An Analysis of Household Characteristics Impacting Food Security and Market Participation in Rural Uganda

by

Micaela M. Wiseman

Bachelor of Science in Electrical Science and Engineering,
Massachusettes Institute of Technology (2013)

Submitted to the Institute for Data, Systems, and Society
in partial fulfillment of the requirements for the degree of

Master of Science in Technology and Policy

at the

MASSACHUSETTS INSTITUTE OF TECHNOLOGY

June 2019

Author_____

Technology and Policy Program
May 16, 2019

Certified by_____

Jarrod Goentzel
Director, Humanitarian Supply Chain Lab
Thesis Supervisor

Accepted by_____

Noelle Eckley Selin
Director, Technology and Policy Program
Associate Professor, Institute for Data, Systems, and Society and Department of Earth, Atmospheric and Planetary Sciences

An Analysis of Household Characteristics Impacting Food Security and Market Participation in Rural Uganda

by

Micaela M. Wiseman

Submitted to the Institute for Data, Systems, and Society
on May 16, 2019, in partial fulfillment of the
requirements for the degree of
Master of Science in Technology and Policy

Abstract

According to the most recent Ugandan census, 40% of rural Ugandan households experience food insecurity and 25% live below the country's poverty line. Of the rural households, 80% are already engaged in some form of agriculture. Since market participation through selling crops is known to decrease both poverty and food insecurity, agricultural marketing presents one of the best economic opportunities for improving Ugandans' livelihoods. The USAID Feed the Future initiative focuses on designing programming to encourage market participation among Ugandan farming households by strengthening the market environment. In this thesis, we identify the household characteristics which are linked to enabling or inhibiting participation in the market in the Ugandan context in order to guide the design of USAID's activities there. We accomplish this by building three different regression models: one around the decision to participate in markets and two using different measures for the level of market participation. We use data from a recent household survey of Ugandan farmers covering a wide range of household characteristics. We received responses from 498 farming households that were randomly selected across 5 regions. We contextualize our results by discussing how the significant household characteristics may be considered when developing market facilitation activities. We also specifically analyze how food security is connected to facets of agriculture and market participation to explore how vulnerable populations may or may not be included in market facilitation efforts. Results from the model of the decision to participate in markets suggest that it is encouraged by the the size of nearby towns, number of available services, and membership in producer organizations, while transportation access had no impact. Results from both models for the level of market participation show that it is impacted by transportation access. Level of food insecurity does not impact market participation, indicating that the vulnerable food insecure population will be included in market facilitation programs. Level of agricultural production had a large impact the decision to participate in markets, the level of participation in markets, and food security. Market facilitation efforts should focus on developing producer organizations, encouraging local businesses which provide agricultural services, and strengthening input supply chains to improve farmers' production capacity.

Thesis Supervisor: Jarrod Goentzel
Title: Director, Humanitarian Supply Chain Lab

Acknowledgments

My research at MIT has been generously supported by the United States Agency for International Development's Mission to Uganda through the USAID/Uganda Feed the Future Market System Monitoring Activity (contract number AID-OAA-A-12-00095). Opinions expressed in this publication do not necessarily reflect the views of the United States Agency for International Development, the United States Government, or MIT. However, I am deeply grateful to USAID/Uganda Feed the Future staff who made this work possible.

Many thanks are due to my advisor for this thesis, Dr. Jarrod Goentzel. I've appreciated his advice and input throughout the entire process. The time he has spent meeting with me, reviewing my drafts, discussing my ideas, and encouraging my intellectual engagement were invaluable in creating this work. I've met few people who can match his level of unwavering optimism and can-do spirit, and I am grateful to have had him in my corner.

In completing this thesis and working with the MSM Uganda team, I owe an incredible debt of gratitude to the inimitable Courtney Blair. From across continents and oceans, over the course of hundreds of hours of phone calls, her patience, insight, and dedication have held me to a standard of excellence for the past two years. I'm thankful to her for not letting me get away with simply declaring something "interesting," but pushing me to articulate why an observation is unexpected, unique, or worthy of discussion.

I'm grateful to Justin Boutilier for his invaluable feedback and consultation on regression methodology. Many thanks go to the rest of the MSM Uganda team and Humanitarian Supply Chain Lab for the discussions and suggestions which have shaped my thinking about this project.

The people of Uganda have my gratitude for their warmth and openness. This analysis would in no way be possible without the willing participation of hundreds of farming families. In addition, Enoch Alioni, Jimmy Kigozi, and Jackson Isiko assisted me greatly in my time spent in Uganda.

The support and encouragement from my fellow TPP classmates has buoyed my spirit throughout this process. I'm thankful for the dear friendship of Veronica Foreman, Katie Picchione, Aurora Edington, and Sruthi Davuluri. Through thick and thin, my life-long friends Sydney Beasley, Lauren Plummer, and Vicki Alfieri have listened to me, comforted me, and believed in me, and I am lucky to have them at my side.

Finally, I wouldn't have been able to complete this graduate program without the unwavering love and support of my parents and family. Mom and dad, thank you for the force of your steadfast belief in me and the many ways in which you've taken care of me. I am forever grateful to have you in my life.

THIS PAGE INTENTIONALLY LEFT BLANK

Contents

THIS PAGE INTENTIONALLY LEFT BLANK

Chapter 1

Introduction

1.1 Agriculture in Uganda

Situated directly on the equator in East Africa, the Republic of Uganda is a vibrant country with immense opportunity for agricultural transformation. The livelihood of many Ugandans depends to some degree on agriculture, whether they farm for subsistence, as their main source of income, or as a supplementary business. According to the most recent Ugandan census, almost 80% of households participate in agriculture in one form or another, and the agricultural sector is responsible for 24% of the nation's GDP (Uganda Bureau of Statistics, 2016). The high engagement in agriculture is in part a result of the country's favorable climate and fertile soil. Temperatures remain warm year round. The southern part of the country experiences two rainy seasons, allowing for two harvests per year. The FAO estimates that Uganda, which is 38% arable land, has the potential to feed its entire population and export food to other nations (FAO, 2018; FAPDA, 2015). The country's main agricultural exports include coffee, cotton, and tea while its top crops for consumption are maize, plantains (matooke), and cassava (FAPDA, 2015).

Even as the nation is ecologically poised to provide a bounty, many people living there experience poverty, hunger, and malnutrition. The 2016-17 Uganda National Household Survey estimates that 10 million Ugandans (21% of the population) live below the national poverty line. This represents an increase of 2 percentage points over poverty estimates from 2012-2013 (Uganda Bureau of Statistics, 2018). Among the rural population, 25% of people are estimated to be living in poverty. For individuals for whom small-scale crop farming is their household's main source of income, the incidence of poverty rises to 29%. For individuals for whom commercial farming is their household's main source of income,

only 14% experience poverty. Regarding food insecurity, 40% of rural households are estimated to not have enough to eat (Uganda Bureau of Statistics, 2018). Taken as a whole, the litany of statistics can begin to paint a picture of a nation experiencing widespread hunger and poverty, but with great potential to leverage its agriculture sector as it moves toward socioeconomic well-being.

As 84% of the population lives in rural areas, agriculture presents one of the best economic opportunities for improving Ugandans' livelihoods (FAO, 2017). For small-holder farmers, market participation has been found to decrease both poverty and food insecurity (Barrett, 2007; Frelat et al., 2016; Montalbano, Pietrelli, & Salvatici, 2018). Engaging with the market encourages market-oriented agricultural production. Households specialize in goods in which they have a comparative advantage and benefit from efficiencies, leading to economic growth. Increased market participation can also improve a household's nutrition through greater dietary diversity (Sibhatu, Krishna, & Qaim, 2015). Stimulating market participation among Uganda's rural small-holder farmers holds great potential to help this population escape crushing poverty traps.

1.2 Market Facilitation

The United States Agency for International Development (USAID) is the arm of the American federal government that provides foreign assistance. Established in 1961, the agency has provided programs around education, agriculture, disaster relief, economic opportunities, and public health, to countries across the world with the goal of reducing poverty globally. The Feed the Future - Value Chain (FTF-VC) program in Uganda is one of these efforts. The initiative focuses on reducing hunger and food insecurity by encouraging agricultural production and market connections. USAID's market facilitation initiative in this vein "focuses interventions at strategic leverage points within a system, such as economic or social structures and incentives, in order to optimize its functionality" (USAID, 2015). The approach takes a system-wide view and uses a "light touch" that aims to strengthen relationships, realign incentives, and introduce new models into the system without taking a direct role. Bolstering the institutions and infrastructure central to connecting with markets will allow for sustainable development that will continue after the programming finishes. Within the agricultural market system, some examples of market facilitation include strengthening the mechanisms through which households access quality inputs, receive training on agricultural techniques, build connections with buyers, and learn about relevant market information such as prices (USAID, 2011). Strengthening of the agricultural market system is seen as an investment in the country's economy and stability.

As USAID seeks to develop programming that is contextually relevant, the organization is continually trying to understand the population that it serves. Because the impact of market facilitation efforts is indirect (in contrast to traditional direct aid interventions, such as providing farmers with equipment),

understanding how it affects households is difficult. Lack of clarity regarding the link between household outcomes and market facilitation interventions can make developing appropriate programming a challenge. In addition, the rates and levels of market participation are largely unknown. Evaluating the ways in which households engage with markets is a crucial piece to illuminating the connection between households and market facilitation interventions.

Such evaluation is one aspect of the Market Systems Monitoring Activity. The purpose of this activity is to develop approaches for assessing the systemic impact of the USAID Uganda FTF-VC market facilitation activities. The team uses systems engineering approaches to understand and measure how the combination of activities under USAID Uganda are enabling systemic change in the agricultural sector. This thesis will help to inform the organization's understanding of the Ugandan small-holder farmer population that it serves. The evidence provided by the analyses in this thesis will enable USAID to direct sustainable initiatives which address specific barriers that households struggle with when connecting with agricultural markets.

1.3 Market Participation

Previous work has explored how various household characteristics impact both the decision to participate in markets and the level of market participation. Several papers build on "double hurdle" models to first see what factors are connected to engagement, and second, conditional on participation, intensity of commercialization is explored. We use a similar approach in Chapter 4 to first investigate features which contribute to market participation generally, and then those which influence level of participation in regards to a specific crop in a specific season.

Goetz (1992) produced the pioneering work on this model. Results show that households in Senegal which sell coarse grains have increased access to information, such as about market prices. An increased amount of grain sold is linked to access to processing technologies and increased market prices.

Key, Sadoulet, and de Janvry (2000) found that among maize producers in Mexico, decreasing transportation costs and connecting sellers with official buyers was associated with increased market participation. Increased quantity sold was associated with membership in a producer organization, selling to an official source, local use of high yielding varieties, access to formal credit, increased price, and the local level of mechanization.

Vance and Geoghegan (2004) also investigated a two stage model for Mexican maize farmers. The decision to sell was connected with a larger area of land owned, receiving government credit, owning a vehicle, and being located in a depot village. Distance to market had a non-linear and significant impact on the decision to sell. Vehicle ownership and living in a depot village with more trade may

reflect decreased costs to information gathering and market accessibility. The area of maize that selling households planted (the proxy for level of market participation in this study) was connected to good soil, government credit, more months of off-farm labor, older household heads, and more household members above 12 years of age.

Heltberg and Tarp (2002) found that market participation for all crops in Mozambique was connected to increased maize yield in the province, being located outside of a risky area, not having outside employment, younger household heads, more trees, more land per household worker, owning transportation, being closer to a railway, and being further from the province capital. The revenue earned from all crops sold is connected to increased price of groundnuts, increased farm size per household worker, more trees owned, tractor ownership, increased maize yield in the province, living outside of a risky area, owning transportation, being closer to a railway, and being further from the province capital.

Alene et al (2008) found that participation in Kenya's maize market was increased by greater availability of extension agents in a district, credit access, increased land per capita, increased adults in the household, living closer to a fertilizer market, having access to transport equipment, and not having off-farm income. In this model, the extent of engagement was modeled by the amount of maize sold. This was associated with more extension agents in a district, increased adults in the household, increased maize prices, being a member of a producer organization, and being close to both fertilizer and maize markets.

The model for market participation developed in Yusuf (2015) found increased farm size and being involved in agriculture as a major occupation increased the likelihood of Nigerian households being net sellers of vegetables instead of net buyers or subsistence. This model then examined level of engagement as measured by the portion of output sold. Increased engagement was associated with older household heads, larger farms, access to credit, lower transportation costs, and shorter distances to markets.

In contrast to studies using the amount sold and revenue generated as a metric for extent of market engagement, Omiti, Otieno, Nyanamba, and McCullough (2009) use the percent of output as a proxy for the level of engagement of Kenyan maize farmers. This model showed that increased intensity of market participation was linked to higher education of the household head, higher output, reduced distance to market, and not having non-farm income.

Lifeyo (2015) actually employs a three-stage model on a single crop to understand which farmers in Malawi will choose to grow beans, which of those growers will choose to sell their beans, and how much those sellers choose to market. This study also employed a net seller, net buyer, and autarkic model. Factors associated with net sellers were more market experience, reduced distance to market, and region. Strangely, ownership of a radio discouraged acting as a net seller. Level of engagement, as

measured by the kg sold of beans, was associated with higher selling prices, male household heads, the availability of market extension services, the amount of land owned, and access to credit. Transporting the crop to market on foot decreased the amount sold compared to other modes of transportation. Increased education was also associated with a decreased quantity sold.

Across these studies, four different metrics for the intensity of market engagement are used:

- amount sold (Alene et al., 2008; Goetz, 1992; Lifeyo, 2017)
- revenue earned (Heltberg & Tarp, 2002)
- area planted (Vance & Geoghegan, 2004, 3)
- proportion sold (Yusuf, Ashagidigbi, & Mustapha, 2015; Omiti, Otieno, Nyanamba, & McCullough, 2009)

The outcome variable chosen for a model heavily influences which independent variables emerge as significant. The conceptual differences between these four different market participation intensity metrics will be reflected in the differing results. Compared to the amount sold, the revenue model may differ based on a household's ability to command higher prices as influenced by increased market information. The area of land a household plants may reflect their intentions to commercialize, but differences in agronomic practices (influenced by experience level, information access, and market infrastructure) can result in wildly different harvest amounts. Households may sell comparable portions of their crops despite large differences in the area planted, amount sold, and revenue earned. This thesis will seek to develop and compare models based on two of these metrics. We will contrast these models and discuss how the choice of an outcome variable influences the factors that are output as important for market participation. How does choosing an output metric for market participation impact which independent variables are significant? This question holds immense weight for organizations like USAID and programs like Feed the Future, which may design market facilitation initiatives around the results of such models. Depending on how organizations choose to analyze systems, different programming areas will be prioritized.

Throughout our discussion, we will be referring to "market development" and "market access." "Market development" encompasses the size and maturity of the local market system. Well-developed markets may have more participants (buyers, sellers, processors, other service providers) and stronger and more established relationships between participants. "Market access" refers to the physical ability to reach a place to sell. This includes the transportation network and access to modes of transport.

1.4 Research Question

This thesis builds on previous literature by researching the dynamics of market participation in the Ugandan context.

What factors impact farmers' decisions to participate in the market?

We want to explore how different characteristics of farming households relate and compare to each other in terms of the impact they have on whether or not Ugandan farming households participate in the market. Through regression modeling on household survey information, we attempt to elucidate the contribution of these dimensions as they impact market participation in our dataset. We want to understand what household characteristics are linked to commercialization choices. One way to do this is to apply a model from previous research and see if the original results hold true in our case. To that end, we re-create a model from the literature and see how the results compare for our dataset.

Throughout this discussion, the word "barrier" will be used to describe characteristics that are connected to a decrease in the probability that a household engages with the market. While some of these barriers may be immutable features of specific households, other can highlight opportunities for market facilitation interventions. Specifically identifying these barriers can help governments and aid organizations determine what programs or regulations can be implemented to reduce the impact of these barriers and encourage the small-holder farmer population to participate in the market. In the regression models built for this research, barriers will be identified as features which have a significant and negative impact on market engagement.

Then word "enabler" will be used to describe characteristics that connected to an increase in the probability that a household engages with the market. Specifically identifying these enablers can help governments and aid organizations determine what programs or regulations can be implemented to increase the impact of these enablers and encourage the small-holder farmer population to participate in the market. In the regression models built for this research, enablers will be identified as features which have a significant and positive impact on market engagement.

The fact that a specific feature may fall into neither of these categories is itself informative. There may be household characteristics or experiences which are believed to influence market engagement, but do not show up as either barriers or enablers. This may suggest that a feature that was perceived to be a barrier does not matter, or that there is some other feature that mitigates the impact. A characteristic that is hypothesized to show up as an enabler but does not may be indicative that it reflects some facet of life that does not influence household decisions to engage with the market.

Conditional on participation, what factors impact the level of market participation?

Two different outcome metrics (dependent variables) for level of market engagement will be explored

to find what factors are connected to the intensity of market participation. Some households are highly engaged in selling what they grow as an economic activity; other households participate minimally. The models built around these metrics can inform our understanding of the household characteristics most closely connected with increased level of market participation. This in turn will be used to suggest market facilitation efforts to increase the level of market participation. We will compare the two models in order to better understand how the different ways of defining this concept determine which barriers and enablers emerge from the model.

Is food security connected market participation or other metrics for agricultural engagement?

In order to explore whether or not market facilitation may exclude some vulnerable populations, we focus on one socioeconomic dimension: food security. We explore how household characteristics related to consumption need, production capacity, and food purchases relate to a household's level of food security. This analysis can help us better understand the dynamics of households experiencing food insecurity and how they may relate to some of the barriers and enablers associated with market participation. A concern with market facilitation efforts is that they may exclude households which do not already meet a certain threshold of assets and connections. Exploring one cross-section of socioeconomic vulnerability can begin to clarify if that concern is valid or not.

1.5 Motivation

The purpose this thesis is to identify household characteristics which are connected to increased market participation of farmers in Uganda. We expect our results to inform setting priorities in resource allocation policy for improved agriculture development. Crafting an empirical model can help provide quantitative evidence to guide the programming decisions made by USAID, our research sponsor. As an organization with limited resources provided by the American tax-payer, the agency seeks to craft interventions that have the greatest potential for impact. In addition to responsible allocation of resources, our hope is that thoughtfully focusing interventions on where they are needed most will have an increased impact on the Ugandan people. This effort is undertaken in service of the families and communities experiencing poverty and food insecurity throughout the region, and we have an obligation to these beneficiaries to use the information at our disposal to design programs from which they stand to gain the most.

The history of foreign aid and Western intervention is fraught with efforts tainted by racism, paternalism, cultural insensitivity, and echoes of colonialism (Easterly, 2006; Kothair, 2006; Moyo, 2009). As many Western nations have explicitly enriched themselves from the colonization and exploitation of sub-Saharan African nations, they now have a responsibility to reverse this damage and provide sustainable forms of reparations to lift back up these peoples. My hope with this thesis is to contribute a very small

piece to that by modeling the experiences of agricultural households in an effort to better understand where to direct these interventions to encourage sustainable economic growth.

1.6 Thesis Overview

This thesis explores the enablers and barriers to market participation of Ugandan farming households by analyzing a novel dataset. Chapter 2 provides an explanation of how the survey for the dataset was created and deployed, while Chapter 3 provides basic descriptive statistics for select features in the dataset. In Chapter 4, regression analyses are used to elucidate what household characteristics are linked to the decision to participate in agricultural markets. We also explore factors related to two different metrics for level of market participation. Chapter 5 focuses in on household food insecurity and how it intersects with market participation. Chapter 6 seeks to explain what the importance of barriers and enablers may mean in terms of market facilitation efforts. The conclusions are useful to governments and development organizations designing interventions with the end goal of improving farmer livelihoods. Identifying these key leverage points will ensure that programs with a systems view are intervening at the most effective points in the small-holder farmer supply chain.

Chapter 2

Data Collection

2.1 Survey Design

In order to gather information on what factors impact the levels of market engagement that farmers exhibit, our team designed and deployed a household survey called the "Farmer Market Engagement Study" (FMES). The survey consisted of seven sections focusing on different dimensions of household characteristics:

- Household demographics
- Finance
- Agronomic practices
- Agricultural inputs
- Production and harvest
- Market linkages
- Access to information

The survey itself can be found in Appendix A. The topics and specific questions were developed by the team at the Humanitarian Supply Chain Lab using invaluable input from the USAID-Uganda mission. Approval was obtained from the Institutional Review Board (IRB) from both MIT COUHES and the Uganda National Council for Science and Technology (UNCST) certifying the survey's compliance with ethical and legal standards. The following sections outline the specific questions which were asked and some of the reasoning behind including them.

2.1.1 Household Demographics

At each household, a respondent was asked some basic information about themselves and the other people living in the household. Each respondent was not necessarily the household head, but instead could be any adult from the household available during the time of the interview. This allowed for easier data collection than requiring the head of household be available at each location.

- **Age:** Though the respondent is not necessarily the head of household, age may be a proxy for experience. It is possible that older people have more years of farming and are therefore more likely to be knowledgeable about successful agricultural practices, typical market prices, and weather patterns. They could be more likely to have built relationships with traders, agents, and buyers over their years interacting in the marketplace.
- **Gender:** Though the respondent is not necessarily the head of household, gender may be associated with risk-aversion and therefore have an influence in the market participation decisions of a household.
- **Highest level of education attained:** Though the respondent is not necessarily the head of household, education may be tied to business savvy and increased ease of accessing information.
- **Household makeup:** The number of children and adults in the household contributes to the understanding of the household's labor capital and food consumption needs. Adults may help with a household's farming, and children may not. The number of children away at boarding school also informs this understanding of the household's expenses.
- **Number of people supported outside the household:** Many households may pay remittances to support family members who live far away. Sending money outside of the household could make a household more likely to engage in economic activities outside of subsistence agriculture.
- **Mobile phone ownership:** Owning a mobile phone can play an important role in connecting households to their communities and marketplaces. It is a source of communication and information. In addition, a mobile phone can give households the ability to open a mobile money account.
- **Access to transportation:** Respondents were asked if their household owned or had access to a bicycle, motorcycle, car, truck, or other form of transportation. Access to transportation can reduces barriers to reaching markets. It may also change a household's perception of what services are nearby and available.
- **Outside income:** Respondents were asked about any income earned outside of the home. They reported what the source of outside income was, what approximate percent (0-25%, 25-50%, 50-75%, 75-100%) the income contributed to total household earnings, and whether it was considered the primary source of income. Having an outside income may influence a household's decision to participate in the market. They may focus their energy on their other business and

only participate in agriculture insomuch as it provides some food for themselves.

- **Outside training/support:** Respondents were asked if the household had participated in or received training or support from the government or an NGO in the past year. If yes, they were asked about the nature of the support and details on what organization provided it. Households who receive extra training and support may be more likely to participate in the market. It may be that the support helped them in some way (accessing credit, learning more efficient farming techniques, etc) that has allowed them to increase their agricultural production and market connections, or it may reflect that more entrepreneurial households that participate in markets are also more likely to seek out support opportunities.

2.1.2 Finances

Households were asked about some of their money habits and their access to financial products. Accessing loans and savings groups can be crucial for farmers. The nature of agriculture requires a large upfront investment (seeds, chemicals, land, labor) and does not see a return for many months. Unpredictable weather and pests can make this initial investment quite risky. Since many small-holder farmers do not have large amounts of capital, loans can help them afford the inputs needed for a new planting season.

- **Bank account:** Households were asked if they had a bank account or not. Having a bank account may be a signal of business savvy, likelihood to access other financial products, and household capital.
- **Membership in a savings group:** Households were asked if they belonged to a village savings and loan association (VLSA), savings and credit cooperative organization (SACCO), or other savings group. Typically, these savings groups work by allowing member to pool their money. Members can borrow from the pool at a low interest rate in times of need. These can serve as a valuable resource for people who may not have good enough credit to borrow money from a bank. The low rates are also beneficial. In theory, membership in a savings group is a signal that a household has increased access to loans.

  Additionally, participating in one of these groups provides a household with a network through the other members. Social connections with other farmers can increase the information (about prices, trainings, techniques, selling opportunities, etc) a household has access to. It may reflect an element of the "connectedness" of a household to markets.
- **Mobile money account:** Mobile money is an electronic wallet service popular in many developing economies. In the system, a phone number acts like a bank account, and users can add money, transfer money to others, and withdraw money. Deposits and withdrawals of cash require access to a mobile money agent. Mobile money can make transactions fast and secure.

- **Borrowing money:** Respondents were asked if anyone in the household had borrowed money in the past year. If yes, they were asked how much, from whom, and the reason for borrowing. If no, they were asked if anyone in the household had attempted to borrow money in the past year and been unsuccessful. If someone had tried and failed to borrow money, they were asked why they were denied. In asking these questions, we are trying to paint a picture of household's access to finance. This type of access can contribute to increased productivity through enabling the purchase and rental of land and other agricultural inputs and services. Learning about why people were barred from accessing finance can inform thinking on how to improve the financial system to work for more households.
- **Credit from a business:** Respondents were asked if anyone in the household had received credit from a business. An example would be an agroinput shop allowing a household to pay for seeds at the end of the season.

  Households who received this type of credit were asked about the kind of business that extended it to them. This kind of credit can be a signal of positive relationships along the supply chain.
- **Lending money:** Respondents were asked if anyone in the household lent money to someone outside the household in the past year. This informs our understanding of the financial flows of a household.

2.1.3 Agronomic Practices

Households were asked a wide range of questions to get more background on their engagement with agriculture. This included land use, livestock, trainings, crop problems, and food security.

- **Acres of land farmed:** Households that farm more land may have a larger harvest and therefore be more likely to sell some of it.
- **Acres of land owned:** This informs our understanding of a household's assets and level of economic security, in addition to its production potential.
- **Acres of land rented or borrowed:** Renting land is associated with more risk than owning. It puts pressure on farmers to achieve a yield on that land that can at minimum cover the cost of the land itself.
- **Previous land rental:** Households were asked if they had rented land in the past. If yes, they were asked about what year (they could choose multiple years, given the options from "before 2007" and then every year from 2007 through 2018). If they rented land anytime after 2012, they were asked how much land and at what price. Households which did not rent land in the past were asked why they did not rent. This can contribute to understanding the financial flows of a household as well as how commonplace land rental is.

- **Rental price:** The cost per acre of land per season gives intuition for some of the start-up costs of agriculture.
- **Hired labor:** Respondents were asked if the household hired any farm labor this or the previous season. If yes, they were asked how many people were hired and the average number of days each worked.
- **Livestock:** Households were asked if they owned livestock. Those that did reported what kind and how many of each. Livestock contribute to a household's assets.
- **Training:** Households were asked if they have ever received any training on farming techniques. Government extension workers and aid organizations put on trainings to educate farmers about best-practices in agriculture, with the goal of increasing yield and overall production. Households that answered affirmatively to receiving training were asked follow-up questions about their most recent training. These details included what year it happened, who provided it, what it was about, and how they learned about the training opportunity.
- **Crop problems:** For each of the two most recent seasons (June-July 2017 and November-December 2017), households were asked if they experienced a range of crop problems. Drought, heavy rains, late rains, army worm, other insects, crop diseases, fire, and other problems all impact a household's yield and expenditures.
- **Months primarily relying on farm for food:** Respondents were asked the number of months out of a typical year that the household eats primarily food grown on the farm. This is meant to be a measure to understand how intensely households depend on agriculture. Some may farm casually as a supplement to their diets, so this number will be low. Other households use their farm as their main source of food and livelihood and will report a higher number of months relying on their farm. Still other households may aspire to meet their household consumption with their farm, but may not own enough assets or be productive enough, so they will report a lower number.
- **Frequency of skipping meals:** Respondents were asked if anyone in the household ever had to skip meals because there was not enough food. This question was meant to roughly assess as household's level of food security. Food security is the ability of a household to feed all members nutritious meals such that they do not regularly or acutely experience hunger. Skipping meals because of a lack of food is a clear indicator of food insecurity. Respondents were given the following options:
  - Never
  - A few times per year
  - Once or twice a month
  - Once or twice a week
  - Other

- **Typically purchase food:** Respondents were asked if their household buys food every year. This question intends to get another dimension of reliance on the farm. In the past, has this household typically been able to rely on what they produce?
- **Purchased food for consumption:** Respondents were asked if the household purchased food for consumption in the past year. This question and the follow-ups were intended to gather more information about household food purchasing patterns to understand some of the dynamics of relying on the farm for sustenance. If respondents reported that the household did purchase food for consumption, follow-up questions were asked:
  - **During which months:** Going through the months of the past year one by one, households were asked if they purchased food. If they did, follow-up questions were asked:
    * **Quantity purchased:** For each month, households reported the quantity of food that they purchased. This was most often reported in kg, sometimes in bags or bunches.
    * **Total cost:** For each month, households reported the total that they spent on food.
    * **Reasons for purchasing:** For each month, households were asked to choose from several reasons to describe why they purchased the food then. The options include:
      · Did not plant this crop
      · Poor harvest
      · Little storage capacity
      · Not enough production because of small land size
      · Sale of harvest due to urgency
      · Other

2.1.4 Agricultural Inputs

This section gathers information about the use of consumable agricultural inputs. While the previous section asked about land and labor, here we seek to learn about seeds, chemicals, and some of the patterns to purchasing these. These kinds of inputs influence the level of productivity that a household can achieve.

- **Nearby input shop:** Respondents were asked if there was an input shop nearby. Agroinput shops provide seeds, chemicals, storage bags, and tools that can help farmers increase their yield. Shop owners can also provide valuable information to farmers about the proper use of products. "Nearby" is a subjective term and can be influenced by the type of transportation a household has access to. "Nearby" on a bicycle is a much smaller distance than "nearby" on a truck. Respondents who said that there was an input shop nearby were asked about the name of the input dealer, the

town/village that they are located in, and if the household had ever purchased anything from there.

A network can take shape from the names and locations of these agrodealers.

Households which reported affirmatively to purchasing inputs from a nearby input shop were then asked why they choose to do so and how often they purchased them.

One of the provided options for describing why a household purchases from an input show is that the shop has a good reputation. Households who specified this option were then asked to provide a short answer describing why they believed the shop has a good reputation. This description informs our understanding of the relationships that exist between farmers and agrodealers. It reveals what characteristics farmers value and trust when purchasing inputs.

Households which reported having an input shop but did not purchase from it were asked why. They were provided the following options to describe their decision:

  - Do not purchase inputs
  - Cannot afford inputs
  - The shop is too far away
  - The shop has a bad reputation
  - They do not trust the shop owner
  - Other

- **Seed sources:** For each of the two seasons in 2017, households were asked where their seeds came from. They could choose from one or more options:

  - Home-saved
  - Purchased
  - NGO
  - Government
  - Other

  Households that indicated that they purchased seeds were asked a follow-up question about where these seeds were purchased from:

  - Input shop
  - Family
  - Friend
  - Neighbor
  - Cooperative
  - Other

- **Agricultural chemical use:** For both seasons in 2017, households were asked a series of questions about their usage of agricultural chemicals. First, for each season, they were asked if they had used chemicals.

  Households which answered affirmatively were then asked what kind (fertilizers, herbicides, pesticides) and where they purchased them from.

  Agricultural chemical use can be an indicator of training, education, and wealth. The use of these chemicals is shown to increase yield.

  Households which answered negatively to using chemicals were asked to specify why they did not use them.
- **Awareness of counterfeits:** Households were asked if they were aware of issues with counterfeit inputs. Counterfeit, adulterated, and poor quality chemicals and seeds have been causing problems in the input market throughout Uganda. The reduced quality has eroded farmers' confidence in these products and dampened rates of adoption (Ashour, Gilligan, Hoel, & Karachiwalla, 2019). The risk of poor return is too high for farmers to invest money in chemicals, especially since households have extremely limited financial resources.

  Households that are aware of counterfeit inputs were asked to provide a short answer about what they do to avoid them. They were also asked how they learned about the problem.
- **Input quality valuation:** Two questions were asked to assess households' stated willingness-to-pay for high quality agricultural inputs.

  The first question had respondents consider a 50g tin of seeds valued at 20,000 UGX (equivalent to $5.35). They were asked how much additional money they would be willing to pay for the seeds to be guaranteed a germination rate of 85-90%.

  The second question had respondents consider 1L of pesticide valued at 20,000 UGX (equivalent to $5.35). They were asked how much additional money they would be willing to pay if the product was certified as genuine and effective.

  The answers to these questions can reveal the extent to which farmers perceive there to be problems with the current products on the market. Expectation that these products are low quality will be reflected in a non-zero willingness-to-pay above the stated price.

2.1.5 Production & Harvest

- **Crops planted:** Households were asked to choose from a list of options and identify which crops were planted for the June-July 2017 season and which were planted for the November-December

2017 season. Depending on the region, some crops serve as cash crops while others are grown primarily for home consumption.

- For each of the top 3 crops planted by acreage in each of the 2 seasons, the following questions were asked.
  - **Acres planted:** Households reported the number of acres that they planted.
  - **Amount harvested:** Households reported how much was harvested.
  - **Storage:** Households were asked if they stored any of the harvest for each crop, and if so, how much. Those who stored were asked about which technologies they used.
  - **Amount sold:** Households were asked about the quantity that they sold.
  - **Price received:** If a household sold a non-zero amount, they were asked what price they received. This could be reported in UGX/kg, UGX/bag, or in total.
  - **Buyer:** For each crop sold, households were asked to provide a short answer describing who they sold it to and where that buyer is located. They were also asked how they came to know this buyer and why they decided to sell to them. Respondents who chose the option that their buyer has a good reputation were asked to provide a short answer describing why they have this good standing. Respondents were also asked if they had a formal arrangement with the buyer and for how many previous seasons they have done business with the buyer (if ever). Respondents were also asked to rank how much they trust this buyer. Households that have a positive, established relationship may receive higher, more stable prices than others. Breaking down the various components of how households connect to buyers and why they choose to continue with them informs our understanding of ways to strengthen those connections.
  - **Selling logistics:** For each crop sold, households were asked if the buyer came to pick up the produce or if the households had to deliver it. Households that delivered their produce were asked what mode of transportation they used and how much it cost them. In infrastructure-poor environments, there may be some benefit to connecting with buyers who are willing to pick up the produce. However, it may impact the price that a household receives.
  - **Reasons for not selling:** For each crop not sold, households were asked why they did not sell it. The following options were presented:
    * No surplus
    * No buyers available
    * Price too low
    * Other

- **Awareness of improved storage technologies:** Households who are aware of improved storage bags/silos may have better access to information. They could have learned about these new technologies from an input dealer, an organization which provides training, or any other connection in the market place.
- **Awareness of quality-differentiated pricing:** Households were asked if they were aware of any farmers who had received better prices for their crops because they were higher quality. In information-poor environments, there is little incentive for providing a higher price for better quality. Farmers may not have the knowledge to assess the quality of their own products. They may have limited information about market pricing in general. Most buyers offer a set price for produce, which disadvantages farmers who expend time and resources to achieve higher quality crops. Awareness may reflect a farmer being more connected to the market and having higher access to information. Households that are aware of opportunities for quality-differentiated pricing may be more likely to produce higher quality crops.
- **Techniques for assessing crop quality:** Respondents were asked to describe how they determined a crop's quality, and what characteristics differentiated a good crop from a bad crop.

2.1.6 Market Linkages

- **Accessible services:** From a long list of agricultural services, households were asked which ones were accessible to them. Households with greater access to services may have a more well developed market environment.

  The list of services includes:

  - Plowing/tilling
  - Soil testing
  - Irrigation
  - Weeding
  - Spraying
  - Pruning
  - Harvesting
  - Drying
  - Grain cleaning
  - Hulling
  - Shelling/threshing
  - Milling
  - Packaging
  - Storage
  - Transportation
  - Extension/training

- **Services paid for:** From the same list of services above, households were asked which they had ever paid for. An increased number of services paid for may reflect a household's wealth, level of agricultural training, or lack of labor capital. It can also reflect a need for specialized equipment and skills, such as those used for milling, hulling, and threshing.
- **Desired use of services:** Households were asked to indicate which of the services they would use contingent on them being affordable. This can help pinpoint the agricultural processes that

farmers most struggle to perform themselves. Every household may conceive of the level of "affordable" differently, but the answers can paint a picture of how farmers value their money vs the time, effort, and equipment needed for various services.

- **Equipment rental:** Households were asked if they had rented any equipment. If yes, they were asked to indicate what kind of equipment.
- **Membership in a producer organization:** Households which are members of a producer organization or cooperative may have greater access to a range of services. Household which do have membership were asked when they joined and what kind of services the group offers. Households which do not have membership were asked if there was a group in the area that they could join if they wanted to. Producer organizations are independent, member-led collectives which can allow farmers to engage in collective marketing that improves economies of scale. They can reduce transportation costs, provide crop processing services, and achieve higher prices for crops than individual farmers could negotiate (Latynskiy & Berger, 2016). This type of group can also provide a household with access to a large network of farmers. Social network effects may increase access to and diffusion of information (Bosc et al., 2002). Membership in such an organization may be an indicator that a household is highly connected with and engaged in the market.
- **Other group membership:** Households indicated whether or not they were a part of any other groups. If so, they specified what kind.

2.1.7 Access to Information

- **Accessible information:** From a list of different types of information, respondents were asked to indicate which the household had access to. They could choose from:
  - Market prices
  - Weather data
  - Opportunities to sell crops
  - New planting, harvesting, or post-harvest handling techniques
  - New products or services
  - Product quality
  - Anti-counterfeit programs
  - Government programs
  - NGO programs

  For each type of information a household had access to, they were asked about how often they received it and from what source they received it:

  - Mobile
  - Newspaper
  - Radio
  - Television
  - Word of mouth
  - Local business
  - Other

  Better access to market information should help farmers learn about opportunities to sell their crops and what prices are fair. The method of communication over which they learn about this information can reflect a household's assets. It also helps assess what modes of communication are the most effective at reaching farming households.
- **Trainings:** Respondents were asked if they had attended trainings on specific topics related to product knowledge, agricultural practices, business practices, financial management, and post-harvest handling and production practices. Respondents explained what the training was and indicated who administered it (an NGO, extension worker, cooperative, government program, input dealer, or buyer).
- **Information gaps:** Respondents were asked to explain what information about agriculture or farming as a business they wished they had access to.
- **Major problems:** Respondents were asked if the household suffered from any traumatic setbacks, such as death or illness in the family in the past two years. Such serious shocks to households can have a large impact on their participation in agricultural and economic activities.
- **Inaccessible necessities:** Respondents were asked to describe any products or services that they felt were necessary to run their farm successfully, but were not currently accessible to the household. This allows respondents to directly express what difficulties they are facing, regardless of if they fit into the check-boxes of the survey.

2.2 Sampling Procedure

The purpose of the survey was to collect information about small-holder farmers, using the household as the unit of observation. To that end, we wanted our sample to be made up only of respondents whose households participated in agriculture by planting crops or owning livestock.

Given transportation and resource constraints, we purposively chose five different districts to deploy the survey in. Districts were chosen with input from USAID to represent Uganda's four regions (North, East, West, and Central) plus a specifically rural district. Districts with unique and non-representative

demographic and agroecological characteristics that differed from the majority of the country were not chosen, such as in the Karamoja region. Security concerns excluded northern districts which bordered South Sudan. The five chosen districts were Gulu, Pader, Ibanda, Mubende, and Iganga, as seen in Figure 2-1.

Within each district, we sought to interview 100 households, for a sample size of 500 farming households.

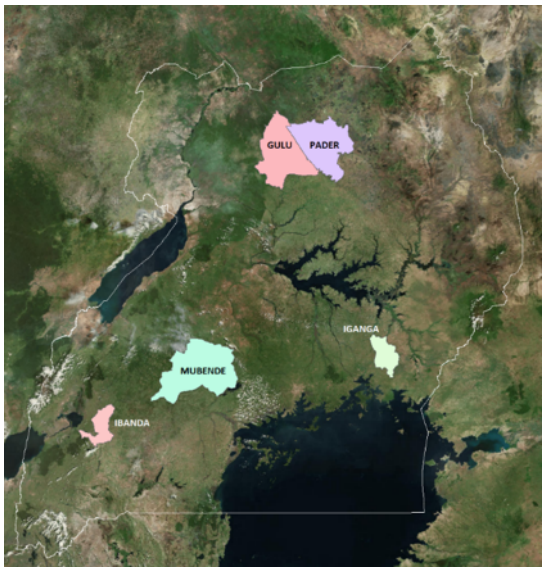

Figure 2-1: The five sampled districts within Uganda: Gulu, Pader, Ibanda, Mubende, and Iganga.

To form a representative sample, we randomly selected farmers within each district, given transportation and resource constraints. A two-stage clustered sampling process was used to identify the households that were interviewed. For the first stage, each district was divided into 2x2 km squares, which became the primary sampling unit. Within the sampled set of squares, the secondary sampling unit was individual farming households, which were selected from a set of buildings identified using satellite imagery.

1. We used GIS shapefiles that were created by the Ugandan Energy Sector GIS Working Group based on 2014 data provided by the Uganda Bureau of Statistics[1]. This dataset provides information that is the most up to date possible and provides information at the sub-county level, the smallest geographic unit of measure used by the census. The shapefile contains:
    - an outline for every sub-county in the country

[1]This data can be found at https://energydata.info/dataset/subcounty-boundaries-2014

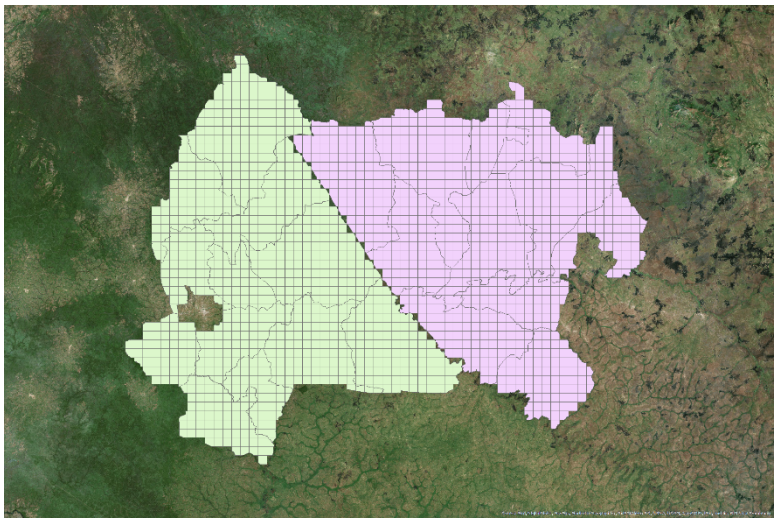

Figure 2-2: The primary stage sampling units. Gulu is highlighted in green, and Pader is highlighted in pink. The urban regions of each district have been removed, as evidenced by the hole in Gulu. Along the border of the two districts, sampling units less than 2 km^2 in area have been removed.

- the name of each sub-county
- the name of the county and district to which each sub-county belongs

2. Since the purpose of the study was to investigate market access among farming households, those households living in dense population centers were purposively excluded. These households are significantly less likely to participate in agriculture due to land constraints. In addition, their proximity to markets results in a wide difference from more rural areas in the way they sell and connect with other market actors. We defined these population centers based on official administrative units: Municipalities and Town Councils. To achieve this, we removed any sub-county in the shapefile that had the phrase "Town Council" in its name. We also removed any sub-county which belonged to a county that had the word "Municipality" in its name.
3. The remaining sub-counties for each chosen district were mapped in ArcGIS. A 2km x 2km grid was placed over each district. Figure 2-2 demonstrates how this looked for Gulu and Pader.
4. Boxes that fell on the border of a district of interest were excluded if more than 2 km^2 (half the area of the box) lay outside the target districts.
5. In order to give each farm household in the district approximately equal probability of being selected, the boxes had to be assigned a non-uniform probability distribution using a probability proportional to size method. In the ideal world, the total number of farming households in each box would be known, and each box would be assigned a probability of:

$$P_{box} = \frac{\#\text{ farming households in box}}{\#\text{ farming households in district}}$$

Since these exact numbers are impossible to know, the farmer population density of a sub-county was used to approximate the number of farmers in a given area. This approach assumes the population of farming households is distributed uniformly across a sub-county. The simplest case is a 2km x 2km box that sits directly in a single sub-county and does not straddle any borders. The farming household density for a sub-county can be calculated using the sub-county area and the number of subsistence farming households in a sub-county as given in the 2014 UBOS census. The census defined a subsistence farming household as one for which subsistence farming was their main source of livelihood. The estimated number of farming households would be the area of the box multiplied by the farming household density of that sub-county.

$$\text{estimated \# of farming households} = \text{box area} * \frac{\#\text{farmers in sub-county}}{\#\text{farmers in district}}$$

Some boxes straddled more than one sub-county such that there were multiple polygons created by the intersection of the box and the borders . The number of farming households in these boxes was calculated by:

$$\text{estimated \# of farming households} = \sum_{n=1}^{i} \text{area of polygon } i * \frac{\#\text{farmers in sub-county } i}{\#\text{farmers in district}}$$

The probability of choosing any given box is the estimated number of farming households in that box divided by the total number of farming households within the district.

$$P_{box} = \frac{\#\text{ estimated farmers in box}}{\#\text{ farmers in district}}$$

6. Once each box had been assigned a weighted probability according to the estimated number of farming households, 30 boxes were chosen at random for each of the districts.
7. The 150 boxes were then inspected using Google Maps satellite imagery, and 13 were eliminated according to the following criteria:
    - Topography: Boxes with more than 50% of their area covered by water or forest cover were eliminated. Two squares in Mubende were covered by a lake and three squares in Ibanda were covered by a forest.
    - Population size: Given our target of 10 interviews per box, those boxes with fewer than

10 visible buildings/compounds were eliminated. Five squares in Mubende, two in Gulu, and one in Pader were sparsely populated and thus eliminated. This may bias the sample toward farmers who are more connected to other market actors, have better infrastructure, and have smaller farm sizes. However, interviewing extremely rural areas that would require extra visits to other areas to make up for the lack of 10 interviews was unfeasible given the constraints in terms of transportation, time, and resources.

8. We used satellite imagery from Google Maps to identify and number the man-made structures in the 2x2 km boxes according to certain criteria. Buildings that were already labeled on Google Maps as churches or commercial buildings were not included.

   This process was completed for the first 22 boxes in each district.

   - Gulu and Pader: In these two rural districts, the secondary sampling units were compounds/homesteads, defined as a cleared area with buildings. Most of the population does not live in structures made with modern building materials. A single household will live on a compound with multiple huts.

     Though it is possible that some of these compounds contained multiple households, such as several families within an extended clan, we determined that the market access characteristics of these households are likely to be the same for the purposes of our study, and therefore the households did not need to be considered separately. Each clearing was marked and numbered, as were any buildings located outside of these groupings. Figure 2-3 provides an example of compounds, with each grouping that was counted circled in yellow.

Figure 2-3: A satellite view of example farmland in Uganda. Units that would have been identified as "households" have been circled.

Figure 2-4: A satellite view of example farmland in Ibanda. Each pin represents an individual household.

- Ibanda, Iganga, and Mubende: In these districts, a single unit was marked as groupings of buildings that seemed part of the same household based on to each other proximity and separation from other buildings. Buildings close together but on separate clearings indicates separate households. Buildings on the same clearing were typically grouped together. All other individual buildings were marked and numbered individually. Figure 2-4 provides an example of counting buildings in this setting.
- Buildings lying on the border of the box boundary were included.

We acknowledge that it is challenging to determine whether a building is a household or a business. Therefore, we included all buildings that did not meet the exclusion criteria above.

9. For the second stage of sampling, uniform random sampling was used to choose 30 households from the list of identified households for each square.
10. For each district, enumerators from a Ugandan data collection firm were given the first 10 chosen households for each of the first 10 chosen boxes. They were instructed to work with the local government officials to pre-screen households and determine that each one participated in agriculture (either growing crops or owning livestock). If a chosen household turned out to be a business, no one was present, or the household did not participate in agriculture, enumerators were instructed to choose a replacement from the next 10 chosen households (indices 11-20) out of the 30 household sample for that square. To accommodate for poor infrastructure that can make traveling even within a box very circuitous, enumerators were free to choose the nearest replacement household instead of moving down the replacements in order.

Visiting households prior to implementing the survey was done to establish the households were comfortable participating. Local officials known as the LC1 (approximately equivalent to a mayor in the United States) were contacted and brought to the pre-screen in order to ensure households that the survey had official approval and would be used appropriately.

11. Farming households were interviewed in the next few days after the initial pre-screen. Enumerators returned to each household and interviewed an available adult. Responses were collected using FieldTask software on tablets and the raw data was provided to our team. Each household was given an honorarium in exchange for their time and participation.

2.3 Data Collection Results

For three of the districts (Mubende, Iganga, and Ibanda), respondents were successfully surveyed in the first 10 squares provided to enumerators.

Square 8 in Gulu and square 6 in Pader had to be replaced by the respective squares 11 in each district. These two squares had to be excluded for operational feasibility; the communities living there were distrustful of outsiders and unwilling to participate in the survey.

In all of the districts, enumerators sometimes struggled to meet the 10 household target. Directions were to visit the next square at the top of the replacements list (ie, indices 11 and greater) and supplement from there. However, enumerators tried to compensate instead by oversampling other squares. The over- and under-sampling should not impact our results. In the end, the breakdown of respondents per square per district is show in Table 2.1.

The final sample contains responses from 498 households. Of these, 496 planted crops in 2017.

Ibanda	
Square ID	# Households
1	12
2	7
3	9
4	10
5	10
6	12
7	8
8	10
9	10
10	12

Gulu	
Square ID	# Households
1	11
2	10
3	8
4	10
5	7
6	12
7	10
9	10
10	9
11	12

Iganga	
Square ID	# Households
1	10
2	10
3	10
4	7
5	10
6	10
7	10
8	10
9	10
10	10

Mubende	
Square ID	# Households
1	10
2	11
3	10
4	11
5	10
6	7
7	10
8	12
9	10
10	10

Pader	
Square ID	# Households
1	14
2	9
3	10
4	10
5	10
7	10
8	10
9	10
10	10
11	9

Table 2.1: Number of respondents in each square.

THIS PAGE INTENTIONALLY LEFT BLANK

Chapter 3

Descriptive Statistics

This section will provide an overview of the dataset in order to provide the reader with intuition. Although our survey contained over 100 questions, we focus here only on features used throughout Chapters 4 and 5.

3.1 Demographics

3.1.1 Household Size

Respondents were asked to report the number of children and the number of adults living in their household. Averages for each district are shown in Figure 3-1. Households across districts have a similar number of adults. Iganga has slighter more children per household than other districts.

In the entire sample, household size has a standard deviation of 3.48.

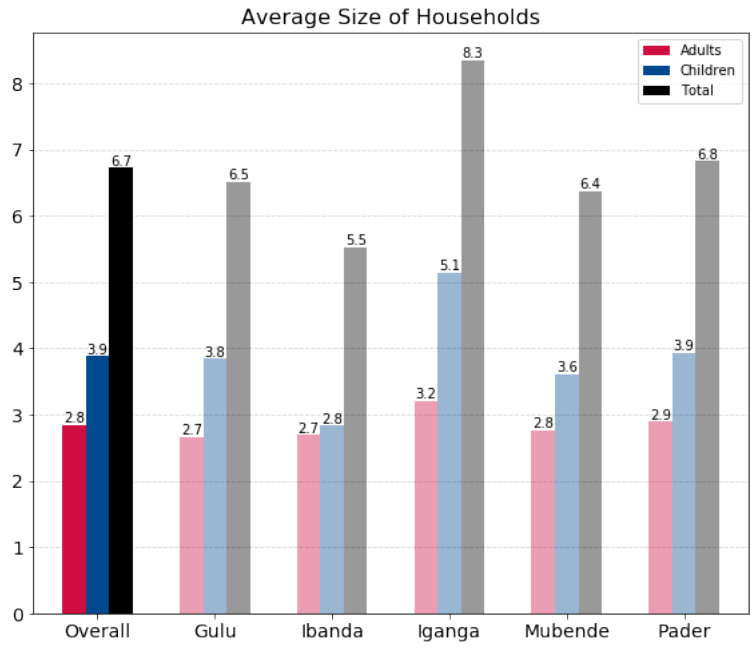

Figure 3-1: The average number of members in households in each district.

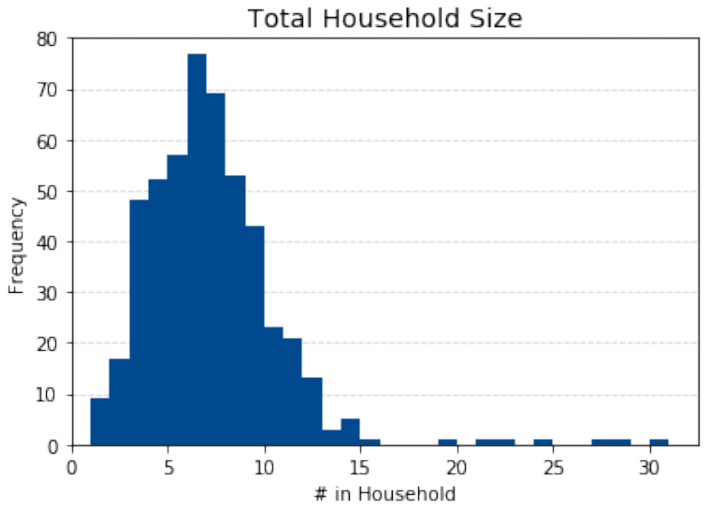

Figure 3-2: The distribution of household size across the entire sample.

3.1.2 Acres Planted

In the survey, households reported the acreage that they planted of various crops in June - July 2017 and November - December 2017. These acreages were summed to calculate the total area of land a household cultivated in 2017.

Across the entire sample, households on average report cultivating 6 acres. Figure 3-3 demonstrates how the average by district differs. Out of all of the regions, Ibanda and Iganga are considered the least rural. This is reflected in these two districts having a smaller average area cultivated than in other more rural areas.

The standard deviation is 4.84 acres.

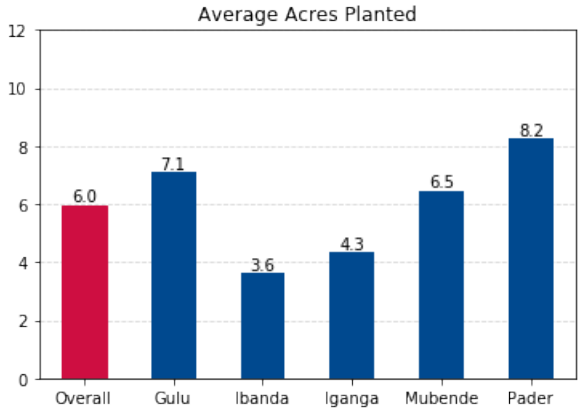

Figure 3-3: The average number of acres planted in 2017 (sum of both seasons) in each district.

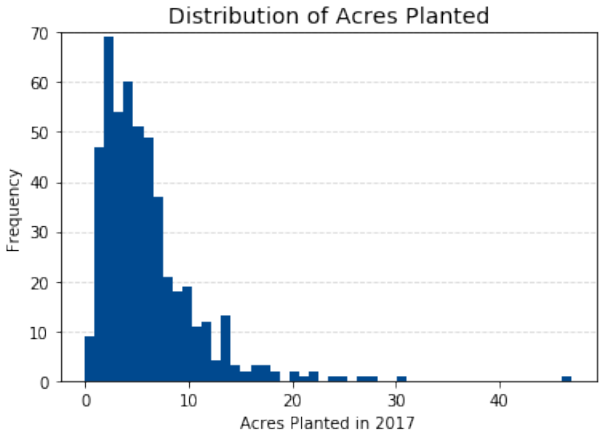

Figure 3-4: The distribution of acres planted.

3.1.3 Mobile Phone Ownership

Overall, 63% of households owned a mobile phone. Pader, one of the most rural districts survey, has a drastically lower rate of mobile phone ownership thatn other districts.

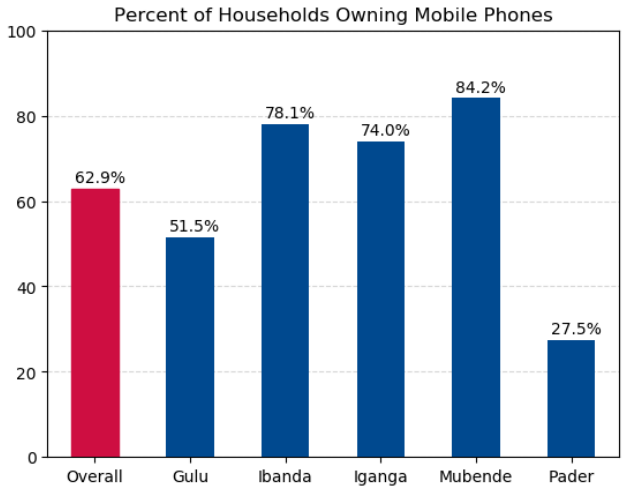

Figure 3-5: The percent of households in each district which own mobile phones.

3.1.4 Transportation Access

In the entire sample, almost 40% of households report lacking any access to private transportation. They do not have a motorcycle, bicycle, car, or truck. Figure 3-6 illustrates the breakdown by district.

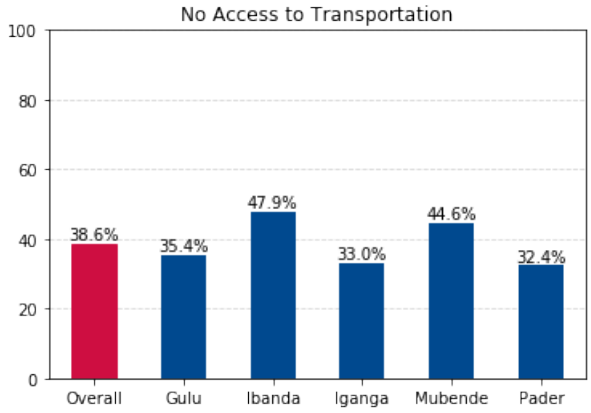

Figure 3-6: The percent of without any access to transportation.

Of the households which do have access to transportation, motorcyles and bicycles are by far the most common. Only 3 households had access to a truck; all were in Mubende. Only 8 households had access to a car; 3 in Mubende, 3 in Iganga, and 2 in Ibanda.

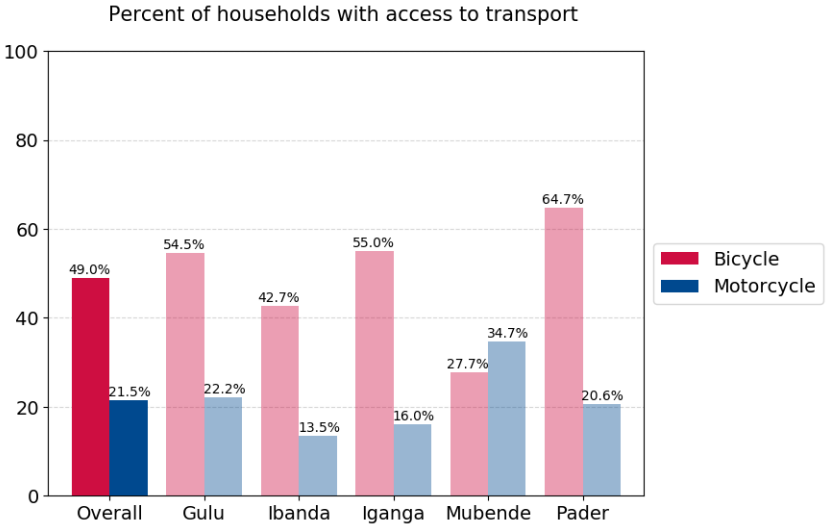

Figure 3-7: The percent of without any access to transportation.

3.1.5 Outside Income

In addition to farming, households may engage in other economic activities which may influence their decisions around participating in agriculture. To capture this impact, households were asked if any had earned income outside of farm in the past year. Those who responded that they did were asked to describe the nature of the work. Finally, households were asked to specify the approximate portion of the household's total income that this outside economic activity is responsible for. They could choose a bucket from 0 - 25%, 25 - 50%, 50 - 75%, or 75- 100%.

Figure 3-8 illustrates the responses. In the entire sample, about one-third of households had outside income. Working outside of the farm is less common in the rural regions of Gulu and Pader than in the relatively more developed areas of Iganga and Mubende.

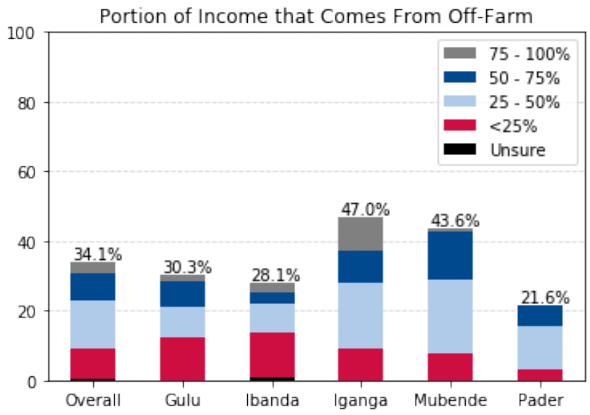

Figure 3-8: The percent of households with income outside of the farm.

In our models, we control for households for whom farming is more of a supplementary income versus a primary economic activity. These households may be less interested in participating in markets and therefore less sensitive to interventions relating to agriculture. We use a dummy variable to indicated if 75 - 100% of a household's income came from outside of the farm. Figure 3-9 explicitly shows the percent of households in each district who fall under this description.

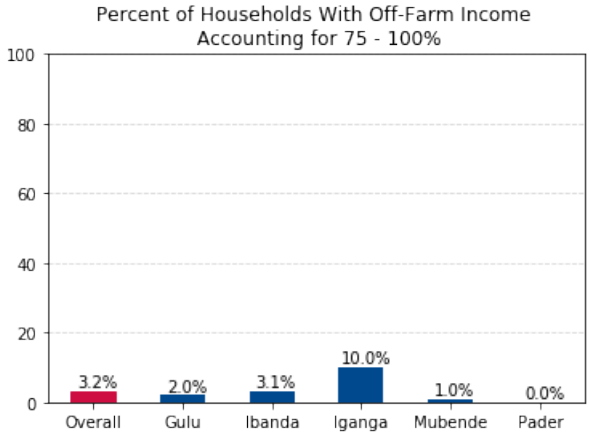

Figure 3-9: The percent with income outside of the farm that accounts for 75 - 100% of the household's income.

3.2 Finances

3.2.1 Received Credit

Overall, 12.7% of households received credit from a business. This may be in the form of seeds received at the beginning of the season and paid off at the end. It could also be a direct loan of cash or services that are paid back later.

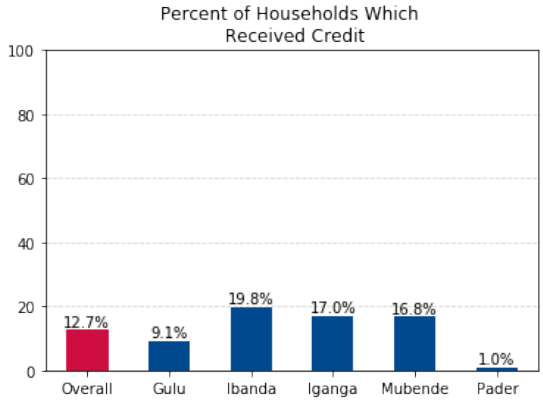

Figure 3-10: The percent of households.

3.3 Agricultural Inputs

3.3.1 Nearby Inputs Shop

Agroinputs shops are small businesses which provide a variety of agricultural products. They sell tools (shovels, buckets, sprayers, hoes), chemicals (fertilizers, herbicides, pesticides), and seeds. In the entire sample, 43% of households reported having an input shop nearby. Of course, "nearby" is a relative shop. A shop 10km away may be "nearby" for a household which owns a motorcycle, but may not be considered "nearby" to the close neighbor with no private transportation.

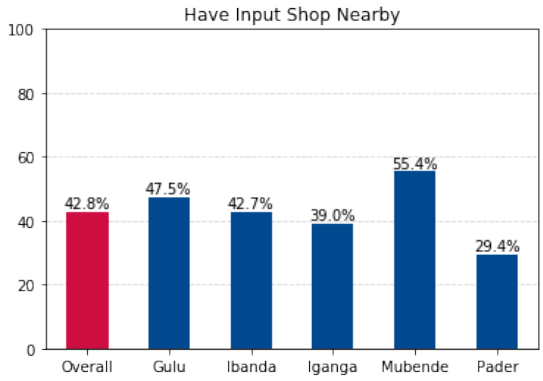

Figure 3-11: The percent with households with an input shop nearby.

3.4 Agronomic Practices

3.4.1 Food Security

Households were asked to indicate how frequently anyone had to skip a meal because there was not enough food.

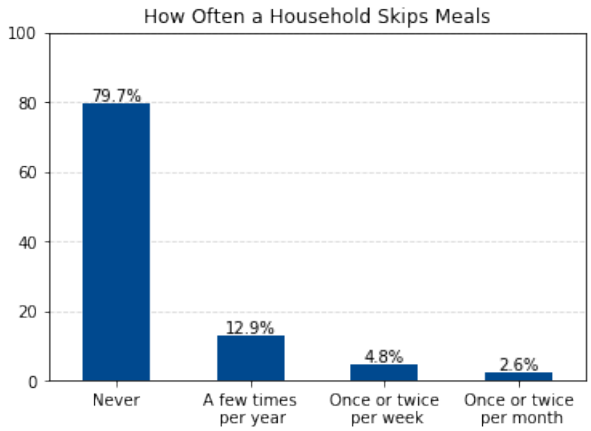

Figure 3-12: The frequency with which households skip meals.

For the purpose of the models discussed in Chapter 4, a dummy variable is used to indicate whether or not a household if food secure. This dummy variable takes on the value "1" if a household indicated that they never skip meals and a value of "0" otherwise.

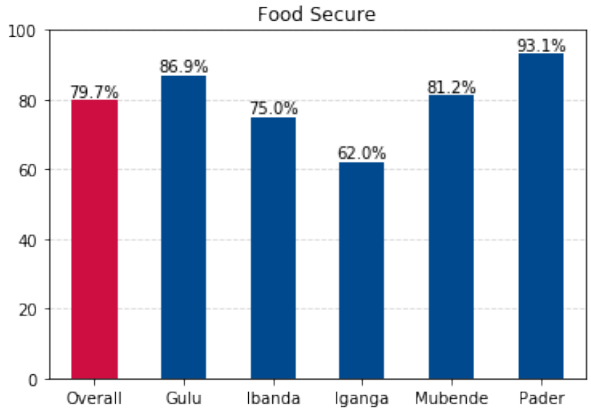

Figure 3-13: The percent of households which never skip meals (and are therefore classified as food secure) in each district.

3.4.2 Months Purchased Food

Households were asked to indicate the months in the past year during which they purchased food. The months that each household chose were totaled. The mean is 1.8.

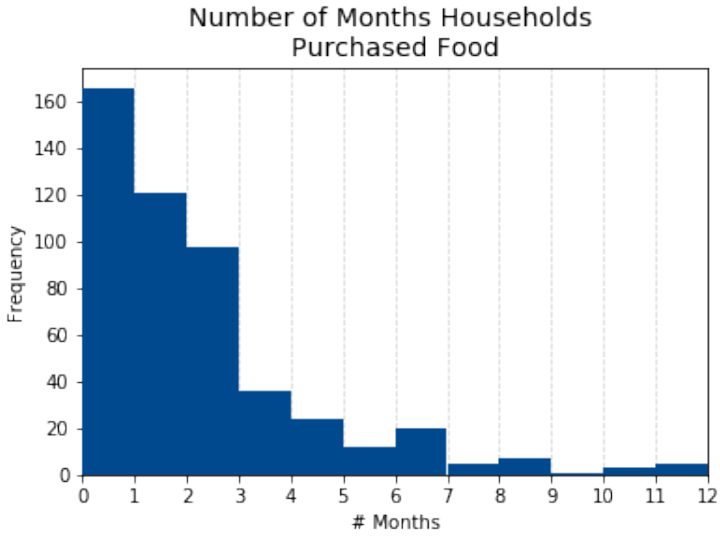

Figure 3-14: The distribution of the number of months in which households purchased food.

3.4.3 Annual Food Expenses

For each month that a household indicated it purchased food in the past year, it was asked to reported the total amount spent. The monthly food expenses were aggregated to come up with a total food expense for the past year.

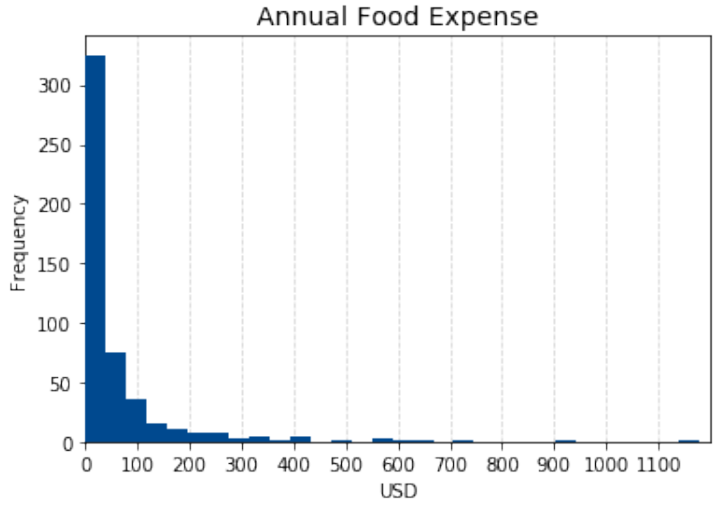

Figure 3-15: The distribution of the number of months in which households purchased food.

3.4.4 Annual Quantity of Food Purchased

For each month that a household indicated it purchased food in the past year, the respondent was asked the quantity of food purchased. The monthly totals were aggregated to come up with a total for the past year.

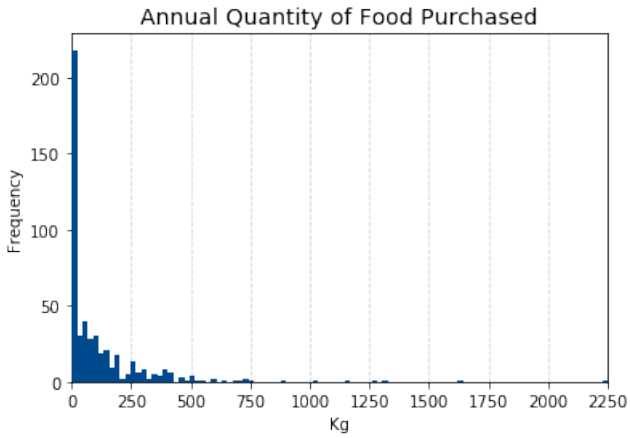

Figure 3-16: The distribution of the amount of food purchased in the last year.

3.5 Production and Harvest

3.5.1 Kilograms Harvested

For both seasons in 2017, each household was asked about their harvest for their top three crops according to the area planted. They were asked to specify the number of kilograms that they harvested for each of these six crop data-points.

The mean value is 1960 kg. Note that this includes a mix of different crops.

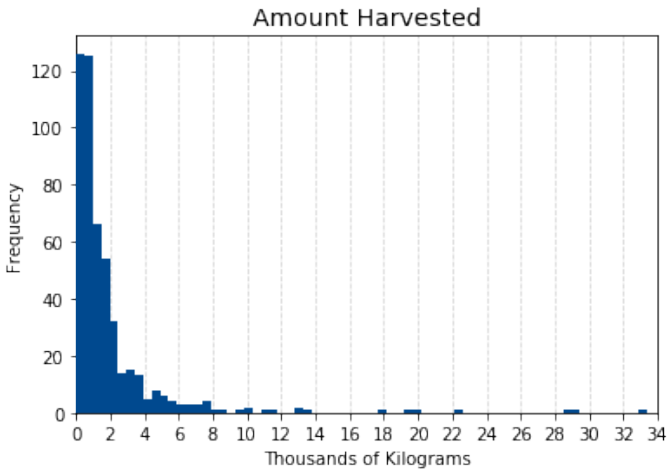

Figure 3-17: The harvest size distribution.

Figure 3-18 shows the distribution for $\ln(\frac{\text{kg harvested}}{\text{household size}})$ for the entire sample. This feature is used as an independent variable in the model in Chapter 4 for the decision to participate in the market. The mean is 5.1 and the standard deviation is 1.3.

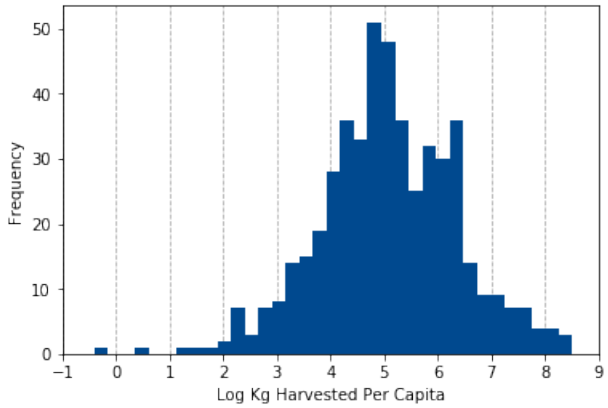

Figure 3-18: The harvest size distribution.

Figure 3-19 shows the distribution of the log of maize harvested, only for the subset of households which sold maize in November - December 2017. This feature is used in the models in Section 4.2 to understand level of market participation. The mean is 6.2 and the standard deviation is 1.2.

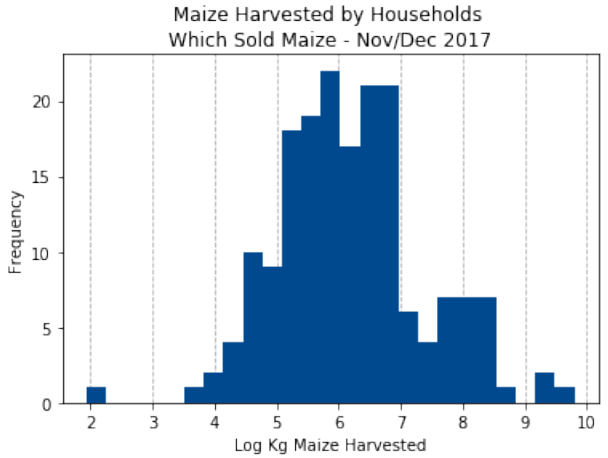

Figure 3-19: The distribution of how much log(maize) was harvested in November/December 2017 by households which sold maize in November/December 2017.

3.5.2 Kilograms Sold

For both seasons in 2017, each household was asked about the volume that they sold for their top three crops according to the area planted. The kg sold value is the sum of the amount sold for these six crops. The mean value is 1043 kg.

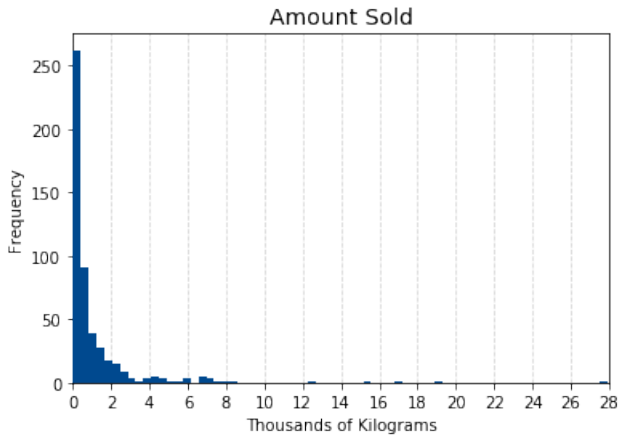

Figure 3-20: The harvest size distribution.

Section 4.2 examines the subset of households which sold maize in November - December 2017. The amount sold is used as a proxy for the level of market participation. To achieve a more normal target distribution, the log of the kg of maize sold is used as the dependent variable. The mean is 5.7 and the standard deviation is 1.3.

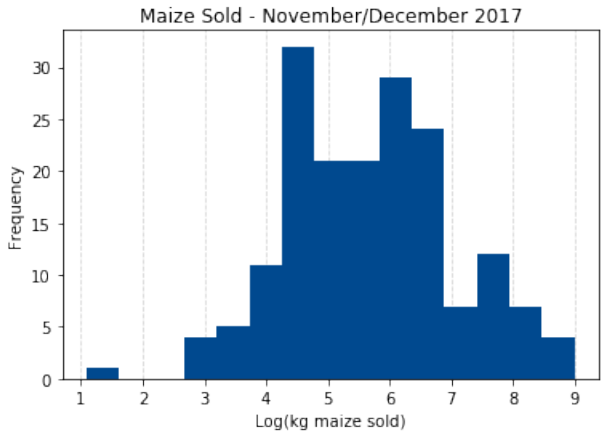

Figure 3-21: The distribution of the log(kg maize sold) for households which sold maize in November - December 2017.

3.5.3 Portion Sold

The second model explored in Section 4.2 used the logit of the portion sold as the dependent variable to explain level of market participation. This model focused on the subset of households which sold maize during November - December 2017. Figure 3-22 shows the distribution of the portion of the maize harvest that households sold.

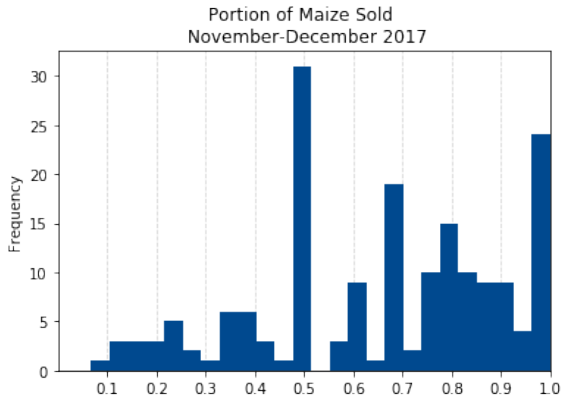

Figure 3-22: The distribution of the portion of maize sold in November - December 2017.

To transform this bounded portion into an unbounded value which can be used as the dependent variable in multiple linear regression, we take the logit.

$$\text{logit}(p) = \ln \frac{p}{1-p}$$

This distribution now looks more normal and takes on a continuous range. The mean is 1.1 and the standard deviation is 1.7. There were 21 households which sold their entire maize harvests, creating a spike on the upper end of the distribution. Since the logit of 1 is infinite, all farmers who sold 100% of their crop were adjusted to 99%, the logit of which is 4.6. Replacement values from 95% to 99.99% were experimented with in the model building process and found to have very minimal impact on the results. For simplicity, 99% was chosen as the final replacement value.

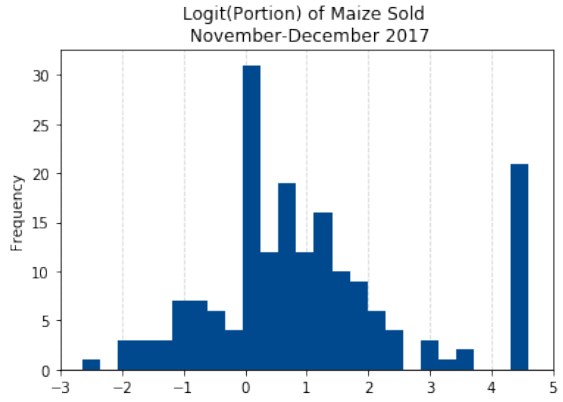

Figure 3-23: The distribution of the logit of the portion of maize sold in November - December 2017.

3.5.4 Mode of Buyer Transportation

For each crop that a household reported selling, they were asked how the crop was transported: by bicycle, motorcycle, truck, on foot, through a cooperative, having the buyer come to pick it up, or through another method.

Of the 2,988 transactions described by respondents, 142 (4.75%) were picked up by a buyer.

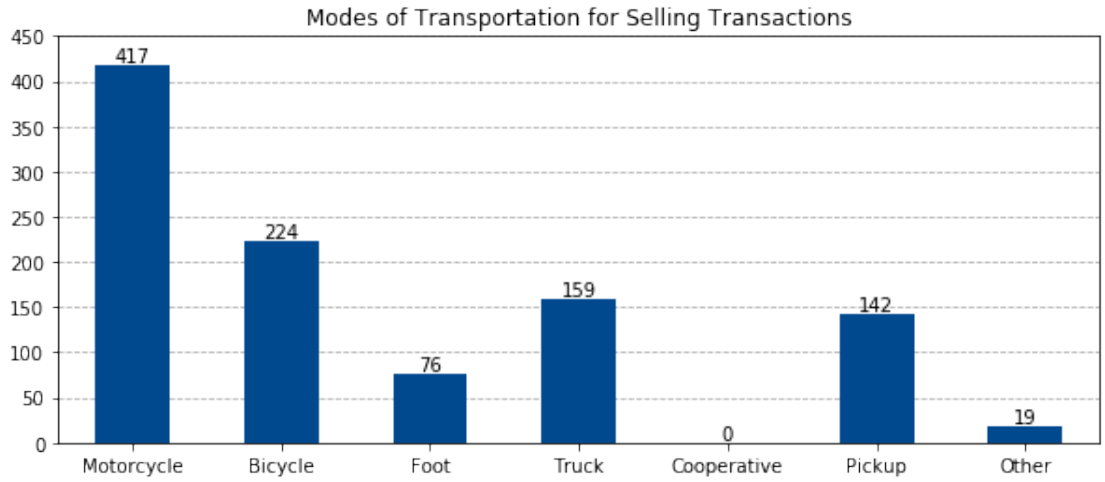

Figure 3-24: The mode of transport for each selling transaction.

3.5.5 Pickup

For the models described in Section 4.2 which focus on maize sold in November - December 2017, we include a variable for whether or not a buyer picked up the harvest. Of the 180 transactions included in this analysis, 23 (12.8%) were picked up by the buyer.

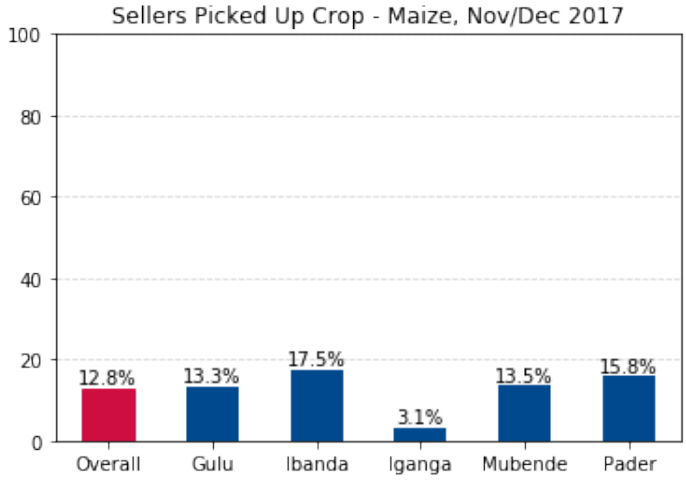

Figure 3-25: The percent of transactions for which maize was picked up.

3.6 Market Linkages

3.6.1 Services Sum

In the survey, households were asked which services from a list they had access to. These included:

- Plowing/tilling
- Soil testing
- Irrigation
- Weeding
- Spraying
- Pruning
- Harvesting
- Drying
- Grain cleaning
- Hulling
- Shelling/threshing
- Milling
- Packaging
- Storage
- Transportation
- Extension/training

The number of services each household indicated having access to was summed and used as a proxy for the development of the surrounding market infrastructure. The mean is 4.9 and the standard deviation is 2.9.

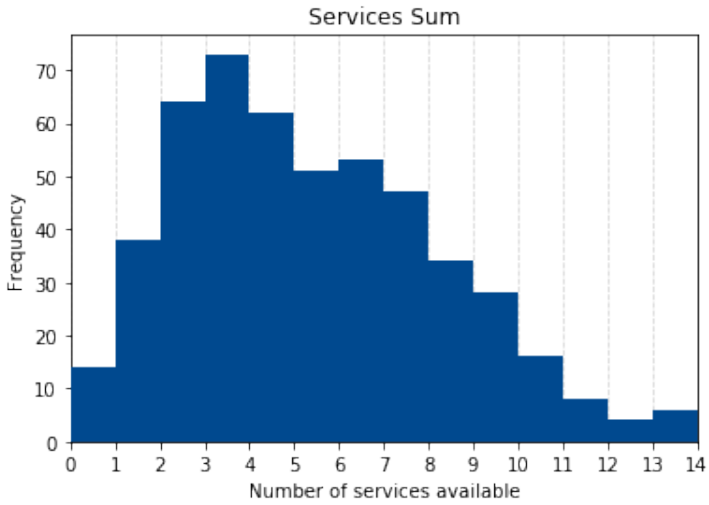

Figure 3-26: The number of services available to households.

3.6.2 Member of Producer Organization

Overall, 8.8% of households belonged to producer organizations.

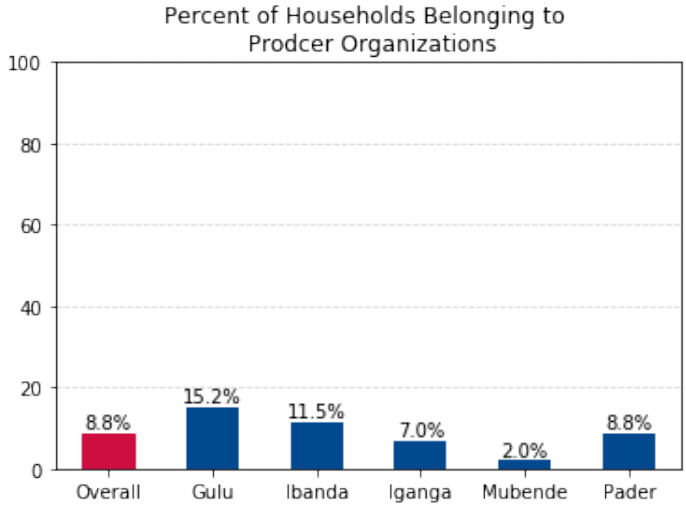

Figure 3-27: The percent of households belonging to producer organizations.

3.7 Geospatial Measures

3.7.1 Kilometers to Nearest Town

To understand the impact of connectedness and transportation infrastructure, we found the minimum distance traveled along roads for each household to reach a town. Within developing countries, it can be challenging to find completed maps or official information on municipalities. Our team decided to use the dataset of towns and cities from OpenStreetMaps (OSM). OSM is an open-source mapping tool whose data is generated by users worldwide. In contrast to Google's proprietary mapping data, OSM is freely available and contains highly detailed information. The dataset for Uganda was downloaded, and the locations of all places labelled as "town" or "city" were overlaid onto a map of the household GPS locations.

There are no hard and fast rules delineating villages, towns, and cities. Guidance provided by OSM defines a city as "the largest urban settlement or settlements within the territory" and a town as "an important urban centre, between a village and a city in size." [1]

[1] https://wiki.openstreetmap.org/wiki/Key:place

Uganda's road network was also downloaded from OSM and overlaid onto the map of towns, cities, and surveyed households. An origin-destination cost matrix was run to determine, for each household, what the minimum distance traveled along roads was until reaching a town/city. For each respondent, the name of the town reached by this minimum distance was also recorded.

This calculated distance may not reflect the reality experienced by households. They may choose to sell at the market of a town slightly further from them or take a different route than the one found by the mapping tool. People may along use more informal routes that may not show up in the OSM road network. However, this distance measurement is a proxy that is helpful for understanding the how accessible market centers may be to households.

3.7.2 Kilometers to Nearest Major Road

Roads labelled as "primary" or "secondary" were selected. Figure 3-30 shows the road network overlaid on satellite imagery, with the location of respondents show in yellow. Primary roads include highways. These classifications are intended to reflect a road's function and importance at the national level. These categories exclude residential roads, footpaths, and tertiary roads. Road networks from Google Maps was not able to be accessed, in addition to being less complete than the OSM version. Visual inspection showed that OSM had many roads in its network that were visible in Google satellite imagery, but not included in the Google road network.

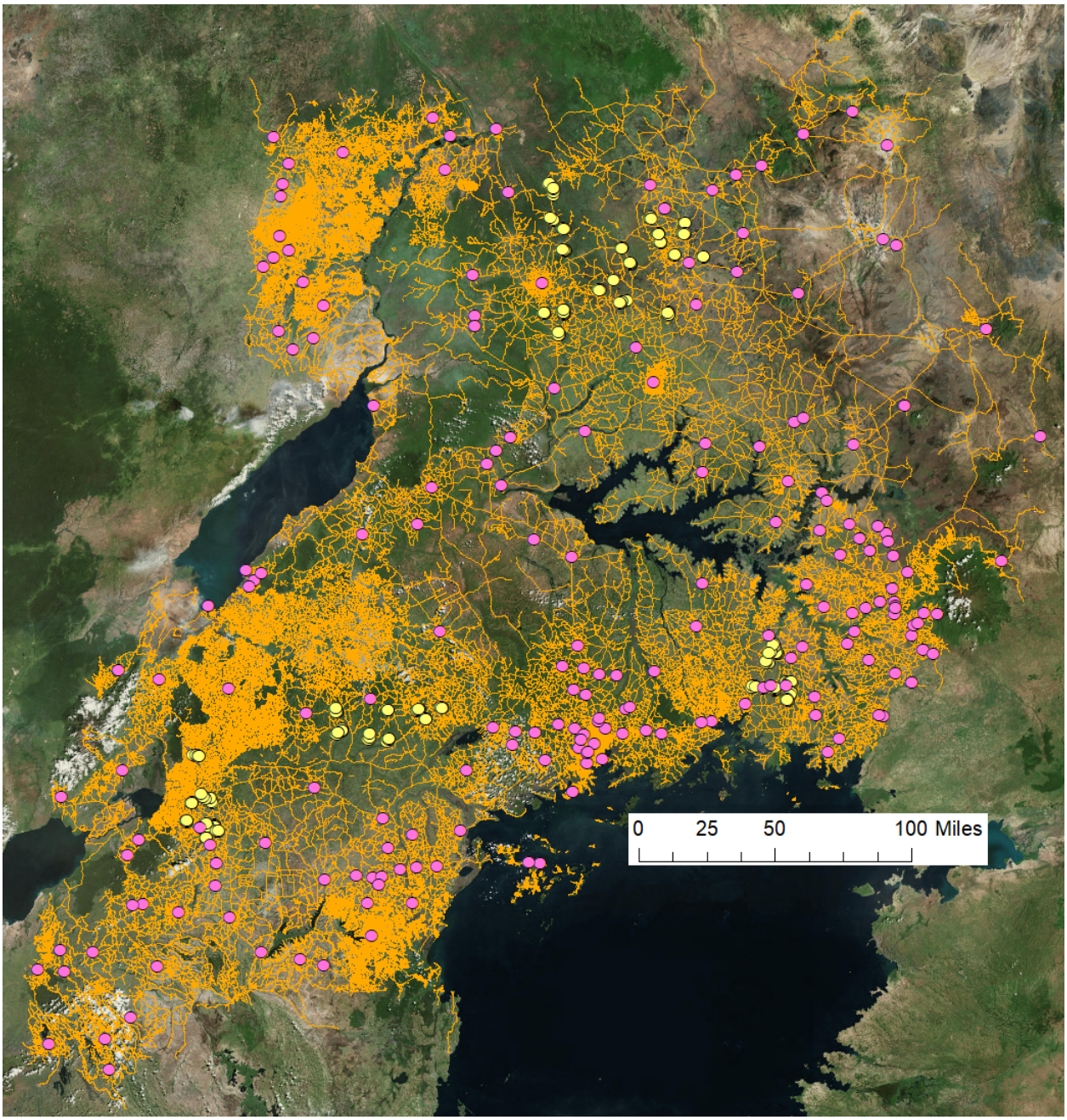

Figure 3-28: The entire OSM road network for Uganda. Respondent locations are shown with yellow markets. Towns and cities are shown with pink markers.

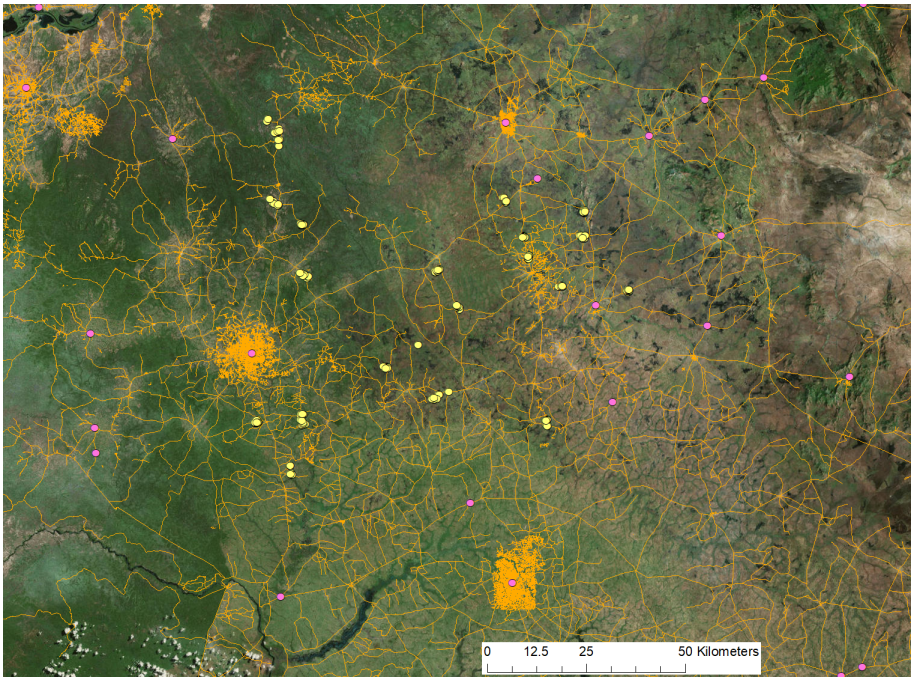

Figure 3-29: A zoomed-in example of the entire road network in relation to towns and respondent households.

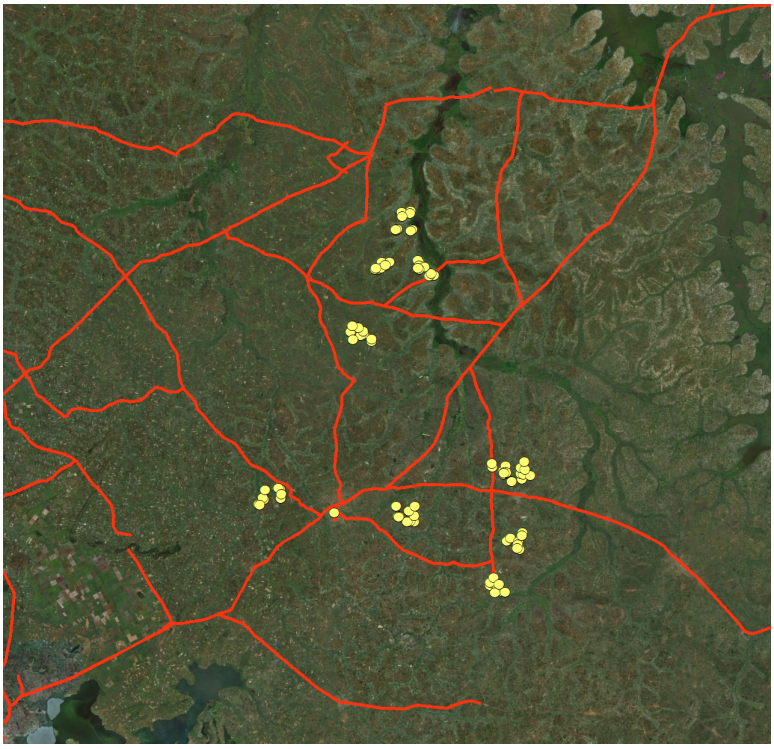

Figure 3-31: An zoomed in example of the road network in relation to households.

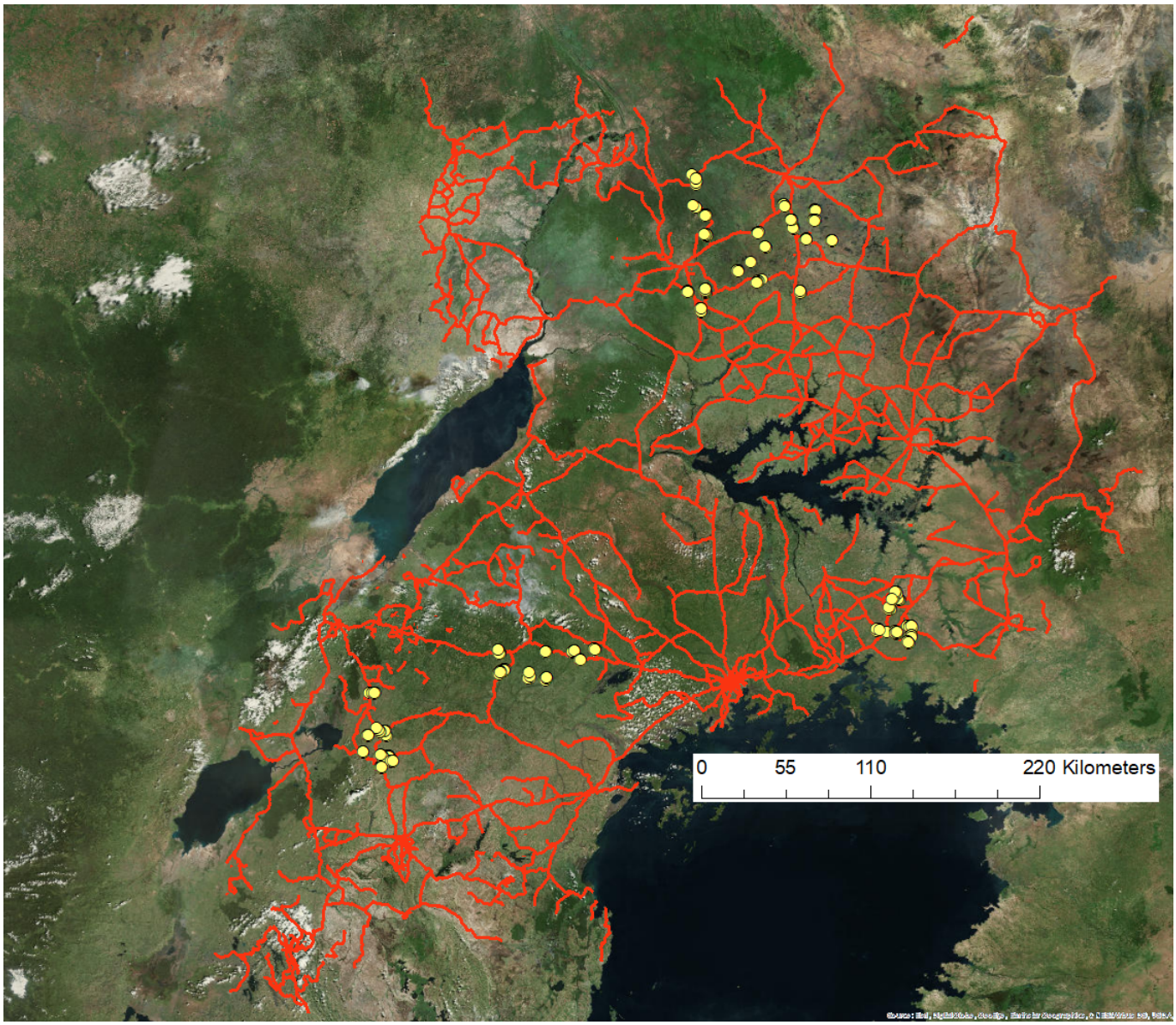

Figure 3-30: The primary and secondary roads of Uganda. Respondent locations are shown with yellow markers.

For each household, the minimum straightline distance to the road network was calculated. The distribution of these distances is shown in Figure 3-32. Half of households live within 1.59 km of a major road.

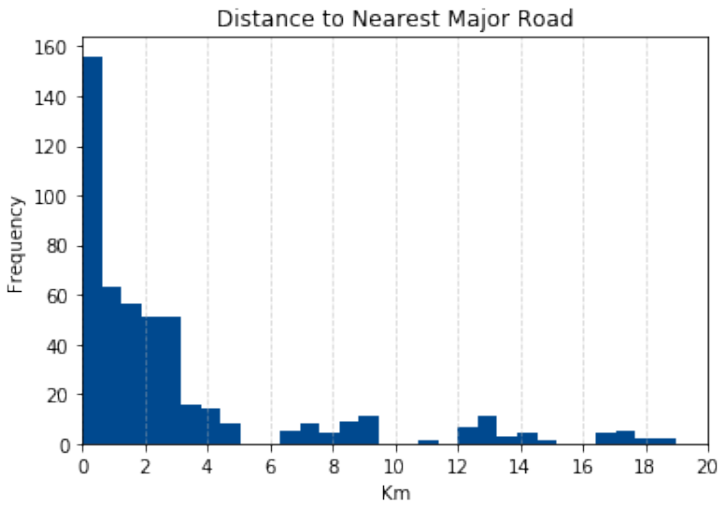

Figure 3-32: Km to nearest major road.

Comparing the distance to reach a major road versus the distance to get to the nearest town, we find a correlation coefficient of just 0.29, suggesting that the measures are not too closely tied. Figure 3-33 shows that no matter how far people must travel to get to town, 85% of them live with 5km of a major road. From this, we can infer that even people who live in more rural areas (further from the nearest town) tend to have access to transportation infrastructure.

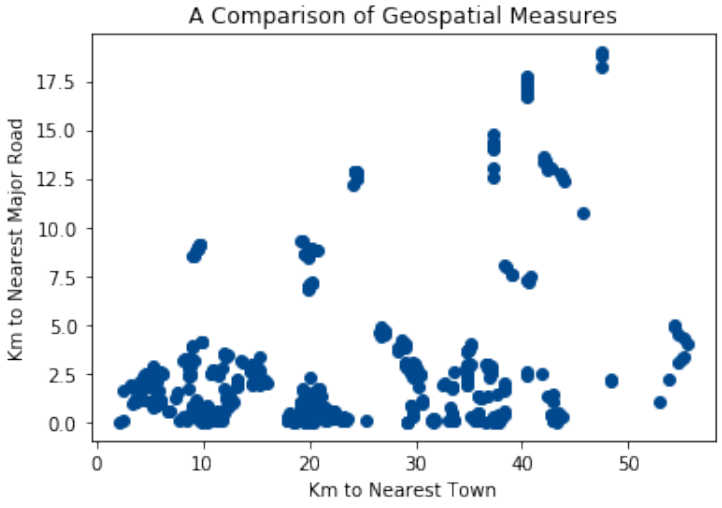

Figure 3-33: Km to nearest town vs km to nearest major road.

3.7.3 Population of Nearest Town

Throughout preliminary analysis, a pattern began to emerge that showed that many of the nearest towns of selling households were known larger market centers. It was hypothesized that the more established and developed market infrastructure of larger towns could be enabling more households to sell. To study this effect, the population for the nearest town to each household was found. We had already extracted the town nearest to each household, according to the OpenStreetMaps dataset. Then, data from the 2014 Ugandan census was cross-referenced to find the population of each of these towns. Table 3.1 details this information.

Town	Population
Gulu	149,802*
Mityana	96,075*
Mubende	95,416*
Iganga	53,870†
Ibanda	31,316†
Kyenjojo,	23,467 †
Kyegegwa	18,729†
Kaliro	16,796†
Busembatia	14,431†
Pader	14,080†
Nakalama	13,010‡
Oyam	11,857 †
Kole	8,833 †
Busesa (Ibaako)	6,093‡
Igorora	5,863†
Achol-Pii	3,417‡
Atiak (Pungole)	2,224‡
Latayi (Latigi)	2,290‡

Table 3.1: Numbers marked with * come from Table 2.6 of the UBOS 2014 NPHC Main Report (Uganda Bureau of Statistics, 2016). Numbers marked with † come from Appendix 3 of the Provisional Results of the National Population and Housing Survey 2014 (Uganda Bureau of Statistics, 2014a). Numbers marked with ‡ come from tables published by UBOS detailing the population of parishes (Uganda Bureau of Statistics, 2014b, 2014c, 2014d, 2014e).

Chapter 4

Regression Modeling

Per our research questions from Chapter 1, we wanted to know what characteristics separate those who sell their harvest and those who don't. There is a confluence of market incentives, infrastructure, and assets for households that sell their crops. "Market engagement" is equivalent to "market participation," and the two terms will be used here interchangeably. Note that we are interested solely in market engagement on the outputs side. That is, does a household sell the harvest that they grow? As previously discussed, we are interested in understanding the characteristics that contribute to households earning income from agriculture. Our scope for market engagement is limited here to the outputs value chains, and is not a function of whether a household purchases food or inputs. Future analyses may investigate how participation in the inputs market overlaps with participation in the outputs market. Market engagement as a concept cannot be measured, and there are many different ways of framing it. This chapter explores several different framings using regression analyses.

First, Section 4.1 will examine household characteristics related to the decision to participate in the market at all in order to answer our first research question: **What factors impact farmers' decisions to participate in the market?** We re-create a model from the literature to see if its results apply to the Ugandan context. Then, we develop our own multiple logistic regression model for our dataset. Focusing only on the households which did participate, Section 4.2 then seeks to understand the dynamics underlying the level of intensity of that participation in order to answer the second research question: **Conditional on participation, what factors impact the level of market participation?**. We discuss two different metrics for this concept and create multiple linear regression models for each one.

While some models in the literature use regional prices as an endogenous variable (Alene et al., 2008; Goetz, 1992; Heltberg & Tarp, 2002; Lifeyo, 2017), we could not incorporate that here. Appropriate price

data for the regions and time frame could not be gathered. For each instance in which a household sold a crop, they reported the price that they received. There is wild variation in these numbers, reflecting the decentralized nature of these rural markets. While regional prices may serve as incentives, the reality of what price a household received is not an independent variable. Households without transportation may have a buyer come pickup their crop for a lower price. Those with increased access to market information may be able to seek out the best prices. It is possible that the relative market price in general motivates households, but for these analyses we examine alternate factors.

4.1 Decision to Participate in the Market

The most basic way of defining market engagement is whether a household sells any of their harvest at any point during a given time period. This binary model is commonly used in market engagement analyses as discussed in Chapter 1 (Alene et al., 2008; Goetz, 1992; Heltberg & Tarp, 2002; Key, Sadoulet, & de Janvry, 2000; Lifeyo, 2017; Omiti et al., 2009; Vance & Geoghegan, 2004, 3; Yusuf et al., 2015). Those classified as "engaged" have sold a crop to some buyer, regardless of the amount of the crop sold. Those who are "not engaged" do not sell their crops.

In the household survey, questions were asked about the two most recent harvest seasons: June-July 2017 and November-December 2017. Respondents were asked to choose from a list of crops indicating what their household grew. For each crop they indicated, they were asked about the acreage planted. For the three crops that were reported with the highest acreage planted, respondents were asked if they sold any of the harvest or not. Across the two seasons and the three top crops in each season for each respondent, we have a total of six datapoints on selling a crop or not. Therefore, a household is classified as engaged if they answered "Yes" at least once for the six datapoints about selling their harvest. A household that sells large amounts from all of its top three crops in both seasons and a household that sells a small fraction of a single crop in only the second season are on equal footing in this model. No overarching question was asked about whether any crop outside of the top three was sold. If there are households which do not sell anything from their top three crops by acreage but do sell from their other crops, they will be classified as "not engaged" under this scheme. Our classification relies upon the assumption that households which are engaged with the market will plant a larger area of the crops that they sell.

We choose to expand our definition of market participation beyond the single crop models used in some of the literature. Observing market participation across all crops is the same approach taken by Heltberg and Tarp (2002). While different crops may involve different dynamics in terms of land needed, preference for home consumption, ability to be stored, and market demand, we are ultimately interested in the overall decision to sell and engage in agriculture as an economic activity. In addition,

examining market participation in aggregate broadly reduces the influence of substitution effects of crops.

We also do not split our sample into buyers, sellers, subsistence, and those who both buy and sell. In our dataset, many households both buy and sell. They may participate as both buyers and sellers because they prioritize the need for cash soon after the harvest but run out of food at a later point. Households may also have a preference for increasing dietary diversity. We are unconcerned with if or why a households chooses to purchase food in this analysis in order to keep the focus on the drivers of commercialization on the outputs side.

Our data set has 498 respondents, 406 who sold any of their top three crops during either season of 2017 and 92 who did not. Of the people who did not sell, 2 did not plant any crops during 2017, 3 did not harvest any of the crops that they planted, and 1 could not remember any of the quantities harvested. These 6 respondents have been left out of this logistic regression model. We want our model to focus on factors which influence decisions to sell crops or not. Households which do not have any harvest to sell implicitly do not have to make this decision. There will be perfect separation between non-selling and non-harvesting. Removing these 6 people, our sample has the following break-down shown in Table 4.1.

Sellers	406
Non-sellers	84
Total	**492**

Table 4.1: The count of sellers and non-sellers within the data.

In analyzing this data, we first explore using a model from the literature as a template in Section 4.1.1. Then, Section 4.1.2 describes the process of building our own model for market participation using domain knowledge of Ugandan farming households in combination with computational methods. Section 4.1.3 discusses the results.

4.1.1 Implementing a Model from the Literature

To understand to what extent our data could be described using models that have already been implemented, models from the literature were reviewed to see if they could be applied to our dataset. Testing a previously created model can provide insight into how results from other developing economies may or may not apply to the market participation dynamics of Ugandan farmers.

Three models that regressed on binary market participation were assessed. Ultimately, all of them were found to include independent variables that our dataset did not contain. For example, Heltberg

Alene Model Factors	**Thesis Model Factors**
Gender of the household head (1 = male)	Gender of respondent (1 = male)
Age of the household head	Age of respondent
Education of the head (6 years or more = 1)	Education of respondent (completed primary school or above = 1)
Hectares of land owned per capita	Hectares of land owned per capita
Adults in the household	Adults in the household
Communication (ownership of radio or telephone; yes = 1)	Mobile phone ownership (yes = 1)
Transport equipment (yes = 1)	Access to private transportation (yes = 1)
Group marketing member (yes = 1)	Member of producer organization (yes = 1)
Zone (Western Province = 1)	District (one-hot encoded)
Livestock ownership (TLU)	Livestock ownership (TLU)
Access to credit for agricultural inputs (yes = 1)	Received credit from a business in the past year (yes = 1)
Off-farm incomes (Ksh)	Quartile bucket of household income from off-farm (ordinal variable)
Distant maize market (>10km = 1)	Above median distance to reach nearest town (>median = 1)
Distant fertilizer market (>15km = 1)	Nearby shop (yes = 1)
Modern maize variety (% adopters/district)	-
Extension coverage (within this household's district, percent of villages with a resident extension agent)	-
Price of maize (Ksh/kg)	-
Pack animals (like a horse or donkey, for transporting goods; yes = 1)	-

Table 4.2: Comparison of the factors used in the Alene model against the most similar factor from our dataset.

and Tarp (2002) use tractor ownership, the prices of various crops, and the risk of an area. Vance and Geoghegan (2004) included agroecological measures such as elevation and soil quality. While it did not completely align with our available variables, the model which examined factors which most closely matched our data came from Alene et al (2008), which studied small-holder participation in maize markets in Kenya.

The dependent variable in the Alene probit model was binary for whether a household sold anything in the maize market or not during 2005. Table 4.2 illustrates the variables that this model used, and lists the variable from our dataset that best align. While the factors do not fully match, there is a large amount of conceptual overlap. The Alene model incorporates the gender, age, and education level of the head of household. For our survey, respondents are not necessarily the head of household, so the gender, age, and education variables from our dataset may contain noise when used as proxies

Animal Type	TLU factor
Cattle	0.7
Sheep	0.1
Goats	0.1
Pigs	0.2
Chickens	0.01

Table 4.3: The conversion factors to calculate a household's tropical livestock units (TLU).

for a household head. The Alene model includes the number of hectares per capita that a household owns in order to capture a household's "potential for production in excess of consumption," but makes no mention of land rental. Therefore, this variable may be capturing the effects of household wealth in addition to production potential. To mirror this, we converted the number of acres of land each household owns to hectares and divided it by the household size. Both datasets include information on the number of adults per household, communication technologies, group marketing participation, and access to transport equipment. To match the control variable for being located in a zone with favorable agro-ecology (western provinces), we use dummy variables for which district a household lives in.

The Alene model aggregates livestock ownership by converting the number and type of animals a household owns into a single metric. However, they do not mention how they calculated this conversion. For our purposes, we use conversion factors from Otte and Chilonda (2002) for livestock in sub-Saharan Africa, shown in Table 4.3

The FMES survey did not specifically ask about access to credit for agricultural inputs, but we do know which households received credit from businesses within the past year. This measure may miss some households who have access to credit but do not take advantage of it and it may over-count households who received credit for something other than agricultural inputs.

The Alene model includes the amount of income that a household earns from off the farm, as measured in Kenyan shillings. While our dataset does not contain information on the actual amount a household earns from off-farm, we do have a rough estimate on the percent of household income earned off-farm. Households indicated whether their off-farm income made up 0%, <25%, 25-50%, 50-75%, or 75-100% of their total household income. Instead of dummy variables, we will use an ordinal variable from 0-4 to indicate which quartile bucket a household falls into.

Alene et al explains that in the Kenyan context, households usually must travel further to obtain fertilizer than to sell their maize. A threshold of 10km is used to determine if a household is far from a maize market and a threshold of 15km it used to determine if a household is far from a fertilizer market. No explanations are offered for these parameters, though the paper mentions that for each

variable, half of the households and near and half are far. This suggests that the median may be serving as the threshold. For our application of this model, we can use the "nearby shop" dummy variable as a proxy for if a fertilizer market is distant or not. The threshold of what constitutes nearby may be different from household to household. While we do have the distance to the nearest town, it is unclear if the 10km threshold fits the definition for "distant" in the Ugandan context. About 25% of Ugandan households are less than 10km from a town; the median distance is 20.2km. Therefore, our version of the Alene model will use a threshold of 20km for a household being classified as distant from a market.

Some of the independent variables used in the Alene model are unavailable in our dataset. Appropriate datasets for the location and number of extension workers, the rate of adoption of modern seed varieties, and the regional prices for crops could not be found for Uganda. In our survey, no households reported owning pack animals, so that variable is removed from our re-creation of this model.

The results of our probit model are shown in Table 4.4.

	Coef.	Std.Err.	z	P>|z|	[0.025	0.975]
Constant	1.4528	0.3602	4.0332	0.0001***	0.7468	2.1588
Respondent Gender	0.1403	0.1631	0.8601	0.3897	-0.1794	0.4599
Respondent Age	-0.0098	0.0051	-1.9395	0.0524*	-0.0197	0.0001
Respondent Education	0.2589	0.1696	1.5266	0.1269	-0.0735	0.5914
Received Credit	0.4639	0.2750	1.6871	0.0916*	-0.0750	1.0028
Nearby Shop	0.2440	0.1498	1.6281	0.1035	-0.0497	0.5377
Hectares Owned Per Capita	0.0434	0.1876	0.2315	0.8169	-0.3242	0.4111
Adults	0.0354	0.0467	0.7570	0.4491	-0.0562	0.1270
TLU	0.0333	0.0421	0.7910	0.4290	-0.0492	0.1158
Off-Farm Income Contribution	-0.0380	0.0650	-0.5850	0.5586	-0.1655	0.0894
Distant Market	-0.2128	0.2000	-1.0642	0.2872	-0.6047	0.1791
Member of Producer Organization	0.4911	0.3178	1.5453	0.1223	-0.1318	1.1139
Mobile Phone Ownership	0.1616	0.1682	0.9608	0.3366	-0.1681	0.4913
Access to Transportation	0.0609	0.1501	0.4058	0.6849	-0.2333	0.3551
District - Gulu	-0.3555	0.2737	-1.2989	0.1940	-0.8919	0.1809
District - Ibanda	-0.3337	0.2829	-1.1798	0.2381	-0.8882	0.2207
District - Iganga	-1.2552	0.2926	-4.2896	0.0000***	-1.8287	-0.6817
District - Pader	-0.6975	0.2750	-2.5363	0.0112**	-1.2365	-0.1585

Table 4.4: The logistic regression model re-created from the literature. *, **, *** represent significance at the 10%, 5%, and 1% levels.

The results find that age of the respondent had a statistically significant effect on decreasing market

participation. Access to credit increased it. Both of these results are consistent with the output of the Alene model. Households in Iganga and Pader had a decreased probability of selling versus Mubende. This is in contrast to the finding in the Alene model that differences in agro-ecological potential of regions did not impact market participation.

The Alene model had found statistically significant impacts for several other variables that were not reflected in our re-creation of the model. They had found that land per capita, the number of adults, and transport equipment were connected with increased market participation. They had also found that increased off-farm income decreased market participation. As our model essentially used ordinal categorical variables to model income, it may be too noisy for an actual impact to come through. It may be that there is an income or percent of income threshold beyond which few households bother selling their crops. The Alene model also found that households more than 15km away from a fertilizer market were less likely to sell, but our proxy for this feature ("nearby shop") has no impact. Finally, the model from the literature found that increased density of extension agents was connected with increased market participation. This is an interesting result with a compelling policy recommendation. However, lack of similar data in Uganda means that our model cannot provide understanding of this factor.

Overall, it is unsurprising that our re-created version of the Alene model offers fuzzier results. Some of our variables do not totally align with their counterparts from the original model and some are completely missing. The Alene version intends to model participation specifically in maize markets, whereas we are attempting to look at participation in markets across crops. There may be factors that are specific to one crop and do not apply to a participation model across a diverse crop portfolio. The differing results may also reflect the differing ways in which transportation networks and market infrastructure are developed in Kenya versus Uganda. However, this exercise can begin to shed light on the dynamics of our dataset. Using transformations and additional factors that are more relevant for our context may bring out the impacts of some of the factors that Alene did find statistically significant. Off-farm income, distance to market, and transportation access were all found to have an impact in this exercise and will be considered for models built from our own data.

4.1.2 Final Model

To get results that may more accurately reflect the Ugandan market participation experience, we develop our own multiple logistic regression model. We believe that some of the features included in the Alene model may not be relevant and there are other additional features that can better explain participation decisions. The model building process is briefly explained below.

For our data, the smaller event category for the dependent variable is non-sellers, with 84 events. Traditional logistic regression modeling has used a rule of thumb indicating that there should be

approximately 10 events (of the smallest event category) per independent variable used in the model in order to keep the risk of over-fitting low (Peduzzi, Concato, Kemper, Holford, & Feinstein, 1996). This restricts us to 8-9 features. However, the work of Vittinghoff and McCulloch (2007) suggest that this may be relaxed to 5-9 events per variable (EPV). In many simulations, they found that models using 5-9 EPV had minimum observed confidence interval coverage and maximum type I error rates similar to models using 10–16 EPV. Based off of this research, it is fine to expand our model to include up to $84/5 \approx 17$ features. The final implementation uses 11 features, with an EPV of 7.6.

In order to determine which features to include in our model, a combination of intuition and computational methods were used. For this process, categorical variables were dummy encoded. Since this regression focuses on the decision to sell or not, all features relating to sellers (such as how much they sold, the price that was received, how they knew the buyer, etc) were removed. Although these factors may influence the decision to sell or not, we do not have any of this counterfactual information for the non-sellers. In addition, some variables were aggregated together. For example, a variable "services sum" was created to sum up the number of different services that a household indicated having access to.

Next, univariate testing was used to determine which independent variables may be related to the dependent variable of selling. As recommended by Hosmer and Lemeshow (2013), each continuous feature was compared against the outcome in univariate logistic regression. The significance of the log-likelihood ratio test of this model is checked. If the p-value is below 0.2, the feature is identified as a candidate for the final model. For each binary variable, Hosmer and Lemeshow recommend checking the significance of a Pearson chi-square test on the contingency table of the outcome variable versus the feature under analysis. Features with a p-value of less than 0.2 for the test's significance are identified as candidates for the final model. The 0.2 threshold is used as an initial screening criterion in order to be conservative and reduce the risk of excluding a variable that is actually important. However, the higher threshold runs the risk of allowing us to include questionable features in the model. The thoughtful application of intuition and subject matter expertise can help to mitigate some of this risk.

From our feature set, 90 were identified as candidates for the final model. This subset is still much larger than the maximum of 17 features our dataset size will allow. Next, intuition was used to choose features from the 90 that may actually be able to explain the dependent variable.

From this purposively chosen subset, we are still unsure which actually impact our dependent variable. Ideally, we want to build a model with the smallest number of features that explain the most variation in the data. To achieve this, an implementation of forward selection and backward elimination was used to construct a model with features that incrementally significantly improved the log-likelihood ratio of the model. Before describing the mechanics of this process, it should be understood that it varies significantly from what is typically associated with "stepwise" regression modeling. Traditional stepwise regression modeling may involve adding a feature to a model and choosing to include it or

remove it based upon the p-value of its coefficient, so that only "significant" features remain. This is a approach can create significantly biased models and is not used here. The process described here relies on checking the significance of the log-likelihood ratio test of a model with a new feature compared against a model without the feature. In that way, it selects for features that provide a significant contribution to improving the fit of the model.

The forward selection and backward elimination algorithm used is based on that described by Hosmer and Lemeshow (2013). First, a model with only a constant is created. From the list of features we may want to include, a model is created for each consisting of that feature and the constant. The p-value of the log-likelihood ratio test between that model and the model with only a constant is computed. The feature with the smallest p-value is added to the model. The process is then repeated, by building a model for each feature left over that has that feature, the constant, and the initially chosen feature. Again, for each of these, the p-value of the log-likelihood ratio test of this model and the previous model is computed. The feature with the smallest p-value that also meets a threshold requirement is chosen to enter the model. After each step where a feature is added to a model, a check must be done to see if any features should be removed from the model. If, in the course of adding in features, one of the older features no longer significantly contributes to the model, it should be removed. To do this, a model is created for each feature already included which removes that feature and checks the significance of the likelihood ratio test for this model against the current model. The features whose exclusion results in a the largest p-value above a threshold for the likelihood ratio test is removed from the model. Whatever variation it was explaining is now covered by the combination of variables included in the model. This process of alternating checking to add variables and then checking to remove them proceeds until no more leftover variables meet the p-value threshold for inclusion and no variables currently in the model meet the p-value threshold for exclusion. Hosmer and Lemeshow recommend a p-value of 0.15 for inclusion and a p-value of 0.20 for exclusion. An alternative way to conclude this stepwise process is by setting a limit on the number of features to include in the model. The final result of this process will be a model that contains only features that have incrementally significantly increased the log-likelihood of our model.

The feature set of this computational model is reviewed as a way to suggest which previously overlooked variables may actually be important. We may wish to include some features that the computational process found did not significantly improve the model fit. However, the non-significant impact of that variable may be an important result. Given a hypothesis that a certain feature impacts the dependent variable in a specific way, that variable should be added back in. A final set of features chosen based on intuition and suggested by the computational model are combined and used to build a multiple logistic regression model.

The resulting model is shown in Table 4.5. Defintions and descriptive statistics for the features used can be found in Chapter 3.

	β	Change in Odds	Std.Err.	z	P>|z|	[0.025	0.975]
Constant	-5.1333		0.8224	-6.2421	0.0000	-6.7451	-3.5215
Log Kg Harvested Per Capita	1.1848	+227.0%	0.1555	7.6179	0.0000***	0.8800	1.4897
Population of Nearest Town in Thousands	0.0132	+ 1.3%	0.0039	3.3442	0.0008***	0.0055	0.0209
Services Sum	0.0984	+10.3%	0.0572	1.7186	0.0857*	-0.0138	0.2106
Nearby Input Shop	0.3536	+42.4%	0.3111	1.1367	0.2557	-0.2561	0.9633
Off-Farm Income 75 - 100% of Total	-1.5674	-79.1%	0.7474	-2.0971	0.0360**	-3.0323	-0.1025
Km to Nearest Town	-0.0043	-0.4%	0.0123	-0.3498	0.7265	-0.0283	0.0197
Km to Nearest Road	0.0254	+2.6%	0.0414	0.6126	0.5402	-0.0558	0.1066
Transportation Access	-0.3825	-31.8%	0.3063	-1.2487	0.2118	-0.9829	0.2179
Food Secure	-0.4024	-33.1%	0.3573	-1.1265	0.2600	-1.1026	0.2978
Mobile Phone Ownership	0.6529	+92.1%	0.2956	2.2085	0.0272**	0.0735	1.2324
Member of Producer Organization	1.4883	+343.0%	0.8191	1.8170	0.0692*	-0.1171	3.0937

Table 4.5: The final logistic regression model. *, **, *** represent significance at the 10%, 5%, and 1% levels.

The final multiple logistic regression model has a pseudo R^2 of 0.306. The chi-squared test of its likelihood ratio test against the null yields a p-value of 1.22e-24, giving us high confidence that this model is better than the null. The Hosmer-Lemeshow test with 10 groupings results in a $\chi^2 = 5.9283$ with a p-value = 0.6553, indicating that we do not reject the null that the model fits well. Even varying the number of groupings used for the Hosmer-Lemeshow test yields robust results for good fit. With 9 groupings, we find $\chi^2 = 2.9799$ with a p-value = 0.8869. With 11 groupings, we find $\chi^2 = 9.2953$ with a p-value = 0.4105. With 12 groupings, we find $\chi^2 = 4.5544$ with a p-value = 0.9189.

For this model, the apparent AUC (area under the receiver operating curve) is 0.86.

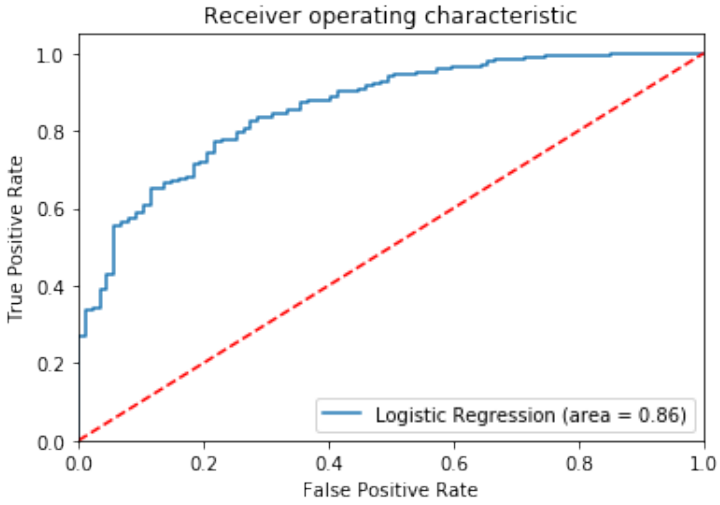

Figure 4-1: The receiving operator curve for the logistic regression model.

4.1.3 Discussion

Log Kg Harvested Per Capita: The level of production per household member is statistically significant at the 1% level. An increase in the mass of the harvest per household member leads to an increase in the likelihood of selling. The more food a family has on hand per person, the more it may be willing to allocate that food away from consumption and toward earnings in the market. Note that this independent variable represents the aggregated kg of the entire mix of crops produced. Although the differing yield and weight of various crops can result in differences across households with different crop portfolios, we believe that the total amount harvested still provides a useful metric of overall production. For this analysis, we have taken the log in order to transform an initially skewed distribution to approximately normal. Therefore, the coefficient must be interpreted in terms of the effect of a one-unit increase in the the log of kg harvested on the log-odds of selling.

Population of Nearest Town in Thousands: At the 1% level, households whose closest town is more populous are more likely to participate in markets. More populous towns will have more traders, more opportunities for transport, and better developed market system infrastructure. The greater number of available traders may mean that farmers can access more competitive and favorable prices for their products. The significance of this variable and its positive coefficient help inform our understanding that more developed markets are linked to market participation. Market development may be an important enabler of participation.

Services Sum: The results show that the more services a household reports are available, the greater their probability of participating in the market. Note that this sum is not necessarily the services that a household took advantage of, but simply that there are people nearby who could perform them if a household was willing to purchase them. Therefore, this measure can be a proxy for the strength of the agricultural system infrastructure around a households. We find that it has a positive impact that is significant at the 10% level. This underscores our finding above about larger towns, showing that more developed markets are linked to more market participation. Market development may be an important enabler of participation.

Nearby Input Shop: The dummy variable for whether or not a household had an inputs shop nearby was not significant in contributing to market participation. This can be another measure of market development as well as market access as it reflects whether or not inputs are both available and accessible. The fact that this does not significantly impact market participation may be a result of market development being encapsulated by the "services sum" variable while market access is captured by the geospatial measures discussed below.

Off-Farm Income 75 - 100% of Total: Having a job or small business off-farm that is responsible for 75 - 100% of a household's income had a significant effect at the 5% level on decreasing a household's market participation. They may use what they grow exclusively for household consumption since their need for cash liquidity is met through alternate economic activities. This barrier does not necessarily require the attention of market facilitation efforts; not all households are interested in commercializing agriculture.

Km to Nearest Town: Interestingly, we find that the distance that must be traveled along roads to reach the nearest town has no impact on market participation. Distance to market and the cost of transportation are not a barrier to selling; people are motivated to sell their crop regardless of how far they are from market centers. This may suggest that there are intermediary points of selling that are not associated with town markets.

Km to Nearest Major Road: Again we find that geospatial characteristics do not impact market participation. This variable is the straight-line km from a household to the nearest major road and can be a proxy for the accessibility of transportation infrastructure. This is not a barrier impacting households' decisions to sell their crops.

Transportation Access: We find that households which report having private transportation access (ie, bicycle, motorcycle, truck, or car) do not have an advantage in participating in markets. Similar to the geospatial measures, this dummy variable reflects an element of market access. Within the Ugandan context, transportation does not serve as an enabler to market engagement. Conversely, households without access to transportation are not excluded from participating in markets.

Food Secure: Households which never skip meals are not more likely to participate in markets than

others. This is an unexpected result that will be explored further in Chapter 5. Instead of eschewing selling in order to eat what they produce, food insecure households still participate in markets.

Mobile Phone Ownership: At the 5% level, households which own a mobile phone are more likely to participate in markets. A phone is a technology which can strengthen a household's connection to the market and ability to access information. It acts as an enabler by reducing the variable costs associated with traveling outside of the home to learn about market prices, available buyers, agricultural products, and transportation opportunities.

Member of Producer Organization: At the 10% level, membership in a producer organization yields a statistically significant increase in a household's likelihood of participating in the market. These groups serve as enablers by providing valuable leadership and organization to small-holder farming communities. Collective marketing can yield more favorable prices from buyers, lower transportation costs, and increased opportunities to learn about agricultural trainings, techniques, and products. In addition to reductions in transaction costs, as discussed in Section 2.1.6, these groups also expand a farmer's social network and opportunity for market-oriented relationships (Bosc et al., 2002). All of these benefits contribute to membership in a producer organization serving as an enabler to market participation.

4.2 Level of Market Participation

Beyond the binary choice of whether or not to participate, another way of thinking about market engagement is the level of engagement of a household. We are interested in understanding what household characteristics are connected to an increased intensity of market participation. Once we understand some of the factors that push/pull households to markets, what can increase how intensely they participate? An increase in the level of market participation will yield an increase in revenue generated from agriculture. To understand some of these factors, we build two linear regression models around two different ways of quantifying the level of engagement. Beyond seeing what factors these models deem significant, we are interested in comparing and contrasting them to gain insight into the ways different metrics for level of engagement may be connected with different factors. Does looking at different metrics suggest different market facilitation strategies be used? If so, how can those strategies be compared and prioritized?

For the two metrics examined, we perform a mono-crop analysis over a single season. Narrowing the focus of the model to one crop can better allow for a better comparison across households. Different crops will have different densities, so it is challenging to compare the farmer who sells 100kg of beans and the one who sells 100kg of cassava. To ensure a consistent comparison across households, we focus on maize sold from the November - December 2017 season. Maize was the crop grown by the

largest share of households, with 351 (70.5%) respondents cultivating it in November - December 2017. This is slightly higher than the 335 (67.3%) households which cultivated it in June - July 2017, so we focus on the former season in order to have a larger dataset. As the country's top crop grown for consumption, maize is versatile in terms of home consumption and selling on well-developed markets (FAPDA, 2015).

Of the households that cultivated maize in November - December 2017, 180 reported selling it. Our analyses of level of market participation will focus only on this subset that participated in the maize market.

In order to better compare the two different metrics of market participation, the same set of independent variables will be used for both models. To understand what features impact the amount a farmers sells, a subset of features was first chosen using subject matter knowledge. This included many of the factors from the logistic regression model for market participation. We used these factors from the previous model in order to inform our understanding of how the decision to participate and level of participation may have similar or different barriers and enablers. Intuition, common sense, and the literature also informed this initial set.

Separate from this list, a stepwise procedure was run on the entire variable set to explore a wider range of factors that may have been overlooked. To do this, a procedure similar to the one described in Section 4.1.2 was used. Forward selection and backward elimination for a linear regression model was performed using the Akaike information criterion (AIC) as the metric for inclusion or exclusion. Models with separate features added or subtracted were compared to determine if features could lower the AIC. An in-depth description of this technique is found in Chatterjee and Hadi (2006, chapter 11). This process is used to help refine our focus on which features to consider for inclusion but we do not use the output as our final model.

Each of the features from the resulting stepwise model was considered for inclusion in the final model. Ultimately, the final model was created by critically analyzing the set of computationally chosen features and the set of features chosen by intuition in order to decide which we were most interested in for their ability to describe the household experience and connect to market facilitation. Table 4.6 summarizes the variables used in the models. Further definitions and descriptive statistics for the features discussed below can be found in Chapter 3.

Beyond the features chosen for these models, there are likely exogenous market factors impacting household selling decisions. Maize grain prices in November 2017 were 12% lower than November 2016 and maize flour prices in November 2017 were 6% lower than November 2016 (WFP, 2017). Regional market trends may have an impact on market participation that we are unable to uncover by focusing on a single season.

Section 4.2.1 first discusses the model for market participation intensity using the log of the amount of

Variable	Average Value
Logit of Portion Sold	1.06
Log Kg Maize Sold	5.70
Log Kg Maize Harvested	6.22
Log Kg Other Harvested	6.74
Household Size	6.40
Population Nearest Town in Thousands	49.17
Services Sum	5.09
Nearby Input Shop	0.49
Off-Farm Income 75-100% of Total	0.02
Km to Nearest Town	23.52
Km to Nearest Road	3.16
Transportation Access	0.60
Pickup	0.13
Food Secure	0.81
Mobile Phone Ownership	0.77
Member of Producer Organization	0.06

Table 4.6: The averages for the features used in the multiple linear regression models. The first two (Logit of Portion Sold and Log Kg Maize Sold) are used as dependent variables. These averages are for the subset of households which sold maize in November - December 2017.

maize sold as the dependent variable. Section 4.2.2 then presents a model using the logit of the percent of maize sold as the dependent variable. Section 4.2.3 discusses and compares the results for these two models.

4.2.1 Amount Sold

In this analysis, we quantify level of market engagement over a continuous range by examining the total amount of maize a household sold in a single season. The continuous range of amount sold will allow for a more nuanced exploration of the factors that impact market engagement. It is intuitive that there is some difference between a farmer who sells massive quantities and a farmer who sells a small amount. What influences the success of the first farmer? Comparing this continuous model to the binary one above, are there some enablers that are "gates" in that they must be present for selling, but the extent and magnitude does not matter? Are there some enablers that work on a scale? As in, do farmers who have a linear quantity of something linearly sell more of their harvest?

	β	Std.Err.	z	P>|z|	[0.025	0.975]
Constant	-0.9365	0.240	-3.900	0.000***	-1.411	-0.462
Log Maize Harvested	1.0676	0.041	26.102	0.000***	0.987	1.148
Log Other Harvested	-0.0228	0.033	-0.700	0.485	-0.087	0.042
Household Size	-0.0141	0.012	-1.176	0.241	-0.038	0.010
Population of Nearest Town in Thousands	0.0008	0.001	0.872	0.384	-0.001	0.003
Services Sum	0.0049	0.013	0.381	0.704	-0.021	0.030
Nearby Input Shop	-0.0319	0.075	-0.427	0.670	-0.179	0.115
Off-Farm Income 75 - 100% of Total	-0.1799	0.250	-0.720	0.473	-0.674	0.314
Km to Nearest Town	0.0063	0.003	1.845	0.067*	-0.000	0.013
Km to Nearest Road	-0.0250	0.009	-2.794	0.006***	-0.043	-0.007
Transportation Access	0.0099	0.081	0.122	0.903	-0.150	0.170
Pickup	0.0829	0.112	0.737	0.462	-0.139	0.305
Food Secure	0.0425	0.093	0.456	0.649	-0.141	0.226
Mobile Phone Ownership	0.1149	0.089	1.290	0.199	-0.061	0.291
Member of Producer Organization	-0.0620	0.154	-0.402	0.688	-0.367	0.243

Table 4.7: The final linear regression model using log(kg maize sold) as the dependent variable. *, **, *** represent significance at the 10%, 5%, and 1% levels. $R^2 = 0.884$; $N = 180$

Our model yields an R^2 of 0.884, indicating that it explains much of the variation in the data. Plotting the actual values against the predictions yield a clear linear pattern. Much of the explanatory power comes from the inclusion of the log of maize harvested. The residuals do not show any pattern and the Breusch-Pagan test yields a high p-value (0.970), indicating homoskedasticity.

4.2.2 Percent Sold

The second way of measuring market engagement that we will explore is the percent of a harvest that a household sells. More commercially oriented households will sell a higher percent, while those who may mostly eat what they grow but have some cash needs will sell a lower percent.

A portion bounded between 0 and 1 poses issues for a fitting an unbounded linear regression model. To mitigate this issue, we take the logit of the percent sold and perform multiple linear regression on that dependent variable. Of the 180 maize sellers, 21 of them sold 100% of their harvest. We adjust this value to 99% in order to be able to take the logit. Models were built replacing 100% with values from 95% to 99.9% to test how choosing this value impacted the output and found only small changes in the coefficient values and no changes in the significance of independent variables.

	β	Std.Err.	z	P>|z|	[0.025	0.975]
Constant	-0.0050	0.815	-0.006	0.995	-1.614	1.604
Log Maize Harvested	0.3001	0.139	2.161	0.032**	0.026	0.574
Log Other Harvested	-0.1704	0.111	-1.538	0.126	-0.389	0.048
Household Size	-0.0523	0.041	-1.282	0.202	-0.133	0.028
Population of Nearest Town in Thousands	0.0023	0.003	0.720	0.473	-0.004	0.009
Services Sum	0.0422	0.044	0.961	0.338	-0.044	0.129
Nearby Input Shop	-0.2756	0.253	-1.089	0.278	-0.775	0.224
Off-Farm Income 75 - 100% of Total	-0.9619	0.849	-1.133	0.259	-2.638	0.714
Km to Nearest Town	0.0244	0.012	2.119	0.036**	0.002	0.047
Km to Nearest Road	-0.0861	0.030	-2.838	0.005***	-0.146	-0.026
Food Secure	0.1748	0.316	0.553	0.581	-0.449	0.799
Transportation Access	-0.0667	0.276	-0.242	0.809	-0.611	0.478
Pickup	-0.1110	0.381	-0.291	0.771	-0.864	0.642
Mobile Phone Ownership	0.2056	0.302	0.680	0.498	-0.391	0.802
Member of Producer Organization	-0.5077	0.524	-0.970	0.334	-1.541	0.526

Table 4.8: The final linear regression model using the logit of the percent of maize sold as the dependent variable. *, **, *** represent significance at the 10%, 5%, and 1% levels. $R^2 = 0.162$; $N = 180$

Our model yields an R^2 of 0.162, indicating that there is a large amount of unexplained variation. Plotting the actual values against the predictions, a pattern barely emerges. This model is not a great fit. The residuals do not seem to show any pattern and the Breusch-Pagan test yields a high p-value (0.326), indicating homoskedasticity. There are factors driving these decisions that we are not able to account for in this analysis.

4.2.3 Discussion

Comparing the model for percent sold to the model for the amount sold, it is interesting that the percent sold model fits so poorly while the other one fits well. One reason may be that fitting models for a bounded range (a proportion from 0 to 1) can be challenging, especially when there are samples which take on one of the limiting values (21 households sell 100% of their harvest). In addition, there may be complex decision making processes behind the portion sold that are not captured by the factors in this model. For example, we see that 31 of the 180 maize sellers (17%) sell exactly 50% of their crop, making it the mode for the percent sold data. Of these households, 16 harvested 200kg, while the others harvested between 80kg and 1600kg. These households are modeled on the same dependent variable in the percent model, even though we expect that they experience very different decision

making factors.

Between the two models for level of market participation, results are consistent. Both models find that the only factors which have a statistically significant impact on increasing level of market engagement are an increase in the amount of maize harvested, a decrease in the distance to the nearest primary road, and an increase in the distance traveled to town. The two different metrics are in complete agreement in identifying barriers and enablers. This concurrence suggests that, should USAID choose to monitor the level of market participation as an indicator of the success of market facilitation, they may find similar results if they choose either of these metrics. In terms of ease of data collection, however, it would make the most sense to gather information on the amount that a household sells. The portion that a household sells requires a second datapoint for the amount that they harvested.

This consistency across these two models should be tested using other types of crops to see if the results hold. Future work could build a model with revenue as the target variable to compare for differences in the relevant factors. In addition, a model could also be built for acres planted, as was used as a proxy for level of market engagement by Vance and Geoghegan (2004). We expect that this model may differ from the others, as it is less direct in reflecting market transactions and may also capture consumption needs.

Log Maize Harvested: Households which harvested more maize tended to sell a larger portion and amount, a finding that is statistically significant at the 1% level. Increased harvest also played a role in the decision to participate in the first place. The consistent importance of this variable suggests that increasing production is crucial to increasing market participation.

Log Other Harvested: The amount of all other crops was included to control for crop substitution dynamics. It may be that households which grew other crops sold those instead of the maize. However, the results do not show any effect.

Household Size: There was no statistically significant impact of household size on the portion or amount of the harvest sold. This was an unexpected result. We expected that, all else being equal, households with more members would sell less than others because they keep more for household consumption. The fact that this is not the case may suggest that households are substituting other crops for consumption in the place of maize.

Population of Nearest Town in Thousands: The size of the nearest town does not impact the amount of maize that households sell, suggesting that level of market development plays a reduced role in level of market engagement. Results from Section 4.1 showed that this factor was significant in encouraging the decision to participate in the market. The level of development of markets encourage participation, but once someone is selling, further development does not impact the amount or portion sold.

Services Sum: The number of services available to a household did not impact the level of intensity of

engagement, even though it influenced the decision to sell. Consistent with the result for the population of the nearest town, the level of market development seems to encourage market participation but has no impact on the level of that participation. It is purely an enabler to selling versus not.

Nearby Input Shop: Having a nearby input shop does not impact the amount of the harvest sold. This suggests that connectedness and level of development do not play a large role in the level of market engagement.

Off-Farm Income 75 - 100% of Total: Employment off the farm that makes up the large majority of a household's cash income did not impact the amount or percent of maize that a household sells. Our model for market participation found that this was a significant variable. Having a large share of income from off the farm may be a "gated" enabler. The fixed transaction costs associated with selling may not be worth it to some households with other economic activities, but once those fixed costs are met, the level of participation is unaffected by outside income.

Km to Nearest Road: A statistically significant effect at the 1% level was found that households closer to roads sell a greater amount and portion of their maize harvest. This result may suggest that the level of effort to transport bags of maize along footpaths and smaller, rougher roads limits the amount that a household chooses to sell. The hassle involved in traversing these very rural transportation networks before hitting a primary road may influence households to eat what they have grown instead of making multiple trips to sell more.

Km to Nearest Town: The distance that a household must travel along the road network to reach the nearest town has a statistically significant effect in both models. Unexpectedly, the coefficient for this variable has a positive sign, implying that the further a household must travel into town, the more they will sell. This result could suggest that households which live further away need to sell more in order to offset transportation costs, or that those who are closer to town may have additional employment options.

Transportation Access: Having access to a bicycle, motorcycle, car, or truck does not impact the amount or portion of maize that a household chooses to sell.

Pickup: Whether or not the buyer picked up the harvest did not have a statistically significant impact on the portion or amount that was sold. We were interested to see if structuring relationships with buyers to be more convenient could increase the level of intensity of market engagement. However, this does not seem to be the case.

Food Secure: Household food security had no impact on the level of market engagement in either model. This is an unexpected result. We saw that the decision to participate was not impacted by food security, suggesting that food insecure households may sell their harvest when they have cash needs. However, it could be expected that food insecure households which sell their crop limit it to a smaller

portion and a smaller amount. This is not what we see reflected in the model. Controlling for the amount harvested, a food insecure household does not sell less than a food secure household. There may be substitution effects in which the food insecure households consume crops other than maize. It may also be the case that the cash needs of these households are so high that they must have an increased level of participation. Further research should explore the dynamics behind the marketing decisions of food insecure households in order to better understand the experience of these vulnerable families and how to better serve them with interventions.

Mobile Phone Ownership: Ownership of a mobile phone did not impact the level of market engagement, even though it does play a role in the decision to participate in the first place.

Member of Producer Organization: Membership in a producer organization does not impact the level of market engagement using either metric, even though it was important to the decision to participate.

Chapter 5

Food Security

5.1 Background

The World Food Summit in 1996 declared that food security occurs when "when all people, at all times, have physical and economic access to sufficient, safe and nutritious food that meets their dietary needs and food preferences for an active and healthy life." This encompasses dimensions of availability, accessibility, affordability, and consistency. At its most basic level, food security describes the risk of people experiencing hunger. In Uganda, food insecurity at the household level is widespread (Uganda Bureau of Statistics, 2018; FAO, 2017). As part of our analysis of the dynamics behind market participation, we wanted to understand the role of household food security. We do not establish a causal relationship, but seek to elucidate how hunger intersects with selling crops. As USAID develops programming, they make efforts so that interventions reach a wide audience and do not exclude the vulnerable or extremely poor. As stated in Chapter 1, **Is there a connection between food security and market participation that should inform the way these efforts are directed?**

To quantify experiences with food insecurity, our survey asked households how often anyone had to skip meals because there was not enough food. Skipping meals is a manifest result of food insecurity (FAO, IFAD, UNICEF, WFP, & WHO, 2018). A household may not have enough money to afford food or alternative food sources may not be accessible.

Households were presented with the following options to best describe the frequency of skipping meals:

- Never

- A few times per year
- Once a month
- Once a week
- Other [1]

The number of households within each category is as follows:

Frequency	# Households
Never	397
A few times per year	64
Once a month	13
Once a week	24

Table 5.1

We took this information and then mapped it to the widely used Household Hunger Scale, developed by the FAO (Ballard, Coates, Swindale, & Deitchler, 2011). The survey that we used does not totally align with that used by the official scale, but the categories used for grouping households are comparable. Under this scheme, households which never skip meals are categorized as "food secure." Households which skip meals only a few times per year are categorized as "moderately food insecure." Those which skip meals with a frequency of once a month or more often are "severely food insecure." In addition, for the analysis presented below, only 459 of the 498 respondents provided acceptable data. The number of remaining respondents in each food security category is shown in Table 5.2.

Level of Food Security	# Households
Food Secure	369
Moderately Food Insecure	60
Severely Food Insecure	30

Table 5.2

There are large differences in the sizes of the groups which we are comparing, but we are careful to employ statistical tests which can accommodate this. We are confident that our samples are large enough to inform our understanding of the general population. It should be noted that this is a fairly coarse measurement of food insecurity. In this sample, we see about 20% of households are food insecure, which contrasts to the UNHS's estimate that 40% of households experience food poverty

[1] Only 6 households choose "Other." They provided a short answer to describe what this frequency was. All of these short answers were appropriately matched to the options given above.

(Uganda Bureau of Statistics, 2018). Although most households never skip meals, they may eat smaller portions or have limited dietary diversity. Frequency of skipping meals is only one component of understanding food security. Given constraints in time and complexity in administering an already long survey, asking about skipping meals was a way to get a general idea of household experiences with food security.

The purpose of this analysis is to better understand the intersection of food security and market participation. First, we looked at the likelihood of selling any part of the harvest across the meal-skipping frequencies. Figure 5-1 shows that there is not a large difference across this measure. Fisher's exact test was used to compare each level of food security against the others to determine if the difference in proportions of households engaging in the market differs. None of these differences are statistically significant. We can say that market participation does not differ significantly with food security. This result is surprising. At first blush, it would seem as if households should first meet their consumption needs, and then sell any leftover harvest. This logic suggests that there should be a steady decline in market participation as food insecurity increases. However, this is not what Figure 5-1 suggests. Our models in Chapter 4 investigating the decision to participate in markets and the level of market participation found that food security had no impact. This chapter further explores characteristics linked to food security to better understand this result.

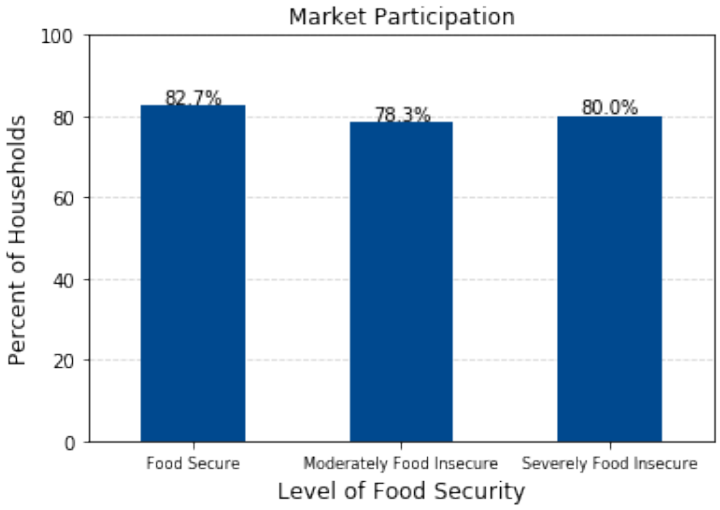

Figure 5-1: The percent of households which sold some of their harvest, as grouped by level of food security.

5.2 Analysis

In addition to asking about the frequency of skipping meals, households were also asked about their harvest, selling patterns, and spending on food in the previous year. For each month, they were asked if they purchased food, what quantity was purchased, and how much it cost. We organize our analysis by different broad categories that contribute to food security: consumption need, production level, market participation, and outside food. For each category, we explore several factors related to it to paint a robust and comprehensive picture of how they may relate to food security.

For all of the variables discussed below, a Wilcoxon–Mann–Whitney two-sample rank-sum test was used to compare the distributions across the categories of food security. Each category was compared against the others to see which had differences in the medians. The shape of the distributions for variables are assumed to be the same across level of food security.

For each factor below, a table is shown with the median value for each level of food security. Boxplots demonstrate the distribution of the factor for each level of food security. A table is presented in which the number in each cell is the median of the cell's column minus the median of the cell's row. The number will always be positive and will always represent how much higher the column category is than the row. A table in which the numbers are concentrated in the lower left corner illustrates a variable which food secure households have more of. A table in which the numbers are concentrated in the upper right corner illustrates a variable which food insecure households have more of. The parentheses contain the p-value for the results of the Wilcoxon–Mann–Whitney test comparing the column as having a greater median than the row. Cells for which the median difference is negative (as calculated by the median of the column minus the median of the row) are denoted with "-" as they cannot have significant results.

5.2.1 Consumption Need

One part of the food security equation is how much food a household needs to avoid hunger. This level of consumption need is primarily set by the number of people living within a household. It may be influenced by other factors not explored here, such as age and level of activity of household members.

Household Size

We look across the categories of food security to see if they have different levels of consumption need as determined by the number of people in the household. The results of the tests show that there is

no difference in the median household size across different levels of household food security. This demonstrates that all households will approximately have the same level of consumption need. The fact that they differ in food security suggests that they may thus differ in their ability to meet that need.

Level of Food Security	Median Household Size
Food Secure	6
Moderately Food Insecure	6
Severely Food Insecure	6.5

Table 5.3: The median households size in each category.

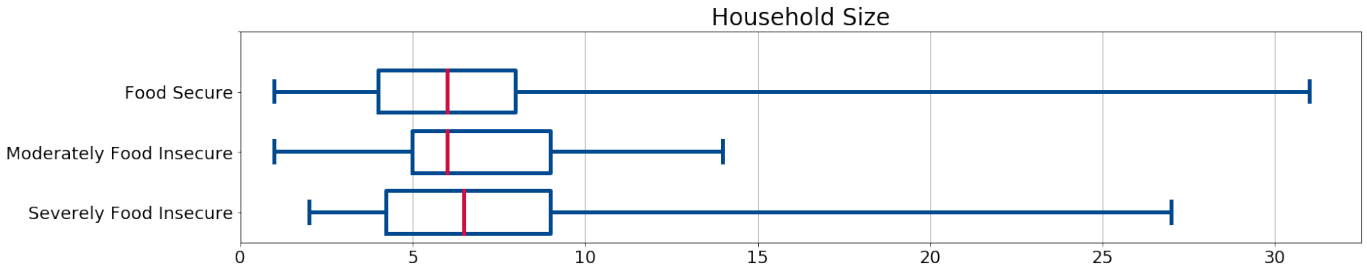

Figure 5-2: The distribution of household size. Households are grouped by their reported frequency of skipping meals.

	Food Secure	Moderately Food Insecure	Severely Food Insecure
Food Secure	-	0 (0.242)	0.5 (0.249)
Moderately Food Insecure	-	-	0.5 (0.427)
Severely Food Insecure	-	-	-

Table 5.4: The difference in median household size, calculated using the median of the category in the column minus the median of the category of the row. In parentheses are the p-values testing whether the median for the category in the column is greater than the median for the category in a row. *, **, *** represent significance at 10%, 5%, and 1% probability levels. There are no differences in the median household size across frequencies of skipping meals.

5.2.2 Production Level

Another part of the food security equation is how much food a household is growing themselves and thus able to eat. We compare these factors to see if increasing food security is associated with increasing

production level, as reflected by the number of acres planted, the amount harvested, and yield. The results shown below demonstrate that food secure households plant more and harvest more than other households. Severely food insecure households achieve lower yields that other households. Production level is strongly tied to level of food security. The results here were found to hold when comparing factors on a per capita basis, which is unsurprising given that the previous section found no differences in household size across level of food security.

Acres Planted

The total number of acres planted for each crop were summed for both seasons in 2017, to yield the total area cultivated by a household in the year. The results of the tests show that households which never skip meals have a higher median number of acres cultivated than others. Table 5.6 shows the significance values. Figure 5-3 shows the distribution of acres for households with different categories of food security. The medians for acres farmed are as follows:

Level of Food Security	Median Acres Planted
Food Secure	5.50
Moderately Food Insecure	3.42
Severely Food Insecure	4.00

Table 5.5: The median acres planted by households in each category

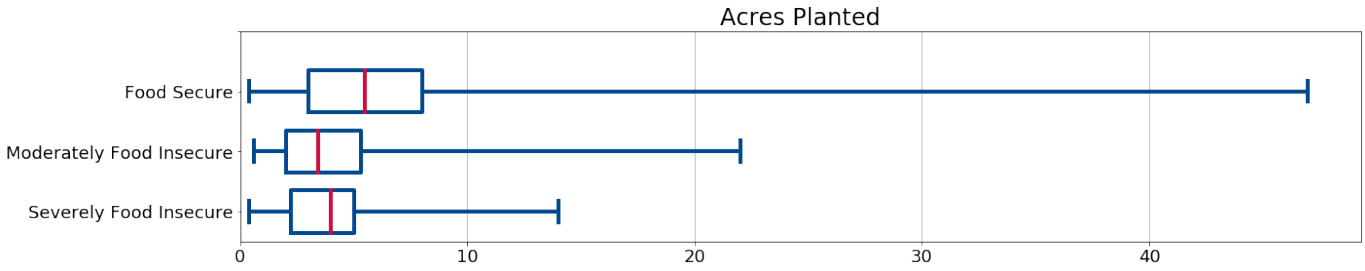

Figure 5-3: The distribution of acres planted. Households are grouped by their level of food security.

	Food Secure	Moderately Food Insecure	Severely Food Insecure
Food Secure	-	-	-
Moderately Food Insecure	2.08 (0.000***)	-	0.58 (0.384)
Severely Food Insecure	1.50 (0.004***)	-	-

Table 5.6: The difference in median acres planted in 2017, calculated using the median of the category in the column minus the median of the category of the row. In parentheses are the p-values testing whether the median for the category in the column is greater than the median for the category in a row. *, **, *** represent significance at 10%, 5%, and 1% probability levels.

Kilograms Harvested

Aggregating the total kilograms harvested from a household's top 3 crops across two seasons, the results show that households which are food secure have a larger harvest than households in all other categories. Households which are moderately food insecure also harvested more than households which are severely food insecure. This is an important result; it connects a household's productivity with its ability to meet its consumption needs. Increased productivity is linked with increased food security. It also reflects the extremely important role that agriculture plays in these households' livelihoods.

Level of Food Insecurity	Median Kg Harvested
Food Secure	1100
Moderately Food Insecure	746
Severely Food Insecure	565

Table 5.7: The median kg harvested by households in each category

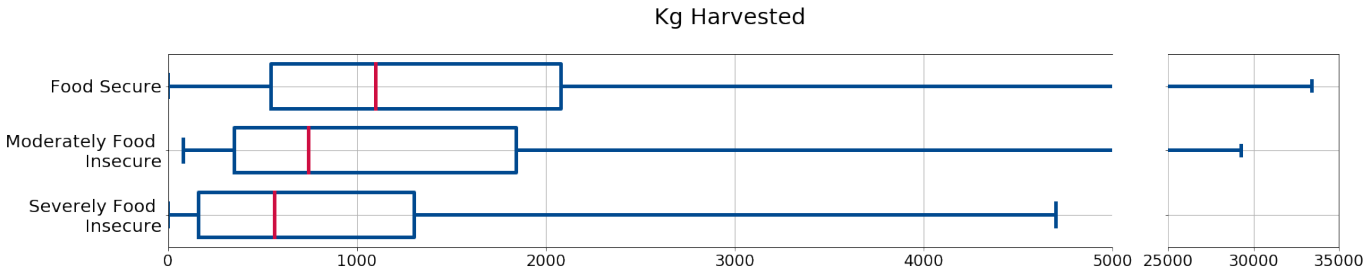

Figure 5-4: The distribution of kg harvested. Households are grouped by their reported frequency of skipping meals.

	Food Secure	Moderately Food Insecure	Severely Food Insecure
Food Secure	-	-	-
Moderately Food Insecure	356.00 (0.016**)	-	-
Severely Food Insecure	537.00 (0.001***)	181.00 (0.048**)	-

Table 5.8: The difference in kilograms harvested, calculated using the median of the category in the column minus the median of the category of the row. In parentheses are the p-values testing whether the median for the category in the column is greater than the median for the category in a row. *, **, *** represent significance at 10%, 5%, and 1% probability levels.

Maize Yield

To further explore the impact of agricultural success on food security, a separate analysis looked only at the 321 households which reported that maize was one of their top 3 crops in the November - December 2017 season. Each of these households reported the number of acres of maize that they planted, in addition to the kilograms that they harvested. This information was used to calculate the yield (kg/acre) of maize that each household achieved. The results generally show that food secure households achieve higher yields than less food secure households. Specifically, severely food insecure households have a statistically significant lower yield than all other households

Level of Food Security	Median Kg/Acre of Maize
Food Secure	326.67
Moderately Food Insecure	400.00
Severely Food Insecure	200.00

Table 5.9: The median number kg/acre planted of maize that households who planted maize harvested in November - December 2017, grouped by households in each category.

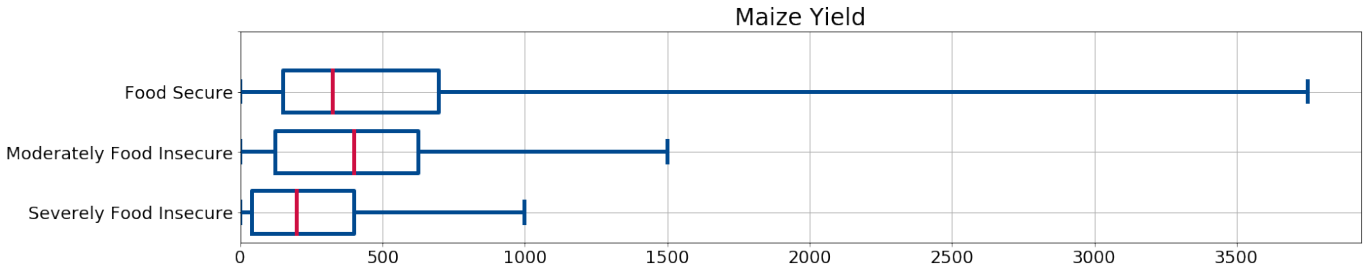

Figure 5-5: The distribution of yield of maize. Households are grouped by their level of food security.

	Food Secure	Moderately Food Insecure	Severely Food Insecure
Food Secure	-	73.33 (0.513)	-
Moderately Food Insecure	-	-	-
Severely Food Insecure	126.67 (0.004***)	200.00 (0.009***)	-

Table 5.10: The difference in median maize yield, calculated using the median of the category in the column minus the median of the category of the row. In parentheses are the p-values testing whether the median for the category in the column is greater than the median for the category in a row. *, **, *** represent significance at 10%, 5%, and 1% probability levels.

5.2.3 Level of Market Participation

This section connects different metrics of level of market participation to level of food security. We examine the two factors used as dependent variables in the models in Chapter 4: kg sold and percent sold. Whereas Chapter 4 focused on a single crop in a single season for these metrics, here we look at the aggregate of all crops. We also look at how food security is connected to the revenue earned from agriculture. The results show that there is no difference across level of food security in the percent of the harvest sold. This is consistent with the results found in Section 4.2. Food secure households sell a larger amount and earn more revenue than food insecure households. This connection between the amount sold and level of food security was not found in Section 4.2. It may be that the inclusion of other factors in the multiple linear regression model eclipsed the significance of food security. Using a single crop in a single season instead of the aggregate kg sold may also explain the differing result.

Kilograms Sold

For each household, the kilograms that they sold from each of their top 3 crops for both seasons in 2017 were totaled. Food secure households sell more than those which are food insecure. It is interesting that this significant result was not reflected in the model discussed in Section 4.2.1. It may be that the other variables included in the model are more strongly tied to the amount sold, and thus eclipse the importance of food security. In addition, the analysis here looks at total kg sold in 2017, while the model in Section 4.2.1 examines the kg sold of maize in a single season. The connection between selling more and food security should be researched further.

Level of Food Security	Median Kg Sold
Food Secure	435.0
Moderately Food Insecure	212.5
Severely Food Insecure	183.0

Table 5.11: The median kg sold by households in each category

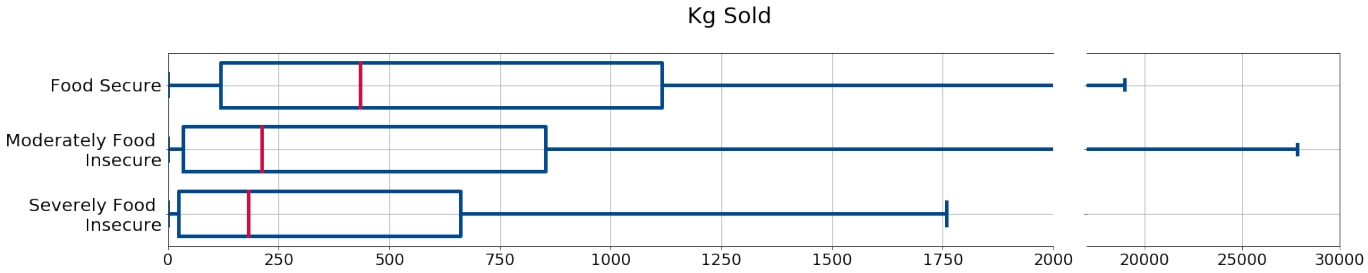

Figure 5-6: The distribution of kg sold. Households are grouped by their level of food security.

	Food Secure	Moderately Food Insecure	Severely Food Insecure
Food Secure	-	-	-
Moderately Food Insecure	222.50 (0.032**)	-	-
Severely Food Insecure	252.00 (0.014**)	29.50 (0.227)	-

Table 5.12: The difference in median kilograms sold, calculated using the median of the category in the column minus the median of the category of the row. In parentheses are the p-values testing whether the median for the category in the column is greater than the median for the category in a row. *, **, *** represent significance at 10%, 5%, and 1% probability levels.

Percent of Harvest Sold

As discussed in Chapter 4, another metric understanding a household's level of engagement with the market is the percent of their harvest which they sell. Here, we look at the percent of the overall harvest sold and find that there is no statistically significant difference across categories of food insecurity.

Level of Food Insecurity	Median Percent of Harvest Sold
Food Secure	42.9%
Moderately Food Insecure	41.2%
Severely Food Insecure	29.0%

Table 5.13: The median percent of their harvest sold by households in each category

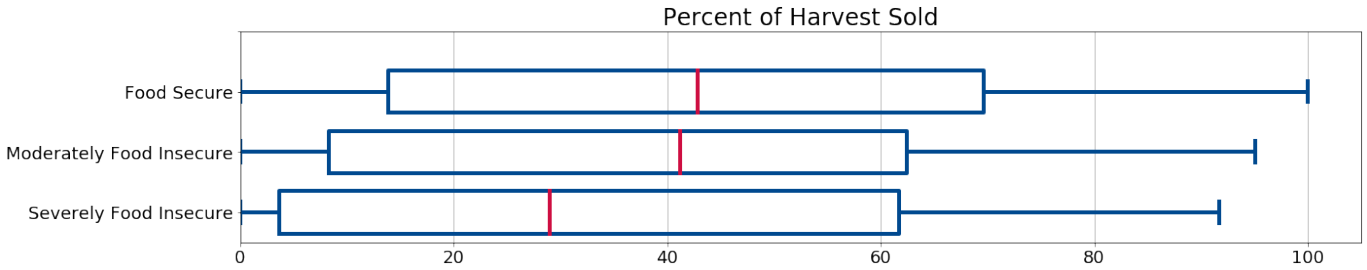

Figure 5-7: The distribution of the percent of harvest each household sold. Households are grouped by their level of food security.

	Food Secure	Moderately Food Insecure	Severely Food Insecure
Food Secure	-	-	-
Moderately Food Insecure	1.7 (0.181)	-	-
Severely Food Insecure	13.9 (0.130)	12.2 (0.373)	-

Table 5.14: The difference in median percent of harvest sold, calculated using the median of the category in the column minus the median of the category of the row. In parentheses are the p-values testing whether the median for the category in the column is greater than the median for the category in a row. *, **, *** represent significance at 10%, 5%, and 1% probability levels.

Annual Revenue

For each household, the amount of revenue that they generated from selling their top 3 crops in both the 2017A (June - July 2017) and 2017B (November - December 2017) was totaled. We see that the most food secure households earned more revenue from agriculture than others. This may reflect that they have higher wealth and better social positioning, not that there is a causal relationship.

Level of Food Security	Median USD Annual Revenue - 2017
Food Secure	$94.50
Moderately Food Insecure	$40.83
Severely Food Insecure	$40.63

Table 5.15: The median revenue earned from selling crops in 2017 by households in each category

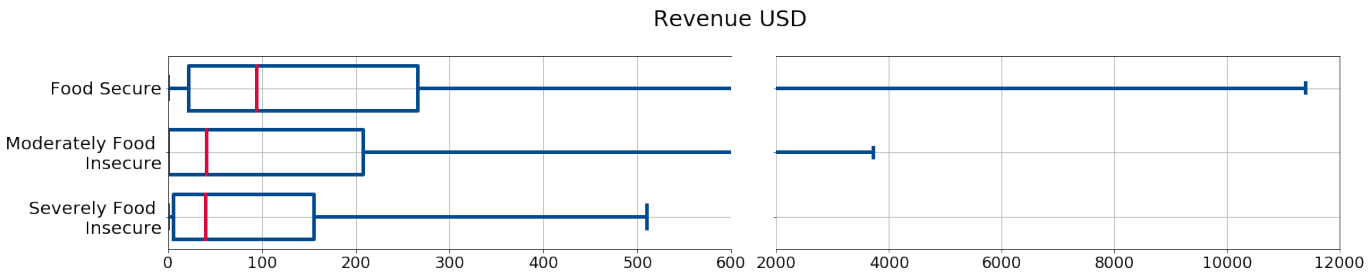

Figure 5-8: The distribution of revenue earned from selling crops in 2017. Households are grouped by their level of food security.

	Food Secure	Moderately Food Insecure	Severely Food Insecure
Food Secure	-	-	-
Moderately Food Insecure	53.67 (0.020**)	-	-
Severely Food Insecure	53.86 (0.022**)	0.20 (0.411)	-

Table 5.16: The difference in median revenue (in USD) earned from selling crops in 2017, calculated using the median of the category in the column minus the median of the category of the row. In parentheses are the p-values testing whether the median for the category in the column is greater than the median for the category in a row. *, **, *** represent significance at 10%, 5%, and 1% probability levels.

5.2.4 Food Purchasing

Complementing the food produced at home, households may also fulfill their consumption need by purchasing food. Food security in developed countries is typically associated with increased purchasing of food (Coleman-Jensen, 2010; St-Germain & Tarasuk, 2018). We look at factors for the amount of food purchased, the amount of money spent on food, the number of months out of the year that food was bought, and the number of months on which households mostly relied on the farm to see how the Ugandan context differs. The results show that food secure households buy less food, spend less money on food, and purchase food for fewer months out of the year than other households.

Food secure households rely primarily on the farm for food for more months out of the year than all other households. Moderately food insecure households rely on the farm for food for more months than severely food insecure households. As level of food insecurity increases, there is a statistically significant connection to an increasing probability of having to purchase food due to a poor harvest. All of these results underscore the importance of agriculture to the households in our survey. Purchasing food is an indicator of food insecurity. Households with higher food security are able to rely on their own production more and purchase less food produced by others.

Amount of Food Purchased Annually

Food secure households had a lower median of the amount of food purchased annually than other households. Across the households which experienced at least some level of food insecurity, there were no differences in the medians.

Level of Food Insecurity	Median Kg Food Purchased Last Year
Food Secure	30.0
Moderately Food Insecure	100.0
Severely Food Insecure	110.0

Table 5.17: The median kg of food purchased last year by households in each category

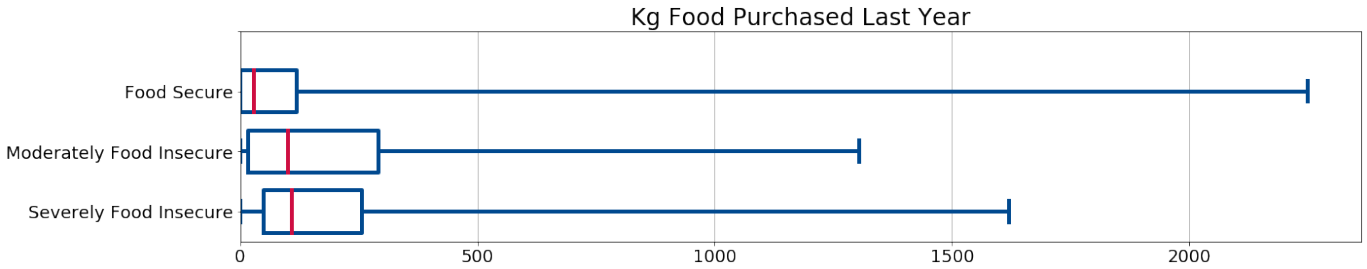

Figure 5-9: The distribution of kg of food purchased in the past year. Households are grouped by their level of food security.

	Food Secure	Moderately Food Insecure	Severely Food Insecure
Food Secure	-	70.0 (0.000***)	80.0 (0.000***)
Moderately Food Insecure	-	-	10.0 (0.211)
Severely Food Insecure	-	-	-

Table 5.18: The difference in median kilograms of food purchased last year, calculated using the median of the category in the column minus the median of the category of the row. In parentheses are the p-values testing whether the median for the category in the column is greater than the median for the category in a row. *, **, *** represent significance at 10%, 5%, and 1% probability levels.

Amount Spent on Food Annually

Food secure households spent a lower median amount on food than other households. This is consistent with the result above that they also purchase a smaller quantity. Across the households which experienced at least some level of food insecurity, there were no differences in the medians.

Level of Food Security	Median USD Spent on Food Last Year
Food Secure	$9.45
Moderately Food Insecure	$40.50
Severely Food Insecure	$42.53

Table 5.19: The median amount spent on food last year, grouped by households in each category

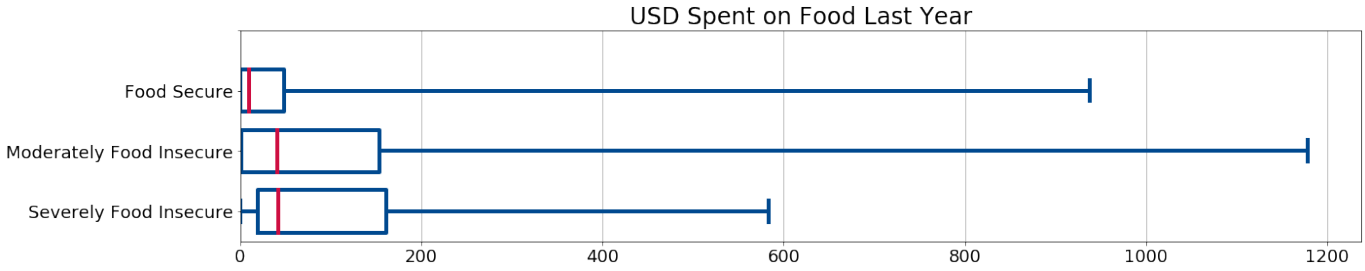

Figure 5-10: The distribution of the amount of money (USD) a household spent on food in the past year. Households are grouped by their level of food security.

	Food Secure	Moderately Food Insecure	Severely Food Insecure
Food Secure	-	31.05 (0.000***)	33.08 (0.000***)
Moderately Food Insecure	-	-	2.02 (0.192)
Severely Food Insecure	-	-	-

Table 5.20: The difference in median amount of money (in USD) spent on food in the past year, calculated using the median of the category in the column minus the median of the category of the row. In parentheses are the p-values testing whether the median for the category in the column is greater than the median for the category in a row. *, **, *** represent significance at 10%, 5%, and 1% probability levels.

Number of Months in Which Food was Purchased

Food secure households purchased food in a smaller number of months than other households. Across the households which experienced at least some level of food insecurity, there were no statistically significant differences in the medians.

Level of Food Security	Median Months Purchased Food Last Year
Food Secure	1
Moderately Food Insecure	2
Severely Food Insecure	2

Table 5.21: The median number on months during which a household purchased food last year, grouped by households in each category

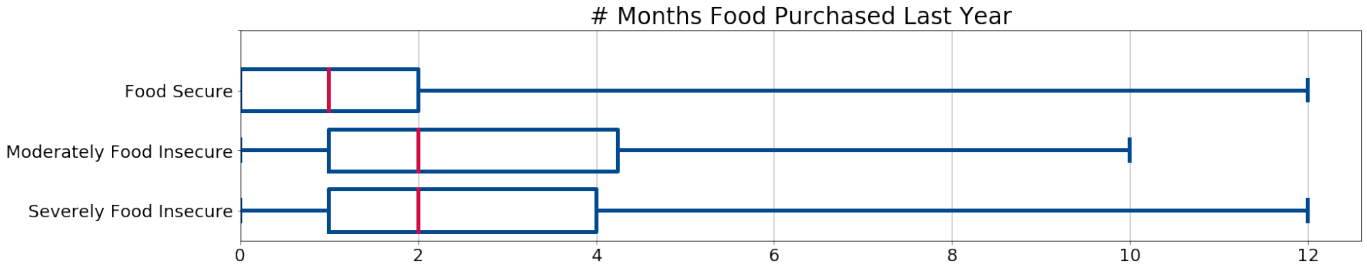

Figure 5-11: The distribution of number of months during which a household purchased food last year. Households are grouped by their level of food security.

	Food Secure	Moderately Food Insecure	Severely Food Insecure
Food Secure	-	1 (0.000***)	1 (0.000***)
Moderately Food Insecure	-	-	0 (0.203)
Severely Food Insecure	-	-	-

Table 5.22: The difference in median number of months during which a household purchased food, calculated using the median of the category in the column minus the median of the category of the row. In parentheses are the p-values testing whether the median for the category in the column is greater than the median for the category in a row. *, **, *** represent significance at 10%, 5%, and 1% probability levels.

Number of Months Relying on the Farm

Food secure households relied on the farm as their primary source of food for more months out of the year than other households. Severely food insecure households relied on the farm for fewer months than moderately food insecure households. These results paint a very consistent picture that increased ability to rely on food produced by the household is connected with increased household food security.

Level of Food Security	Median Months Primarily Relied on Farm Last Year
Food Secure	11
Moderately Food Insecure	10
Severely Food Insecure	6

Table 5.23: The median number on months during which a household primarily relied on the farm for food last year, grouped by households in each category

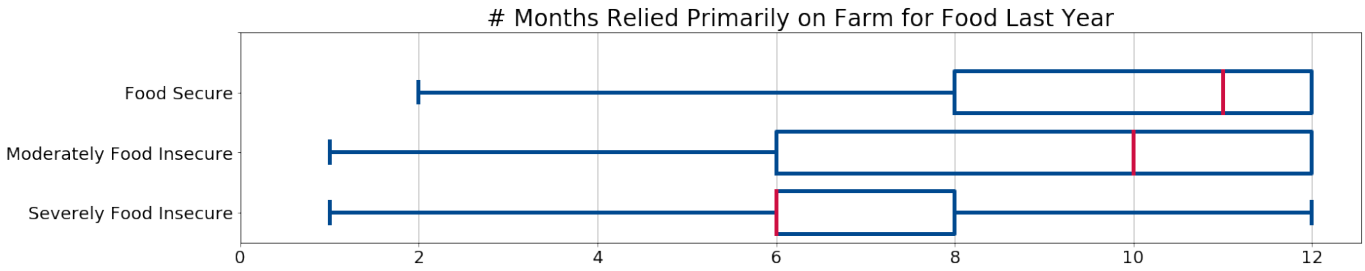

Figure 5-12: The distribution of number of months primarily relying on the farm for food in the past year. Households are grouped by their level of food security.

	Food Secure	Moderately Food Insecure	Severely Food Insecure
Food Secure	-	-	-
Moderately Food Insecure	1 (0.001***)	-	-
Severely Food Insecure	5 (0.000***)	4 (0.002***)	-

Table 5.24: The difference in median number of months relying on the farm, calculated using the median of the category in the column minus the median of the category of the row. In parentheses are the p-values testing whether the median for the category in the column is greater than the median for the category in a row. *, **, *** represent significance at 10%, 5%, and 1% probability levels.

Poor Harvest

Looking at the reasons households gave for purchasing food in a given month, households which experienced the most severe food insecurity were much more likely to indicate a poor harvest as all other households, as shown in Figure 5-13. To demonstrate that this difference in the probability of purchasing grain due to poor harvest is significant, Fisher's exact test was used to compare the probabilities to each other. The massive increase of the severely food insecure is statistically significant compared to the other levels. In fact, within the severely food insecure category, 100% of households which skipped meals once a week or more reported having a poor harvest. This speaks to how crucial home production is for food security; poor harvests are the reason for food insecure households purchasing food.

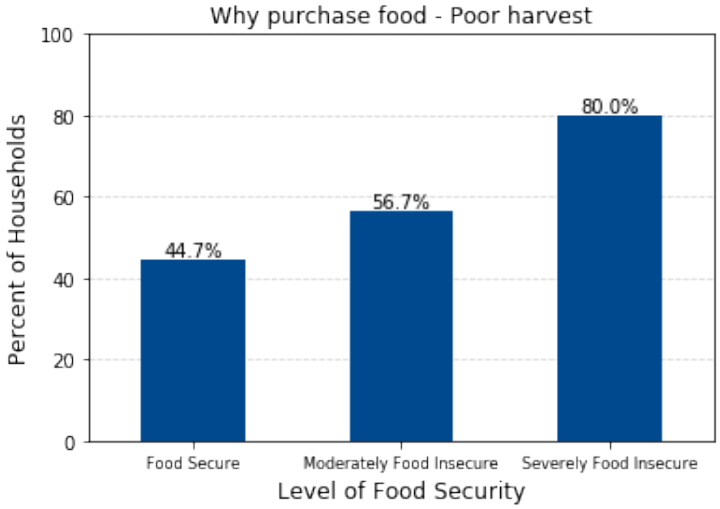

Figure 5-13: The percent of households which cited a poor harvest as the reason for purchasing food in any month in the past year.

5.3 Results

From the results discussed above, it is clear the in the Ugandan context, household food security is linked to increased agricultural production. Across all levels of food security, households have approximately the same level of consumption need. However, food insecure households farm less land and have smaller harvests than food secure households, precluding them from meeting their consumption needs through their own farming efforts. We see this conclusion reflected back to us in the results showing that the most severely food insecure households buy more food, spend more money on food, purchase it in more months out of the year, and rely on their farms fewer months of the year. Note that this result runs counter-intuitive to the experience of food insecurity in many developed countries. In the West, purchasing larger quantities of food and spending more money on it is typically a sign of increased wealth and ability to meet consumption needs (Coleman-Jensen, 2010; St-Germain & Tarasuk, 2018). In Uganda, increased spending on food seems to reflect a desperation to meet consumption needs.

In addition, not only are food secure households using more land, they are more productive with each of those acres. Section 5.2.2 demonstrated that the severely food insecure households achieved much lower maize yields than others. This connection between decreased productivity and food insecurity may also reflect lower levels of education and wealth.

We find that the connection to level of market participation may need further research. Percent sold did not differ across level of food security, consistent with the model in Section 4.2. However, the results for the tests here show that increased food security is linked with an increased amount sold and thus increased revenue earned from selling. The model in Section 4.2 did not find a significant result connecting food security to amount sold. It may be that the other factors in that model overshadow the connection to food security, or that level of food security impacts the amount sold in aggregate but not the amount of maize sold in November - December 2017.

The result that food security is not a prerequisite for market participation may reflect other household economic decisions. Households may have a need for liquidity that urges them to sell. This may be to pay for medicine, school fees, or burial fees. The lack of an association between the decision to participate in markets and food security suggests that market facilitation efforts will not exclude this vulnerable population and that they may in fact have much to gain from this programming. Further research should examine other dimensions of socioeconomic vulnerability to see if they do impact market participation.

THIS PAGE INTENTIONALLY LEFT BLANK

Chapter 6

Discussion

6.1 Findings and Policy Implications

The results from the analysis of the household characteristics which impact market participation by selling crops can help inform future market facilitation efforts across Uganda. USAID can incorporate the findings presented here as they develop new activities for the Feed the Future program.

Section 6.1.1 discusses the results found in Section 4.1.2 to answer our first research question: **What factors impact farmers' decisions to participate in the market?**

Section 6.1.2 then discusses the results found in Section 4.2 to answer our second research question: **Conditional on participation, what factors impact the level of market participation?**

Section 5.3 has already answered our third research question: **Is food security connected market participation or other metrics for agricultural engagement?**

6.1.1 Market Participation

We find that the extent of market development has a significant impact on the decision to participate while market access does not. In this model, market development is explained by the size of nearby towns and the number of services available while market access is explained by the distance to the nearest road, distance to the nearest town, and transportation access.

The importance of market development suggests that market facilitation interventions may be most fruitful when aimed at developing small businesses which provide agricultural services (such as weed-

ing, plowing, harvesting, and drying) and increasing competition among buyers. This may include educational programs on business management and low-cost loans for start-up capital. In addition, there may be a benefit to segmenting market facilitation efforts based on the size of nearby towns. Households in rural areas that are closest to smaller towns may get a greater benefit from programs focused on market development than households which are closest to a large town (where markets may already be well-developed).

The finding that market access does not enable market participation is in contrast with the findings in Alene (2008) and Key et al (2000). Households which are further from roads, further from towns, and do not have access to private transportation are not barred from participating in markets. No matter how far away people are from a town, they will find a way to sell what they have grown. This result suggests that market facilitation interventions could be shifted away from focusing on access and reallocated elsewhere.

The model also shows that food secure households were not more likely to participate in markets. This finding can allay concerns that market facilitation efforts exclude vulnerable populations. Households without enough food to eat may choose to sell some of their crops as they prioritize other needs for cash (medical bills, school fees, burial fees). Not only do food insecure households stand to benefit from market facilitation efforts, but commercialization programs could be aimed at this specific population.

The model for market participation decisions also showed that membership in a producer organization served as an enabler to households selling their crops. This link is not necessarily causal. For example, it may be that households which are already engaged in markets are more likely to join these groups in the hopes of reducing their transaction costs. However, there is still benefit in promoting such organizations. If they are effective at reducing transaction costs, that may encourage a household which otherwise would perceive the threshold cost of entering the market as too high. The linkages between members can provide information and the possibility of more market connections.

Latynskiy and Berger (2016) found that lack of liquid and physical assets and low commitment among members hindered the effectiveness of producer organizations in Uganda. Bernard et al (2008) similarly demonstrated that low managerial capacity and limited financial resources constrained the impact of these groups in Senegal and Burkina Faso. USAID should design market facilitation efforts to address these challenges so as to strengthen these organizations in areas where they exist. In areas where there are no local producer organization, programming should seek to educate and mobilize farmers to form them, providing them with access to resources and leadership training to set the group on a path to self-sufficiency. Further study should seek to understand why some households choose not to join these organizations even when they exist nearby.

Mobile phone ownership was also found to increase the probability of selling. This reflects the impor-

tance of access to market information. Households with phones can more easily learn about prices, buyers, transportation options, and service providers. Market facilitation efforts should therefore include components of outreach to increase available market information.

6.1.2 Level of Market Participation

Using different metrics for market participation, the two models yielded the same significant features with the same signs. In this instance, building a model with one dependent variable versus the other would not impact the barriers or enablers identified. This consistency confirms that the amount sold metric and the proportion sold metric are measuring a similar concept. If one had to be chosen as an indicator of the impact of market facilitation, it would make more sense to collect information on the amount sold as it requires only one measurement. Calculating the proportion sold requires the additional measurement of the amount harvested. This makes collecting information on the amount sold cheaper, faster, easier, and less prone to error.

These analyses found that increasing distance from a major road decreased the level of market participation, while increasing distance traveled to town increased the level of market participation. It may be that increasing distance from a major road means more travel along footpaths and poor quality roads, reducing the amount an individual wants to carry or an intermediary will travel to buy crops. However, once they reach a main road and have to travel to town, longer travel time may mean a higher cost, which requires a larger sale to offset. This result is analogous to the one found in Heltberg and Tarp (2002). That study found that the aggregate agricultural revenue earned by farmers in Mozambique decreased the further that they were from a railway (similar to our distance to road variable) and increased the further away they were from the provincial capital (similar to our distance to town variable). One explanation proposed is that those who are closer to town may be more likely to participate in seasonal work than those further away. Seasonal work may not be captured by the variable for the percent of off-farm income, which may only reflect current and consistent employment. Those who can access seasonal employment substitute it for market participation. This result should be explored further, especially to understand the interaction of these two features.

Finally, the factor that is most significantly linked to market participation decisions, the level of market participation, and level of food security is production level. Households which harvest more are more likely to be food secure, more likely to participate in the markets, and participate in those markets more intensely. The consistent importance of production levels suggests that USAID should strongly consider programming that strengthens the input supply chain to farmers. Future research should more closely examine this input path to better understand how it connects to production and thus market participation. Market facilitation efforts in this vein may be designed around agroinput shops and financial products to increase opportunities to buy and rent land. Agroinput shops provide tools,

improved seed varieties, chemicals, and information to their customers. Issues with counterfeiting and adulteration have hindered adoption of chemicals and improved seeds across Uganda (Bold, Kaizzi, & Yanagizawa-Drott, 2017; Ashour et al., 2019). Market facilitation efforts to improve the quality of these productions and build back consumer trust have the potential to increase agricultural production levels.

The connection between chemical use (herbicides, pesticides, fertilizers) and greater production should be given additional critical thought. Increased education around and credit for chemicals may be able to help households with small land-holdings increase the productivity of that land. However, concerns about the long-term impacts of these products should be carefully considered. Products and farming techniques that pose fewer health risks and are more sustainable economically and environmentally should be explored as alternatives. If loans are to be offered to purchase chemicals, they could also be offered to hire laborers to help with more labor-intensive organic farming techniques.

6.2 Future Research Opportunities

The work presented here examines only a small part of the agricultural supply chains which impact small-holder farmers in Uganda. There are myriad opportunities for future research to build upon our investigation.

With access to appropriate price data for the region, time period, and crops, future work seek to understand the ways that prices interact with other household characteristics to drive market participation. There may be price thresholds whereby certain households only participate when they believe the price will be high enough.

The poor fit of our model regressing on the percent of the harvest sold suggests that further research should be conducted to better understand this dynamic. While the production level of a household does influence what portion of it they choose to sell, there is still a large amount of unexplained variation in the data. In light of our analysis that shows food security is not connected to market participation, this unexplained variation may not be surprising. Clearly, households make marketing decisions based on priorities beyond simply production and consumption levels. Further analysis could improve the understanding of these priorities. Crops other than maize and seasons other than November - December 2017 should be modelled to see if the barriers and enablers found in our multiple linear regression models are consistent across the model focus.

Future research may want to more explicitly explore how indicators of vulnerability impact market participation. Our model found that food insecure households engaged in markets at the same rate as other households. When vulnerability is defined using a different framework, however, it is possible

that there may be a connection to market participation. Households experiencing extreme poverty, extreme isolation, and disability/illness should be researched further to understand if and how they participate in markets.

Well-functioning market systems do not appear out of nowhere, but are built overtime out of mutually-beneficial connections between agribusinesses and farmers. Throughout this thesis, we touch upon the role of agribusinesses as they relate to offering credit and services to farmers. Analyzing these connections can further develop our understanding of the Ugandan market system and appropriate market facilitation interventions. To that end, information from the household survey can be used in conjunction with data collected about nearby agribusinesses to understand the role they play in market infrastructure. Research should be conducted to further clarify market incentives that align increased market participation with the interests of agribusinesses. As mentioned above, more research should be directed at the inputs supply chain to better understand how strengthening the inputs market can impact the output market. This can include the availability, accessibility, and affordability of services (weeding, harvesting, processing) and agroinput shops.

THIS PAGE INTENTIONALLY LEFT BLANK

Appendix A

Appendix A: Household Survey

Jarrod Goentzel, Director, Humanitarian Response Lab
Massachusetts Institute of Technology, 77 Massachusetts Ave.,
Cambridge, MA 02139 USA
goentzel@mit.edu +1 617-253-2053

FARMER HOUSEHOLD SURVEY QUESTIONNAIRE

0.1	DATE	*[DATE]*
0.2	START TIME	*[TIME]*
0.3	DISTRICT	*[SINGLE CODE]* *GULU* *IBANDA* *IGANGA* *MUBENDE* *PADER*
0.4	VILLAGE	*[SHORT ANSWER]*
0.5	INTERVIEWER ID	*[NUMBER]*
0.6	INTERVIEWER NAME	*[SHORT ANSWER]*

[INTERVIEWER SCRIPT]

Thank you for agreeing to participate in this study. We are conducting research for USAID in order to understand how farmers like you engage with the market. We will ask some basic questions about your household, and then ask about your farm and your business partners. We appreciate your participation. Let's get started.

1. DEMOGRAPHIC QUESTIONS

1.1	Name	*[SHORT ANSWER]*
1.2	Age	*[NUMBER]*
1.3	Gender	*MALE* *FEMALE*
1.4	Highest level of education attained	*[SINGLE CODE]* *NO FORMAL EDUCATION* *SOME PRIMARY* *COMPLETED PRIMARY* *SOME O-LEVEL* *COMPLETED O-LEVEL* *SOME A-LEVEL* *COMPLETED A-LEVEL* *SOME UNIVERSITY* *COMPLETED UNIVERSITY* *ADULT LITERACY ONLY (NO FORMAL EDUCATION)* *OTHER [SPECIFY]* *DON'T KNOW* *REFUSED*
Interviewer script		*Now I will ask some questions about*

Jarrod Goentzel, Director, Humanitarian Response Lab
Massachusetts Institute of Technology, 77 Massachusetts Ave.,
Cambridge, MA 02139 USA
goentzel@mit.edu +1 617-253-2053

		*your household. Your household includes everyone who usually **lives here, sleeps here, and eats from the same source.** Please include children, relatives, or orphans, even if they are not at home at the time of interview, but do not count temporary visitors. Please include children who may be away at school.*
1.5	How many adults live in the household?	*[NUMBER]* *DON'T KNOW* *REFUSED*
1.6	How many children under 18 live in the household?	*[NUMBER]* *DON'T KNOW* *REFUSED*
	IF 1.6 > 0	
1.6A	Do any of the children live away at boarding school?	*YES* *NO* *DON'T KNOW* *REFUSED*
	IF YES @ 1.6A	
1.6B	How many?	*[NUMBER]* *DON'T KNOW* *REFUSED*
1.7	Do you financially support any other people who do not live in the household?	*YES* *NO* *DON'T KNOW* *REFUSED*
	IF YES @ 1.7	
1.7A	How many?	*[NUMBER]* *DON'T KNOW* *REFUSED*
1.8	Do you own a mobile phone?	*YES* *NO* *DON'T KNOW* *REFUSED*
1.9	Do you own or have access to a bicycle or other means of transportation?	*[MULTICODE]* *BICYCLE* *MOTORCYCLE* *CAR* *TRUCK* *OTHER [SPECIFY]* *NONE* *DON'T KNOW*

Jarrod Goentzel, Director, Humanitarian Response Lab
Massachusetts Institute of Technology, 77 Massachusetts Ave.,
Cambridge, MA 02139 USA
goentzel@mit.edu +1 617-253-2053

		REFUSED
1.10	Did your household earn income from outside the farm in the past year?	*YES* *NO* *DON'T KNOW* *REFUSED*
	IF YES @ 1.10	
1.10A	What portion of the household's income comes from work outside the farm?	*[SINGLE CODE]* *LESS THAN 25%* *25-50%* *50-75%* *75-100%* *DON'T KNOW* *REFUSED*
1.10B	What are the other sources of household income?	*[SHORT ANSWER]* *DON'T KNOW* *REFUSED*
1.10C	What do you consider your primary source of income?	*[SHORT ANSWER]* *DON'T KNOW* *REFUSED*
1.11	Have you participated in or received support or training from any NGO or government programs in the past year?	*YES* *NO* *DON'T KNOW* *REFUSED*
	IF YES @ 1.11	
1.11A	What kind of support or knowledge? Who provided it?	*[SHORT ANSWER]* *DON'T KNOW* *REFUSED*
1.11B	Who provided it?	*[SHORT ANSWER]* *DON'T KNOW* *REFUSED*

2. FINANCE QUESTIONS

2.1	Does anyone in the household have a bank account?	*YES* *NO* *DON'T KNOW* *REFUSED*
2.2	Does anyone in the household belong to a VSLA, SACCO, or savings group?	*YES* *NO* *DON'T KNOW* *REFUSED*
2.3	Does anyone in the household have a mobile money account?	*YES* *NO*

Jarrod Goentzel, Director, Humanitarian Response Lab
Massachusetts Institute of Technology, 77 Massachusetts Ave.,
Cambridge, MA 02139 USA
goentzel@mit.edu +1 617-253-2053

		DON'T KNOW *REFUSED*
2.4	Has anyone in the household borrowed money in the past year?	*YES* *NO* *DON'T KNOW* *REFUSED*
	IF YES @ 2.4	
2.4A	What was the amount borrowed?	*[NUMBER]* *DON'T KNOW* *REFUSED*
2.4B	What was the money borrowed for?	*[MULTI CODE]* *AGRICULTURAL INPUTS* *SCHOOL FEES* *WEDDING* *BURIAL* *OTHER [SPECIFY]* *DON'T KNOW* *REFUSED*
2.4C	What was the source of the money?	*[MULTI CODE]* *RELATIVE* *FRIEND* *INFORMAL LENDER* *BANK* *MICROFINANCE* *SACCO* *VSLA* *OTHER [SPECIFY]* *DON'T KNOW* *REFUSED*
	IF NO @ 2.4	
2.4D	Did anyone in the household try to borrow money or get a loan in the past year and not succeed?	*YES* *NO* *DON'T KNOW* *REFUSED*
	IF YES @ 2.4D	
2.4E	Why not?	*[MULTI CODE]* *NO COLLATERAL* *MISSING PAPERWORK* *OTHER [SPECIFY]* *DON'T KNOW* *REFUSED*
2.5	Has anyone in the household received credit from a business in the past year, such as inputs that are paid for at the end of the season?	*YES* *NO* *DON'T KNOW* *REFUSED*

Jarrod Goentzel, Director, Humanitarian Response Lab
Massachusetts Institute of Technology, 77 Massachusetts Ave.,
Cambridge, MA 02139 USA
goentzel@mit.edu +1 617-253-2053

	IF YES @ 2.5	
2.5A	What was the source of this credit?	*[MULTI CODE]* *INPUT DEALER* *TRADER* *COOPERATIVE* *OUTGROWER SCHEME* *OTHER [SPECIFY]* *DON'T KNOW* *REFUSED*
2.6	Has anyone in the household lent money to someone outside the household in the past year?	*YES* *NO* *DON'T KNOW* *REFUSED*

3. AGRONOMIC PRACTICES

3.1	How much land does your household farm?	*[NUMBER] [ACRES/HA]* *DON'T KNOW* *REFUSED*
3.2	How much of this land does your household own?	*NUMBER [ACRES/HA]* *DON'T KNOW* *REFUSED*
3.3	How much of this land are you renting or borrowing?	*NUMBER [ACRES/HA]* *DON'T KNOW* *REFUSED*
3.4	Has the household ever rented land for farming in the past? (this question should be asked with an emphasis on the past, even if the family currently rents all or some of their land)	*YES* *NO* *DON'T KNOW* *REFUSED*
	IF YES @ 3.4	
3.4A	When was this?	*YEAR* *DON'T KNOW* *REFUSED*
	IF YEAR @ 3.4A > 2012	
3.4B	How many acres/hectares?	*[NUMBER] [ACRES/HA]* *DON'T KNOW* *REFUSED*
3.4C	At what price?	*NUMBER [TOTAL / PER ACRE / PER HA / OTHER]* *DON'T KNOW* *REFUSED*

Jarrod Goentzel, Director, Humanitarian Response Lab
Massachusetts Institute of Technology, 77 Massachusetts Ave.,
Cambridge, MA 02139 USA
goentzel@mit.edu +1 617-253-2053

	IF NO @ 3.4	
3.4D	Why haven't you rented land?	*[MULTI CODE]* *DON'T NEED* *CAN'T AFFORD* *NOT AVAILABLE* *OTHER [SPECIFY]* *DON'T KNOW* *REFUSED*
3.5	How much does it usually cost per season to rent an acre in this area?	*[NUMBER] [PER ACRE / PER HA]* *DON'T KNOW* *REFUSED*
3.6	Did you hire anyone to work on your farm this past season or the season before?	*YES* *NO* *DON'T KNOW* *REFUSED*
	IF YES @ 3.6	
3.6A	How many people did you hire?	*[NUMBER]* *DON'T KNOW* *REFUSED*
3.6B	How many days did each person work on average?	*[NUMBER]* *DON'T KNOW* *REFUSED*
3.7	Do you have any livestock?	*YES* *NO* *DON'T KNOW* *REFUSED*
	IF YES @ 3.7A	
3.7A	What kind?	*[MULTI CODE]* *CATTLE* *GOATS* *PIGS* *SHEEP* *CHICKENS* *DONKEYS* *OTHER [SPECIFY]* *DON'T KNOW* *REFUSED*
	LOOP FOR EACH ANSWER @ 3.7A	
3.7B	How many?	*[NUMBER]*
3.8	Have you ever received any formal training on farming techniques?	*YES* *NO* *DON'T KNOW* *REFUSED*
	IF YES @ 3.8	

Jarrod Goentzel, Director, Humanitarian Response Lab
Massachusetts Institute of Technology, 77 Massachusetts Ave.,
Cambridge, MA 02139 USA
goentzel@mit.edu +1 617-253-2053

3.8A	When was the most recent training?	*YEAR* *DON'T KNOW* *REFUSED*
3.8B	From whom?	*[MULTI CODE]* *NGO* *EXTENSION WORKER* *COOPERATIVE* *GOVERNMENT PROGRAM* *INPUT DEALER* *BUYER* *OTHER [SPECIFY]* *DON'T KNOW* *REFUSED*
3.8C	What was the training about?	*[SHORT ANSWER]* *DON'T KNOW* *REFUSED*
3.8D	How did you hear about the training?	*[MULTI CODE]* *PERSONALLY INVITED* *WORD OF MOUTH* *HEARD ON RADIO* *SELF-INVITATION* *OTHER [SPECIFY]* *DON'T KNOW* *REFUSED*
3.9	Did you have any of these problems with your crops last season (ending November-December 2017)? (Pease ask, "Did you have ____ " for each option)	*[MULTI CODE]* *DROUGHT* *HEAVY RAINS* *LATE RAINS* *FALL ARMY WORM* *OTHER INSECTS* *CROP DISEASES* *FIRE* *OTHER [SPECIFY]* *NONE* *DON'T KNOW* *REFUSED*
3.10	Did you have any of these problems with your crops the season before (ending June-July 2017)? (Please ask, "Did you have ____ " for each option)	*[MULTI CODE]* *DROUGHT* *HEAVY RAINS* *LATE RAINS* *FALL ARMY WORM* *OTHER INSECTS* *CROP DISEASES* *FIRE* *OTHER [SPECIFY]*

Jarrod Goentzel, Director, Humanitarian Response Lab
Massachusetts Institute of Technology, 77 Massachusetts Ave.,
Cambridge, MA 02139 USA
goentzel@mit.edu +1 617-253-2053

		NONE *DON'T KNOW* *REFUSED*
3.11	For most years, how many months out of the year does the household eat primarily food grown on the farm?	*[NUMBER]* *DON'T KNOW* *REFUSED*
3.12	In the last year, did your household purchase food for consumption?	*YES* *NO* *DON'T KNOW* *REFUSED*
	IF YES @ 3.12, ASK 3.12A-3.12E. **3.12B-D should be looped in a subgroup for each month checked.**	
3.12A	During which months did you buy food this past year?	*[MULTI CODE]* *JANUARY* *FEBRUARY* *MARCH* *APRIL* *MAY* *JUNE* *JULY* *AUGUST* *SEPTEMBER* *OCTOBER* *NOVEMBER* *DECEMBER* *DON'T KNOW* *REFUSED*
	LOOP 3.12B-D FOR EACH MONTH SELECTED	
3.12B	How much did you purchase?	*[NUMBER] [KG / SPECIFY UNIT]* *DON'T KNOW* *REFUSED*
3.12C	How much did you pay in total?	*[NUMBER] [UGX / SPECIFY UNIT]* *DON'T KNOW* *REFUSED*
3.12D	Why did you purchase the food?	*[MULTI CODE]* *DIDN'T PLANT THIS CROP* *POOR HARVEST* *LITTLE STORAGE CAPACITY* *NOT ENOUGH PRODUCTION BECAUSE OF SMALL LAND SIZE* *SALE OF HARVEST DUE TO URGENCY* *OTHER [SPECIFY]* *DON'T KNOW*

Jarrod Goentzel, Director, Humanitarian Response Lab
Massachusetts Institute of Technology, 77 Massachusetts Ave.,
Cambridge, MA 02139 USA
goentzel@mit.edu +1 617-253-2053

		REFUSED
3.12E	Do you buy food every year?	*[SINGLE CODE]* *YES* *ONLY IF HARVEST IS BAD* *ONLY IN EXCEPTIONAL CASES* *OTHER [SPECIFY]* *DON'T KNOW* *REFUSED*
3.13	Does anyone in your household ever have to skip meals because there is not enough food?	*[SINGLE CODE]* *ONCE OR TWICE A WEEK* *ONCE OR TWICE A MONTH* *A FEW TIMES PER YEAR* *NEVER* *OTHER [SPECIFY]* *DON'T KNOW* *REFUSED*

4. AGRICULTURAL INPUTS

4.1	Is there an input shop nearby where you can purchase inputs?	*YES* *NO* *DON'T KNOW* *REFUSED*
	IF YES @ 4.1	
4.1A	Name of input dealer	*[SHORT ANSWER]* *DON'T KNOW* *REFUSED*
4.1B	Town / Village	*[SHORT ANSWER]* *DON'T KNOW* *REFUSED*
4.1C	Have you ever purchased inputs from this shop?	*YES* *NO* *DON'T KNOW* *REFUSED*
	IF YES @ 4.1C	
4.1D	How often do you buy from them?	*[SINGLE CODE]* *EVERY TIME I BUY INPUTS* *SOMETIMES* *JUST ONCE* *OTHER [SPECIFY]* *DON'T KNOW* *REFUSED*
4.1E	Why did you decide to purchase from	*[MULTI CODE]*

Jarrod Goentzel, Director, Humanitarian Response Lab
Massachusetts Institute of Technology, 77 Massachusetts Ave.,
Cambridge, MA 02139 USA
goentzel@mit.edu +1 617-253-2053

	this input shop?	*NEARBY* *REFERRED BY SOMEONE* *SAW ADVERTISEMENT* *HAS GOOD REPUTATION* *NO PARTICULAR REASON* *OTHER [SPECIFY]* *DON'T KNOW* *REFUSED*
	IF HAS GOOD REPUTATION @ 4.1E	
4.1F	Why does this shop have a good reputation?	*[SHORT ANSWER]* *DON'T KNOW* *REFUSED*
4.1G	Do you trust the shop owner?	*[SINGLE CODE]* *1: NOT AT ALL* *2: NOT REALLY* *3: NOT SURE* *4: SOMEWHAT* *5: YES, DEFINITELY* *DON'T KNOW* *REFUSED*
	IF NO @ 4.1C	
4.1H	Why haven't you purchased inputs from them?	*[MULTI CODE]* *DON'T PURCHASE INPUTS* *CAN'T AFFORD INPUTS* *SHOP TOO FAR* *SHOP HAS BAD REPUTATION* *DON'T TRUST OWNER* *OTHER [SPECIFY]* *DON'T KNOW* *REFUSED*
	IF NO @ 4.1	
4.2	For the crops that were planted last season, where did your seeds come from?	*[MULTI CODE]* *HOME-SAVED* *PURCHASED* *NGO* *GOVERNMENT* *OTHER [SPECIFY]* *DON'T KNOW* *REFUSED*
	IF PURCHASED @ 4.2	
4.2A	Where did you purchase them?	*[MULTI CODE]* *INPUT SHOP* *FAMILY* *FRIEND* *NEIGHBOR*

Jarrod Goentzel, Director, Humanitarian Response Lab
Massachusetts Institute of Technology, 77 Massachusetts Ave.,
Cambridge, MA 02139 USA
goentzel@mit.edu +1 617-253-2053

		COOPERATIVE *OTHER [SPECIFY]* *DON'T KNOW* *REFUSED*
4.3	For the crops that were planted for the first season of 2017, where did your seeds come from?	*[MULTI CODE]* *HOME-SAVED* *PURCHASED* *NGO* *GOVERNMENT* *OTHER [SPECIFY]* *DON'T KNOW* *REFUSED*
	IF PURCHASED @ 4.3	
4.3A	Where did you purchase them?	*[MULTI CODE]* *INPUT SHOP* *FAMILY* *FRIEND* *NEIGHBOR* *COOPERATIVE* *OTHER [SPECIFY]* *DON'T KNOW* *REFUSED*
4.4	Did you use any agricultural chemicals last season, which ended in November-December 2017?	*YES* *NO* *DON'T KNOW* *REFUSED*
	IF YES @ 4.4	
4.4A	What did you use?	*[MULTI CODE]* *FERTILIZER* *PESTICIDE* *HERBICIDE* *OTHER [SPECIFY]* *DON'T KNOW* *REFUSED*
4.4B	Where did you purchase it?	*[MULTI CODE]* *INPUT SHOP* *FAMILY* *FRIEND* *NEIGHBOR* *COOPERATIVE* *OTHER [SPECIFY]* *DON'T KNOW* *REFUSED*
	IF NO @ 4.4	
4.4C	Why didn't you use any?	*[MULTI CODE]*

Jarrod Goentzel, Director, Humanitarian Response Lab
Massachusetts Institute of Technology, 77 Massachusetts Ave.,
Cambridge, MA 02139 USA
goentzel@mit.edu +1 617-253-2053

		COULDN'T AFFORD *DON'T NEED* *NOT AVAILABLE* *DON'T SEE VALUE* *TOO FAR AWAY* *OTHER [SPECIFY]* *DON'T KNOW* *REFUSED*
4.5	Did you use any agricultural chemicals the season before, which ended in June-July 2017?	*YES* *NO* *DON'T KNOW* *REFUSED*
	IF YES @ 4.5	
4.5A	What did you use?	*[MULTI CODE]* *FERTILIZER* *PESTICIDE* *HERBICIDE* *OTHER [SPECIFY]* *DON'T KNOW* *REFUSED*
4.5B	Where did you purchase them?	*[MULTI CODE]* *INPUT SHOP* *FAMILY* *FRIEND* *NEIGHBOR* *COOPERATIVE* *OTHER [SPECIFY]* *DON'T KNOW* *REFUSED*
	IF NO @ 4.5	
4.5C	Why didn't you purchase any?	*[MULTI CODE]* *COULDN'T AFFORD* *DON'T NEED* *NOT AVAILABLE* *DON'T SEE VALUE* *TOO FAR AWAY* *OTHER [SPECIFY]* *DON'T KNOW* *REFUSED*
4.6	Are you aware of the problem with counterfeit inputs?	*YES* *NO* *DON'T KNOW* *REFUSED*
	IF YES @ 4.6	

Jarrod Goentzel, Director, Humanitarian Response Lab
Massachusetts Institute of Technology, 77 Massachusetts Ave.,
Cambridge, MA 02139 USA
goentzel@mit.edu +1 617-253-2053

4.6A	How did you learn about this problem?	*[MULTI CODE]* *FRIENDS OR NEIGHBORS* *INPUT DEALER* *NGO* *GOVERNMENT* *NEWSPAPER* *RADIO* *TELEVISION* *OTHER [SPECIFY]* *DON'T KNOW* *REFUSED*
4.6B	What do you do to avoid counterfeit inputs?	*[SHORT ANSWER]* *DON'T KNOW* *REFUSED*
4.7	Let's say a particular tin of seeds costs 20,000/= for 50g. How much more would you pay for a tin that was guaranteed to have a germination rate of 85-90%?	*[SINGLE CODE]* *1000 UGX* *2000 UGX* *5000 UGX* *10000 UGX* *OTHER [SPECIFY]* *DON'T KNOW* *REFUSED*
4.8	Let's say 1L of pesticides costs 20,000/=. How much more would you pay for 1L of pesticides that were certified as genuine and effective?	*[SINGLE CODE]* *1000 UGX* *2000 UGX* *5000 UGX* *10000 UGX* *OTHER [SPECIFY]* *DON'T KNOW* *REFUSED*

5. PRODUCTION & HARVEST

Now I want to ask you about your harvests. First we will talk about the most recent harvest, from November-December 2017. Then we will talk about the one before that, which started in June-July 2017.

	LOOP ENTIRE SECTION TWICE: ONCE FOR JUNE 2017 HARVEST SEASON, ONCE FOR DECEMBER 2016 HARVEST SEASON	
5.1	Which crops did you plant?	*[MULTI CODE]* *MATOOKE* *SWEET BANANA*

Jarrod Goentzel, Director, Humanitarian Response Lab
Massachusetts Institute of Technology, 77 Massachusetts Ave.,
Cambridge, MA 02139 USA
goentzel@mit.edu +1 617-253-2053

		CASSAVA *MAIZE* *BEANS* *SORGHUM* *FINGER MILLET* *RICE* *IRISH POTATOES* *SWEET POTATO* *ORANGE SWEET POTATOES* *GROUNDNUTS* *SIMSIM* *SOYA BEANS* *PIGEON PEAS* *COWPEAS* *OTHER [SPECIFY]* *DON'T KNOW* *REFUSED*
	LOOP FOR EACH CROP SELECTED @ 5.1	
5.2	How much (in acres) did you plant?	*[NUMBER]* *DON'T KNOW* *REFUSED*
	LOOP 5.3 FOR TOP 3 CROPS BY AREA PLANTED FROM 5.2	
5.3	How much did you harvest?	*[NUMBER] [KG / 100KG BAGS / SPECIFY UNIT]* *DIDN'T MEASURE* *DON'T REMEMBER* *OTHER [SPECIFY]* *DON'T KNOW* *REFUSED*
5.4	Did you store any of your harvest?	*YES* *NO* *DON'T KNOW* *REFUSED*
	IF YES *@ 5.4, LOOP 5.4A-B*	
5.4A	How much did you store?	*[NUMBER] [KG / 100KG BAGS / PERCENTAGE OF HARVEST / SPECIFY UNIT]* *DIDN'T MEASURE* *DON'T REMEMBER* *OTHER [SPECIFY]* *DON'T KNOW* *REFUSED*
5.4B	What did you use to store the crop?	*[MULTI CODE]* *GRAIN STORAGE BAGS*

Jarrod Goentzel, Director, Humanitarian Response Lab
Massachusetts Institute of Technology, 77 Massachusetts Ave.,
Cambridge, MA 02139 USA
goentzel@mit.edu +1 617-253-2053

		IMPROVED STORAGE BAGS *PLASTIC SILO* *METAL SILO* *OTHER [SPECIFY]* *DON'T KNOW* *REFUSED*
5.4C	Are you aware of improved storage bags / silos?	*YES* *NO* *DON'T KNOW* *REFUSED*
5.5	Did you sell any of your harvest?	*YES* *NO* *DON'T KNOW* *REFUSED*
	IF YES *@ 5.5, LOOP 5.5A-N.* ***If NO @ 5.5, skip to 5.5O***	
5.5A	How much did you sell?	*[NUMBER] [KG / 100KG BAGS / PERCENTAGE OF HARVEST / SPECIFY UNIT]* *DIDN'T MEASURE* *DON'T REMEMBER* *OTHER [SPECIFY]* *DON'T KNOW* *REFUSED*
5.5B	What price did you receive?	*[NUMBER] [UGX PER KG / SPECIFY UGX PER UNIT]* *DON'T KNOW* *REFUSED*
5.5C	Who did you sell it to?	*[SHORT ANSWER]* *DON'T KNOW* *REFUSED*
5.5D	Where are they located?	*[SHORT ANSWER]* *DON'T KNOW* *REFUSED*
5.5E	How did you come to know this buyer?	*[MULTI CODE]* *CAME TO ME* *WORD OF MOUTH* *REFERRED BY SOMEONE* *SAW ADVERTISEMENT* *OTHER [SPECIFY]* *DON'T KNOW* *REFUSED*
5.5F	Why did you decide to sell to this buyer?	*[MULTI CODE]* *NEARBY* *BEST PRICE*

Jarrod Goentzel, Director, Humanitarian Response Lab
Massachusetts Institute of Technology, 77 Massachusetts Ave.,
Cambridge, MA 02139 USA
goentzel@mit.edu +1 617-253-2053

		ONLY BUYER *CAME TO PICK CROP* *HAS GOOD REPUTATION* *NO PARTICULAR REASON* *OTHER [SPECIFY]* *DON'T KNOW* *REFUSED*
	IF HAS GOOD REPUTATION @ 5.5F	
5.5G	Why does this buyer have a good reputation?	*[SHORT ANSWER]* *DON'T KNOW* *REFUSED*
5.5H	Did you have a contract or formal arrangement to sell to this buyer?	*YES* *NO* *DON'T KNOW* *REFUSED*
5.5I	Have you sold to this buyer in previous seasons?	*ALWAYS* *SOMETIMES* *NEVER* *OTHER [SPECIFY]* *DON'T KNOW* *REFUSED*
	IF ALWAYS / SOMETIMES @ 5.5I	
5.5J	How many years have you sold to them?	*[NUMBER]* *DON'T KNOW* *REFUSED*
5.5K	Do you trust this buyer?	*[SINGLE CODE]* *1: NOT AT ALL* *2: NOT REALLY* *3: NOT SURE* *4: SOMEWHAT* *5: YES, DEFINITELY* *DON'T KNOW* *REFUSED*
5.5L	Did you bring the produce to the buyer, or did they come to you?	*[SINGLE CODE]* *DELIVERED TO BUYER* *THEY CAME TO PICK PRODUCE* *OTHER [SPECIFY]* *DON'T KNOW* *REFUSED*
5.5M	What was the mode of transportation?	*[MULTI CODE]* *ON FOOT* *BICYCLE* *MOTORCYCLE* *TRUCK* *BUYER PICKED UP FROM FARM*

Jarrod Goentzel, Director, Humanitarian Response Lab
Massachusetts Institute of Technology, 77 Massachusetts Ave.,
Cambridge, MA 02139 USA
goentzel@mit.edu +1 617-253-2053

		THROUGH COOPERATIVE *OTHER [SPECIFY]* *DON'T KNOW* *REFUSED*
5.5N	How much did transport cost you?	*NUMBER [UGX]* *DON'T KNOW* *REFUSED*
	IF NO @ 5.5	
5.5O	Why didn't you sell any of your harvest?	*[MULTI CODE]* *NO SURPLUS* *NO BUYERS AVAILABLE* *PRICE TOO LOW* *OTHER [SPECIFY]* *DON'T KNOW* *REFUSED*
5.6	Do you know any farmers that have received better prices for better quality crops?	*YES* *NO* *DON'T KNOW* *REFUSED*
5.7	How do you determine crop quality? What do you look for to know if crops are good quality or bad quality?	*[SHORT ANSWER]* *DON'T KNOW* *REFUSED*

6. MARKET LINKAGES

6.1	Which of these services are accessible to you? (If unclear, prompt with "Is there someone near you who could provide these services if you wanted or needed them?")	*[MULTI CODE]* *PLOWING/TILLING* *SOIL TESTING* *IRRIGATION* *WEEDING* *SPRAYING* *PRUNING* *HARVESTING* *DRYING* *GRAIN CLEANING* *HULLING* *SHELLING/THRESHING* *MILLING* *PACKAGING* *STORAGE* *TRANSPORTATION* *EXTENSION/TRAINING*

Jarrod Goentzel, Director, Humanitarian Response Lab
Massachusetts Institute of Technology, 77 Massachusetts Ave.,
Cambridge, MA 02139 USA
goentzel@mit.edu +1 617-253-2053

		OTHER [SPECIFY] *DON'T KNOW* *REFUSED*
6.2	Have you ever paid for any of these services?	*[MULTI CODE]* *PLOWING/TILLING* *SOIL TESTING* *IRRIGATION* *WEEDING* *SPRAYING* *PRUNING* *HARVESTING* *DRYING* *GRAIN CLEANING* *HULLING* *SHELLING/THRESHING* *MILLING* *PACKAGING* *STORAGE* *TRANSPORTATION* *EXTENSION/TRAINING* *OTHER [SPECIFY]* *DON'T KNOW* *REFUSED*
6.3	If you could afford to, which of these services would you use?	*[MULTI CODE]* *PLOWING/TILLING* *SOIL TESTING* *IRRIGATION* *WEEDING* *SPRAYING* *PRUNING* *HARVESTING* *DRYING* *GRAIN CLEANING* *HULLING* *SHELLING/THRESHING* *MILLING* *PACKAGING* *STORAGE* *TRANSPORTATION* *EXTENSION/TRAINING* *OTHER [SPECIFY]* *DON'T KNOW* *REFUSED*
6.4	Have you ever rented equipment, such as spray pumps, processing equipment, or a	*YES* *NO*

Jarrod Goentzel, Director, Humanitarian Response Lab
Massachusetts Institute of Technology, 77 Massachusetts Ave.,
Cambridge, MA 02139 USA
goentzel@mit.edu +1 617-253-2053

	vehicle?	*DON'T KNOW* *REFUSED*
	IF YES @ 6.4	
6.4A	What kind of equipment?	*[SHORT ANSWER]* *DON'T KNOW* *REFUSED*
6.5	Are you a member of a producer organization / cooperative?	*YES* *NO* *DON'T KNOW* *REFUSED*
	IF YES @ 6.5	
6.5A	When did you join?	*[YEAR]* *DON'T KNOW* *REFUSED*
6.5B	What services does the group provide?	*[MULTI CODE]* *INPUTS FOR PURCHASE* *BULK MARKETING* *EXTENSION AGENTS* *ADVOCACY FOR ORDINANCES* *LOANS* *BURIAL FUNDS* *SCHOOL FEES* *OTHER [SPECIFY]* *DON'T KNOW* *REFUSED*
	IF NO @ 6.5	
6.5C	Is there one in this district or sub-county that you could join?	*YES* *NO* *DON'T KNOW* *REFUSED*
6.6	Are you a member of any other groups?	*YES* *NO* *DON'T KNOW* *REFUSED*
	IF YES @ 6.6	
6.6A	What type of group?	*[SHORT ANSWER]* *DON'T KNOW* *REFUSED*

7. ACCESS TO INFORMATION

Jarrod Goentzel, Director, Humanitarian Response Lab
Massachusetts Institute of Technology, 77 Massachusetts Ave.,
Cambridge, MA 02139 USA
goentzel@mit.edu +1 617-253-2053

7.1	Which of the following types of information do you have access to?	MARKET PRICES WEATHER DATA OPPORTUNITIES TO SELL CROPS NEW PLANTING/HARVESTING/PHH TECHNIQUES NEW PRODUCTS OR SERVICES PRODUCT QUALITY ANTI-COUNTERFEIT PROGRAMS GOVERNMENT PROGRAMS NGO PROGRAMS *DON'T KNOW* *REFUSED*
	LOOP FOR EACH TYPE OF INFORMATION SELECTED IN 7.1:	
7.1A	How do you get this information?	*[MULTI CODE]* *MOBILE* *NEWSPAPER* *RADIO* *TELEVISION* *WORD OF MOUTH* *LOCAL BUSINESS* *OTHER [SPECIFY]* *DON'T KNOW* *REFUSED*
7.1B	How often do you receive this information?	*[SINGLE CODE]* *ONE TIME ONLY* *DAILY* *WEEKLY* *MONTHLY* *SOMETIMES* *NEVER* *OTHER [SPECIFY]* *DON'T KNOW* *REFUSED*
7.2	Have you attended trainings on any of the following topics? Category: Product Knowledge 1. Identifying genuine inputs 2. Safe use of chemicals and herbicides	*[SHORT ANSWER]* *NONE OF THESE* *DON'T KNOW* *REFUSED*
7.2A	Who provided these trainings?	*[SHORT ANSWER]* *OTHER [SPECIFY]*

Jarrod Goentzel, Director, Humanitarian Response Lab
Massachusetts Institute of Technology, 77 Massachusetts Ave.,
Cambridge, MA 02139 USA
goentzel@mit.edu +1 617-253-2053

	1. NGO 2. Extension worker 3. Cooperative 4. Government program 5. Input dealer 6. Buyer	*DON'T KNOW* *REFUSED*
7.3	Have you attended trainings on any of the following topics? Category: Agricultural Practices 1. No-till planting 2. When to plant/harvest crops 3. Agricultural machinery & using machines 4. Environmental mitigation practices 5. Good agronomic practices 6. Integrated pest management 7. Labor-saving technologies 8. Intercropping	*[SHORT ANSWER]* *NONE OF THESE* *DON'T KNOW* *REFUSED*
7.3A	Who provided these trainings? 1. NGO 2. Extension worker 3. Cooperative 4. Government program 5. Input dealer 6. Buyer	*[SHORT ANSWER]* *OTHER [SPECIFY]* *DON'T KNOW* *REFUSED*
7.4	Have you attended trainings on any of the following topics? Category: Business Practices 1. Record keeping for inventory/sales 2. Farming as a family business 3. Leadership 4. Entrepreneurship 5. Business plan development 6. Literacy 7. Numeracy 8. Marketing 9. Pricing 10. Business registration	*[SHORT ANSWER]* *NONE OF THESE* *DON'T KNOW* *REFUSED*

Jarrod Goentzel, Director, Humanitarian Response Lab
Massachusetts Institute of Technology, 77 Massachusetts Ave.,
Cambridge, MA 02139 USA
goentzel@mit.edu +1 617-253-2053

7.4A	Who provided these trainings? 1. NGO 2. Extension worker 3. Cooperative 4. Government program 5. Input dealer 6. Buyer	*[SHORT ANSWER]* *OTHER [SPECIFY]* *DON'T KNOW* *REFUSED*
7.5	Have you attended trainings on any of the following topics? Category: Financial Management Practices 1. Bookkeeping 2. Saving 3. E-payments 4. Credit access 5. Insurance	*[SHORT ANSWER]* *NONE OF THESE* *DON'T KNOW* *REFUSED*
7.5A	Who provided these trainings? 1. NGO 2. Extension worker 3. Cooperative 4. Government program 5. Input dealer 6. Buyer	*[SHORT ANSWER]* *OTHER [SPECIFY]* *DON'T KNOW* *REFUSED*
7.6	Have you attended trainings on any of the following topics? Category: Post-Harvest Handling and Production Practices 1. Harvesting 2. Post-harvest handling techniques 3. Grading 4. Quality control	*[SHORT ANSWER]* *NONE OF THESE* *DON'T KNOW* *REFUSED*
7.6A	Who provided these trainings? 7. NGO 1. Extension worker 2. Cooperative 3. Government program	*[SHORT ANSWER]* *OTHER [SPECIFY]* *DON'T KNOW* *REFUSED*

Jarrod Goentzel, Director, Humanitarian Response Lab
Massachusetts Institute of Technology, 77 Massachusetts Ave.,
Cambridge, MA 02139 USA
goentzel@mit.edu +1 617-253-2053

	4. Input dealer 5. Buyer	
7.3	What information about agriculture or farming as a business do you not have access to that you wish you did?	*[OPEN-ENDED]*

8. OPEN-ENDED QUESTIONS

8.1	Did your household experience any major problems this year or last year? Such as an illness or a death in the family?	*[OPEN-ENDED]*
8.2	Are there any products or services that you need to successfully run your farm that you don't have access to here?	*[OPEN-ENDED]*

9. HONORARIUM

Thank you for your participation in this survey. When this study is completed, we will make the results of our research available to your community.

We would like to give you 20,000 UGX in appreciation for your time today.

9.1	Do you have a mobile money account that we can send the money to?	*YES* *NO* *DON'T KNOW* *REFUSED*
	IF YES @ 9.1	
9.1A	Enter mobile money number	*[PHONE NUMBER]*
9.1B	Confirm mobile money number	*[PHONE NUMBER]*

Bibliography

Alene, A., Manyong, V., Omanya, G., Mignouna, H., Bokanga, M., & Odhiambo, G. (2008). Smallholder market participation under transactions costs: Maize supply and fertilizer demand in Kenya. *Food Policy*, *33*(4), 318–328.

Ashour, M., Gilligan, D. O., Hoel, J., & Karachiwalla, N. (2019). Do Beliefs About Herbicide Quality Correspond with Actual Quality in Local Markets? Evidence from Uganda. *The Journal of Development Studies*, *55*(6), 1285–1306.

Ballard, T., Coates, J., Swindale, A., & Deitchler, M. (2011). *Household Hunger Scale: Indicator Definition and Measurement Guide.* Food and Nutrition Technical Assistance II Project, FHI 360. Washington D.C.

Barrett, C. (2007). Smallholder Market Participation: Concepts and Evidence from Eastern and Southern Africa. *Food Policy*, *33*(4), 299–317. doi:http://dx.doi:10.1016/j.foodpol.2007.10.005

Bernard, T., Collion, M.-H., de Janvry, A., Rondot, P., & Sadoulet, E. (2008). Do Village Organizations Make a Difference in African Rural Development? A Study for Senegal and Burkina Faso. *World Development*, *36*(11), 2188–2204.

Bold, T., Kaizzi, J., K.C.and Svensson, & Yanagizawa-Drott, D. (2017). Lemon Technologies and Adoption: Measurement, Theory and Evidence from Agricultural Markets in Uganda. *The Quarterly Journal of Economics*, *132*(3), 1055–1100.

Bosc, P.-M., Eychenne, D., Hussein, K., Losch, B., Mercoiret, M.-R., & Rondot, P. M.-W., P. ... Sadie. (2002). *The role of rural producer in the world bank development strategy*. The World Bank. Washington D.C.

Chatterjee, S. & Hadi, A. S. (2006). *Regression analysis by example*. New York: Wiley.

Coleman-Jensen, A. (2010). U.S. Food Insecurity Status: Toward a Refined Definition. *Social Indicators Research*, *95*(2), 215–230.

Easterly, W. R. (2006). *The white man's burden : Why the west's efforts to aid the rest have done so much ill and so little good*. New York: Penguin Press.

Fafchamps, M. & Hill, R. V. (2005). Selling at the Farmgate or Traveling to Market. *American Journal of Agricultural Economics*, *87*, 717–734.

FAO. (1996). *Rome Declaration on World Food Security and World Food Summit Plan of Action*. Document WFS 96/3.

FAO. (2017). Selected Indicators for Uganda. Retrieved April 24, 2019, from http://www.fao.org/faostat/en/#country/226

FAO. (2018). *Uganda and FAO: Fostering sustainable development and better livelihoods*. Food and Agriculture Organization of the United Nations.

FAO, IFAD, UNICEF, WFP, & WHO. (2018). *The State of Food Security and Nutrition in the World 2018:Building climate resilience for food security and nutrition.* FAO. Rome.

FAPDA. (2015). Uganda Country Fact Sheet on Food and Agriculture Policy Trends.

Frelat, R., Lopez-Ridaura, S., Giller, K., M., H., Douxchamps, S., Djurfeldt, A., … van Wijk, M. (2016). Drivers of household food availability in sub-Saharan Africa based on big data from small farms. *Proceedings of the National Academies of Sciences*, *113*(2), 458–463.

Garloch, A. (2012). *Pushing the Poverty Frontiers of Inclusive Value Chain Development*. United States Agency for International Development (USAID).

Garloch, A. (2015). *A Framework for a Push/Pull Approach to Inclusive Market Systems Development*. United States Agency for International Development (USAID).

St-Germain, A.-A. F. & Tarasuk, V. (2018). Prioritization of the essentials in the spending patterns of Canadian households experiencing food insecurity. *Public Health Nutrition*, *21*(11), 2065–2078.

Goetz, S. (1992). A Selectivity Model of Household Food Marketing Behavior in Sub-Saharan Africa. *American Journal of Agricultural Economics*, *74*(2), 444–452.

Heltberg, R. & Tarp, F. (2002). Agricultural supply response and poverty in Mozambique. *Food Policy*, *27*, 103–124.

Key, N., Sadoulet, E., & de Janvry, A. (2000). Transactions Costs and Agricultural Household Supply Response. *American Journal of Agricultural Economics*, *82*(2), 245–259.

Kothair, U. (2006). An agenda for thinking about 'race' in development. *Progress in Development Studies*, *6*(1), 9–23.

Latynskiy, E. & Berger, T. (2016). Networks of Rural Producer Organizations in Uganda: What Can be Done to Make Them Work Better? *World Development*, *78*, 572–586.

Lifeyo, Y. (2017). *Market Participation of Smallholder Common Bean Producers in Malawi*. International Food Policy Research Institute (IFPRI). MaSSP Working Paper 21.

Montalbano, P., Pietrelli, R., & Salvatici, L. (2018). Participation in the market chain and food security: The case of the Ugandan maize farmers. *Food Policy*, *76*, 81–98. doi:http://dx.doi:10.1016/j.foodpol.2018.03.008

Moyo, D. (2009). *Dead Aid: Why Aid is Not Working and How There is a Better Way for Africa*. Penguin Books.

Omiti, J. M., Otieno, D. J., Nyanamba, T. O., & McCullough, E. B. (2009). Smallholder market participation under transactions costs: Maize supply and fertilizer demand in Kenya. *African Journal of Agricultural and Resource Economics*, *3*(1), 57–82.

Otte, M. & Chilonda, P. (2002). *Cattle and Small Ruminant Production Systems in sub-Saharan Africa - A Systematic Review*. Food and Agriculture Organization of the United Nations.

Peduzzi, P., Concato, J., Kemper, E., Holford, T., & Feinstein, A. (1996). A simulation study of the number of events per variable in logistic regression analysis. *Journal of Clinical Epidemiology*, *49*(12), 1373–1379.

Sibhatu, K., Krishna, V., & Qaim, M. (2015). Production diversity and dietary diversity in smallholder farm households. *Proceedings of the National Academies of Sciences*, *112*(34), 10657–10662.

Sturdivant, R. X. (2013). *Applied logistic regression*. Wiley.

Uganda Bureau of Statistics. (2014a). *The National Population and Housing Census 2014 – Provisional Results*. Retrieved from %7Bhttp://www.ubos.org/onlinefiles/uploads/ubos/NPHC/NPHC%5C%202014%5C%20PROVISIONAL%5C%20RESULTS%5C%20REPORT.pdf%7D

Uganda Bureau of Statistics. (2014b). *Total Population by Sex, Total Number of Households and proportion of Households headed by Females by Subcounty and Parish, Central Region, 2014*. Retrieved from %7Bhttps://www.ubos.org/wp-content/uploads/publications/03_2018Population_by_Parish_Census_2014_Central_Region.pdf%7D

Uganda Bureau of Statistics. (2014c). *Total Population by Sex, Total Number of Households and proportion of Households headed by Females by Subcounty and Parish, Eastern Region, 2014*. Retrieved from %7Bhttps://www.ubos.org/wp-content/uploads/publications/03_2018Population_by_Parish_Census_2014_Eastern_Region.pdf%7D

Uganda Bureau of Statistics. (2014d). *Total Population by Sex, Total Number of Households and proportion of Households headed by Females by Subcounty and Parish, Northern Region, 2014*. Retrieved from %7Bhttps://www.ubos.org/wp-content/uploads/publications/03_2018Population_by_Parish_Census_2014_Northern_Region.pdf%7D

Uganda Bureau of Statistics. (2014e). *Total Population by Sex, Total Number of Households and proportion of Households headed by Females by Subcounty and Parish, Western Region, 2014*. Retrieved from %7Bhttps://www.ubos.org/wp-content/uploads/publications/03_2018Population_by_Parish_Census_2014_Western_Region.pdf%7D

Uganda Bureau of Statistics. (2016). *The National Population and Housing Census 2014 – Main Report*.

Uganda Bureau of Statistics. (2018). *Uganda National Household Survey 2016/2017 Report*.

USAID. (2011). *Understanding Facilitation*. United States Agency for International Development (USAID).

USAID. (2015). *The Facilitation Approach at USAID: A Discussion Paper*. United States Agency for International Development (USAID).

Vance, C. & Geoghegan, J. (2004). Modeling the Determinants of Semi-Subsistent and Commercial Land Uses in an Agricultural Frontier of Southern Mexico: A Switching Regression. *International Regional Science Review*, *27*, 326–347.

Vittinghoff, E. & McCulloch, C. (2007). Relaxing the rule of ten events per variable in logistic and Cox regression. *American Journal of Epidemiology*, *165*(6), 710–718.

WFP. (2017). *Monthly Market Monitor - Uganda* (tech. rep. No. 44). The World Food Programme.

Yusuf, S. A., Ashagidigbi, W. M., & Mustapha, T. I. (2015). Choice and Level of Dry Season Vegetable Market Participation Under Tropical Conditions. *International Journal of Vegetable Science*, *21*, 272–280.