ONLINE APPENDIX

Market Access and Quality Upgrading: Evidence from Four 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] Catholic University of the Sacred Heart, Milan, Italy (email: selene.ghisolfi@unicatt.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).

A. 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 (see online Appendix G Table 1).

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 is 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 1 in Online Appendix G provides summary statistics.

B. Measuring the quality of maize at the farmgate

I. Visually verifiable quality (defects)

We assessed the quality of maize grain from 99 smallholder commercial maize farmers in nine villages in Sample Frame 3 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).

II. Lab verified quality (defects)

One randomly selected bag was purchased from the 99 households described above, of which 82 (43 from the control group) were tested in the lab. In addition, one randomly selected bag was purchased from each of 30 households sampled from an additional six villages in Sample Frame 3 (see online Appendix H, Figure 1), 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 4.5 mm 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.

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

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 (maize kernel including the pericarp) was ground to a fine flour using a three-step process: (i) mechanically 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 hours.

A finely ground maize 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

[5] Nguyen (2013).

the sample 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 within 5 min. A positive result for aflatoxin at the cut-off level being tested (≥ 10 ppb) was determined when only the control line was visible within 5 min.

IV. 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 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.

For grade 1 maize the thresholds are: moisture levels no higher than 13%, no live insects in the bag, and a maximum (when pooling the individual thresholds for non-grain substances and defective grain) of 4.05% non-grain substances and defective grain. For grade 2 maize the thresholds are: moisture levels no higher than 13%, no live insects in the bag, and a maximum of 8.6% non-grain substances and defective grain. For grade 3 maize the thresholds are: moisture levels no higher than 13%, no live insects in the bag, and a maximum of 10.85% non-grain substances and defective 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 to either grade 1, 2 or 3. Reject maize is maize that cannot be sorted or treated to grade 1, 2 or 3.

C. Returns to quality experiment

In December 2018, we enrolled, at the time of harvest, 99 maize farming households from nine villages in Sample Frame 3. Original households drawn into the sample who did not give consent to participate, or that could not be tracked, were replaced by households from a replacement list. The households were randomly assigned into a treatment and a control group of equal size. At enrollment, a short survey was administered. Table 3 in online Appendix G compares pre-harvest outcomes between treatment and control group.

Table 4 in online Appendix G 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. 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 had 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. As a result, one bag each from 82 households was tested in the lab. The attrition rates for bags tested from individuals were similar across assignment groups (see online Appendix G, Table 4)

The second round of 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. The attrition rates for sales data available from households were similar across assignment groups (see online Appendix G, Table 4).

D. Premium for quality maize

Below we describe the framework used to determine the minimum premium farmers need to be paid to produce high quality maize 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-defective maize kernels while a share α of a unit of low quality maize contains waste and defective 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 needs to pay in order for the farmer to be willing to produce/sell 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 financial 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 defective kernels and waste in low quality maize, valued at the premium price.

With estimates of p_L and α, one can determine the minimum price the farmer needs to be paid for high quality as that p_H, which solves the farmer's participation constraint (1) when it is binding, i.e. $p_H = p_L/(1-\alpha)$. 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 pretreatment pilot data. Specifically, we continuously collected price information for maize from all nearby trading centers (p^{TC}) and as local prices closely follow these trading prices, but are lower, we assume $p_L = \gamma p^{TC}$, where $\gamma < 1$. 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, 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 essentially grade 3 level maize; i.e., approximately 11% waste and defective kernels, while average quality in the market was assumed to contain 25% waste and defects, giving a value for α of 0.14. 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 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.

With these parameter values for α and γ, we can calculate the premium for high quality maize relative to local village prices as $(p_H - p_L)/p_L = (\frac{p_L}{1-\alpha} - p_L)/p_L = \frac{\alpha}{1-\alpha} = 0.16$ Relative to trading center prices, the premium is $(p_H - p^{TC})/p^{TC} = (p_H - \frac{p_L}{\gamma})/\frac{p_L}{\gamma} = (\frac{p_L}{1-\alpha} - \frac{p_L}{\gamma})/\frac{p_L}{\gamma} = (\gamma - (1-\alpha))/(1-\alpha) = 0.047$.

The average premium paid by the high quality buyer relative to commercial trader/trading centre prices was 5% across all seasons, which was equivalent to an average premium relative to expected local prices (which were estimated to be 10% below the trading centre price) of 15%. Ex-post (i.e. compared to local prices collected as part of the household survey at the same time), the premium amounted to 15%.

E. Mechanisms and TFP

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 could not measure, increased, and that market access affected how well a given bundle of inputs was used; i.e., it increased 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. To assess this possibility, we 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 products 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 the relative importance of the 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, which is normalized to one for simplicity. Finally, we assume that harvest volume, Y, is well-described by the following Cobb-Douglas production function:

$$Y = \theta L_F^{\alpha_1} L_H^{\alpha_2} E^{\alpha_3} , \tag{2}$$

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

Equation (2) can be estimated either as a log-linear or a linear model. As not all farmers use both own and hired labor, we focus attention on the linear version and write:

$$Y^d = \alpha_0^d + \alpha_1^d L_F^d + \alpha_2^d L_H^d + \alpha_3^d A^d + \alpha_4^d \mu^d + \theta^d + \tilde{\varepsilon}^d \ , \tag{3}$$

where superscript $d \in \{0,1\}$ denotes assignment groups (0 control, 1 treatment), , and $\tilde{\varepsilon}^d$ is a zero-mean error term assumed to be independent of the regressors x_j^d and θ^d.

To consistently estimate the parameters, $\boldsymbol{\alpha^d} = [\alpha_1^d, \alpha_2^d, \alpha_3^d, \alpha_4^d]$, of the measured inputs, $\boldsymbol{x^d} = \left[L_F^d, L_H^d, A^d, \mu^d\right]$, in (3) we need to assume that these inputs are independent of TFP (θ^d), given the treatment status. This is the key condition in the sequential ignorability assumption of Imai et al. (2010, 2011) that allows one to use experimental data to analyze mechanisms. 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 $\boldsymbol{\alpha}$. We cannot reject the null hypothesis of independence of the increments (see online Appendix G, Table 12 Panel B).[6] We can thus obtain consistent estimates of the impact of observed inputs on output by regressing output on the vector of measured inputs, $\boldsymbol{x^d}$, and a dummy for treatment assignment, δ^d,

$$Y^d = \delta^d + \boldsymbol{\alpha x^d} + \varepsilon^d \ , \tag{4}$$

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):

$$E(Y^1 - Y^0) = (\delta^1 - \delta^0) + \textstyle\sum_k \alpha_k \, E[x_k^1 - x_k^0] \ , \tag{5}$$

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.[7] The vector of observed inputs, $\boldsymbol{x}$, can explain the treatment effect on harvest volumes only if they affect harvest ($\boldsymbol{\alpha} \neq 0$) and, on average, are affected by the experiment, so that $E[x_k^1 - x_k^0] \neq \mathbf{0}$. Both these conditions hold in our experimental data (online Appendix G, Table 12, Panel B).

Online Appendix G, Table 12 Panel A reports the estimated treatment effects and the contributions of measured and unmeasured inputs (TFP). Output is 344 kilograms higher in and slightly more than half the increase in output comes through TFP (52%). Of the measured inputs, the largest contributions come from land and hired labor (which contribute 26% and 16%, respectively, to the increase in harvest). The contributing effects of family labor, fertilizer and hybrid seeds are also positive, albeit smaller. In sum, increases in measured inputs account for 48% of the treatment effect on harvest, while improvements in TFP account for the majority of the increase in output.

[6] Alternatively, if we are willing to assume independence of the observed and unobserved inputs (in both treatment and control), the test of H_0: $\widehat{\boldsymbol{\alpha}}^{\mathbf{1}} = \widehat{\boldsymbol{\alpha}}^{\mathbf{0}}$ 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).

[7] 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).

F. Spillover effects on sales to the local market

I. Estimating market shares and prices in levels and differences

To estimate trader- and assignment-specific market shares and prices as well as their changes between treatment and control, we use data on all sales in follow-up seasons in treatment and control. We define for sale i in follow-up season t an indicator $s_{j,i,t}$ for each trader type j, which takes on the value 1 if sale i in follow-up season t is made to trader type j and zero otherwise. We then regress these indicators on a constant and a dummy $D_{i,t}$ that takes the value 1 if sale i in period t takes place in the treatment group and zero otherwise

$$s_{j,i,t} = \alpha_j + \gamma_j D_{i,t} + \varepsilon_{i,t,j} \ , \tag{6}$$

running one regression per trader type $j = \{OT, LT, CT, HT\}$. The average market share for trader type j in control, $\bar{s}_{0,j}$, is then estimated as $\hat{\alpha}_j$ and the average market share for trader type j in treatment, $\bar{s}_{1,j}$, is estimated as $\hat{\alpha}_j + \hat{\gamma}_j$. The change in market shares equals $\Delta\bar{s}_j = \hat{\gamma}_j$.

To estimate the average price paid by other traders in treatment and its difference relative to control, we regress the price for sale i in follow-up season t on a constant and the interaction of the indicators $s_{j,i,t}$ for $j = OT, HT$ and the treatment dummy $D_{i,t}$:

$$p_{i,t} = \alpha + \sum_{j=OT,HT} \beta_j D_{i,t} \times s_{j,i,t} + \varepsilon_{i,t}. \tag{7}$$

The average price in control $\bar{p}_0$ is then estimated as $\hat{\alpha}$ and the average price for trader type $j = \{OT, HT\}$ in treatment, $\bar{p}_{1,j}$, is estimated as $\hat{\alpha} + \hat{\beta}_j$. The difference relative to control therefore equals $\Delta\bar{p}_{j,0} = \hat{\beta}_j$. To estimate average prices separately for local and commercial traders and their differences between treatment and control, we regress the price for sale i in follow-up season t on a constant and the interaction of the indicators $s_{j,i,t}$ for $j = LT, CT, HT$ and the treatment dummy $D_{i,t}$.

In online Appendix G Table 13, we show that the market shares ($\bar{s}_{d,j}$) and prices ($\bar{p}_{d,j}$) of local and commercial traders, respectively, are balanced at baseline across treatment and control.

II. Recovering the causal spillover effect

For farmer k in follow-up season t, let $H^d_{k,t} = h$ be an indicator for the farmer's decision to sell to the high quality buyer ($h = 1$) or to other traders in the market ($h = 0$) if they are assigned to group $D_k = d$.[8] By design, $H^0_{k,t} = 0$, because no-one in the control group could sell to the high quality buyer. Further, let $p_{k,t}(h, d)$ denote the potential price for farmer k:s sale in follow-up season t if the destination for the farmer's sale were $H_{k,t} = h$ and their assignment group were $D_k = d$.

We assume the potential price for farmer k:s sale in follow-up season t in the control group is a linear function of a farmer-specific effect (θ_k), a time fixed effect (λ_t) and a random error,

$$p_{k,t}(0,0) = \alpha + \theta_k + \lambda_t + \varepsilon_{k,t} \ , \tag{8}$$

where $E(\varepsilon_{k,t}|\theta_k, t) = 0$ with the farmer-specific effects normalized to $E(\theta_k) = 0$. Given randomization and the assumptions we have made,

$$E\big(p_{k,t}(0,0)\big|\theta_k, H_{k,t}, D_k, t\big) = E\big(p_{k,t}(0,0)\big|\theta_k, t\big) \, . \tag{9}$$

The farmer's potential price in treatment is

$$p_{k,t} = \rho_{j,k,t} + p_{k,t}(0,0) \ , \tag{10}$$

for $j = OT, HT$.

With this structure, the causal effect of the high quality buyer entering on prices for sales to other traders can be recovered from the data using a difference-in-differences

[8] To simplify the exposition, we assume that each farmer makes only one sale per season (which is true for the large majority of farmers).

approach. We estimate the difference-in-differences using normalized prices $\tilde{p}_{k,t}$ in each period, where from each $p_{k,t}$ we subtract the average control group price, and then divide the difference by it $\left(\tilde{p}_{k,t} = \frac{p_{k,t}-\bar{p}_{0,t}}{\bar{p}_{0,t}}\right)$. This removes the aggregate variation λ_t in each period. To purge the data of unobserved time-invariant farmer heterogeneity, we then consider the difference between $\tilde{p}_{k,t}$ and the farmer's normalized baseline price.

$$\Delta_{DiD} = E\left(\tilde{p}_{k,t} - \tilde{p}_{k,BL} \middle| D_k = 1, H_{k,t} = 0\right) = E\left(\frac{\rho_{OT,k,t}}{\bar{p}_{0,t}} \middle| D_k = 1, H_{k,t} = 0\right) = \Delta\tilde{p}_{OT,causal} \quad (11)$$

Equation (11) gives the causal effect in each period in percentage changes relative to the control group.

We estimate Δ_{DiD} by regressing the difference $\Delta\tilde{p}_{k,t} = (\tilde{p}_{k,t} - \tilde{p}_{k,BL})$ in treatment data on the indicators $s_{j,k,t}$ $(j = OT, HT)$, which equal 1 if treatment farmer k sold to market j in follow up period t and zero otherwise

$$\Delta\tilde{p}_{k,t} = \eta_{OT} s_{OT,k,t} + \eta_{HT} s_{HT,k,t} \,. \quad (12)$$

The estimated coefficient $\hat{\eta}_{OT}$ is a consistent estimator of the causal market spillover effect. Finally, the difference between the trader-specific price change between treatment and control and the causal effect $\Delta\bar{p}_{j,0} - \hat{\eta}_j$ for $j = OT, HT$ provides a consistent estimate of the selection effect.

III. Calculating the contribution of the spillover effect to the average treatment effect

We decompose the difference between prices in treatment and control (in percent relative to control group), which by virtue of randomization measures the average causal effect on prices, into the causal effect on prices for sales to the local market and the causal effect for sales to the high quality buyer.

$$\begin{aligned} \Delta\tilde{p} \xrightarrow{p} E\left(\tilde{p}_{k,t}(.,1) - \tilde{p}_{k,t}(0,0)\right) &= \Delta\tilde{p}_{causal} \\ &= \Pi\left(H_{k,t} = 0 \middle| D_k = 1\right) \times \Delta\tilde{p}_{OT,causal} \\ &+ \Pi\left(H_{k,t} = 1 \middle| D_k = 1\right) \times \Delta\tilde{p}_{HT,causal} \end{aligned} \quad (13)$$

Hence, the share of the average causal effect on prices explained by the market spillover effect is $\frac{\Pi(H_{k,t} = 0|D_k = 1)\times\Delta\tilde{p}_{OT,causal}}{\Delta\tilde{p}_{causal}}$, which we estimate as $\frac{\bar{s}_{OT,1}\times\Delta\hat{\tilde{p}}_{OT,causal}}{\Delta\tilde{p}}$. The average treatment effect on sales is 12.8% and the term $\bar{s}_{OT,1} \times \Delta\hat{\tilde{p}}_{OT,causal}$ is estimated as 4%. Hence, the ratio is 32%.

G. Tables

Table 1. Local and commercial traders

Variable	Mean
Interactions: local vs. commercial traders	
Market share of local traders (%)	78.2
Market share of local traders: season 3 (%)	68.4
Market share of local traders: season 4 (%)	82.6
Market share of local traders: season 5 (%)	78.6
Market share of local traders: season 6 (%)	82.1
Market share of local traders: season 7 (%)	79.4
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 (%)	3.90
Sold to one buyer only in the season (%)	89.6
Sold to two buyers in the season (%)	9.60
Sold to three buyers in the season (%)	0.90
Share sold (%)	82.1
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 (%)	36.7
Sold to the same buyer in two out of five seasons (%)	30.6
Sold to different buyers in each season (%)	2
Prices	
Normalized price paid by local traders	1
Normalized price paid by commercial traders	1.08
Sample	
Sales	420
Household-seasons	335
Households	78

Note: See main text for details.

Table 2. Correlations between lab and visually verifiable defects and lab verified defects and probability of aflatoxin

Specification	(1)	(2)	(3)
Outcome variable:	Lab verified quality	Aflatoxin >10ng/g	
Visually verifiable quality	3.84		
	(0.75)		
	[0.000]		
Lab verified quality		0.10	1.06
		(0.04)	(0.54)
		[0.015]	[0.048]
Constant	1.87	-0.22	-6.67
	(0.18)	(0.09)	(2.12)
Observations	43	101	101
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.

Table 3. Pre-harvest 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	
	(0.23)	(0.18)	(0.10)	(0.06)	(0.08)	
	[0.95]	[0.39]	[0.91]	[0.25]	[0.21]	
Constant	2.33	1.29	0.52	0.07	0.76	
	(0.14)	(0.12)	(0.07)	(0.04)	(0.06)	
Joint test p-value						[.51]
Observations	99	99	99	99	99	99
R-squared	0.22	0.17	0.23	0.12	0.39	0.13

Note: OLS regressions with village fixed effects. Robust standard errors in parenthesis and p-values in brackets. All variables are collected pre-harvest (but post-planting) in the fall 2019 season. 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 4. 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 for the returns to quality experiment. Attrition rate is the share of households, out of all enrolled, not surveyed/tested at follow-up. Robust standard errors. P-values in brackets.

Table 5. Samples: Market access and extension experiments

	All	Treatment	Control
***Panel A*. Market access experiment**			
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.	1198	692	506
Baseline & follow-up panel: households	180	104	76
Baseline & follow-up panel: clusters	20	12	8
***Panel B*. Extension service experiment**			
Baseline panel: household-season obs.	495	252	243
Baseline panel: households	172	88	84
Baseline panel: clusters	18	9	9
Follow-up panel: household-season obs.	458	235	223
Follow-up panel: households	164	82	82
Follow-up panel: clusters	18	9	9
Baseline & follow-up panel: household-season obs.	931	471	460
Baseline & follow-up panel: households	164	82	82
Baseline & follow-up panel: clusters	18	9	9

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

Table 6. Attrition: Market access and extension experiments

	(1) All	(2) Treatment	(3) Control	(4) Difference	(5) Obs.
***Panel A*. Market access experiment**					
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]	
***Panel B*. Extension service experiment**					
Households attritted during follow-up	0.047	0.068	0.024	0.044	172
				[.192]	
Household-season re-survey rate: follow-up	0.931	0.955	0.907	0.049	492
				[.119]	

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. Extension experiment: summary statistics and balance at baseline

	(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	0.23	0.42	172	0.26	0.19	0.071	.253
Main decision maker: completed primary school	0.36	0.48	172	0.41	0.31	0.100	.203
Number of household members	5.49	2.32	172	5.23	5.76	-0.535	.155
Distance to district capital (km)	22.7	8.14	172	22.5	22.8	-0.280	.945
Distance to main road (km)	11.6	6.88	172	12.1	11.0	1.060	.473
Panel B. Farm enterprise characteristics							
Maize acreage	1.98	1.41	495	1.88	2.09	-0.219	.392
Expenses (USD)	113.0	129.8	330	111.1	115.0	-3.795	.863
Harvest (ton)	2.22	1.88	473	2.20	2.24	-0.049	.881
Yield (ton/hectare)	2.57	1.24	473	2.66	2.49	0.162	.455
Share sold	0.85	0.21	469	0.84	0.86	-0.028	.364
Price kilogram (USD)	0.14	0.06	450	0.14	0.14	0.003	.342
Harvest value (USD)	310.6	306.3	473	302.5	318.9	-12.99	.793
Profit I (USD)	154.6	188.9	329	157.3	151.8	6.672	.795
Joint balance test I							.274
Joint balance test II							.125
Joint balance test III							.096

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, equipment, transport 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. 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 452); all farm enterprise outcomes in test III (seasons 2-3; sample size 317).

Table 8. 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	
***Panel A*. Market access experiment**									
Maize acreage	1.75	2.27	2.44	2.16	1.57	1.38	1.76	1.60	544
Expenses (USD)		133.9	159.2	146.3		131.2	181.2	158.0	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	498
Price kilogram (USD)	0.25	0.21	0.16	0.20	0.06	0.04	0.023	0.057	470
Harvest value (USD)	495.1	450.5	367.9	433.2	495.7	425.8	356.4	425.9	499
Profit I (USD)		316.6	208.7	263.7		336.7	224.3	291.8	363
***Panel B*. Extension service experiment**									
Maize acreage	1.74	2.12	2.08	1.98	1.39	1.33	1.49	1.41	495
Expenses (USD)		125.3	101.0	113.0		136.8	121.7	129.8	330
Harvest (ton)	1.88	2.33	2.39	2.22	1.59	2.01	1.95	1.88	473
Yield (ton/hectare)	2.32	2.59	2.78	2.57	1.20	1.28	1.19	1.24	473
Share sold	0.82	0.86	0.86	0.85	0.25	0.22	0.15	0.21	469
Price kilogram (USD)	0.21	0.15	0.074	0.14	0.031	0.020	0.017	0.060	450
Harvest value (USD)	409.5	349.6	186.4	310.6	369.1	311.8	171.3	306.3	473
Profit I (USD)		224.3	86.2	154.6		219.4	119.1	188.9	329

Note: Households in the baseline panel sample. Season 1-3 represents the order of seasons (spring 2017, fall 2017, spring 2018 in the market access experiment, and fall2017, spring 2018, fall 2018 in the extension service experiment). 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.

Table 9. Baseline balance: Control and quasi-control

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
	Sample			*Means*		*Difference in means*	
Variable	Mean	Std.	Obs.	QC	C	Coeff.	*p*
***Panel A*. Household characteristics**							
Main decision maker: female	0.13	0.34	275	0.12	0.17	-0.058	.488
Main decision maker: completed primary school	0.41	0.49	275	0.42	0.39	0.038	.522
Number of household members:	6.00	2.50	275	5.91	6.24	-0.335	.337
Distance to district capital (km)	29.7	9.53	273	29.5	30.2	-0.685	.857
***Panel B*. Farm enterprise characteristics**							
Maize acreage	2.38	1.46	275	2.37	2.41	-0.033	.936
Expenses (USD)	118.0	132.0	275	117.3	120.0	-2.783	.931
Harvest (ton)	2.30	1.88	275	2.35	2.19	0.157	.734
Yield (ton/hectare)	2.30	0.97	275	2.34	2.20	0.148	.410
Share sold	0.87	0.15	271	0.88	0.82	0.059	.133
Price kilogram (USD)	0.077	0.020	267	0.079	0.074	0.005	.107
Harvest value (USD)	178.8	157.2	271	186.8	157.9	28.859	.393
Profit I (USD)	62.5	113.9	271	71.9	37.9	33.992	.021
Joint balance test I							.771
Joint balance test II							.216

Note. Households in Sample Frame 1, control group (C) and all households in Sample Frame 3 (quasi-controls (QS)). Panel A: measured at first baseline round. Panel B: measured in Season 4. Standard errors are clustered at the village level. Expenses is expenses on inputs and hired labor. Harvest value includes own-produced consumption, valued at community-specific market value. Profit I is the difference between harvest value and expenses. The joint balance tests report *p*-values from testing whether the baseline outcomes predict enrollment into quasi-control, with profit dropped due to collinearity: all household characteristics in test I; farm enterprise outcomes in test II (sample size 267).

Table 10. Impact on investment: Incidence of hybrid seeds and fertilizer use

	(1) Any improved seeds	(2) Any fertilizer
***Panel A*. Market access experiment**		
Access to a market for quality maize	0.038	0.040
	(.250)	(.054)
	[.274]	[.065]
Observations	658	658
R-squared	0.08	0.03
Mean control	0.13	0.03
***Panel B*. Extension service experiment**		
Extension service	0.0028	0.014
	(.964)	(.490)
	[.966]	[.527]
Observations	447	447
R-squared	0.18	0.01
Mean control	0.24	0.04

Note. ANCOVA specification. 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 11. Impact on productivity and income: Trimmed sample

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*. Market access experiment**				
Access to a market for quality maize	0.018	0.0076	280.1	107.4	73.1	11.7	66.9	93.1
	(.001)	(.967)	(.153)	(.029)	(.072)	(.367)	(.035)	(.019)
Observations	612	670	650	651	651	633	628	456
Mean for control	0.15	2.21	1755.87	778.44	272.17	100.34	172.28	120.07
				***Panel B*. Extension service experiment**				
Extension service	0.002	-0.055	-98.4	26.6	-15.8	-8.04	-3.05	13.6
	(.649)	(.698)	(.592)	(.648)	(.680)	(.393)	(.915)	(.654)
Observations	415	451	439	437	441	439	432	432
Mean for control	0.17	1.81	1603.29	858.11	293.53	89.13	206.92	109.83

Note. Trimmed sample: non-negative continuous variables are trimmed at the top 1% observations in each season. Variables that can take both positive and negative values, i.e. profit, are trimmed at the top- and bottom 1% observations in each season. Clustered-by-village standard errors with *p*-values in parenthesis.

Table 12. Mechanisms

Specification	(1) Linear model
Panel A	
$E[Y_1 - Y_0]$	343.9
	[.012]
Decomposing $E[Y_1 - Y_0]$	
ΔTFP: $\Delta\theta/E[Y_1 - Y_0]$	0.519
	[.000]
ΔLand: $\Delta A/E[Y_1 - Y_0]$	0.264
	[.016]
ΔFamily labor: $\Delta L_F/E[Y_1 - Y_0]$	0.027
	[.722]
ΔHired labor: $\Delta L_H/E[Y_1 - Y_0]$	0.158
	[.276]
ΔSoil enhancing investment: $\Delta\mu/E[Y_1 - Y_0]$	0.032
	[.575]
ΔMeasured inputs total: $\Delta x/E[Y_1 - Y_0]$	0.481
	[.001]
Panel B	
Test of independence	0.327
	[.890]
Test I of mediators	125.1
	[.000]
Test II of mediators	14.47
	[.013]
Observations	464

Note. Sample: households in the last three follow-up seasons in the market access experiment. Panel A: overall treatment effect and the relative magnitudes of measured inputs and unmeasured inputs (TFP). Clustered-by-village standard errors with p-values in parenthesis. 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 tests the H_0 that the vector of observed inputs do not explain harvest volumes. Test II of mediators tests the H_0 that the vector of observed inputs are not affected by treatment.

Table 13. 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	.339
Commercial traders	0.36	0.28	0.085	.339
Prices:				
Local traders	560.9	568.6	-7.689	.551
Commercial traders	565.3	573.9	-8.597	.707

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

Table 14. Impact on productivity adjusting for selection

	(1) Harvest	(2) Yield
Selling to the high quality buyer	0.453	0.173
	[.132]	[.147]
Selling to other traders	-0.084	0.072
	[.597]	[.401]

Note. Each column reports the causal effect estimated regressing the difference between the (normalized) outcome (harvest in column 1 and yield in column 2) at follow-up and baseline on an indicator whether the household sold to the high quality trader in the follow-up season or not; i.e. specification (17) in online Appendix section G. The unit of observation is the household (658 observations). Clustered-by-village standard errors with p-values in brackets.

H. Figures

Figure 1. Design of the program

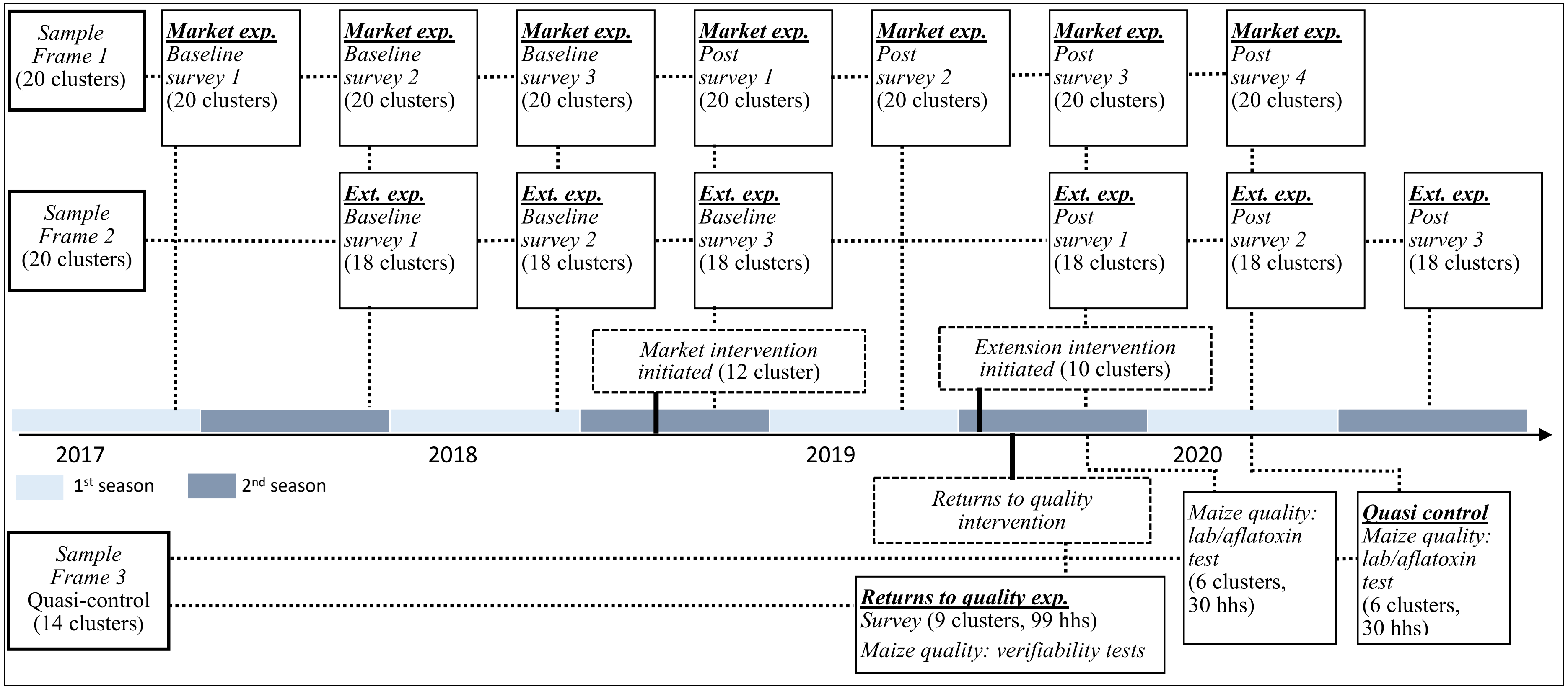
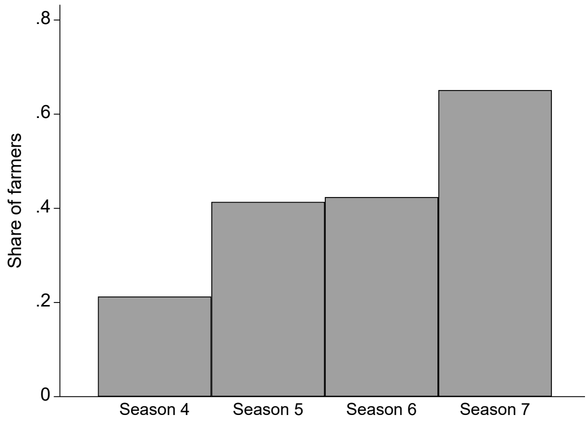
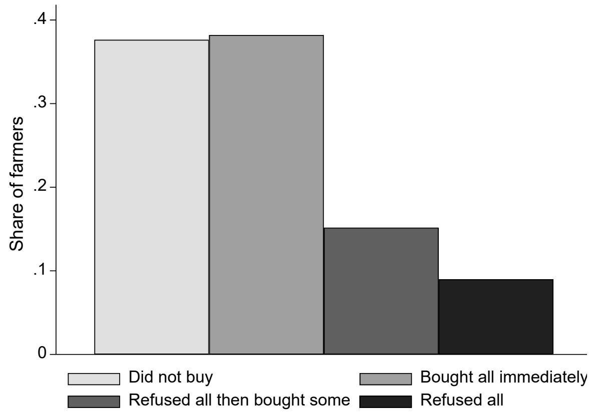


Figure 2. Quality upgrading: agro-company’s interactions with farmers

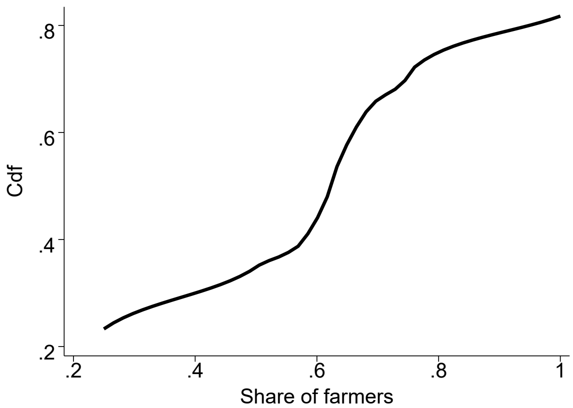
Panel A. Share of farmers selling to agro-company



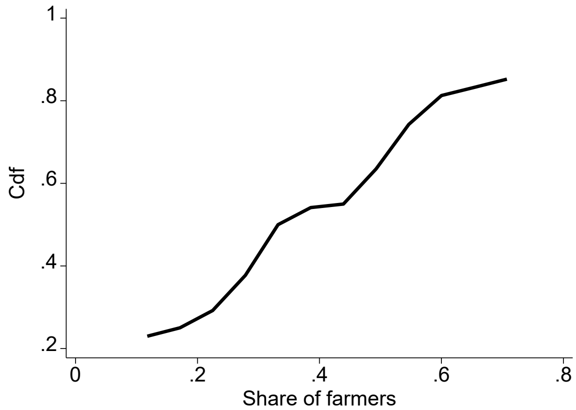
Panel B. Agro-company’s buying pattern



Panel C. Distribution: market shares (Season 7)

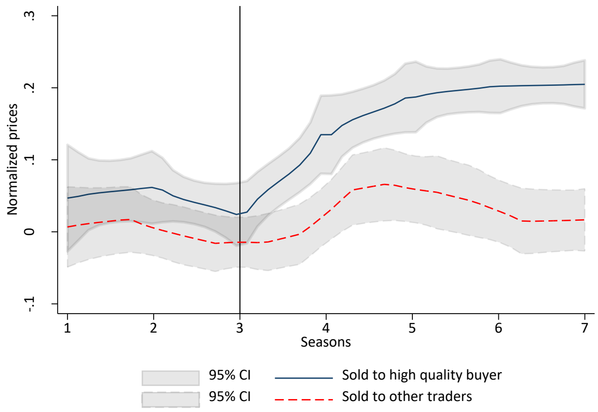


Panel D. Distribution: market share (all seasons)



Note. Panel A: household survey data; share of farmers selling to the premium quality buyer per follow-up season. Panel B: data from agro trading company matched to trial sample households. Panel C: household survey data; share of farmers selling to the premium quality buyer, distribution across villages in last follow-up season (season 7). Panel D: household survey data; share of farmers selling to the premium quality buyer, distribution across villages, average across all four follow-up seasons.

Figure 3. Normalized prices in the group selling to the high quality trader and other traders



Note. Normalized prices in the group of farmers selling to the high quality trader (sold more than once at follow-up) and other traders. Season 1-3 are baseline seasons and 4-7 follow-up seasons.

Additional references

Heckman, James and Rodrigo Pinto. 2015. "Econometric Mediation Analyses: Identifying the Sources of Treatment Effects from Experimentally Estimated Production Technologies with Unmeasured and Mismeasured Inputs". *Econometric Reviews* 34(1-2): 6-31.

Imai, Kosuke, Luke Keele and Teppei Yamamoto. 2010. "Identification, Inference and Sensitivity."Analysis for Causal Mediation Effects *Statistical Science* 25(1): 51-71.

Imai, Kosuke, Luke Keele, Dustin Tingley, and Teppei Yamamoto. 2011. "Unpacking the black box of causality: Learning about causal mechanisms from experimental and observational studies." *American Political Science Review* 105(4): 765-789.

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.