the average market share of local traders from 78% in control to 48% in treatment (Δms^{LT} = 0.308; p = .002). Commercial traders' market share decreased from 21% in control to 13% (Δms^{LT} = 0.08, p = .213). That is, farmers in the treatment group primarily switched from selling maize to local traders to selling (high quality) maize to the agro-trading company, although in relative terms, both local and commercial traders lost about 40% of their market share to the high quality trader.

Commercial traders in the treatment group paid 2 pp less than the average price in the control group, but we cannot reject the null hypothesis that the prices are equal (p = .508). The (average) price paid by local traders in the treatment group, on the other hand, is 5.6 pp higher (p = .039) than the average price in the control group.

What do these local market equilibrium effects imply? The random assignment of villages ensures that we can directly measure the causal effect of the market access intervention on the (average) price. Estimates of differences in prices for various types of buyers, however, are not necessarily causal, since a farmer's decision of whom to sell to is endogenous.

To disentangle the causal effect from the selection effect, we decompose the treatment effect on local and commercial trader prices using a simple potential outcomes framework. The potential price paid for a sale i in follow-up season t to trader type j, $p_{i,t}^{d_i,j}$, depends first on whether or not the high quality trader enters, $d_i = \{1,0\}$, which is the exogenous treatment assignment. In the treatment group, farmers can either sell to the high quality trader, HT, or continue to sell to local, LT, and/or commercial traders, CT. In the control group, farmers can only sell to local LT, and/or commercial traders, CT.[31] The decision to whom to sell is (as

[30] Given the large variation in average price between seasons, we report normalized prices; i.e., $p_{it} = (p_{it}^{n} - \bar{p}_t^0)/\bar{p}_t^0$, where p_{it}^{n} is the nominal price and $\bar{p}_t^0$ is the average price in the control group in season t.

[31] In what follows, we assume that a sale that goes to local traders in the control group, would go to either the high quality buyer or to local traders in the treatment group, and similarly for sales to commercial traders. Thus, we rule out that entry of the high quality buyer causes a switching of sales from local to commercial traders.

above) denoted by a dummy variable $z_{i,t}^{j}$, which takes on the value 1 if a sale goes to trader type j and zero otherwise. With this notation, we write the difference between the average price selling to a trader of type $j = LT, CT$ in the treatment group and the average price in the control group as:

$$\Delta p^{j} = E\left(p_{i,t}^{1,j} \middle| D_i = 1, z_{i,t}^{j} = 1\right) - E\left(p_{i,t}^{0} \middle| D_i = 0\right) . \tag{13}$$

We can decompose the average price difference as

$$\Delta p^{LT} = \Delta p_{causal}^{LT} + \Delta p_{causal}^{LT} = \tag{14}$$
$$\left[E\left(p_{i,t}^{1,j} \middle| D_i = 1, z_{i,t}^{j} = 1\right) - E\left(p_{i,t}^{0,j} \middle| D_i = 1, z_{i,t}^{j} = 1\right)\right]$$
$$+\left[E\left(p_{i,t}^{0,j} \middle| D_i = 1, z_{i,t}^{j} = 1\right) - E(p_{i}^{0} | D_i = 0)\right] .$$

The first difference in (14) measures the causal or `welfare' effect, i.e., how much prices for sales to trader type $j = LT, CT$ in treatment differ from prices for sales to these trader types if the high quality buyer had not entered. The second difference is the selection effect. For example, if farmers who continue selling to local or commercial traders when the high quality buyer enters are positively selected in terms of price; i.e., in the absence of the market access intervention they would receive higher prices than the average farmer, this selection effect is positive.

The average price $E\left(p_{i,t}^{0,j} \middle| D_i = 1, z_{i,t}^{j} = 1\right)$ in the selection term is counterfactual, because we cannot observe what sales in treatment to trader type $j = LT, CT$ would have earned if the high quality trader had not entered. However, the availability of baseline data makes it, in principle, possible to estimate this term. To make progress, we first introduce an additional conditioning variable that is invariant across all sales-season observations for a given farmer k, namely, the number of seasons the farmer who originated sale i sold to the high quality buyer across all follow-up seasons, denoted $\sum_t w_{k,t}^{HT}$, where $w_{k,t}^{HT}$ is a dummy equal to 1 if farmer k sells to the high quality buyer in season t and zero otherwise. The counterfactual price can then be written as

$$E\left(p_{i,t}^{0,j} \middle| D_i = 1, z_{i,t}^{j} = 1\right) = \tag{15}$$
$$\sum_s E\left(p_{i,t}^{0,j} \middle| D_i = 1, z_{i,t}^{j} = 1, \sum_t w_{k,t}^{HT} = s\right) \times \Pi_i(\sum_t w_{k,t}^{HT} = s | D_i = 1, z_{i,t}^{j} = 1) ,$$

where $\Pi_i(\sum_t w_{k,t}^{HT} = s | D_i = 1 = 1, z_{i,t}^{j} = 1)$ is the probability that sale i to trader type j originated from a farmer k, who sells to the high quality trader in s of the follow-up seasons.

We now discuss how to estimate the terms in (15). First, $\Pi_i(\sum_t w_{k,t}^{HT} = s | D_i = 1, z_{i,t}^{d,j} = 1)$ is calculated as the share of all sales to a trader of type j in treatment at follow-

up, which originate from farmers who have sold s times to the high quality buyer during the experiment.[32] Second, to estimate $E\left(p_{i,t}^{0,j} \middle| D_i = 1, z_{i,t}^{d,j} = 1, \sum_t w_{k,t}^{HT} = s\right)$ with the help of baseline data, we need to make two assumptions: (i) we assume that the normalized potential prices are time invariant, so that $p_{i,t}^{0,j} = p_i^{0,j}$, (ii) we assume that selection does not happen at the sale-season level, but rather at the farmer level and that the number of seasons the farmer sold to the high quality buyer, $\sum_t w_{k,t}^{HT}$, is a sufficient statistic for selection. Together, these assumptions imply that $E\left(p_{i,t}^{0,j} \middle| D_i = 1, z_{i,t}^{1,j} = 1, \sum_t w_{k,t}^{HT} = s\right) = E\left(p_i^{0,j} \middle| D_i = 1, \sum_t w_{k,t}^{HT} = s\right)$. We estimate these conditional expectations both parametrically and non-parametrically by regressing baseline prices for sales to traders of type $j = LT, CT$ in the treatment group on either the number of times the farmer sold to the high quality buyer at follow-up, $\sum_t w_{k,t}^{HT}$, or on a set of dummies, which take on the value 1 if the farmer sold s times to the high quality buyer at follow-up and zero otherwise.

The results are reported in Table A7, panels A and B, and show that farmers who sell to the high quality trader are positively selected on baseline price. More precisely, sales to local traders at baseline from farmers who never sell to the high quality trader during the experiment fetch a price that is 3 pp lower than the average price (in the control group). Sales to local traders at baseline from farmers who sell to the high quality trader in all follow-up seasons, on the other hand, earn 4 pp more than the average price, a difference of 7 pp (Panel A, column 1). This translates into an average effect of an additional sale to the high quality buyer on baseline prices of 2 pp (p =.06, Panel A, column 2). For sales to commercial traders, the results are similar.

Conversely, farmers who mainly sell to local or commercial traders (once the high quality trader has entered) are negatively selected on baseline price ($\Delta p_{select}^{LT} = -0.025$ and $\Delta p_{select}^{CT} = -0.023$). These findings are consistent with a model, where farmers are heterogeneous with respect to an "ability" dimension, which increases the price they can receive from local and commercial traders as well as their likelihood of upgrading and selling to the high quality buyer.

The causal treatment effect on prices for sales to local and commercial traders is therefore larger than the average differences reported in Table 12, panel A. Combining the

[32] As local and commercial traders lose the same percentage of market share, $\Pi_i(\sum_t w_{k,t}^{HT} = s | D_i = 1, z_{i,t}^{d,j} = 1)$ is in practice independent of whether the sale goes to a local or a commercial trader. We therefore estimate $\Pi_i(\sum_t w_{k,t}^{HT} = s | D_i = 1, z_{i,t}^{HT} = 0)$, which is the probability that a sale that does not go to the high quality trader originates from a farmer who sells to the high quality trader in s seasons.

estimates of Δp^j in (12) and Δp^j_{select} in (14), we conclude that sales to local traders in treatment earned 8.2 pp (p =.01) more than they would have in the absence of the high quality trader entering. The causal treatment effect on sales to commercial traders, on the other hand, is zero (see Table 12, panel B).

Turning from farmer to trader behavior, the difference in prices paid by local traders in the two assignment groups could be driven by at least two mechanisms: (i) intensive margin changes; i.e., existing buyers increase their prices in response to the entry of the high quality buyer; (ii) extensive margin changes; i.e., the entry of the high quality buyer drives the traders who offered the lowest prices out of the market. While the exact mechanism is less important from a farm household perspective, the traders' response sheds light on how local markets function.

We use baseline sales data collapsed at individual trader-level to test for selective exit of local traders (see supplementary appendix S8). Specifically, we estimate the probability of exit as a function of the average price paid at baseline, assignment group and their interaction (see Table A8 for results). The regression shows that more traders exit in treatment than in control, that traders who exit paid lower prices at baseline, and that this selection effect is uniform across assignment groups. Thus, we find no evidence of selective exit in response to the high quality trader entering and conclude that incumbent local traders in treatment villages most likely raised their prices.[33] This suggests, in turn, that local traders (on average) earned positive rents before the entry of the high quality buyer and hence that the village markets are not perfectly competitive.

In summary, the analysis shows that treatment households who sold to local traders earned higher prices and that incumbent local traders paid higher prices than they would have in the absence of the high quality trader entering. As the premium for quality was set relative to trading center prices, this local market equilibrium effect reduced the relative price of higher quality maize, which in turn may have weakened the incentives for quality upgrading.

How important, quantitatively, is the local market equilibrium effect in explaining the increase in the average price in the market access group? To answer this question, we perform a simple back-of the envelope calculation. We solve for the causal treatment effect

[33] Entry of traders (several of which may have been active in the markets in the first two baseline season when we did not record names and types of traders) is an additional channel; i.e., the price increase in treatment could be driven by 'new traders' entering who set higher prices. In the data, the price difference between treatment and control is the same for 'new traders' (8 pp) as for incumbent traders (6 pp) (see appendix S8). This result lends support to our assumption to treat incumbent and 'new traders' as a uniform group. Note that we do not observe 'new traders' at baseline, thus we cannot separately estimate the causal effect on sales to 'new' traders.

of selling to the high quality trader Δp_{causal}^{HT} by substituting Δp_{causal}^{LT} and Δp_{causal}^{CT} into the expression for the average price change between treatment and control Δp in equation (12):

$$\Delta p_{causal}^{HT} = \bar{p}_{causal}^{1,HT} - \bar{p}^{0} = \frac{\Delta p - \sum_{j=LT,CT} \overline{ms}^{1,j}\left(\bar{p}_{causal}^{1,j} - \bar{p}^{0}\right)}{\overline{ms}^{1,HT}}. \quad (16)$$

We then calculate a counterfactual price change, which is the average treatment effect on prices assuming local [commercial] traders in the treatment group pay the same as local [commercial] traders in the control group; i.e., $\Delta\tilde{p} = \sum_{j=LT,CT} \overline{ms}^{1,j}\left(\bar{p}^{0,j} - \bar{p}^{0}\right) + \overline{ms}^{1,HT}\left(\bar{p}_{causal}^{1,HT} - \bar{p}^{0}\right)$. The difference $\Delta p - \Delta\tilde{p}$ measure the market equilibrium effect. We estimate $\Delta\tilde{p}$ to be 0.082 while Δp using sales data is 0.121. Thus 33% of the average price in the treatment compared to the control group is driven by the higher prices paid by local traders in response to the entry of the high quality buyer.

6.4.7. Extension service

To examine whether impacts of the market access program are driven only by its extension component, and to estimate the direct effects of hands-on-training on cultivating a well-known crop, we ran a parallel trial where the company only implemented the extension service component (see Bold et al, 2020 for details).

The results from this parallel trial show no impacts on harvest value, price, and yield, suggesting that the economically significant impacts observed in the market access-cum extension trial materialize only when farmers also had the opportunity to sell their produce to a buyer who rewards quality. Similarly, treatment farmers increased neither overall expenses nor specific components emphasized in the training, such as expenses on modern inputs and post-harvest practices.

These results do not rule out that the extension service program had an impact in the access to market intervention. In fact, one interpretation of the productivity increase we document (see section 6.4.5) is that better knowledge about best practice pre- and post-harvest processes was put to use with market access.[34] But they make it unlikely that this kind of supply intervention by itself would significantly and sustainably change how farmers operate.

[34] Previous research has documented positive, albeit small, impacts of providing accessible, tailored, and timely information through hands-on training on demonstration plots (see for example, Duflo et al., 2007; Hanna et al., 2012; and Islam and Beg, 2020). Related to this, Bernard (2017) show that Senegalese onion producers adopted quality-enhancing inputs when they expected market structure to change from rewarding volume to rewarding quality. For recent reviews of the literature on extension service, see Macours (2019), Magruder (2018), and Takahashi, et al. (2020).

7. Discussion

We interpret our results as demonstrating a proof of concept: improving smallholders' access to markets where high-quality produce is rewarded and more generally linking farmers to value chains has a large potential. Our work, however, also highlights challenges. First, while farmers increased their use of modern inputs – one pathway to increased productivity – adoption of these technologies remained low. A number of promising, and potentially interlinked constraints, focusing on the supply side, have recently been studied in the literature (see review in de Janvry, et al., 2017), including missing markets for risk (Karlan et al., 2014) and behavioral constraints (Duflo et al., 2011). Assessing the impact of coupling such supply interventions with market access or value chain inclusion is, we believe, an important area for future research.

The second challenge is operational. The market access intervention was implemented by an integrated domestic buyer to circumvent many of the potential agency- and information problems that plague the market for (lower quality) maize.[35] After factoring out all evaluation costs, the agro trading company broke even in two of the four buying seasons.[36] Adding farmer profit, joint surplus in these two seasons was therefore strictly positive. Three structural features of the product and the economy constrained the company's ability to increase revenues. First, as quality is more difficult to determine once the grain is milled, customers need to learn about higher quality through consuming it. As a consequence, it takes time to build up a reputation for high quality maize flour and a domestic customer base willing to pay a premium for it. Second, the price elasticity of quality among large segments of consumer market is low. This in turn is a consequence of low consumer awareness of the benefits of food safety and the inability of the government to publicize, test, or enforce quality standards at all stages of the value chain. Third, though quality maize can be exported at a (high) premium, a seller needs to incur large (fixed) costs (related to establishing contacts with international buyers and producing at the necessary scale) to enter the export market. These fixed costs help explain why the formal export market is dominated by few large actors.

[35] As a reference, in high income countries, agricultural outputs are produced in closely aligned segments of the value chain by actors exploiting economies of scale. In Uganda and most of Sub-Saharan Africa, on the other hand, markets for agricultural outputs are segmented, and the low quality segment is characterized by a multitude of layers of small actors. For example, maize in Uganda often passes through several sets of traders, or aggregators, before reaching mills located in urban centers (Daly et al., 2016).

[36] The loss in the last season was caused by the lockdown, and specifically the closing of boarding schools (one of the company's main customers), following the COVID-19 pandemic.

Other features of the business model raised costs. Specifically, the company's business model was not one of pure profit-maximization. Unlike other vertically integrated firms on the market, it offered to buy maize from smallholders managing to sell maize of high quality, rather than identifying areas with largeholder farmers producing relatively high quality maize. While this strategy raised costs and hence decreased company profits, it was likely crucial to achieving the large increases in farmer surplus we document. That is, the market intervention had such large positive impacts on income precisely because it provided market access to the poorest farmers who are currently excluded from global val ue chains.

Overall, the limited profitability of the vertically integrated model used here, and the revenue and costs constraints the company faced, provide important clues as to why market integration of large swathes of the rural population, and for many of the agricultural products they produce, is challenging – despite its potential. On the other hand, if a market for quality maize that smallholders could access is not financially viable, one could consider using subsidy money to generate such a market. While we cannot properly compare the costs and benefits of market access to the various multifaceted programs to help the very poor (see e.g. Bandiera et al., 2017; Banerjee et al., 2015), the effects on income we document suggest a market access program is a candidate worth investigating more closely.

References

ATAI. 2018. “Evidence for transformation: Framing a research agenda in agriculture for development.” CEGA, UC Berkeley.

Antras, Pol. (2015). Global production: Firms, contracts, and trade structure. Princeton University Press.

Ashraf, Nava, Xavier Giné and Dean Karlan. 2009. “Finding Missing Markets (And a Disturbing Epilogue): Evidence from an Export Crop Adoption and Marketing Intervention in Kenya”. American Journal of Agricultural Economics vol. 91 (4): 973-990.

Atkin, David, Amit K. Khandelwal, and Adam Osman. 2017. “Exporting and firm performance: Evidence from a randomized experiment.” The Quarterly Journal of Economics 132(2):551-615.

Bernard, Tanguy, Alain de Janvry, Samba Mbaye, and Elisabeth Sadoulet. 2017. “Expected Product Market Reforms and Technology Adoption by Senegalese Onion Producers.” American Journal of Agricultural Economics vol. 99 (4): 1096-1115.

Blouin, Arthur and Rocco Macchiavello.2019. Strategic default in the international coffee market. The Quarterly Journal of Economics, 134(2):895-951.

Bold, Tessa, Selene Ghisolfi, Frances Nsonzi, and Jakob Svensson. 2020. “Extension service and quality upgrading: experimental evidence from smallholder farmers in Uganda.” Mimeo, IIES, Stockholm University.

Burlig, Fiona, Louis Preonas, and Matt Woerman. 2020. “Panel data and experimental design.” Journal of Development Economics, vol. 144: 102458.

Casaburi, Lorenzo and Reed Tristan. 2020. “Interlinked Transactions and Competition: Experimental Evidence from Cocoa Markets.” Working Paper.

de Almeida, Luis, Robert Williams, Dirce M. Soares, Harry Nesbitt, Graeme Wright, and William Erskine. 2019. “Aflatoxin levels in maize and peanut and blood in women and children: The case of Timor-Leste.” Scientific Reports vol. 9, 13158.

de Janvry Alain and Elisabeth Sadoulet. 2019. “Transforming developing country agriculture: Removing adoption constraints and promoting inclusive value chain development.” FERDI Working paper P253.

de Janvry, Alain, Elisabeth Sadoulet, and Tavneet Suri. 2017. “Field Experiments in Developing Country Agriculture”. Handbook of Economic Field Experiments. Vol. 2. 427-466.

De Loecker, Jan and Pinelopi Koujianou Goldberg. 2014. “Firm Performance in a Global Market.” Annual Review of Economics, 6(1): 201–227.

Deutschmann, Joshua W., Maya Duru, Kim Siegal, and Emilia Tjernström. 2019. “Can Smallholder Extension Transform African Agriculture?” NBER Working Paper No. 26054.

Duflo, Esther, Michael Kremer, and Jonathan Robinson. 2007. “Why Don’t Farmers Use Fertilizer? Experimental Evidence from Kenya.” Unpublished.

Duflo, Esther, Michael Kremer, and Jonathan Robinson. 2011. “Nudging farmers to use fertilizer: theory and experimental evidence from Kenya”. American Economic Review 101 (6), 2350–2390.

Daly, Jack, Danny Hamrick, Gary Gereffi, and Andrew Guinn. 2016. Maize value chains in East Africa. Report F-38202-RWA-1, IGC, London.

Hanna, Rema, Sendhil Mullainathan, and Joshua Schwartzstein. 2012. “Learning Through Noticing: Theory and Experimental Evidence in Farming”. NBER Working Paper 18401.

Hansman, Christopher, Jonas Hjort, Gianmarco León-Ciliotta, and Matthieu Teachout. 2020. “Vertical Integration, Supplier Behavior, and Quality Upgrading among Exporters.” Journal of Political Economy, 128(9): 3570-3625.

Hell, Kerstin, Pascal Fandohan, Ranajit Bandyopadhyay, Sebastian Kiewnick, Richard Sikora, Peter J. Cotty. 2008. “Pre- and postharvest management of aflatoxin in maize: an African perspective.” Mycotoxins: detection methods, management, public health and agricultural trade. Leslie, J. F., Bandyopadhyay, R., Visconti, A. (eds). CAB International.

Nguyen, Van Hoan. 2013. Technical specifications for maize. V 13.0. WFP viewed at: https://documents.wfp.org/stellent/groups/public/documents/manual_guide_proced/wfp258162.pdf

Hoffmann, Vivian, Samuel Mutiga, Jagger Harvey, Michael Milgroom, and Rebecca Nelson. 2013. “A Market for lemons: Maize in Kenya.” Working paper, University of Maryland.

Hanna, Rema, Sendhil Mullainathan, and Joshua Schwartzstein. 2012. “Learning Through Noticing: Theory and Experimental Evidence in Farming”. NBER Working Paper 18401.

Islam, Mahnaz and Sabrin Beg. 2020. “'Rule-of-Thumb' Instructions to Improve Fertilizer Management: Experimental Evidence from Bangladesh.” Economic Development and Cultural Change (forthcoming).

Karlan, Dean, Robert Osei, Isaac Osei-Akoto, and Christopher Udry. 2014. “Agricultural Decisions after Relaxing Credit and Risk Constraints,” Quarterly Journal of Economics 129, 597–652.

Macauley, Harold. 2015. Background Paper: Cereal Crops: Rice, Maize, Millet, Sorghum, Wheat. African Development Bank.

Macchiavello, Rocco and Josepa Miquel-Florensa. 2019. “Buyer-Driven Upgrading in GVCs: The Sustainable Quality Program in Colombia.” CEPR Working paper No. 13935.

Macours, Karen. 2019. “Farmers’ Demand and the Traits and Diffusion of Agricultural Innovations in Developing Countries.” Annual Review of Resource Economics vol. 11:1: 483-499.

Magruder, Jeremy R. 2018. “An Assessment of Experimental Evidence on Agricultural Technology Adoption in Developing Countries.” Annual Review of Resource Economics vol. 10: 299-316.

McKenzie, David. 2012. “Beyond baseline and follow-up: The case for more T in experiments.” Journal of Development Economics vol. 99(2): 210-221.

Morjaria, Ameet and Martin Sprott. 2018. Ugandan Arabica coffee value chain opportunities. IGC Report F-43410-UGA-1.

Ranum, Peter, Juan Pablo Peña-Rosas, and Maria Nieves Garcia-Casal. 2014. “Global maize production, utilization, and consumption.” Annals of the New York Academy of Sciences, 1312: 105-112.

Rosenzweig, Mark R. and Christopher Udry. 2020. “External Validity in a Stochastic World: Evidence from Low-Income Countries.” The Review of Economic Studies, vol. 87 (1): 343-381.

Sserumaga, Julius P., Alejandro Ortega-Beltran, John M.Wagacha, Charity K.Mutegi, and Ranajit Bandyopadhyay. 2020. “Aflatoxin-producing fungi associated with pre-harvest maize contamination in Uganda.” International Journal of Food Microbiology vol. 313, 108376.

Takahashi, Kazushi, Rie Muraoka, and Keijiro Otsuka. 2020. “Technology adoption, impact, and extension in developing countries’ agriculture: A review of the recent literature.” Agricultural Economics vol. 51: 31-45.

UBOS. 2019. Uganda National Household Survey 2016/2017 Report. Uganda Bureau of Statistics, Kampala, Uganda.

USDA. 2019. United States Department of Agriculture, (2019), Data retrieved from National Agricultural Statistics Service.

World Bank. 2007. World Development Report 2008. Agriculture for Development. Washington, DC.

World Bank. 2020. World Development Report 2020: Trading for Development in the Age of Global Value Chains. Washington, DC: World Bank.

Table 1. Maize quality: Summary statistics

Quality measure:	Mean	Std	Min	Max	Obs
Visually verifiable quality (#)	2.5	1.5	0	7.0	355
Lab verified quality (%)	25.9	33.9	4.3	100	104
Wet maize (%)	28.2	-	0	1	110
Grade 1: EAS (%)	0	-	-	-	103
Grade 2: EAS (%)	3.9	-	-	-	103
Grade 3: EAS (%)	10.7	-	-	-	103
Ungraded/reject: EAS (%)	85.4	-	-	-	103

Note. Unit of observation is a bag. Visually verifiable quality is the number of detected defects out of 10 in a bag of maize. Lab verified quality is grams of defects per 200g maize sample drawn from bags bought in the field (in %). Wet maize is a binary indicator for maize with a moisture level above 13%. Grade 1-3 and ungraded/reject maize are East African Quality Standard (EAS) classifications for maize quality, with grade 1 having the most stringent thresholds for defects. See supplementary appendix S3 for details on measurement.

Table 2. Returns to quality: quality outcomes in treatment and control groups

Variable	Treatment (1)	Control (2)	Difference (3)
Visually verifiable defects: bags for sale (# defects; max 10)	0.05 (.29) {267}	2.46 (1.51) {355}	-2.16 [.000] {622}
Visually verifiable defects: average (by household) (# defects; max 10)	0.05 (.25) {49}	2.23 (1.49) {50}	-1.99 [.000] {99}
Visually verifiable defects: random bag for sale (# defects; max 10)	0.04 (.20) {49}	2.37 (1.67) {49}	-2.07 [.000] {98}
Lab verified defects (%)	6.54 (1.79) {39}	32.2 (39.5) {43}	-19.7 [.000] {82}
Wet maize (0;1)	0 – {49}	0.140 – {50}	-0.162 [.002] {99}
Grade 1: EAS (%)	12.8 – {39}	0 – {43}	10.9 [.031] {82}
Grade 2: EAS (%)	71.8 – {39}	30.2 – {43}	40.3 [.000] {82}
Grade 3: EAS (%)	15.4 – {39}	14.0 – {43}	-0.92 [.915] {82}
Ungraded/reject: EAS (%)	0 – {39}	55.8 – {43}	-50.3 [.000] {82}

Note. Visually verifiable defects is the number defects (out of 10) verified by trained enumerators. Lab verified defects is grams of defects per 200g maize as measured in the lab. Specification: Columns (1) and (2) are mean outcomes, with standard deviations in parenthesis and number of observations in braces. Column (3) is difference in mean outcomes from OLS regressions with randomization strata (villages), with p-value on the null hypothesis of equal means in brackets.

Table 3. Returns to quality: treatment effects

Outcome variable:	Price	Price	Deductions	Net-price
Specification	(1)	(2)	(3)	(4)
Model	OLS	IV	OLS	IV
Treatment	2.95		0.18	
	(9.87)		(.63)	
	[.766]		[.781]	
Visually verifiable defects		-1.50		-0.50
		(5.04)		(.54)
		[.766]		[.927]
Constant	530.5	530.3	3.75	511.2
	(14.0)	(14.1)	(1.17)	(15.1)
	[.000]	[.000]	[.002]	[.000]
Observations	116	116	116	116
Households	94	94	94	94
R-squared	0.91		0.22	

Note. Unit of observation is a maize sale. Price (p) is in UGX per kilogram. Visually verifiable defects is average number of defects in a bag, collapsed at the sales level. Deductions is defined as $(y - z)/y$, where y is the weight of maize sold as measured by enumerators and z is the agreed upon (or by the buyer determined) sales volume, expressed in percent. Net-price is price net of any weight deductions; i.e., pz/y. Visually verifiable defects is the average number of defects (max 10) in a bag put for sale. All specifications include week fixed effects and randomization strata (villages). Clustered by household standard errors in parentheses and p-values in brackets.

Table 4. Minimum detectable effects

Variables	(1) *MDE: ANCOVA*	(2) *MDE: ANCOVA (% SD)*	(3) *MDE: ANCOVA (% MEAN)*
Maize acreage	0.65	32	29
Expenses (monetary)	51	33	48
Expenses (incl. own labor)	62	35	37
Harvest	682	35	36
Yield	104	27	13
Price	30	14	6
Harvest value	103	29	36
Profit (monetary expenses)	70	24	42
Profit (incl. own labor)	71	26	40

Note. Minimum detectable effects (MDE) in the ANCOVA model (equation (3)), with a power of 80% and significance level of 5%, based on data from control group. MDE as share of the standard deviation (column 2) and as share of the mean (column 3). Expenses (monetary) is expenses on inputs and hired labor. Expenses (incl. own labor) also includes family labor, valued at community specific wages. Harvest value includes own-produced consumption, valued at community-specific market values. Profit (monetary expenses) is the difference between harvest value and monetary expenses. Profit (incl. own labor) is the difference between harvest value and all costs (including own labor).

Table 5. Sample

	All	Treatment	Control
Baseline panel: household-season obs.	544	316	228
Baseline panel: households	189	110	79
Baseline panel: clusters	20	12	8
Follow-up panel: household-season obs.	677	391	286
Follow-up panel: households	180	104	76
Follow-up panel: clusters	20	12	8
Baseline & follow-up panel: household-season obs.	1,198	692	506
Baseline & follow-up panel: households	180	104	76
Baseline & follow-up panel: clusters	20	12	8

Note. Number of household-season observations, households, and clusters in the baseline and follow-up panels.

Table 6. Attrition

	(1) All	(2) Treatment	(3) Control	(4) Difference	(5) Obs.
Households attritted during follow-up	0.048	0.055	0.038	0.017	189
				[.547]	
Household-season re-survey rate: follow-up	0.940	0.940	0.941	-0.001	720
				[.963]	

Note. Share of households attritted and share of households not surveyed in follow-up seasons. Column (4) is difference in mean outcomes across assignment groups, with p-value on the null hypothesis of equal means in brackets. Standard errors are clustered at the village level.

Table 7. Baseline: summary statistics and balance

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
	Sample			*Means*		*Difference in means*	
Variable	Mean	Std.	Obs.	T	C	Coeff.	*p*
			Panel A. Household characteristics				
Main decision maker: female	0.18	0.39	189	0.19	0.16	0.026	.776
Main decision maker: completed primary school	0.41	0.49	189	0.42	0.39	0.026	.765
Main decision maker: days/month in agriculture	24.8	5.27	189	24.4	25.3	-0.962	.347
Number of household members:	6.15	2.57	189	6.15	6.15	-0.006	.987
of which below 5	0.96	0.92	189	0.98	0.92	0.058	.659
Distance to district capital (km)	29.8	9.53	189	29.5	30.2	-0.703	.873
Distance to main road (km)	12.1	8.30	189	12.2	11.9	0.325	.872
			Panel B. Farm enterprise characteristics				
Maize acreage	2.16	1.60	544	2.41	2.48	0.047	.900
Expenses (USD)	133.2	145.3	363	149.5	154.2	2.805	.937
Harvest (ton)	2.10	1.94	499	2.32	2.40	0.096	.844
Yield (ton/hectare)	2.09	1.04	499	2.25	2.15	0.101	.651
Share sold	0.82	0.24	499	0.85	0.84	-0.012	.699
Price kilogram (USD)	0.20	0.06	470	0.16	0.16	0.001	.796
Harvest value (USD)	433.6	426.2	498	363.7	375.8	40.33	.697
Profit I (USD)	277.0	303.5	362	212.9	221.6	24.05	.733
Joint balance test I							.964
Joint balance test II							.591
Joint balance test III							.631
Joint balance test IV							.795

Note. Households in the baseline panel sample. Panel A: measured at first baseline round. Panel B: pooled data over the three baseline rounds. Difference in means conditioning on season fixed effects in Panel B. Standard errors are clustered at the village level. Expenses is expenses on inputs and hired labor. Data on hired labor was not collected in season 1. Harvest value includes own-produced consumption, valued at community-specific market value. Profit I is the difference between harvest value and expenses. Standard errors are clustered at the village level. The joint balance tests report *p*-values from testing whether the baseline outcomes predict enrollment into treatment, with profit dropped due to collinearity: all household characteristics in test I; farm enterprise outcomes except expenses in test II (seasons 1-3; sample size 470); all farm enterprise outcomes in test III (seasons 2-3; sample size 336); all household characteristics and enterprise outcomes in test IV (first season for household characteristics and season 3 for farm enterprise outcomes; sample size 170).

Table 8. Impact on investment

Specification	(1) Expenditure on seeds & fertilizer	(2) Expenditure on all inputs	(3) Any improved seeds	(4) Any fertilizer	(5) Pre-harvest expenses	(6) Post-harvest expenses	(7) Post-harvest expenses (labor only)	(8) Post-harvest expenses other costs)
Access to a market for quality maize	2.34 (.048) [.005]	4.33 (.060) [.015]	0.038 (.246) [.273]	0.031 (.144) [.159]	16.3 (.275) [.296]	5.94 (.255) [.271]	5.87 (.143) [.153]	0.47 (.797) [.804]
Observations	658	658	657	657	464	464	464	464
R-squared	0.31	0.33	0.08	0.03	0.20	0.26	0.22	0.22
Mean for control	3.76	13.27	0.13	0.03	53.79	30.38	15.64	14.73

Note. Clustered-by-village standard errors with *p*-values in parenthesis. *p*-values from Fisher-permutations test based on 10,000 permutations of the treatment assignment in brackets.

Table 9. Impact on post-harvest quality upgrading activities

Specification	(1) Proper drying	(2) Maize sorted	(3) Maize winnowed
Access to a market for quality maize	0.24 (.000) [.001]	0.14 (.002) [.001]	0.15 (.033) [.047]
Observations	640	464	464
R-squared	0.21	0.03	0.04
Mean for control	0.35	0.13	0.19

Note. Clustered-by-village standard errors with *p*-values in parenthesis. *p*-values from Fisher-permutations test based on 10,000 permutations of the treatment assignment in brackets.

Table 10. Impact on productivity and income

Specification	(1) Price	(2) Maize acreage	(3) Harvest	(4) Yield	(5) Harvest value	(6) Monetary expenses	(7) Profit (monetary expenses)	(8) Profit (incl. own hours)
				Panel A. Full sample				
Access to a market for quality maize	0.017 (.001) [.005]	0.046 (.829) [.838]	239.3 (.308) [.350]	111.7 (.036) [.044]	78.7 (.079) [.099]	18.3 (.305) [.321]	63.2 (.060) [.070]	97.8 (.028) [.030]
Observations	617	677	658	658	658	640	640	464
R-squared	0.70	0.27	0.29	0.09	0.32	0.33	0.23	0.18
Mean for control	0.15	2.29	1887.5	793.1	286.7	106.5	177.6	121.0
				Panel B. Trimmed sample				
Access to a market for quality maize	0.017 (.001)	0.008 (.967)	280.1 (.153)	106.4 (.030)	76.6 (.069)	14.6 (.278)	71.9 (.024)	84.9 (.020)
Observations	612	670	650	650	650	632	627	454
Mean for control	0.15	2.21	1755.9	779.6	267.36	96.8	167.5	120.7

Note. Clustered-by-village standard errors with *p*-values in parenthesis. *p*-values from Fisher-permutations test based on 10,000 permutations of the treatment assignment in brackets. Trimmed sample: for non-negative continuous variables we trimmed the top 1% observations in each season. For variables that can take both positive and negative values, we trimmed the top- and bottom 1% observations in each season. Binary outcome variables are not trimmed.

Table 11. Mechanisms

Specification	(1) Linear model	(2) Log-linear model
Panel A		
$E[Y_1 - Y_1]$	328.0	0.117
	[.027]	[.050]
Decomposing $E[Y_1 - Y_1]$		
ΔTFP: $\Delta\theta/E[Y_1 - Y_1]$	0.562	0.870
	[.001]	[.000]
ΔLand: $\Delta A/E[Y_1 - Y_1]$	0.192	-0.041
	[.024]	[.036]
ΔFamily labor: $\Delta L_F/E[Y_1 - Y_1]$	0.042	-0.051
	[.645]	[.218]
ΔHired labor: $\Delta L_H/E[Y_1 - Y_1]$	0.167	0.153
	[.231]	[.103]
ΔSoil enhancing investment: $\Delta\mu/E[Y_1 - Y_1]$	0.037	0.068
	[.466]	[.234]
ΔMeasured inputs total: $\Delta x/E[Y_1 - Y_1]$	0.438	0.130
	[.009]	[.208]
Panel B		
Test of independence	0.39	1.98
	[.852]	[.129]
Test I of mediators	139.8	168.1
	[.000]	[.000]
Test II of mediators	13.2	10.3
	[.021]	[.066]
Observations	383	386

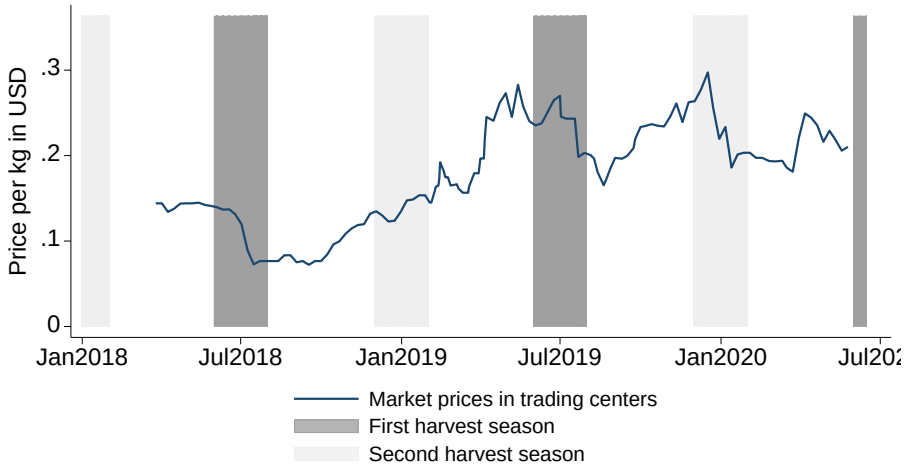
Note. Sample: households in the last three follow-up seasons. Panel A: overall treatment effect and the relative magnitudes of measured inputs and unmeasured inputs (TFP). Panel B: test of independence test the null hypothesis that the experimentally-induced increments in unmeasured inputs are independent of the experimentally induced increments in measured inputs, assuming autonomy. Assuming independence of the observed and unobserved inputs (in both treatment and control), the test is equivalent to testing autonomy; i.e., the parameters of the production function are the same in the control and the treatment group (see Heckman and Pinto, 2015). Test I of mediators test the H_0 that the vector of observed inputs do not explain harvest volumes. Test II of mediators test the H_0 that the vector of observed inputs are not affected by treatment.

Table 12. Trader prices and market shares

Outcomes	Estimate [p-value]
Panel A. Difference in market shares and prices	
Difference in market shares of local traders: Δms^{LT}	-0.308 [.000]
Difference in market shares: commercial traders: Δms^{CT}	-0.081 [.213]
Difference in local trader prices vs. control mean: $(\bar{p}^{1,LT} - \bar{p}^{0})$	0.056 [.039]
Difference in commercial trader prices vs. control mean: $(\bar{p}^{1,CT} - \bar{p}^{0})$	-0.021 [.580]
Panel B. Causal difference in prices: T vs. C	
Difference in local trader prices: Δp^{LT}_{causal}	0.082 [.009]
Difference in commercial trader prices: Δp^{CT}_{causal}	0.002 [.956]

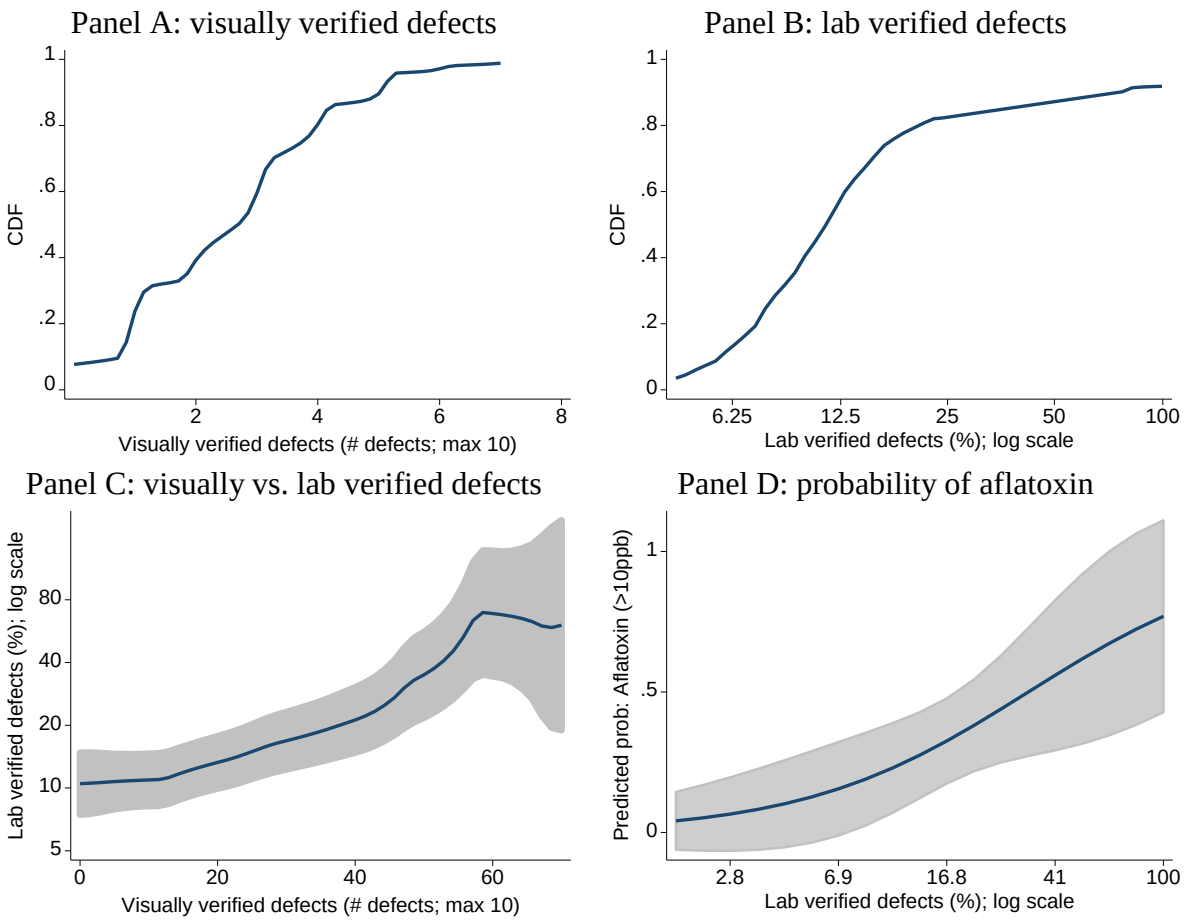
Note. See text for details. The estimates in panel A are from an estimation of the system of regressions (10)-(11) (see section 6.4.6). Unit of observation is household-sale (844 observations). Clustered-by-village standard errors with *p*-values in brackets. The estimates in panel B adjust for the selection effect (see section 6.4.6). Their *p*-values (in brackets) are obtained by jointly estimating the regression in panel A and the selection effect (Table A7, Panel B, columns (1) and (3)).

Figure 1. Market prices



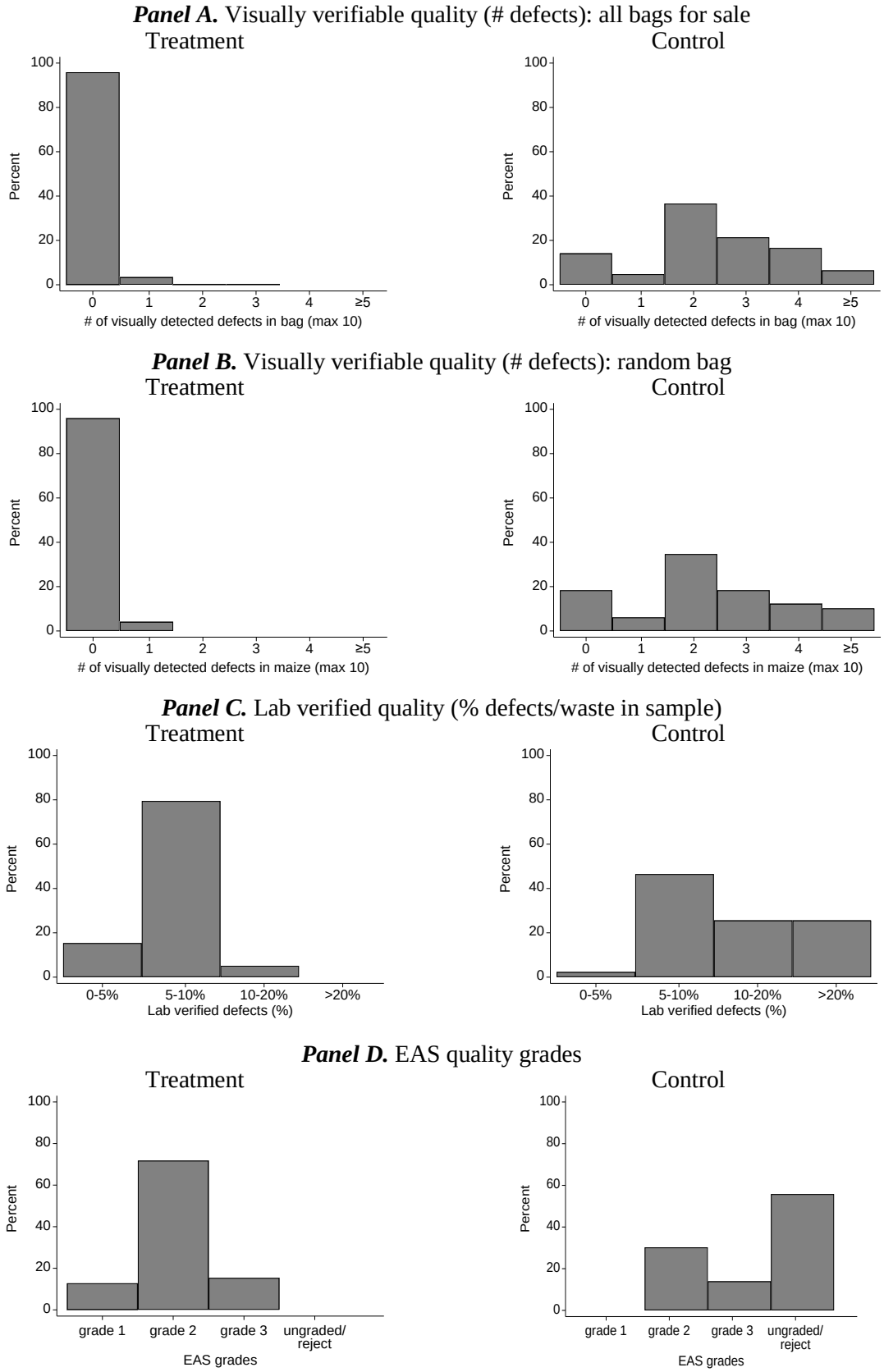
Note. Data sources: Market prices in trading centers in the study are collected by the research team. Average price from five trading centers.

Figure 2. Maize quality and verifiability of quality



Note. Panel A: CDF of visually verifiable defects. Panel B: CDF of lab verified defects (log scale). Panel C: Smoothed values from a local polynomial regression of lab verified defects on visually verifiable defects. Grey shaded area represents the 95% confidence interval. Panel D: Predicted probability of aflatoxin above the UNBS cut-off as a function of share of defects in the bag (lab verified defect), from specification (3), Table A2, with the grey shaded area representing the 95% confidence interval. The unit of observations is a bag. See supplementary appendix S2 for details on the tests and samples used.

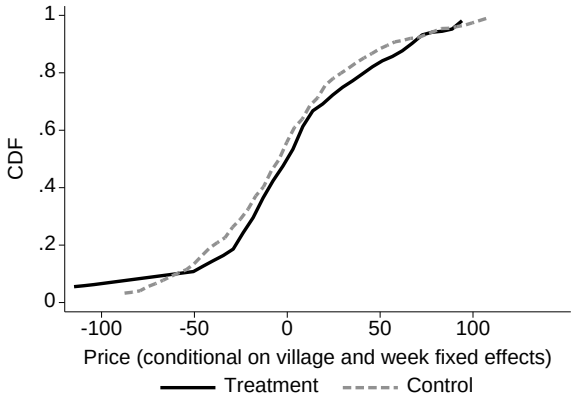
Figure 3. Returns to quality: quality outcomes in treatment and control



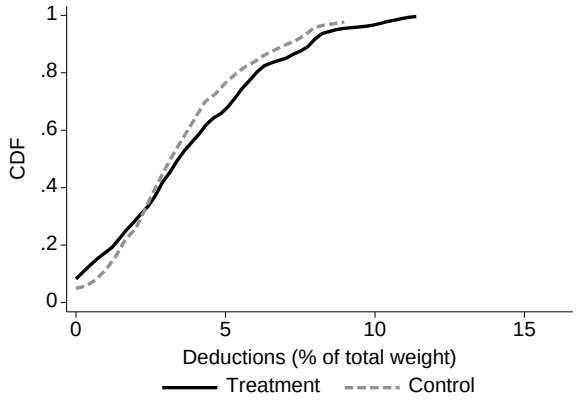
Note: Visually verifiable quality is the number of visually detected defects (max 10) in maize bags: Panel A depicts the average across all bags and Panel B depicts results for the randomly selected bag tested for quality in the lab. Lab verified quality (Panel C) is percent of defects and waste in bags. Panel D reports grades based on the East African Quality Standard (EAS) classification. See appendix S3 for details on the testing protocols.

Figure 4. Returns to quality

Panel A. Time of sale and price

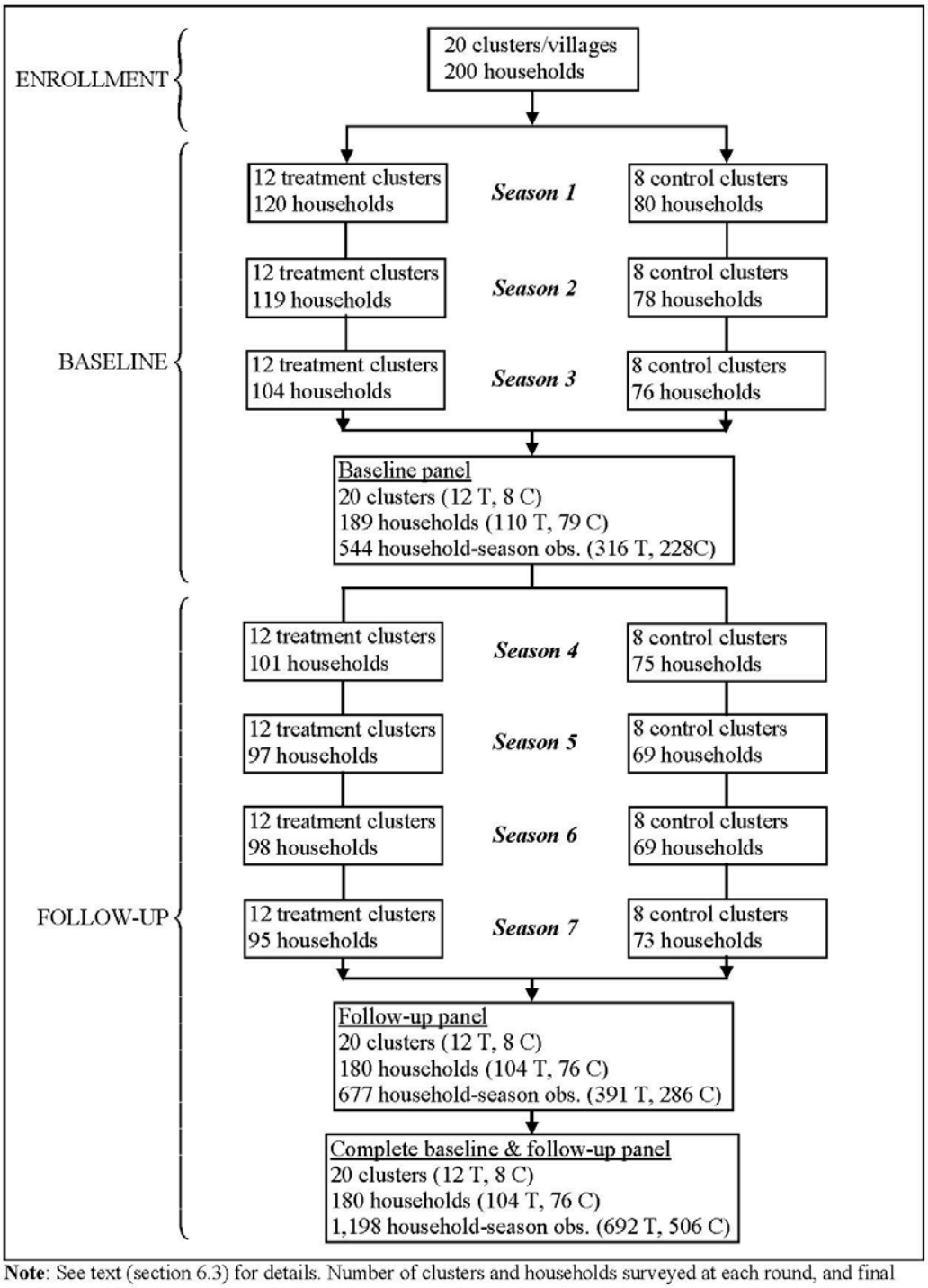


Panel B. Deductions



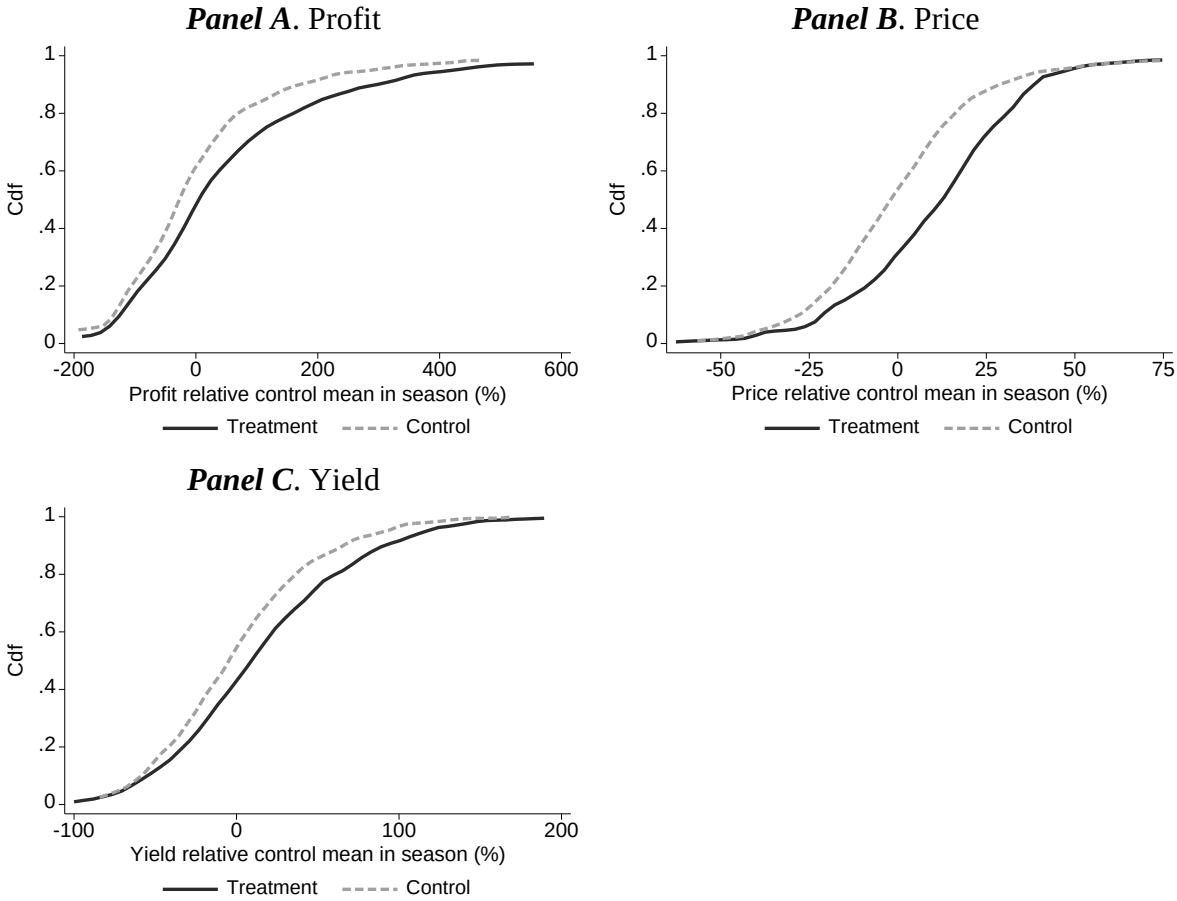
Note. Panel A: cumulative distribution functions of price (conditional on village and week fixed effects) in the assignment groups. Panel B: cumulative distribution functions of deductions in the assignment groups, with deduction defined as $(y - z)/y$, where y is the weight of maize sold as measured by enumerators and z is the agreed upon (or by the buyer stated) sales volume. The Kolmogorov-Smirnov D statistic on the test of equality of the treatment and control distributions is 0.11 [.84] in Panel A and 0.17 [p = .39] in Panel B.

Figure 5. Trial design: access to market for quality maize



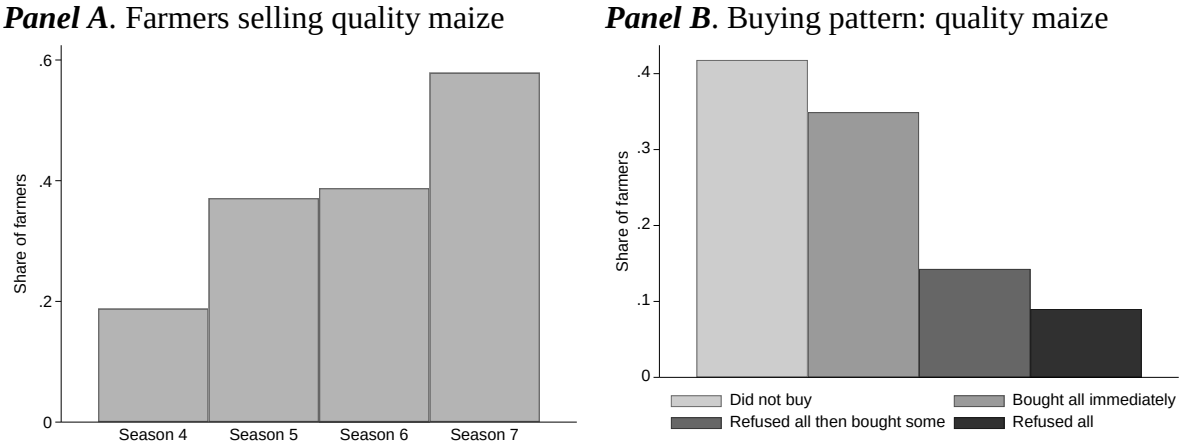
Note: See text (section 6.3) for details. Number of clusters and households surveyed at each round, and final baseline and follow-up panels with of clusters, households, and household-by-season observations.

Figure 6. Effect of access to a market for quality maize on profit



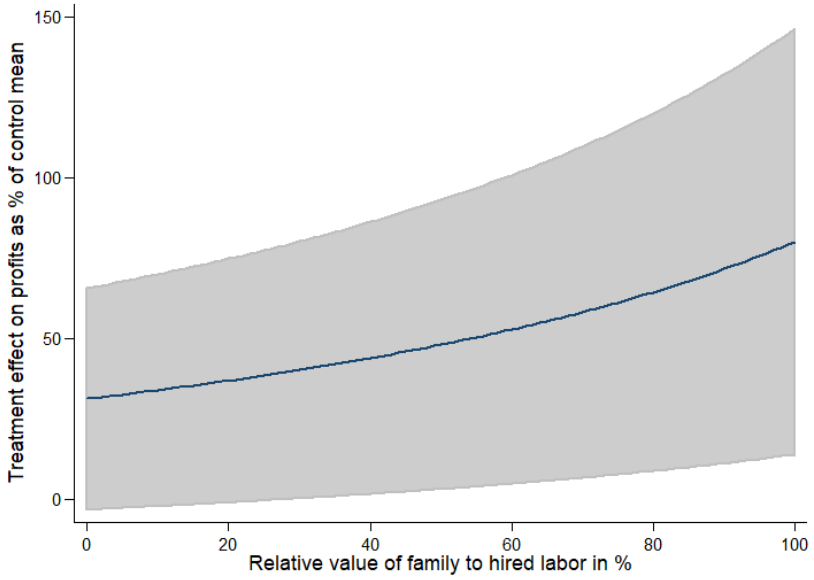
Note. All outcomes are expressed as percent of the control group mean by season. The Kolmogorov-Smirnov D statistic is 0.17 (p = .000) for profit, 0.34 (p = .000) for price, and 0.14 (p = .002) for yield.

Figure 7. Quality upgrading: agro-company's interactions with the farmers



Note. Data sources: agro trading company (trial sample households).

Figure 8. Profit as a function of the value of family labor



Note. The y-axis shows the treatment effect on profits expressed as percentage of the control mean (solid blue line) and its 95% confidence interval (grey shaded area) as a function of the relative value of family to hired labor on the x-axis, which ranges from 0 to 100%.

For Online Publication

Supplementary appendix

Market Access and Quality Upgrading: Evidence from Three Randomized Experiments

Tessa Bold,[1] Selene Ghisolfi,[2] Frances Nsonzi,[3] Jakob Svensson[4]

[1] IIES, Stockholm University, Stockholm, Sweden; CEPR (tessa.bold@iies.su.se).
[2] Leap, Bocconi University, Italy (ghisolfi@unibocconi.it).
[3] Department of Food Technology and Nutrition, Makerere University, Kampala, Uganda; PNDK Solutions, Kampala, Uganda (nsonzif@gmail.com).
[4] IIES, Stockholm University, Stockholm, Sweden; BREAD; CEPR (jakob.svensson@iies.su.se).

S1. Context

To describe the village markets for maize, we use data from the control group in the access to a market for quality maize experiment (see sections 3 and 6). Households here were surveyed at the end of the season for seven consecutive seasons. Detailed data on who farmers sold to were collected in the last five seasons. In total, the sample consists of 420 sales observations from 335 households-by-season observations from 78 households over five seasons.

Let s_{ijt} denote the share of the total maize sold by household i in season t to seller j. We define the market share of the type j seller in season t, as $ms_{jt} = \sum_i s_{ijt}$. The normalized price per kilogram of maize sold defined as $p_{it} = p_{it}^n / \bar{p}_t^0$ where p_{it}^n is the nominal price and $\bar{p}_t^0$ is the average price in season t. Table A1 provides summary statistics.

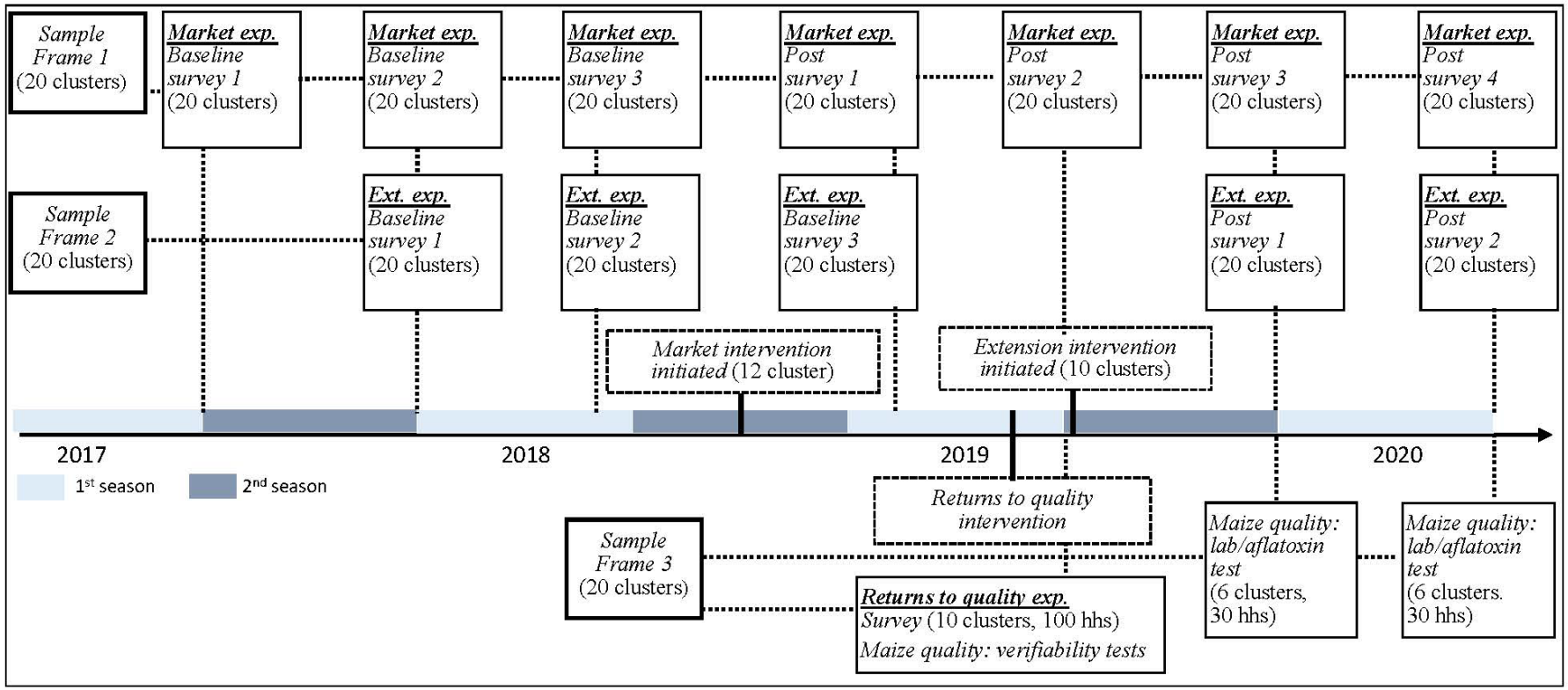
Table A1. Local and commercial traders

Variable	Mean
Interactions: local vs. commercial traders	
Market share of local traders (%)	78.6
Market share of local traders: season 3 (%)	68.4
Market share of local traders: season 4 (%)	82.1
Market share of local traders: season 5 (%)	82.6
Market share of local traders: season 6 (%)	79.4
Market share of local traders: season 7 (%)	78.6
Sold to a commercial trader at least once in five seasons (%)	51.3
Sale pattern over 5 seasons	
Sold once in the season (%)	78.8
Sold twice in the season (%)	17.3
Sold three or more times in the season (%)	3.9
Sold to one buyer only in the season (%)	89.6
Sold to two buyers in the season (%)	9.6
Sold to three buyers in the season (%)	0.9
Repeated interactions with local buyers (five seasons)	
Sold to the same buyer in all seasons (%)	12.2
Sold to the same buyer in four out of five seasons (%)	18.4
Sold to the same buyer in three out of five seasons (%)	32.7
Sold to the same buyer in two out of five seasons (%)	34.7
Sold to different buyers in each season (%)	2.0
Prices	
Normalized price paid by local traders	1
Normalized price paid by commercial traders	1.08

Note: See main text for details.

S2. Sample frames and research design

Figure A1. Design of the program



S3. Measuring the quality of maize at the farmgate

A. Visually verifiable quality (defects)

Maize grain was assessed from 99 sampled (using Sample Frame 3 – see Figure 2) smallholder commercial maize farmers (of which 50 were control households) at the time and point of sale. The mean number of bags in a lot of maize was 7 bags. Each bag in a lot was analyzed. Maize grain samples (300-350 g) were drawn from the top, middle and bottom of each bag with a grain sampling spear. Each sample was visually checked by trained enumerators for the defects listed in the East African Standard on Maize grain (East African Community, 2011). The following 10 defects (using a binary score; observed=1, not observed=0, were recorded: dirty grain, cobs, stones, dust, insects (live or dead), and broken, immature, damaged, rotten, and moldy grain. The moisture content in the bag was also determined using a portable grain moisture meter (AgraTronix MT-16).

B. Lab verified quality (defects)

One randomly selected bag was purchased from the 99 households described above, of which 82 were tested in the lab. In addition, one randomly selected bag each was purchased from another 30 households sampled from Sample Frame 3 (see Figure 2), and surveyed over two seasons. In total, 142 samples were tested at the PNDK lab in Kampala.

Three samples (300-350 g) were each drawn with a grain sampling spear from the top, middle and bottom of the sampled bag and thoroughly mixed to make one representative sample of the bag (total weight: 1000 g). Samples were analyzed using the methodology detailed in the Technical specifications for maize of the World Food Program.[5] A sub-sample of 200 g of maize was weighed into a glass beaker and sorted over a 4.5mm round hole sieve. The sieve was placed over a plastic basin to collect the small-sized particles. The broken grains, immature and shriveled grains, some foreign matter, and some inorganic matter e.g. stones, passed through the sieve due to their small particle size. They were each hand sorted into separate plastic containers and weighed. Pest-damaged grains, rotten and diseased grains, large stones, some foreign matter, some inorganic matter, and discolored grains remained on the sieve. They were each hand sorted into separate plastic containers and weighed.

C. Detection of Aflatoxin at 10 ppb using the AflaCheck® Mycotoxin Testing

Sub-samples of maize tested for the amount of defects were also tested for aflatoxin using the AflaCheck® mycotoxin testing system in accordance with the manufacturer's recommendations at the PNDK lab in Kampala.

About 500 g of hand sorted maize grain sample was ground to a fine flour using a three-step process: (i) mechanical grinding the maize grain to a coarse flour; (ii) pulverization to a fine flour; and (iii) sieving the flour to retain only the fine maize flour. The fine flour sample was packed in plastic containers, stored at room temperature and analyzed within 24 h.

A finely ground maize flour sample (5.00 g) was measured into an extraction tube to which 10 mL of 70% methanol (v/v) was added using a 10 mL measuring cylinder to test for aflatoxin at 10 ppb. The Extraction Tube was covered and shaken thoroughly for about 2 min. Thereafter the sample suspension was allowed to settle for about 5 min.

Strip test dilution tubes (1 mL vials) were placed in a paper strip test rack. 250 µL of distilled water were added to the dilution tubes with a 250 µL strip test pipettor. 250 µL of the sample

[5] Nguyen (2013).

supernatant in the extraction tube were then pipetted into this strip test dilution tube and the solution thoroughly mixed.

To test for aflatoxin, an AflaCheck® Strip was added to the strip test dilution tube containing the solution. The test was allowed to develop for about 10 min. A negative result for aflatoxin at the cut-off level being tested (< 10 ppb) was determined when both the test line and control line were visible after 10 min. A positive result for aflatoxin at the cut-off level being tested (≥ 10 ppb) was determined when no test line was visible after 10 min.

D. East African Quality Standard (EAS) classification

The East African Quality Standard (EAS) classifies maize into three broad quality categories based on moisture level and amount of non-grain substances and defected grain: graded maize, under-grade maize and reject maize. Graded maize (quality maize) is further categorizes into three grades: grades 1, 2 and 3, with grade 1 having the most stringent thresholds for defects. .

For grade 1 maize the thresholds are: moisture levels no higher than 13% and a maximum of 4.1% non-grain substances and defected grain. For grade 2 maize the thresholds are: moisture levels no higher than 13% and a maximum of 8.6% non-grain substances and defected grain. For grade 3 maize the thresholds are: moisture levels no higher than 13% and a maximum of 10.9% non-grain substances and defected grain. Maize that does not meet the criteria of grades 1, 2 and 3 and is not rejected is considered under-grade. Under-grade maize can in principle be sorted or treated for either grade 1, 2 or 3. Reject maize is maize which does not meet the criteria of grades 1, 2 and 3 with high levels or more severe levels of defects.

S4. Correlations between lab and visually verifiable defects and lab verified defects and probability of aflatoxin

Table A2.

Specification	(1)	(2)	(3)
Outcome variable:	Lab verified quality	Aflatoxin >10ng/g	
Visually verifiable quality	3.84		
	(.75)		
	[.000]		
Lab verified quality		0.10	1.08
		(.04)	(.55)
		[.014]	[.048]
Constant	1.87	-0.22	-6.75
	(.18)	(.09)	(2.15)
Observations	43	103	103
R-squared	0.31	0.50	

Note: OLS regressions (1)-(2), logit regression (3) with season fixed effects. Unit of observation is a maize bag. Lab verified quality is grams of defects per 200g maize (%), expressed in logs. Visually verifiable quality is the number of defects (out of 10) detected in the bag in the field. Aflatoxin >10ng/g is a dummy variable indicating an aflatoxin level above the limit imposed by the Uganda National Bureau of Standards (10ng/g). Robust standard errors are in parentheses and *p*-values in brackets.

S5. Returns to quality experiment

In the end of 2018, we enrolled 99 maize farming household from Sample Frame 3. After stratifying by village, the household were randomly assigned into a treatment and a control group of equal size. At enrollment, a short survey was administered. Table S3 compares pre-harvest outcomes between treatment and control group. None of the collected covariates show any statistically significant differences across treatment assignment and a χ^2-test of the differences between assignment groups across the collected covariates confirms that the samples are balanced.

Table S4 describes the trial sequence and samples used. At the first follow-up, all 99 households were revisited when they had bagged but not yet sold their maize and data on visually verifiable quality was collected (see section S2). One randomly selected bag was also purchased from each farmer and brought to Kampala for further quality testing. A bag was only purchased if the farmer has more than one bag to sell, which happened in 98 out of 99 cases. Not all bags purchased were tested in the lab because of administrative constraints. Specifically, bags bought for testing on a few occasions could not be handled according to the protocol since the field staff in charge of storing and transporting the bags for testing were occupied with the buying operation. As a result, one bag each from 82 households were tested in the lab.

The second follow-up (visits) took place after a household reported it had sold all or part of their maize. At this second follow-up, data on prices and sales volume were collected. In total, data from 116 sales by 94 households were recorded.

Table A3. Baseline balance: returns to quality experiment

Specification	(1)	(2)	(3)	(4)	(5)	(6)
Covariate	Acreage	Expected harvest	Expected to harvest in the 2nd half of season	Use modern seeds	Use of chemicals	Joint balance test
Treatment	-0.01	-0.16	0.01	0.07	-0.10	[.14]
	(.23)	(.18)	(.10)	(.06)	(.08)	
	[.95]	[.39]	[.91]	[.25]	[.22]	
Constant	2.79	1.04	0.99	0.06	0.84	
	(.39)	(.26)	(.06)	(.14)	(.14)	
Observations	99	99	99	99	99	99
R-squared	0.22	0.17	0.23	0.12	0.39	0.08

Note: OLS regressions with randomization stratas (villages). Robust standard errors in parenthesis and p-values in brackets. Specifications: (1) is acreage of land used for planting maize; (2) is expected harvest of maize (tons/acreage); (3) is a binary indicator taking the value 1 if the household expected to harvest their maize in the 2nd half of the season (i.e. in the first week of February 2019 or later); (4) used modern seed (hybrid or OPV seeds); (5) used chemicals (pesticides and/or herbicides); (6) joint balance tests report the p-value from jointly testing whether the variables in columns (1)-(5) predict enrollment into treatment.

Table A4. Sample: returns to quality experiment

Sample	All	Treatment	Control	Attrition rate T vs. C (households)
Enrolled: Households	99	49	50	
Follow-ups:				
I. Visual quality: # Households	99	49	50	
I. Visual quality: # Bags	622	267	355	
II. Lab quality: # Household & bag	82	39	43	-0.06 [.40]
III. Price & sales: # Households	94	47	47	0.02 [.67]
III. Price & sales: # Sales	116	60	56	

Note: Sample sizes. Attrition rate is the share of households, out of all enrolled, not surveyed/tested at follow-up.

S6. Premium for quality maize

Below we describe the framework used to determine the premium as a function of observable outcomes (amount of defects and prices in the trading centers).

Assume farmers can sell one unit (or bag) of maize of either low or high quality. High quality maize contains only non-defected maize kernels while a share α of a unit of low quality maize contains waste and defected kernels. Assume further that the cost of producing one unit of low quality maize is κ. Let the price of low quality maize be p_L. Then, if a farmer sells one unit of low quality maize, the farmer's profit is simply $p_L - \kappa$.

Consider now a profit maximizing buyer who wants to buy high quality maize at a price p_H. What is the minimum premium, $(p_H - p_L)$, which the buyer then needs to pay for high quality maize?

To solve this problem we make two assumptions: (i) the farmer can turn low quality maize into high quality maize by sorting away defects and waste; (ii) the cost of doing so is zero. These two assumptions imply that a farmer selling one unit of high quality maize would earn $p_H(1-\alpha) - \kappa$ and the farmer would be willing to do so if

$$p_H(1-\alpha) - \kappa \geq p_L - \kappa \tag{1}$$

The minimum price that needs to be offered is found when the participation constraint (1) binds; i.e., $p_H - p_L = \alpha p_H$. In other words, the minimum premium is the share of defected kernels and waste in low quality maize, valued at the premium price.

With estimates of p_L and α, one can determine p_H. However, while local village prices, and the difference in the share of defects in high vs. low quality maize, are in principle observable, they are observable with a lag (and for α require laboratory equipment). We therefore determined the premium based on variables we could observe in real time (i.e., prices in trading centers) and an assumption about α based on pre-treatment pilot data. Specifically, we continuously collected price information for maize from all nearby trading center (p^{TC}) and as local prices closely follow these trading prices, but are lower, we assume $p_L = \gamma p^{TC}$, where $\gamma < 1$. Further, and again based on pre-treatment pilot data, we predicted that maize with no visually verifiable defects, and a moisture level below 13%, would contain 10-15% waste and defected kernels, while average quality in the market was assumed to contain 25-28% waste and defects. That is, we assumed quality in the market was low but that farmers could, using traditional methods for drying, sorting, and cleaning, produce and sell maize of essentially grade 3 quality (using the East African Quality Standard (EAS) grading system), which the company could process further to grade 1 or 2 quality maize.[6]

Rewriting the participation constraint, and substituting for p_L, we can write the premium as a percentage of trading center prices, $\omega(p^{TC})$, as $\omega(p^{TC}) = (p_H - p^{TC})/p^{TC} = \frac{\gamma}{1-\alpha} - 1$. Based on pre-treatment pilot data, we set $\gamma = 0.9$; i.e., we assumed that local prices, on average, are 10% lower than prices in the trading centers. Further, we estimated α, the share of a unit of maize that farmers sort away when turning low quality into high quality maize, by calculating the difference in the midpoints of the two spans ($\alpha = 0.265 - 0.125 = 0.14$), yielding a

[6] The East African Quality Standard (EAS) for maize classifies maize into three broad categories (graded maize, under-grade maize and reject maize) based on moisture level and amount of defective kernels and waste. Graded maize (quality maize) is further categorizes into three grades: grades 1, 2 and 3, with grade 1 having the most stringent thresholds for defects. The thresholds for grade 3 maize are a moisture level ≤ 13% and total defects and waste ≤ 11%.

minimum premium relative trading center prices, $\omega(p^{TC})$, of 5% and a premium relative local prices, $\omega(p_L) = (p_H - p_L)/p_L = \frac{1}{\gamma}(\omega(p^{TC}) + 1 - \gamma)$, or approximately 15%.

The quality premium we choose should be viewed as a lower bound of a "market-based" quality premium for several reasons. First, in the model, the participation constraint binds. Second, if we relax the assumption that the cost of sorting and cleaning away waste and defected kernels is costless; for instance by letting the cost of producing high quality maize be $\beta\kappa$, with $\beta > 1$, the premium increases to $\alpha p_H + (\beta - 1)\kappa$. Third, we did not factor in that more waste, defected kernels and too high moisture levels increase the risk that the maize will become unsuitable for consumption, for instance due to too high levels of aflatoxin (see section 4). Fourth, our estimate of α turned out to be too low. Specifically, data from lab testing of bags bought by the company (see section 6.4.2. for details) show that the average bag bought by the company contained 8% of waste & defected kernels; i.e. the company bought on average grade 2 maize. As a comparison, evidence from our assessment of maize quality (see Table 1) shows that, on average, the bags tested for quality contained 26% waste and defected kernels. Setting $\alpha = 0.18$ yields a minimum premium of 10% rather than 5% as we used. Finally, as discussed in detail in section 6.4.6., local traders in the treatment villages offered higher prices in response to the entry of the quality buyer, resulting in a de facto lower premium relative local prices.

S7. Farm enterprise characteristics at each baseline round

Table A5. Farm enterprise characteristics at baseline: summary statistics

Specification	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
	Mean				*Standard deviation*				*Obs.*
Variable	Season 1	Season 2	Season 3	Pooled	Season 1	Season 2	Season 3	Pooled	
Maize acreage	1.75	2.27	2.44	2.16	1.57	1.38	1.76	1.60	544
Expenses (USD)		115.6	151.5	133.2		116.5	168.6	145.3	363
Harvest (ton)	1.87	2.03	2.35	2.10	1.60	1.77	2.29	1.94	499
Yield (ton/hectare)	1.98	2.07	2.21	2.09	0.96	1.04	1.09	1.04	499
Share sold	0.88	0.76	0.84	0.82	0.14	0.30	0.21	0.24	499
Price kilogram (USD)	0.25	0.21	0.16	0.20	0.06	0.04	0.02	0.06	470
Harvest value (USD)	495.1	450.4	368.8	433.6	495.7	425.6	357.1	426.2	498
Profit I (USD)		334.8	216.5	277.0		349.6	232.2	303.5	362

Note: Households in the baseline panel sample. Expenses is expenses on inputs and hired labor. Data on hired labor was not collected in season 1. Harvest value includes own-produced consumption, valued at community-specific market value. Profit I is the difference between harvest value and expenses. Obs. is number of observations in the pooled sample.

S8. Local market equilibrium effects

Baseline

We collected data on whom the farmers sold to from the third season and forward. Table A6 reports market shares and average prices paid by local traders and commercial traders in season 3; i.e., the last baseline season.

Table A6. Market shares and prices at baseline

	(1)	(2)	(3)	(4)
	Means		*Difference in means*	
Variable	T	C	Coeff.	*p*
Market shares:				
Local traders	0.64	0.72	-0.085	.338
Commercial traders	0.36	0.28	0.085	.338
Prices:				
Local traders	560.9	568.6	-7.69	.550
Commercial traders	565.3	573.9	-8.60	.705

Note. Market shares and prices are derived from household sales data in season 3 (last baseline season). p-values (*p*) with standard errors clustered at the village level.

Selection into selling to local and commercial traders

In Table A7, panel A, we use sales data to predict normalized price in treatment in the last baseline season for a sale to a trader of type $j = LT, CT$ as a function of the number of follow-up seasons s the farmer from whom the sale originated sold to the high quality trader. In columns (1) and (3), we regress prices on a set of dummies for the number of follow-up seasons s the farmer sold to the high quality trader. In columns (2) and (4), we regress prices on the number of follow-up seasons s a farmer sold to the high quality trader. In Table A7, Panel B, we combine the regression coefficients from Panel A with estimates of the share of sales to either local or commercial traders at follow-up that originate from farmers who sold to the high quality trader s times at follow-up, $\widehat{\Pi}_i\left(\sum_t w_{k,t}^{HT} = s \middle| d_i = 1, z_{i,t}^{HT} = 0\right)$, to calculate the selection into selling to local and commercial traders in treatment villages at follow-up.

Intensive vs. extensive margins

We use sales data collapsed at the trader-market-season level to test for selective exit of traders. Specifically, we estimate a linear probability model with exit of local trader as dependent variable and price at baseline (season 3), a treatment indicator, and an interaction term, as independent variables. The results are reported in Table A8.

Entry

We use sales data collapsed at the trader-market-season level to measure the price paid by incumbent local traders and new entrants. That is, we here exploit variation within the group of local traders. Note that since we have no data on name and type of traders in the first two baseline seasons, we cannot rule out that (some) traders in the new entrants group are strictly speaking not new entrants. As reported in Table A9, the difference in prices paid by incumbent local traders in the treatment ($p^{1,LTI}$) vs. the control ($p^{0,LTI}$) group is 6.0 pp (p = .179). The differences in prices paid by new entrants in the treatment ($p^{1,LTN}$) vs. the control ($p^{0,LTN}$) group is 8.2 pp (p = .033). We cannot reject that these price differences are the same for incumbents and new traders in the treatment group: $(p^{1,LTI} - p^{0,LTI}) - (p^{1,LTN} - p^{0,LTN}) = 0.022$ (p = .686).

Table A7. Testing for selection

Specification	(1)	(2)	(3)	(4)
	Price received in last baseline round in treatment			
	Sale to local trader		Sale to commercial trader	
Panel A				
Sold to high quality buyer in:				
0 follow-up seasons	-0.025		-0.000	
	(.023)		(.051)	
	[.266]		[.995]	
1 follow-up season	-0.086		-0.100	
	(.030)		(.063)	
	[.004]		[.111]	
2 follow-up seasons	0.024		-0.029	
	(.029)		(.063)	
	[.422]		[.641]	
3 follow-up seasons	0.005		0.069	
	(.039)		(.081)	
	[.892]		[.398]	
4 follow-up seasons	0.039		0.070	
	(.039)		(.079)	
	[.320]		[.381]	
Times sold to high quality buyer		0.019		0.023
		(.010)		(.021)
		[.060]		[.263]
Panel B				
Selection into selling to trader of type *j*				
Δp^{j}_{select}	-0.025	-0.021	-0.023	-0.017
	(.210)	(.298)	(.495)	(.607)
Observations	85	85	48	48
R-squared	0.12	0.04	0.08	0.03
Mean for control	0.000	0.000	0.000	0.000

Note. Panel A: The dependent variable is the normalized price in treatment in the last baseline season for a sale to a trader of type $j = LT, CT$. In columns (1) and (3) price for a sale to trader of type $j = LT, CT$ is regressed on a set of dummies for the number of follow-up seasons s a farmer sold to the high quality trader. In columns (2) and (4) price for a sale to trader of type $j = LT, CT$ is regressed on the number of follow-up seasons s a farmer sold to the high quality trader. Robust standard errors in parentheses and *p*-values in brackets. Price is normalized by the control group mean in the last baseline season. Unit of observation is sales in last baseline season in treatment to trader type j. Panel B: Δp^{j}_{select} is the selection effect of selling to a trader of type j, which is a non-linear combination of the coefficients estimated in Panel A and the share of sales to trader type j at follow-up that originate from farmers who sold to the high quality trader s times at follow-up, $\widehat{\Pi}_i\left(\sum_t w^{HT}_{k,t} = s \middle| d_i = 1, z^{HT}_{i,t} = 0\right)$ (see section 6.4.6 for details). Δp^{j}_{select} and its *p*-value are estimated by estimating the regression in panel A and the probabilities $\Pi_i\left(\sum_t w^{HT}_{k,t} = s \middle| d_i = 1, z^{HT}_{i,t} = 0\right)$ as a simultaneous system. .0703221 s.e
.0794187 and p 0.381

Table A8. Selective exit

Variable	Exit
p_j	-0.393
	[.196]
Market access	0.218
	[.029]
$p_j \times$ Market access	0.035
	[.953]
Constant	0.243
	[.007]
R^2	0.06
Observations	101

Note. Household-sale data collapsed at the local trader-village level. Linear probability model. Exit is a dummy variable indicating that the trader was active at baseline (season 3) but not in any follow-up season. p_j is (normalized by control group mean) price at baseline. Market access is the assignment to treatment indicator. Clustered-by-village standard errors with *p*-values in brackets.

Table A9. Prices paid by new vs. incumbent local traders

Outcomes	Estimate [p-value]
Incumbent local traders in treatment vs. control: $(p^{1,LTI} - p^{0,LTI})$	0.060
	[.179]
New local traders in treatment vs. control: $(p^{1,LTN} - p^{0,LTN})$	0.082
	[.033]
$(p^{1,LTI} - p^{0,LTI}) - (p^{1,LTN} - p^{0,LTN})$	-0.022
	[.686]

Note. Household-sale data collapsed at the local trader-village-season level for four follow-up seasons (328 observations). Price is normalized by the control group mean in each season. Clustered-by-village standard errors with *p*-values in brackets.

References

Nguyen, Van Hoen. 2013. Technical specifications for maize. V13.1. World Food Program. https://documents.wfp.org/stellent/groups/public/documents/manual_guide_proced/wfp261422.pdf. Viewed online on: 25 September 2018.

East African Community. 2011. East African Standard: Maize grains – Specification. East African Community, Arusha, Tanzania.