Market Access, Trade Costs, and Technology Adoption: Evidence from Northern Tanzania*

Shilpa Aggarwal† Brian Giera‡ Dahyeon Jeong§ Jonathan Robinson¶
Alan Spearot‖

November 30, 2017

Abstract

We study how geographical remoteness affects agricultural productivity via access to input and output markets using novel, self-collected data on the supply chains for chemical fertilizer and maize in all of the 570 villages in the Kilimanjaro region of Northern Tanzania. In reduced form, adoption of fertilizer for villages in the 3rd and 4th highest remoteness quartiles is 31-47% lower than the least remote quartile. Villages in the highest remoteness quartiles are 36% less likely to interact with an output buying intermediary. We find evidence that reduced access to markets (either directly, or through middlemen) is an important contributing factor to this disparity. Using point-to-point travel costs for the universe of villages in Kilimanjaro, we find that the standard deviation of travel cost-adjusted fertilizer prices is 15% of the mean, and that 20% of the villages face fertilizer prices that are 30% higher than the lowest-cost village; on the output side, the best available price of maize for 40% of villages is 30% lower than the best price overall. To quantify these effects on the profitability of fertilizer and thus adoption we develop a spatial model of agro-retailer pricing, farmer investment, and intermediary activity, and find that a counterfactual 50% reduction in travel costs along the supply chain accounts for 16% of the reduced form relationship between remoteness and adoption of fertilizer. Targeted counterfactuals suggest that access to local input and output markets are equally important for this effect, while there is little effect of the costs to bring inputs to retailers from urban hubs.

JEL Codes: F14, O12, O13, O18, Q12

Keywords: market access, inputs, technology adoption, transport costs, roads

*We are grateful to Sheah Deilami-Nugent and David Roughgarden for outstanding work coordinating field activities, and BRAC Tanzania and IPA Tanzania for administrative support. We thank Patrick Olobo for his collaboration. Anouk Foure, Petra Gram, Matthew Krupoff and Alexandra Murphy provided excellent research assistance. We thank Lauren Falcao Bergquist and seminar participants at many seminars and conferences for helpful comments. We acknowledge funding from the IGC (grant 1-VRC-VTZA-VXXXX-89307), the National Science Foundation (SES 1658942) and Innovations for Poverty Action. The data collection approval was approved by the IRBs of UCSC and the Indian School of Business, and by the Tanzania Commission for Science and Technology (COSTECH).

†Indian School of Business, email: shilpa_aggarwal@isb.edu

‡Amazon Research, email: bgiera14@gmail.com

§University of California, Santa Cruz, email: dajeong@ucsc.edu

¶University of California, Santa Cruz and NBER, email: jmrtwo@ucsc.edu

‖University of California, Santa Cruz, email: acspearot@gmail.com

1 Introduction

It is widely believed that poor access to markets – due mainly to poor transportation infrastructure – limits agricultural productivity in rural areas of developing countries, by making it harder to access productivity-enhancing inputs like fertilizer and to obtain high prices for harvest output (World Bank, 2008; 2017).[1] However, while remoteness no doubt limits market access, there is little research to rigorously quantify its effect, especially in the case of input adoption.

In this paper, we rigorously document market access among farmers in the Kilimanjaro region of Northern Tanzania. Our data collection exercise spans the entire supply chain in all 570 villages in the Kilimanjaro region, including (1) surveys with a sample of 550 farmers in 115 randomly selected villages; (2) a census of all 395 agro-retailers, or "agrovets," in the region, and follow-up surveys with 369 agro-retailers that sell fertilizer; (3) collection of information on road quality, travel times, and travel costs to all villages from their respective local markets as well as from 3 major urban centers, and travel times and costs between each market and each major urban center; (4) driving times and distances pulled from Google Maps API for the universe of bilateral village pairs, as well as for pairs of villages and major urban centers in the region; and (5) interviews with maize-buying agents and distributors. This region includes a great deal of heterogeneity in remoteness – varying from villages just a few kilometers from the major town where fertilizer is distributed to villages located in remote mountains 200 km away – and so provides enough variation to examine fairly substantial changes in travel time to the urban hub.

The first part of our paper is a reduced form investigation of the correlation between input usage, market access and remoteness. We define remoteness as travel time from the regional hub, Moshi, and examine remoteness by quartile. We find that input adoption is much lower in remote areas: adoption of fertilizer in the 3rd and 4th highest cost quartiles is 21-32 percentage points lower than in the lowest cost quartile (equivalent to 31-47% in percentage terms). Effects on quantities are even larger (equivalent to a 67% decline). The effects of remoteness on adoption of improved/hybrid seeds are of a similar magnitude. All adoption effects are qualitatively similar when farmer attributes and soil conditions are controlled for.

We focus on access to input and output markets as an explanation for this gap. On the input side, as a summary statistic, we calculate travel-cost adjusted prices for fertilizer for the universe of villages in the region (constructed by adding observed retailer prices and point-to-point costs of travel). We find substantial heterogeneity in this measure: the standard deviation in travel cost-adjusted prices is 15% of the mean, and 20% of villages face "delivered" prices 30% higher than the lowest-cost village. Our data suggests that those living in villages located at greater-than-median remoteness have to travel 4-5 kilometers (56-73%) farther to access their lowest delivered price, and at the agrovet, they are sold fertilizer that is marked-up 33-42% more than in less remote areas.

[1] Transportation infrastructure is particularly underdeveloped in Africa as the continent has only 137 kilometers of roads per 1000 square kilometers of land area, with only a quarter paved. In contrast, the average for developing countries outside the region is 211 kilometers of roads per 1000 square kilometers, with more than half paved (World Bank, 2010). For comparison, the US has 679 kilometers per 1000 square kilometers, with nearly 2/3 paved.

On the output side, the story is similar: to measure general selling conditions, we construct the travel cost-adjusted selling price for maize sales, which is the price available at every weekly market minus the cost of getting there and back. Then, for each village, we find the maximum of this adjusted price. The standard deviation in the best travel cost-adjusted selling price of maize is also 15% of the mean, and for 40% of villages, this maximum is 30% lower than the village with the best adjusted selling price. Independent of prices, sellers in remote areas have to travel 1.7-3.9 kilometers (88-195%) farther to reach their primary weekly market. We also document the likelihood of a visit from an output-buying intermediary, and find that villages in the highest remoteness quartile are 21 percentage points less likely (36%) to have an intermediary visit their village.[2]

To quantify the impact of transport costs and other factors on input adoption, we develop a spatial model of the market for fertilizer and maize. In the model, the decision to adopt fertilizer is based on local output prices, idiosyncratic farmer productivity, the distribution of input prices, and idiosyncratic shocks. Transportation costs affect the distribution of prices by increasing the cost for the farmer to reach a particular agrovet, increasing the costs of agents to reach villages to buy output, and increasing the costs of agrovets to acquire fertilizer. To account for the fact that we may not fully capture the bundle of inputs that a farmer purchases at a given location (or other heterogeneity in pricing power by an agrovet), we allow for an unobserved agrovet-specific scalar on the retail price of fertilizer. This yields a first-order condition for each agrovet that is a function of transport-adjusted demand, and its firm-specific ability to maintain a high price.

We aggregate fertilizer demand in standard multinomial logit form, and essentially execute a Berry (1994) inversion in "reverse" to recover agrovet-specific effects that rationalize mark-ups (which we measure). Given the structure of the model, we can perfectly match agrovet pricing, and then conditional on this pricing, we use adoption decisions and observed agent activity to calibrate implied local productivity of fertilizer, conditional on other local attributes (output prices and average farm size). Finally, the maize market is pinned down by a local market clearing condition that is a function of local demand, demand from the regional hub (via agents), and local supply, which ultimately depends on adoption of fertilizer.

We use the model to run two primary counterfactuals. In the first, we iteratively lower trade costs along the supply chain to zero to understand the role of transportation in adoption decisions. Evaluating the relationship of remoteness to *predicted* adoption decisions at each counterfactual trade cost, we then compare this to the relationship of remoteness to *observed* adoption decisions. We find that a 50% reduction in trade costs along the supply chain accounts for 16% of the reduced form association between adoption and remoteness. In the second counterfactual, we evaluate trade shocks specific to each part of the supply chain. We find little effect of changes in the cost of shipping fertilizer from the distributor to local retailer. However, we find sizeable elasticities of adoption to transport costs when changing local costs (farmer traveling to an agrovet) and agent costs (agent traveling from the hub to the village). In both cases, the elasticity of adoption to transport costs

[2] Maize-buying intermediaries - "agents" - typically visit villages soon after harvest and buy maize in bulk to profit from either inter-temporal or spatial arbitrage opportunities.

is higher in more remote villages.

This paper sits at the intersection of trade and development economics, and we hope to provide value to both literatures. Our primary question considers the impact of remoteness on the price, availability, and adoption of fertilizer by rural farmers. Sub-Saharan Africa has lagged far behind the rest of the developing world in agricultural technology adoption (World Bank 2007) despite evidence that improved technologies could generate large *yield* increases (i.e. Duflo, Kremer and Robinson 2008; Beaman et al. 2013; Stewart et al. 2005; Udry and Anagol 2006). The *profitability* of these technologies thus depends on the relative prices of fertilizer and crop output, and on the size of the yield increase. The literature is more divided on whether these technologies are profitable, with some papers finding large returns (i.e. Duflo, Kremer and Robinson 2008) and others lower or even negative returns (i.e. Beaman et al. 2013). While this previous literature has mostly focused on measuring yield increases, profitability is equally affected by access to technology and sales opportunities, the focus of this paper. Our results quantify the extent to which profitability, and thus adoption, will tend to be lower in more remote locations.

Our paper is also differentiated from much of the development literature by focusing on market access, rather than on demand side explanations like farmer knowledge and learning spillovers (Foster and Rosenzweig, 1995; Conley and Udry 2010; Bandiera and Rasul, 2006; Emerick, 2017), credit, liquidity or insurance constraints (Bardhan and Mookherjee, 2011; Maitra et al., 2017; Karlan et al., 2015), or behavioral explanations (Duflo, Kremer and Robinson, 2011; Hanna, Mullainathan, and Schwartzstein, 2014).[3] Our work is most closely related to Suri (2011), who shows that many Kenyan farmers with high gross returns to hybrid seeds choose not to adopt them because the fixed costs of obtaining seeds are too high, presumably due to travel costs. Our paper is differentiated by focusing on heterogeneity in market access, rather than on heterogeneity in returns.

Our paper is related to a rapidly growing literature about the effect of roads or other infrastructure improvements on development outcomes and on the spatial distribution of economic activity.[4] Many of these papers evaluate large-scale infrastructure programs as natural experiments, or by employing structural techniques, and thus provide causal evidence on the effect of *roads* on various outcomes. The key difference in our paper is that we focus narrowly on the specific effect of transportation costs on market access (i.e. the actual time and money costs of transportation and the presence of intermediaries and the prices they charge) in isolation, without changing other margins.[5] Building roads may change many outcomes other than just prices, including consumption diversity (Aggarwal, 2017), human capital investment (Adukia et al., 2016, Aggarwal 2017), migration (Morten and Oliveira 2016), occupation choice (Asher and Novosad 2016), as well as many others such as electrification, proximity to health care, etc.[6] By contrast, our goal is to focus solely

[3] See Foster and Rosenzweig (2010) and Jack (2013) for reviews of this literature.

[4] A partial listing of papers includes Aggarwal (2017), Alder (2017), Adukia et al. (2016), Asher and Novosad (2016), Banerjee et al. (2012), Bird and Straub (2016), Bryan and Morten (2017), Gertler et al. (2014), Ghani et al. (2016), Khanna (2016), Shamdasani (2016), and Storeygard (2016). See Donaldson (2016) for a review.

[5] Technological advances may make it possible to decouple market access from traditional road infrastructure. For example, Rwanda has a "droneport" already under construction just outside the city of Kibuye, where drones capable of transporting cargo of up to 20 kilos over a distance of 100 kms already exist.

[6] Indeed, several papers in this literature use overall economic development (as proxied by night lights) to capture

on the effect of remoteness on intermediary entry and pricing, with special emphasis on chemical fertilizer.[7]

Our work is related to a voluminous trade literature. Within this literature, price differentials across space can be attributed to three primary components – marginal trade costs (e.g. Donaldson, forthcoming; Eaton and Kortum, 2002; Keller and Shiue, 2007; Sotelo, 2016), spatially varying mark-ups (Atkin and Donaldson, 2015; Asturias et al., 2017), and the organization of intermediaries (Allen and Atkin, 2016; Dhingra and Tenreyro, 2017; Bergquist, 2017; Casaburi and Reed, 2017). Simply quantifying these price differences is important for the literature, as there is a dearth of data studying rural markets, and in particular, access to inputs. We collected price and sales information by firm, input-type and brand - essentially "scanner" data - including wholesale prices for these items, which facilitates an exact measure of retail mark-ups. Further, our unique transportation surveys allow us to calculate the exact cost of acquiring inputs for all possible locations in our sample, providing a comprehensive mapping of input and output market access within the region.

Our paper is closely related to Atkin and Donaldson (2015), who estimate trade costs in a situation in which an oligopolist intermediary buys products at wholesale prices, transports them to distant markets and sells them directly to consumers. By contrast, we are interested in how trade costs affect the buying decisions of final consumers (in this case, farmers), entry of output buying intermediaries, as well as pricing decisions by retailers. Though not directly comparable since they are at different points in the supply chain, our average ad-valorem "trade costs" of farmers procuring fertilizer turn out to be similar to those of the intermediaries in Atkin and Donaldson (2015). Our costs, however, are calculated over a much shorter trip.[8] Using our quantitative model, we find that farmers are particularly sensitive to the costs of reaching retailers. However, we find a muted effect of distribution costs on adoption decisions. On the output side of the market, we collect novel descriptive measures of intermediary behavior, in particular, the entry of output buying "agents." Allen and Atkin (2016) models a similar channel, where a perfectly competitive, heterogeneous group of traders travel from market to market exploiting all available arbitrage opportunities.[9] Different from the data used in their work, we measure intermediation directly at the level of the farmer - whether crops are sold, and if so, in what quantity and at what price. Indeed we have found an active supply network for maize that is run by intermediaries, we find that many farmers

the all-pervasive nature of the impacts generated by road construction. See, for example, Alder (2017), Khanna (2016), and Storeygard (2016).

[7]In the specific context of agricultural inputs, both Aggarwal (2017) and Shamdasani (2016) find evidence of increased input adoption in the wake of a pan-Indian rural road construction program. However, the impact documented by both of these papers is reduced form in nature and neither is able to establish either the impact of transport costs on decreasing adoption or the channels through which road construction encourages adoption.

[8]Specifically, to find the best travel-adjusted price for fertilizer, our results suggest that for the typical village, the best option is 10km away. In Atkin and Donaldson, ad-valorem estimates are calculated based on the cost difference of a trip to the most remote location (500 miles away) relative to the least remote location (50 miles away), which is approximately a 720km difference.

[9]In Allen and Atkin (2016), when a particular market has excess supply, less efficient intermediaries enter that "route" to exploit the new arbitrage opportunity. In their work, they use this model to quantify the role of revenue volatility in crop choice, and use a highway project in India to evaluate how crop choice affects the gains from integration.

are not served by them, and that distance to the nearest market and nearest town significantly reduces the probability of being served. Indeed, we also find that adoption is particularly sensitive to the trade costs incurred by output-buying intermediaries, especially in remote villages.

Finally, much of the trade literature, which has documented larger gains from integration when there are input-output relationships (e.g. Yi, 2001; Costinot and Rodriguez-Clare, 2014; Sotelo, 2016) has only evaluated economies under the assumption of monopolistically competitive or purely competitive sectors at a fairly aggregate level (e.g. international trade by industry).[10] By contrast, our model is based on a standard discrete-choice logit model in which farmers choose the best agrovet from which to purchase fertilizer. In contrast to the previous literature, such a model allows for reductions in transport costs to be absorbed in part by increased mark-ups by the retailer. Largely, we find that the mark-up channel has only a modest effect on adoption only when the source of the shock is to wholesale prices.[11]

The rest of this paper proceeds as follows. Section 2 provides background and context on our study region, and lays out the sampling strategy that was adopted for this project. Section 3 explains the data, and documents summary statistics about the various data-collection units. Section 4 presents our main results. We put our findings in the context of a spatial model, which is presented and calibrated in Section 5, and run policy counterfactuals. Section 6 concludes with a discussion.

2 Background and Sampling Strategy

2.1 Background on fertilizer and maize price dispersion

There has been a lot of academic and policy interest in understanding why fertilizer usage in Africa lags behind other regions like South Asia. The profitability of fertilizer can be written as $\frac{p_{output} * \Delta y}{p_{fertilizer}}$, where p is the price and Δy is the increase in yield. A lot of work has gone into studying Δy and in interventions to increase it, for example via agricultural extension services, the development of new input varieties, or by matching input choices to soil characteristics through soil testing. Of course, profitability is also similarly affected by changes in input and output market prices, but there has been less work on this issue (which is the focus of this paper).

In this subsection, we assemble a dataset of prices of a variety of commodities and farming inputs, and present descriptive evidence on dispersion. We do this using 5 secondary datasets

[10] Our work is closely related to Sotelo (2016) develops a model of regional trade in agriculture and road quality in Peru to study the impact of road and output shocks on regional welfare and crop choice. Our work differs in its focus on local intermediaries and how their presence affects the landscape of market access.

[11] Yi (2001) provides an influential take on the role of vertical relationships in the growth of vertical trade that is germane to our work. Intuitively, if inputs are traded from one country to another, and then final goods are traded back to the origin country, the role of distance is amplified by the multiple stages of production. That is, since borders must be crossed more than once, the costs of distance are amplified by the number of times the good crosses the border prior to consumption. Our field work has identified that economy in rural Tanzania is similar to this setting, where inputs are sourced from larger cities, and output, if sold at all, is trade back to these same cities.

which include a total of 1,512 locations[12] in 56 countries [13], as well as price data we collected in Northern Tanzania between March and August 2016 with 251 retailers of various sorts (shops, agro-input dealers, and maize traders) in 82 markets. Results are presented in Table 1. To quantify price dispersion, we first run the following regression

$$log(p_{mcjt}) = \gamma_c + \gamma_j + \gamma_t + \epsilon_{mjt} \tag{1}$$

where p_{mjt} (log) prices in market m for product j at time t in country c, and the γ terms are country, product, and time fixed effects. As a measure of spatial price variation, we calculate the standard deviation of the resulting residual. In the secondary datasets, the standard deviation is 0.45 for all products, 0.34 for maize, and 0.12 for fertilizer; in our Tanzania data, the figures are 0.22, 0.14, and 0.09. The somewhat lower standard deviation in our data could be indicative of reduced measurement error, or that prices vary less within the geographic concentrated area of Kilimanjaro we focus on.

Second, we run dyadic regressions to look at price gaps. These regressions are motivated by an assumption of free entry, which implies that an arbitrageur will enter a market to arbitrage prices between markets m and m' if $\mid (p_m - p_{m'}) \mid \geq c_{mm'}$, where p is the price and $c_{mm'}$is the cost of transport between markets. Following Engel and Rogers (1996) and other papers[14], this free entry condition motivates the following regression:

$$log(\mid p_{mjt} - p_{m'jt} \mid) = \theta log(c_{mm'}) + \gamma_m + \gamma_{m'} + \gamma_j + \epsilon_{mm'jt} \tag{2}$$

For each dyad, we regress the absolute difference in log prices on two measures of distance: (1) kilometers between locations in Columns 1, 4, and 7, and (2) driving time between locations in Columns 2, 5, and 8 (both calculated via Google maps). We cluster standard errors by both the destination and origin market (Cameron, Gelbach and Miller 2012). In each of the datasets for which we have constructed driving times, we find significant, positive coefficients, suggesting that price gaps are larger between more distant markets. The coefficients are economically meaningful: a doubling of travel costs would increase price gaps by about 1-3% in the secondary datasets. In Tanzania, we find that doubling distances would increase price gaps by a similar amount.

Finally, we can use this data to provide some descriptive evidence on road upgrading. We conjecture that price gaps should respond to the time it takes to travel from point to point, and not the geographic distance (since the time and other costs of traveling to sell items should be what

[12]These are not necessarily all unique locations. Though we have cleaned these datasets, there are some misspellings, different names for the same markets, and also differing levels of granularity in the datasets.

[13]We include the following datasets: (1) prices of 6 staple crops in 41 major market centers in 8 East African countries from 1997-2015, collected by RATIN; (2) prices of 25 commodities from 276 markets in 53 countries in from 2013-2015, collected by Africafoodprices.io; (3) prices of 4 major varieties of fertilizer (Urea, DAP, CAN, and NPK complex 17-17-17) in 129 markets in 7 East African countries collected by AMITSA; (4) prices of 5 major varieties of fertilizer (Urea, CAN, DAP, and NPK 17 17 17) in 18 countries from 2010-16 in Africafertilizer.org; and (5) prices of a number of commodities in 38 countries from 1992-2016 collected by the WFP.

[14]In particular, see papers on the effect of cell phones on price dispersion, for example Aker (2010), Aker and Fafchamps (2015), and Jensen (2007).

is important). To examine this, we regress price gaps on both distance and duration in Columns 3, 6, and 9. Consistent with priors, we find that duration is significant, whereas distance is not – which suggests that improving road quality would reduce these gaps.

2.2 Background on fertilizer market and Kilimanjaro region

This primary study took place in the Kilimanjaro region[15] of Northern Tanzania. There are 570 villages in the region, and according to the 2012 census of Tanzania, the total population of the area is 1.6 million, about three-quarters of which is rural (National Bureau of Statistics, 2013). Our data collection covers the entire area of the region, a substantial area of 13,250 square-kilometers, roughly equivalent to the state of Connecticut in the USA or the country of Montenegro. Within Tanzania, Kilimanjaro is a relatively prosperous region, and agricultural productivity is relatively high. Using data provided by the 2012-13 wave of the National Panel Survey, we find that farmers in Kilimanjaro reported maize yields that are about 30 percent higher than the national average. Roads within Kilimanjaro are also marginally better than in Tanzania on the whole - according to numbers reported by the government of Tanzania, the paved trunk road density in Kilimanjaro is 2.2 percent of the total land area in the region (i.e. there are 2.2 kilometers of roads per 100 square kilometers of area), as opposed to only 0.7 percent for Tanzania as a whole. The density of the total network of trunk and regional roads is 7.4 percent in Kilimanjaro, but only half as much (3.7 percent) for the entire country of Tanzania (TanRoads and PMO-RALG, 2014).[16] The relative density of other minor roads is likely similar, although these numbers are harder to obtain. From an objective standpoint however, the road network in Kilimanjaro is quite poor. For instance, the density of the road network in the United States is 68 percent; the OECD average is 134 percent.[17]

Kilimanjaro has two growing seasons: a longer, more productive "long rains" season, which runs from March to June, and a less productive "short rains" season from October to January. Input usage tends to be much higher in the long rains, and some farmers decide not to plant in the short rains at all. Our main outcomes are based on behavior in the long rains.

We worked off of the list of villages included in the documents pertaining to the 2012 census of Tanzania, and did data-collection in the universe of villages (570 villages) listed as being in the Kilimanjaro region. As discussed in more detail below, we conducted surveys in a subset of villages, and did a comprehensive census of agro-input retailers in the entire region.

Virtually all fertilizer is imported in Tanzania. While some developing countries (such as India) produce chemical fertilizer domestically, production capacity is virtually non-existent in sub-Saharan Africa, and therefore, many sub-Saharan African nations import the entirety of their fertilizer requirements (FAOSTAT Online database, 2016; Hernandez and Torero, 2011).[18] As a

[15] Tanzania has 31 regions in all, including 5 in Zanzibar.

[16] The Roads Act, 2007 (No. 13 of 2007) defines a a trunk road as one that is primarily (i) a national route that links two or more regional headquarters or (ii) an international through route that links regional headquarters and another major or important city or town or major port outside Tanzania. A regional road is a secondary national road that connects (i) a trunk and district or regional headquarters; (ii) a regional headquarters and district headquarters.

[17] Information compiled from various online resources.

[18] Tanzania has some limited domestic production capacity in the form of an Arusha-based company called Minjingu

result, for these countries, transport costs from port to farm will directly affect prices. At present, we do not document the costs from port to distributor, but collect costs from that point on. In particular, we focus on the rural costs of intermediation and the costs at which farmers acquire inputs from retailers. To our knowledge, documenting the latter costs is entirely novel within the literature.

2.3 Sampling Strategy

The goal of this project is to construct a dataset representative of the entire region of Kilimanjaro. The main categories of data we set out to measure were: (1) surveys of farmers, fertilizer retailers, and maize buying agents; (2) transportation costs; and (3) prices. We initially set out to measure prices of a variety of goods. However, many villagers do not purchase most of their goods in their local village, and instead travel to local markets which operate one or several days a week. We decided to use these markets as the unit at which we would measure prices.

Thus, to construct our sample, we first assigned every village in our sample to a market catchment area. This was done by visiting ward offices (the ward is the lowest administrative level in Tanzania) and asking the ward officer to list the market that people from that village frequented. We use this market information in two main ways. First, we randomly selected markets for inclusion in the price collection from this list. Second, it was not feasible to travel individually from every village to a particular point to measure transport costs. Instead, we measure transport costs, requiring routes to go through the market center – we measure distances from every village to its closest market, and from every market to the main road. A map of the villages in our sample is included as Figures 1 and 2.

The geography of Kilimanjaro region provides for a setting with potential wide variation in transportation costs to Moshi. Closest to Moshi are semi-urban and rural districts surround the city. While many villages may be located off main roads, their location is proximate to the main supply points in the region. In the northeastern part of Kilimanjaro, near the border with Kenya, many villages are by straight-line distance not far from Moshi, but the the presence of Mt. Kilimanjaro is complicating for travel. Further removed from Moshi are villages near the town of Same, itself connected to Moshi by the main trunk road within the region. However, even along this road, travel times are not trivial, and many villages are located within, or on the other side of, the Pare Mountains (to the northeast of Same). Overall, the region provides substantial geographic variation that we now document in terms of the costs of travel.

3 Data and summary statistics

We have four main sources of data we use in this draft: agrovet surveys, farmer surveys, transport surveys, and maize price surveys. All were collected from January 2016 to October 2017.

Mines and Fertilizer Ltd. Only a handful of retailers in our sample sell this brand of fertilizer, however.

3.1 Agrovet surveys

First, we conducted a census of all agrovets in the region, finding a total of 395 agrovets. Of these agrovets, 376 sell fertilizer, which will be our primary sample to conduct the detailed agrovet survey. We then revisited these agrovets to conduct a longer survey which took about 2 hours to complete. In total, 369 agrovets completed the survey. The survey asked questions about varieties of fertilizer sold, prices, quantities, and the wholesale costs of acquiring stock from the distributor. The survey took care to differentiate fertilizer types by distributor, brand, and type – thus the level of granularity should be akin to the barcode-level. The survey also included a number of questions about costs of travel to the distributor, as well as some background characteristics.

3.2 Farmer surveys

We conducted farmer surveys in a randomly sampled subset of 115 villages. Within a selected village, enumerators were instructed to first find a landmark.[19] Once the village center was identified, the enumerators randomly picked a direction to begin their fieldwork, and selected every third homestead, or the next homestead after five minutes of walking, whichever came first. Overall, we enrolled an average of 4.8 farmers in 115 villages. The survey itself included questions on input usage and prices, maize sales, harvest output, and related outcomes. The survey also included some household and demographic questions.[20]

3.3 Measuring transport costs

We measured transportation costs in several ways. First, we collected GPS location for every village in Kilimanjaro,[21] from which we calculated driving times and distances using the Google API (via the statistical program R). Second, we conducted surveys of transportation operators in every village in our sample, which were either motorbike taxis ("Boda Bodas"), or consumer van taxis ("Dala Dalas"). In each village, we asked up to 3 operators how much it cost to travel to the major towns in Kilimanjaro (Arusha and Moshi), the capital city (Dar es Salaam) and the market center.

Third, enumerators recorded information on road quality and travel times as part of their field work. There are several major paved roads in Kilimanjaro. While not up to developed country standards, these roads are better maintained and most are paved. They are typically 2 lane roads. To get to a village, it is typically necessary to turn off one of these main roads and then travel for some time on unpaved feeder roads and village roads. To measure travel times, field officers used the following protocol. On a GPS unit, they recorded the point at which they had to turn off the main road, and then recorded the travel time, distance, and road quality on the road to the market

[19] These landmarks included a primary/secondary school within the village (1st choice), local church within the village (2nd), Boda stand within the village (3rd).

[20] This sampling strategy was used primarily because of budget and time constraints. In ongoing field work, we instead sample farmers randomly from a list provided by local village officials.

[21] We cross-checked these GPS coordinates, and filled in a handful of missing values, using a dataset of postal geocodes from www.geopostcodes.com.

center associated with the village. Once reaching the market, enumerators took a second form of transportation to the village, recording again distance, travel time, and road quality. We use this data to correlate costs of travel with road quality, and to estimate the percentage of roads which are paved (to inform later counterfactuals).

3.4 Agent surveys, store surveys and logbooks

To measure market access on the output side, we collected several surveys. Farmers who sell maize will either do so locally (typically to a vendor who then sells to other consumers or, more rarely, the farmer may retail it directly) or to an intermediary known as a "maize agent." Agents visit villages just after harvest and offer to buy maize in bulk. Agents then organize transportation of the maize to other locations. The bulk of this maize is transported to the major local towns in the region (Arusha and Moshi) to be sold to large maize warehouses known as "stores." The largest stores have capacity for tens of thousands of bags of maize. Store owners sell maize in bulk to major buyers in other locations (as well as to local vendors and to consumers). For example, much of the maize in Kilimanjaro is transported north to Nairobi, Kenya.

Interviewing agents and stores is challenging, because there is no registry of these types of businesses and because agents are itinerant, moving from village to village. To construct a sample of agents and to get information on maize flows, we asked a selected subset of the largest maize stores to keep a logbook. Each store was given a bound book in which they were asked to record each major transaction, recording the price, quantity and the method of transportation. In order to collect information about agents, stores were asked to record the name and phone number of each agent that they purchased maize from, as well as the location from where they came. We also surveyed stores, asking questions about maize volumes, prices, and transportation costs, as well as some background and demographic questions.

Using the list of agents and their contact information from the store logbooks, we called agents to schedule appointments for a survey. This survey included questions similar to the stores, but also asked questions about where agents traveled to buy maize, what price they paid in different locations, methods of transportation, and storage of purchased maize.

While we intend to use this data in future iterations of this paper, we do not use results from these surveys in the current draft.

3.5 Price collection

This project was initially centered around collecting prices in rural markets, but then evolved towards rigorously documenting market access for farmers, and centering the project around the surveys listed above. However, we did collect price information for two periods of period of time. In the first wave, we enrolled 82 locations into a price-collection protocol, conducted from March-August 2016. To enroll participants, we visited each market and selected several types of retailers for project inclusion, including fertilizer retailers ("agrovets"), maize sellers, and retail shops. Each respondent was called once per month and asked about current retail and wholesale prices for each item in a pre-selected list of standardized goods (e.g., 200 ml box of Azam juice). Respondents were compensated for participation by mobile money transfer. We do not utilize this price data in this current draft but intend to do so in future iterations.

In the second wave of price collection, we visited markets post-harvest in September and October of 2017. During the visits, enumerators sampled maize sellers to document pre- and post-harvest prices for maize during recent seasons. This price collection is used below in measuring the post-harvest selling conditions for farmers.

3.6 Summary statistics

Summary statistics are provided in Table 3 for villages (Panel A) and roads (Panel B). The average village has 2,842 people, and is located 5.7 kilometers from the nearest market center. A round-trip to the market center takes about 40 minutes according to surveys (20 minutes according to Google maps), and costs about US $1.60. The average distance to the nearest major town of Moshi is about 65 km, and a round-trip there would take just under 3 hours and cost about $4.70. These travel costs are substantial for poor farmers making a few dollars a day.

There is also substantial variation in travel costs to these cities in the region, from towns just outside Moshi to remote villages in the mountains in Same District in the South of Kilimanjaro. The standard deviation of travel costs to Moshi is about 80% of the mean, while the minimum travel cost is about $0.30 and the maximum is $22. In this context, it is reasonable to consider counterfactuals of even very large increases in travel costs.

Panel B shows information on the quality of the rural roads connecting markets and villages. Roads are about 1/3 paved, 1/3 dirt, and 1/3 gravel, and travel times according to google are fairly slow: 30.6 km/hour on rural roads compared to 49.5 km/hr on the main roads.[22]

Table 4 presents summary statistics on farmers. Fifty-five percent of farmers use fertilizer, substantially higher than the national average reported in the Tanzania NPS. Conditional on using, farmers tend to use about 55 kilograms (close to the FAO recommendation for 1 acre of land – so much less than FAO recommendations given that the average farmer has 2.7 acres of land). Most farmers (83%) use improved seeds of some sort. A minority of farmers (38%) sell maize, and about half of these sales are to agents (the rest are local sales). Finally, agricultural productivity appears very low – the average yield is only 430 kg per acre. Total output is only 700 kg, worth only about $160 at average post-harvest prices. These yields are too low to survive on alone, so most farmers have other sources of income – average income from other sources is $300 per year.

[22]However, note from panel A that travel times on google, at least on rural roads, are about half the travel times experienced by enumerators.

4 Main results

4.1 Specification

From the above data sources, we are able to construct transportation costs to every village in our sample, using either survey transport costs or Google maps. Our main empirical specification then becomes:

$$m_{fvt} = \sum_{q} \theta_q q_{vt} + \epsilon_{fvt} \quad (3)$$

where m_{vt} is a measure of market access and q_{vt} is the quartile in measured costs from the regional hub city of Moshi (see Figures 1 and 2 for a map of the area, where Moshi is marked with a star). We choose to present the results in quartiles rather than a log-linear specification because travel costs in our sample are particularly relevant for the most remote regions. However, results are qualitatively very similar in log-linear form.

We focus on 3 core sets of results: (1) farmer outcomes measured from farmer surveys in 115 villages, including adoption, output, and interactions with maize-buying intermediaries ("agents"), shown in Table 5; (2) market access outcomes including input access measured in all 570 villages and village-level output access measured in the same 115 villages as above, shown in Table 6; and (3) "agrovet" outcomes for the 351 agrovets that we identified in the study who complete surveys, shown in Table 7.

4.2 Farmers

Table 5 shows farmer results. Panel A shows dramatically lower input usage in more remote areas: usage declines by 21-32 percentage points in the most remote quartiles, relative to the quartile closest to Moshi. On a base of 68% in the nearest quartile, these translate to percentage declines of 30-47%. Impacts on quantities are even larger, equivalent to about a 67% decline. Effects on improved seeds are qualitatively similar, though baseline usage rates are considerably higher. Interestingly, however, farmers who do purchase inputs travel no further in remote areas: the average farmer travels about 6 km to buy inputs though this is driven by a few large values. Figure 3 shows the CDF of distances traveled, showing that the median farmer buys inputs in her own village, and that only 20% travel more than 10 km. In any case, the lack of a correlation between incurred distance and remoteness implies that farmers who do buy in remote areas tend to buy locally.

Table 5 Panel B focuses on the output side. Farmers in remote areas are less likely to sell maize to any type of buyer. This is likely driven by reduced access to "agents" – maize-buying intermediaries who travel from major towns to buy maize from farmers, to be shipped to larger urban centers. Farmers in the most remote quartile are 20 percentage points less likely to have been visited by an agent, equivalent to a 2/3 reduction in percentage terms. On the other hand, farmers in remote areas are far more likely to buy maize: the probability of buying maize is roughly twice as high in remote areas.

Finally, Panel C examines productivity. The results in Panel A imply that yields should be considerably lower in rural areas, and that is exactly what we find: yield per acre is 50% lower in the most remote areas. While the specific magnitude of this number is very large (we will re-examine this with data currently being collected), the sign of the effect is consistent with expectations.

In Appendix Tables 1 and 2, we show results from the same specifications as in Panels A, B, and C of Table 5, but with the inclusion of controls. Specifically, we control for crop-suitability and potential production capacity, i.e. soil characteristics (from the FAO's data on Global Agro-Ecological Zones) in Appendix Table 1, and soil as well farmer characteristics (the latter taken from our own farmer surveys) in Appendix Table 2. While some of the results get attenuated, the effect sizes are qualitatively robust to the inclusion of these controls.

4.3 Market Access

Table 6 shows statistics on market access. Panel A focuses on input access. Since we visited *all* agro retailers and villages in the region, and have GPS coordinates for all of them, we have measures of input access for all villages (not just those in which farmers were surveyed). Somewhat counter-intuitively, we find that remote areas are *more* likely to have an agrovet: while only 20% of villages in the lowest cost quartile have an agrovet, this increases to 38% in the most remote region. The explanation for this is that, since travel costs are large, agrovets cannot sell to farmers in remote areas without locating near them – and evidently it is profitable to do this in at least some cases. However, heterogeneity here is key: while 38% of villages in remote areas have agrovets, the other 62% are likely very far away from retailers and costs are likely very high for them.

To shed light on this, we summarize access by calculating a travel cost-adjusted price of fertilizer for every village as follows:

$$r_v^{tc} = \min_j \{r_j + c_{jv}\} \tag{4}$$

where r_j is the price at agrovet j and c_{jv} is the cost of transporting a bag of fertilizer from agrovet j to village v. We assume that farmers are free to travel to any agrovet from which they want to purchase, but must incur a transportation cost (which we estimate by combining transport surveys and google distances). For now, we assume that a farmer must make 3 one-way trips to purchase the bag (a round-trip for herself plus one additional trip for the fertilizer itself – this is based on qualitative field reports), and we impute a cost of travel of 175 Tsh per km (about \$0.076). We are in the process of refining this measure with new surveys. We assume that only the monetary cost of travel matters, and that travel costs are fungible with pecuniary costs for fertilizer. We assume farmers buy a 50 kg bag of fertilizer (farmers buy on average less than this, so this will tend to understate travel costs). In future work, we aim to refine this (in particular because preliminary exploratory work strongly suggests that farmers behave as if travel costs are more expensive than the cost of fertilizer, and thus tend to travel less far than we would predict), but for now we argue that treating travel and fertilizer costs as fungible is likely a lower bound on true costs.

For maize prices, we adopt a similar approach, but instead construct the maximum travel cost-adjusted selling price for maize:

$$p_v^{tc} = \max_m \{p_m - c_{mv}\} \tag{5}$$

Here, p_m is the price of maize post-harvest for market m, and $c_{mv}$is the cost of traveling from village v to market m. We assume that farmers sell 240kgs of maize (i.e. 2 units of a 120kg bag). The cost of carrying the maize bags are calibrated from the farmer survey.

Figure 4 plots CDFs of village-level "best" prices of inputs and output, adjusting for travel costs, and show tremendous heterogeneity in prices across villages. In Panel A, for maize prices, we observe that farmers receiving the lowest price, receive less than half of what is received by the highest, and a full 40% of the sample receives 70% or less of this highest possible price. Panel B shows the distribution of fertilizer prices, and the maximum observed price is 200% of the minimum, and 20% of villages face prices 30% higher than the minimum. Panel C presents a stark implication of these sharp disparities by constructing a profitability index of the ratio of the best output price to the best input price for each village: for 20% of villages, fertilizer is only half as profitable as it is for those with the highest profitability index.

In Figure 5, we attempt to unpack the disparity in prices faced by our sample by correlating it with distance. We find that the villages located farthest away get about 10% lower prices for output, and pay a 40% higher price for fertilizer, translating into a 30% lower profitability index of fertilizer.

In regression form (Table 6, Panel A), we find that the average village in the 3rd and 4th quartiles face fertilizer prices 15% higher than in the closest quartile, a substantial difference. These villagers must also travel further to obtain good prices (implying that prices in the rural areas are higher in general – which we will show in the next table).

Finally, Panel B shows access to output markets. These measures are less crisp than the input side (since the maize buying agents are itinerant and their presence is hard to capture, except with farmer surveys). However, we find much lower access in remote areas: villages in the most remote quartile are 21 percentage points less likely to have *any* farmer report an agent visit in the surveys (36%). We find some weak evidence that prices are actually higher in remote areas (which would be consistent with lower supply of maize due to less production), though we hesitate to put too much weight on this here. For now we take this evidence as suggestive that farmers in remote areas are less likely to find a buyer for their maize (which is also consistent with the dramatic reduction in sales in Table 5).

4.4 Agrovets

Finally, Table 7 shows results for agrovets. These results deserve some caution in interpreting, since they are conditional on the decision to enter a particular market. First, Figure 6 shows histograms of retail prices, wholesale prices, and markups, showing quite a bit of variation in retail prices and markups but less on the wholesale price. Markups are also adjusted for the self-reported travel costs to distributors in the agrovet survey. The lack of variation in wholesale prices is because most agrovets pick up fertilizer from the wholesaler and transport it back themselves.

Turning to the regressions, we find little effect of distance on the types or quantities of fertilizer sold. However, consistent with the previous results, we find that remote retailers charge higher prices: for the most common variety of fertilizer (Urea) agrovets in the most remote quartiles have 7-12% higher prices. Looking across all fertilizer types (bottom of panel), the results are similar. This is due to a combination of higher markups and/or higher marginal costs of accessing inputs – most agrovets travel to town to buy fertilizer, so it is no surprise that wholesale prices do not vary much by remoteness. Once we adjust mark-ups for travel costs, the mark-up increment for Urea in the remotest quartile becomes statistically indistinguishable from zero, although it continues to be high in magnitude. For all fertilizers taken together, the mark-up effect persists even after the travel cost adjustment.

5 Model and Counterfactuals

Above, we have provided a number of mostly descriptive results that show a consistent lack of market access in remote regions. Specifically, farmers in remote regions tend to travel farther to reach an agrovet, get a good price for fertilizer, and are less likely to have an output buying intermediary visit their village. This evidence is consistent with a story in which reduced sales opportunities and costly input procurement lead to far lower levels of input adoption.

Of course, remote regions may also differ in other dimensions. They may be poorer, hold different levels of land, or possibly live in areas not suited for fertilizer. Thus, absent an experiment that varies market access, a model is required to account for other factors that may be affecting adoption, and then used to run counterfactuals over different transport costs.

Below, we combine techniques from empirical industrial organization and trade to develop a spatial model of agro-retailer pricing, farmer investment, and agent activity. To quantify the impact of transportation costs on farmers, using the data collected from farmers and agrovets, we evaluate counterfactuals related to transportation costs along the supply chain.

To begin, we outline a model of fertilizer adoption and agrovet pricing.

5.1 Farmers, Fertilizer Adoption, and Village Output

Suppose that there are I villages, indexed by i, with L_i workers/consumers who each own k_i units of land. On this land, all individuals farm maize, and may purchase fertilizer to improve farm productivity. Farmers maximize wealth, and they use this wealth to fund other consumption. In this subsection, we evaluate a model in which farmers choose whether or not to buy fertilizer to increase yields on their land, and if so, where they should be purchasing their fertilizer to improve productivity.

Effective land holdings are $\tilde{\theta}_i k_i$, where $\tilde{\theta}_i = \theta_i$ when fertilizer is used, and $\theta_i = 1$ otherwise. Conditional on their technology choice, farmers also hire l_i workers at local wage w_i. Output of the farm, y_i, is governed by a Cobb-Douglas function of labor and effective land, with weight β on labor and $1-\beta$ on land: $y_i = l_i^{\beta}\left(\tilde{\theta}_i k_i\right)^{1-\beta}$. If the farmer sells output, they can sell at price p_i.

Holding effective land fixed, farms maximize profits by optimally choosing labor. Doing so yields the following variable profit function:

$$\pi_i = \left((1-\beta)\beta^{\frac{\beta}{1-\beta}}\right) p_i^{\frac{1}{1-\beta}} w_i^{-\frac{\beta}{1-\beta}} k_i \theta_i \tag{6}$$

Here, we have assumed that fertilizer is used. If it isn't used, replace $\theta_i = 1$. To simplify analysis, we group terms and write variable profits as:

$$\pi_i = A_i \theta_i \tag{7}$$

where $A_i = \left((1-\beta)\beta^{\frac{\beta}{1-\beta}}\right) p_i^{\frac{1}{1-\beta}} w_i^{-\frac{\beta}{1-\beta}} k_i$. Lastly, variable profits are linked to labor costs and revenues, respectively, via $w_i l_i = \beta \pi_i$ and $v_i = \frac{1}{1-\beta}\pi_i$, which will be used extensively when deriving equilibrium conditions.

For farmers who buy fertilizer, suppose that they choose to buy fertilizer from agrovet j at a price r_j. To travel to and from agrovet j, the farmer must pay F_{ij}. Recognizing that there may be other economic reasons that a farmer may go to location j (other items available, near other stores, higher reliability), we assume that farmers face an agrovet-specific cost sensitivity, δ_j, to the delivered costs of fertilizer, $F_{ij} + r_j$. We also assume an idiosyncratic error ε_{ij} for location j by farmer i that is independent of other factors. Thus, for a farmer that chooses location j, wealth is written as:

$$W_{ij} = A_i \theta_i - F_{ij} - r_j \delta_j + \varepsilon_{ij} \tag{8}$$

The usefulness of having δ_j in the model will become apparent shortly. Essentially, it will act as a "residual" in the mark-up equation that allows us to perfectly match mark-ups, conditional on adoption decisions and imputed market shares.[23]

For those farmers who do not adopt, their wealth is a simple function of their endowment and idiosyncratic error. Precisely:

$$W_{i0} = A_i + \varepsilon_{i0} \tag{9}$$

To characterize the discrete choice model in terms of probabilities, we assume that all idiosyncratic ε's are distributed Gumbel; thus, the model yields the standard multinomial logit choice

[23] Typically in an empirical IO model, brand quality would be an additive term that is eventually backed-out of the data by contraction. If δ_j were additive, it would not effect the demand elasticity other than through the implicit effect on market shares. We have tested the model under that assumption and we cannot match observed mark-ups to the data. Thus, we use a multiplicative approach, which is similar to the location-specific scale in Cosar, Grieco, Li and Tintelnot (2017).

probabilities. Recognizing that only differences in wealth matter within multinomial logit, we can easily show that the probability farmer i chooses agrovet j is written as:

$$\lambda_{ij} = \frac{\exp\left(\alpha_i - F_{ij} - r_j \delta_j\right)}{1 + \sum_s \exp\left(\alpha_i - F_{is} - r_s \delta_s\right)} \tag{10}$$

where $\alpha_i = A_i(\theta_i - 1)$. A useful transformation of this for the empirical analysis is:

$$\lambda_{ij} = \frac{\exp\left(-\left(F_{ij} + r_j\right)\delta_j\right)}{\sum_s \exp\left(-F_{is} - r_s \delta_s\right)} \mu_i \tag{11}$$

where μ_i is the probability that farmer i adopts fertilizer, and is defined as:

$$\mu_i = \frac{\sum_s \exp\left(\alpha_i - \left(F_{is} + r_s\right)\delta_s\right)}{1 + \sum_s \exp\left(\alpha_i - F_{is} - r_s \delta_s\right)} \tag{12}$$

Moving forward, we must aggregate variable and total profits for the farmer to be used later in general equilibrium conditions. Expected variable profits (defined as not including investment costs) for the farmer are written as:

$$\mathbb{E}\pi_i = A_i \theta_i \mu_i + A_i(1 - \mu_i)$$

Rearranging, we have:

$$\mathbb{E}\pi_i = \alpha_i \mu_i + A_i \tag{13}$$

When including the costs of fertilizer and optimal choice of agrovet, as shown in Train (2003)[24], expected profits can be written as

$$\mathbb{E}W_i = \log\left(1 + \sum_s \exp\left(\alpha_i - \left(F_{is} + r_s\right)\delta_s\right)\right) + \Upsilon_i$$

where Υ_i is a constant from integration. Imposing the definition of farmer-level adoption, we have:

$$\mathbb{E}W_i = \log\left(\frac{1}{1 - \mu_i}\right) + \Upsilon_i \tag{14}$$

With the farmer's problem described, we now move to the pricing optimization for agrovets.

5.2 Agrovet Pricing

Keeping with the i index from above, we now assume that a farmer i is representative of the village, which has L_i similar farmers. Any characteristics for each village i will be averaged across surveyed farmers from that village.

[24] This result is originally derived in Williams (1977) and Small and Rosen (1981)

For an arbitrary agrovet j, profits are written as follows,

$$\Pi_j = \left(r_j - r_j^o - \tau_j^o\right) \sum_i L_i \lambda_{ij}$$

where r_j^o is the price at which agrovet j purchases fertilizer from a distributor, and τ_j^o is the transport cost in procuring fertilizer from the distributor. The simplicity of logit is apparent when deriving the effects of r_j on λ_{ij}.

$$\frac{d\lambda_{ij}}{dr_j} = -\delta_j \lambda_{ij} \left(1 - \lambda_{ij}\right)$$

Thus the first order conditions for agrovet j can be written as:

$$\frac{\partial \Pi_j}{\partial r_j} = \sum_i L_i \lambda_{ij} - \left(r_j - r_j^o - \tau_j^o\right) \sum_i L_i \delta_j \lambda_{ij} \left(1 - \lambda_{ij}\right) = 0$$

Rearranging, we can solve for the absolute mark-up, $r_j - r_j^o - \tau_j^o$:

$$r_j - r_j^o - \tau_j^o = \frac{\sum_i L_i \lambda_{ij}}{\delta_j \sum_i L_i \lambda_{ij}(1 - \lambda_{ij})} \tag{15}$$

In 15, we see that level mark-ups are a function of village sizes, the probability each village buys from j, and the agrovet-specific cost sensitivity δ_j.

5.3 Agents

Above, we characterized the decisions on the input-side of the supply chain for maize; that is, the procurement, and equilibrium pricing of, fertilizer. The primary issue for farmers was their proximity to agrovets, and for the agrovets, the distribution of demand and any nearby competitors. As described in earlier sections, there is also an active market for output intermediaries, or "agents", who travel to villages and purchase maize output to sell, usually in larger cities. These agents are important for understanding the overall affect of roads on fertilizer adoption. Intuitively, if agent activity increases, this pushes up the value of output, and hence, the incentives to pay the fixed costs of using fertilizer.

To formalize the agent's problem, we first assume the each agent has one unit of capacity, and that using this unit of capacity as many times as necessary, they enter all profitable markets. There are N total agents. The agent can sell a unit of maize back to the larger (unmodeled) market at a price p_a. This selling price is unobserved and will be absorbed in estimation by a constant. If an agent travels to village i, the agent meets a seller with probability κ, and if they meet, buys a unit of maize at price p_i. To travel to and from village i, the agent must pay a travel cost $c_0 + c_1 d_i$, where c_0 and c_1 are coefficients and d_i is the cost of traveling to village i.

All together, the (expected) profit function of the agent is written as:

$$\Pi_i^a = (p_a - p_i)\kappa - c_0 - c_1 d_i$$

Since we do not know every transaction price for maize, both at the village and at the destination market, we assume that they are measured with error around the sampled prices in each market. Thus, we can re-write expected profits as:

$$\Pi_i^a = (p_a - p_i)\kappa - c_0 - c_1 d_i + \varepsilon$$

If we assume that noise term ε is distributed Gumbel, then we get the standard logit probability for entering village i:

$$\Pr\left(\Pi_i^a > 0\right) = \frac{\exp\left((p_a - p_i)\kappa - c_0 - c_1 d_i\right)}{1 + \exp\left((p_a - p_i)\kappa - c_0 - c_1 d_i\right)}$$

Thus, the expected sales (revenues) to agents in village i is written as:

$$V_i^a = p_i \frac{\exp\left((p_a - p_i)\kappa - c_0 - c_1 d_i\right)}{1 + \exp\left((p_a - p_i)\kappa - c_0 - c_1 d_i\right)} \kappa N$$

To estimate this model, we note that since the selling price is unobserved, but assumed to be the same across all agents, we can write a linear specification within the probability of visiting village i,

$$V_i^a = p_i \frac{\exp\left(\beta_0 - \beta_1 p_i - \beta_2 d_i\right)}{1 + \exp\left(\beta_0 - \beta_1 p_i - \beta_2 d_i\right)} \kappa N \tag{16}$$

where β_1 and β_2 are defined positively to indicate the predictions of the model.

5.4 Market Clearing

To study the impact of transport shocks, we now define a market clearing condition for maize and labor used for maize production. So far, we have defined the output of maize for each farm, and also the demand for maize by maize agents. A final component of demand is local demand for consumption. Taking a simplified approach to this problem, we assume that local income is generated by farm profits (expected wealth), W_i, wages paid to workers, $w_i l_i$ through maize farming, and exogenous other income, I_i. This income is used for consumption of maize and an outside good, where the share of expenditures on maize is γ.

With local demand defined, the market clearing condition for maize (written as expenditures at local prices) is written as:

$$V_i^a + \gamma \left(\mathbb{E}W_i + w_i l_i + I_i \right) L_i = V_i^f L_i$$

Using the relationships that $w_i l_i = \frac{\beta}{1-\beta}\pi_i$ and $V_i^f = \frac{1}{1-\beta}\pi_i$, we can rearrange as the product market clearing condition as:

$$V_i^a + \gamma L_i \mathbb{E}W_i + \gamma I_i L_i = B\mathbb{E}\pi_i L_i \tag{17}$$

where $B = \frac{(1-\gamma\beta)}{1-\beta} > 0$.

For the labor market, we make the strong assumption that the mass of workers available in village i for maize farming is equal to village size.[25] Thus, labor payments from maize farming must equal the total value of labor demand for maize farming:

$$w_i L_i = \frac{\beta}{1-\beta} L_i \mathbb{E}\pi_i$$

Imposing the formula for $\mathbb{E}\pi_i$, we have:

$$w_i = \frac{\beta}{1-\beta}\alpha_i \left(\mu_i + \frac{1}{\theta_i - 1} \right) \tag{18}$$

We will use this equation to implicitly define the relationship between α_i and p_i. Using the formula for α_i, and rearranging, we can write:

$$\alpha_i = (1-\beta)\, p_i \left((\theta_i - 1) K_i \right)^{1-\beta} \left(\mu_i + \frac{1}{\theta_i - 1} \right)^{-\beta} \tag{19}$$

Since the agent's problem is defined by the p_i's, but the farmer's and agrovet's problem are defined directly by the α_i's, this equation will be important for implementing the counterfactuals. Key to this equation will be calibrating the θ_i's.

5.5 Calibration

To setup the model for counterfactuals, we will first calibrate the model in four steps. In step one, we will use J agrovet pricing equations and I adoption probabilities to solve for J δ_j's and I α_i's using these equations. In the second step, we'll use the labor market clearing condition to calibrate the link between α_i's and $p_i's$ using a choice for θ_i in each village. In the third step, we will use the formula for agent demand for local maize to estimate parameters for the agents problem. In the the final step, we will solve for remaining parameters as residual variation in the maize market

[25] One way to justify this assumption is to assume that planting and harvest periods are intertemporally specific, meaning that everybody works in maize farming during the important time of the year, but may have other income at other times of the year. This allows us to abstract from any substitution between farm and non-farm production, which while important, complicates the model considerably).

clearing conditions.

Step 1 - Agrovet Pricing and Farmer Adoption Since multinomial logit does not allow for zero probabilities for any option, we first need to smooth out the ones and zeros we have for adoption data to probabilities of adoption. For simplicity, we simply assign a 0.01 for surveyed villages without adoption, and 0.99 for surveyed villages with full adoption. These adoption probabilities will be shown as "baseline" in later figures. In future drafts, we will integrate an estimation of these probabilities simultaneously within the pricing equations.

With these adoption probabilities, we can proceed to backing out unobserved parameters in the model. First, we begin with δ's. Rearranging (15), and imposing the second definition of selecting agrovet j (from 20), we get:

$$\delta_j = \left(\frac{1}{r_j - r_j^o - \tau_j^o}\right) \frac{\sum_i L_i \frac{\exp(-F_{ij} - r_j\delta_j)}{\sum_s \exp(-F_{is} - r_s\delta_s)} \mu_i}{\sum_i L_i \frac{\exp(-F_{ij} - r_j\delta_j)}{\sum_s \exp(-F_{is} - r_s\delta_s)} \mu_i \left(1 - \frac{\exp(-F_{ij} - r_j\delta_j)}{\sum_s \exp(-F_{is} - r_s\delta_s)} \mu_i\right)} \tag{20}$$

There are J of these equations, and we solve these equations using a non-linear solver in R. As long as the data are scaled appropriately, the system converges to a solution very quickly.

Next, using the μ_i's and the δ_j's, we can simply solve for the unobserved α_i's, which again are interpreted as farmer-specific increase in profits from using a bag of fertilizer. To see how, note that the probability of fertilizer adoption is written as:

$$\mu_i = \frac{\exp(\alpha_i) \sum_s \exp(-F_{is} - r_s\delta_s)}{1 + \exp(\alpha_i) \sum_s \exp(-F_{is} - r_s\delta_s)} \tag{21}$$

Rearranging, we have:

$$\alpha_i = \log\left(\frac{\mu_i}{1 - \mu_i} \frac{1}{\sum_s \exp(-F_{is} - r_s\delta_s)}\right) \tag{22}$$

We now have solved for I α's and J δ's, exactly matching the observed mark-ups and imputed adoption probabilities.

Step 2 - Labor Market Clearing

As derived above in (19), by manipulating the labor market clearing condition for village i we get the following

$$\alpha_i = (1 - \beta)\, p_i\, ((\theta_i - 1)\, k_i)^{1-\beta} \left(\mu_i + \frac{1}{\theta_i - 1}\right)^{-\beta} \tag{23}$$

Here, there are a number of parameters that need to be estimated. First is β which is the share of labor in the farm's production function. Since we do not have a good estimate of this, we will simply assume it takes on a value of 1/2 for now. For the rest of the terms, we will use the village-level maize price and average village acreage to represent p_i and k_i. The μ_i's are adoption rates, and we'll use our farmer surveys for this. The remaining term is θ_i, which is the village-level benefit of fertilizer. For each village, we will solve the above non-linear equation to pin-down θ_i. This term can be interpreted in many ways, from the natural benefit of fertilizer based on soil quality and climate, to the village-level competence at using fertilizer. We will be agnostic about its precise interpretation, and instead treat it as a parameter to calibrate the model.

Step 3 - Parameters for the Agent's Model

From the agents problem total revenues from selling to agents in village i are again written as:

$$V_i^a = p_i \frac{\exp(\beta_0 - \beta_1 p_i - \beta_2 d_i)}{1 + \exp(\beta_0 - \beta_1 p_i - \beta_2 d_i)} \kappa N$$

To estimate the terms within the exponentials, we will use a logit model predicting the probability that an agent enters a particular market. To represent d_i we will use the costs of travel from Moshi to each village, as routed through the market relevant for village i. For the price, we use the village sales price, or the average sales price within the market catchment area if it is missing. After running the model, we will use the observed village prices, surveyed village sales, and the total size of the village to agents to calibrate κN.[26]

Step 4 - Unobserved terms in Market Clearing

Finally, using (13), (14) and (16), the market clearing conditions can be written as:

$$V_i^a + \gamma L_i \log\left(\frac{1}{1-\mu_i}\right) + \gamma L_i (\Upsilon_i + I_i) = B\alpha_i \left(\mu_i + \frac{1}{\theta - 1}\right) L_i \tag{24}$$

To calibrate this equation, we need to take stand on the share of expenditures that go to maize. In this case, we again simply assume that γ is 1/2 (like the labor share in production). For village size, L_i we take data from the 2012 Census of Tanzania. With these choices, and using the calibration and estimates from steps 1-3, we can uniquely pin-down the residual components of the product market clearing condition, $\Upsilon_i + I_i$.

5.6 Counterfactuals

We now use the model to evaluate the effects of transport costs on fertilizer adoption. The goal of this exercise is two-fold. First, we wish to isolate the effects of transportation costs on the reduced form results in section three. This should indicate whether the source of variation driving the results is due to transport or other unobserved factors that vary with distance. Second, we wish to

[26]For the latent variable logit model, the coefficient on output price is -8.95 (p = 0.02158) and the coefficient on cost of travel from Moshi to the village is -0.167 (p=0.055). Both prices and costs are divided by 2000 (approximately the conversion to USD) to match the scaling choice in the other components of the calibration.

evaluate how transport cost shocks at different points along the supply chain affect adoption. In particular, do reductions in trade costs have larger effects in particular areas, and if so, on which portions of the supply chain are these reductions most effective?

We first examine the role of overall trade costs on fertilizer adoption. After calibrating the model, we hold fixed all terms not related to transport costs, and then iteratively reduce transport costs on all parts of the supply chain by one percent from observed values to map out the effects on adoption. At each value of trade costs, we also estimate the reduced form relationship between remoteness (log of travel time to Moshi) and adoption using predicted adoption at these reduced trade costs. Along with adoption, the key statistic of interest is the share of the reduced form effect that is accounted for by this reduction in trade costs.

To begin, we evaluate an across-the-board (agent, distributor, and farmer) reduction in transport costs on adoption, as well the share of the reduced form effect that is captured by the reduction in transport costs. From the sample average of approximately 60%, by reducing transport costs by 1/2, which would bring transport costs more in-line with developed country standards, adoption increases to 73%, or about 16% above baseline.

To evaluate how much this increase in adoption accounts for the reduced form effects as described in section three, we compare a regression of prediction adoption on log distance to the regional hub (Moshi) with a regression of baseline adoption on the same distance measure. We hypothesize that as the reduction in transport costs is more pronounced, this will account for a larger share of the reduced form effect. Indeed, with a 50% reduction in transport costs, the relationship between predicted adoption and remoteness is 16% less pronounced. For completeness, we extend this analysis out to 100% reduction in transport costs, which accounts for 80% of the reduced form effect. While this latter large number is likely a result of the logit assumptions on demand, the mid-range estimate summarizes an economically meaningful effect of transport costs on adoption.

Lastly, we run counterfactuals that are specific to different areas of the supply chain. As a reminder, in the model, transportation costs are incurred from distributor to agrovet (for fertilizer), by the farmer in traveling from the village to each agrovet, and then by the agent in reaching villages to buy maize and return to the regional hub. To evaluate the responsiveness of each channel, we impose a small transportation shock on each part of the supply chain and report the elasticity of adoption to these shocks. Further, we summarize these effects within quartiles of remoteness from the regional hub. These results are presented in Table 8, where we report the elasticity of adoption to small changes in transport costs in three areas. "Local" transport shocks are those representing the costs of the farmer reaching each possible agrovet in the region. "Agent" represents the costs of the agent in traveling from the regional hub in Moshi to each village. "Distribution" represents the per-bag cost of each agrovet transporting fertilizer from the distributor. Finally, "All" represents changing transport costs on all aspects of the supply chain by the same small percentage.

The results in Table 8 suggest that the costs of transport induce a higher adoption response in remote regions. Focusing on Panel A (average elasticity across villages within each group), a comprehensive change in transport costs along all aspects of the supply chain produces an adoption elasticity to trade costs of -0.966. These effects are concentrated in the two most remote quartiles of remoteness. When evaluating each transportation shock separately, there is very little response of adoption to the distribution channel. However, there is a sizeable, negative elasticity of adoption to cost shocks to local transport costs and agent transport costs, with these shocks more pronounced in more remote regions.

6 Discussion

This paper has presented novel evidence on the availability of inputs for farmers in Kilimanjaro, Tanzania, and especially, the role the remoteness plays in access to productive technologies. Being more remote, as measured by transportation costs and distance from the regional hub in Moshi, is associated with lower adoption of fertilizer, a decreased likelihood of visits by output buying intermediaries, higher retail fertilizer prices, and higher "local" transport costs for farmers to incur in the process of purchasing fertilizer

As Kilimanjaro is a relatively prosperous region, this begs the question as to the role of remoteness in less developed regions. For a preview of how our results may extend to other regions and areas of Africa, in Table 9, we have assembled data from the World Bank LSMS-ISA household panel surveys for Ethiopia, Niger, Nigeria, Malawi, Tanzania, and Uganda, to study how remoteness affects fertilizer adoption. In the LSMS, measures of remoteness include distance to the main market, and distance to a population center. Using both measures of remoteness, we find a negative association between remoteness and technology adoption. However, since we cannot associate these adoption decisions with prices or precise measures of transport costs, we plan to continue our full field work in other regions of Tanzania. Further, we plan to do extensive work evaluating the sourcing decisions of output buying intermediaries, and how the presence of output buying intermediaries ultimately affects the decision by farmers to adopt fertilizer.

References

[1] Adukia, Anjali, Sam Asher, and Paul Novosad (2016). "Educational Investment Responses to Economic Opportunity: Evidence from Indian Road Construction." Unpublished

[2] Aggarwal, Shilpa (2017). "Do Rural Roads Create Pathways out of Poverty? Evidence from India." Unpublished.

[3] Alder, Simon (2017). "Chinese Roads in India: The Effect of Transport Infrastructure on Economic Development." Unpublished.

[4] Allen, Treb and Costas Arkolakis (2016). "The Welfare Effects of Transportation Infrastructure Improvements." Unpublished.

[5] Asher, Sam and Paul Novosad (2016). "Market Access and Structural Transformation: Evidence from Rural Roads in India." Unpublished.

[6] Asturias, Jose, Manuel Garcia-Santana, and Roberto Ramos (2017). "Competition and the welfare gains from transportation infrastructure: Evidence from the Golden Quadrilateral of India." Unpublished.

[7] Atkin, David and David Donaldson (2015). "Who's Getting Globalized? The Size and Implications of Intra-National Trade Costs." NBER Working Paper No. 21439

[8] Bandiera, Oriana and Imran Rasul (2006). "Social networks and technology adoption in northern Mozambique." *Economic Journal* 116 (514): 869-902.

[9] Banerjee, Abhijit, Esther Duflo, and Nancy Qian (2012). "On the Road: Access to Transportation Infrastructure and Economic Growth in China," Unpublished.

[10] Bardhan, Pranab and Dilip Mookherjee (2011). "Subsidized Farm Input Programs and Agricultural Performance: A Farm-Level Analysis of West Bengal's Green Revolution, 1982-1995." *American Economic Journal: Applied Economics*, 3(4): 186-214.

[11] Baum-Snow, Nate (2007). "Did Highways Cause Suburbanization?" *Quarterly Journal of Economics*, 122, 775–805.

[12] Beaman, Lori, Dean Karlan, Bram Thuysbaert, and Christopher Udry (2013). "Profitability of Fertilizer: Experimental Evidence from Female Rice Farmers in Mali." *American Economic Review* 103 (3): 381-86.

[13] Bergquist, Lauren (2017). "Pass-Through, Competition, and Entry in Agricultural Markets: Experimental Evidence from Kenya." Unpublished.

[14] Berry, Steven (1994). "Estimating Discrete-Choice Models of Product Differentiation." *RAND Journal of Economics* 25 (2): 242-262.

[15] Bird, Julia and Stephane Straub (2014). "The Brasilia experiment: road access and the spatial pattern of long-term local development in Brazil." Unpublished.

[16] Cagley, Jessica, Mary Kay Gugerty, and Robert Plotnick (2009). "Political Economy of Fertilizer Policy in Tanzania." Prepared for the Farmer Productivity Team of the Bill and Melinda Gates Foundation.

[17] Casaburi, Lorenzo and Tristan Reed (2017). "Competition in Agricultural Markets: An Experimental Approach." Unpublished

[18] Conley, Timothy and Christopher Udry (2010). "Learning about a new technology: Pineapple in Ghana." *American Economic Review* 100 (1): 35-69.

[19] Costinot, Arnaud and Andres Rodriguez-Clare (2014). "Trade Theory with Numbers: Quantifying the Consequences of Globalization." In *Handbook of International Economic*s, Volume 4, Chapter 4. Editors: Gita Gopinath, Elhanan Helpman, and Kenneth Rogoff. Elsevier.

[20] Dhingra, Swati and Silvana Tenreyro (2017). "Piggy-Back Exporting, Intermediation, and the Distributional Gains from Trade in Agricultural Markets." Unpublished.

[21] Donaldson, David (2015). "The Gains from Market Integration." *Annual Review of Economics* 7: 619-647.

[22] Donaldson, David (Forthcoming) "Railroads of the Raj: Estimating the impact of transportation infrastructure." *American Economic Review*, Forthcoming

[23] Duflo, Esther, Michael Kremer, and Jonathan Robinson (2008). "How high are rates of return to fertilizer? Evidence from field experiments in Kenya." *American Economic Review* 98 (2): 482-488.

[24] Duflo, Esther, Michael Kremer, and Jonathan Robinson (2011). "Nudging farmers to use fertilizer: theory and experimental evidence from Kenya." *American Economic Review* 101 (6): 2350-2390.

[25] Eaton Jonathan and Samuel Kortum (2002). "Technology, Geography, and Trade." *Econometrica* 70: 1741–1779

[26] Emerick, Kyle (2017). "The Cost of Favoritism in Network-based Markets." Unpublished.

[27] Foster, Andrew and Mark Rosenzweig (1995). "Learning by Doing and Learning from Others: Human Capital and Technical Change in Agriculture." *The Journal of Political Economy* 103(6): 1176-1209

[28] Foster, Andrew and Mark Rosenzweig (2010). "Microeconomics of Technology Adoption." *Annual Review of Economics* 2: 395-424.

[29] Gertler, Paul, Tadeja Gracner, Marco Gonzalez-Navarro, and Alex Rothenberg (2014). "Road Quality and Local Economic Activity: Evidence from Indonesia's Highways." Unpublished.

[30] Ghani, Ejaz, Arti G. Goswami, and William R. Kerr (2016). "Highway to success: The impact of the Golden Quadrilateral project for the location and performance of Indian manufacturing." *The Economic Journal* 126 (591): 317-357.

[31] Gollin, Doug, David Lagakos, and Michael Waugh (2014a). "The Agricultural Productivity Gap." *Quarterly Journal of Economics* 129 (2): 939-993.

[32] Gollin, Doug, David Lagakos, and Michael Waugh (2014b). "Agricultural Productivity Differences across Countries." *American Economic Review* 104 (5): 165-170.

[33] Gollin, Doug and Richard Rogerson (2014). "Productivity, Transport Costs, and Subsistence Agriculture." *Journal of Development Economics* 107: 38-48.

[34] Hanna, Rema, Sendhil Mullainathan, and Josh Schwartzstein (2014). "Learning through noticing: Theory and evidence from a field experiment." *Quarterly Journal of Economics*, 129(3), 1311-1353.

[35] Hernandez, Manuel and Maximo Torero (2011). "Fertilizer Market Situation. Market Structure, Consumption and Trade Patterns, and Pricing Behavior." International Food Policy Research Institute (IFPRI), Washington, DC, USA.

[36] Jack, B. Kelsey (2013). "Market Inefficiencies and the Adoption of Agricultural Technologies in Developing Countries." White Paper. Agricultural Technology Adoption Initiative, JPAL (MIT) and CEGA (UC Berkeley)

[37] Karlan, Dean, Robert Osei, Isaac Osei-Akoto, and Christopher Udry (2015). "Agricultural Decisions after Relaxing Credit and Risk Constraints." *Quarterly Journal of Economics*, 129(2), 597-652.

[38] Keller, Wolfgang and Carol Shiue (2007). "Markets in China and Europe on the Eve of the Industrial Revolution." *American Economic Review* 97 (4): 1189–1216.

[39] Khanna, Gaurav (2016). "The Road Oft Taken: The Route to Spatial Development." Unpublished.

[40] Maitra, Pushkar, Sandip Mitra, Dilip Mookherjee, Alberto Motta, and Sujata Visaria (2017). "Financing Smallholder Agriculture: An Experiment with Agent-Intermediated Microloans in India." *Journal of Development Economics* 127: 306 – 337.

[41] Morten, Melanie and Jaqueline Oliveira (2016). "The Effects of Roads on Trade and Migration: Evidence from a Planned Capital City." NBER Working Paper 22158.

[42] Redding, Stephen, and Daniel Sturm. (2008). "The Costs of Remoteness: Evidence from German Division and Reunification." *American Economic Review* 98 (5): 1766-97.

[43] Rothenberg, Alex (2013). "Transport Infrastructure and Firm Location Choice in Equilibrium: Evidence from Indonesia's Highways." Unpublished.

Sotelo, Sebastian (2016). "Domestic Trade Frictions and Agriculture", working paper.

[44] Stewart, W.M., D.W. Dibb, A.E. Johnston, and T.J. Smyth (2005). "The contribution of commercial fertilizer nutrients to food production." *Agronomy Journal* 97 (1): 1-6

[45] Storeygard, Adam (2016). "Farther on down the Road: Transport costs, trade and urban growth in Sub-Saharan Africa." *Review of Economic Studies* 83 (3): 1263-1295.

[46] Suri, Tavneet (2011). "Selection and Comparative Advantage in Technology Adoption." *Econometrica* 79 (1): 159-209

[47] Udry, Christopher and Santosh Anagol (2006). "The Return to Capital in Ghana." *American Economic Review* 96 (2): 388-393.

[48] World Bank (2010). "World Development Report 2008: Agriculture for Development." Washington, DC: World Bank.

[49] World Bank (2010). "Africa's Infrastructure: A time for Transformation." Africa Development Forum. Washington, DC.

[50] World Bank (2017). "Enabling the Business of Agriculture 2017." Washington, DC: World Bank. doi:10.1596/978-1-4648-1021-3. License: Creative Commons Attribution CC BY 3.0 IGO

[51] Yi, Kei-Mu (2003). "Can Vertical Specialization Explain the Growth of World Trade?" *Journal of Political Economy* 111 (1): 52-102.

Table 1. Input and output market price dispersion across countries

	(1) Secondary Datasets[1]	(2) Tanzania Data[2]
Residual standard deviation in log prices for:[3]		
All products	0.45	0.22
Maize only	0.34	0.14
Fertilizer only	0.12	0.09

Notes:

[1]Datasets include RATIN (prices of major crops across 41 major markets in 5 countries - Kenya, Tanzania, Uganda, Burundi, and Rwanda - over the 1997-2015 time period), Africafoodprices.io (25 products over 276 markets in 53 countries), AMITSA (the Regional Agricultural Input Market Information and Transparency System for East and Southern Africa, which includes information on 9 fertilizer varieties in 95 markets in 8 countries), prices of 5 major varieties of fertilizer (Urea, CAN, DAP, and NPK 17 17 17) in 18 countries from 2010-16 in Africafertilizer.org; and prices of a number of commodities in 38 countries from 1992-2016 collected by the WFP.

[2]Maize prices are from a survey of market sellers in 98 markets conducted in October 2017. Fertilizer prices are from surveys of agro-input retailers in 2017.

[3]Calculated from a regression of log prices on product, country, and time fixed effects.

Table 2. Dyadic price dispersion

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
	Dependent variable: Absolute log price difference								
Panel A. Secondary Datasets									
Log (distance)	0.03***		0.000	0.03***		0.000	0.01***		0.010
	(0.002)		(0.010)	(0.002)		(0.015)	(0.002)		(0.014)
Log (travel time)		0.03***	0.03***		0.04***	0.04**		0.01***	0.000
		(0.002)	(0.011)		(0.003)	(0.017)		(0.002)	(0.016)
Products	All	All	All	Maize	Maize	Maize	Fertilizer	Fertilizer	Fertilizer
Dependent variable mean	0.21	0.21	0.21	0.20	0.20	0.20	0.11	0.11	0.11
Dependent variable sd	0.20	0.20	0.20	0.17	0.17	0.17	0.13	0.13	0.13
Observations	4,752,196	4,752,196	4,752,196	675,880	675,880	675,880	38,364	38,364	38,364
Number of locations	1335	1335	1335	1335	1335	1335	1335	1335	1335
Countries	49	49	49	43	43	43	18	18	18
Panel B. Northern Tanzania									
Log (distance)	0.01***		-0.030	0.03***		-0.10**	0.003*		0.007
	(0.003)		(0.020)	(0.011)		(0.050)	(0.002)		(0.017)
Log (travel time)		0.01***	0.04*		0.04***	0.16**		0.004	-0.004
		(0.004)	(0.025)		(0.016)	(0.069)		(0.002)	(0.019)
Products	All	All	All	Maize	Maize	Maize	Fertilizer	Fertilizer	Fertilizer
Dependent variable mean	0.16	0.16	0.16	0.21	0.21	0.21	0.13	0.13	0.13
Dependent variable sd	0.14	0.14	0.14	0.18	0.18	0.18	0.10	0.10	0.10
Observations	22,386	22,376	22,376	6,873	6,873	6,873	15,064	15,056	15,056
Number of locations	82	82	82	65	65	65	60	60	60

Notes: Regressions include product, month and year fixed effects. All regressions are within country. Travel time and distances calculated from Google maps. See Table 1 and text for discussion of datasets.

Two-way clustered standard errors in parentheses. ***, **, and * indicate significance at 1%, 5%, and 10%.

Table 3. Summary statistics on villages in Kilimanjaro region

Panel A. Villages (N = 570)	
Population	2842.15
	(1882.22)
Distance to nearest market center (km) - Google	5.69
	(9.10)
Time for round-trip journey to nearest market center (mins) - Google	21.28
	(32.48)
Time for round-trip journey to nearest market center - surveys	40.43
	(33.03)
Cost of round-trip journey to nearest market center (USD) - surveys	1.59
	(1.94)
Distance to Moshi (km) - Google	65.76
	(52.52)
Round-trip travel time to Moshi (mins) - Google	177.23
	(117.92)
Round-trip cost of travel to Moshi USD - surveys	4.69
	(3.76)
Panel B. Road Quality (N = 570)	
Measurement of roads in field	
Percent of road that is:	
Paved	0.27
Dirt	0.35
Gravel	0.37
Cost of trip from market center to village (paid by enumerator)	0.91
	(1.19)
Google estimates	
Travel speed on major roads - km/hr (Google)	49.5
Travel speed on feeder roads and rural roads - km/hr (Google)	30.6

Notes: Standard deviations in parentheses.

Table 4. Summary statistics on farmers

Farmer characteristics	
Age	49.82
	(13.99)
Female	0.51
Married	0.78
	(0.41)
Years of education	7.36
	(2.54)
Home has thatch roof	0.09
	(0.29)
Walls of home are mud	0.25
	(0.44)
Floors of home are mud	0.39
	(0.49)
Has cell phone	0.87
	(0.34)
Has bank account	0.20
	(0.40)
Has mobile money account	0.81
	(0.39)
Acres of land	2.63
	(2.90)
Household size	5.19
	(2.18)
Has market business	0.22
	(0.41)
Annual total income from non-farming (USD)	305.35
	(797.05)
Input usage, sales and output	
Used chemical fertilizer in 2015 long rains season	0.55
If yes, quantity used	55.62
	(70.47)
Used improved seeds in 2015 long rains season	0.83
If yes, quantity used	8.79
	(8.85)
Sold maize after 2015 long rains season	0.38
Sold maize to an agent	0.19
Acres of land	2.63
	(2.90)
Total harvest output in 2016 long rains (kg)	698.75
	(625.30)
Output per acre	432.76
	(440.97)
Value of harvest output (USD, at mean 2015 long rains prices)	157.18
	(140.65)

Notes: Standard deviations in parentheses.

Table 5. Relationship between remoteness and fertilizer adoption, maize sales, and harvest output

	(1)	(2)	(3)	(4)	(5)
	Mean of dependent variable	Regression by quartile			
		Mean in lowest-cost quartile	Regression coefficient for:[1]		
			2nd	3rd	Highest cost
Panel A. Input usage					
Used chemical fertilizer in 2015 long rains	0.55	0.68	-0.01	-0.32***	-0.21***
			(0.06)	(0.06)	(0.06)
Quantity of chemical fertilizer used	27.38	44.65	-11.88*	-29.58***	-29.43***
	(56.94)	(73.80)	(6.48)	(6.76)	(6.70)
Used improved seeds in 2015 long rains	0.82	0.90	-0.02	-0.14***	-0.13***
			(0.05)	(0.05)	(0.05)
Quantity of improved seeds used	6.49	7.91	-0.47	-2.18**	-3.25***
	(8.56)	(10.40)	(0.99)	(1.03)	(1.02)
If used inputs, distance traveled to agrovet (km)	5.94	6.05	-1.48	-1.18	2.40
	(13.49)	(8.94)	(2.44)	(2.79)	(2.63)
Panel B. Output markets					
Sold maize after 2016 long rains	0.38	0.47	-0.01	-0.14**	-0.21***
			(0.06)	(0.06)	(0.06)
Quantity sold (kg)	248.20	351.60	-59.44	-145.34***	-220.35***
	(435.70)	(512.80)	(49.91)	(52.04)	(51.54)
Agent visited homestead	0.24	0.31	-0.01	-0.06	-0.20***
			(0.05)	(0.05)	(0.05)
Number of agents visited	0.60	0.71	0.11	-0.07	-0.52***
	(1.32)	(1.25)	(0.15)	(0.16)	(0.16)
Sold maize to an agent after 2016 long rains	0.19	0.21	0.05	-0.04	-0.12**
			(0.05)	(0.05)	(0.05)
Quantity sold to agents (kg)	127.80	150.20	23.18	-25.48	-94.62**
	(324.40)	(360.90)	(37.50)	(39.10)	(38.73)
Farmer ever buys maize	0.39	0.26	0.05	0.23***	0.26***
			(0.06)	(0.06)	(0.06)
Quantity purchased in typical year	127.80	61.67	24.97	107.41***	141.96***
	(244.40)	(162.30)	(27.71)	(28.89)	(28.61)
Panel C. Productivity					
Total harvest output in 2016 long rains (kg)	696.40	874.90	-130.34*	-219.02***	-379.34***
	(623.90)	(668.50)	(71.03)	(74.06)	(73.35)
Harvest output per acre in 2016 long rains	432.30	644.30	-220.09***	-310.52***	-327.08***
	(442.10)	(514.10)	(49.42)	(51.63)	(51.03)
Value of harvest output at average regional price	156.70	196.80	-29.32*	-49.27***	-85.33***
	(140.30)	(150.40)	(15.98)	(16.66)	(16.50)

Notes: N = 563 farmers in 115 villages. In Column 1, standard deviations are in parentheses. In Columns 2-4, standard errors in parentheses. *, **, and *** indicate significance at 10%, 5%, and 1%.

Table 6. Relationship between remoteness and village-level market access

	(1)	(2)	(3)	(4)	(5)
	Mean of dependent variable	Regression by quartile			
		Mean in lowest-cost quartile	Regression coefficient for:[1]		
			2nd	3rd	Highest cost
Panel A. Access to input markets (N = 570)					
Has an agrovet in village	0.26	0.20	0.04	0.04	0.18***
			(0.05)	(0.05)	(0.05)
Has at least 1 agrovet within 5 km of village	0.70	0.74	-0.04	-0.10*	-0.04
			(0.06)	(0.06)	(0.06)
Has at least 1 agrovet within 10 km of village	0.87	0.93	-0.05	-0.07	-0.11***
			(0.04)	(0.04)	(0.04)
Number of agrovets within 5 km of village	4.04	4.69	-0.44	-1.21**	-0.98*
	(4.73)	(4.21)	(0.58)	(0.57)	(0.57)
Number of agrovets within 10 km of village	8.51	8.90	0.95	-0.27	-2.20*
	(9.51)	(6.02)	(1.16)	(1.15)	(1.15)
Distance to nearest agrovet	3.96	3.38	0.28	1.04	1.02
	(5.31)	(3.64)	(0.65)	(0.64)	(0.64)
Minimum travel-cost adjusted price	21.77	20.07	1.25***	2.71***	2.90***
for 50 kg of Urea[1]	(3.07)	(1.67)	(0.35)	(0.34)	(0.34)
Distance to obtain minimum travel-cost	9.47	7.02	0.77	5.15***	3.94***
adjusted price (km)	(10.83)	(6.59)	(1.31)	(1.29)	(1.29)
Cost of travel to obtain minimum travel-cost	2.16	1.60	0.17	1.17***	0.90***
adjusted price (USD)	(2.47)	(1.50)	(0.30)	(0.29)	(0.29)
Panel B. Access to output markets (N = 109)					
At least one agent visited village	0.55	0.59	0.08	0.06	-0.21*
			(0.14)	(0.14)	(0.12)
Average number of agents visiting farmers	0.66	0.87	-0.14	0.06	-0.58**
	(0.92)	(1.15)	(0.25)	(0.26)	(0.22)
Average village output price (per kilogram)	0.23	0.21	0.02	0.03	0.03**
	(0.06)	(0.05)	(0.02)	(0.02)	(0.02)
Distance to the nearest market (km)	3.60	2.01	0.69	1.77***	3.93***
	(5.24)	(3.04)	(0.62)	(0.61)	(0.61)
Maximum travel-cost adjusted price	0.34	0.33	0.02**	0.05***	0.00
for 240kg of maize (USD)[2]	(0.05)	(0.03)	(0.01)	(0.01)	(0.01)

Notes: Panel A comes from a listing exercise of every agrovet and village in the region and so includes all 570 villages in the Kilimanjaro region. Panel B is obtained from farmer surveys which were conducted in a randomly selected subset of 109 villages. In Column 1, standard deviations are in parentheses. In Columns 2-4, standard errors in parentheses. Lowest quartile is omitted from regressions. *, **, and *** indicate significance at 10%, 5%, and 1%.

[1]Travel costs imputed from transport surveys and Google maps. We assume farmers buy a 50 kg bag in one trip (enough for 1 acre), the modal amount observed in our data, and must incur the cost of 3 trips to the retailer (a round-trip for herself, plus a trip for the bag of fertilizer).

[2]Travel costs imputed from transport surveys and Google maps. We assume farmers sell two 120 kg bags in one trip.

Table 7. Relationship between remoteness and fertilizer retailer sales, prices, and other characteristics

	(1)	(2)	(3)	(4)	(5)
	Mean of dependent variable	Regression by quartile			
		Mean in lowest-cost quartile	Regression coefficient for:[1]		
			2nd	3rd	Highest cost
Panel A. Varieties					
Sells Urea fertilizer	0.98	0.99	-0.02	-0.02	-0.01
			(0.02)	(0.02)	(0.02)
Sells DAP fertilizer	0.62	0.80	-0.07	-0.01	0.05
			(0.07)	(0.07)	(0.08)
Sells CAN fertilizer	0.15	0.35	-0.13**	-0.04	-0.02
			(0.05)	(0.05)	(0.06)
Sells NPK fertilizer	0.12	0.28	-0.05	-0.14***	0.07
			(0.05)	(0.05)	(0.05)
Sells other types of fertilizer[1]	0.12	0.28	-0.05	-0.14***	0.07
			(0.05)	(0.05)	(0.05)
Bags of Urea sold	105.40	343.60	-97.33**	-68.62*	-24.55
	(270.70)	(540.80)	(40.62)	(41.00)	(41.84)
Total bags of fertilizer sold	182.10	620.20	-170.30**	-115.49	-58.38
	(465.00)	(946.50)	(69.79)	(70.44)	(71.89)
Panel B. Prices and markups					
Most common fertilizer (Urea)					
Retail price for 50 kilograms, Urea only	23.73	21.93	0.51	2.67***	1.52***
	(3.27)	(2.41)	(0.56)	(0.55)	(0.55)
Wholesale price for 50 kilograms, Urea	20.01	19.78	0.00	0.40*	0.12
	(1.37)	(1.13)	(0.25)	(0.24)	(0.24)
Markup for Urea	0.19	0.11	0.01	0.09***	0.07***
	(0.14)	(0.09)	(0.03)	(0.02)	(0.02)
Markup (travel-cost adjusted) for Urea	0.15	0.092	0.01	0.07***	0.03
	(0.11)	(0.09)	(0.02)	(0.02)	(0.02)
All fertilizers[2]					
Retail price for 50 kilograms, all types	25.37	23.86	0.49	2.08***	1.69***
	(5.53)	(5.34)	(0.36)	(0.52)	(0.49)
Wholesale price for 50 kilograms, all types	21.64	20.76	0.33*	0.54**	0.34
	(4.37)	(4.32)	(0.19)	(0.24)	(0.24)
Markup, all types	0.18	0.14	0.01	0.07***	0.06***
	(0.13)	(0.09)	(0.01)	(0.02)	(0.02)
Markup (travel-cost adjusted), all types	0.14	0.12	0.01	0.05***	0.04**
	(0.11)	(0.09)	(0.01)	(0.02)	(0.02)

Notes: N = 351. In Column 1, standard deviations are in parentheses. In Columns 2-4, standard errors in parentheses. *, **, and *** indicate significance at 10%, 5%, and 1%.

[1]Other types of fertilizer include local varieties SA, Yara, and Minjingu.

[2]Regressions include fixed effects for brand and type of fertilizer.

Table 8. Counterfactuals - Elasticities of Fertilizer Adoption to Transport Shocks

	(1)	(2)	(3)	(4)	(5)
		Village Remoteness Quartiles			
	Overall	1st	2nd	3rd	4th
Panel A. Mean Elasticities					
Elasticity with respect to changes in:					
All transport costs simultaneously	-0.966	-0.409	-0.844	-1.328	-1.255
Farmers' transport costs to agrovets only	-0.522	-0.458	-0.399	-0.75	-0.503
Agents' transport costs from Moshi to villages only	-0.275	0.068	0.113	-0.523	-0.697
Transport costs from the distributor to the retailer	-0.023	-0.011	-0.007	-0.028	-0.043
Panel B. Median Elasticities					
Elasticity with respect to changes in:					
All transport costs simultaneously	-0.494	-0.242	-0.24	-0.591	-1.02
Farmers' transport costs to agrovets only	-0.272	-0.325	-0.212	-0.286	-0.269
Agents' transport costs from Moshi to villages only	-0.129	0.038	0.002	-0.327	-0.735
Transport costs from the distributor to the retailer	-0.008	-0.004	-0.003	-0.016	-0.026

Table 9. Adoption in LSMS-ISA surveys

	(1)	(2)
	Dependent variable: used chemical fertilizer	
Distance to nearest major market (km)	-0.027***	
	(0.005)	
Distance to nearest population center (km)		-0.019*
		(0.010)
Dependent variable mean	0.32	0.32
Independent variable mean	3.23	3.21
Independent variable sd	1.27	1.02
Observations	35,938	35,938
Individuals	26,653	26,653

Notes: Regressions include World Bank LSMS-ISA household panel surveys in Ethiopia, Niger, Nigeria, Malawi, Tanzania, and Uganda. Standard errors clustered at the enumeration area level are in parentheses. ***, **, and * indicate significance at 1%, 5%, and 10%.

Figure 1. Map of Survey Region and Villages

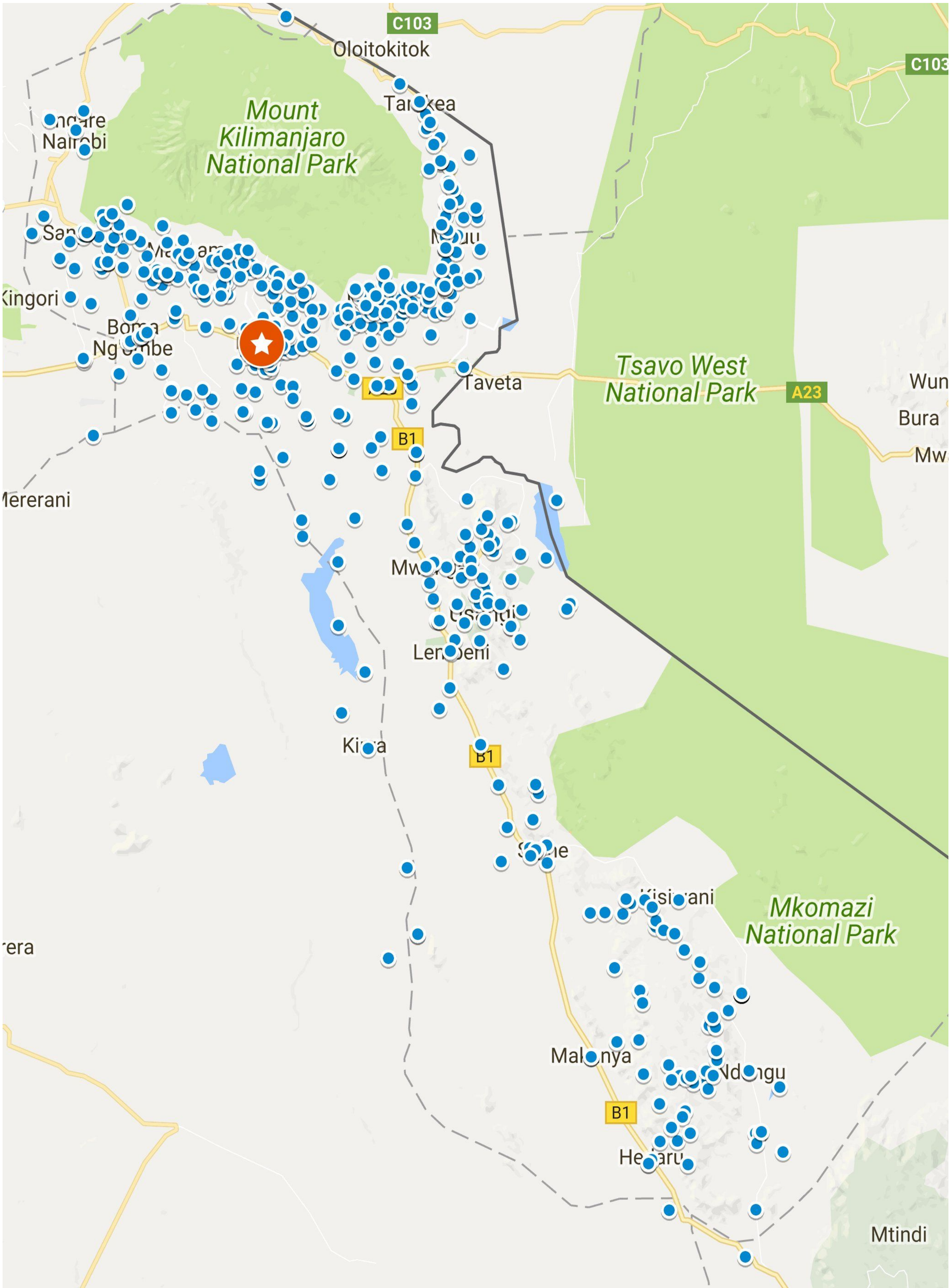


Figure 2. Map of Survey Region and Villages

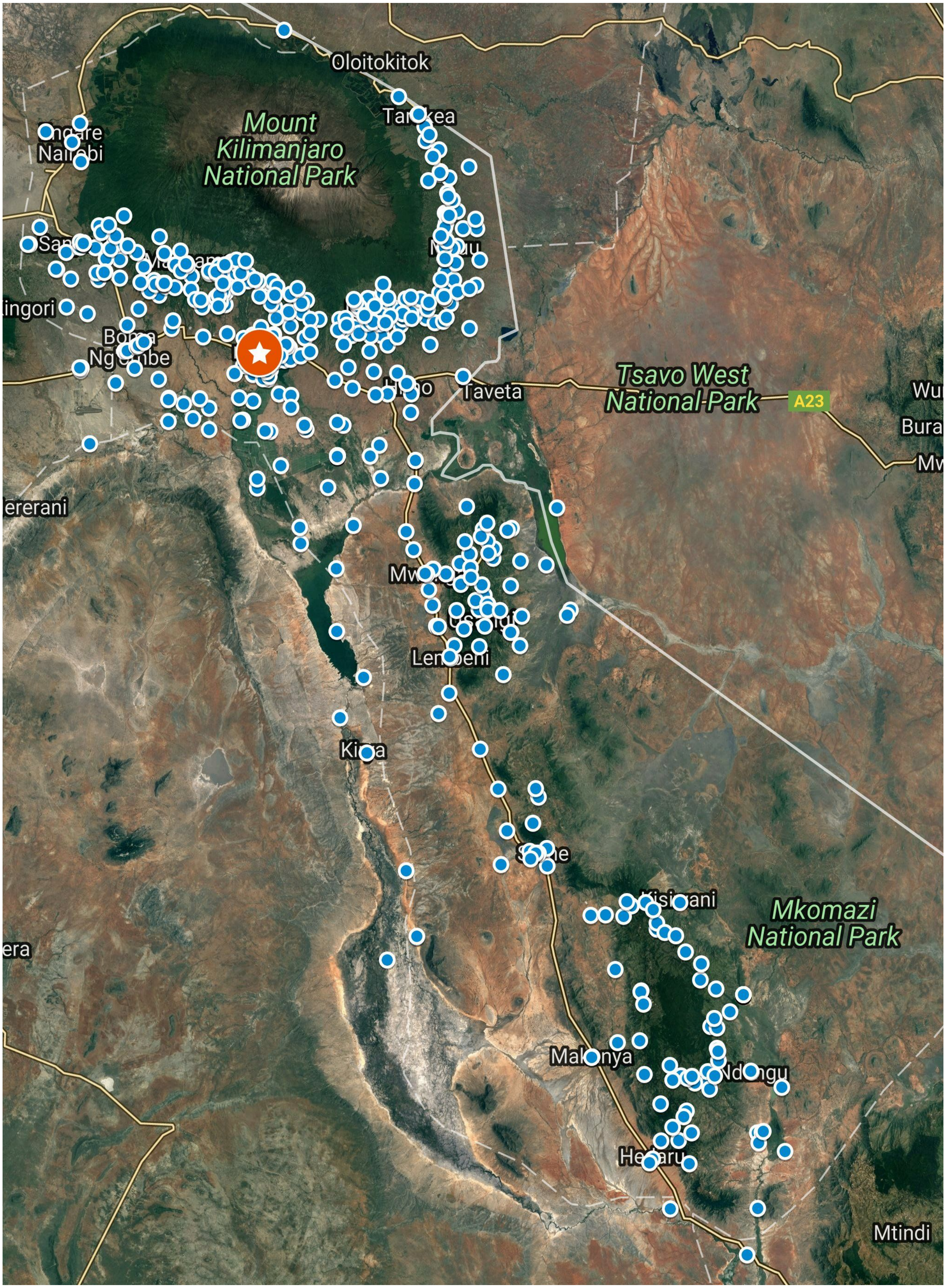
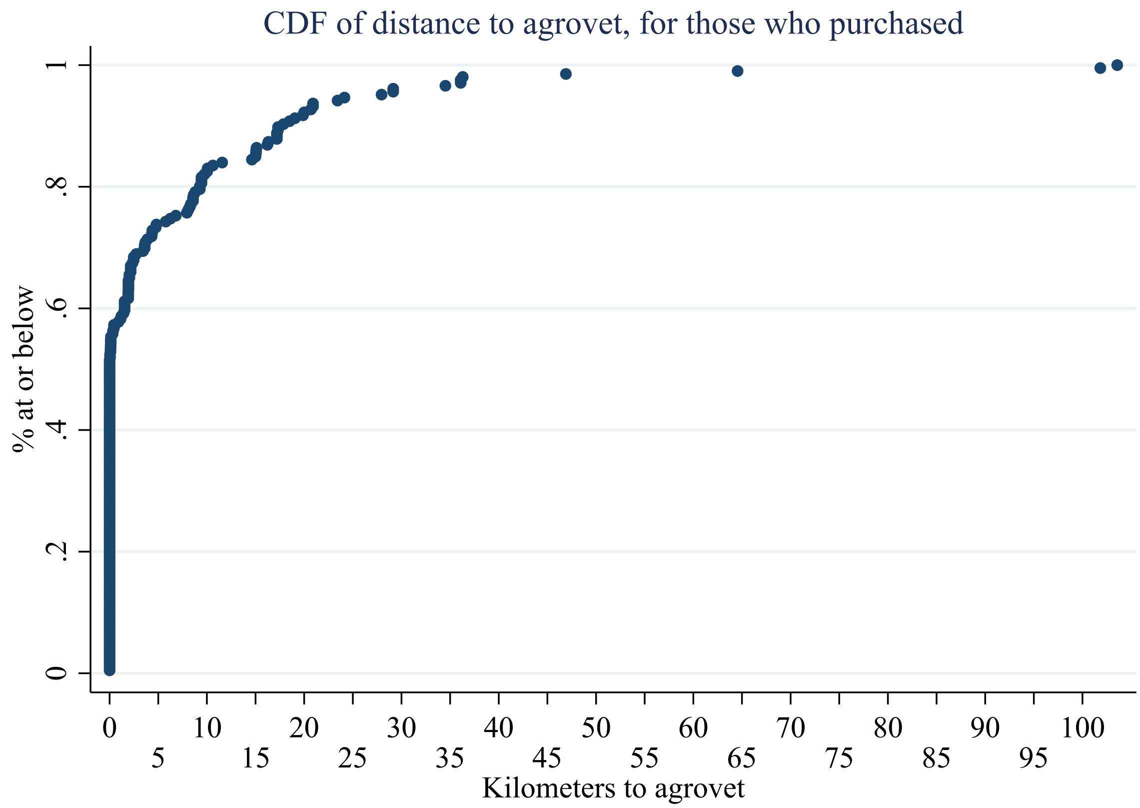
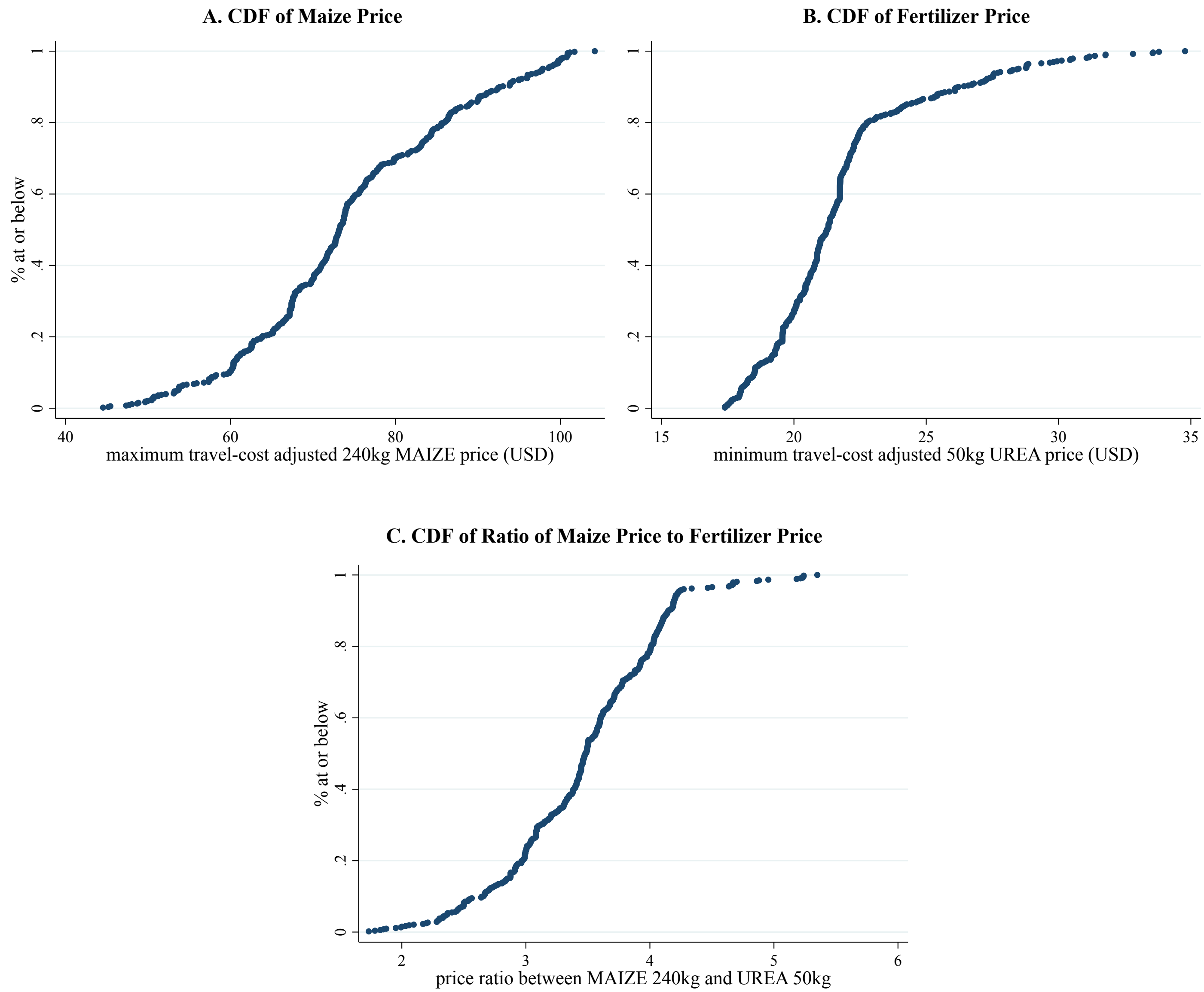


Figure 3. CDF of distance farmers travel to purchase inputs



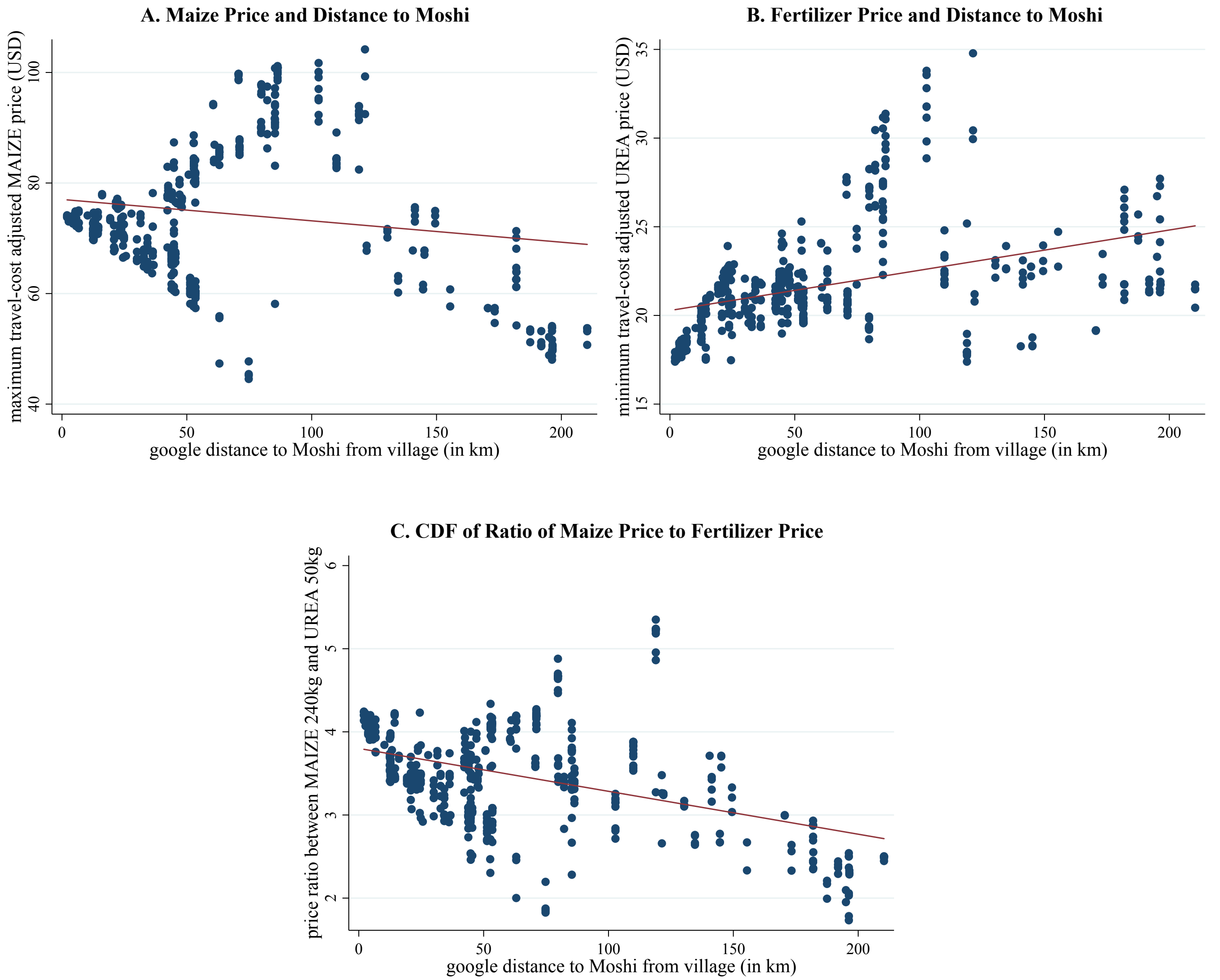
Notes: Each point represents a farmer.

Figure 4. CDF of travel-cost adjusted prices across villages



Notes: Each observation represents a village. Travel-cost adjusted prices are calculated through observed prices from an agrovet survey, a maize price survey at markets and transport cost information collected from interviews with transport operators.

Figure 5. Relationship between prices and distance to Moshi



Notes: Each observation represents a village. Travel-cost adjusted prices are calculated through observed prices from an agrovet survey, a maize price survey at markets and transport cost information collected from interviews with transport operators.

Figure 6. Retail prices, wholesale prices, and markups for Urea

Notes: Prices are for a 50kg bag of fertilizer. Markup is computed by "[retail price / (wholesale price + travel cost)] - 1."

Appendix Table 1. Relationship between remoteness and fertilizer adoption, maize sales, and harvest output

	(1)	(2)	(3)	(4)	(5)
		Regression by quartile			
	Mean of dependent variable		Regression coefficient for:[1]		
		Mean in lowest-cost quartile	2nd	3rd	Highest cost
Panel A. Input usage					
Used chemical fertilizer in 2015 long rains	0.55	0.68	-0.03	-0.25***	-0.16**
			(0.06)	(0.07)	(0.07)
Quantity of chemical fertilizer used	27.38	44.65	-12.48*	-26.36***	-26.24***
	(56.94)	(73.80)	(6.60)	(7.97)	(8.31)
Used improved seeds in 2015 long rains	0.82	0.90	-0.03	-0.15***	-0.15**
			(0.05)	(0.06)	(0.06)
Quantity of improved seeds used	6.49	7.91	-0.45	-2.93**	-3.66***
	(8.56)	(10.40)	(1.00)	(1.20)	(1.25)
If used inputs, distance traveled to agrovet (km)	5.94	6.05	0.53	-1.71	2.85
	(13.49)	(8.94)	(2.49)	(3.45)	(3.31)
Panel B. Output markets					
Sold maize after 2016 long rains	0.38	0.47	0.00	-0.04	-0.11
			(0.06)	(0.07)	(0.07)
Quantity sold (kg)	248.20	351.60	-44.35	-78.23	-135.97**
	(435.70)	(512.80)	(50.13)	(60.50)	(63.10)
Agent visited homestead	0.24	0.31	0.01	0.02	-0.10
			(0.05)	(0.06)	(0.06)
Number of agents visited	0.60	0.71	0.14	0.07	-0.33*
	(1.32)	(1.25)	(0.16)	(0.19)	(0.19)
Sold maize to an agent after 2016 long rains	0.19	0.21	0.06	0.03	-0.03
			(0.05)	(0.06)	(0.06)
Quantity sold to agents (kg)	127.80	150.20	33.66	4.87	-45.13
	(324.40)	(360.90)	(37.70)	(45.50)	(47.46)
Farmer ever buys maize	0.39	0.26	0.05	0.16**	0.21***
			(0.06)	(0.07)	(0.07)
Quantity purchased in typical year	127.80	61.67	28.74	106.56***	154.48***
	(244.40)	(162.30)	(27.96)	(33.75)	(35.20)
Panel C. Productivity					
Total harvest output in 2016 long rains (kg)	696.40	874.90	-112.43	-149.35*	-281.30***
	(623.90)	(668.50)	(70.93)	(85.61)	(89.29)
Harvest output per acre in 2016 long rains	432.30	644.30	-204.65***	-216.19***	-224.21***
	(442.10)	(514.10)	(49.48)	(59.81)	(62.23)
Value of harvest output at average regional price	156.70	196.80	-25.29	-33.59*	-63.28***
	(140.30)	(150.40)	(15.96)	(19.26)	(20.08)

Notes: N = 563 farmers in 115 villages. In Column 1, standard deviations are in parentheses. In Columns 2-4, standard errors in parentheses. *, **, and *** indicate significance at 10%, 5%, and 1%. All specifications include crop suitability index and potential production capacity obtained from FAO-GAEZ.

Appendix Table 2. Relationship between remoteness and fertilizer adoption, maize sales, and harvest output

	(1)	(2)	(3)	(4)	(5)
	Mean of dependent variable	Regression by quartile			
		Mean in lowest-cost quartile	Regression coefficient for:[1]		
			2nd	3rd	Highest cost
Panel A. Input usage					
Used chemical fertilizer in 2015 long rains	0.55	0.68	0.02	-0.18**	-0.08
			(0.06)	(0.07)	(0.07)
Quantity of chemical fertilizer used	27.38	44.65	-11.62*	-22.52***	-21.04**
	(56.94)	(73.80)	(6.68)	(8.21)	(8.64)
Used improved seeds in 2015 long rains	0.82	0.90	-0.01	-0.11*	-0.09
			(0.05)	(0.06)	(0.06)
Quantity of improved seeds used	6.49	7.91	-0.73	-3.28***	-3.57***
	(8.56)	(10.40)	(0.95)	(1.17)	(1.23)
If used inputs, distance traveled to agrovet (km)	5.94	6.05	0.89	-2.40	3.13
	(13.49)	(8.94)	(2.65)	(3.88)	(3.65)
Panel B. Output markets					
Sold maize after 2016 long rains	0.38	0.47	0.01	0.00	-0.06
			(0.06)	(0.07)	(0.07)
Quantity sold (kg)	248.20	351.60	-42.30	-53.13	-102.54
	(435.70)	(512.80)	(49.29)	(60.59)	(63.74)
Agent visited homestead	0.24	0.31	0.01	0.02	-0.07
			(0.05)	(0.06)	(0.06)
Number of agents visited	0.60	0.71	0.12	0.06	-0.29
	(1.32)	(1.25)	(0.16)	(0.19)	(0.20)
Sold maize to an agent after 2016 long rains	0.19	0.21	0.06	0.03	-0.03
			(0.05)	(0.06)	(0.06)
Quantity sold to agents (kg)	127.80	150.20	33.33	2.85	-46.09
	(324.40)	(360.90)	(36.91)	(45.37)	(47.72)
Farmer ever buys maize	0.39	0.26	0.00	0.10	0.13*
			(0.06)	(0.07)	(0.07)
Quantity purchased in typical year	127.80	61.67	9.41	80.33**	125.71***
	(244.40)	(162.30)	(27.54)	(33.85)	(35.61)
Panel C. Productivity					
Total harvest output in 2016 long rains (kg)	696.40	874.90	-108.84	-134.47	-256.97***
	(623.90)	(668.50)	(67.86)	(83.41)	(87.74)
Harvest output per acre in 2016 long rains	432.30	644.30	-160.53***	-177.88***	-176.29***
	(442.10)	(514.10)	(48.84)	(59.99)	(63.08)
Value of harvest output at average	156.70	196.80	-24.48	-30.25	-57.80***
regional price	(140.30)	(150.40)	(15.26)	(18.76)	(19.74)

Notes: N = 563 farmers in 115 villages. In Column 1, standard deviations are in parentheses. In Columns 2-4, standard errors in parentheses. *, **, and *** indicate significance at 10%, 5%, and 1%. All specifications include crop suitability index, potential production capacity obtained from FAO-GAEZ, and farmer-level controls such as wall and roof materials, assets and income.