Search Costs, Intermediation, and Trade: Experimental Evidence from Ugandan Agricultural Markets

Lauren Falcao Bergquist[1]*, Craig McIntosh[2], and Meredith Startz[3]

[1]Yale University and NBER
[2]University of California, San Diego
[3]Dartmouth College and NBER

July 2023

PRELIMINARY DRAFT

Abstract

Search costs may be a barrier to market integration in developing countries, harming both producers and consumers. We present evidence from the large-scale experimental rollout of a mobile phone-based marketplace intended to reduce buyer-seller search and matching costs for agricultural commodities in Uganda. We find that market integration improves substantially: trade increases and price dispersion falls. This reflects price convergence across relative surplus and deficit markets, with no change in average prices. Interpreting our experimental variation through the lens of a trade model, we correct our reduced form estimates to account for equilibrium effects on control markets via trade connections. Our results suggest that the intervention reduced fixed trade costs between treated markets by 28% and increased average trade flows across all markets by 2%. Contrary to the stated goals of the marketplace, but consistent with the existence of economies of scale in search or other trade costs, almost all activity on the platform is among larger traders, with very little use by smallholder farmers. Nevertheless, the benefits of improved arbitrage by traders appears to pass through to farmers in the form of higher revenues in surplus markets, as trader entry increases and measured trader profits decrease in response to falling search costs.

*We gratefully acknowledge funding from USAID/BASIS, USAID/DIL, the Bill & Melinda Gates Foundation and the UK Foreign, Commonwealth & Development Office, awarded through the Agricultural Technology Adoption Initiative (ATAI), the International Growth Center, the Policy Design and Evaluation Lab, and an anonymous donor. We thank IPA, Stephanie Annyas and Laza Razafimbelo for excellent project management. This trial was registered as AEARCTR-0000749, and is covered by UCSD IRB #140744. All errors are our own.

1 Introduction

The integration of agricultural markets is an issue of central welfare importance in developing economies. On the production side, it matters to the many households that rely on farming as a major source of income. On the consumption side, it influences the purchasing power of the poor, who spend a large fraction of their income on food. Trading frictions that limit arbitrage between relative surplus and deficit areas can therefore have large welfare costs (Barrett, 2008; Rashid and Minot, 2010). These frictions thwart otherwise profitable trades, leading to low producer prices, high consumer prices, and large geographic and intertemporal price dispersion.

While high transport costs are an important source of trading frictions, particularly in sub-Saharan Africa (Teravaninthorn and Raballand, 2009), there is growing evidence that price gaps are substantially larger than can be explained by transportation alone (Atkin & Donaldson, 2015). Evidence from the spread of mobile phones suggests that search frictions may be an important contributor – as access to phones grew in developing countries, agricultural market integration improved and geographic price dispersion fell (Aker, 2010; Jensen, 2010; Allen, 2014). However, integration remains highly incomplete even in places with with high mobile phone penetration (Moser et al., 2009; Porteous, 2019). Furthermore, it is not clear what type of information drives search frictions – efforts to disseminate price information to farmers, for instance, have had mixed results, and overall, very limited impact (Camacho and Conover, 2011; Fafchamps and Minten, 2012; Nakasone, 2014; Mitra et al., 2018; Hildebrant et al., 2020).

We investigate the role of a different type of search friction: search for potential transaction partners. We conduct a large-scale randomized control trial of a mobile marketplace for food crops in Uganda, which links potential buyers and sellers through a simple SMS-based platform. The design of the platform and experiment enable us to isolate the causal effect of reducing buyer-seller search frictions, separate from other factors that often come bundled with new technology or agricultural interventions. Our at-scale randomization enables us to generate and measure impacts at the market level.

Access to the platform substantially increased trade between treated markets and led to convergence in prices across surplus and deficit areas. We reach two major conclusions. First, buyer-seller matching frictions are a quantitatively important barrier to market integration, even in this setting where mobile phones are ubiquitous, and they can be alleviated through technologies like the platform we evaluate. Second, scale economies in trade are crucial to understanding search frictions. Even though reducing search costs decreases scale economies, most Ugandan farmers are still too small to benefit by engaging in inter-market trade directly. Instead, the platform improves arbitrage among wholesale traders, and the benefits of this pass through to farmers and consumers via impacts on equilibrium prices in their home markets.

The mobile trading platform we evaluate – called Kudu – acts as a clearinghouse, in which those buying and selling agricultural commodities can "match" on their phones. Users submit bids and asks, and the platform directly connects potential matches. Kudu is free, and was offered to both farmers and traders. In-village support services were provided to ensure ease of access. Introduction was randomized across 110 Ugandan subcounties, covering about 12% of the country and a population of millions. We track outcomes over three years, equivalent to six harvest seasons. We gathered data on market-level prices in 236 markets every two weeks. We also collect multiple survey rounds with a representative sample of traders in the study markets to analyze how the intervention drives their trading behavior, prices, and profits. Finally, we collect surveys of farming households to study the impacts of the platform on farmer revenues and welfare.

Using this data, we find that access to Kudu increases trade between treated markets on both the extensive and intensive margins – it increases the probability that treated markets trade with each other, the number of traders who trade between them, and the volumes traded. Market prices increase in relative surplus areas that are treated and decrease in relative deficit areas. As a result, average price gaps between treated market decrease. Usage of the platform is almost entirely among traders; only the very largest farmers use the system. However, farmers in areas with relative surplus at baseline benefit from the equilibrium effects on prices. These farmers see significant increases in revenues and quantities sold, even though most do not use the platform.

The scale of the experimental intervention enables us to address a new set of questions about the impacts of reducing search frictions in a connected trading network. Clustered randomization at the market level lets us measure effects on equilibrium outcomes such as prices and aggregate trade flows. Randomization also circumvents causal inference issues related to the endogenous location and quality of infrastructure (such as roads, rail, telegraph and phone coverage, etc) that are common challenges when studying trade costs. However, control markets are affected by the intervention because they trade with treated markets. This means simple treatment-control comparisons do not yield valid estimates of treatment effects, a problem that is inherent to causal inference in a trade setting.

To solve this problem and correctly identify the magnitude of the impact on trade generated by the platform, we interpret the experimental impact on trade flows through the lens of a trade model. To capture the crucial role played by the extensive margin, we allow for fixed trade costs of trade, including the costs that buyers and sellers incur to connect with one another across markets. Kudu lowers these search costs. Holding market-level factors constant, we isolate the direct effect on trade costs using experimental variation in bilateral treatment status between markets. We estimate that Kudu drove a 28% reduction in fixed trade costs between treated markets and increased aggregate trade flows across all markets by 2%. This is about half of the treatment effect on aggregate trade flows implied by a naive treatment-control comparison that does not account for equilibrium effects on control markets.

Combining the treatment effect on fixed trade costs with other estimated model parameters, we can measure the total effect of the platform on market-level outcomes such as prices. These estimates account for the influence of other markets on both treatment and control market outcomes, and therefore capture the total impact of the platform as it was actually rolled out across a large area. There are distributional effects of the platform, with farmers gaining and consumers losing in surplus markets, and the opposite in deficit markets. However, arbitrage improves across deficit and surplus areas, and there are net gains from trade, such that total welfare in the study area increases due to the reduction in search costs caused by the platform.

The rest of the paper is structured as follows: Section 2 discusses the setting and study design. Section 3 presents reduced-form effects on market-level outcomes. Sections 4, 5, and 6 present the structure, estimation, and results of the model respectively. Section 7 lays out the distributional effects of the intervention on traders and farmers. Section 8 concludes.

2 Setting and Study Design

We conducted an at-scale, clustered randomized RCT that operated in the field for three years. The study operates in 110 subcounties in Uganda that our implementing partners selected as promising for the platform roll-out. These 110 subcounties, which are the unit for random assignment of the intervention, represent a mix of major surplus or minor deficit areas for maize production (see Figure B.1 for a map from FEWS-NET of surplus maize areas in Uganda presented alongside a map of the study subcounties).

2.1 Market Integration and Search Behavior

Market prices for maize, our core study crop, show strong variation both over space and over time across our study area (see Table A.1).[1] Figure 1 presents the absolute value in the gap in prices across each pair of markets in our sample (black solid line), as collected by our market surveys (see Section 2.4 for additional detail on the data collection). We observe the average gap in price for maize is about 110 Ugandan shillings (UGX) per kg, or about 15% of average price. Unsurprisingly, these price gaps rise with the distance between the markets. A major driver of this price dispersion observed across markets is transportation costs, which in the region are among the highest in the world (Teravaninthorn and Raballand, 2009). Transportation costs alone, however, cannot explain the full gap in prices observed across markets. Figure 1 also plots estimated transport cost as a function of distance (dotted line), as reported in our trader surveys.[2] We observe price gaps that are roughly 50% higher than transport costs, suggesting pervasive violations of the law of one-price.

As is common in rural markets in the region, the farmers in our study area are smallholders and do very limited marketing of their crops. They are on average small net producers, though our sample contains net consumers as well. Farmers grow an average of 835kg/year of maize, spread out over two harvest season. They sell on average 58% of this harvest, typically making two sales

[1] Maize is the most commonly grown and consumed crop in our study area. As we will describe later, it was also the crop most commonly traded on the platform. Our market survey also follows beans, another non-perishable staple, and two perishable crops, tomatoes and bananas (green bananas are steamed and eaten as the most important staple starch crop in many parts of Uganda).

[2] Traders reported the roundtrip transport costs along each of their five most commonly travelled routes, and the vehicle size typically used. From this data, we construct an estimate of per kg transport costs, which we then estimate as a function of the km traveled.

per year (one per season). Farmers sell very close to home, with 66% selling exclusively at farm-gate. The remainder sell in very nearby markets, which are on average just 2km away from their farm-gate. Farmers typically sell to a single trader a year, with whom they have transacted in the past.

In contrast, traders are both much larger and do much more active marketing. The median trader buys and sells 25 tons of maize per year, more than 50x the size of the median farmer. 62% of traders use 5-10 ton trucks to move their inventory; the remainder use motorcycles or bicycles. Traders typically operate from a local market, making many small purchases within their local subcounty, aggregating, and selling downstream to fewer, larger markets. The typical market is home to on average six traders, though there is variation in market size in our sample. Churn among traders is low, with the median market seeing a little over one new and one exiting trader per year. Traders appear to work with large margins; at baseline traders bought maize at an average of 12.7 cents/kilo and sold at 16.4 cents/kilo, a buy-sell margin of 29%. From baseline monthly revenues of \$2,243, traders report an average monthly profit of \$297.[3] By comparison, average total monthly household expenditure in our farmer sample is \$65.

Mobile phones are an important technology used in traders' search and matching, with traders reporting calling a seller or other contact ahead of time to get information about price or availability in 54% of their transactions; however, more traditional search-by-visiting is still common, with traders reporting that 47% of their transactions occur without the trader calling ahead for any prior information. At baseline, about 50% of traders report using radio broadcasts as a market discovery tool, and 10% some kind of SMS service. Use of these services is lower among farmers, with only 7% of farmers reporting listening to radio broadcasts for market information and 2% receiving information via SMS. Moreover, these services typically only offer simple price alerts; none to our knowledge provides direct connections to buyers or sellers.

[3] Consistent with these figures, Bergquist and Dinerstein (2020) estimate that the median trader in their sample in Kenya retains 12% of total revenues in profits.

2.2 Kudu: a Market-Matching Platform

The introduction and rapid spread of mobile phones across sub-Saharan Africa has generated much excitement, offering the promise of dramatically reducing search costs. Indeed, the rollout of cell-phone towers in the early 2000s has been shown to have substantially reduced price dispersion in grain markets (Aker, 2010; Aker and Mbiti, 2010). Building off this success, recent efforts have attempted to move beyond the passive reduction in search costs facilitated by easier bilateral communication via mobile phones, and into more active facilitation of search on mobile platforms design for agriculture.

The first generation of these initiatives focused on the dissemination of price information to farmers via mobile phone. However, price information alone has had mixed results (Camacho and Conover, 2011; Fafchamps and Minten, 2012; Nakasone, 2014; Mitra et al., 2018; Hildebrant et al., 2020). Moreover, most existing studies are randomized at the individual or village level, while prices are often determined in general equilibrium across a wider geographic area, complicating inference.[4]

Kudu, in contrast to existing price information platforms, offers a full match-making services. The system, which was developed at Kampala's Makerere University and in partnership with researchers at Microsoft Research, is designed to make it easier for buyers and sellers of agricultural commodities to connect. Users can post asks (sale offers) and bids (purchase offers) onto Kudu either using a smartphone or by registering their location and then using a basic feature phone to send messages to the platform via free-form text message or a USSD drop-down menu. A call-center also collects asks and bids by phone. Based on the price, quantity, and location of the buyer and seller, the system then matches particular buyers and sellers as proposed trading partners.[5] Users are then contacted by SMS and informed that the match has occurred, along with the contact

[4]For example, Svensson and Yanagizawa (2009) find evidence that broadcasting prices via radio lead to higher farmgate prices in Uganda; however, a follow-up paper suggests that once accounting for general equilibrium effects, average farmer revenues impacts are minimal (Svensson and Yanagizawa-Drott, 2012). Hildebrant et al. (2020) also find evidence of spillovers from a price alert system in Ghana.

[5]There were two processes by which buyers and sellers could be matched. First was the Kudu algorithm that cleared the market each day, attempting to maximize what its algorithm calculated to be the global gains from trade, using a penalty function decreasing in the price difference between the bid and ask and increasing in distance. Second was a hand-matching process conducted by employees who could view a dashboard of the business on the platform and attempt to match trades manually. For more technical details on the Kudu platform, see Newman et al. (2018).

information of the other party. Buyers and sellers can then directly connect and arrange for the sale, though commonly, a Kudu employee will help facilitate this communication, reaching out by phone to both parties to gauge their interest in the deal and coordinate next steps.[6]

To randomly deploy Kudu, we partnered with AgriNet, a private sector agribusiness firm, to employ and train 210 Commission Agents (CAs) to serve as the on-the-ground agents of the project, promoting the mobile marketplace to local farmers and traders. However, CAs were not reliable promoters of the project, perhaps because they were recruited from pools of existing traders in the area and therefore lacked the incentive to promote the platform to others (though as we will describe below, they were regular users of Kudu for their own trading operations). Ultimately, Kudu hired salaried staff, not drawn from the local trader population, to promote the platform, which was much more successful in generating broad awareness of Kudu.

In addition, we promoted Kudu via text message every two weeks to all treatment traders (and CAs), as well as a randomly selected three-quarters of the treatment farmer households in the study.[7] Promotional information included an advertisement and information on how to trade on the platform, either by registering directly on Kudu or by contacting a local Kudu staffer, whose contact information was provided. The text message also included price information for maize, beans, matooke, and tomatoes at the recipient's local market, downstream hub market, and closest major market, as well as five randomly sampled treatment markets each week. [8]

Finally, understanding that contractual risk might hamper the introduction of impersonal, technologically-mediated trade, we introduced a system of transport guarantees to compensate

[6] The manual matching process could also deal with failed matches in a flexible way (moving on immediately to the next counterparty), while the Kudu algorithm required them to go back into the matching pool for the next iteration.

[7] This randomization, which was conducted at the household level (blocked on subcounty), was set up in order to generate exogenous variation in direct use of Kudu. The goal was to separately estimate the direct impact of using Kudu vs. the indirect impact of living in an area in which others were using Kudu and therefore being exposed to GE effects of others' use, but not directly using the platform oneself. In practice, this second randomization offers a negligible first stage for direct use, as farmer take-up is close to zero. Instead, treatment effects observed for farmers are mostly a result of shifting GE forces. See Section 7 for greater discussion.

[8] The five random sampled markets were the same for all recipients, such that the entire network received price information about the same randomly selected markets each round. This randomization was designed to test the impact of price information alone on these markets. As described in Section 8, we see no effect of this randomly select price information. We also randomly rolled in the full price information messages (including price information on one's own local markets) to a random subset of control subcounties towards the end of our study, and see no impact of this intervention. We therefore see the price information component of the intervention as non-pivotal. See Section 8 for further detail.

randomly selected users for losses should transactions not be executed as promised on the Kudu platform. These offers were randomized, in order to separately test the impact of contractual risk reduction. Appendix XX provides greater detail.

2.3 Randomization

The introduction of Kudu was randomly assigned at the subcounty level. In the Ugandan administrative hierarchy, this is several levels above the village, as villages are organized into parishes and parishes into subcounties. At the time of our study, the typical subcounty in Uganda had about 50,000 residents. This is therefore at "at-scale" randomization, conducted at this level with the goal of being able to observe treatment-control differences in general equilibrium outcomes such as prices and trade flows.

To conduct the randomization, we first conducted a census of all markets within our 110 study subcounties. We listed all markets that were permanent (e.g. not meeting only on specific days of the week) and featured both buying and selling of maize (as opposed to wet markets where fruits and vegetables are only sold). This process identified 236 trading centers, hereafter referred to as markets. Markets were then classified as "hubs" (major local commercial centers) and "spokes" (more remote local markets) (see Figure B.2 for a map of markets). We also at this time conducted a market survey to measure prices of four crops (maize, beans, matooke, and tomatoes) at baseline.

We first blocked the randomization of the 110 subcounties in our sample on whether the subcounty contained a hub (17%) or not (83%), and we stratified by a subcounty-level price index (mean of the z-scores of the prices of each of the four crops at the markets in each subcounty).

2.4 Survey Sampling and Timeline

The intervention ran for three years, starting in 2015 and concluding in 2018. This time period spans six major agricultural seasons. Figure B.3 presents a timeline for the project, and Figure B.4 provides a CONSORT diagram of study recruitment and attrition for each type of data.

We collect three core types of data for this project, using the 236 markets in our study as the primary sampling units. The first of these datasets is a high-frequency market survey. This survey gathered information in each market every two weeks by calling a key market informant, typically a trader whose store was based in the market. We collected data on the buying and selling price, availability, and average quality of four major food crops (maize, nambale beans, matooke bananas, and tomatoes). We also surveyed 20 hub markets in adjacent, non-study districts to provide an additional measure of potential spillover effects, as well as in the four 'super-hub' markets of Uganda.[9] The total number of markets reporting the biweekly market survey is thus 260, of which 236 form the core experimental sample. The market survey was collected for the three years during which the intervention ran.

The second dataset collected is a survey of traders in each study market. We first conducted a census of traders who were based in that market and who bought and sold at least one study crop. For markets that had fewer than 10 traders identified in the census, we surveyed all traders; for markets with more than this, we randomly sampled 10 traders. These traders were administered a baseline survey in 2015, prior to the initiation of any treatment, a midline survey in 2016 after one year of treatment, and an endline in 2018 after three years of treatment. The trader analysis is weighted to make it representative of all traders in study markets.

Finally, to understand the impact of the platform on farmers, we drew in a sample of agricultural households. We first listed all villages located in the subcounty.[10] We then selected the village containing the market (which are typically more urban) and randomly sampled one of the remaining villages within the same parish (which tend to be more rural). For these two villages, we then listed all the households based on administrative records held by the village chairperson, and randomly sampled households from these lists. We randomly sampled 8-9 farming households located within each village containing the market and another 4 in each rural village that does not contain the market. We imposed two eligibility criteria: (i) the household had to be engaged in agriculture, and (ii) the household had to have sold some quantity of any of the four crops included in the study in the previous year. In practice, these eligibility criteria excluded few households, with over 90% of households qualifying for study inclusion. Study households completed a baseline in 2015 and

[9] These super hubs are the capital, Kampala, plus three border markets that trade grain with neighboring countries: Kabale on the border with Rwanda, Busia on the border with Kenya, and Arua which trades to the DRC and South Sudan.

[10] In Uganda, these villages are called Local Council 1, or LC1s.

an endline survey in 2018 covering agricultural activities, farmgate prices, and marketed surpluses. Farmer analysis is weighted to make it representative of all farming households in sampled study village.

2.5 Attrition and Balance

We now present the attrition and balance for each of the three types of data captured in the study: the market surveys, the trader surveys, and the household surveys. For the market survey, we have 88% of the attempted (market x survey round) observations.[11] For the trader midline, we were able to survey 1,358 of the 1,457 baseline traders (93.2%). For both the trader and household endlines, we ran standard panel tracking, and then conducted an intensive tracking exercise that attempted to follow up with a random sample of attritors. The trader endline originally located 1,248 traders (85.7%), after which we randomly sampled 20% of attritors (41 individuals) for intensive tracking, and successfully located 37 of these (92.7%). The weighted tracking rate in the trader endline is therefore 98.6%. The household endline originally located 2,744 of the 2,971 baseline respondents, and we then randomly sampled 17% or 39 households for intensive tracking. 31 of these households were successfully intensively tracked (79.5%), giving us a weighted household tracking rate of 98.7%.[12]

Table A.4 examines the balance of the market survey for the two main study crops (maize and beans) and the core variables in the market surveys (buying and selling price, number of traders, and quality rating). Table A.5 uses the market survey data in dyadic form and examines the baseline balance of the experiment on price dispersion between markets. The experiment is well balanced at the market level. For the trader and household analysis, balance is analyzed using the sample still present at endline and is weighted using the attrition weights so as to mirror the structure of the

[11] As a robustness check for attrition, we present appendix tables that show the main market survey data using interpolation; given the long panel (83 rounds of market surveys) and the highly interspersed nature of the missing observations this provides a reasonable check on the extent to which market survey attrition may influence our results.

[12] Appendix Figure B.5 and Tables A.2 and A.3 present tests comparing attrition in the treatment to the control across the three data types. Among all the tests that we conduct only the intensive tracking rate in the trader survey appears differential, and given that this arises from finding 14 out of 14 control versus than 24 out of 27 treatment traders in the intensive tracking, this has relatively little influence on study-level effects. Overall, weighted attrition rates are very low and the overall unweighted attrition rate from the combination standard and intensive tracking is similar across treatment arms for all data types.

outcome analysis. Table A.6 analyzes the baseline attributes of traders across seventeen different attributes and finds no evidence of baseline imbalance. Table A.7 conducts the same exercise for households, finding two out of seventeen outcomes significantly different at the 10% level and one at the 5% level, in line with what we would expect by random chance. We therefore proceed to the analysis section with confidence that the study is both representative and well-balanced.

2.6 Platform Usage

Over the three years that the Kudu platform was operational as a part of this project, it received 29,308 unique asks and 31,177 unique bids. Maize accounts for 67% of asks on the platform, though 19 total crops were successfully traded, with the next most common being soya, rice, and beans. Figure B.6 shows the smoothed quantity of new bids and asks posted on the platform per day, with supply climbing steadily through the first year to reach a steady maximum of about 200 tons per day, and demand following a similar time path to reach average levels somewhat less than twice supply.[13] Figure B.7 shows the spatial distribution of asks, indicating study market centers across the country posting upwards of 1,000 asks each. Among those posting asks to sell on Kudu, 45% were study traders, 14% were AgriNet CAs, and 6% were study farmers. For those posting bids to buy, the corresponding percentages are 48% and 11% for study traders and CAs, and less than 1% of sellers are study farmers.[14]

Kudu instructs buyers to post their reservation bid prices and sellers to post their reservation ask prices; however, in practice the price recommended by the platform in the event of a match is the seller's price, so this is only incentive compatible for the buyer (though prices could in practice be renegotiated once the trading partners meet in person, so the suggested price was not binding). It was assumed at the launch of the platform that there would be a sizable gap between the two, ideally substantial enough for the platform to broker trades with comfortable margins for all parties, such it could eventually charge commissions in order to make the platform self-sustaining financially.

[13]Standing up supply and demand simultaneously was an issue at inception of the project; an initial surge of asks in the first season overwhelmed demand, but then a drive to encourage buying on the platform was highly successful and for the remainder of the project the total demand on the platform exceeded supply.

[14]The remainder of bids and ask came mostly from within the treated study areas, but from users outside of our study sample.

However, in practice, seller ask prices were often higher than buyer bid prices, and in fact higher than local market and even hub or superhub market prices.[15] Qualitative interviews suggested that sellers often posted prices that reflected strategic price offers, fishing for higher prices, much in the way that one would typically make offers in more traditional, in-person negotiations. Buyers' average bid prices, on the other hand, track hub market prices very well. Figure B.8 plots these values over time for maize, and Figure B.9 provides a box-and-whisker plot of bid and ask prices within each season, in which we can see that the median bid price is typically at or below the 25th percentile of ask prices. In practice, Kudu accommodated this sort of negotiating behavior by allowing for a small negative overlap of prices, taking prices as a signal of willingness to buy or sell, but allowing those that were close but not quite overlapping to match.

Nonetheless, about 7,300 tons of grain were successfully transacted, worth about $2.3 million USD. Figure B.10 shows the cumulative sales over the platform during the duration of the study.

Overall, take-up of Kudu among treated study traders was high, with 80% of treated traders posting to the platform at least once and 22% successfully trading on the exchange. However, take-up was much lower among farmers, with only 26% of treated households posting to the platform and less than 2% successfully transacting on the platform. Therefore, Kudu was a system primarily used by intermediaries rather than farmers directly, a point to which we return below.

3 Reduced Form Effects on Trade Flows and Market Prices

We now turn the reduced form effects of the platform on trade flows and market prices. One note on terminology: we use the phrase "reduced form effects" to refer to the observed (post-treatment) difference between treatment and control markets (or more precisely, treated and control trade routes, meaning pairs of markets). As we will address explicitly in Section 4, this may not capture the full treatment effect of the intervention, as prices and trade flows in control markets may be affected through general equilibrium forces. However, as we will describe later, this reduced form variation will be key to estimating our model and ultimately quantifying the aggregate treatment

[15] In recognition of this, Kudu developed a feedback system that sent a message back to unmatched sellers stating "You would have had to ask for X price in order to match on Kudu." However, this failed to align prices, as average ask prices remained above bids the for the duration of the study.

effects. We therefore present this variation here.

Figures 2-3 presents these reduced form effects on several outcomes: whether any trade is occurring between subcounties, the number of traders engaged in trade between subcounties, the volume of trade flowing between subcounties, and price dispersion between markets. The first three of these outcomes is drawn from our panel survey of traders, in which we asked detailed questions about their trading behavior at the subcounty level (the level of randomization). The last is drawn from our market-level price surveys and is therefore at the market dyad level. The lefthand side of Figures 2-3 presents non-parametric local Fan regressions of each outcome on the distance, separately for treatment routes (pairs of subcounties or markets in which both destination and origin are in the treatment group and therefore potentially see a Kudu-driven reduction in trade costs along the route) and control routes (all other routes). Distance is measured as the road distance of the shortest route connecting the two.

Before examining the reduced form difference between treatment and control routes, we first note some important patterns observed on control routes. In the upper left panel, we see that while the probability of any trade is high for nearby subcounties, this diminishes rapidly with distance. The probability of any trade occurring between the subcounties is close to zero beyond 200km distance. Consistent with this, the number of traders (second row) and total trade volumes (third row) also falls quickly with distance. These increasing trade costs with distance lead to notably higher price gaps between markets located at further distances, as shown in the bottom row. This salient extensive margin of trade is consistent with the presence of fixed costs to trade, which will feature prominently in our model in Section 4. The notable decline in extensive margin trade with distance suggests these fixed costs grow with distance, a feature which will shape the economic geography of treatment effects.

When we compare these patterns with those along treated routes, we see a higher probability of any trade (first row) and a larger number of traders engaged in trade (second row) along treated routes, suggesting greater trade on the extensive margin. We also see larger volumes of trade flows (third row) and lower price dispersion (fourth row). Notably, all effects are concentrated among relatively nearby markets, suggesting strong heterogeneity by distance. Beyond about 200km, markets do not trade directly, and the introduction of Kudu does not appear to alter patterns of trade.

To gauge the statistical significance of these differences, the righthand panel of Figures 2-3 presents the "reduced form effect," i.e. the difference between treatment and control line in the lefthand figure, in black. It also presents the bootstrapped 90% (in light gray) and 95% confidence intervals (in dark gray).[16]. We see effects for routes of relatively nearby markets are strongly statistically significant, tapering off to precise zeros around 200km.

We also presents results in regression form in Tables 1 and 2, running the following specification:

$$Y_{dr} = \alpha + \beta_1 T_d + \beta_2 D_d + \varepsilon_{dr} \tag{1}$$

Here, Y_{dr} is the outcome of interest in subcounty or market dyad d in round r, pooling all post-treatment survey rounds in the same analysis. For the first three columns, the outcome of interest is whether any trade is reported, the number of traders trading, and the volume of trade flowing between subcounty dyads, as reported by traders in the traders midline and endline. In the fourth column, the outcome is the inverse hyperbolic sine transformation of the absolute value of the gap between prices across each possible market dyad (d) in round (r), which is two-week intervals as measured in the market survey. These outcomes are regressed on T_d, a dummy for both members of the pair being in treated subcounties and D_d, a measure of the shortest road distance between the pair. Standard errors are clustered two-way by each subcounty (the unit of randomization).[17]

We first turn to Table 1, which presents results for the full sample of dyads. While we see point estimates consistent with the effects observed in Figures 2-3 – higher probability of any trade, larger number of traders operating between subcounties, greater trade volumes, and lower price dispersion – these effects are only marginally or are not significant. However, recall in Figures 2-3 that we note a striking pattern by distance. Effects are strongly concentrated among nearby markets. In fact, we

[16] The bootstrap is clustered by the sending subcounty. Two-way clustering by both sender and receiver subcounties is a work in progress. Appendix Figure B.11 presents p-values from randomization inference, which are below 0.1 for nearby market pairs, indicating significant results consistent with the main figure. Regression results in Tables 1 and 2 present results with two-way clustering by both sender and receiver subcounty.

[17] Dyads in the same subcounty are dropped both here and in Figures 2-3 , as they mechanically have the same treatment status.

see almost no difference in trade flows beyond 200km, the point at which the probability of direct trade drops close to zero. Table 2 explores this further, estimating Equation 1 separately for market pairs above and below 200km (which is close to the median distance observed in our sample). We again see that effects for all outcomes are concentrated in nearby subcounties and markets, with significantly higher probability of trade occurring, a larger number of traders operating between subcounties, and higher trade volumes. We also see marginally significantly lower price dispersion by 15% (p-value of 0.12). In contrast, we see almost no effect of the platform on direct trading outcomes (extensive margin connections, number of traders, and trade volumes) for markets that are at above median distance. Although point estimates suggest that price convergence results may persist along a farther distance, effects sizes are much smaller and no longer statistically significant.

The platform therefore generates meaningful differences in trade flows and price dispersion among nearby markets. However, it falls to live up to the often-touted promise of such digital marketplaces to *directly* connect remotely-located farmers and markets with urban consumers.

3.1 Unpacking Price Convergence

What is driving the observed price convergence? Figure 4 presents reduce form effects on price levels in relative surplus vs. relative deficit areas, as measured by average marketed surplus per farmer at baseline. First, we note in the lefthand panel that, as expected, prices in the control group are higher in relative deficit areas and lower in relative surplus areas. However, we see a less steep relationship in the treatment group, as with relatively lower prices prices in treated deficit markets and higher relative prices in treated surplus markets. The righthand panel presents this reduce form effect, along with the 90% and 95% bootstrapped confidence intervals. We see that prices are weakly lower in deficit areas and statistically significantly higher in surplus areas.

Table 3 presents similar results in regression form. We see in Column 1 that the overall effect on average price levels is a statistical zero. This is consistent with the netting out of two competing effects seen in the previous figure (the density in the right-hand panel of Figure 4 shows that for the median trading center, the average price effect is roughly zero). Column 2 presents heterogeneity by baseline average marketed surplus. We again see that prices are weakly lower in relative deficit areas (as evidenced by the negative coefficient on the treatment term) and higher in relative surplus areas (as evidenced by the significant and positive coefficient on the interaction term). With an average baseline marketed surplus of about 1 ton, these effects almost exactly offset each other for the median market. Column 3 divides our sample into areas of relative surplus and deficit, with the cutoff defined by being above or below the average amount of 1 ton. First, we note that prices are significantly lower in surplus areas overall, as expected. Second, we see that, with the introduction of Kudu, these surplus area experience significantly higher prices than they would have otherwise, while deficit areas experience weakly lower prices.

4 Experimenting in Equilibrium

The reduced form empirical patterns presented in the previous section are key to understanding Kudu's impact on trade. They show that the intervention did generate experimental variation in aggregate outcomes between treatment and control markets and routes, and tell us whether and how different types of outcomes were impacted. However, these comparisons alone are not sufficient to correctly estimate the magnitude of treatment effects. Control markets and routes are also likely to have been affected via their trade connections to treated markets, and so they do not provide the right counterfactual comparison.

Importantly, this is not the type of spillover that could be avoided with tighter control of compliance with treatment assignment, or by randomizing over larger units. Rather, it is a challenge inherent to studying trade costs, which by nature involve linkages between treated and control units. We can still make use of the experimental variation to estimate treatment effects, but need to interpret it carefully. To do so, we make use of a trade model.

In this section, we describe the model that will guide our estimation. The model serves two purposes. First, it will guide how to separate direct effects on treated units from indirect equilibrium effects. This will enable us to use the experimental variation to correctly estimate the treatment effect of access to Kudu on trade between pairs of treated markets. Second, combined with estimation of some additional model parameters, it will enable us to measure the total impact of access to Kudu across many markets on market-level outcomes, including the indirect effects of exposure to changes in other places.

In order to capture the key empirical patterns of the previous section, the model must have two main features. First, it must allow for the role of the extensive margin of trade. At baseline, there are many market pairs that do not trade with one another. Furthermore, we see that exposure to treatment increases both the probability of any trade and the number of traders serving a given route. Second, there must be a role for frictions in matching between buyers and sellers, which are alleviated by Kudu. Below, we outline a model that has these two key realistic features, while remaining as simple and comparable to standard quantitative trade models in the literature as possible.

4.1 Model setup

4.1.1 Geography

There are locations $i, j \in \{1, ..., J\}$, each with a continuum of consumers of measure Z_i and a continuum of traders of measure N_i. Consumers in each location are endowed with L_i of a homogenous crop and income D_i from other activities, which we can think of as an "outside" good.

Moving the crop between locations incurs a multiplicative variable cost, τ_{ij} and a fixed cost F_{ij}, both denominated in terms of the outside good. The outside good is freely and costlessly traded, and its price is normalized to one so that it acts as a numeraire.

4.1.2 Demand

Consumers have Cobb-Douglas demand over the crop and the outside good, so that they spend a constant share of their income, Y_i, on the crop: $E_i = \alpha Y_i$.

Consumers must purchase the crop from traders selling in their home market. They have idiosyncratic match values with individual traders, indexed ω, representing realistic factors like, language, ethnicity, availability and location within the market, reliability, and so on. These match values, $\varepsilon(\omega)$, are i.i.d draws from a mean zero Gumbel distribution. Consumers make a discrete choice over sellers, choosing to buy the crop from one to maximize their utility.

Utility when purchasing the crop from trader ω is:

$$u(\omega) = \mathbf{ln}e - \mathbf{ln}p(\omega) + \mu\varepsilon(\omega)$$

where $p(\omega)$ is the price charged by trader ω, e is per capita expenditure on the crop and $\varepsilon(\omega)$ is the match value with trader ω

This ideal variety utility function gives rise to aggregate demand facing each trader in each location that is isomorphic to CES with monopolistic competition. Consumer welfare from the crop can therefore be summarized as $\frac{E_i}{P_i}$, where P_i is a standard CES retail price index, defined below. The crop itself is homogenous, not differentiated. However, sellers are imperfectly substitutable from the perspective of buyers, and therefore have market power.

4.1.3 Traders

Each location is home to a continuum of traders, with measure N_i. Traders buy the crop in their home market and can resell the crop anywhere after paying trade costs.

Traders have heterogenous productivity, represented by an operating cost (i.e. inverse productivity) $a(\omega)$, drawn from a distribution with CDF $G(a)$. Let this inverse productivity distribution be Pareto with shape parameter k and minimum a_L, so that:

$$G(a) = 1 - \frac{a_L^k}{a^k}$$

Traders can purchase as much of the crop as they want in their local market at the wholesale market price, p_i. They can resell in any market by paying the trade costs between their home market, i, and the destination market, j.

4.1.4 Search and matching frictions

All agents perfectly observe prices and trade costs in all locations.

However, in order to sell in a non-local market, selling agents must find specific buyers in that market. We attribute part of the fixed cost of selling in a destination market (F_{ij}) to a search and matching cost, S_{ij}, that a seller pays to match with potential buyers there. In reality, this is likely to be the cost of traveling to the market to find buyers, or of making phone calls to find out what the market is like and who is buying on a given day. Once this cost has been paid, the seller is "visible" to buyers in that market. The measure of sellers visible in any market j is Ω_j, indexed ω. Buyers then draw idiosyncratic match values with each visible seller, as described above.

4.2 Equilibrium

4.2.1 Trader optimization

Faced with the demands and costs described above, traders decide which destination markets to sell in, and what price to charge. In any destination market j, a trader ω whose home market is i faces demand:

$$x_{ij}(\omega) = \frac{p_{ij}(\omega)^{-\sigma} E_j}{P_j^{1-\sigma}}$$

where $p_{ij}(\omega)$ is the price that trader ω charges when selling the crop from i in destination j, E_j is aggregate expenditure on the crop in destination j, P_j is the price index in j, and σ is the own-price elasticity of demand $\sigma = \frac{\mu-1}{\mu}$.

The profit maximizing price charged by a traders serving a route ij is:

$$p_{ij}(\omega) = \left(\frac{\sigma}{\sigma - 1}\right) \tau_{ij} p_i a(\omega)$$

where the marginal cost depends on p_i, the price at which the crop was purchased in the home market, $a(\omega)$, the trader-specific productivity, and τ_{ij}, the variable cost of trading from i to j. As usual under CES-form demand with monopolistic competition, the optimal price features a constant markup over marginal cost. The productivity term $a(\omega)$ can be interpreted as the relative marginal cost of operation, so that more productive traders have lower costs and charge lower prices in all markets, and will therefore face higher demand. The most productive trader is the one with the lowest inverse productivity, a_L.

We can also define the retail price index in market i as $P_i^{1-\sigma} = \sum_j N_j \int_{a_L}^{a_{ji}^*} p_{ji}(a)^{1-\sigma} dG(a)$.

A trader's operating profits from serving market j are:

$$\pi_{ij}(\omega) = \left(\frac{1}{\sigma}\right)\left(\frac{\left(\frac{\sigma}{\sigma-1}\right)\tau_{ij}p_i a(\omega)}{P_j}\right)^{1-\sigma} E_j - F_{ij}$$

A trader will serve any market where they can earn non-negative operating profits. Therefore, there is a cutoff productivity level a_{ij}^* such that any trader in i with inverse productivity less than or equal to this level, $a(\omega) \leq a_{ij}^*$ will serve destination j,and otherwise will not.

The cutoff productivity is:

$$a_{ij}^* = \left(\frac{P_j}{\left(\frac{\sigma}{\sigma-1}\right)\tau_{ij}p_i}\right)\left(\frac{\sigma F_{ij}}{E_j}\right)^{\frac{1}{1-\sigma}}$$

4.2.2 Bilateral trade flows

We can characterize trade flows in terms of three margins: any trade, number of traders, and value of trade.

There will be some trade on a given ij route if the highest productivity (/lowest cost) trader can earn non-negative profits:

$$\left(\frac{1}{\sigma}\right)\left(\frac{\left(\frac{\sigma}{\sigma-1}\right)\tau_{ij}p_i a_L}{P_j}\right)^{1-\sigma} E_j > F_{ij} \tag{2}$$

Note that this does not depend on any agent-specific factors – only equilibrium features of each market (prices and expenditure), trade costs, and the highest possible trader productivity.

The number of traders who will serve a route ij is:

$$N_{ij} = \begin{cases} N_i \int\limits_{a_L}^{a_{ij}^*} dG(a) & \text{for } a_{ij}^* \geq a_L \\ 0 & \text{otherwise} \end{cases}$$

Under the assumption that $G(a)$ is Pareto, this implies that, for $a_{ij}^* \geq a_L$:

$$N_{ij} \quad = \quad N_i - N_i a_L^k P_j^{-k}\left(\left(\frac{\sigma}{\sigma-1}\right)\tau_{ij}p_i\right)^k (\sigma F_{ij})^{\frac{-k}{1-\sigma}} E_j^{\frac{k}{1-\sigma}} \tag{3}$$

The value of trade on a route ij is:

$$M_{ij} = \begin{cases} N_i \int\limits_{a_L}^{a^*_{ij}} r_{ij}(a)\, dG(a) & \text{for } a^*_{ij} \geq a_L \\ 0 & \text{otherwise} \end{cases}$$

where $r_{ij}(a) = \left(\frac{\left(\frac{\sigma}{\sigma-1}\right)\tau_{ij} p_i a(\omega)}{P_j} \right)^{1-\sigma} E_j$ is the revenue or value of trade per trader.

Under the assumption that $G(a)$ is Pareto, this implies that, for $a^*_{ij} \geq a_L$:

$$M_{ij} \quad = \quad N_i \left(\frac{\left(\frac{\sigma}{\sigma-1}\right)\tau_{ij} p_i}{P_j} \right)^{1-\sigma} E_j \frac{k a_L^{1-\sigma}}{1-k-\sigma} \left(\left(\frac{P_j}{a_L \left(\frac{\sigma}{\sigma-1}\right)\tau_{ij} p_i} \right)^{1-k-\sigma} \left(\frac{\sigma F_{ij}}{E_j} \right)^{-k} - 1 \right) \tag{4}$$

4.3 Taking the model to data

In order to use the model to guide the estimation of treatment effects from the experiment, we will first make some more specific assumptions about the form of trade costs and then derive estimating equations that correspond to the observed trade outcomes in our data.

4.3.1 Specification of trade costs

Traders in treated markets have access to Kudu, which allows them to match with buyers in other traded markets via the platform. We will interpret this access as (potentially) reducing the search and matching component of the fixed cost of trade. Kudu enables a seller to be visible to buyers in a distant market by simply posting on the platform, rather than having to travel or make calls.

To facilitate estimation, we will specify the form of variable and fixed trade costs as follows:

$\tau_{ij}^{1-\sigma} \equiv d_{ij}^{-\gamma} e^{u_{ij}}$

$F_{ij} \equiv \exp\left(\phi + \theta \mathbf{ln} d_{ij} + \beta_1 \mathbf{1}^K_{ij} + \beta_2 \mathbf{ln} d_{ij} \mathbf{1}^K_{ij} - v_{ij} \right)$

where $u_{ij} \sim N\left(0, \sigma^2_u\right)$ and $v_{ij} \sim N\left(0, \sigma^2_v\right)$

and $\mathbf{1}^K_{ij} = \begin{cases} 1 & \text{if both i and j are treated} \\ 0 & \text{else} \end{cases}$

and d_{ij} is the distance between i and j

4.3.2 Estimating equations

In estimating the model, we are primarily interested in estimates of β_1 and β_2, which describe the impact of the intervention on bilateral trade costs. However, it will also be useful to get estimates of the other model parameters, in order to interpret magnitudes and consider counterfactual policy scenarios.

Before describing these, it will be useful to note that we can state the three equations describing the margins of trade above as:

$$\text{Any trade:} \qquad T_{ij} = \begin{cases} 1 & \text{for } a_{ij}^* \geq a_L \\ 0 & \text{otherwise} \end{cases}$$

$$\text{Number of traders:} \qquad N_{ij} = \begin{cases} N_i \int\limits_{a_L}^{a_{ij}^*} dG\left(a\right) & \text{for } a_{ij}^* \geq a_L \\ 0 & \text{otherwise} \end{cases}$$

$$\text{Value of trade:} \qquad M_{ij} = \begin{cases} N_i \int\limits_{a_L}^{a_{ij}^*} r_{ij}\left(a\right) dG\left(a\right) & \text{for } a_{ij}^* \geq a_L \\ 0 & \text{otherwise} \end{cases}$$

Using the assumptions above about the form of trade costs, we can derive three estimating equations, as follows.

Any trade

We take logs of Equation 2 to get:

$$\ln\tfrac{1}{\sigma} + \left(1-\sigma\right)\ln\tfrac{\sigma}{\sigma-1} + \left(1-\sigma\right)\ln a_L + \left(1-\sigma\right)\ln p_i + \left(\sigma-1\right)\ln P_j + \ln E_j + \left(1-\sigma\right)\ln\tau_{ij} - \ln F_{ij} \quad > \quad 0$$

Now using the parameterization of τ_{ij} and F_{ij}, and collecting terms, we can get:

$$\zeta_0 + \zeta_j + \frac{\left(1-\sigma\right)}{\sigma_\eta^2}\ln p_i + \left(\gamma-\theta\right)\ln d_{ij} - \beta_1 \mathbf{1}_{ij}^K - \beta_2 \mathbf{ln} d_{ij} \mathbf{1}_{ij}^K + \eta_{ij} > 0 \tag{5}$$

where ζ_j is a destination market fixed effects, and $\eta_{ij} = v_{ij} + u_{ij}$ so that $\eta_{ij} \sim N\left(0, \sigma_\eta^2\right)$ where $\sigma_\eta^2 = \sigma_v^2 + \sigma_u^2$

Dividing through by σ_η^2 and making use of the normal distribution of η_{ij}, we get an equation for the probability of there being any trade from i to j:

$$\rho_{ij} = \Phi\left(\hat{\zeta}_0 + \hat{\zeta}_j + \frac{(1-\sigma)}{\sigma_\eta^2}\ln p_i + \frac{(\gamma-\theta)}{\sigma_\eta^2}\ln d_{ij} - \frac{\beta_1}{\sigma_\eta^2}\mathbf{1}_{ij}^K - \frac{\beta_2}{\sigma_\eta^2}\mathbf{ln}d_{ij}\mathbf{1}_{ij}^K\right) \tag{6}$$

where hatted variables are equal to their counterparts in 5, divided by σ_η^2, e.g. $\hat{\zeta}_0 = \frac{\zeta_0}{\sigma_\eta^2}$.

This yields our first moment condition:

$$\mathbf{Pr}\left(T_{ij}=1\right) - \Phi\left(\hat{\zeta}_0 + \hat{\zeta}_j + \frac{(1-\sigma)}{\sigma_\eta^2}\ln p_i + \frac{(\gamma-\theta)}{\sigma_\eta^2}\ln d_{ij} - \frac{\beta_1}{\sigma_\eta^2}\mathbf{1}_{ij}^K - \frac{\beta_2}{\sigma_\eta^2}\mathbf{ln}d_{ij}\mathbf{1}_{ij}^K\right) = 0$$

Number of traders

Estimation of a log-linear equation describing the number of traders serving a route will be biased by selection on the any trade margin – the set of routes with non-zero traders have unobserved features that are systematically different from those with zero traders. To avoid this problem, we rearrange Equation 4 to consider the fraction of traders based in market i that do not serve market j:

$$\ln\left(1-\tfrac{N_{ij}}{N_i}\right) \quad = \quad k\left(\ln\tfrac{\sigma}{\sigma-1} + \ln a_L - \tfrac{1}{1-\sigma}\ln\sigma\right) + k\ln p_i - k\ln P_j + \tfrac{k}{1-\sigma}\ln E_j + k\ln\tau_{ij} - \tfrac{k}{1-\sigma}\ln F_{ij}$$

Now using the parameterization of trade costs from the previous section, and collecting terms we get:

$$\ln\left(1-\frac{N_{ij}}{N_i}\right) = \varphi_0 + \varphi_j + k\ln p_i + \frac{k\left(\gamma-\theta\right)}{1-\sigma}\ln d_{ij} - \frac{k\beta_1}{1-\sigma}\mathbf{1}_{ij}^K - \frac{k\beta_2}{1-\sigma}\mathbf{ln}d_{ij}\mathbf{1}_{ij}^K + \varphi_{ij}$$

where $\varphi_{ij} = \frac{k}{1-\sigma}\left(v_{ij}+u_{ij}\right)$ is a mean zero normally distributed error term.

Value of trade on route

Both the any trade margin and the composition of heterogenous traders serving a route are important for determining the value of trade on a route, and we need to account for both factors when taking the model to the data.

Starting from Equation 4 and taking logs, we get:

$$\ln M_{ij} \;=\; (1-\sigma)\ln\tfrac{\sigma}{\sigma-1} + (1-\sigma)\ln p_i + \ln N_i + (\sigma-1)\ln P_j + \ln E_j + (1-\sigma)\ln\tau_{ij} + \ln V_{ij}$$

where $V_{ij} = AW_{ij}$ and $W_{ij} = \max\left\{\left(\frac{a^*_{ij}}{a_L}\right)^{1-k-\sigma} - 1, 0\right\}$ and $A = \frac{ka_L^{1-\sigma}}{1-k-\sigma}$ is a constant.

Now using the parameterization of τ_{ij} from the previous section and collecting terms we get:

$$\ln M_{ij} = \psi_0 + \psi_j + (1-\sigma)\,p_i + \gamma\ln d_{ij} + u_{ij} + \ln V_{ij} \tag{7}$$

Note that $\ln M_{ij}$ is only defined when $T_{ij} = 1$. Conditioning on $T_{ij} = 1$ means that both $\ln V_{ij}$ and u_{ij} are correlated with $\ln d_{ij}$, and so if they are in the unobserved error term, estimates of γ will be biased. Therefore, we need estimates of these variables to include on the right side of the estimating equation. We can do this through a Heckman type correction for selection on both the any trade and trader composition margins:

$\ln\left\{exp\left[\delta\left(\hat{z}^*_{ij} + \hat{\bar{\eta}}^*_{ij}\right)\right] - 1\right\}$ is a consistent estimate for $\mathbb{E}\left[\ln W_{ij}|., T_{ij} = 1\right]$, where where $\delta \equiv \frac{\sigma_\eta(k-\sigma+1)}{\sigma-1}$, $\hat{z}^*_{ij} = \Phi^{-1}\left(\hat{\rho}_{ij}\right)$ and $\hat{\rho}_{ij}$ is the predicted value from Equation 6, and δ is a coefficient to be estimated.

$B\frac{\phi\left(\hat{z}^*_{ij}\right)}{\Phi\left(\hat{z}^*_{ij}\right)}$ is a consistent estimate for $\mathbb{E}\left[\ln u_{ij}|., T_{ij} = 1\right]$, where $\hat{z}^*_{ij}$ is defined as above and $B = \frac{\text{corr}(u_{ij},\eta_{ij})}{\left(\frac{\sigma_u}{\sigma_\eta}\right)}$ is a coefficient to be estimated.

5 Model Results

5.1 Estimating Equations

The above model generates the following three estimating specifications, each corresponding to an outcome observed in Figure 2 and therefore mapped to the reduced form variation driven by the experiment.

$$\mathbf{Pr}\left(T_{ij} = 1\right) = \Phi\left(\hat{\zeta}_0 + \hat{\zeta}_j + \frac{(1-\sigma)}{\sigma^2_\eta}\ln p_i + \frac{(\gamma-\theta)}{\sigma^2_\eta}\ln d_{ij} - \frac{\beta_1}{\sigma^2_\eta}\mathbf{1}^K_{ij} - \frac{\beta_2}{\sigma^2_\eta}\mathbf{1}^K_{ij}\mathbf{ln}d_{ij} - \hat{\phi}_i\right) \tag{8}$$

Equation 8 states that the probability of any trade from market i to market j is lower when the wholesale price in the sending market is high a more expensive source (mediated by how price sensitive consumer are, via σ), when markets are further apart in distance (mediated by the elasticities of fixed and variable trade costs with respect to distance, θ and $-\gamma$ respectively), and by whether the route is treated with Kudu ($\mathbf{1}_{ij} = 1$) (mediated by the treatment effect of Kudu on fixed costs, β_1). We also allow this treatment effect of Kudu to vary by distance β_2. Finally, all of these effects are mediated by σ_η^2, the combined variance of the idiosyncratic shocks to fixed and variable trade costs, which determines whether the above shifts to the profitability of trading are sufficient to induce the market into trading.

$$\ln\left(1 - \frac{N_{ij}}{N_i}\right) = \varphi_0 + \varphi_j + k\ln p_i + \frac{k\left(\gamma - \theta\right)}{1-\sigma}\ln d_{ij} - \frac{k\beta_1}{1-\sigma}\mathbf{1}_{ij}^K - \frac{k\beta_2}{1-\sigma}\mathbf{1}_{ij}^K\mathbf{ln}d_{ij} + \varphi_{ij} \tag{9}$$

Equation 9 states that the share of traders in market i *not* serving market j is lower when the source market is cheaper ($\ln\ p_i$), the distance is shorter, and when the route is treated. The shape of the trader productivity distribution, κ, governs how many new traders are pulled across the threshold of profitability when costs fall.

$$\ln M_{ij} = \psi_0 + \psi_j + (1-\sigma)\ln p_i + \gamma\ln d_{ij} + B\frac{\phi\left(\hat{z}_{ij}^*\right)}{\Phi\left(\hat{z}_{ij}^*\right)} + \ln\left\{exp\left[\delta\left(\hat{z}_{ij}^* + \hat{\bar{\eta}}_{ij}^*\right)\right] - 1\right\} \tag{10}$$

Finally, Equation 10 presents the value of trade flowing from market i to market j. The last two terms address selection into this equation due to treatment-induced extensive margin effects (via a Heckman-style correction a la Helpman et al. (2008)), accounting, respectively, for the fact that as treatment reduces trade costs, the set of routes with any trade changes, as does the composition of traders serving those routes. The remainder of the equation states that, once we have controlled for those extensive margins, the intensive margin of trade only depends on variables costs, which are themselves a function of sender price and distance.

5.2 Identification

Note that in order to identify the impact of Kudu on trade costs and run counterfactuals from Equations 8-10, we must have unbiased estimates of σ, σ_η^2, and κ, as well as β_1 and β_2.[18] Why? First, the effect of treatment on any trade (Equation 8) and the share of traders (Equation 9) is mediated by these parameters, so if we wish to separate out the direct effect of Kudu on fixed costs (β_1 and β_2), we must identify these parameters. For example, note the coefficient on the treatment indicator in Equation 8 is a function of both β_1, the reduction in fixed cost driven by Kudu, and the variation in the idiosyncratic shocks to trade costs, σ_η^2, which determines the probability of a market being moved above the threshold of having any profitable trade as a result of a β_1 reduction in trade costs. Therefore, to identify β_1 from this coefficient, we must also identify σ_η^2. Similar logic holds for Equation 9: to identify the reduction in fixed cost driven by Kudu, β_1, from the coefficient on the treatment indicator, we must also identify how a given reduction in trade costs increases the share of traders for whom trade is profitable, which will be mediated by κ, the shape of the trader productivity distribution, and σ, the elasticity of demand. The second reason we need unbiased estimates of these parameters is that we intend to run counterfactuals, to identify the full aggregate impact of Kudu. As Kudu changes trade flows along directly treated routes, this will translate into price and expenditure changes through the network, which will be mediated by these key parameters.

We therefore need unbiased estimates of the coefficients on both the treatment indicator ($\mathbf{1}_{ij}^K$) and sender price ($\ln p_i$) in the above Equations 8-10. We get unbiased estimates of the coefficients on the treatment indicator as a direct consequence of randomization. The coefficients on sender price is more challenging, as we have not randomized price. However, we can get exogenous variation in price from *randomized exposure* to treatment in other subcounties, following the seminal work of Miguel and Kremer (2004). In our context, we define a measure of exposure to treated subcounties that has parallels to the market access literature (Donaldson and Hornbeck, 2016). Formally, let:

[18] Unbiased estimates of θ and γ are not needed, as the impact of distance on non-Kudu fixed and variable trade costs is policy invariant.

$$\mu_i \equiv \sum_j \mathbb{1}_{\mathbb{j}}^{\mathbb{T}} \frac{X_j}{d_{ij}}$$

where μ_i is the exposure of subcounty i, defined as the sum of exposure to all other treated subcounties j ($\mathbb{1}_{\mathbb{j}}^{\mathbb{T}} = 1$), weighted by subcounty j's market size (as proxied by j's total bilateral trade flows with all other markets, X_j) and by the inverse of distance between i and j (d_{ij}). As noted by Miguel and Kremer (2004), and generalized and formalized by Borusyak and Hull (2023), even if treatment is random, *exposure* is only random conditional on economic geography and other features of the baseline economic environment. Therefore, we follow Borusyak and Hull (2023) and run 1,000 placebo randomization draws. For each placebo draw, we construct the exposure measure for each subcounty. We then demean our realized exposure measure μ_i by the average of the exposure measure under the 1,000 placebo randomization draws, $\bar{z}_i$, such that $\tilde{\mu}_i \equiv \mu_i - \bar{z}_i$. The resulting measure $\tilde{\mu}_i$ therefore captures variation in exposure that arises purely from the realized randomization draw, which is exogenous.

In Table 4, we estimate the impact of exposure to treatment in other subcounties on own price using:

$$lnp_i = \alpha + \beta\tilde{\mu}_i + \varepsilon_i$$

Column 1 shows a significant impact on own price of randomized exposure to treatment in other subcounties.[19] However, we cannot from theory alone sign the direction of expected effects of this exposure, as treatment has a different effect in surplus subcounties, where it acts as an exogenous, positive shock to demand, and deficit subcounties, where it acts as an exogenous, positive shock to supply. In Column 2, we therefore construct this randomized variation in exposure to treated surplus and treated deficit subcounties separately. We see that (random) exposure to treated surplus subcounties reduced own price, as demand is diverted to treated subcounties and away from one's

[19] Standard errors are clustered by subcounty, the level of randomization. However, Borusyak and Hull (2023) suggest that this may be insufficiently conservative, as variation in exposure may be correlated at a geographic level greater than that of randomization. We therefore follow Borusyak and Hull (2023) and present randomization inference p-values in the notes of Table 4. Although the coefficient in Column 1 (for which theory has no predicted sign) is no longer significant under randomization inference, results from Columns 2 and 3 (for which theory does predict effects) remain statistically significant under this more rigorous standard.

own subcounty. Conversely, (random) exposure to treated deficit subcounties increase own price, as supply is diverted to treated subcounties, albeit this latter effect is not significant. In Column 3, we interact own treatment status with these randomized exposure measures. We see the diversion effects are most pronounced among controls, as evidence by the significant coefficients on exposure to treated surplus subcounties and exposure to treated deficit subcounties. Own treatment seems somewhat protective against this diversion, mitigating both coefficients (in absolute value).

These results, in addition to providing reduced form evidence of the presence of spillovers, are useful for generating additional exogenous variation in price. We use the specification from Column 3 in Table 4 as an instrument for price in Equations 8-10. This allows experimental identification of σ, σ_η^2, and κ, as well as β_1 and β_2. Intuitively, a route's own treatment status identifies the direct treatment effect on trade costs (β_1 and β_2), while randomized exposure identifies the spillover effects that operate through impacts on market-level changes in prices and expenditure (which are mediated by the parameters in the coefficient on lnp_i of σ, σ_η^2, and κ).

5.3 Results and Parameter Estimates

Here, we present the results of each of the above estimating equations, as well as the implied parameter estimates from a joint estimation using a generalized method of moments (GMM) estimator.

Tables 5-7 present results from each estimating equation. For all equations, we present specifications with and without distance interactions, though we will see it does ultimately not make much of a difference for most of our parameter estimates. Table 5 shows that, as expected, trade is less likely when the sender price is higher and for markets that are more distant from one another. Importantly, trade is more likely when the route has access to Kudu. We see similar effects for the share of traders in Table 6 and trade volumes in Table 7.[20]

Table 8 presents the implied parameters estimated jointly via a GMM estimator. Our estimator uses three sets of moment conditions, corresponding to the three estimating equations. First, we formulate equation 8 as a maximum likelihood problem and form moments from the first-order

[20]Note that due to the outcome being the share of traders in market i who are *not* serving market j, the intuition about the sign of the coefficients is flipped in Table 6.

conditions of this problem. Second, we formulate equations 9 and 10 as nonlinear least-squares problems and form moments from the first order conditions of these problems. Parameters that appear in multiple estimating equations have more than one corresponding moment condition. Thus, we are over-identified.

To increase computational efficiency, we use a nested-loop approach to find the parameter vector that minimizes the GMM objective. In the outer loop, we search over values for structural primitives of the model (σ, σ_η, θ, γ, β_1, $beta_2$, B). In the inner loop, we solve for the set of fixed effects that exactly match the corresponding moments, conditional on the trial values of the structural primitives.

Table 8 presents results. We see sensible estimates for our key parameters. Importantly, the parameter estimates on β_1 suggest that the direct impact of Kudu is to reduce the size of fixed costs by 27.9% on average across all markets, a sizable reduction in total fixed costs.

We now turn to what these implied parameters mean for the aggregate impact of Kudu.

6 Aggregate Impact of Kudu

6.1 Corrected impacts on trade flows

Armed with these experimentally estimated parameters, we can now implement the full model. To do so, we bring in data on baseline production, expenditure, number of traders, and distances across markets, in addition to the parameters estimated in the previous section.[21] We can now run counterfactuals, turning on and off the 27.9% reduction in trade costs along treated routes. This allows us to estimate total treatment effects, comparing trade flows in the presence of Kudu to trade flows in a counterfactual without Kudu, rather than to control routes, which may be invalid counterfactuals for treated routes. Notably, we can make comparisons to counterfactuals for both treated and control routes, and therefore calculate the "treatment effect on treated routes" and "treatment effect on control routes." It is necessary to consider both when estimating aggregate impacts.

[21] As well as two temporarily calibrated parameters (σ and a_L), which in future drafts we aim to estimate in our data.

Results are presented in Figure ??. We see the "naive reduced form effect" in black, the treatment effect one would estimate if simply comparing the (post-treatment) treated route - control route outcomes. In contrast, the red line presents the full treatment effect on treated routes and the blue line the full treatment effect on control routes. We see the naive reduced form implies too-large treatment effects in two ways: first, it slightly overestimates the treatment effect on treated routes, because it compares outcomes along treated routes to control routes, which see some trade diversion and therefore negative impacts. Control routes are thus an invalid counterfactual for what treatment routes would look like in the absence of Kudu. Second, this diversion from control routes represents a real loss in trade volumes for those routes, which a naive reduced form estimate would fail to capture.

Figure ?? presents the implications for effects on average trade flows across both treatment and control routes, taken together, with the implied aggregate effects under the naive reduced form estimate in black and the GE-corrected estimates in purple. There are two take-aways. First, the true, GE-corrected estimates are still positive and substantial: Kudu increases trade on both the extensive and intensive margin, on average, even once accounting for the negative spillovers to control routes. These effects are large, given that Kudu is a lightweight, cheap intervention. On average, Kudu increases total trade flows by 2%. However, secondly, we see that these implied aggregate impacts are substantially less than one would estimate using the naive reduced form results alone. For example, the impacts on total trade volumes are only about 43% – less than half – than implied under a naive reduced form.

6.2 Market-level impacts

Correctly estimating treatment effects on trade costs and therefore on bilateral trade is an important first step in understanding the role of search costs. However, the objects of interest to policymakers are not trade or trade costs per se, but rather how those translate into market-level outcomes relevant to welfare, such as prices.

Figure ?? shows wholesale prices for all markets in our sample, including both treatment and control markets. We see that treatment increases wholesale prices for surplus areas, but decreases them for deficit areas, on average. This implies there are winners and losers from the Kudu-induced reduction in trade costs, with net producers in surplus areas and net consumers in deficit areas gaining, and net consumers in surplus areas and net producers in deficit areas losing. However, because surplus areas contain more net producers and deficit areas more consumers, average welfare improves in both locations.

7 Distributional and welfare impacts

7.1 Implications of scale economies for take-up

We emphasized in Section 4 that scale economies can explain why reducing matching frictions induces an extensive margin impact on trade flows. Here, we show how scale economies can also explain why the take-up of Kudu is concentrated among traders, with limited direct use by smallholder farmers. Figure 5 presents the distribution of transaction size for farmers (black line) and traders (grey line). We see that traders are substantially larger than farmers. Using the model estimated in the previous section, we can calculate, for each trading route, the minimum "threshold" size necessary to make trade along that route profitable, given the fixed trade costs. The long dash presents the average threshold size in the absence of Kudu, while the short dash presents the – now lower – average threshold size with Kudu.

We note two implications of this figure. First, the decline in the threshold driven by Kudu shifts the minimum size for transacting to the left within the trader size distribution, suggesting that Kudu should lead to greater entry among traders. We have already seen in Table 2 that Kudu induces an increase in trader entry, with point estimates suggesting a 28% increase in entry overall – and a larger, significant 43% increase in entry for shorter distance routes, for which the threshold is lower and therefore a larger fraction of traders are closer to the cutoff. Further, Figure 5 implies that these new entrants induced into trade by Kudu should be smaller than incumbent traders. This is indeed what we observe. Column 1 of Table 9 suggests that traders operating on treated routes are on average smaller, in terms of variable profits, than those that operate on control routes. Column 2 suggests that, in particular, treatment allows a drop in the minimum size trader along a route.

The second notable feature of Figure 5 is that the Kudu-induced reduction in the threshold size still lies far to the right of the farmer size distribution. Even with a 28% reduction in fixed costs, farmers are simply still too small to make direct engagement in cross-market trade profitable in the face of scale economies to trade.[22] This suggests that – despite the commonly cited motivation of using technology to "cut out the middleman" and allow smallholders to directly connect to markets – size economies will make this a tall order.

7.2 Welfare impacts for traders

What is the impact of Kudu on traders? Table 10 presents effects on average trader profits.[23] The main specification pools the post-treatment survey rounds and runs:

$$Y_{ir} = \alpha + \beta Treat_i + \gamma Y_{i0} + \delta_r + X_i + \varepsilon_{dr} \tag{11}$$

In which Y_{it} is outcome Y for trader i in round r (either midline or endline), $Treat_i$ is a dummy for being a treated trader, Y_{i0} is the baseline level of the outcome variable, δ_r is a dummy for survey round, and X_i is a vector of controls.[24] Treatment effects are given by the coefficient β. We see that traders located in treated subcounties see a significant reduction in measured profits, by about 14% of their average value.

[22] We show here the *average* minimum transaction size across all routes. We do see about 2% of farmers take-up. These are the largest farmers, and they tend to engage in very short distance cross-market trade, for which the threshold is lowest.

[23] Since 92% of sample is comprised of maize traders, for variables for which we must specify the crop – i.e. volumes, markups, and prices – we present result for maize.

[24] Our pre-analysis plan specified that we would include baseline controls that are most predictive of the outcome. We do this by identifying controls to include via a double lasso procedure, set to predict endline profits, our main trader outcome. Those covariates considered were: gender, age, education, length of time in business, number of subcounties in which purchase, number of subcounties in which sell, profits, net revenues, annual costs, annual revenue, monthly costs, and markups, quantities purchased and sold, prices at which purchased and sold, revenue, net revenue, and cost per kg for maize and beans, all as measured at baseline. Those selected by the lasso procedures and therefore included in X_i are: baseline profits, baseline annual costs, and baseline monthly costs.

7.3 Welfare impacts for farmers

We saw that farmers are on average too small to adopt Kudu, even with a reduction in the minimum threshold size required for cross-market trade. However, this does not mean that farmers are not impacted by Kudu. As traders take-up Kudu and cross-market arbitrage improves, shifting market prices can impact farming households.

We run the following specification in Table 11:

$$Y_i = \alpha + \beta Treat_i + \gamma Y_{i0} + X_i + \varepsilon_i \tag{12}$$

in which Y_i is outcome Y for farmer i at endline, $Treat_i$ is a dummy for being a treated farmer, Y_{i0} is the baseline level of the outcome variable, and X_i is a vector of controls.[25] Treatment effects are given by the coefficient β. We see no statistically significant effect on total revenues, maize revenues, maize volumes sold, or price received for maize sold, on average. Point estimates are positive and, in some cases, quite large (for example, the point estimate on total revenues is 9.7% of average revenue), but estimates are imprecise.

Recall, however, that Kudu induced no average change in market prices. Rather, Kudu increased prices in surplus areas and reduced prices in deficit areas. These shifts in market prices can impact farming households, even if they do not directly use Kudu or engage in cross-market trade. Figure 6 therefore displays treatment effects on farmer revenue by relative surplus status. In the lefthand panel we see, as expected, that revenue increase with surplus status in both treatment and control. However, the righthand panel, which shows the treatment effect, along with the 90% and 95%, shows that treatment significantly increases farmer revenues in surplus areas (and decreases revenues in deficit, those this effect is not statistically significant).

In sum, although very few farmers directly use the platform, we see evidence of impacts on

[25]Our pre-analysis plan specified that we would include baseline controls that are most predictive of the outcome. We do this by identifying controls to include via a double lasso procedure, set to predict endline total revenue, our main farmer outcome. Those covariates considered were: gender, age, high level of educational attainment, number of household members, revenues, quantity sold, land holdings size, quantity harvested, number of times told, any sales at the market, percent sold to market, distance to market, distance to Kampala, total value of all assets, expenditures in the last 30 days, food expenditure in the past 30 days, value of inputs used in the last year, all as measured at baseline. Those selected by the lasso procedures and therefore included in X_i are: baseline revenues, baseline quantity sold, and baseline value of inputs used in the last year.

farmers through shifting market-level forces. Although this does not operate through the mechanism that policymakers may have had in mind – that mobile marketplaces like Kudu will enable the smallest, poorest farmers to directly access a wider market and bypass intermediaries – it appears that farmers can benefit from the passed through effects of arbitrage by those intermediaries. [26]

8 Why did the intervention work?

The Kudu intervention comprises a number of potentially distinct mechanisms through which impacts might have been achieved. We can exploit sub-experiments and heterogeneity analysis to investigate the mechanisms through which the intervention worked.

We have several ways to focus on the impact of price information alone. First, we conducted a sub-experiment that randomly rolled tranches of control markets into the SMS Blast treatment. In each of the 12 market survey rounds between October 21, 2016 and March 24, 2017 we rolled in three control markets to the SMS Blast, treating all study traders and farmers by sending them all of the price information on the platform that was shared with the treatment throughout the study. Then, subsequent to the household and trader endline surveys, we rolled in an additional 36 control trading markets to the SMS Blast and so observe four final rounds of market surveys with this system in place. Because this roll-in did not include promotion to Kudu or any on-the-ground support in using the platform, its impact isolates information effects. The most well-powered way of analyzing this roll-in experiment is to use the high-frequency market survey, since we can then match the biweekly monadic market prices and dyadic price gaps against the timing of the introduction of information to the market. In Table A.10 we provide two-way fixed effects estimates of the roll-in to initial control markets using monadic data and find no significant average price changes. This impression is confirmed visually in Figure B.12, which shows average prices for each of the roll-in tranches and shows that they continue to track control market prices very closely.[27]

[26] It is worth noting that the pass-through of market price shifts to farmgate will occur under any model of competition between traders and farmers, ranging from perfect competition to monopsony. Of course, the *degree* of pass-through – a key feature governing the *magnitude* of farmer gains from a primarily intermediary-used intervention – will be mediated by the degree of competition (and the shape of farmer supply). This is something we aim to explore in future work.

[27] This figure plots average maize buying prices across time, breaking the roll-in into five aggregated tranches that enter the treatment from the pooled control counterfactual.

In Table A.11 we use dyadic data for pairs untreated in the initial experiment to look at the effect of the roll-in on price dispersion, using biweekly data with round fixed effects and two-way clustering at the level of each sub-county. Unlike the core results where strong price convergence was found, here if anything we find evidence of divergence, even for nearby markets; however none of the effects are significant despite having similar power to the main results. This sub-experiment therefore conforms with the broader literature in finding no large improvements stemming from the information-only component of a market price interventions.

A second and quite unique window on information-only effects is provided by a sub-experiment that we named the 'Random Blast'. For this intervention, we randomly selected one treatment market for each round of the Blast, took the price information for that randomly selected market, and we sent that information to the entire treated network. For each biweekly round, then, we have thousands of traders, buyers, and farmers being simultaneously told about prices in a single market. By defining leads and lags of the month in which a given market is included, this gives us a very high-powered way of examining whether prices, the number of traders, or price dispersion in that market changes around the time its price information was shared nationally. In Figure B.13 we present these impacts in an event study format, and show that no lead or lag is significant for any outcome. This is perhaps particularly surprising in the bottom panel, which uses dyadic data only for treated market pairs (meaning that every market included in this estimation receives the Random Blast) and finds no evidence of panel price convergence with the included market around the time its information is shared. Particularly in a system that was already operating credibly at national scale, this provides quite a concrete demonstration that information alone is not moving market prices.

A third approach is to consider where we would expect heterogeneity in impacts if the core role of the Kudu platform was providing information. An information intervention should have impacts through *learning*, and given that the overall geography of surplus/deficit areas is presumably well understood by traders, learning would most obviously arise from high-frequency deviations from expected regional market price averages. We form this measure of an informational shock as follows. First, we calculate the average price for each month-of-year across all years for each district. Then, for each round of the market survey we calculate average deviation in that round and district from the typical price at that location and time of year. Moving into the dyadic data structure, we then calculate the absolute value of these deviations across the dyadic pair, so this represents the component of the price gap between two districts that would not have been anticipated given the typical seasonal differences and hence provides a measure of the informational innovation present in the price discovery from the treatment. We can then interact these price gaps with market-level treatment status to examine whether being treated causes faster price convergence for those treated pairs where the Blast information is revealing larger unexpected price gaps. We look at this interaction starting with the contemporaneous shocks, and then lag the shocks by one, two, and three periods to allow for the fact that arbitrage will take time to lower these price gaps. Table A.12 shows that while the contemporaneous 'shock' dispersion is as we would expect a very strong driver of overall price dispersion, not only does this relationship fade over the course of three lags, but at no time interval is the interaction with treatment significant. Hence, the treatment is not causing faster convergence in markets where the price information shows that the transient price gaps are largest.

If price information did no cause the effects observed, then where do the treatment effects come from? Several stylized facts are consistent with the mechanism embedded in the model that the intervention lowers the costs to finding new trading partners and hence acts like a reduction to the fixed costs of trade. First, the core heterogeneity we have uncovered has to do with increasing the flow of trade from net surplus to net deficit regions of the country, a form of heterogeneity that would have been well understood, and so this increase suggests cost reductions on a margin other than pure information. Second, intervention pushes traders into new markets and induces trade with new counterparts. Finally, we can use our survey data to show that these new trading relationships tend to be durable and outlast the initial deal formed through Kudu. From our survey data, we find that 43% of the traders who initially matched on Kudu report transacting again with that same individual off-Kudu – and at large volumes, with these repeat transactions in aggregate accounting for 7x the volumes initially conducted on-Kudu. This is consistent with the fact that the total volume of trade caused by the intervention based on trader-reported volumes is 4x of the volume of trade between treated markets that is observed on Kudu. Hence, while this conclusion is necessarily somewhat speculative, the bulk of evidence suggests that our impacts arise from lower market linkage costs and not from price information.

9 Conclusion

This study shows that search and matching frictions continue to inhibit trade in African agricultural markets. A trading platform to facilitate connections between buyers and seller resulting in greater trade flows on both the extensive and intensive margins. These increases in trade flows reduced price dispersion, increasings prices in surplus areas and decreasing prices in deficit. Results are consistent with the platform reducing the fixed cost of trade by 28%. These impacts yield benefits for net producers in surplus areas, net consumers in deficit areas, and overall average welfare.

Accounting for equilibrium effects is key to correct estimation of the impacts of trade cost interventions, even those randomized "at-scale." Control markets and routes are affected by trade and price effects in general equilibrium, and therefore controls no longer serve as valid counterfactuals. However, experimental variation – when interpreted through the lens of a quantitative general equilibrium model - can yield impacts. In our context, aggregate impacts on trade flows are 43% of those suggested by naive "reduced form" comparisons of treatment and control routes.

Finally, consistent with scale economies, the platform was used almost exclusively by intermediaries. This is consistent with scale economies in trade, which make interventions targeted at engaging farmers directly in cross-market trade unlikely to succeed, on the margin. However, farmers and consumers can benefit indirectly from market price changes due to trader adoption of such platforms. We see evidence that revenues increases significantly for farmers in surplus areas. Of course, the *degree* of pass-through – a key feature governing the *magnitude* of farmer gains from a primarily intermediary-used intervention – will be mediated by the degree of competition (and the shape of farmer supply). This is something we aim to explore in future work.

References

Aker, Jenny C, "Information from markets near and far: Mobile phones and agricultural markets in Niger," *American Economic Journal: Applied Economics*, 2010, *2* (3), 46–59.

— **and Isaac M Mbiti**, "Mobile phones and economic development in Africa," *Journal of economic Perspectives*, 2010, *24* (3), 207–32.

Allen, Treb, "Information Frictions in Trade," *Econometrica*, 2014, *82* (6), 2041–2083.

Barrett, Christopher B, "Smallholder market participation: Concepts and evidence from eastern and southern Africa," *Food policy*, 2008, *33* (4), 299–317.

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

Borusyak, Kirill and Peter Hull, "Non-Random Exposure to Exogenous Shocks," *Econometrica*, 2023.

Camacho, Adriana and Emily Conover, "The Impact of Receiving Price and Climate Information in the Agricultural Sector," *IDB Working Paper Series, No. IDB-WP-220*, 2011.

Donaldson, Dave and Richard Hornbeck, "Railroads and American Economic Growth: A Market Access Approach," *Quarterly Journal of Economics*, 2016.

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

Helpman, Elhanan, Marc Melitz, and Yona Rubinstein, "Estimating Trade Flows: Trading Partners and Trading Volumes," *The Quarterly Journal of Economics*, 2008, *123* (2), 441–487.

Hildebrant, Nicole, Yaw Nyarko, Giorgia Romangnoli, and Emilia Soldani, "Price Information, Inter-Village Networks, and "Bargaining Spillovers": Experimental Evidence from Ghana," *Working Paper*, 2020.

Jensen, Robert T, "Information, efficiency, and welfare in agricultural markets," *Agricultural Economics*, 2010, *41*, 203–216.

Miguel, Edward and Michael Kremer, "Worms: Identifying Impacts on Education and Health in the Presence of Treatment Externalities," *Econometrica*, 2004.

Mitra, Sandip, Dilip Mookherjee, Maximo Torero, and Sujata Visaria, "Asymmetric information and middleman margins: An experiment with Indian potato farmers," *Review of Economics and Statistics*, 2018, *100* (1), 1–13.

Nakasone, Eduardo, "The Role of Price Information in Agricultural Markets: Evidence from Rural Peru," *Working Paper*, 2014.

Newman, Neil, Lauren Falcao Bergquist, Nicole Immorlica, Kevin Leyton-Brown, Brendan Lucier, Craig McIntosh, John Quinn, and Richard Ssekibuule, "Designing and evolving an electronic agricultural marketplace in Uganda," in "Proceedings of the 1st ACM SIGCAS Conference on Computing and Sustainable Societies" 2018, pp. 1–11.

Rashid, Shahidur and Nicholas Minot, "Are staple food markets in Africa efficient? Spatial price analyses and beyond," Technical Report 2010.

Svensson, Jakob and David Yanagizawa, "Getting prices right: the impact of the market information service in Uganda," *Journal of the European Economic Association*, 2009, *7* (2-3), 435–445.

_ **and David Yanagizawa-Drott**, "Estimating Impact in Partial vs. General Equilibrium: A Cautionary Tale from a Natural Experiment in Uganda," *Working Paper*, 2012.

Teravaninthorn, Supee and Gael Raballand, "Transport Prices and Costs in Africa," *World Bank*, 2009.

Figures

Figure 1: **Price Differentials and Transport Costs.** The y-axis presents the absolute difference in prices across each market dyad (pair) in the sample. The solid black line presents the gap observed in prices across each pair of markets in our sample. The dotted line presents estimated transport costs. To generate this prediction, we asked surveyed traders to report the costs of traveling roundtrip along each of their five most commonly travelled routes and the vehicle size typically used. From this data, we construct an estimate of per kg transport costs, which we then estimate as a function of the km traveled. These transport costs represent an upper bound on the price dispersion that should be observed if transport costs are the only trade costs. The gray area represents excess price dispersion, the portion of price dispersion that cannot be explained by transportation costs.

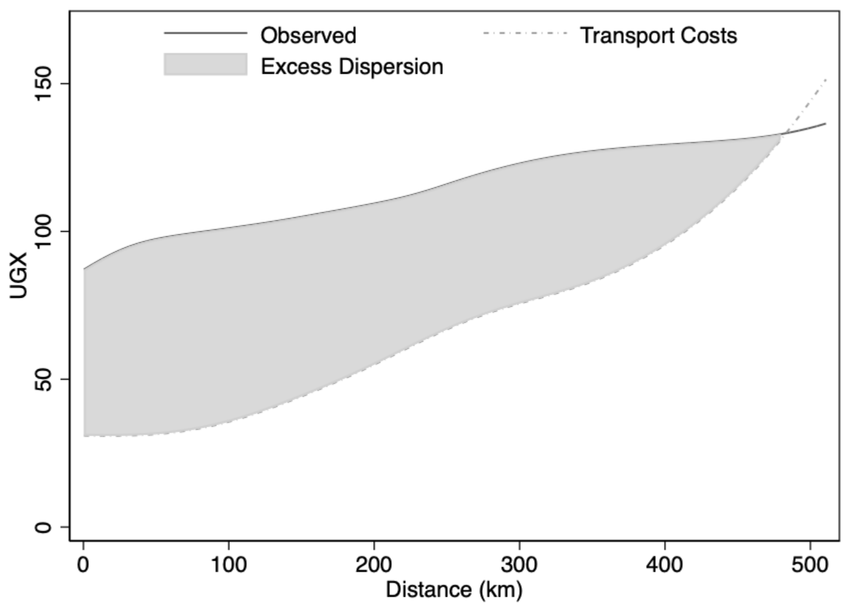


Figure 2: **Reduced form effects on trade linkages, number of traders, and trade volumes**. The top lefthand panel presents a non-parametric local Fan regression of the probability of trade along any treated route (dotted line), in which both markets are treated and therefore are connected by Kudu, vs. control routes (solid line), as a function of distance. The top righthand panel presents the "reduced form effect," i.e. the difference between treatment and control line in the lefthand figure, along with the bootstrapped 90% (light gray) and 95% confidence intervals (dark gray). Subsequent rows present similar results for the number of traders actively trading along a route (second row) and trade volumes (third row).

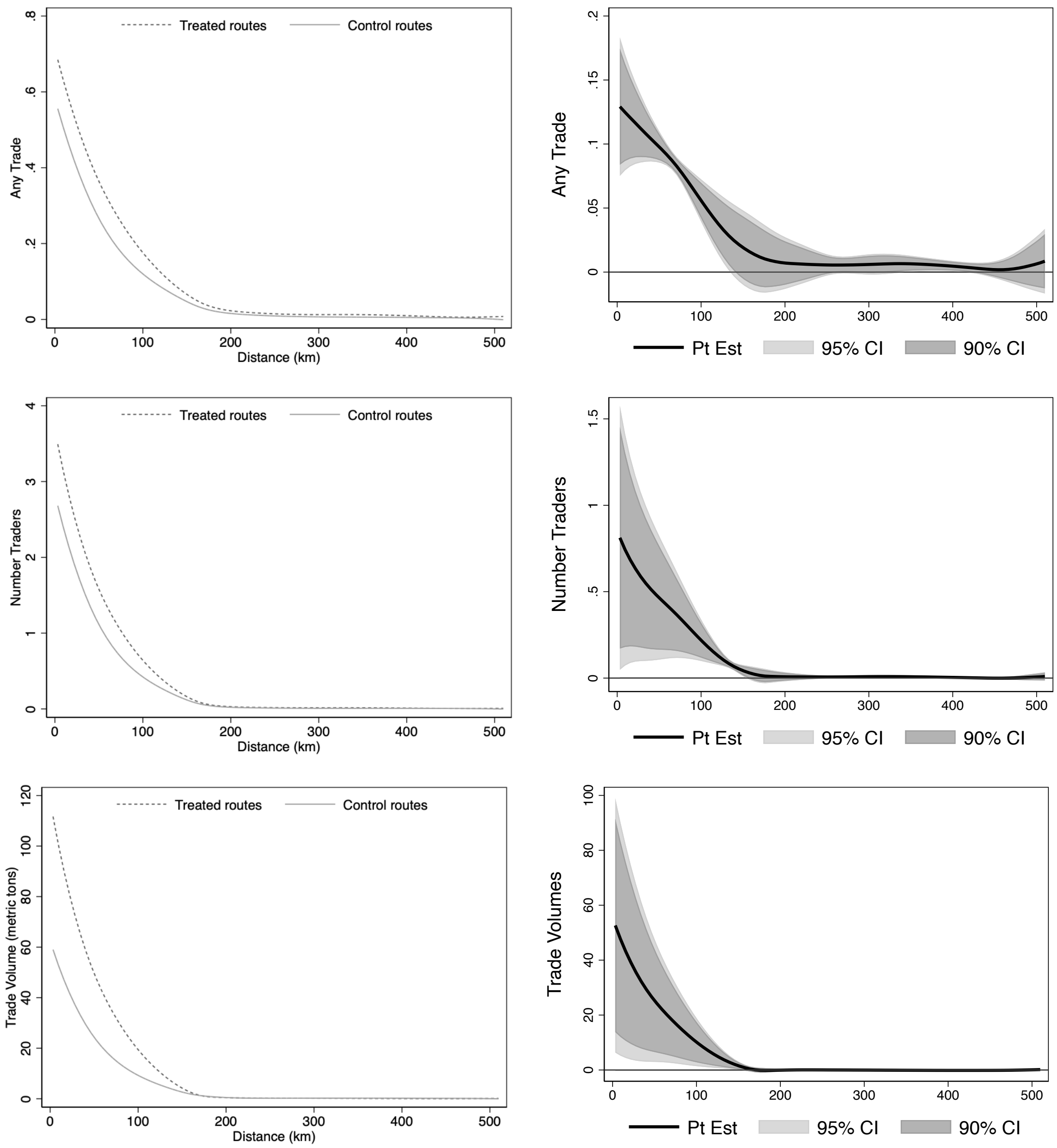


Figure 3: **Reduced form effects on price gaps**. The top lefthand panel presents a non-parametric local Fan regression of the price gaps across markets along any treated route (dotted line), in which both markets are treated and therefore are connected by Kudu, vs. control routes (solid line), as a function of distance. The top righthand panel presents the “reduced form effect,” i.e. the difference between treatment and control line in the lefthand figure, along with the bootstrapped 90% (light gray) and 95% confidence intervals (dark gray).

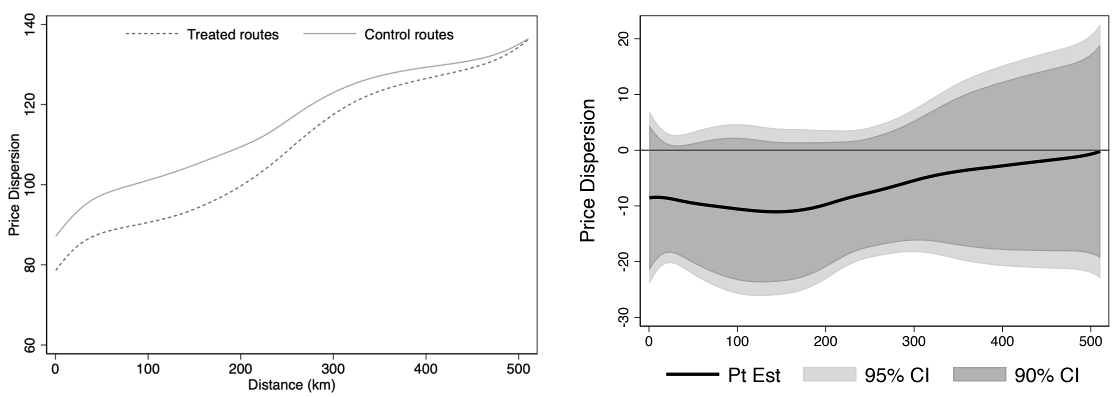


Figure 4: **Reduced form effects on price levels by relative surplus vs. deficit areas**. The left panel shows the level of prices in treatment vs. control markets, with respect to the average market surplus per farmer, as measure in ton at baseline. The right panel shows the difference between the two (the treatment effect), along with the 90% and 95% confidence intervals from a bootstrap estimation.

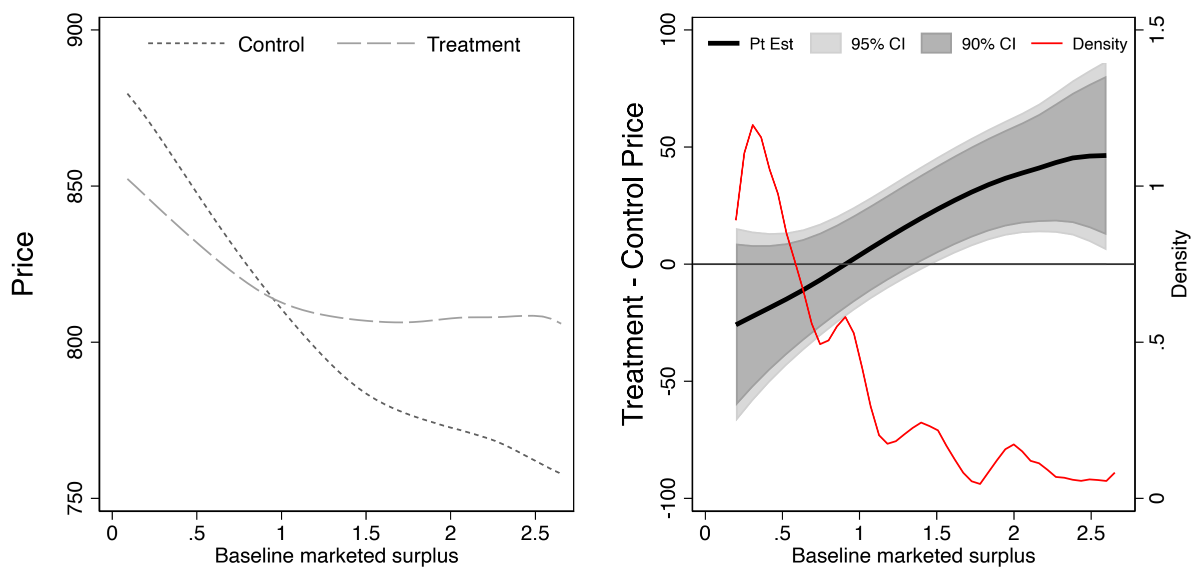


Figure 5: **Threshold transaction size**

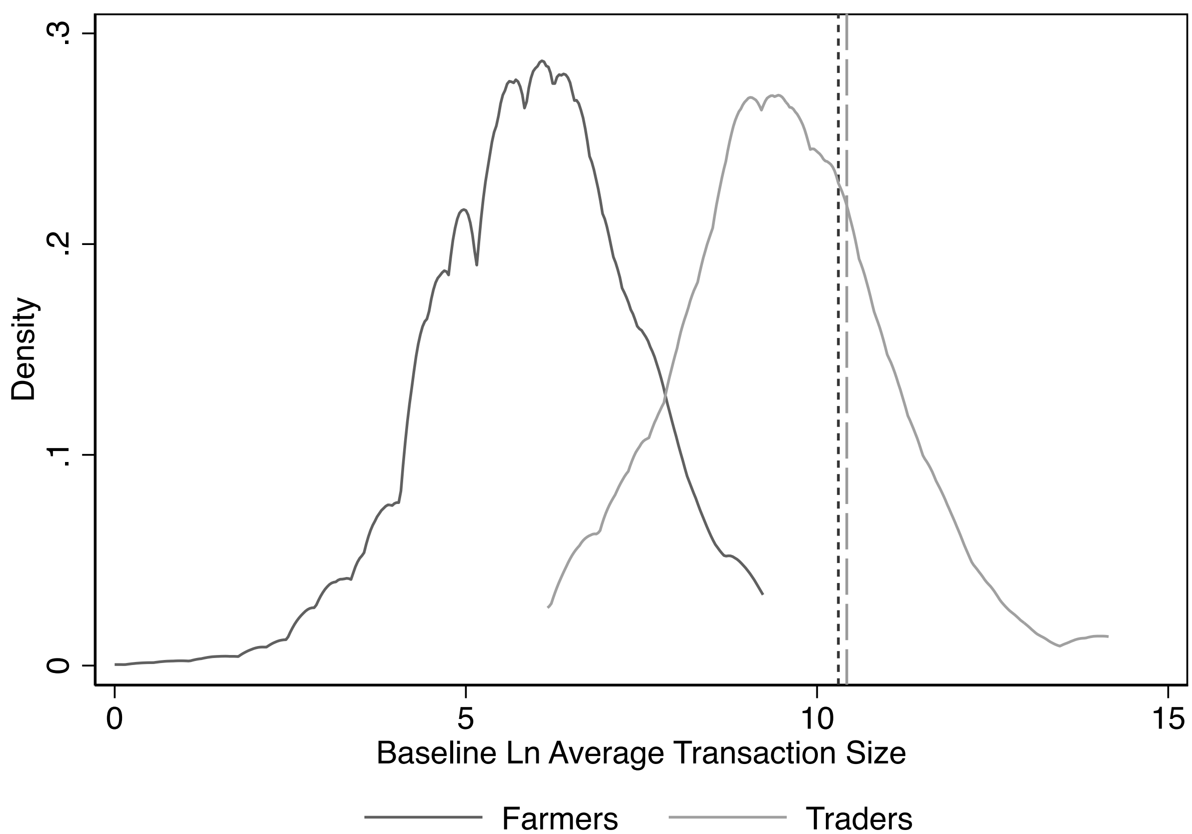
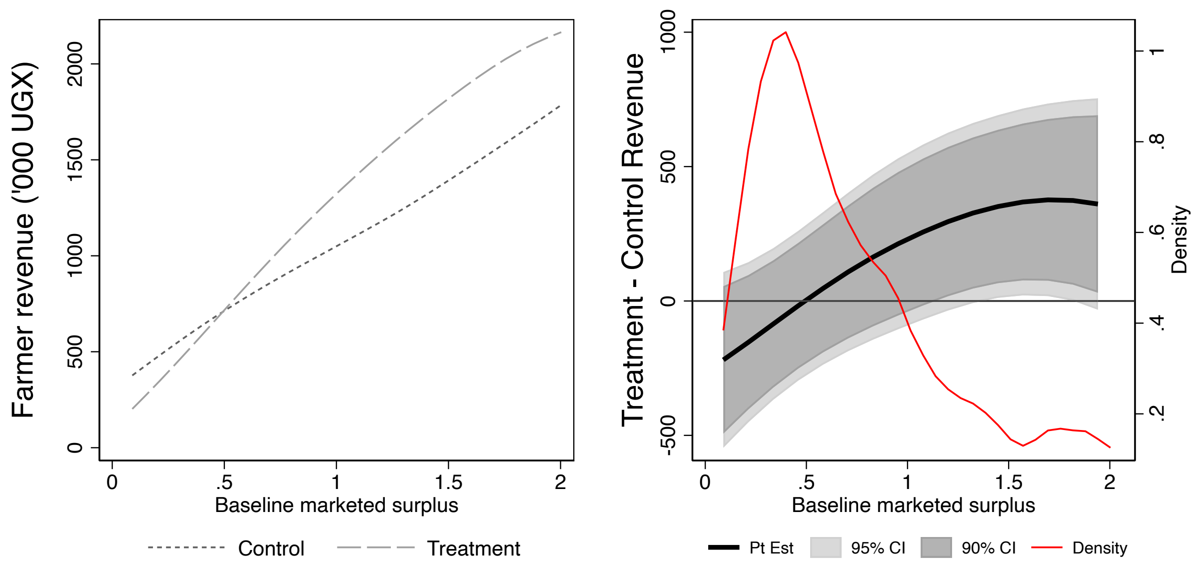


Figure 6: **Farmer revenue impacts by surplus vs. deficit**



Tables

Table 1: **Reduced form effects on trade linkages, number of traders, and trade volumes, and price dispersion**

	Any Trade	Number Traders	Volume (tons)	Price Dispersion
Both treated	0.02	0.08	4.56	-0.09
	(0.01)	(0.07)	(2.78)	(0.07)
Dist (10km)	-0.01***	-0.02***	-0.53***	0.01***
	(0.00)	(0.00)	(0.15)	(0.00)
Observations	5565	5565	5565	713713
Mean DV	0.06	0.20	4.82	4.56
P-val Both treat	0.25	0.26	0.10	0.21

Table 2: **Reduced form effects on trade linkages, number of traders, and trade volumes, and price dispersion by distance**

	Below 200km				Above 200km			
	Any Trade	Num Traders	Vol (tons)	Price Disp	Any Trade	Num Traders	Vol (tons)	Price Disp
Both treated	0.07**	0.37**	18.13**	-0.15	0.00	0.00	-0.07	-0.06
	(0.03)	(0.17)	(9.09)	(0.10)	(0.00)	(0.01)	(0.14)	(0.08)
Dist (10km)	-0.03***	-0.13***	-3.40***	0.02***	-0.00*	-0.00*	-0.01*	0.01***
	(0.00)	(0.02)	(0.95)	(0.01)	(0.00)	(0.00)	(0.00)	(0.00)
Observations	1722	1722	1722	224286	3843	3843	3843	489427
Mean DV	0.16	0.62	15.16	4.28	0.01	0.01	0.19	4.68
P-val Both treat	0.03	0.04	0.05	0.12	0.55	0.54	0.59	0.43

Table 3: **Reduced form effects on price levels by relative surplus vs. deficit areas**

	(1)	(2)	(3)	(4)
Treatment	-4.801	-20.78	-11.58	
	(11.56)	(16.40)	(13.74)	
Baseline marketed surplus		-28.66***		
		(7.716)		
Treat*Baseline marketed surplus		19.23*		
		(9.995)		
Surplus dummy			-48.60***	-47.88***
			(9.868)	(9.834)
Treat*Surplus dummy			29.39*	17.81*
			(15.10)	(9.369)
Treat*Deficit dummy				-10.26
				(13.62)
Observations	15211	15161	15211	15211
Mean DV	831.6	831.6	831.6	831.6
Mean Baseline Marketed Surplus		0.907		
Percent Surplus			0.275	0.275
P-Val Treat*Deficit=Treat*Surplus				0.0640
R2	0.844	0.847	0.847	0.847

Table 4: **Impact of experimental exposure** Exposure is defined as $\mu_i \equiv \sum_j \mathbb{1}_{\mathbb{j}}^{\mathbb{T}} \frac{X_j}{d_{ij}}$, where μ_i is the exposure of subcounty i, defined as the sum of exposure to all other treated subcounties j ($\mathbb{1}_{\mathbb{j}}^{\mathbb{T}} = 1$), weighted by subcounty j's market size (as proxied by j's total bilateral trade flows with all other markets, X_j) and by the inverse of distance between i and j (d_{ij}). As noted by Miguel and Kremer (2004), and generalized and formalized by Borusyak and Hull (2021), even if treatment is random, *exposure* is only random conditional on economic geography and other features of the baseline economic environment. Therefore, we follow Borusyak and Hull (2021) and run 1,000 placebo randomization draws. For each placebo draw, we construct the exposure measure for each subcounty. We then demean our realized exposure measure μ_i by the average of the exposure measure under the 1,000 placebo randomization draws, $\bar{z}_i$, such that $\tilde{\mu}_i \equiv \mu_i - \bar{z}_i$. The resulting measure $\tilde{\mu}_i$ therefore captures variation in exposure that arises purely from the realized randomization draw, which is exogenous. Column 1 regresses ln price on this exposure measure. Column 2 constructs this same procedure but separately for exposure to surplus and deficit market. Column 3 includes treatment interaction terms, allowing the impact of exposure to vary by own treatment status. The table presents standard errors clustered by subcounty, the unit of randomization. Table notes present p-values from randomization inference.

	Ln Price	Ln Price	Ln Price
Exposure to T markets	-0.10***		
	(0.02)		
Exposure to T surplus markets		-0.47***	-0.64***
		(0.16)	(0.22)
Exposure to T deficit markets		0.21	0.71***
		(0.13)	(0.26)
Treat x Exposure to T surplus markets			0.26
			(0.32)
Treat x Exposure to T deficit markets			-0.70**
			(0.29)
Treat			0.01
			(0.02)
RI p-value: Exposure to T markets	0.82		
RI p-value: Exposure to T surplus markets		0.04	0.10
RI p-value: Exposure to T deficit markets		0.59	0.07
RI p-value: Treat*Exposure to T surplus markets			0.39
RI p-value: Treat*Exposure to T deficit markets			0.11
RI p-value: Treat			0.88

Table 5: **Any trade**

	Any trade	Any trade
ln p_i	-2.691*	-2.703*
	(1.400)	(1.415)
ln d_ij	-1.167***	-1.172***
	(0.056)	(0.058)
Treat_ij	0.175*	0.104
	(0.093)	(0.415)
Treat_ij*ln d_ij		0.017
		(0.093)
Observations	17176	17176
Mean DV	0.034	0.034
Fixed effects	Receiver	Receiver

Table 6: **Fraction trader**

	Fraction traders	Fraction traders
ln p_i	0.038*	0.032*
	(0.020)	(0.018)
ln d_ij	0.017***	0.015***
	(0.002)	(0.001)
Treat_ij	-0.002	-0.051**
	(0.001)	(0.025)
Treat_ij*ln d_ij		0.009**
		(0.004)
Observations	20462	20462
Mean DV	-0.005	-0.005
Fixed effects	Receiver	Receiver

Table 7: **Trade volumes**

	Log value per N	Log value per N
ln p_i	-4.29	-4.30
	(3.16)	(3.16)
ln d_ij	-0.46**	-0.46**
	(0.22)	(0.22)
B	0.82***	0.82***
	(0.27)	(0.27)
δ	0.85	0.86
	(0.60)	(0.60)
Observations	780	780
Distance interaction	No	Yes
Fixed effects	Receiver	Receiver

Table 8: **Parameter estimates**

	No interaction	**Distance interaction**	**Parameter definition**
β_1	-0.279	-0.165	Treatment effect on fixed costs
β_2	N/A	-0.027	Treatment effect on fixed costs x distance
σ	5.290	5.300	Elasticity of demand wrt price
γ	0.460	0.460	Elasticity of variable trade costs wrt distance
θ	1.400	1.404	Elasticity of fixed trade costs wrt distance
κ	0.039	1.326	Shape parameter of trader productivity
σ_η^2	1.594	1.591	Variance of variable + fixed cost error terms

Table 9: **New entrants on average smaller**

	Profits (1000 UGX)	
	Mean	Min
Treat_ij	-1910.76**	-2049.28***
	(936.17)	(773.07)
ln d_ij	1189.85**	2054.22***
	(475.42)	(462.14)
Observations	780	780
Mean DV	7389.80	5337.57

Table 10: **Effects on trader profits**

	Profits ('000)
Treat	-1025.1*
	(555.5)
Observations	2592
Mean DV	7279

Table 11: **Effects on farmer revenues, volumes sold, and prices**

	Revenues Total ('000)	Revenues Maize ('000)	Qnt Sold Maize	Price Maize
Treat	99.2	72.0	61.3	18.0
	(90.9)	(68.5)	(118.2)	(14.0)
Observations	2775	2775	2769	1959
Mean DV	1019	672	1040	631
R2	0.32	0.31	0.33	0.02
Controls	Yes	Yes	Yes	Yes

Appendix A Tables

Table A.1: **Analysis of Variance in Market Prices**. Each coefficient in this table reports the R-squared from a different dummy variable fixed effects regression of prices in the panel market survey. The first column uses Trading Center FE (and so measures cross-sectional variation in prices), the second column month-of-year fixed effects (and so measures the degree of typical seasonality), the third column includes fixed effects for each round of the market survey (and so measures the extent of pure time-series market-level price variation), and the fourth column includes both TC and round fixed effects.

	Trading Center	Month of Year	Survey Round	TC and Round
Maize	0.04	0.18	0.84	0.87
Beans	0.26	0.15	0.34	0.55
Matooke	0.55	0.01	0.06	0.60
Tomato	0.30	0.04	0.09	0.39

Table A.2: **Trader Survey Attrition**.

	Treat	Control	Obs	T-C	
				diff	*p-val*
Baseline trader completes midline	0.92	0.94	1,457	-0.02	0.35
Tracked in original endline exercise	0.87	0.85	1,457	0.02	0.38
Found in Intensive Tracking	0.89	1.00	41	-0.11	0.08
Baseline trader completes endline	0.89	0.87	1,457	0.03	0.18

Notes: Analysis uses the full baseline sample of traders to study different definitions of attrition. Standard errors are clustered by subcounty.

Table A.3: **Household Survey Attrition**.

	Treat	Control	Obs	T-C	
				diff	*p-val*
Tracked in standard exercise	0.92	0.93	2,971	-0.01	0.44
Tracked in intensive tracking	0.74	0.85	39	-0.11	0.35
Successfully tracked	0.93	0.94	2,971	-0.01	0.31

Notes: Analysis uses the full baseline sample of farming households to study different definitions of attrition. Standard errors are clustered by subcounty.

Table A.4: **Market Survey Balance**

	Treat	Control	Obs	T-C	
				diff	*p-val*
Maize buying price	506.06	494.02	232	12.04	0.37
Maize selling price	625.92	628.50	232	-2.58	0.86
Number of maize traders	8.82	8.84	233	-0.03	0.98
Maize quality	1.58	1.61	232	-0.04	0.57
Beans buying price	1,566.81	1,573.55	214	-6.74	0.92
Beans selling price	1,884.94	1,994.98	214	-110.04	0.16
Number of beans traders	5.14	5.01	233	0.13	0.88
Beans quality	1.45	1.50	214	-0.06	0.51

Notes: Analysis uses the cross-sectional average of outcomes from the two pre-treatment market survey waves to examine balance of the market survey. Standard errors are clustered by subcounty.

Table A.5: **Market Survey Balance in Price Dispersion**

	Maize	Beans	Bananas	Tomatoes
One treated	0.0566	0.0349	0.0571	0.00314
	(0.0564)	(0.0704)	(0.0895)	(0.0817)
Both treated	0.0905	0.0104	0.103	0.0567
	(0.103)	(0.115)	(0.158)	(0.145)
Mean DV	4.592	5.978	8.265	3.478
N	26218	21129	20196	26149

Notes: Analysis uses dyadic averages in the two pre-treatment rounds of the market survey to examine balance in price dispersion. Standard errors are clustered by subcounty.

Table A.6: **Trader Survey Balance**

	Treat	Control	Obs	T-C *diff*	T-C *p-val*
Female	0.07	0.06	1,281	0.01	0.64
Age	37.16	37.39	1,281	-0.23	0.76
Education	7.68	7.32	1,281	0.36	0.24
Age of business	10.86	10.92	1,178	-0.07	0.92
# of subcounties in which bought	1.15	1.12	1,281	0.03	0.46
# of subcounties in which sold	1.27	1.31	1,281	-0.03	0.65
Net revenue, mz & bn	21,946,001.68	28,474,012.42	1,275	-6,528,010.74	0.54
Business costs per month	6,290,868.45	6,050,540.21	1,281	240,328.24	0.80
Annual Revenue	47,550,250.45	45,657,411.81	1,278	1,892,838.64	0.82
Annual Costs	43,068,736.38	40,790,579.76	1,281	2,278,156.62	0.72
Volume buy (kgs), mz	112,323.01	100,580.90	1,281	11,742.10	0.63
Volume buy (kgs), bn	6,174.67	4,936.33	1,281	1,238.34	0.49
Volume sold (kgs), mz	157,676.55	161,821.69	1,281	-4,145.14	0.94
Volume sold (kgs), bn	6,667.08	5,906.31	1,281	760.77	0.71
Trade maize	0.92	0.94	1,281	-0.02	0.42
Trade beans	0.28	0.25	1,281	0.03	0.54
Annual profits	5,617,367.84	5,717,231.92	1,274	-99,864.08	0.92

Notes: analysis conducted using the endline sample, with weights reflecting survey sampling and intensive tracking. Standard errors are clustered by subcounty, the unit of assignment. The first two columns give the means in the control and treatment group respectively. The third column gives the total number of observations across the two groups. The last two columns give differences in means and the corresponding p-value.

Table A.7: **Household Survey Balance**

	Treat	Control	Obs	T-C	
				diff	*p-val*
Number HH members	6.17	6.19	2,775	-0.03	0.86
Female	0.37	0.39	2,775	-0.01	0.66
Age	41.81	41.81	2,774	0.00	1.00
Highest grade completed	7.60	7.08	2,775	0.51	0.07
Food expenditure (month)	93,295.98	79,337.38	2,743	13,958.61	0.04
Land size (acre)	5.65	5.88	2,529	-0.23	0.60
Qtny sold, total (annual, kg)	1,133.87	1,056.52	2,775	77.35	0.70
Qtny harvest, total (annual, kg)	1,862.26	1,751.21	2,775	111.05	0.68
Number times sell	3.12	2.75	2,775	0.37	0.10
Percent of time sold at market	0.29	0.28	2,775	0.01	0.84
Sell in market	0.36	0.33	2,775	0.03	0.59
Distance to market	2.02	2.21	2,420	-0.20	0.63
Distance to Kampala	175.00	172.15	2,437	2.85	0.80
Assets (UGX)	2,508,859.59	2,297,790.91	2,775	211,068.68	0.62
Total exp (monthly, UGX)	219,099.68	191,827.79	2,775	27,271.89	0.09
Input exp (annual, UGX)	275,318.30	304,037.63	2,775	-28,719.33	0.46
Revenue, total (annual UGX)	637,169.99	555,801.61	2,775	81,368.38	0.43

Notes: analysis conducted using the endline sample, with weights reflecting survey sampling and intensive tracking. Standard errors are clustered by subcounty, the unit of assignment. The first two columns give the means in the control and treatment group respectively. The third column gives the total number of observations across the two groups. The last two columns give differences in means and the corresponding p-value

Table A.8: **Impact on Levels in Maize Markets**.

	Buy Price	Sell Price	Sell-Buy Margin	No. of Traders	Quality
Treatment	-1.781	-1.326	0.478	0.158	-0.0271
	(11.65)	(13.45)	(5.664)	(0.481)	(0.0483)
Mean DV	721.9	820.6	98.71	5.013	1.642
N	236	236	236	236	236

Notes: Analysis uses the cross-sectional average of all post-treatment waves of the market survey to study the impact of the intervention on market-level outcomes. Standard errors are clustered at the subcounty level.

Table A.9: **Impact on Out-of-sample Volumes Purchased and Sold**.

	(1)	(2)
	Out-of-sample buy	Out-of-sample sell
Treat	-4346	-12472
	(3150)	(7605)
Observations	2898	2898
Mean of DV	13397	45617
R squared	0.05	0.09
Controls	Yes	Yes

Notes: The outcome variable is volumes purchased (Column 1) and sold (Column 2) in out-of-sample markets by traders, regressed on trader treatment status. Standard errors are clustered at the subcounty level.

Table A.10: **Maize: Impact of the SMS Blast Roll-in.**

	Buy Price	Sell Price	Sell Buy Margin
Roll in Treatment	1.843	3.364	-0.127
	(7.196)	(9.234)	(0.446)
Mean DV	719.4	819.7	4.861
N	6928	6936	7445

Standard errors in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Notes: This analysis uses monadic market data in panel format with fixed effects for market and market survey wave, and analyzes the roll-in of the SMS Blast to control markets during the second half of the study. Standard errors are clustered at the subcounty level.

Table A.11: **Impacts of SMS Roll-in on Price Dispersion**

	All	Below 200km	Above 200km
Both treated by Rollin	0.0513	0.0975	0.0295
	(0.0825)	(0.118)	(0.0700)
Dist (10km)	0.0145***	0.0197*	0.00947**
	(0.00191)	(0.0101)	(0.00387)
Mean DV	4.641	4.358	4.767
N	220669	67899	152770

Notes: This analysis uses dyadic market data in panel format with fixed effects for market survey wave, and analyzes the roll-in of the SMS Blast to control markets during the second half of the study. The sample for this analysis consists only of dyads in which neither market was treated in the original experiment, and 'Both Treated' means that both markets in the dyad were included in the roll-in treatment sample. Standard errors are two-way clustered at the subcounty level.

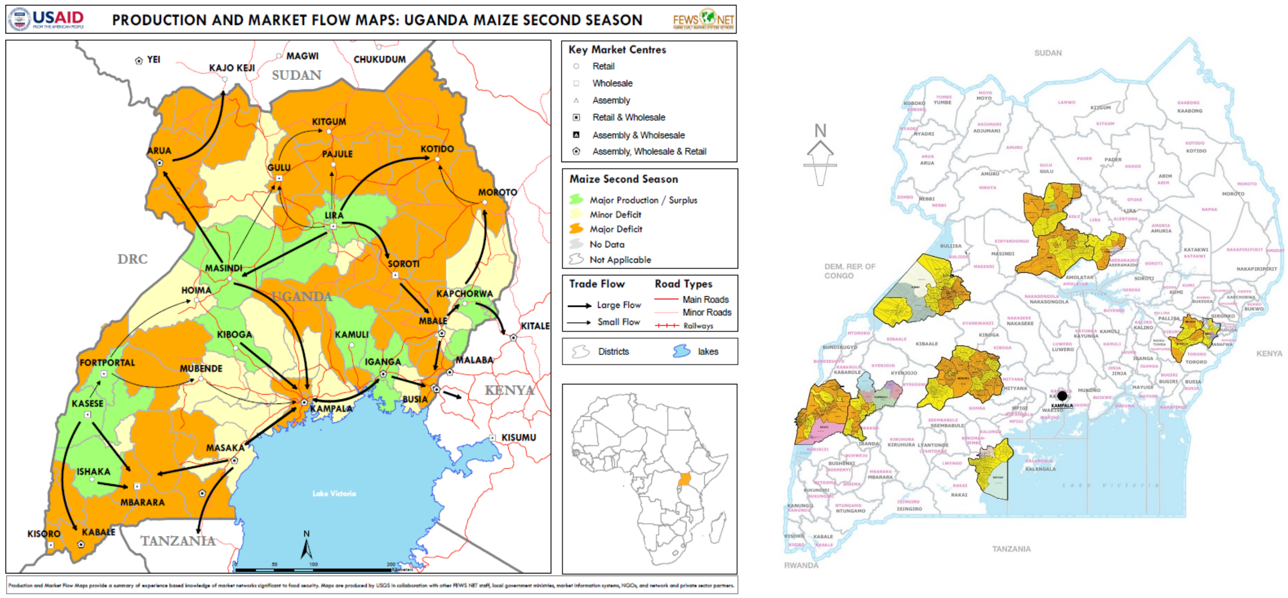
Table A.12: **Heterogeneity by Price Deviations from Normal**

	Contemporaneous	First Lag	Second Lag	Third Lag
Both Treated x Shock	0.0113	-0.0123	0.0141	0.00593
	(0.0416)	(0.0376)	(0.0239)	(0.0175)
Dyadic shock dispersion	0.447***	0.162***	0.0257	0.0187
	(0.0351)	(0.0278)	(0.0168)	(0.0167)
Both treated	-5.051	-3.627	-4.698	-4.178
	(3.317)	(3.260)	(3.160)	(3.227)
Dist (10km)	0.478***	0.709***	0.823***	0.835***
	(0.0796)	(0.0793)	(0.0788)	(0.0798)
Mean DV	104.4	104.8	105.0	105.2
N	1440955	1416167	1390239	1366089

Notes: This analysis uses dyadic market data in panel format with fixed effects for market survey wave, and interacts the treatment variables with a measure of how far away from regional monthly averages prices are in that wave. Standard errors are two-way clustered at the subcounty level.

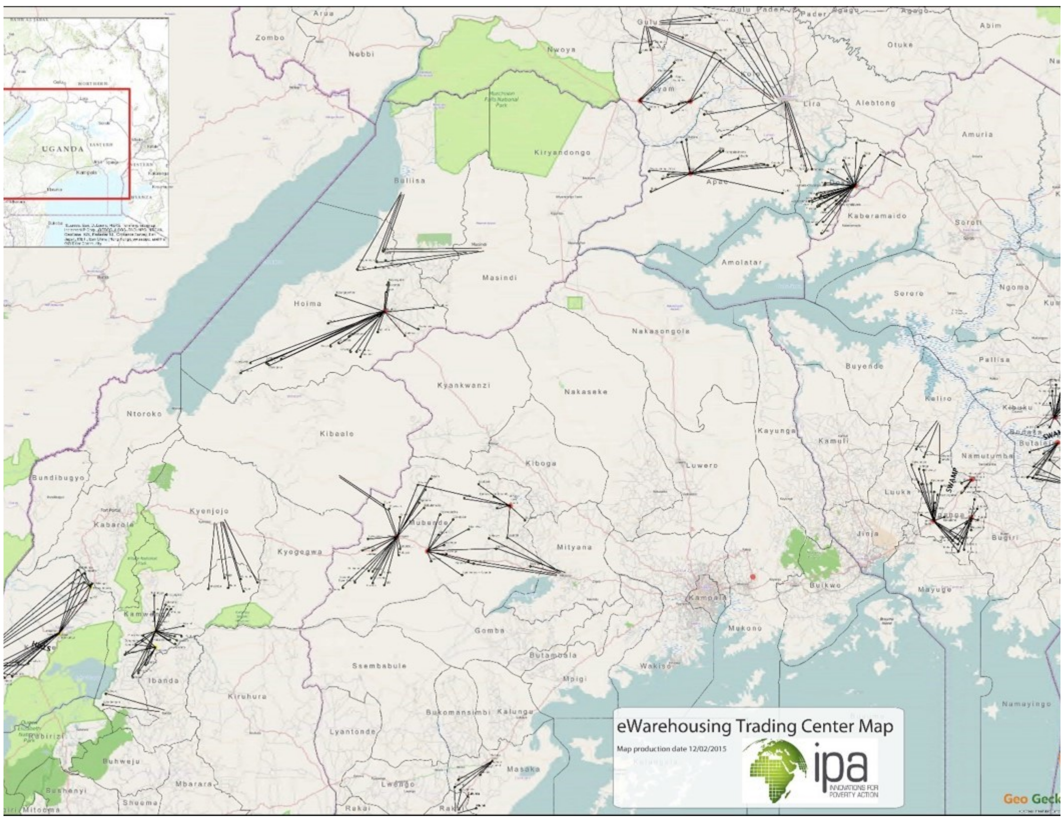
Appendix B Figures

Figure B.1: **Maps of the Study Area**



Notes: The left-hand panel is USAID's FEWS-Net map of Surplus Maize Areas of Uganda, and the right-hand panel shows the 11 study districts.

Figure B.2: **Map of Hub and Spoke structure of trading centers**



Notes: Figure depicts each study spoke TC connected by a straight line to its respective hub market.

Figure B.3: **Study Timeline**

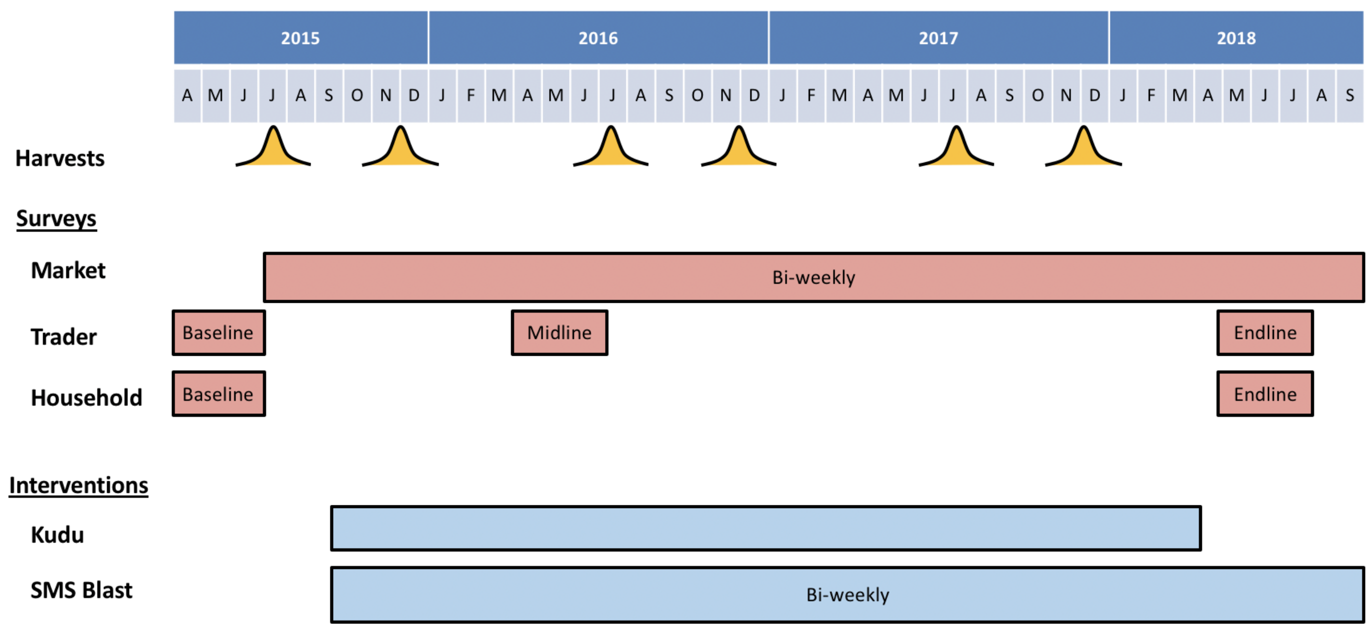


Figure B.4: **CONSORT Diagram of Study Recruitment and Attrition**

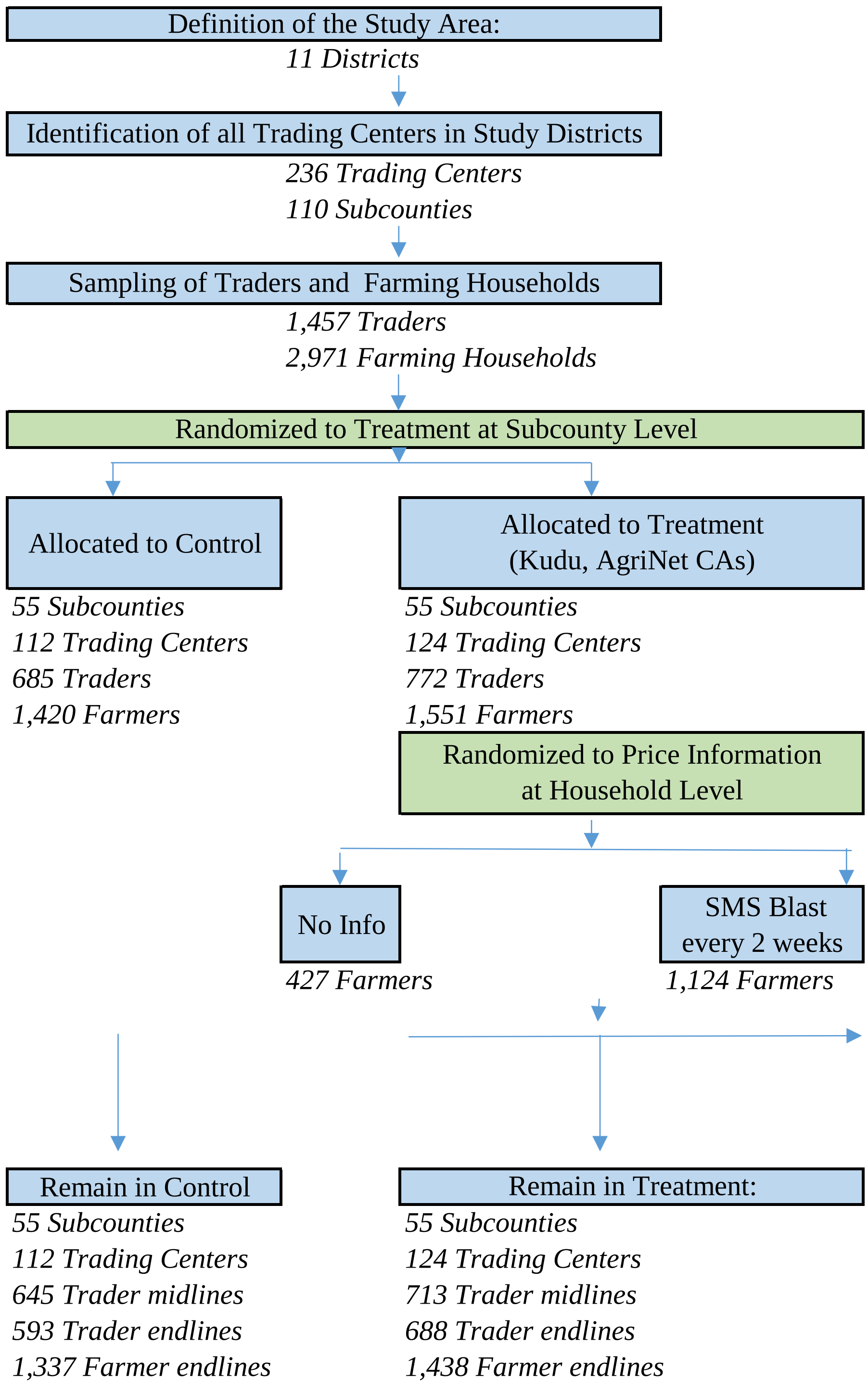
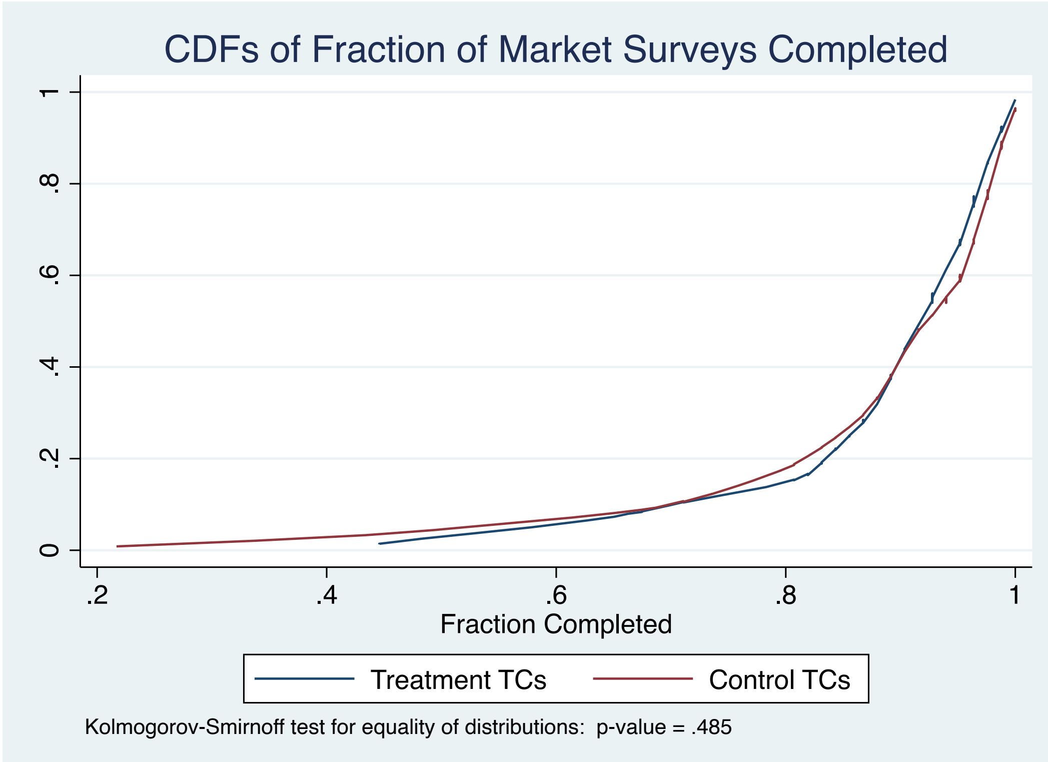


Figure B.5: **Attrition from the Market Survey**



Notes: The figure shows the Cumulative Density Functions (CDFs) of the fraction of intended surveys (83) that were completed for each TC, separating out the treatment and control TCs. The KS test fails to reject that the two distributions are the same.

Figure B.6: **Amount of Grain newly Posted to Kudu per Day**

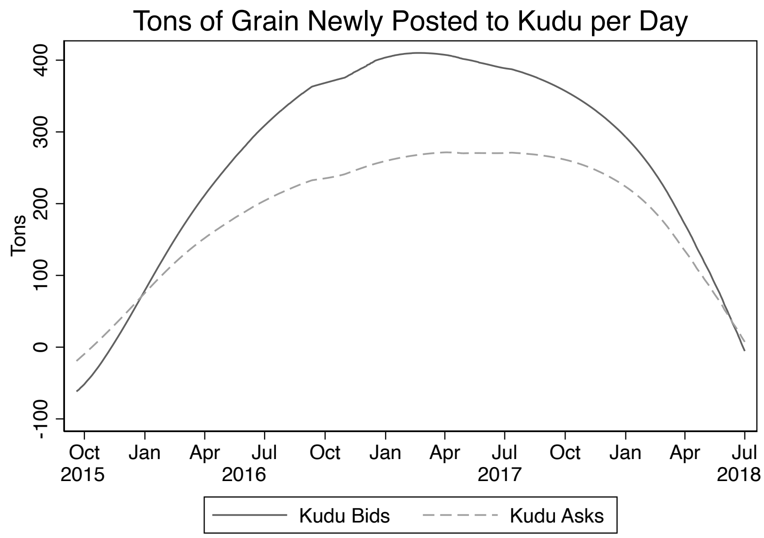


Figure B.7: **Volume of Bids and Asks**

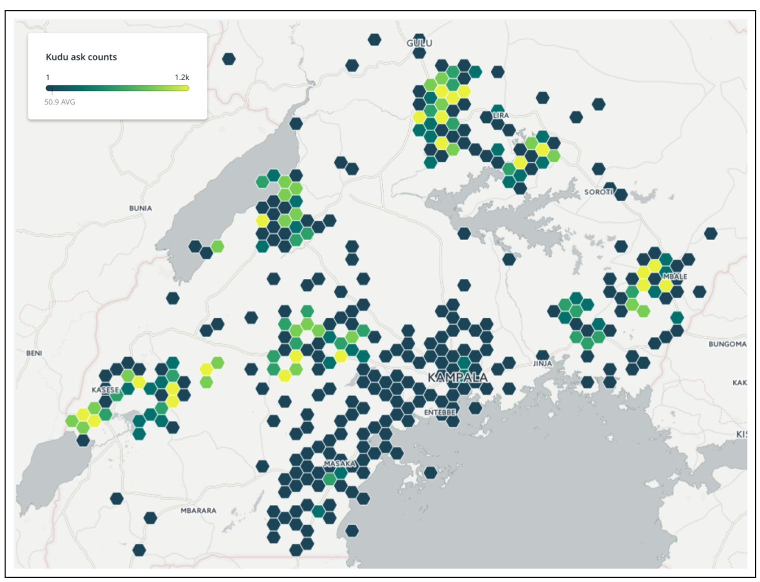


Figure B.8: **Maize Prices in Kudu vs. Market Survey**

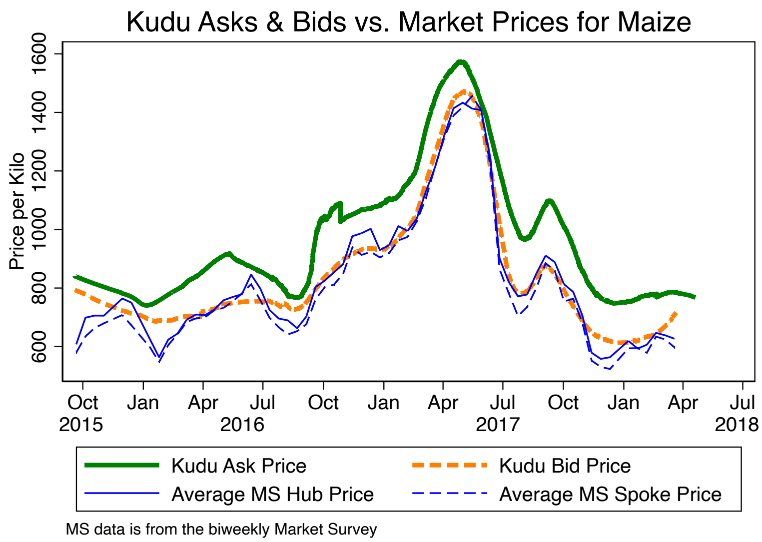


Figure B.9: **Distribution of Ask and Bid Prices, by Season**

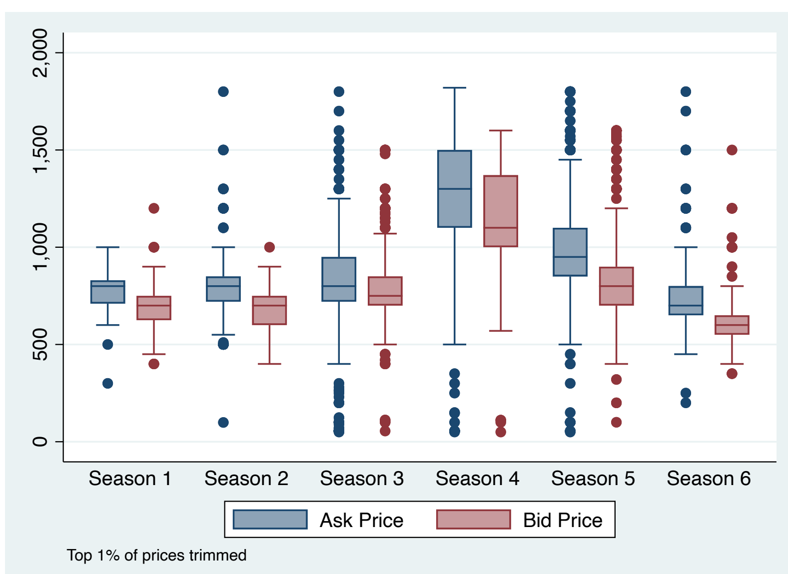


Figure B.10: **Cumulative Sales on Kudu**

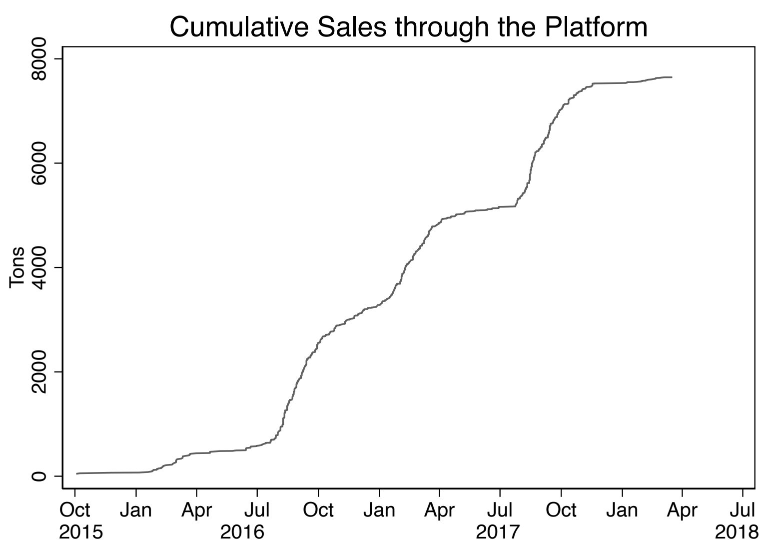


Figure B.11: **Randomization inference p-values**. Randomization inference p-values for the treatment effects shown in the right-hand side panel of Figure 2. P-values drawn from 1,000 placebo randomization draws. One sided p-values shown in black and two-sided in red.

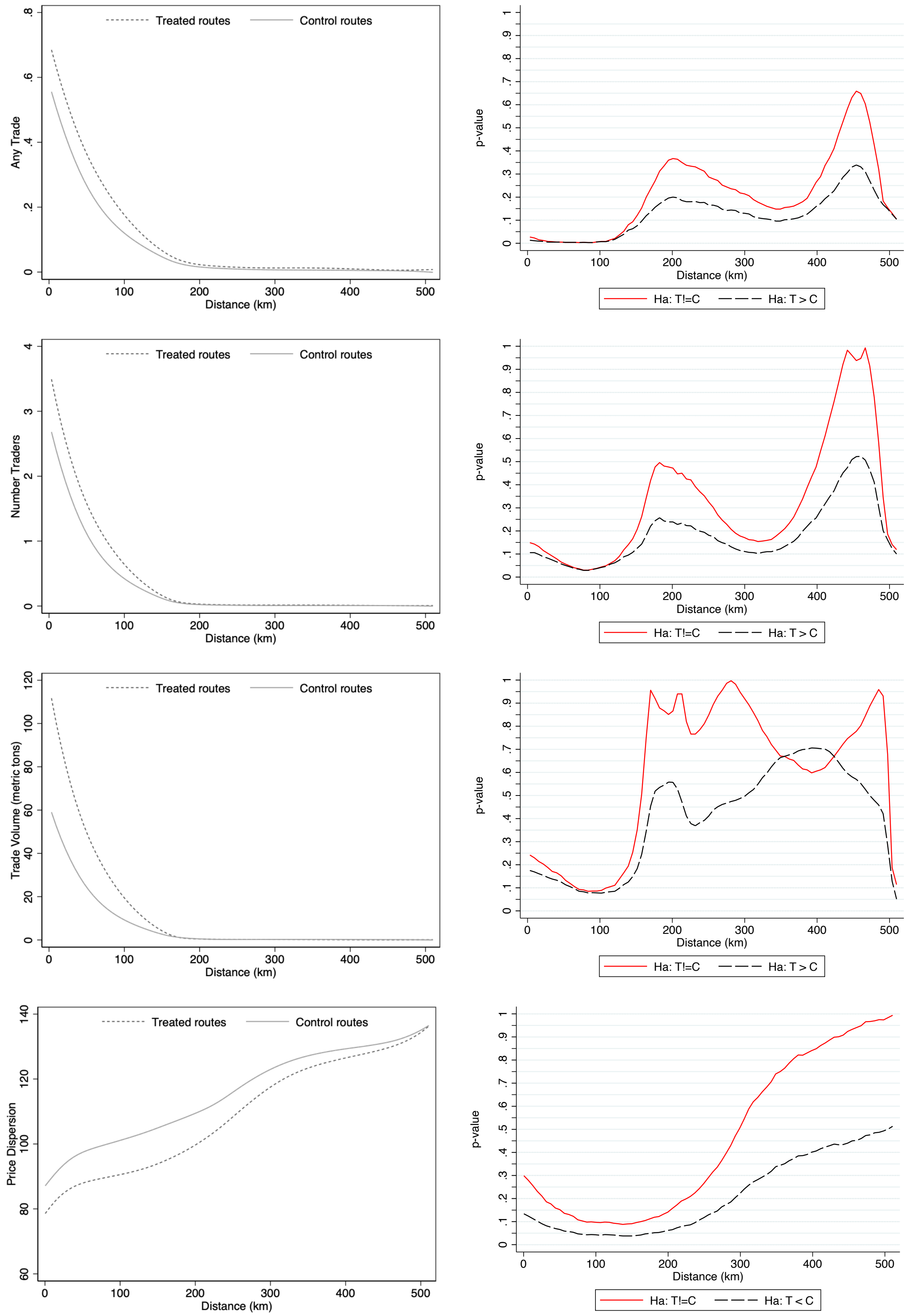
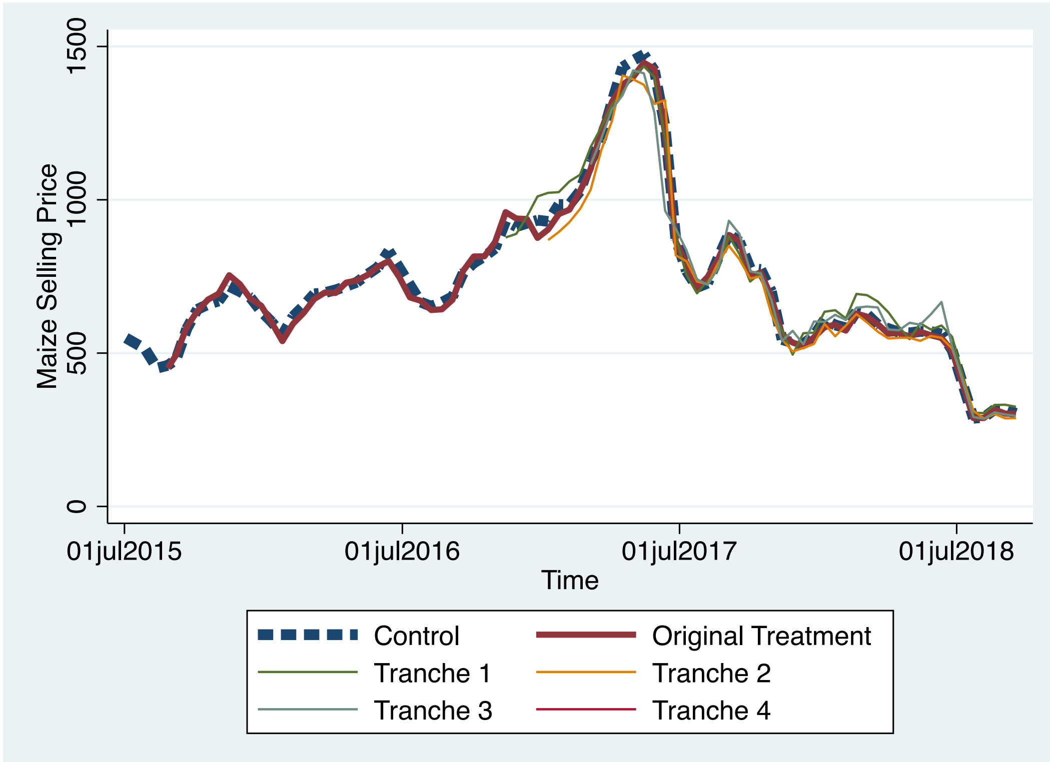
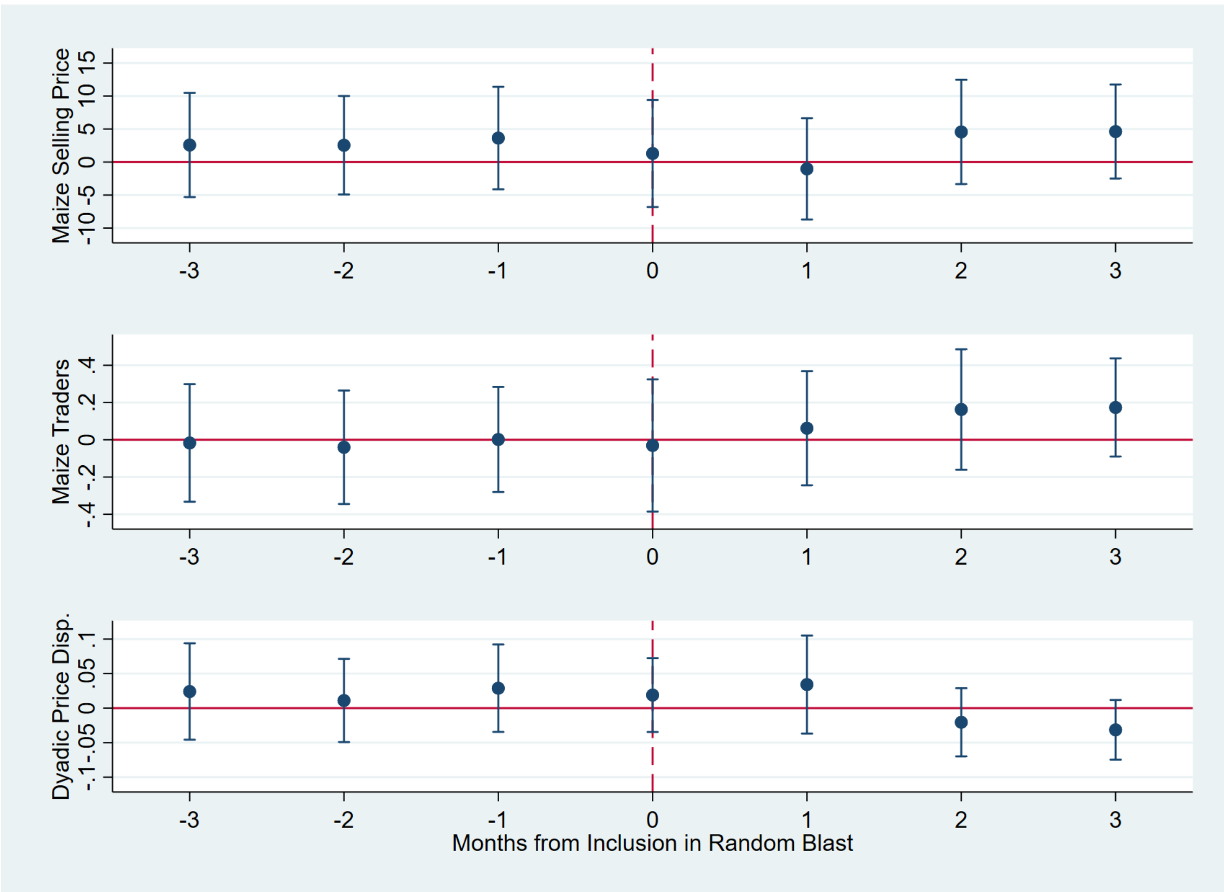


Figure B.12: **Impact of the Roll-in of Price Information**



Notes: This figure shows the average maize selling price at the TC/market survey level. All TCs not yet receiving the Mobile Price Information service at a moment in time are group together into the Control, and then each tranche is broken out as a separate average once it is rolled in to the information service.

Figure B.13: **Impact of the Random Blast**



Notes: This figure illustrates the monthly leads and lags of the inclusion of a given market in the 'Random Blast', where maize prices for that market were sent out to every market in the Kudu system. The first panel examines maize selling prices, and the second the number of maize traders in the random blast market at the monadic level. The third panel examines dyadic price dispersion between the Random Blast market and all other treatment markets that received information on that market's price.

Appendix B.1 Additional pre-committed analyses

For transparency, we describe here any additional analyses referred to in our pre-analysis plan, written in 2015, that we did not present above. First, we had intended to conduct an experiment to test credit constraints among traders by offering loans to a randomly selected subset of Commission Agents. We conducted a pilot for this experiment in the first season, issuing 62 short-term working capital loans to a group randomly selected from 124 CAs who expressed a desire for credit. In the end, the repayment rate on these loans was poor (78%) and our partner decided not to move this experiment to the intended scale, so we do not analyze it. Our PAP also specifies a set of hypotheses about convergence between spokes and hubs, and the differential effect of treatment for spokes in which the hub is and is not treated. In the end we were only able to map 84% of our spokes to hubs, and the analysis conducted within this reduced sample is typically inconclusive, suggesting that the trading networks may be more complex than our simple hub-and-spoke mapping supposed.