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

February 2024

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 25% 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 and James Allen, Russell Morton, and Jared Stolove for excellent research assistance. 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. Barriers that limit arbitrage between agricultural surplus and deficit areas can therefore have substantial welfare costs (Barrett, 2008; Rashid and Minot, 2010). High trade costs thwart otherwise profitable exchanges, leading to lower producer prices, higher consumer prices, and large geographic and intertemporal price dispersion.

Trade costs in sub-Saharan Africa are some of the highest in the world, and while part of this is due to very high transportation costs (Teravaninthorn and Raballand, 2009), there is also growing evidence that price gaps are much larger than can be explained by transportation alone (Atkin and 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 – for instance, efforts to disseminate price information to farmers have had mixed results, and overall, 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 information 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 allow us to isolate the causal effect of reducing buyer-seller search frictions. The at-scale randomization enables us to generate and measure impacts at the market level. We interpret this experimental variation through the lens of a model in order to correctly estimate treatment effects in the presence of trade links and GE forces that create spillovers between treatment and control markets.

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 fixed costs of search decreases scale economies, most Ugandan farmers are still too small to engage 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 tracked outcomes over three years, equivalent to six harvest seasons. We gathered data on market-level prices in 236 markets every two weeks. We also collected 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 surveyed farming households to study the impacts of the platform on farmer production, revenues, and welfare.

We find that access to Kudu increases trade between treated markets on both 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 markets 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, even though most do not use the platform. A separate treatment arm that provided price information rather than connections to new transaction partners had no impact, confirming that the channel for these impacts is likely buyer-seller matching, rather than price search.

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 a large geographic level lets us measure effects on equilibrium outcomes such as prices and aggregate trade flows. Randomization also circumvents causal inference issues that are common challenges when studying trade costs, related to the endogenous location and quality of infrastructure (such as roads, rail, telegraph and phone coverage, etc). 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 costs of trade. These can include the search costs that buyers and sellers incur to connect with one another across markets, such as time spent calling contacts, or the time and travel costs of going to another market in person to look for buyers. Kudu lowers these search costs by making it easier to connect with new potential trade partners. We isolate Kudu's direct effect on trade costs using experimental variation in bilateral treatment status between markets, holding market-level factors constant. We estimate that Kudu drove a 25% reduction in fixed trade costs between treated markets, a substantial reduction.

In order to capture the total impact of Kudu, we estimate additional parameters of the trade model, making use of randomized variation in exposure to treatment in other markets. Combining our main estimate of the direct treatment effect on fixed trade costs with these other estimated parameters, we can measure the total effect of the platform. 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. We find that Kudu 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.

This approach also allows us to measure impacts on market-level outcomes such as prices. We find that there are distributional effects of the platform, with farmers gaining and consumers losing in surplus markets, where prices rise, and the opposite in deficit markets, where prices fall. 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 RCT that operated in the field for three years. The study took place 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 in our study area (see Table A.1).[1] Figure 1 presents the absolute value of the gap in prices between each pair of markets in our sample (black solid line), as collected by our market surveys (see Section 2.4 for detail on data collection). The average price gap for maize is about 110 Ugandan shillings (UGX) per kg, or roughly 15% of average price. Unsurprisingly, price gaps rise with the distance between the markets.

[1] Maize is the most commonly grown and consumed crop. 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 matooke (green bananas steamed and eaten as a staple starch crop).

A major driver of these price gaps is transportation costs, which are among the highest in the world (Teravaninthorn and Raballand, 2009) in this region. 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 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 operate on a much larger scale and do much more active marketing, despite the fact that our frame of traders are based in smaller markets and hence we exclude the major national traders who are based in the country's largest cities. The median study 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 subcounty, aggregating, and selling downstream to fewer, larger markets. The median 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 one new and one exiting trader over the three year study period. At baseline, traders bought maize at an average of 12.7 cents/kg and sold at 16.4 cents/kg, 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

[2] Traders reported round trip 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 non-parametric function of the km traveled.

[3] Consistent with these figures, Bergquist and Dinerstein (2020) estimate that the median trader in their sample

$65.

Mobile phones are an important technology used in traders' search and matching, with traders reporting calling a seller, buyer, or other contact ahead of time to get information about price or availability in 54% of their transactions. More traditional search-by-visiting is also common, with the remainder of transactions occuring 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. These services typically only offer simple price alerts; none to our knowledge provides direct connections to buyers or sellers.

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 designed 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]

in Kenya retains 12% of total revenues in profits.

[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. Where barriers to information are larger (such as cross-border business by small-scale traders, there is more evidence that information-only interventions can be effective

Kudu, in contrast to existing price information platforms, offers matches between individual market participants. The system, which was developed at Kampala's Makerere University 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.[5] 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.[6] Users are contacted by SMS to inform them that the match has occurred and provide them with the contact 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. If interested, buyers and sellers meet in person to exchange the goods and make payment; Kudu has no mechanism for processing payments.[7]

To promote take-up of Kudu, salaried staff were sent to treatment markets, to serve as the on-the-ground agents of the project, promoting the mobile marketplace to local farmers and traders. [8] In addition, we promoted Kudu via text message every two weeks to treated traders and farmers

(Wiseman, 2023).

[5]Kudu instructs users to post their reservation prices. The (non-binding) price recommended by the platform in the event of a match is the seller's price, so this is incentive compatible for the buyer. Perhaps because of this, seller ask prices were often higher than buyer bid and market prices. Qualitative interviews suggested that sellers often posted strategic offers, much in the way they would in more traditional, in-person negotiations. Buyers' bid prices, on the other hand, track market prices very well (see Figure B.7 and Figure B.8). 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. Kudu also did not enforce these price recommendations, as it does not have a mechanism for processing payments; rather, payment is collected in cash when trading partners meet in-person to exchange the goods. This means that in practice prices could be renegotiated flexibly.

[6]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).

[7]Though as part of this project, research staff assisted Kudu in setting up a robust system of phone surveys for tracking completed transactions. This was useful not just for data collection, but also for updating the pool of still available asks and bids.

[8]We initially partnered with a private sector agribusiness firm to employ and train 210 Commission Agents (CAs) to promote Kudu on the ground. However, CAs were not reliable promoters, perhaps because they were recruited

in our sample.[9] 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.[10]

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; 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 an "at-scale" randomization, conducted with the goal of being able to observe treatment-control differences in outcomes such as prices and trade flows that are determined in equilibrium at the market level.

We first implemented a census of all markets within our 110 study subcounties. We listed all markets that were permanent (i.e. 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 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

from among local traders and therefore lacked the incentive to promote the platform to others, although they were regular users of Kudu for their own trading operations. Ultimately, Kudu hired salaried staff, not drawn from the local trader population, which was much more successful in generating broad awareness of Kudu.

[9] This promotional text message was send to all treatment traders, as well as a randomly selected three-quarters of treatment farmer households. This farmer-level 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 all farmers are mostly a result of shifting GE forces. See Section 7 for further discussion.

[10] This message also included price information for maize, beans, matooke, and tomatoes in users' local market, downstream hub market, and closest major market, as well as five randomly sampled treatment markets each two-week round. However, as we describe in Section 3.5, we also conducted a separate experiment in which we sent only price information, and see no impact. We therefore see the price information component of the intervention as non-pivotal.

(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.

We collect three core types of data, illustrated in Figure B.4, which provides a CONSORT diagram of study recruitment and attrition for each type of data. . The first is a high-frequency market survey, which gathered information in each of the 236 markets 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).

The second dataset 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, 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.[11]

Finally, to understand the impact of the platform on farmers, we surveyed a sample of agricultural households. We first listed all villages located in study subcounties. For each market in our study subcounties, we then selected the village containing the market (which is typically more urban) and randomly sampled one of the remaining villages within the same parish (which tends 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 sampled 8-9 farming households in each market village and another 4 in each rural village. 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

[11] We also re-conduct the census of traders present in the markets at endline to study entry into the trading industry. We find no treatment effects on exit or entry into the trading business (see Table 5.

practice, these eligibility criteria excluded few households, with over 90% of households qualifying for study inclusion. Study households completed a baseline in 2015 and an endline survey in 2018 covering agricultural production activities and agricultural marketing and sales. Farmer analysis is weighted to make it representative of all farming households in the sampled study villages.

2.5 Attrition and Balance

We now present attrition and balance analysis for each of the three types of data captured in the study: the market surveys, trader surveys, and household surveys. For the market survey, we have 88% of the attempted (market by survey round) observations. For the trader midline, we were able to survey 1,358 of the 1,457 baseline traders (93.2%). The trader endline originally located 1,248 traders (85.7%), after which we randomly sampled 20% of attritors (41 individuals) for intensive tracking. We successfully located 37 of these (92.7%), bringing the weighted tracking rate for the trader endline to 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 core variables measured (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 gaps across 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 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

[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, and documents that attrition is balanced across treatment. Among all the tests that we conduct, only the intensive tracking rate in the trader survey appears differential, with more control traders found than treatment, but given that this arises from finding 14/14 control traders versus than 24/27 treatment traders who were intensively 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.

10% level and one at the 5% level, in line with what we would expect by random chance.

2.6 Platform Usage

Over the three years of this study, Kudu 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. Bids (offers to buy) and asks (offers to sell) posted on the platform each day climbed through the first year to reach a steady maximum of about 400 tons of per day in bids and 200 tons in asks.[13] Figure B.6 shows the spatial distribution of asks, indicating study market centers across the country posting upwards of 1,000 asks each. Among the individuals posting asks to sell on Kudu, 45% were sampled study traders, 14% were agents of an agribusiness firm contracted to promote Kudu, and 6% were sampled study farmers. For those posting bids to buy, the corresponding percentages are 48% and 11% for sampled traders and agents and less than 1% of sellers are sampled farmers. The remainder of bids and ask came mostly from within the treated study areas, but from individual users who were not sampled in our data collection. Overall, Kudu saw 7,300 tons of grain successfully transacted during the study period, worth a total value of about $2.3 million USD. Figure B.9 shows the cumulative sales over the platform during the duration of the study.

Within our study sample, 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. Appendix Table A.8 presents predictors of take-up for both samples. We see that scale, in terms of quantity sold or traded, is an important predictor of adoption, a point to which we will return later.[14]

[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] Interestingly, the only other significant predictor of adoption is gender. We find female farmers are less likely to use Kudu. We see weaker, although similarly signed effects among traders, though there are very few female traders overall.

3 Reduced Form Impacts

3.1 Effects on Trade Flows and Market Prices

We now turn to the reduced form effects of the platform on trade flows and market prices. We use the phrase "reduced form effects" to refer to the observed post-treatment differences between treatment and control markets (or 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, this reduced form variation will be key to estimating our model and ultimately quantifying full treatment effects.

Figures 2-3 present these reduced form effects on several outcomes at the "route" or market-pair level: whether any trade is occurring between the two subcounties, the number of traders engaged in trade between subcounties, the volume of trade flowing between subcounties, and price gaps between markets. The first three of these outcomes are drawn from our panel survey of traders. The last is drawn from our market-level price surveys. The lefthand side of Figures 2-3 presents non-parametric local Fan regressions of each outcome on the distance between the subcounties, separately for treatment and control routes. Treatment routes are those between 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 are 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 top left panel of Figure 2, 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 is close to zero beyond 200km distance. Consistent with this, the number of traders (second row) also falls quickly with distance. The salience of this extensive margin of trade is consistent with the presence of fixed costs, which will feature prominently in our model in Section 4. Total trade volumes (third row) also decline in distance, while price gaps are larger between markets located at further distances, as shown in Figure 3.

When we look at these outcomes along Kudu-treated routes, we observe a higher probability of any trade (first row of Figure 2) and a larger number of traders engaged in trade (second row), suggesting greater trade on the extensive margin. We also see larger volumes of trade flows (third row). These increases in trade flows come with reduced price gaps, as shown in Figure 3. Notably, these 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 middle and righthand panels of Figures 2-3 present results from a randomization inference procedure, in which 1,000 placebo randomized treatment assignments are drawn. In the middle panel, the "reduced form effect" – i.e., the difference between treatment and control line for the realized randomization in the lefthand figure – is shown in black. The grey lines show the same "effect" for each of the 1,000 placebo treatment assignments. Long and short-dashed lines indicate the 90th and 95th confidence intervals of these placebo "effects." The righthand panel shows the resulting p-values from this randomization procedure, with the p-values from a two-sided test in red and a one-sided test in black.[15] Consistent with the lefthand panel, which suggests treatment effects are concentrated among relatively nearby markets, we see statistically significant effects for routes at short distances, with randomization inference p-values for both the one and two-sided tests near or below 0.1 (for some outcomes and distances, well below). Beyond 200km, we see in the middle panel that effects taper off to precise zeros, with the corresponding randomization inference p-values steeply increasing in the righthand panel.

We present the same results in regression form in Tables 1-2, running the following specification:

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

[15] For the one-sided test, the null hypothesis being rejected is that the trade flow outcomes (any trade, number of traders, and trade volumes) tested in Figure 2 are greater in control than in treatment and that price gaps tested in Figure 3 are greater in treatment than in control.

where Y_{dr} is the outcome of interest along route (dyad) d in round r, pooling all post-treatment survey rounds in the same analysis. Outcomes are regressed on T_d, a dummy for whether the route is treated and D_d, a measure of the shortest road distance between the pair of markets. Standard errors are clustered two-way by each subcounty (the unit of randomization). We see in Table 1, which presents results for the full sample of routes, point estimates that are consistent with the effects observed in Figures 2-3, but effects are only marginally or not significant. However, Table 2 documents that effects are concentrated along short distance routes, with significantly higher probability of trade occurring, a larger number of traders operating between subcounties, higher trade volumes, and lower price gaps. In contrast, we see almost no effect of the platform on direct trading outcomes for further distance routes.

The platform therefore generates meaningful differences in trade flows and price gaps among nearby markets. However, it does not live up to the often-touted promise of such digital marketplaces to directly connect remotely-located producers and markets with distant urban consumers.

3.2 Unpacking the Price Convergence Impacts

What is driving the reduction in price gaps between treated market pairs? Figure 4 presents reduced form effects on price levels in relative surplus versus relative deficit areas, where surplus and deficit are defined based on average marketed surplus per farmer at baseline. First, we note in the lefthand panel that, as would be 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, with relatively lower prices in treated deficit markets and higher in treated surplus markets. The righthand panel presents this reduced 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, where we again see that prices are weakly lower in relative deficit areas and higher in relative surplus areas. With an average baseline marketed surplus of about one ton, these effects almost exactly offset each other for the median market.

3.3 Trader Impacts

We saw in Section 3.1 that there are a greater number of traders active along treated routes. Who are these new entrants? Table 4 documents that traders operating along treated routes are smaller on average than those operating on control routes, as measured by baseline profits, volume traded, and value traded (Columns, 1, 3, and 5).[16] Columns 2, 4, and 6 suggest that, in particular, treatment allows a drop in the minimum size trader along a route. This is consistent with Kudu lowering the fixed cost of serving treated routes and inducing in smaller entrants, a point to which we return in Section 4.

In Table 5, we turn to Kudu's effect on trader business outcomes, regressing firm-level outcomes on the trader's home market treatment status. We find that while treated traders' total quantity traded goes up (Column 1), the buy-sell margin at which they trade goes down (Column 2). While these results are imprecisely measured, they are consistent with Kudu enabling traders to find new markets and expand their trading volumes, but – as other traders are also exposed to Kudu – the aggregate increases in trade flows that drive higher prices in surplus areas and lower prices in deficit areas cut into traders' buy-sell margins.

These changes however, are not sufficient to alter exit or entry into the trading industry as a whole. We see in Column 3 no effect on the probability that treated traders are still in business in subsequent surveys rounds. We also re-conduct our trader census at endline, to order to measure whether treatment affects entry by new traders into the business. We see in Column 4 no effects on the number of new traders in business in treated markets.

[16]It is critical that these measures of size are from our baseline, as we aim to shed light into how treatment alters the *selection* of traders into treated trading routes.

3.4 Farmer Impacts

We next turn to looking at outcomes for farming households. We see in Table 6 no statistically significant effect on total revenues, maize revenues, maize volumes sold, or price received for maize sold, for farmers living in treated subcounties, 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.[17] Nonetheless, these results suggest that the intervention did not drive overall prices to farmers, and consequently had modest effects on revenue and quantity sold.

Recall, however, that Kudu induced no average change in market prices; rather, it increased them in surplus areas and reduced them in deficit areas. Figure 5 therefore displays treatment effects on farmer revenue by relative surplus status. In the lefthand panel we see that revenue increases with surplus status in both treatment and control. However, the righthand panel shows that treatment significantly increases farmer revenues in surplus areas (and decreases revenues in deficit, although this effect is not statistically significant). This suggests that shifts in market prices can impact farming households, even though we know that most did not directly use Kudu or engage in cross-market trade.

We also test whether Kudu affected production, either on average or differentially by surplus status. One might imagine that as farmers receive higher prices in surplus areas, this could drive increases in production, for example, through increased input usage. The converse might be true in deficit areas. However, as shown in Appendix Table A.9, we find no significant impact of Kudu on harvest levels (although point estimates go in the expected direction – positive in surplus areas and negative in deficit – they are far from significant). This suggests that the increase in trade we observe arises mainly from changes in the allocation of existing production, as maize is moved away from places of low marginal utility and into areas of high marginal utility, rather than through increases in productive specialization. It is of course possible that such production changes require more time to adjust, beyond the three years encompassed by our study. We therefore view the gains from Kudu observed here as a lower bound for those that could emerge in the long-run.

[17] Note that the farmer sales price data is available only for those who report sales (either at farmgate or at market) but we do not find a treatment effect of Kudu on reporting a price or on the location of sale).

3.5 Unpacking matching frictions

The unique feature of our platform is the direct creation of new market linkages. We also, however, provided randomized price information in a number of different ways in the three years of the study, and so have the ability to speak to the literature that examines price-only impacts. Using these sub-experiments we present additional evidence that price information alone has limited impact on market integration in our setting (perhaps because cell phones are already fairly ubiquitous) and further positive evidence in favor of the role of buyer-seller matching frictions. These findings influence the way we model the treatment effect of Kudu in the following section of the paper.

We have two experiments designed to test the role of price information alone. First, we randomly rolled out (only) the SMS-based price information alerts to a random subset of control markets.[18] We rolled in three control markets in each of the 12 market survey rounds between October 2016-March 2017 (roughly the second half of the study period), and then, subsequent to the household and trader endline surveys, we rolled in an additional 36 control trading markets and observe a final four 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 the effect of price information alone. The most well-powered way of analyzing this price information-only experiment is to use the high-frequency market survey, as we can then match the biweekly market pair price gaps to the timing of the introduction of price information to the market. In Table A.10, we present the impact of this sub-experiment, comparing price gaps among markets rolled into the price-only intervention to pure controls, before and after the intervention using panel data with round fixed effects. Unlike the introduction of Kudu, which drove strong price convergence, here we find no evidence of a reduction in gaps in prices across markets that receive only price information (even for those at short distances).

A second sub-experiment shedding light on the role of price information alone comes from within-treatment variation. Recall that as part of the price information sent to treated traders and farmers in the main experiment, we randomly selected five treated markets in each round,

[18] Recall these price alerts included information on 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. See Section 2 for more detail.

and broadcasted information about prices in those markets to the entire treated network. Though more short-run, this intervention was quite powerful, as it sent the market's price information to thousands of traders and farmers simultaneously. However, this price-only intervention also had no effect (see Figure B.10).

These two sub-experiments suggest that price information did not drive the main treatment effects observed in the study. This is further corroborated by the fact that Kudu increased the flow of trade from net surplus to net deficit regions of the country, a form of heterogeneity that is relatively time-invariant and well-understood by professional maize traders, the main users of Kudu. Conversely, we find no evidence that Kudu reduced high-frequency deviations from expected average prices, as one would expect if the type of search cost alleviated were price information (see Table A.11).

Instead, our evidence suggest that Kudu worked by reducing buyer-seller matching frictions, introducing users to specific other users who were available and willing to trade, and thereby reducing the cost of searching for potential transaction partners.[19] Therefore, rather than supplanting traditional relationship-based trading networks, Kudu appears to support the formation of new trading relationships, by reducing matching frictions and introducing new trading partners. We now present a model with this logic at its core.

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 generated experimental variation in aggregate outcomes between treatment and control markets, as well as 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

[19] In fact, our survey data suggests that many of the new trading relationships formed by Kudu were quite durable, outlasting the initial deal formed on Kudu. 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 accounting for 7x the volume of the deal initially conducted on-Kudu. This is consistent with the fact that the total volume of trade generated by the intervention (based on trader surveys of trading patterns) is 4x the volume of trade between treated markets that is observed on Kudu.

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 experimentation with trade costs, which by nature involves 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.

In this section, we describe the model that will guide our estimation. The model serves two purposes. First, it will inform how we separate direct effects on treated units from indirect equilibrium effects. This will enable us to use the experimental variation to correctly estimate the direct treatment effects of access to Kudu. 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.[20]

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 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.

[20]We see no treatment effects on harvest levels, either on average or by surplus status (see Appendix Table A.9). For simplicity and to match this empirical reality, we therefore model the environment as an endowment economy.

4.1.3 Traders

Each location is home to a continuum of traders, with measure N_i.[21] 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

[21] We take N_i as exogenous, as we see no treatment effects on entry or exit of traders into the trading industry. See Table 5 for details.

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 $\sigma = \frac{\mu-1}{\mu}$.[22]

The profit maximizing price charged by a traders serving a route ij features a constant markup over marginal cost:

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

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. 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.

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:

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

[22]The price index in market i is defined as $P_i^{1-\sigma} = \sum_j N_j \int_{a_L}^{a_{ji}^*} p_{ji}(a)^{1-\sigma} dG(a)$.

4.2.2 Bilateral trade flows

We can characterize trade flows in terms of three margins: any trade on a route, the number of traders serving the route, and the value of trade on the route.

There will be some trade on a given ij route if the most productive 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 minimum of the trader operating cost distribution.

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

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

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:

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

where r_{ij} is the revenue or value of trade per trader.

Again, 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 make more specific assumptions about the form of trade costs and then derive estimating equations that correspond to the 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.

We specify variable trade costs as $\tau_{ij}^{\sigma-1} \equiv d_{ij}^{\gamma} e^{-u_{ij}}$, where d_{ij} is the distance between markets i and j. Fixed trade costs are: $F_{ij} \equiv \exp\left(\phi + \theta \mathbf{ln} d_{ij} + \beta_1 \mathbf{1}_{ij}^{K} + \beta_2 \mathbf{ln} d_{ij} \mathbf{1}_{ij}^{K} - v_{ij}\right)$ where $u_{ij} \sim N\left(0, \sigma_u^2\right)$ and $v_{ij} \sim N\left(0, \sigma_v^2\right)$, and $\mathbf{1}_{ij}^{K}$ is an indicator variable equal to one if both i and j are treated.

We are primarily interested in estimates of β_1 and β_2, which describe the impact of Kudu 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.

4.3.2 Estimating equations

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

Any trade

Taking logs of Equation 2 and using the parameterization of τ_{ij} and F_{ij}, we get:

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

where ζ_j is a destination market fixed effect, 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:

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

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:

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

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:

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

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

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

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, taking logs, and making use of the parameterization of τ_{ij}, 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 $B = \frac{\mathrm{corr}(u_{ij},\eta_{ij})}{\left(\frac{\sigma_u}{\sigma_\eta}\right)}$ is a coefficient to be estimated.

5 Model Results

5.1 Estimating Equations

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

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

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 higher (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.

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

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.

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

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 total trade flows 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.[23] 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

[23] 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.

that has parallels to the market access literature (Donaldson and Hornbeck, 2016). Formally, let:

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

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 7, 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.[24] 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

[24]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 7. 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.

subcounties reduces own price, as demand is diverted to treated subcounties and away from one's own subcounty. Conversely, (random) exposure to treated deficit subcounties increases 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 7 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 implied parameters estimated jointly via a generalized method of moments (GMM) estimator. Our estimator uses four sets of moment conditions, corresponding to the three estimating Equations 8-10 and a first stage equation for price (corresponding to Column 3 from Table 7). First, we formulate Equation 8 as a maximum likelihood problem and form moments from the first-order conditions of this problem. Second, we formulate Equations 9-10 as nonlinear least-squares problems and form moments from the first order conditions of these problems. Third, we formulate the first stage for price as an OLS problem and form moments from the first order conditions from that problem.[25][26]

Table 8 presents results. We present specifications with and without the distance interaction,

[25] Parameters that appear in multiple estimating equations have more than one corresponding moment condition. Thus, we are over-identified.

[26] 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 ((β_1, β_2, σ, γ, θ, κ, σ_η^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.

though we will see it does ultimately not make much of a difference for most of our parameter estimates. We see sensible estimates for our key parameters. The estimate of β_1 suggests that the direct impact of Kudu is to reduce the size of fixed costs by 25% 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.[27] We can now run counterfactuals, turning on and off the 25% 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.

Results are presented in Figure 6. 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

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

routes represents a real loss in trade volumes for those routes, which a naive reduced form estimate would fail to capture.

Figure 7 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.

As with trade flows, we can use our model to calculate total impacts on prices that incorporate the spillover effects on control markets and hence provide a more complete picture of general equilibrium effects. Figure 8 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. These model-derived price changes are qualitatively similar to those coming from the reduced-form experimental heterogeneity, but are not identical because the nature of the spillover from treatment to control is itself different across surplus and deficit, leading to differential contamination. They nonetheless support the overall story of price convergence across surplus and deficit areas, as well as increasing farmer revenues in surplus areas.

6.3 Welfare impacts

The results of the prior section make it clear that 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. Figure 10 presents these welfare results.

7 Policy Implications

7.1 Implications of scale economies for take-up

Platforms like Kudu are typically motivated by policymakers as a way to 'cut out the middleman' and improve farmer welfare by connecting the rural poor directly to high-priced urban markets. Our results suggest that a rethink of this logic is order. We emphasized in Section 4 that scale economies can explain why reducing matching frictions induces an extensive margin impact on trade flows. Scale economies can also explain why the take-up of Kudu is concentrated among traders, with limited direct use by smallholder farmers. Figure 9 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 25% 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 9 implies that these new entrants induced into trade by Kudu should be smaller than incumbent traders.

This is indeed what we observe as shown in Table 4.

The second notable feature of Figure 9 is that the Kudu-induced reduction in the threshold size still lies far to the right of the farmer size distribution. Even with a 25% 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.[28] 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.

Hence, farmers are on average too small to adopt Kudu, even with a reduction in the minimum threshold size required for cross-market trade. Consistent with this, only 2% of farmers trade on Kudu. In Appendix Table A.12, we also show that Kudu does not affect the number of traders to which farmers sold in the 12 months prior to endline, the number of *new* traders to which farmers sold, or whether farmers made any sales at market (rather than farmgate). Overall, Kudu does not appear to alter farmers' traditional marketing channels.

However, given the substantial general equilibrium effects of the platform, 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. 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. [29]

7.2 Impacts under scaled implementation

Experiments on market-level interventions inherently ask an artificial question because they require that we leave untreated units to serve as counterfactuals. Any full-scale implementation of this

[28]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.

[29]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.

type of intervention would involve treating all the units in a market, meaning that the most policy-relevant causal effect is the comparison of the outcome when all markets are treated to that when none are. As shown in other recent papers that combine a structural model with an experiment, the model allows us to infer what the impact of the intervention would have been had it been universally implemented (Franklin et al., 2023).

An example of what can be achieved with this approach is shown in Figure 11. This figure plots the model-derived trade flows for treated clusters and for untreated clusters as we shift the intensity of treatment from 0% to 100% of markets. Given that the actual saturation at which we conducted the study is 50%, the reduced-form treatment effects presented in Section 3 are given by the green bracket (raw difference between treatment and control). However, given that the study is generating GE spillovers on the control, the corrected treatment effect presented in Section 6 is given by the red bracket (total effect on the treatment). These two quantities are different by the spillover effect on the control (blue bracket).

More interestingly, however, the model also allows us to form counterfactuals involving other saturations that were not directly observed. Because all values in this figure are defined relative to the pure control outcome, the 0% saturation outcome where all units are untreated is 0 and forms the comparison for all study-level effects. The opposite case is the universal 100% saturation outcome, at which all markets are treated. The difference between these two outcomes is then the total study-level treatment effect (purple bracket), which we believe to be the most policy-relevant comparison that can be made but one that is empirically infeasible to measure. We can also think about study-level impacts at any other saturation, where we compare the full treatment effects across both treated and control units to the 0% outcome (as represented by the dotted line that lies between the treatment and control outcomes).[30]

Several features that are deeply intuitive arise from this picture. The treatment has the effect of allowing treated traders to benefit from arbitrage opportunities that are not available to the control. The fewer markets are treated, the greater are the opportunities available to those who

[30] If we are examining monadic outcome, this is the weighted average of the treated and control outcome where the weight is the saturation. Because this figure examines a dyadic outcome and considers dyads with one market treated as controls, that means that the overall effect is weighted by the probability of a treatment-treatment pairing (.25 at 50% saturation).

are treated. Hence, at 10% treatment saturation traders in treated markets dramatically increase trade, and while these come at the expense of control markets the overall magnitude of the spillover to controls is small. As treatment saturations increase the average benefit of being treated declines dramatically, but because more treated traders increase trading volumes the overall spillover to the control increases as well. Nonetheless, the sum of the treatment and spillover effects is always positive. Interestingly, while the reduced form treatment effect that would be realized by an experimental comparison of treatment to control declines monotonically in treatment saturations, the true total effect of the intervention (the height of the dotted line) increases monotonically, illustrating the extent to which reduced form comparisons can be misleading in general equilibrium. The total study level effect at full treatment saturation represents the largest overall increase in trade that can be generated by the intervention.

8 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 25%. 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.

Atkin, David and Dave Donaldson, "Who's Getting Globalized? The Size and Implications of Intra-national Trade Costs," *NBER Working Paper 21439*, 2015.

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.

Franklin, Simon, Clement Imbert, Girum Abebe, and Carolina Mejia-Mantilla, "Urban public works in spatial equilibrium: Experimental evidence from Ethiopia," Technical Report 2023.

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.

Moser, Christine, Christopher Barrett, and Bart Minten, "Spatial Integration at Multiple Scales: Rice Markets in Madagascar," *Agricultural Economics*, 2009, *40* (3), 281–94.

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.

Porteous, Obie, "High Trade Costs and Their Consequences: An Estimated Dynamic Model of African Agricultural Storage and Trade.," *American Economic Journal: Applied Economics*, 2019, *11* (4), 327–366.

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.

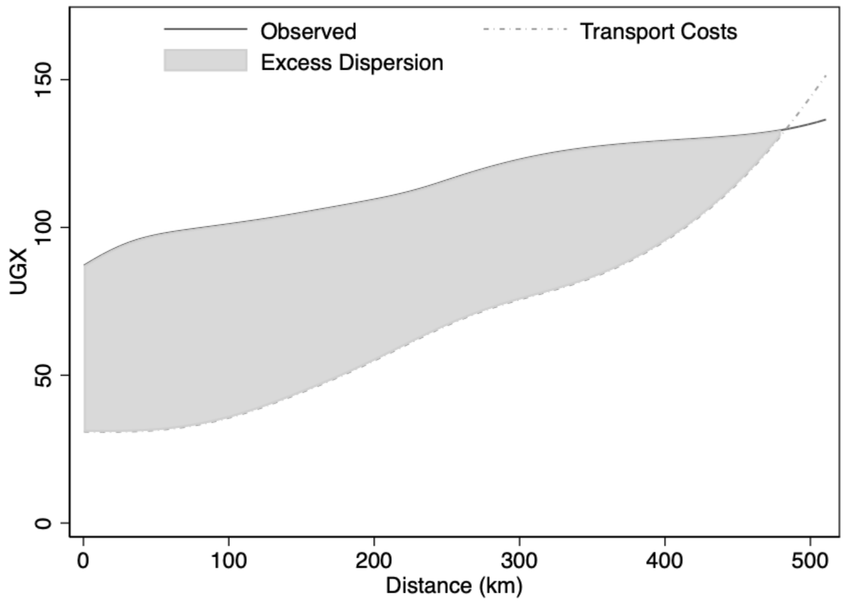
_ **and David Yanagizawa-Drott**, "Estimating Impact in Partial vs. General Equilibrium: A Cautionary Tale from a Natural Experiment in Uganda," *Working Paper*, 2012.

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

Wiseman, Eleanor, "Border Trade and Information Frictions: Evidence from Informal Traders in Kenya," 2023.

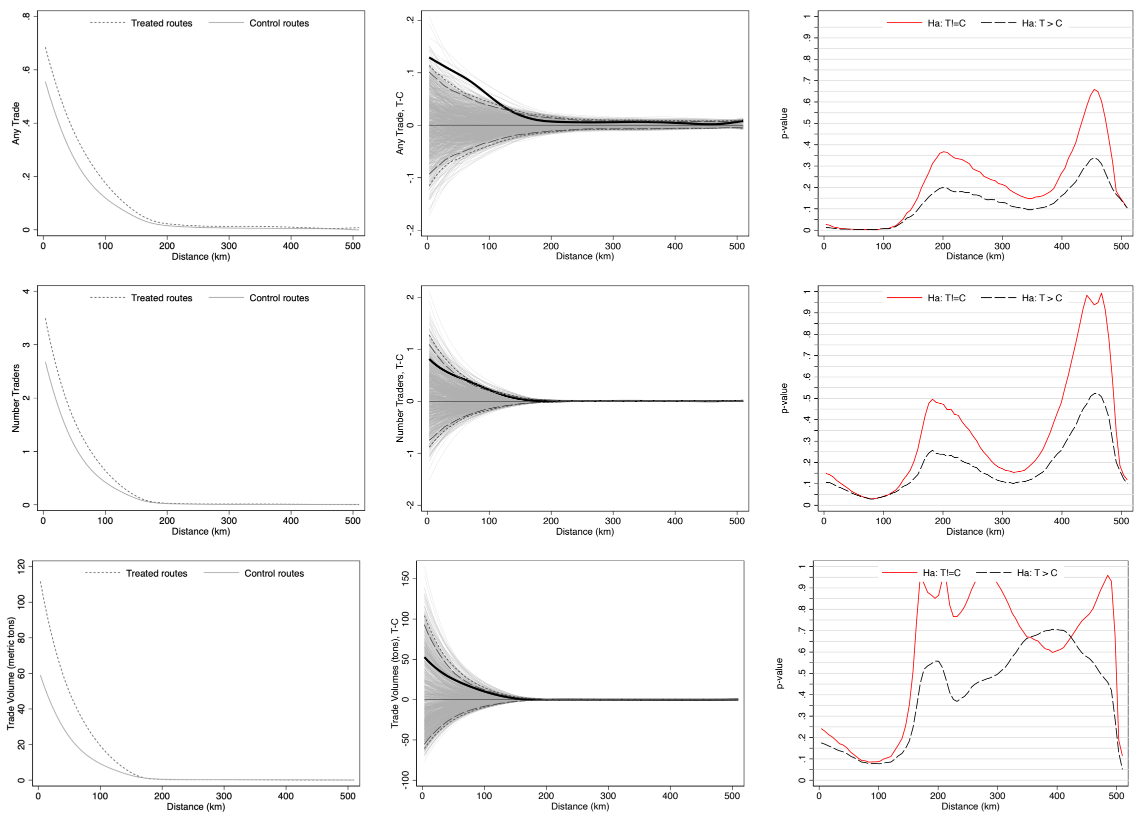
Figures

Figure 1: **Price gaps and transport costs.**



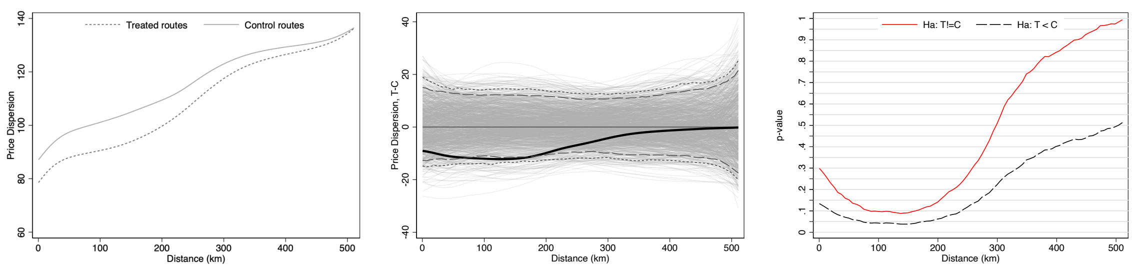
Note: The y-axis presents the absolute difference in prices across each market dyad (pair) in the sample. The solid black line presents the gap observed in prices across each pair of markets in our sample. The dotted line presents estimated transport costs. To generate this prediction, we asked surveyed 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**.



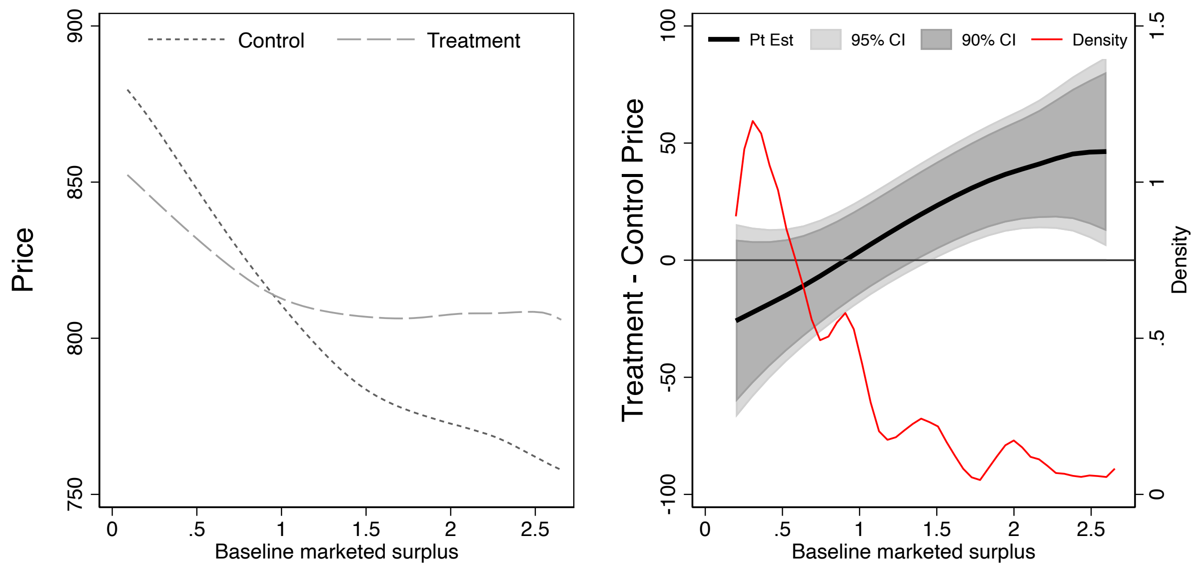
The top lefthand panel presents a non-parametric local Fan regression of the probability of trade along any treated route (dotted line), in which both markets are treated and therefore are connected by Kudu, vs. control routes (solid line), as a function of distance. The middle and righthand panels present results from a randomization inference procedure, in which 1,000 placebo randomized treatment assignments are drawn. In the middle panel, the "reduced form effect," i.e. the difference between treatment and control line in the lefthand figure, from the realized randomization is shown in black. The grey lines show the same effect for each of the 1,000 placebo treatment assignments. Long and short-dashed lines indicate the 90th and 95th confidence intervals of these placebo "reduced form effects." The righthand panel shows the resulting p-values from this randomization procedure, with the p-values from a two-sided test in red and a one-sided test in black. For the one-sided test, the null hypothesis being rejected is that the outcome is greater in control than in treatment. 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**.



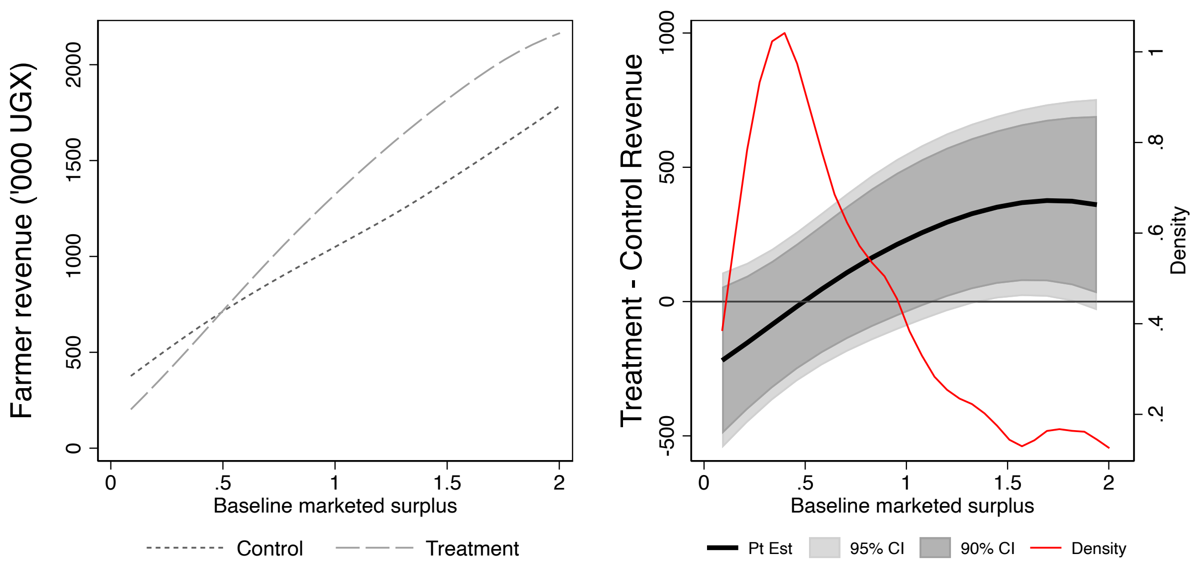
Note: The top lefthand panel presents a non-parametric local Fan regression of the price gaps across markets along any treated route (dotted line), in which both markets are treated and therefore are connected by Kudu, vs. control routes (solid line), as a function of distance. The middle and righthand panels present results from a randomization inference procedure, in which 1,000 placebo randomized treatment assignments are drawn. In the middle panel, the "reduced form effect," i.e. the difference between treatment and control line in the lefthand figure, from the realized randomization is shown in black. The grey lines show the same effect for each of the 1,000 placebo treatment assignments. Long and short-dashed lines indicate the 90th and 95th confidence intervals of these placebo "reduced form effects." The righthand panel shows the resulting p-values from this randomization procedure, with the p-values from a two-sided test in red and a one-sided test in black. For the one-sided test, the null hypothesis being rejected is that the outcome is greater in treatment than in control.

Figure 4: **Market price effects in surplus vs. deficit areas**.



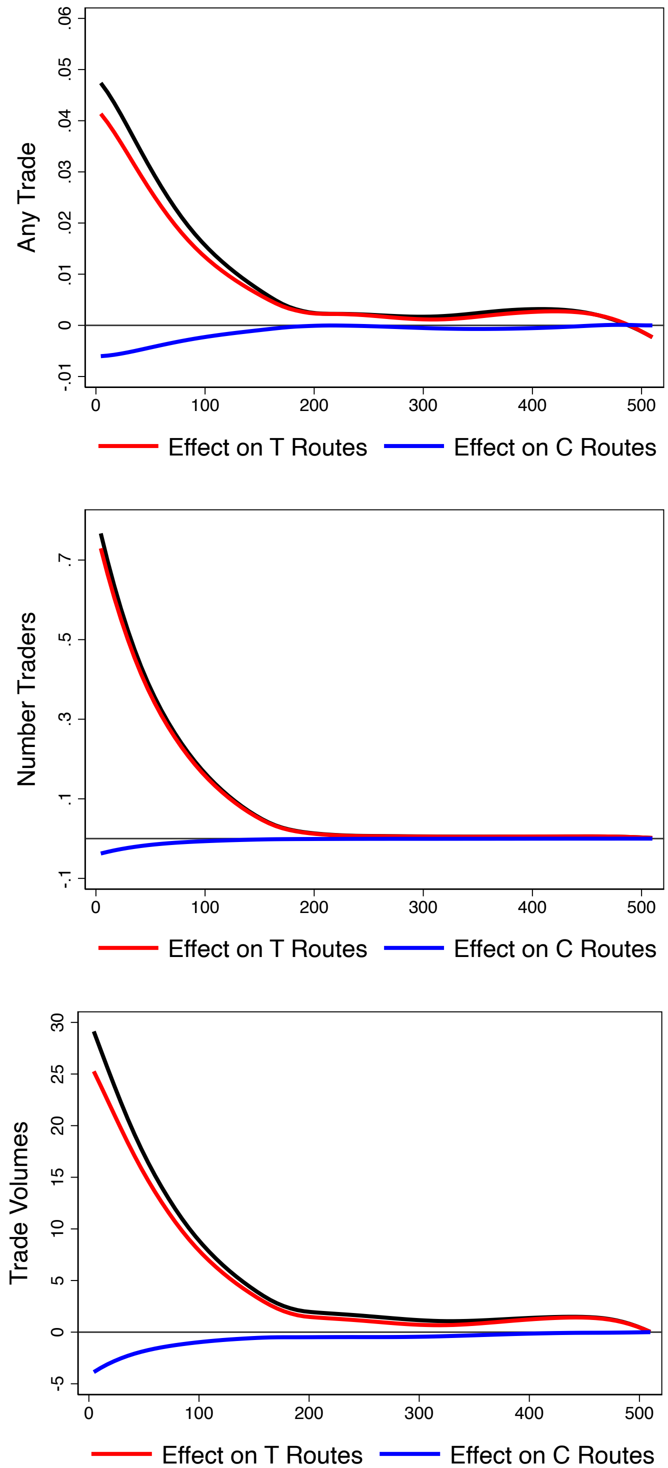
Note: The left panel shows the level of market survey prices in treatment vs. control markets, with respect to the average market surplus per farmer, as measured in tons at baseline. The right panel shows the difference between the two (the treatment effect), along with the 90% and 95% confidence intervals from a bootstrap estimation and the density of observations.

Figure 5: **Farmer revenue effects in surplus vs. deficit areas**



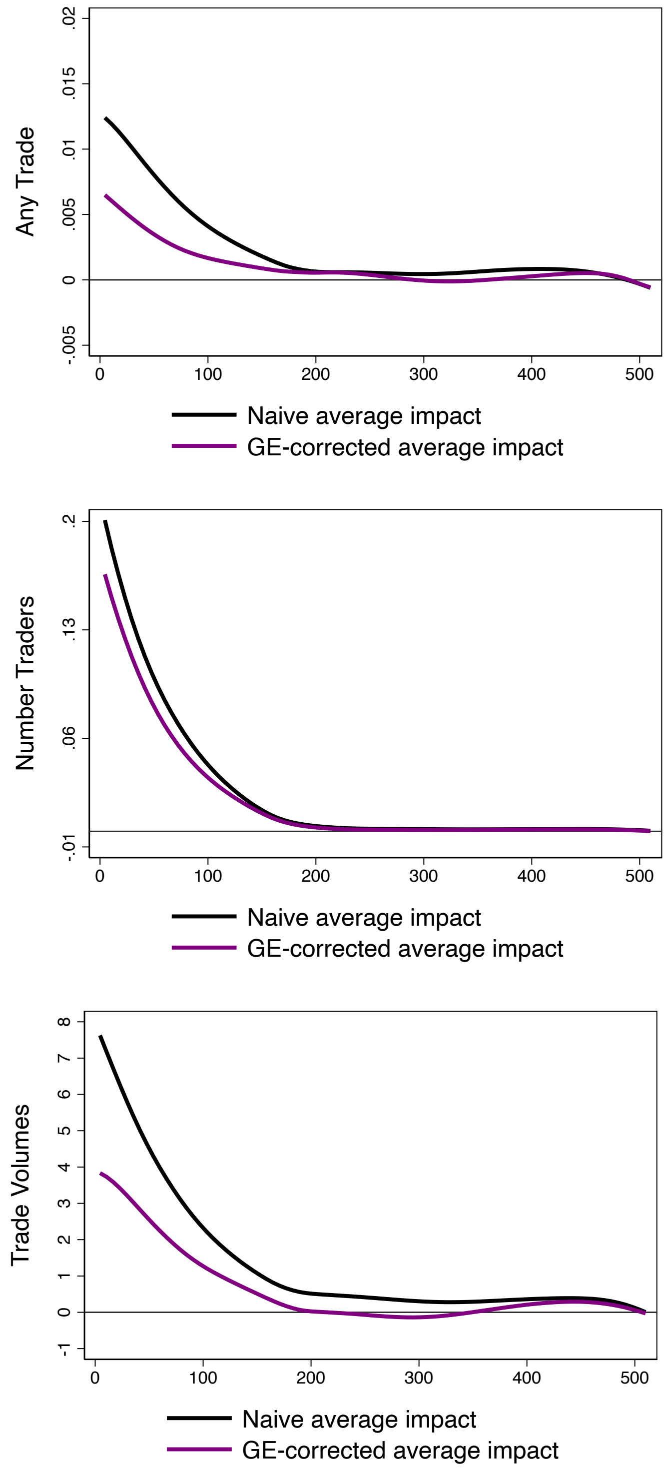
Note: The left panel shows the level of farmer revenue from the household survey in treatment vs. control markets, with respect to the average market surplus per farmer, as measured in tons at baseline. The right panel shows the difference between the two (the treatment effect), along with the 90% and 95% confidence intervals from a bootstrap estimation and the density of observations.

Figure 6: **GE-corrected treatment effects: treated and control route trade flows**.



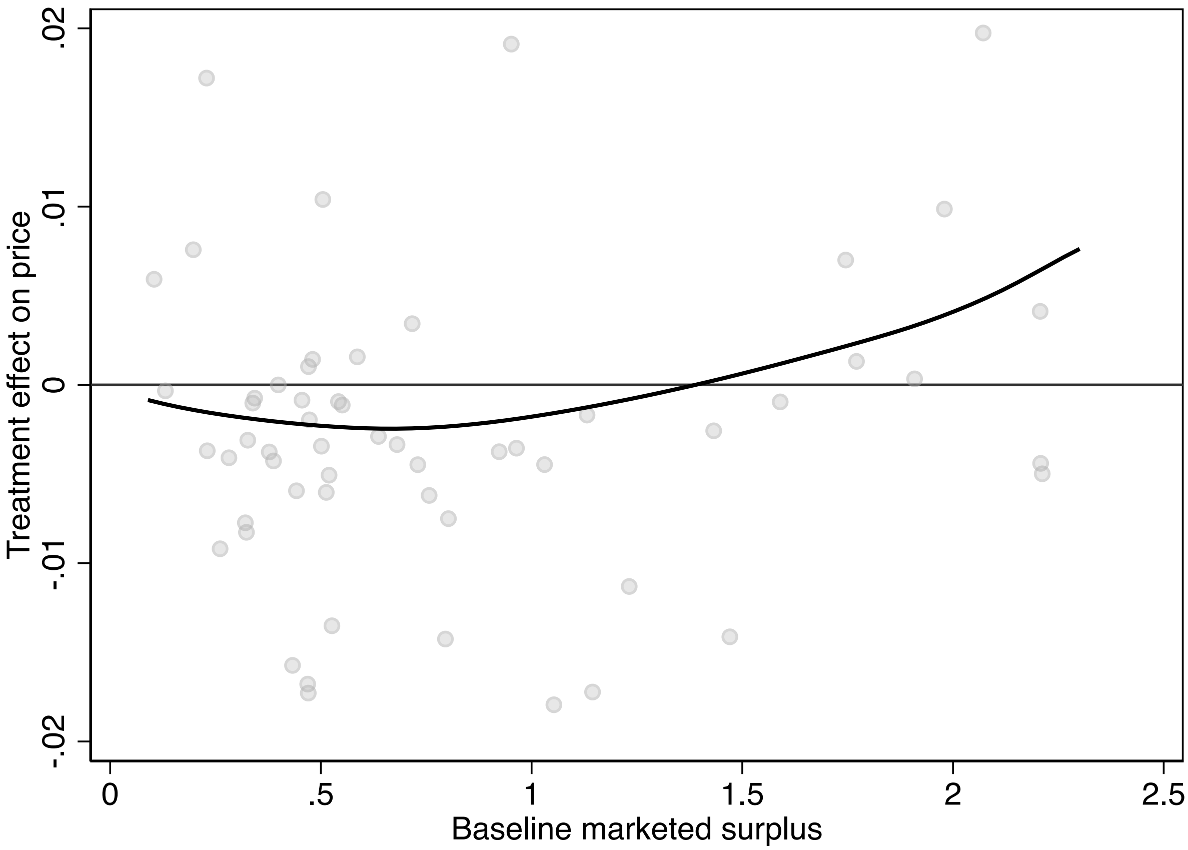
Note: These figures use the structural model to estimate the full general equilibrium effect of the treatment on outcomes in the treatment (red), the control (blue), and the difference between these (black). The panels of the figure examine dyadic measures of any trade, number of traders, and trade volumes, and each outcome is plotted across dyadic distance.

Figure 7: **GE-corrected treatment effects: aggregate trade flows**.



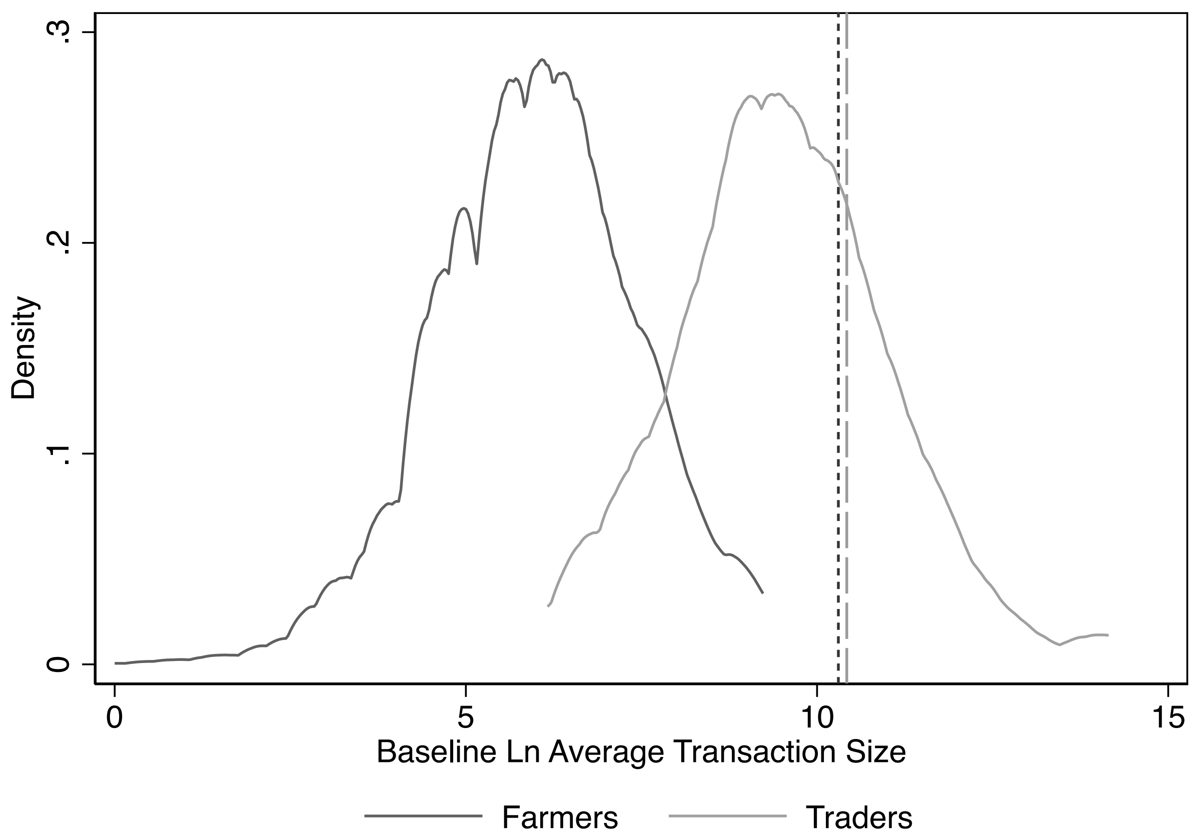
Note: These figures compare the reduced-form treatment control differences (darker line) to the model-derived estimates of the total average effect of the treatment (lighter line). The panels of the figure examine dyadic measures of any trade, number of traders, and trade volumes, and each outcome is plotted across dyadic distance.

Figure 8: **Model-derived price effects**



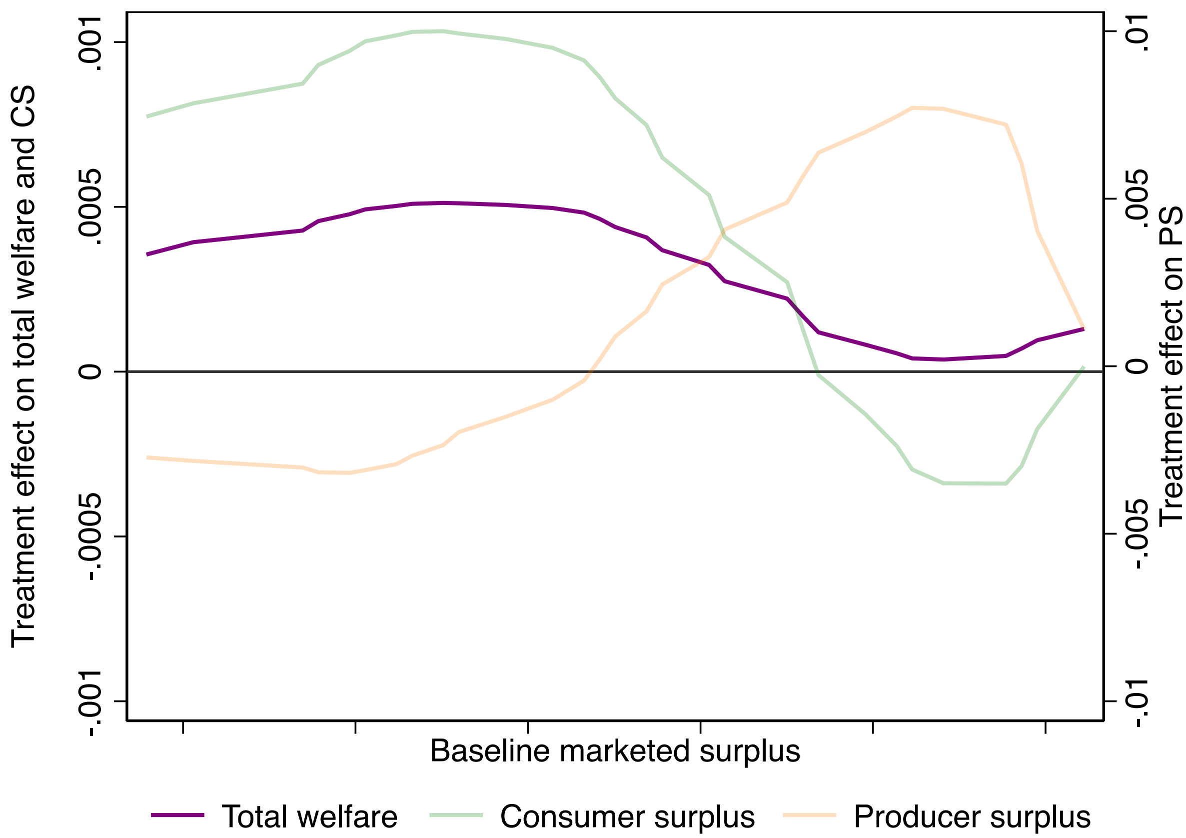
Note: Figure represents the average model-derived impact of the intervention on prices in surplus versus deficit areas. Market-level measure of surplus vs. deficit is taken from the real data and the price variable is the model-derived total impact on prices arising from the intervention.

Figure 9: **Threshold transaction size**



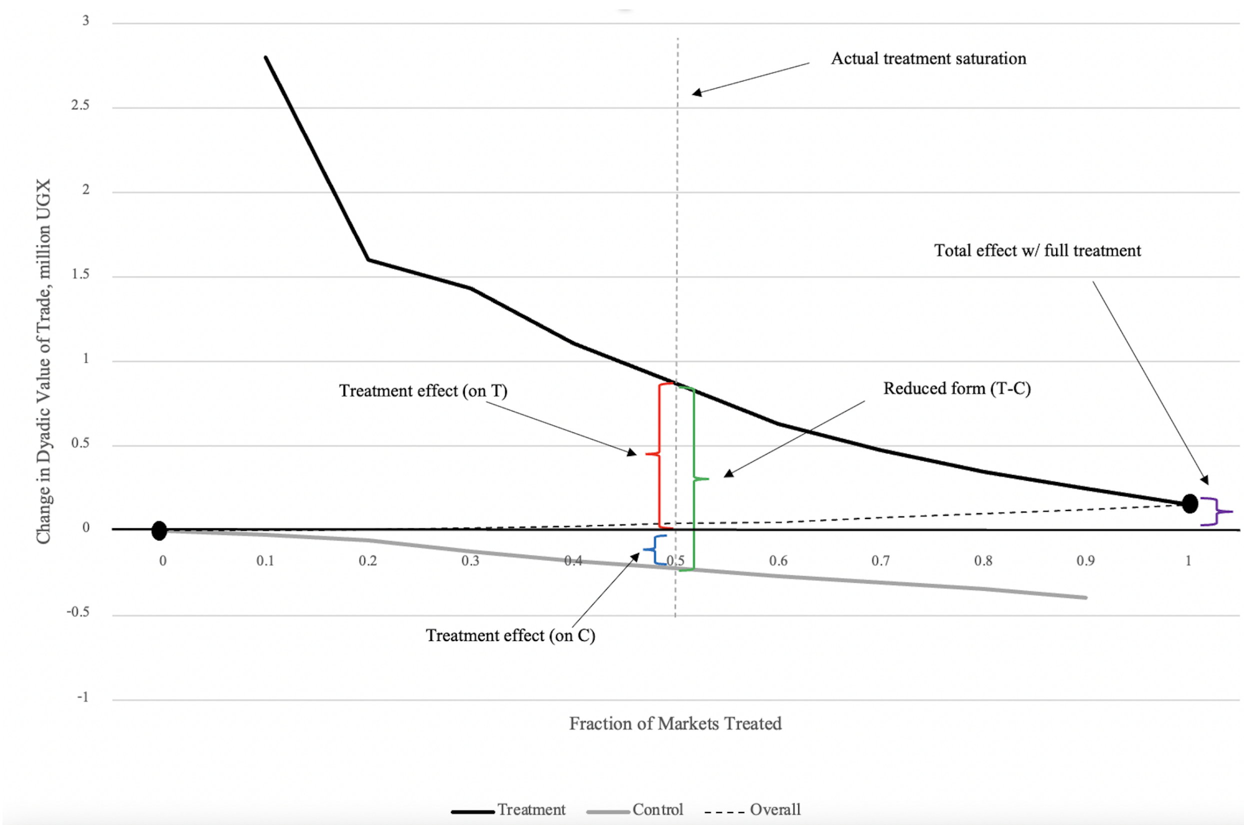
Note: The densities represented in this picture are the log transaction size for farmers (left distribution) and traders (right distribution). The vertical lines represent the model-derived threshold size for trade in the absence of the intervention (dashed vertical line) and in the presence of the intervention (dotted vertical line).

Figure 10: **Welfare effects**



Note: Figure plots the total model-derived impacts of the intervention on producer surplus (yellow), consumer surplus (green), and overall welfare (purple). All three quantities are represented relative to baseline marketed surplus.

Figure 11: **Model-derived effects of treatment intensity on trade volumes**



Note: Figure plots the model-derived total effects on trade volumes as treatment intensity is varied away from the 50% observed in the study. The black line represents the average impact in treated markets, the gray line the average impact in control markets, and the dotted line the average impact across all study locations as the fraction of markets assigned to treatment is increased.

Tables

Table 1: **Reduced form effects on dyadic trade outcomes.**

	Any Trade	Number Traders	Volume (tons)	Price Gaps
Both treated	0.02	0.05	3.03	-6.36
	(0.01)	(0.05)	(1.90)	(5.09)
Dist (10km)	-0.00***	-0.02***	-0.51***	1.00***
	(0.00)	(0.00)	(0.14)	(0.12)
Observations	11236	11236	11236	1443397
Mean DV	0.05	0.19	4.71	116.30

Notes: The first three columns use panel trader surveys, and the final column the bimonthly market price survey, to study dyadic trade outcomes between subcounty pairs. All regressions include the distance between markets as a control. Column 1 examines the binary outcome of any trade on the dyad, Column 2 the number of traders operating between those markets, and Column 3 the volume of maize traded. The final column analyzes absolute price dispersion across each dyad for each round of the market survey. Regressions use all post-treatment data, include fixed effects for the data wave, and standard errors are two-way clustered by each of the subcounties in the dyad, the unit of assignment.

Table 2: **Dyadic trade outcomes by distance**

	Below 200km				Above 200km			
	Any Trade	Num Traders	Vol (tons)	Price Gaps	Any Trade	Num Traders	Vol (tons)	Price Gaps
Both treated	0.06^{**}	0.26^{*}	12.24^{**}	-10.30^{*}	0.00	0.00	-0.05	-4.29
	(0.03)	(0.14)	(6.09)	(5.31)	(0.00)	(0.01)	(0.10)	(5.62)
Dist (10km)	-0.03^{***}	-0.13^{***}	-3.19^{***}	0.76^{*}	-0.00^{*}	-0.00^{*}	-0.01	0.85^{***}
	(0.00)	(0.02)	(0.93)	(0.42)	(0.00)	(0.00)	(0.00)	(0.21)
Observations	3429	3429	3429	456811	7807	7807	7807	986586
Mean DV	0.16	0.61	15.00	97.43	0.01	0.01	0.18	125.03

Notes: Regressions follow the same specification as Table 1, but the sample is partitioned by dyadic distance. The first four columns analyze dyads within 200 km of each other (roughly the median), while the last four columns analyze only dyads further part than this.

Table 3: **Reduced form price effects in surplus and 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

Notes: Table uses monadic data from the bimonthly market survey to analyze treatment effects on average maize selling price. Post-treatment surveys are used with fixed effects at the survey wave level and standard errors clustered by subcounty, the unit of assignment. Column 2 interacts treatment with the baseline marketed surplus in the subcounty. Column 3 interacts treatment with a dummy for being in the top half of the surplus distribution, and Column 4 separately identifies treatment effects in the surplus and deficit areas.

Table 4: **New entrants on average smaller**.

	Profits (mil UGX)		Tons traded		Value traded (mil UGX)	
	Mean	Min	Mean	Min	Mean	Min
Treat_ij	-1.91**	-2.05***	-76.04	-98.83	-39.37	-48.92
	(0.94)	(0.77)	(81.42)	(77.49)	(34.69)	(32.80)
ln d_ij	1.19**	2.05***	170.98*	186.26**	73.59*	81.21**
	(0.48)	(0.46)	(89.11)	(88.49)	(38.50)	(38.21)
Observations	780	780	780	780	754	754
Mean DV	7.39	5.34	255.38	189.14	126.35	91.94
P-value	0.04	0.01	0.35	0.21	0.26	0.14

Notes: Columns 1, 3, and 5 display the average trader size among all traders who are active along a route post-treatment. Importantly, size is measured by *pre-treatment* proxies at baseline. This table therefore shows how treatment induces into treated routes new types of traders (specifically, smaller traders). In Columns 1-2, size is measured by baseline profits of the trading enterprise. In Columns 3-4, it is measured by total tons traded (across all routes, at baseline) and in Columns 5-6 by total value traded (across all routes, at baseline). Columns 1, 3, and 5 show the average size across all traders active on the route, while Columns 2, 4, and 6 show the minimum size among all traders active on the route.

Table 5: **Effects on trader business outcomes**

	Trading Volumes (tons)	Buy-Sell Margins (UGX)	Still in Business	Number New Traders
Treatment	129.76	-10.96	-0.02	-0.01
	(106.01)	(9.28)	(0.02)	(0.14)
Observations	2370	2268	2863	236
Mean of DV	242.94	135.01	0.86	1.05
R squared	0.04	0.02	0.01	0.00

Notes: Table analyzes overall impacts on traders. The first three columns use the two post-treatment trader surveys to analyze total trading volume, margins between buying and selling prices for traders, and whether a baseline trader is still in business. The fourth column uses market-level data on the number of new traders who have entered. Standard errors are clustered at the subcounty level.

Table 6: **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

Notes: Analysis uses the farmer endline survey to analyze treatment effects. Columns 1-3 use the full sample, and Column 4 uses only those with non-zero maize sales. Controls are selected from baseline covariate set using double lasso. Standard errors are clustered at the subcounty level.

Table 7: **Impact of experimental exposure**.

	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

Notes: 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

Table 8: **Parameter estimates**

	No interaction	**Distance interaction**
β_1	-0.25	-0.31
	(0.15)	(0.16)
β_2	N/A	-0.08
		(0.11)
σ	2.32	2.30
	(0.49)	(0.59)
γ	-0.51	-0.48
	(0.14)	(0.19)
θ	3.45	3.50
	(0.64)	(0.65)
κ	0.01	0.01
	(0.00)	(0.00)
σ_η^2	3.91	3.89
	(0.30)	(0.30) ‘
B	1.72	1.68
	(0.14)	(0.28)

Notes: Parameter estimates from a joint estimation of Equations 8-10 using generalized method of moments (GMM). 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 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 (β_1, β_2, σ, γ, θ, κ, σ_η^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. Standard errors (in parentheses) are clustered by sender and receiver subcounty.

Appendix A Tables

Table A.1: **Analysis of Variance in Market Prices**.

	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

Notes: 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 fixed effects (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 two-week round of the survey (and so measures the extent of overall time-series variation), and the fourth column includes both trading center and round fixed effects..

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

	Treat	Control	Obs	T-C *diff*	T-C *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 with a measure of subsequent attrition as the dependent variable. The first row examines midline attrition, the second row endline attrition prior to intensive tracking, the third row attrition during intensive tracking, and the final row overall endline attrition. Standard errors are clustered by subcounty.

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

	Treat	Control	Obs	T-C *diff*	T-C *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 households with a measure of subsequent attrition as the dependent variable. The first row examines attrition prior to intensive tracking, the second row attrition during intensive tracking, and the final row overall endline 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 market-level 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: **Dyadic market survey balance in absolute price gaps**.

	Maize	Beans	Bananas	Tomatoes
Both treated	10.56	22.11	116.2	1.916
	(7.715)	(31.59)	(354.4)	(2.753)
Mean DV	131.5	559.4	6361.7	72.37
N	26218	21129	20196	26149

Notes: Analysis uses the dyad-level averages across the two pre-treatment rounds of the market survey to examine balance in absolute price gaps within dyad pairs. Standard errors are two-way clustered by each of the subcounties in the dyad, the unit of assignment.

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 in baseline trader data but using only the attrited endline sample, with weights reflecting survey sampling and intensive tracking as used in the main analysis. 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. Standard errors for the p-values are clustered by subcounty, the unit of assignment.

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 in baseline household data but using the attrited endline sample, with weights reflecting survey sampling and intensive tracking as in the main analysis. 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. Standard errors for the p-values are clustered by subcounty, the unit of assignment.

Table A.8: **Predictors of take-up**.

	Farmers		Traders	
	Ever used Kudu	Completed deal on Kudu	Ever used Kudu	Completed deal on Kudu
Tons sold (log)	0.111***	0.019	0.128***	0.067**
	(0.032)	(0.066)	(0.032)	(0.034)
Age	-0.003	-0.001	-0.004	0.001
	(0.003)	(0.007)	(0.006)	(0.006)
Education	0.005	-0.009	0.003	-0.016
	(0.008)	(0.022)	(0.016)	(0.012)
Female	-0.314***	-0.272*	-0.009	-0.080
	(0.078)	(0.151)	(0.214)	(0.265)
Observations	1322	1322	716	716
Mean DV	0.26	0.02	0.80	0.22

Notes: Columns 1-2 present the predictors of take-up among treated farmers, while 3-4 present those among treated traders, estimated via probit (with standard errors clustered at subcounty level). The outcome in Columns 1 and 3 is ever posting to Kudu, while the outcomes in Columns 2 and 4 are successfully completing a transaction on Kudu. Regressors are all measured at baseline and are: (log) quantity of maize sold or traded, age of respondent, years of education completed, and a dummy for being female.

Table A.9: **Impacts on farmer's production**.

	Harvests Maize KG	Harvests Maize KG
Treat	-15.54	-45.93
	(131.7)	(97.16)
Surplus		577.6**
		(250.7)
Treat*Surplus		108.9
		(329.6)
Observations	2768	2768
Mean of DV	1380.2	1380.2
R squared	0.343	0.357

Notes: Analysis uses endline farmer surveys to measure impacts on production. Column one presents treatment effects on maize harvests (in kg), while column 2 presents an interaction of treatment and baseline surplus status of the area. Standard error are clustered by subcounty, the level of randomization.

Table A.10: **Impact of price information alone on price gaps**.

	All	Below 200km	Above 200km
Both treated by price info-only	4.17	6.89	2.88
	(8.31)	(10.14)	(7.32)
Dist (10km)	0.96***	1.02	0.83**
	(0.17)	(1.02)	(0.40)
Observations	220669	67899	152770
Mean DV	127.54	110.44	135.14

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 price information-only intervention 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.11: **Heterogeneity by unexpected price deviations**.

	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 table examines whether dyadic connections through the platform induce faster convergence of prices when unexpected shocks hit. To calculate an unexpected dyadic price deviation, we first compute the average price for each district x month-of-year. Then, for each round of the market survey, we calculate the 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 the difference in these deviations across the dyadic pair. This then 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 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 price information is revealing larger than expected 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.

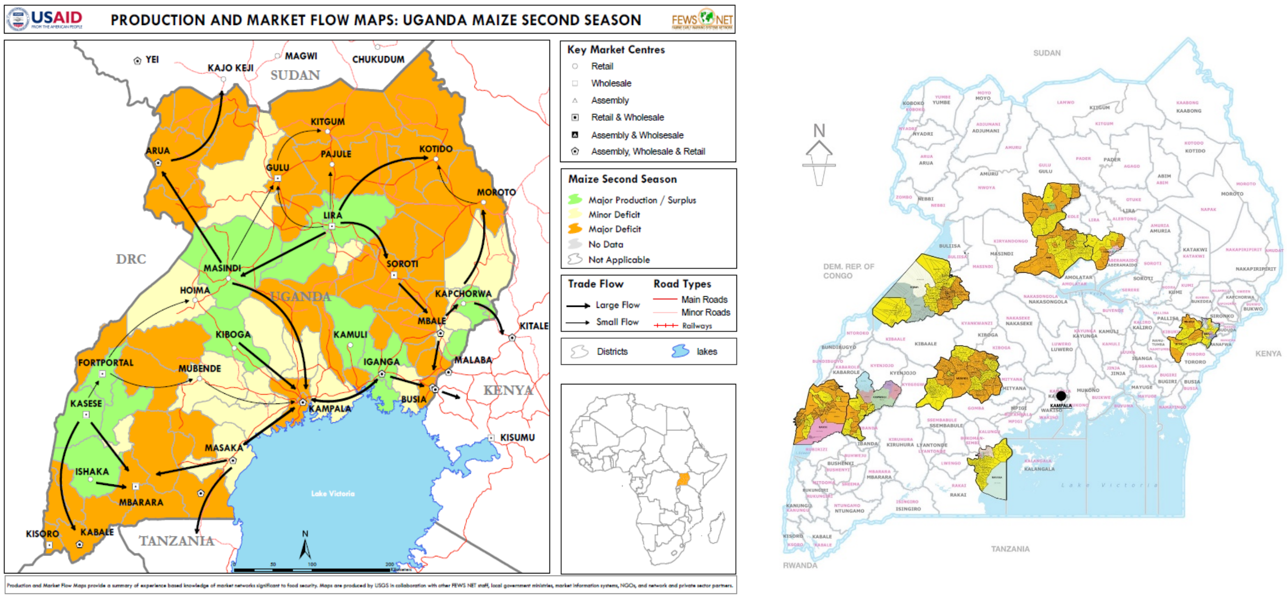
Table A.12: **Impacts on marketing**.

	Number Traders		Number New Traders		Any Sales at Mkt	
Treat	-0.02	-0.02	0.02	0.05	0.00	0.01
	(0.08)	(0.09)	(0.03)	(0.03)	(0.03)	(0.03)
Surplus		0.39***		0.16***		-0.05
		(0.10)		(0.05)		(0.03)
Treat*Surplus		-0.01		-0.09		-0.04
		(0.13)		(0.09)		(0.05)
Observations	2768	2768	2768	2768	2774	2774
Mean of DV	1.05	1.05	0.22	0.22	0.26	0.26
R squared	0.05	0.07	0.01	0.01	0.04	0.05

Notes: Columns 1, 3, and 5 present treatment effects on the number of traders to which the farmer sold in the last 12 months, the number of *new* traders to which the farmer sold in the last 12 months, and whether the farmer sold anything at market in the last 12 months. Columns 2, 4, and 6 present an interaction of treatment and baseline surplus status of the area. Standard error are clustered by subcounty, the level of randomization.

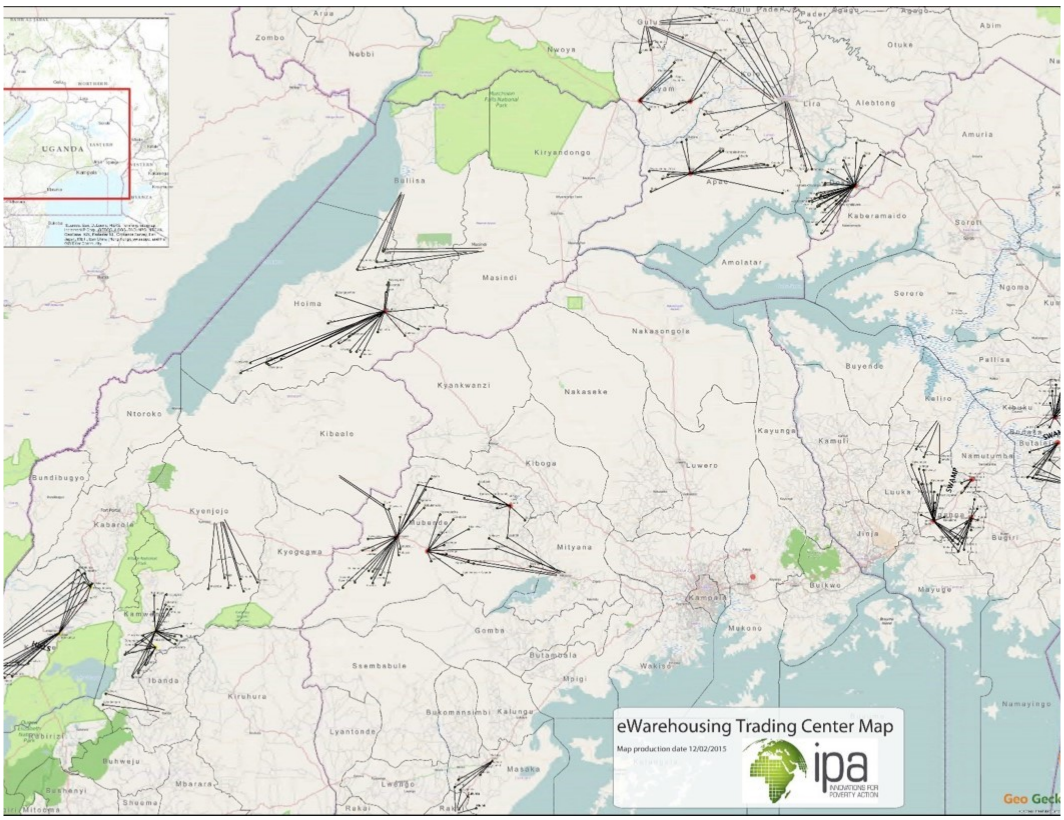
Appendix B Figures

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



Note: 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**

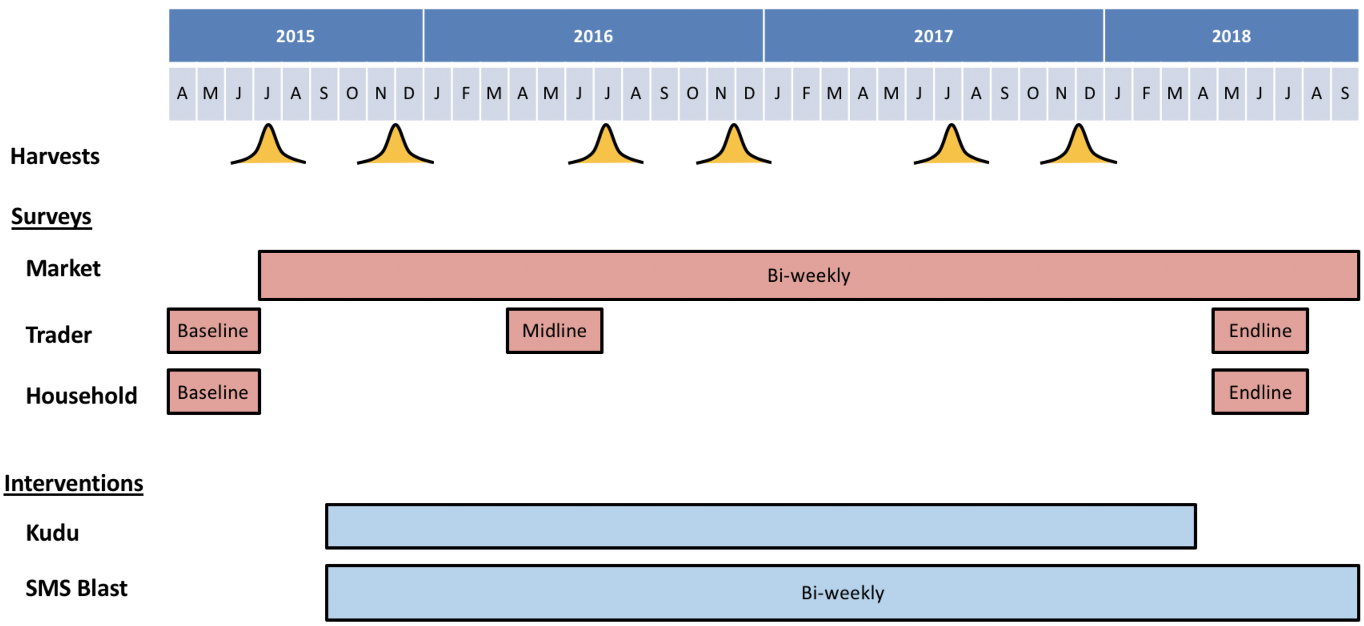


Figure B.4: **CONSORT Diagram of Study Recruitment and Attrition**

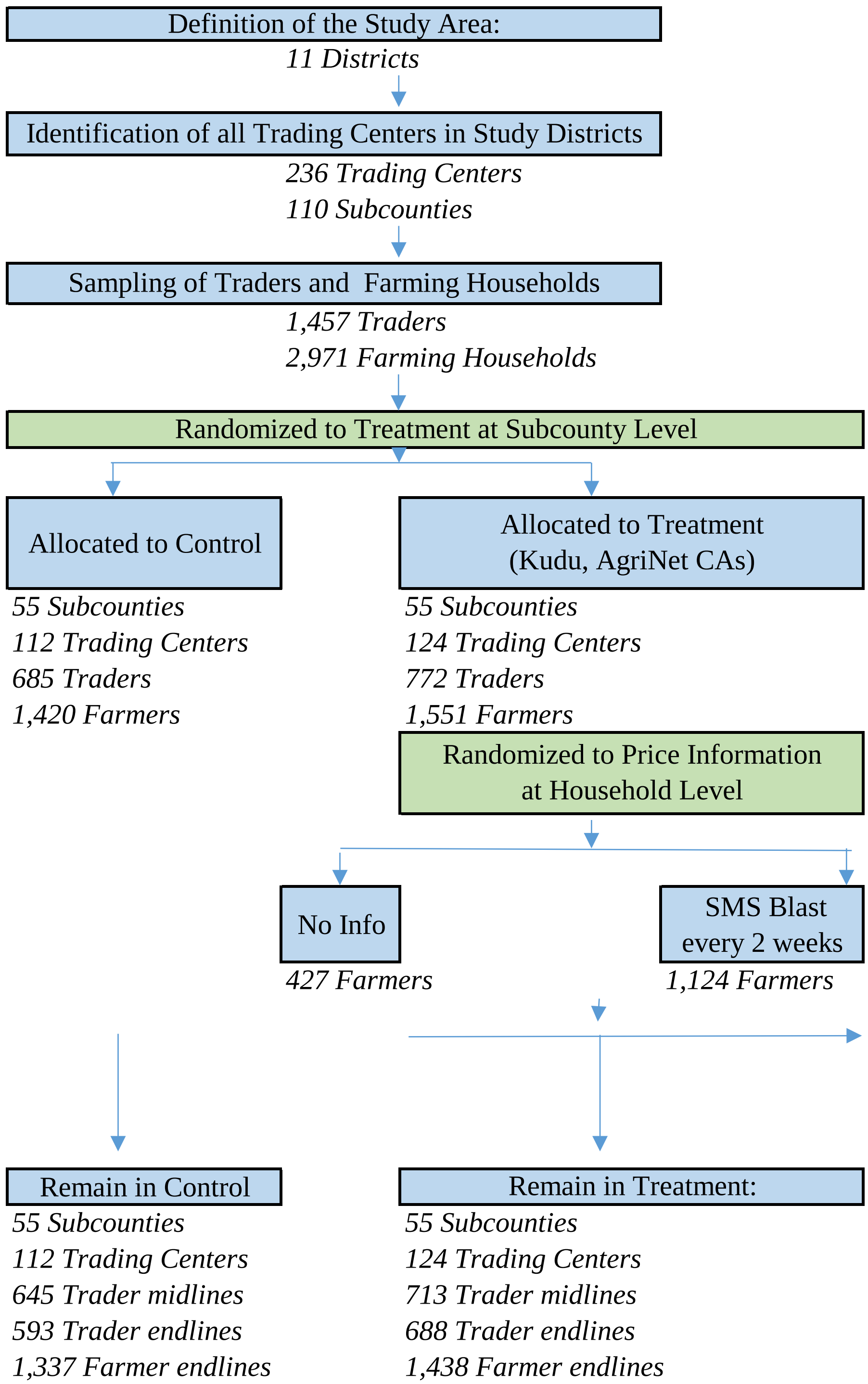
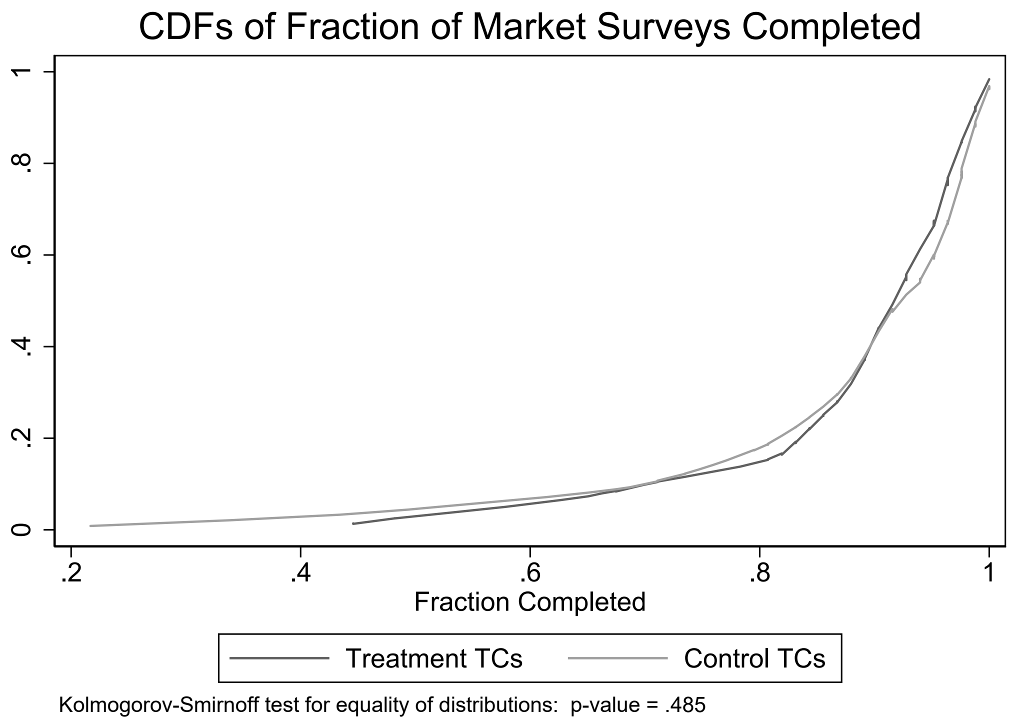
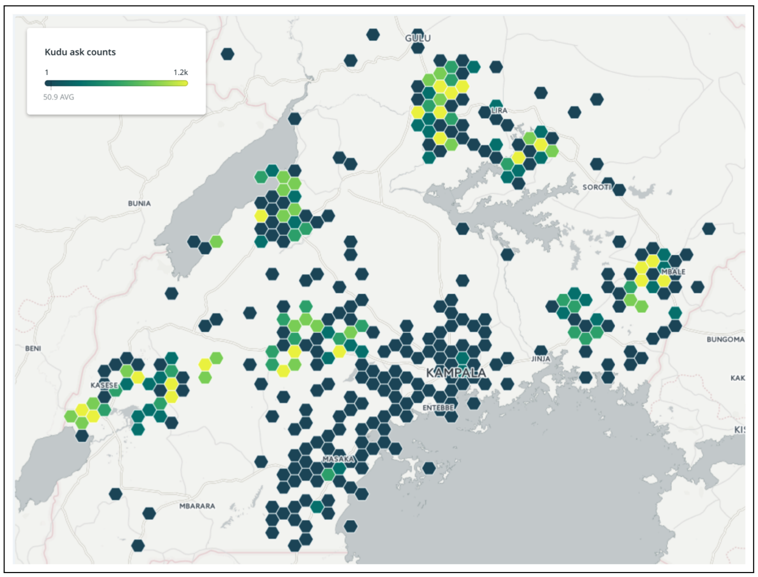


Figure B.5: **Attrition from the Market Survey**



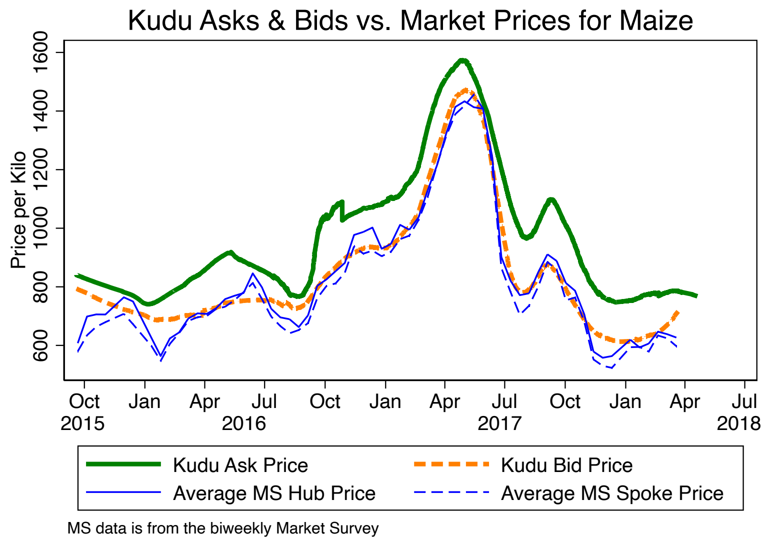
Notes: The figure shows the Cumulative Density Functions (CDFs) of the fraction of intended surveys (83) that were completed for each TC, separating out the treatment and control TCs. The KS test fails to reject that the two distributions are the same.

Figure B.6: **Volume of Bids and Asks**



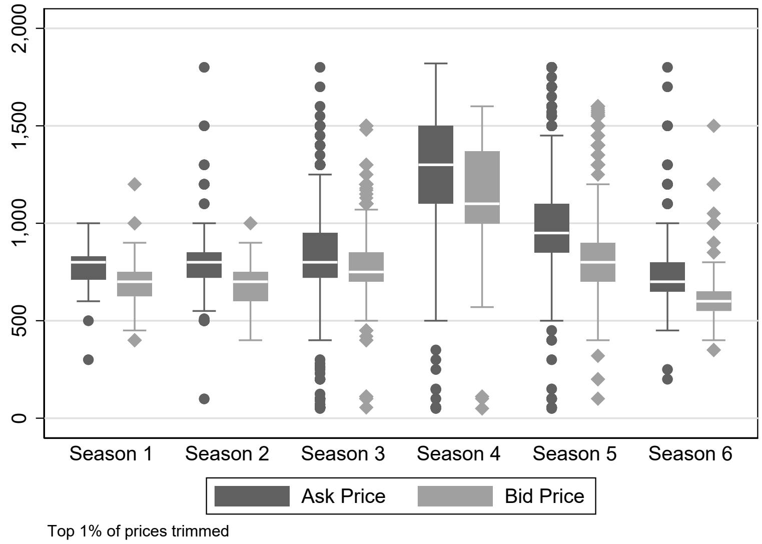
Note: figure represents the spatial density of bids and asks over the study period.

Figure B.7: **Maize Prices in Kudu vs. Market Survey**



Note: figure presents a comparison of average prices across time from the Kudu platform versus from the bimonthly market survey data.

Figure B.8: **Distribution of Ask and Bid Prices, by Season**



Note: figure presents a box-and-whisker plot of the distribution of bid and ask prices across Kudu over the six seasons of the study.

Figure B.9: **Cumulative Sales on Kudu**

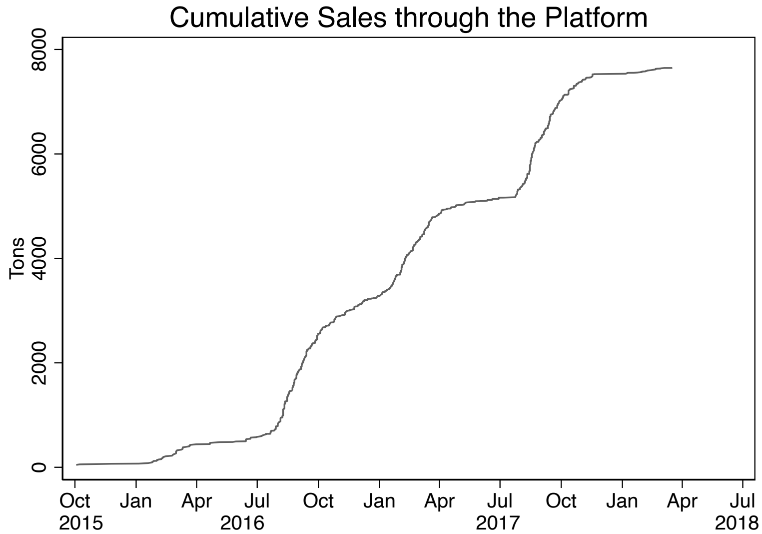
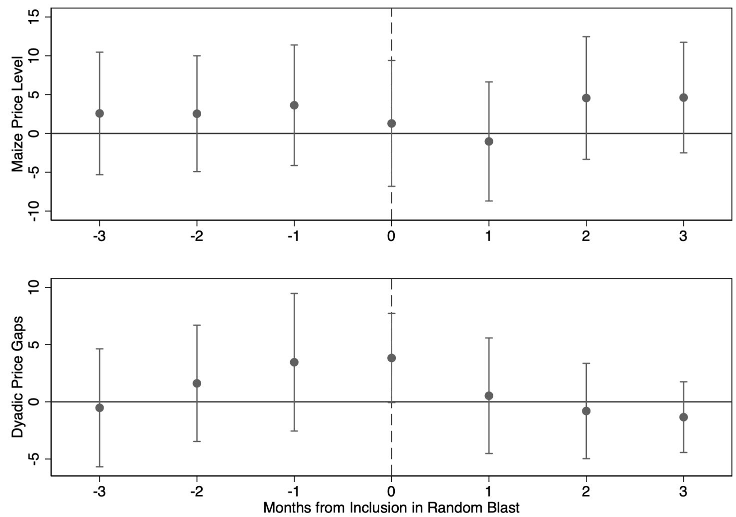


Figure B.10: **Impact of the Random Blast**. This figure presents an event study of the impact of the "Random Blast," in which prices for that randomly-selected market were sent out to every market in the Kudu system (i.e. every other treated market). The top panel presents monthly leads and lags of the treatment on price levels. The bottom panel examines impacts on price gaps 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.