Three Essays on Agriculture, the Environment, and Peer Networks in Western Kenya

A Dissertation Presented to the Faculty of the Graduate School of Cornell University in Partial Fulfillment of the Requirements for the Degree of Doctor of Philosophy

David Michael Murphy
August 2018

Three Essays on Agriculture, the Environment, and Peer Networks in Western Kenya

David Michael Murphy, Ph.D.

Cornell University, 2018

This dissertation investigates selected issues related to rural poverty in Sub-Saharan Africa and evaluates whether information – such as that regarding a farmer's soil nutrient levels or agroforestry practices – can provide rural households with the needed skills and knowledge to break the cycle of poverty. The setting of the research is western Kenya, a densely populated area where agriculture is the primary economic activity. Many of the farms in the area are very small, often only one acre, and the lack of effective fertilizer and other input use has contributed to crop yields far below potential. The essays in this dissertation seek to determine whether information regarding agroforestry practices and soil nutrient levels can be effective tools for economic development, and, additionally, analyze the effect. the effect of information spread through peer networks on agricultural productivity.

The first essay (Chapter 2) focuses on the potential of agroforestry to reverse environmental degradation in Kenya by increasing tree coverage and providing a renewable source of household fuelwood. The results show that the various sources of fuelwood, such as collecting fuelwood from off the farm and producing it on the farm through agroforestry, are not readily substituted by households in western Kenya, primarily due to gender-specific labor roles within households. As a result, there is an effective limit on the extend of agroforestry in the area unless gender norms for household labor change. In Chapter 3, we analyze the effect of soil information transfers on household agricultural input demand by sampling soils from small-scale farms in western Kenya. After testing the soils and returning the soil nutrient results

to farmers, we used experimental auctions both before and after the farmers received the results to measure changes in agricultural input demand. We find that soil testing and input recommendations do have a significant effect on farmer input demands, though the results are heterogeneous by input type and gender. Chapter 4 analyzes peer network effects on agricultural productivity in western Kenya. We find that female peers tend to increase male farmer's productivity, and hypothesize that this is due to the higher levels of social capital (proxied by network centrality) found among men in the village samples, which increases their bargaining power with those less central in the networks (often women). The three essays point to the important role of gender in resource allocation, social dynamics, and factors influencing agricultural productivity in western Kenya. Overall, this dissertation shows both the opportunities of information diffusion for further economic development in Kenya, but also the challenges posed by cultural norms that limit the spread of productivity enhancing information.

Biographical Sketch

David Michael Augustus Murphy was born in New Haven, Connecticut and raised in Natick, Massachusetts and Geneva, New York. He graduated from the State University of New York at Geneseo in 2009, *summa cum laude*, with bachelor's degrees in economics and international relations. After serving two years in the Peace Corps in Armenia, he attended the Fletcher School at Tufts University, obtaining a master's degree in development economics and environmental and energy policy, before beginning his Ph.D. program at Cornell University.

To my wife, Ani, and daughters, Ekaterina and Alexandra.

Acknowledgments

I first would like to thank David Lee, my Dissertation Committee chair, for his guidance and support not only directly related to the dissertation itself, but throughout my entire PhD program. His encouragement, counsel, field visits, and countless hours spent reviewing survey instruments and manuscript drafts were crucial to the completion of this dissertation. I am also grateful to the other members of my Committee: Jura Liaukonyte, Johannes Lehmann, and Janice Thies. They gave generously of their time, guidance and expertise in helping me develop the content of this research, and provided many useful suggestions in improving this dissertation. I further want to thank Johannes Lehmann, who served as Lead PI of the project "Fueling Local Economics and Soil Regeneration: Biofuels and Biochar Production for Energy Self-sufficiency and Agricultural Sustainability," through Cornell's Atkinson Center for a Sustainable Future, which substantially funded my research. I also want to thank Janice Thies, who provided significant support for this research through the National Institute of Food and Agriculture Project: "Improving Bean Yields by Reversing Soil Degradation and Reducing Soil Borne Pathogens on Small-holder Farms in Western Kenya."

Many other individuals also made significant contributions to the research reported in this dissertation and they have my sincere gratitude. Julia Berazneva served as a mentor during the first years of my PhD program and allowed me to use data she collected for Chapter 2 of this dissertation. I am very grateful for her guidance and support, particularly in the difficult early days of the PhD program. I also wish to thank Dries Roobroeck at the International Institute of Tropical Agriculture (IITA) in Nairobi who provided logistical support for the field research, organized and interpreted the soil analysis results for the Kenyan farms, and made other valuable contributions to the research in Chapter 3. Finally, Edmundo Barrios, formerly of the World Agroforestry Centre (ICRAF) in Nairobi and now with the Food and Agricultural Organization of the United Nations, together with Johannes Lehmann, enabled me to participate in the Global Environmental Facility project in Kenya, crucial for the results found in Chapter 4, and I greatly appreciate this support.

I also wish to thank Harry Kaiser, for inspiring the methods used in this research, as well as Chris Barrett for his important insights. Linda Sanderson provided valuable administrative support throughout the period of my dissertation research. Seminar participants at Cornell University, SUNY Geneseo, Texas Lutheran University, the University of Evansville, IITA (Nairobi office), and at the 2015 and 2017 annual AAEA meetings and the 2017 SEA annual meetings also provided valuable feedback on this research.

I want to thank the many organizations whose generous support made this research possible. These include: a U.S. Borlaug Fellowship in Global Food Security; Yossie Hollander, the Fondation des Fondateurs, and the David R. Atkinson Center for a Sustainable Future at Cornell University; the NIFA and GEF programs noted above; and finally, several grant programs at Cornell: the Peter Rinaldo Sustainable Development Fund in the Dyson School; an International Research Travel Grant from the Mario Einaudi Center for International Studies; and a Richard Bradfield Research Award and "Frosty" Hill and Andrew W. Mellon research grants from the College of Agriculture and Life Sciences. Significant support was also provided by the International Institute for Tropical Agriculture and the World Agroforestry Centre in Kenya, whose generous staff provided their expertise and resources towards this research.

I would also like to thank the members of my field staff whose assistance was key to successfully completing the research reported here: James Agwa Otieno, Samuel

Omollo, Electine Khisa Nanyama, Benjamin Harim Awiti, Martin Wakoli, Ezekiel Avedi, and Henry Agallo. I am also grateful for the assistance of numerous Kenyan village elders, chiefs, and guides who are too numerous to name, but without whom we could not have collected the data used in this dissertation.

Last, but certainly not least, I wish to thank my wife, Ani, for supporting me during the past five years of my research. She sacrificed so much to provide for our family during these years, both in the United States and in Kenya, and this dissertation was only possible with her love, encouragement, and assistance.

Numerous others have supported my work over these last few years but are too many to name. I thank everyone for their support and assistance in this endeavor.

Contents

Chapter 1: Introduction

The World Bank estimates that 67.5 percent of the population in Sub-Saharan Africa (SSA) earn the equivalent of less than $3.20 dollars per day (PPP), indicating extensive levels of poverty across the continent (The World Bank, 2018). Farm yields in Africa are generally low and far below potential (Zingore *et al.*, 2015), a major factor limiting growth in rural incomes. In addition, climate change has increasingly led to unpredictable weather patterns, and together with deforestation, poor land management, and damaging environmental practices, has increased soil erosion in many places (García-Ruiz *et al.*, 2017). This soil degradation, exacerbated by the lack of effective input use, has contributed to stagnating crop yields (Zingore *et al.*, 2015). Due to these various obstacles, rural households often struggle to increase their incomes and accumulate assets. Instead, decreasing yields lower farmers' incomes, which leads to a reduced capacity to purchase agricultural inputs, further decreasing yields. This " poverty trap" leads to a downward spiral of poverty from which it is often difficult for rural households to escape (Dasgupta, 1997).

This objective of this dissertation is to investigate selected problems related to rural poverty in Sub-Saharan Africa and evaluate potential tools that policymakers and institutions can use – such as providing better information regarding agroforestry practices and on-farm soil nutrient levels – which can provide rural households with the needed skills and knowledge to break the cycle of poverty. The chapters in this dissertation focus on the use of information, on environmental degradation related to deforestation and soil nutrient deficiencies, and on social inequities, especially with respect to gender, that can potentially lead to unequal trajectories of economic growth within rural African communities. Whether these resource and social obstacles can be overcome depends in large part on access to information, which is often limited by constraints including financial costs, cultural practices, and the structure of local social networks.

1.1 Rural poverty and information diffusion in Sub-Saharan Africa

A large volume of studies over several decades has shown that information can greatly shape agricultural and environmental behavior in general, and in SSA, specifically (see Foster and Rosenzweig (2010) and Nakasone *et al.* (2014) for reviews). Historically, information has often been communicated from institutions to farmers in SSA is through extension field agents, though evidence has been mixed as to the efficacy of this system (Anderson and Feder, 2004). Non-governmental organizations have also played an increasing role in many countries where traditional extension services have foundered. With the technological revolution in communications, there has also been an expansion of efforts to transfer information through mobile phones (Aker, 2011), which many people in rural areas of SSA now own. Recent evidence has shown that these efforts may be effective in encouraging farmer adoption of improved pesticides (Cole and Fernando, 2012) and in increasing agricultural productivity (Casaburi *et al.*, 2014). Farmer field schools, organized by governments or NGOs, often train participants in the use of new technologies with the hopes that these farmers will then communicate improved information to others in their villages. Studies have shown that field schools can be effective at technology diffusion (Davis *et al.*, 2012), though not necessarily in every situation (Feder *et al.*, 2004a,b).

Underlying the farmer field school paradigm is the general consensus, supported by an extensive literature, that individuals often learn most effectively from their peers.

Proximity to those who live nearby enables individuals to experience first-hand the success of new technologies or practices (Kondylis *et al.*, 2017). An individual's peers also have a "safety credibility," providing increased trust in the information they provide (Rogers, 1995). Indeed, studies have shown that information diffusion is typically more effective when communicated through peers than through extension agents (Conley and Udry, 2010; Krishnan and Patnam, 2014). However, as demonstrated in this dissertation, information diffusion through peer connections is not always efficient, and can create inequities between those more central in village social networks and those on the periphery, while providing relatively greater benefits to those with more advantageous social connections (Chapter 4). In addition, conformity effects within a village can prevent adoption of a practice or technology, biasing individuals towards traditional methods (Moser and Barrett, 2006). Moreover, entrenched cultural practices or norms, such as strict gender divisions in household labor activities, can limit the impact of new information and prevent new technologies and practices from reaching optimal scales (Chapter 2). However, for farmers in rural SSA to accumulate assets in the face of soil degradation and unpredictable weather patterns, they must incorporate new and improved information regarding optimal agricultural practices and technologies. In the chapters of this dissertation, we focus on two practices in particular: agroforestry and the use of organic inputs.

Over the past century, forests have degraded substantially in Sub-Saharan Africa. In a recent study, Aleman *et al.* (2018) estimate that closed canopy forest coverage has decreased on the continent by 22 percent since 1900, with 290,000 hectares of forest lost in the first decade of the 21st century alone (Food and Agriculture Organization: UN, 2010). Increasing demands for cropland and livestock are major contributors to deforestation in SSA, although fuelwood collection and charcoal production for household energy use is the single largest cause (Hosonuma *et al.*, 2012). Not only does forest degradation increase the effort required by households to collect fuelwood, but it also decreases soil health by increasing erosion and decreasing the soil's ability to retain water (Karltun *et al.*, 2013). Agroforestry, or the planting of trees on-farm as a crop, is a potential solution that has become more common in recent years. Agroforestry enables farmers to obtain fuelwood for household energy consumption, and also increases soil organic matter and soil moisture, and decreases nutrient leaching (Mbow *et al.*, 2014). While agroforestry has many benefits, traditional customs and practices, including strict divisions in on-farm labor roles between genders, can limit the adoption and scale of agroforestry in SSA (Chapter 2).

Nutrient deficiencies are common in soils across SSA, where inputs into the soil have not kept pace with agricultural intensification (Drechsel *et al.*, 2001). One major result is low crop yields: on average, maize yields in SSA have been claimed to be no more than sixteen percent of their potential (Nkonya *et al.*, 2016). Yet, in many areas of Kenya, inorganic fertilizer use is well established. In fact, in some regions, inorganic fertilizers such as DAP (diammonium phosphate) are used in excess of profitable levels (Sheahan *et al.*, 2013). Moreover, many soils in the area are unresponsive to inorganic fertilizer use due to low levels of organic matter (Vanlauwe *et al.*, 2001). Studies find that farmers in SSA must increase their use of organic inputs to recover soil nutrient levels and increase crop yields (Mateete *et al.*, 2010). Organic inputs, such as crop residues, farmyard manure, and compost are traditionally used in farming, but often at insufficient levels due to their high opportunity cost (Place *et al.*, 2003; Berazneva *et al.*, 2017).

In addition to exploring methods to increase soil fertility and access to renewable sources of household energy, a major theme running through the essays in this dissertation relates to gender differentials: I explore the impacts of differences between men and women in such matters as household labor roles, their valuation of agricultural inputs, and crop yields. In western Kenya, as in many areas of SSA, women are responsible for both agricultural and household work, often creating a large time burden on women known as the "double workday" (Kes and Swaminathan, 2006). Female-managed plots are generally allocated fewer resources within the household (Udry *et al.*, 1995), have lower crop yields (Slavchevska, 2015), and women are less likely to have adopted improved agricultural practices on plots that they manage (Ndiritu *et al.*, 2014). Additionally, women tend to be less centrally positioned within their village social networks, limiting their ability to effectively collect and incorporate agricultural information on their farms (Chapter 4).

In the essays in this dissertation, I use data collected from western Kenya to analyze the effect of agricultural information on farmer management strategies, their demands for agricultural inputs, and crop productivity. I first examine the elasticity of demand for household fuelwood sources (Chapter 2) to determine if improved information regarding agroforestry practices and alternatives is likely to be effective in decreasing environmentally damaging fuelwood collection practices. In Chapter 3, I detail an experiment where we provided detailed, personalized information to respondents regarding various agricultural inputs to determine whether their demands change for these inputs based on the information provided. The third essay (Chapter 4) focuses on the effect of social capital, measured by the centrality of household members within their social network, on agricultural productivity, and, specifically, investigates the gender differences in these peer effects. These essays collectively add to several different strands of the development economics literature. They contribute to the literatures on household labor, resource management, information provision and technology adoption, social capital, gender and development, and the effects of peer linkages on economic growth. Moreover, the second essay adds significantly to the literature on experimental auctions in developing countries, which are increasingly being used as a tool to measure willingness to pay for various products, including fertilizer, in SSA.

1.2 Research area and data

The data used in this thesis come from two separate sources. Data for the first essay (Chapter 2, on agroforestry in western Kenya) were collected by Julia Berazneva, PhD, for her dissertation "Reconciling Food, Energy, and Environmental Outcomes: Three Essays on the Economics of Biomass Management in Western Kenya" (Cornell University, 2015) and are used with her permission here. These data were collected in 2011-2012 from households in the western Kenyan highlands that comprise areas of the former provinces of West Kenya, Nyanza, and the Rift Valley. From 15 villages, 315 randomly selected households were chosen to participate in the study. These 15 villages were divided into three research "blocks," and the villages in each block were in relative proximity to the others. Farms of the participants in this study had their soils tested for nutrients and the trees on their farms were counted and measured. Participants were also asked questions related to household production activity, socio-demographic characteristics, and labor market participation, among other topics. Additional details regarding this data set are provided in the 2015 Berazneva dissertation (Chapter 1) and in Chapter 2 of this dissertation.

I collected the data used in the latter two essays (Chapters 3 and 4) from farmers surveyed in four counties of western Kenya: Bungoma, Busia, Kakamega, and Nandi (Figure 3.A.2). The former three counties are in the former Western province of Kenya, while Nandi County is adjacent to Kakamega county and in the former Rift Valley Province. Agriculture is the primary economic activity in these counties. Most farms are small (less than five acres) and managed by individual households. Over the past century, maize has emerged as the primary food crop grown in these areas, while beans, bananas, and cassava are also widely cultivated. Major cash crops include sugarcane (Bungoma, Busia, and Kakamega counties), tobacco (Busia and western Bungoma counties), and tea (Nandi County). While many farmers attempt to grow enough crops to sell to market after harvest, many others practice subsistence farming or are net buyers of foodstuffs, relying on off-farm wage income to supplement their household agricultural production. Household incomes are generally low: in 2016, the year these surveys were conducted, an agricultural day laborer earned the equivalent of 200 to 300 Kenyan Shillings for a full day's work (roughly two to three U.S. dollars at the exchange rates then prevailing). Since it is rare for more than one spouse to work off-farm, per-person incomes in households with small landholdings in these rural areas are very low.

This research in this dissertation builds on two projects currently in implementation through Cornell University in partnership with the International Institute of Tropical Agriculture (IITA) and the World Agroforestry Center (ICRAF). One of these projects, "Improving bean yields by reversing soil degradation and reducing soil borne pathogens on small-holder farms in western Kenya," began in 2012 and is funded by USDA's National Institute of Food and Agriculture. This project has tested various agricultural inputs on seventy different researcher and farmer-managed plots across Kakamega, Bungoma, and Busia counties in Kenya. Researchers in this project have analyzed how various agricultural inputs affect crop yields, nutrient levels, and pathogen suppression. The inputs tested include biochar, vermicompost, mazao fertilizer, agricultural inoculants, and various combinations of these products (these inputs are defined and described in detail later in this dissertation). The results have found significant increases in bean yields through the use of these organic inputs. A second project, "Sharing Knowledge on the Use of Biochar for Sustainable Land

Management" focuses on biochar-related information diffusion and is funded through the Global Environmental Facility. This project, which took place in Kenya in 2016-2017, held workshops and demonstrations in Nandi County to educate farmers about biochar. Farmers were divided into various treatments of biochar and other inputs and provided with biochar for their experimental plots. The project found farmers' knowledge of biochar production and use increased substantially, and biochar combined with DAP fertilizer produced the greatest yield increases among participating farmers.

For the research reported in Chapters 3 and 4 below, we identified participants by acquiring village-level household lists from area chiefs in the areas of these research projects. We then randomly selected participants from these village lists, and visited and interviewed each household head and his/her spouse between July and November 2016 (Bungoma, Busia, and Kakamega counties), and March to April 2017 (Nandi county). In total, we sampled 992 individuals in 612 households. When visiting the households, local enumerators asked each participant a wide range of questions including regarding agricultural production and input use, household assets, and demographic characteristics. To analyze the effects of soil nutrient information transfers on farmer behavior, we also collected soil samples from the primary maize plots on each household's farm. With help from IITA, these soil samples were analyzed to determine the levels of various soil nutrients as well as the pH and texture of the soils. To investigate issues related to information diffusion and the effect of information on farmer behavior (Chapter 3), we included an experimental auction module and a peer network model after Conley and Udry (2010) with the survey instrument that was used to gather data for Chapters 3 and 4. At the time of the survey, we used handheld GPS devices to measure the precise locations of the homesteads and plots, as studies have shown that farmers in SSA generally to not provide accurate estimates about the size of their land (Carletto *et al.*, 2015).

1.3 Overview of dissertation

The first essay in this dissertation, entitled "Fuelwood Source Substitution, Gender, and Shadow Prices in Western Kenya" (Chapter 2), analyzes households' choices of fuelwood sources, which have ramifications for fuelwood scarcity and deforestation in western Kenya. Fuelwood scarcity is a major environmental problem in much of the developing world, and often places a major burden particularly on women and children in the rural areas of these countries. Consequently, many governments, donors and non-governmental organizations have encouraged on-farm fuelwood production and agroforestry practices. Whether, however, fuelwood from different sources can be easily substituted is an important empirical question, as the degree of substitutability can depend on local markets and households' resource endowments and incomes. If limited substitution exists in this context, then this will decrease the impact of programs seeking to increase fuelwood sourced from agroforestry. In this essay, I examine the substitution among three fuelwood sources in rural western Kenya: fuelwood collected off-farm, fuelwood produced on-farm, and that which is purchased. Using household-specific shadow prices for fuelwood and male and female wages, I find that strict gender divisions in household labor result in limited substitution between fuelwood sources. Among the implications of these findings are that programs and policies promoting agroforestry will have limited success without first addressing the structural differences in labor markets.

The research in the second essay, "Underground Knowledge: Estimating the Impacts of Soil Information Transfers through Experimental Auctions" (Chapter 3), analyzes the impacts of soil information transfers on farmer behavior to learn whether these transfers can be an effective method to aid in recovering soil nutrients in Sub-Saharan Africa (SSA). Soil degradation more often than not necessitates the use of fertilizers to increase soil fertility and improve crop yields. However, rural smallholders usually do not have sufficient information about their soil nutrient levels to make profit-maximizing decisions about fertilizer usage. This leads to sub-optimal combinations of inputs, and frequently, further soil and environmental degradation, food insecurity, and a reduced ability to increase household incomes. In this essay, I report on a two-round experimental auction approach (after BDM) to test whether providing soil test information and fertilizer recommendations to farmers affected their behavior and ability to optimize their input choices. I auctioned packages of inorganic and organic inputs,[1] dividing farmers into different soil fertility information treatments, and analyzed the data using triple difference estimation methods. I find that providing soil fertility information has significant effects on farmers' demands for agricultural inputs: recommendations to farmers to use inorganic fertilizer increase willingness to pay by 61% compared to the baseline, while recommendations to use organic fertilizers lead to more nuanced effects that depend on the gender of the respondent. Overall, the research reported in this essay strongly suggests that soil information transfers can enable more effective fertilizer optimization among farmers and can potentially be a cost-effective way to reverse localized ecological degradation and increase crop yields.

In the third essay in the dissertation, "Social Capital and Gendered Peer Networks," I use peer network data coupled with agricultural production data collected from surveys of rural households in western Kenya to analyze the relationship between an individual's position in a social network and agricultural productivity. Women in

[1] The inputs auctioned were biochar, vermicompost, diammonium phosphate (DAP), cow manure, and combinations of these inputs. Additional details are included in Chapter 3.

SSA are often located at the periphery of social networks, which limits their accumulation of social capital, leading to a weakened bargaining position for information exchange and potential negative effects on economic outcomes. Using social network data from nearly 1,000 individuals from four counties, I construct various measures of network centrality and demonstrate the relative periphery of female farmers in these social networks. Then, using linear-in-means estimation, I find evidence consistent with differential levels of bargaining power between genders in the sample: among men, we find that increasing the share of female peers corresponds with increases in agricultural productivity (controlling for a multitude of other factors), while no benefits accrue to women based on the gender share of their peers. These results show that men are advantaged in their placement within social structures in these villages, which increases their ability to obtain information. This study adds an important economic dimension to the existing literature on gendered social networks in SSA and points to the detrimental effects of lower social capital accumulation among women on their economic outcomes.

These essays together highlight both the possibilities and obstacles to productivity-enhancing information diffusion in SSA. While information about agroforestry can provide benefits both to farmers and their environment, Chapter 2 of the dissertation shows that cultural norms can lead to sub-optimal allocation of land for on-farm fuelwood collection. In Chapter 3, results from experimental auctions for agricultural inputs show that soil information transfers have strong impacts on farmer behavior. In Chapter 4, however, we see that the location of individuals within their social network is related to the quantity of information they receive, affecting their agricultural productivity. Overall, this dissertation shows that the role of new information in increasing the economic opportunities for farmers in SSA is extremely important, but our results also identify the obstacles impeding the diffusion of information and how these limitations can potentially be addressed. Cultural norms, in particular, often limit the flow of information. While these norms typically change slowly, the general direction in SSA is towards more open and equitable societies that aid in the spread of improved knowledge regarding agricultural practices and energy sources. The results from this dissertation thus contribute towards our understanding of information diffusion in rural SSA, with potential applications to other developing country regions as well. In the final chapter of the dissertation, I summarize the conclusions and impacts from these essays, as well as opportunities for future research.

Chapter 2: Fuelwood Source Substitution, Gender, and Shadow Prices in Western Kenya

Much of the world's population, especially the poor in rural areas of developing countries, rely on biomass (crop residues, animal dung, and fuelwood) for basic household energy requirements. In rural Sub-Saharan Africa (SSA), for example, 80 percent of the population depends on biomass for daily cooking fuel, with most of the biomass coming from fuelwood (International Energy Agency, 2014). This dependency on fuelwood carries many implications for the environment and for households' livelihoods, gender roles, and health.

The environmental impacts of fuelwood use include greenhouse gas (GHG) emissions and deforestation. In the year 2000, net residential GHG emissions in SSA totaled 79 million metric tons of carbon (MtC), 61 percent of which were due to fuelwood use (Bailis *et al.*, 2005). These emissions are projected to increase. Under a "business as usual" scenario, cumulative residential GHG emissions in SSA are estimated to reach 6.7 billion tons of carbon by 2050, or 134 MtC per year – the equivalent of more than four large coal-fired power plants operating at full capacity over the period (World Wide Fund for Nature, 2007). Fuelwood off-farm collection, along with charcoal production, also contributes to widespread deforestation (Hosonuma *et al.*, 2012). From all sources, the last decade witnessed 13 million hectares of trees lost every year globally (Food and Agriculture Organization: UN, 2010), including 290,000 hectares in Africa (Joint Research Center: The European Commission, 2013). Other environmental concerns include the loss of animal habitat and decreases in soil nutrients and moisture, leading to desertification (World Meteorological Organization, 2010).

In addition to environmental concerns, the use of fuelwood as an energy source places a particular burden on women in the household, given that women in SSA are often responsible for both fuelwood collection and food preparation. Increasing scarcity of fuelwood means increasing collection times. This adds to the labor burden of women, as traditional roles such as raising children, cooking, and other household tasks create a "double workday" and mean that women often work much longer hours than their male spouses (Kes and Swaminathan, 2006). Moreover, smoke from all biomass sources (including fuelwood) is associated with millions of deaths per year in SSA due to respiratory diseases (Lim *et al.*, 2012). As incomes increase, households are unlikely to quickly switch in large numbers to more modern fuels such as kerosene or LPG (Cooke *et al.*, 2008). Instead, households often engage in "fuel stacking," gradually adding new sources of energy while continuing to consume traditional biomass such as fuelwood (Masera *et al.*, 2000; Van Der Kroon *et al.*, 2013).

Renewable forestry management has frequently been viewed as a potential remedy for these related problems. On-farm fuelwood production and agroforestry, for example, can reduce the environmental impacts of fuelwood and charcoal use (Mbow *et al.*, 2014) and mitigate household search costs associated with deforestation. Since the 1970s, many research and non-governmental organizations have focused on promoting agroforestry in SSA,[2] with many projects paying particular attention to transferring agroforestry skills to women (Bradley and Huby, 1993; Maathai, 1993; Kiptot and Franzel, 2012). These projects have been influential in shifting on-farm tree management from non-fuelwood uses to fuelwood usage and in increasing the absolute number of trees on-farm.

The main goals of this paper are to investigate 1) whether household fuelwood

[2] In Kenya, for example, the Green Belt Movement and the Stockholm Environmental Institute are two of the best known organizations. The World Agroforestry Centre (ICRAF) is also very active in the promotion of agroforestry in the area studied here.

sources (fuelwood collected off-farm, that produced on-farm, or purchased) are close substitutes or differentiated products, and 2) whether gender roles persist in fuelwood on-farm production and off-farm collection. Few studies have analyzed whether multiple fuelwood sources themselves are close substitutes to one another using shadow prices. The answer to this question, however, can have important implications for policies centered on reducing forest degradation or promoting agroforestry to produce a renewable fuelwood source. Most of the empirical literature examining household energy needs thus far has focused on understanding the substitution between aggregate fuelwood consumption (or consumption from a single source) and other biomass options such as agricultural residues, and has relied on data from South Asia, with only a few studies of fuelwood demand in SSA (table 2.1).

We examine fuelwood substitution and gender roles in the context of western Kenya. In the area of our study, fuelwood markets are imperfect and household production and consumption decisions are non-separable.[3] Following Heltberg *et al.* (2000) and Palmer and MacGregor (2009), we modify the agricultural household model to focus on the substitution among different fuelwood sources and on the role of a household's labor endowment. Empirically, we first estimate shadow prices for different fuelwood sources using household-specific male and female wages. Controlling for potential selection bias and endogeneity, we then estimate demand equations for different sources of fuelwood: fuelwood collected off-farm, fuelwood produced on-farm, and that bought at the market. The data used in the empirical estimation come from a recent detailed production and consumption survey of over 300 households in the western Kenyan highlands (Berazneva *et al.*, 2017). Since the majority of existing fuelwood demand studies focus on South Asia, our analysis offers new evidence

[3]We test for and confirm non-separability in the household energy market following Dillon and Barrett (2017).

of fuelwood consumption patterns in East Africa.

We show that cross-price elasticities between fuelwood sources are very low (ranging from 0.02 to 0.24), suggesting that Kenyan households do not readily substitute between fuelwood sources. As expected, we also find that own-price demand elasticities for non-purchased fuelwood are negative and inelastic (-0.55 to -0.61). As the implicit cost increases for a particular source of fuelwood, there is only limited substitution with other fuelwood sources. This limited substitution is, we suggest, partially explained by gender roles. The data show that women are primary collectors of fuelwood off-farm and men are primary producers of fuelwood on-farm. This gender division is also reflected in the econometric results, with female and male shadow wages tied to off-farm fuelwood collection and on-farm fuelwood production, respectively. It appears that the lack of labor substitutability contributes to limited opportunities to substitute between fuelwood sources.

This paper is related to a rich body of research in economics that examines household energy decisions. As fuelwood scarcity increases, households react to the rising implicit cost of obtaining fuelwood in various ways: they substitute other fuels, purchase fuelwood from the market, plant trees on their own farm, adopt higher efficiency stoves, or increase off-farm collection times (see Cooke *et al.* (2008) for a review of the literature). The empirical evidence in support of these hypotheses, however, has been mixed. Several studies that look at the use of fuelwood, crop residues, and animal dung find no evidence of substitution (Pattanayak *et al.*, 2004; Palmer and MacGregor, 2009; Damte *et al.*, 2012), while others find evidence of complementarity between fuelwood and cut grass and leaf fodder (Cooke, 1998a) and animal dung (Mekonnen, 1999). Other studies analyze fuelwood sourcing from forest reserves. Cooke (2014), for example, shows that the level of restrictions on fuelwood collection in the community managed forests in South Asia determines the quantity that is collected from other sources; while in Uganda, Miteva *et al.* (2017) find proximity to forest resources to increase the likelihood of fuelwood collection (and the likelihood of purchasing fuelwood increases as the distance to market decreases). Amacher *et al.* (1993), Amacher *et al.* (1996, 1999) and Pattanayak *et al.* (2004) find that owning more efficient stoves leads to a significant decrease in fuelwood consumption, although Heltberg *et al.* (2000) find no such effect. Finally, the response of labor supply to increases in the scarcity of fuelwood (and in the implicit cost of fuelwood) is always positive, but the evidence is mixed as to whether the magnitude is greater than or less than that of the own-price elasticity (Amacher *et al.*, 1996; Cooke, 1998a; Heltberg *et al.*, 2000; Palmer and MacGregor, 2009).

To our knowledge, no existing study has specifically focused on the substitution among rural households' three major sources of fuelwood – fuelwood collected off-farm, produced on-farm, and purchased. In perfectly functioning fuelwood and labor markets, the costs of the fuelwood coming from different sources would be equal (given the same quality of fuelwood demanded). Market imperfections, however, can create divergences between the household-specific implicit or shadow prices of different fuelwood sources and the market price, and can lead to source-specific own-price and cross-price elasticities. Several studies that estimate the demand for off-farm collection and on-farm fuelwood production do not, however, estimate cross-price elasticities to measure their substitution (Amacher *et al.*, 1993; Heltberg *et al.*, 2000). They also use collection time as a proxy variable for the shadow price, and in the case of Heltberg *et al.* (2000) combine fuelwood produced on-farm with crop residues and animal dung. In contrast, we estimate own-price elasticities for three separate fuelwood sources using shadow prices and market prices, and then analyze the substitution patterns among the sources given by cross-price elasticities. An understanding of household substitution among fuelwood sources can help reveal whether households treat fuelwood as a homogeneous product, as is often implicitly assumed in the literature, or whether it is a differentiated product based on its source. If indeed households do differentiate among fuelwood based on its source and if strong preferences exist for certain fuelwood sources, policies that promote agroforestry will likely be ineffective unless the factors influencing these preferences are first addressed.

The rest of this paper is organized as follows. In Section 2 we describe the background to the research area and data collected in western Kenya in 2011-2012. Section 3 presents a non-separable agricultural household model that takes into account the various fuelwood sources and household labor endowments. In Section 4 we describe our empirical strategy, which includes maximum likelihood estimation of the Heckman estimators to control for selection bias in the imputed wages and fuelwood source groups, and two-stage least squares estimation to control for endogeneity in the shadow prices. We present our results in Section 5 and highlight their management and policy implications in Section 6.

2.1 Background

Forests cover less than seven percent of Kenya's land area, yet they make a significant contribution to the national economy and provide many direct and indirect goods and services to its people (Republic of Kenya: Ministry of Environment, 2014). Historically, Kenyan forests have been cleared both to create land for agriculture and for the sale and subsistence use of forest products. In recent years, deforestation has been largely driven by the latter, as the private consumption of forest products doubled between 2000 and 2010 (Crafford *et al.*, 2012). The rate of deforestation has averaged about 5,000 hectares per year in the Kenyan montane forests (Crafford *et al.*, 2012) and has had substantial effects on many aspects of the Kenyan environment and economy. Evidence suggests, for example, that deforestation has raised ambient surface temperatures and increased the incidence of malaria (Yasuoka and Levins, 2007), augmented river sedimentation and harmed fish habitats (Simonit and Perrings, 2011), and reduced water flow used for irrigation and energy production by hydropower plants (Crafford *et al.*, 2012), among other impacts. The impacts of deforestation have been estimated to cost the Kenyan economy 5.8 billion Kenyan shillings (69 million US dollars) in 2010 (Crafford *et al.*, 2012).

Roughly 80 percent of Kenyan households and businesses still depend on fuelwood as a primary energy source (Republic of Kenya: Ministry of Environment, 2014). The Kenyan government and many non-governmental organizations have promoted private tree cultivation on household lots in an effort to curb further deforestation (see, for example, Kenya Forest Service, 2009; Mathu, 2011). As a result of these longstanding policies and programs, fuelwood in rural Kenya is often collected both from off-farm sources and from private farm woodlots. In many villages, fuelwood is also purchased either from neighbors or in local markets. The labor division in fuelwood sourcing is strict. Similar to women in other SSA countries, women in Kenya are engaged both in "productive" activities, such as fuelwood and water collection, and "reproductive" activities, such as cooking, cleaning, and childcare (Kes and Swaminathan, 2006). Men, on the other hand, are generally engaged only in "productive" activities, both on-farm (growing crops and trees, rearing livestock) and off-farm wage labor. The "double workday" for women often means that women work longer hours than men, which limits their opportunity for participation in the off-farm labor market (Kes and Swaminathan, 2006).

Qualitative studies from the early 1990s show strong cultural taboos against women participating in on-farm tree management (Chavangi and Adoyo, 1993; Kiptot and Franzel, 2012; Mugure and Oino, 2013). In Kakamega County in western Kenya, for example, the belief exists that "if a woman plants a tree, she will become barren" (Chavangi and Adoyo, 1993, pg. 66). This differs from practices in South Asia: while Amacher *et al.* (1993) and Heltberg *et al.* (2000) find that women and children are the primarily collectors of fuelwood off-farm and men are the primary collectors on-farm, Kohlin and Amacher (2005) and St. Clair (2016) show significant contributions of men to off-farm fuelwood collection in India and Nepal, respectively. In data collected for this study, 94 percent of primary fuelwood collectors off-farm are women and 67 percent of on-farm woodlots are managed by men. We reflect these gender differences in fuelwood collection in our theoretical model below and subsequently empirically test whether male or female shadow wages are correlated with a particular method of fuelwood acquisition.

The household data used in our analysis were collected in 2011-2012 in 15 villages in Kakamega, Kericho, Kisumu, Siaya, Uasin Gishu, and Vihiga counties of Kenya (Berazneva *et al.*, 2017).[4] The full survey included 21 randomly sampled households in each village and covered a wide range of Living Standards Measurement Survey (LSMS) components. Importantly, the survey included a detailed module on household energy consumption and production from all available sources. Households were asked about their energy use from fuelwood, agricultural residues, charcoal, kerosene, LPG, and electricity during each month of the 2011 calendar year, as well the sourcing of energy from on-farm, off-farm, and market.

The vast majority of households in the sample (98 percent) use fuelwood as a primary cooking energy and most acquire their fuelwood from more than one location. In the research area, land is privately owned. While the majority of households reports collecting fuelwood from neighboring farms or unfarmed area in their communities,

[4]Three villages were randomly selected from each of the five 10-kilometer blocks, originally used in the Western Kenya Integrated Ecosystem Management Project that was implemented by the Kenya Agricultural Research Institute and the World Agroforestry Center in 2005-2010.

paying access fees is not common. A small number of households reports collecting fuelwood from government forest reserves. These forest reserves do have restrictions on fuelwood collection activities, but anecdotal evidence suggests that they are not often enforced. Only a small portion of households (about 15 percent) grows trees on dedicated woodlots that are generally close to the homestead; most households, however, grow trees on their farms - along the farm boundaries, plot edges, and scattered throughout plots. The majority of trees are planted by households (very few are native species left from land clearing). Main species are *Eucalyptus saligna*, *Cupressus lusitanica*, *Markhamia lutea*, *Grevillea robusta*, *Persea Americana*, *Psidium guajava*, *Mangifera indica*, and *Sesbania sesban*, among others, and many are planted with several goals in mind: fuelwood production, erosion control, property boundaries, shade around the homestead, etc.

In our analysis, following the methodology of Acharya and Barbier (2002) and Palmer and MacGregor (2009), if a household uses fuelwood from different sources, we consider the household to be present in each of the three source groups. For example, a household that obtains fuelwood both from off-farm and on-farm sources is considered to be both in the off-farm fuelwood collection group and the on-farm production fuelwood group. As a consequence, as in table 2.2, the total number of observations in the three fuelwood groups added together is greater than the total number of households in the sample.

Several differences among source groups are immediately apparent (table 2.2). Fuelwood buyers, for example, on average have larger households, higher annual incomes (though not per capita incomes), more education, and less land area than the other two groups. All of these differences are to be expected. Households with greater incomes can more readily afford to buy fuelwood, and smaller land areas mean less room for on-farm woodlots. Off-farm collectors, on the other hand, have the lowest mean income of the three groups, have younger household heads, a lower number of trees on-farm, smaller land parcels, and have a lower asset index.[5] In addition, lower incomes among fuelwood off-farm collectors limit fuelwood purchases and fewer on-farm trees imply lower fuelwood production from private woodlots. Finally, on-farm fuelwood producers have larger landholdings, a greater absolute number of trees, a larger herd size, lower wages for men and women, and a smaller share of income earned off the farm. Larger landholdings suggest lower opportunity costs for on-farm woodlots, all else being equal, as more land is available for tree cultivation. A larger herd size and smaller share of income earned off-farm suggest that on-farm fuelwood producers expend more labor hours working on-farm. This may lead to lower opportunity costs of on-farm production as farmers may be able to practice tree management concurrent with other on-farm activities.

2.2 Theoretical household model

In rural Kenya, as elsewhere in SSA, a typical household consumes much of its own production. As a result, and given likely imperfections in markets for both labor and goods, market wages may not reflect household opportunity costs when it comes to off-farm collection and on-farm fuelwood production (Skoufias, 1994; Amacher *et al.*, 1996). While hired labor is used in this part of Kenya for agricultural activities, no households in our sample hired labor for fuelwood acquisition. In a constrained labor market, labor allocated to private energy collection is thus subject to an unobserved shadow wage that forms the basis for households' production decisions (Strauss, 1986). In our dataset, 63 percent of households consuming fuelwood purchased none from the market.

[5]Following Sahn and Stifel (2003), an asset index is derived from a factor analysis on household durables and housing quality (table 2.A.1 in the Appendix).

The opportunity cost, as measured by the shadow price or shadow cost, of fuelwood for these households can therefore be substantially different from the market price. Strauss (1986), Jacoby (1993), and Skoufias (1994) were among the first to develop the concept of shadow wages and shadow prices in a general agricultural context, and Amacher *et al.* (1996) were first to apply it specifically to fuelwood. Heltberg *et al.* (2000) and Palmer and MacGregor (2009) extended the non-separable agricultural household model to focus on traditional energy substitutes and, in the case of Palmer and MacGregor (2009), the substitution between fuelwood collected and purchased. We build on their model and include three different fuelwood sources (fuelwood collected off-farm, fuelwood produced on-farm, and that purchased), as well as account for the substitution of fuelwood with traditional fuels (e.g., agricultural residues) and other alternatives (e.g., kerosene).

More formally, let a representative agricultural household maximize a monotonic, continuous, quasi-concave utility function U:

$$\underset{C_E, C_X, C_L^M, C_L^F}{MaxU} = U(C_E, C_X, C_L^M, C_L^F; \boldsymbol{z}^h), \tag{1}$$

where C_E stands for consumed goods requiring energy inputs, C_X represents all other consumed goods, C_L^M is leisure consumed by men in the household, C_L^F is leisure consumed by women in the household, and $\boldsymbol{z}^h$ is a vector of household characteristics that affect consumption.

Household goods C_E are produced according to function θ using a mixture of energy types and technology:

$$C_E = \theta(C_{FW}, C_B, C_A; S). \tag{2}$$

Here, C_{FW} represents fuelwood consumed, which can be from fuelwood collected off-farm, produced on-farm, or purchased. C_B stands for other traditional biomass fuels such as crop or animal residues usually produced on farm, C_A represents the consumption of more advanced fuels such as kerosene, and S represents stove technology.

Based on our data from Kenya, in our model we assume women are the primary collectors of fuelwood off-farm and men are the primary producers of fuelwood on-farm. We also assume that male and female labor is not perfectly substitutable. Therefore, the consumption of leisure in the model is divided between women and men, $C_L^{F,M}$, and is given by:

$$C_L^{F,M} = L^{F,M} - l_{AG}^{F,M} - l_{off}^{F,M} - l_{FW}^{F,M}, \tag{3}$$

where L is the total endowment of labor, l_{AG} is labor devoted to agricultural activities, l_{off} is off-farm labor, and l_{FW} is labor allocated to fuelwood on-farm production or off-farm collection. Fuelwood production on-farm (P) and fuelwood collection off-farm (C) are given by continuous, quasi-concave functions of household labor:

$$q_{FW}^P = f_{FW}^P(l_{FW}^M; \boldsymbol{z}_{FW}^P), \tag{4}$$

$$q_{FW}^C = f_{FW}^C(l_{FW}^F; \boldsymbol{z}_{FW}^C), \tag{5}$$

where q_{FW} is the quantity of fuelwood produced on-farm or collected off-farm and $\boldsymbol{z}_{FW}$ includes other household characteristics.

For simplicity[6] we assume that all fuelwood collected off-farm, produced on-farm, or purchased by the household is consumed such that

$$q_{FW}^B = C_{FW} - q_{FW}^P - q_{FW}^C \geq 0, \tag{6}$$

where q_{FW}^B is the quantity of fuelwood bought by a household. Net consumption is positive for buyers and equal to zero for non-buyers. The agricultural production, q_{AG}, is a function of male and female labor (l_{AG}), agricultural residues used for soil fertility management and animal feed (q_B), agricultural inputs such as land (a_{AG}),

[6]Only twelve households in the sample (four percent) sell fuelwood.

and other household endowments ($\boldsymbol{z}_{AG}$), as follows:

$$q_{AG} = f_{AG}(l_{AG}^{M}, l_{AG}^{F}, q_B, a_{AG}; \boldsymbol{z}^{AG}), \tag{7}$$

$$q_B = \alpha q_{AG} - a_{AG}, \tag{8}$$

where α is the proportion of the agricultural production that results in residues, so that q_B is the amount of residues left after use for agricultural production.

The household budget constraint is given by Equation 9:

$$P_X C_X + P_{FW}(C_{FW} - q_{FW}^{P} - q_{FW}^{C}) + P_A C_A = P_{AG} q_{AG} + w_M l_{off}^{M} + w_F l_{off}^{F} + V, \tag{9}$$

where P_X, P_{FW}, P_A, P_{AG} are the prices of the respective goods, $w_{M,F}$ are wage rates for men and women, and V represents other household income such as remittances.

Assuming an interior solution and substituting equations 2 and 3 into equation 1, we thus have the following Lagrangian:

$$\begin{aligned}\mathcal{L} = {} & U[\theta(C_{FW}, C_B, C_A; S), C_X, L^M - l_{AG}^{M} - l_{off}^{M} - l_{FW}^{M}, L^F - l_{AG}^{F} - l_{off}^{F} - l_{FW}^{F}; \boldsymbol{z}^h] \\ & - \lambda[P_X C_X + P_{FW}(C_{FW} - q_{FW}^{P} - q_{FW}^{C}) + P_A C_A - P_{AG} q_{AG} - w^M l_{off}^{M} - w^F l_{off}^{F} - V] \\ & - \mu_{AG}[q_{AG} - f_{AG}(l_{AG}^{M}, l_{AG}^{F}, \alpha q_{AG} - q_B; \boldsymbol{z}_{AG})] - \mu_{FW}^{P}[q_{FW}^{P} - f_{FW}^{P}(l_{FW}^{M}; \boldsymbol{z}_{FW}^{P})] \\ & - \mu_{FW}^{C}[q_{FW}^{C} - f_{FW}^{C}(l_{FW}^{F}; \boldsymbol{z}_{FW}^{C})] + \eta[C_{FW} - q_{FW}^{P} - q_{FW}^{C})].\end{aligned} \tag{10}$$

where λ, μ, and η are the multipliers on the budget, production, and consumption constraints. We also assume that the shadow prices of fuelwood and agricultural production (e.g., yields) are positive ($\mu_{AG}, \mu_{FW} > 0$).

Selected first-order conditions for utility maximization are given as:

$$\frac{\partial \mathcal{L}}{\partial C_{FW}} = \frac{\partial U}{\partial \theta}\frac{\partial \theta}{\partial C_{FW}} - \lambda P_{FW} + \eta = 0, \tag{11}$$

$$\frac{\partial \mathcal{L}}{\partial C_X} = \frac{\partial U}{\partial C_X} - \lambda P_X = 0, \tag{12}$$

$$\frac{\partial \mathcal{L}}{\partial q_{AG}} = \lambda P_{AG} + \mu_{AG}\left[\alpha\frac{\partial f_{AG}}{\partial q_{AG}} - 1\right] = 0, \tag{13}$$

$$\frac{\partial \mathcal{L}}{\partial q^P_{FW}} = \lambda P_{FW} - \mu^P_{FW} - \eta = 0, \tag{14}$$

$$\frac{\partial \mathcal{L}}{\partial q^C_{FW}} = \lambda P_{FW} - \mu^C_{FW} - \eta = 0, \tag{15}$$

$$\frac{\partial \mathcal{L}}{\partial l^{F,M}_{AG}} = \mu_{AG}\frac{\partial f_{AG}}{\partial l^{F,M}_{AG}} - \frac{\partial U}{\partial C^{F,M}_L} = 0, \tag{16}$$

$$\frac{\partial \mathcal{L}}{\partial l^{F,M}_{off}} = \lambda w^{F,M} - \frac{\partial U}{\partial {C_L}^{F,M}} = 0, \tag{17}$$

$$\frac{\partial \mathcal{L}}{\partial l^F_{FW}} = \mu^C_{FW}\frac{\partial f^C_{FW}}{\partial l^F_{FW}} - \frac{\partial U}{\partial C^F_L} = 0, \tag{18}$$

$$\frac{\partial \mathcal{L}}{\partial l^M_{FW}} = \mu^P_{FW}\frac{\partial f^P_{FW}}{\partial l^M_{FW}} - \frac{\partial U}{\partial {C_L}^M} = 0, \tag{19}$$

$$C_{FW} - q^P_{FW} - q^C_{FW} \geq 0. \tag{20}$$

Rearranging the first-order conditions produces a number of important relationships. Equations 11, 14, and 15, for example, suggest that the marginal utility of fuelwood consumption is equal to the shadow price of fuelwood:

$$\frac{\partial U}{\partial \theta}\frac{\partial \theta}{\partial C_{FW}} = \lambda\left(P_{FW} - \frac{\eta}{\lambda}\right) = \mu^P_{FW} = \mu^C_{FW}, \tag{21}$$

where μ^P_{FW} and μ^C_{FW} are, respectively, the shadow prices of on-farm fuelwood production and off-farm fuelwood collection, which in equilibrium are equal. In practice, however, these shadow prices can differ due to household preferences, lack of substitutability of labor between male and female household members, and environmental factors, among other reasons.

Rearranging equations 16 through 19 we have:

$$\frac{\partial U}{\partial {C_L}^F} = \mu^C_{FW} \frac{\partial f^C_{FW}}{\partial l^F_{FW}} = \mu_{AG} \frac{\partial f_{AG}}{\partial l^F_{AG}} = \lambda w^F, \tag{22}$$

$$\frac{\partial U}{\partial {C_L}^M} = \mu^P_{FW} \frac{\partial f^P_{FW}}{\partial l^M_{FW}} = \mu_{AG} \frac{\partial f_{AG}}{\partial l^M_{AG}} = \lambda w^M. \tag{23}$$

Equations 22 and 23 demonstrate that the marginal utility of leisure is equal to the marginal value product of labor in fuelwood production on-farm/off-farm collection and the marginal value product of labor in agriculture, which also depends on the household specific wage rate for men and women. The non-separability of households' on-farm and off-farm fuelwood consumption decisions thus implies that household labor activities are subject to household-specific unobserved opportunity costs or shadow prices. In particular, the household consumption of fuelwood depends on the household-specific shadow price of fuelwood (for non-buyers), which is further divided into the shadow prices of on-farm production, μ^P_{FW}, and off-farm collection, μ^C_{FW}.

From the first-order conditions we also obtain the following reduced-form equations for the quantity of fuelwood produced on-farm, q^P_{FW}, collected off-farm, q^C_{FW}, and purchased, q^B_{FW}:

$$\left.\begin{array}{l} q^P_{FW} \\ \\ q^C_{FW} \\ \\ q^B_{FW} \end{array}\right\} = f(P_{FW}, P_X, P_{AG}, P_A, w^{F,M}, \boldsymbol{z}^h, \boldsymbol{z}^{P,C}_{FW}, L, S, V). \tag{24}$$

It is not clear whether the price of fuelwood, P_{FW}, must be the market price in the case of fuelwood collected off-farm or produced on-farm in a labor constrained market. More likely, the price is endogenous and a function of shadow prices. The wage rate,

$w^{F,M}$, is also not exogenous but a function of implicit household wage rates.

2.3 Estimation strategy

In order to estimate the demand equations (Equation 24), we first need to estimate shadow prices of fuelwood from different sources that take into account the opportunity costs of production on-farm and collection off-farm. The *shadow price of off-farm fuelwood collection*, for example, captures the time spent collecting fuelwood off-farm as well as the opportunity cost of labor, so that the increased time it takes to collect each kilogram of fuelwood, or the higher the opportunity cost of the labor, the more expensive each unit of fuelwood becomes (Mekonnen, 1999). The variable used to represent the shadow price of collecting fuelwood has varied in the literature. Cooke (1998a,b), Mekonnen (1999), and Baland *et al.* (2010), for example, use the opportunity cost of labor (shadow wage) multiplied by the time spent collecting each unit of fuelwood. The data on amounts of fuelwood collected, however, are often difficult to obtain, so Amacher *et al.* (1993), Heltberg *et al.* (2000), and Palmer and MacGregor (2009), for example, use the time spent collecting fuelwood as a proxy. This variable, however, does not capture the value of time and often leads to underestimates of the elasticity of demand for fuelwood (see table 2.1).

Following Cooke (1998a,b), Mekonnen (1999), and Baland *et al.* (2010), we define the average shadow price of fuelwood collected from off-farm for household i in the collecting group C as:

$$\mu_i^C = \left(\frac{H_i^C}{q_i^C}\right)\omega_i^F, \tag{25}$$

where H is the monthly number of hours spent collecting fuelwood, q is the monthly amount of fuelwood collected, and ω is the household-specific opportunity cost of female labor, i.e., female shadow wage.[7] Since the shadow wage is given in Kenyan Shillings per hour, the unit value of the shadow price is Kenyan shilling per kilogram (KES/Kg).

Similarly, we define the *shadow price of fuelwood produced on-farm*. In western Kenya, as elsewhere in SSA, on-farm fuelwood production is often a by-product or co-product of growing trees for timber and other purposes (Buck *et al.*, 1999). Since producing fuelwood on-farm does not necessarily require felling trees, data on the time spent producing fuelwood on-farm only are not available. We approximate the number of hours spent producing fuelwood on-farm by the number of on-farm trees and the time necessary to cultivate and manage an individual tree:

$$\mu_i^P = \left(\frac{\gamma^P T_i^P}{q_i^P}\right)\omega_i^M, \tag{26}$$

where γ is the average number of hours needed to grow one tree, T is the number of on-farm trees, q is the amount of fuelwood per month produced by household i in the on-farm producing group P, and ω is the household-specific opportunity cost of male labor – the male shadow wage. The value for γ comes from the Kenya Forestry Research Institute (KEFRI) estimates for growing Eucalyptus trees (Oballa *et al.*, 2010). Eucalyptus is very common in the research area and is a primary choice for agroforestry in Kenya – being among the most popular tree species planted on household farms in this area (Scherr, 1995; Henry *et al.*, 2009).[8] In the survey, households that purchased fuelwood were asked their individual price paid for a particular quantity. Because this market price may be correlated with unobserved household characteristics, we follow a 2SLS strategy as outlined below.

[7] In this section, the subscript FW for fuelwood is dropped.

[8] Using KEFRI's data, γ is approximately equal to 1.6 hours of work per tree over its life. 1.6 hours is equivalent to 0.01 man months, assuming eight hours in a work day and twenty days in a work month.

Estimating both Equations 25 and 26 also requires shadow wages for men and women. Since not all households in the sample engage in off-farm wage labor, we estimate the household-specific shadow wages for men and women following the methodology to account for self-selection proposed by Heckman (1979) and Olsen (1980) and used by Cooke (1998a,b) in a similar setting.[9] We estimate these shadow wages using maximum likelihood for greater efficiency, especially important due to the sample size restrictions in the data.[10] Results for these estimations are in table 2.A.2 in the Appendix.

There are several additional aspects to our estimation of demand elasticities. First, in order to estimate cross-price elasticities, it is necessary to proxy shadow prices and market prices for households that do not participate in all groups. For example, households that only collect fuelwood off-farm do not have estimates for shadow prices for fuelwood produced on-farm or prices for fuelwood bought at the market. To create these full sample variables, we follow the strategy suggested by Mekonnen (1999) and use the village-specific maximum values for hours spent collecting, the number of on-farm trees and market prices, when household-specific values are absent. The household's decision to participate (or not) in each of the fuelwood groups (producing on-farm, collecting off-farm, and purchasing) must reflect the household-specific cost of participation. So, if the household is not observed in a particular group, it is likely the case that its cost of participation is greater than the cost of any other participating household in the village.

[9] Overall, 37 percent of men and 19 percent of women in the sample are engaged in wage labor. See also Binder and Scrogin (1999), Levison *et al.* (2008), and DeGraff and Levison (2009) for examples of this methodology in the development literature.

[10] Our exclusion restrictions necessary to control for selection bias include the dependency ratio and the distance from the village center. The dependency ratio here is defined as number of children under 15 and elderly over 65 divided by number of adults between 15 and 65. Individuals are more likely to enter the off-farm workforce if they live closer to the village center, but the majority of households (74 percent in this sample) live on inherited land so that their farm location is thus exogenously determined.

Second, we are concerned that households may self-select into their respective fuelwood source group(s), which may bias our results. To account for this selection bias, we add a variable to the second stage estimation of the demand elasticities following Olsen (1980) and Wooldridge (2002).[11] Third, we take precautions against endogeneity due to the likelihood of simultaneity and omitted variable bias arising from the shadow price variable and report results from the two-stage least squares (2SLS) estimation. We borrow an instrumental variable estimation strategy from the demand literature. To instrument for the market price, we use the average of all (except own) market prices in the sample. As in Hausman *et al.* (1994), the key assumption is that random household-level factors influencing the market price are independent of other households.

As for the likely endogenous shadow prices for on-farm producers and off-farm collectors, we match each household with five other households in the sample outside their own block (households in the same village and two neighboring villages) based on the most similar shadow price. For off-farm collectors, we use the average of the hours spent collecting of the matched households as an instrument on the shadow price of fuelwood collecting of the first household, while for on-farm producers, we use the average number of trees owned as the instrument on the shadow price of on-farm production of the first household.[12] By excluding households in the own-block in the matching, we prevent possible validity issues arising from geographic proximity

[11]table 2.A.3 in the Appendix presents the first-stage results of the jointly estimated linear probability models that show the likelihood of participating in a particular fuelwood source group. Exclusion variables for these estimations are number of parcels for off-farm collecting households, land slope for on-farm producing households, and distance to the nearest town for purchasing households.

[12]For example, if household 1 has a shadow price of off-farm collecting fuelwood of x, we find five other households outside the same geographic area (block) that have the most similar shadow price of off-farm collecting to x. We then take the average of hours spent collecting of these five matched households, and use that as an instrument on the first household. By design, the instrument is relevant, and because these matched households are far away geographically, the instrument is valid.

of some villages.[13] First-stage IV results are in table 2.A.4 in the Appendix.

In order to allow for inter-household comparisons, continuous variables (e.g. shadow prices, shadow wages, land area, etc.) are scaled by adult equivalent units, which accounts for differing numbers and ages of participants in each household (Cavendish, 2002).[14] Our results can then be interpreted as per-capita monthly[15] values. We estimate the household-specific demand equations for fuelwood collected off-farm, q_i^C, produced on-farm, q_i^P, and bought in the market, q_i^B:

$$q_i^j = \beta_0 + \beta_1\mu_i^j + \beta_2\nu_i^k + \beta_3\varphi_i + \beta_4\omega_i^M + \beta_5\omega_i^F + \beta_6\boldsymbol{X}_i + \zeta + \varepsilon_i^j, \tag{27}$$

$$q_i^B = \beta_0 + \beta_1 P_i^B + \beta_2\nu_i^C + \beta_3\psi_i^P + \beta_4\omega_i^M + \beta_5\omega_i^F + \beta_6\boldsymbol{X}_i + \zeta + \varepsilon_i^B. \tag{28}$$

In Equation 27, superscripts j, k represent either collecting off-farm or producing on-farm ($j \neq k$), μ_i is the shadow price of either collecting off-farm or producing on-farm, ν and φ are the shadow prices and market prices with full observations, ω^M and ω^F are the average shadow wages for men and women, $\boldsymbol{X}_i$ are household variables that influence fuelwood use, ζ are geographic fixed effects (at the block level), and ε is the error term. In equation 28, P_i is the market price paid by a household that purchases fuelwood and ν^C and ψ^P are shadow prices for fuelwood collected off-farm and produced on-farm with full observations, respectively. Given possible automated regressor bias, we bootstrap all standard errors.

[13]Because households may know one another between villages in proximity (within the research block), we exclude all households from within a household's own block to ensure that households have no connection with one another outside the similarity in their shadow price values

[14]Unlike Cavendish (2002), we do not have data on amount of time during the year that each individual lived at home, so cannot account for this in our adult equivalent unit adjustment. We do however account for number of individuals, their ages, and gender. We thank an anonymous reviewer for suggesting this adjustment.

[15]Monthly fuelwood values are annual quantities of fuelwood divided by twelve. This provides an average monthly value that avoids problems of seasonality.

2.4 Results

We first estimate household-specific shadow prices of collecting fuelwood off-farm and producing fuelwood on-farm, following Equations 25 and 26, and using household-specific wages for men and women. Our results, reported in table 2.3, suggest imperfections in fuelwood and labor markets in western Kenya. The median off-farm collection shadow price (1.53 KES/kg) is below the median on-farm production shadow price (5.21 KES/kg), which is in turn below the median market price (5.85 KES/kg). This ordering is consistent with the traditional agricultural household model. When shadow prices for a particular fuelwood source approach and exceed the market price, the household switches to purchasing from the market (given that one is available) (Key *et al.*, 2000). On-farm production shadow prices above that of off-farm collection shadow prices can have several explanations, including higher time requirements for managing woodlots on-farm compared to collecting from off-farm, as well as the opportunity cost of planting woodlots instead of food crops.

We then estimate demand equations (Equations 27 and 28) for different fuelwood sources, including the respective shadow prices, fuelwood market price, and wages as right-hand side variables. Table 2.4 shows the regression results from ordinary least squares estimation, as well as the results from two-stage least squares estimations that control for the endogeneity of shadow prices and potential selection bias. The results across the three estimations (OLS, 2SLS, and 2SLS+Olsen Estimator) are quite similar. Although the tests of endogeneity and selection bias suggest exogeneity in both off-farm collection and on-farm fuelwood production regressions, for the sake of caution, we use coefficient values from the 2SLS-Olsen regressions when interpreting our results (third column for each fuelwood source group). The coefficients on shadow prices in table 2.4 can be interpreted as elasticities as all variables are used in log form.

As expected, own-price elasticities are negative, inelastic, and statistically significant at a p-value of five percent or less across all groups and specifications. Moreover, they are very similar across non-purchased fuelwood sources: own-price elasticities range from -0.48 to -0.61 for off-farm fuelwood collectors and -0.50 to -0.55 for on-farm fuelwood producers for all specifications. The inelastic own-price elasticity means that increases in fuelwood costs lead to less than equi-proportionate decreases in the amount of fuelwood obtained from that source, suggesting that fuelwood is a necessity good for the households in our sample. Our own-price elasticity values are somewhat higher in magnitude than elasticities found in other studies (table 2.1). Geography can play a large role in the elasticity differences in fuelwood demand. Amacher *et al.* (1996) and Amacher *et al.* (1999), for example, find large differences in own-price elasticities between hill and plain-dwelling populations in Nepal. Most of the existing studies on this subject are primarily from South Asia, and none are from household samples in East Africa or Kenya specifically.

Cross-price elasticities are also inelastic, but positive. The very low cross-price elasticities suggest that substitution between fuelwood sources is low. Although fuelwood is often considered to be a homogeneous product, households in western Kenya do not readily substitute between fuelwood sources. For example, an increase of ten percent in the shadow price of on-farm production of fuelwood (corresponding to an increase of 1.40 KES/kg), leads to a decrease in fuelwood produced on-farm of 10.13 kg per month, evaluated at the mean. Moreover, substitution to a different source is low: an increase of ten percent in the shadow price of on-farm production leads to the increase of fuelwood both bought and collected off-farm by only 5.19 kg per month, evaluated at the mean. This shadow price increase, therefore, leads to a net decrease of 4.19 kg per month in fuelwood consumed, holding other fuelwood costs constant. This low quantity is surprising, given the mean consumption of about 210 kg per month, and illustrates the lack of substitutability between fuelwood groups.

This lack of substitution between fuelwood sources demonstrated by our results can be explained in part by the gender division in household labor. We find that female shadow wages in the demand equations are statistically significant with respect to fuelwood collected and male shadow wages are likewise statistically significant with respect to fuelwood produced on-farm (table 2.4). The coefficients in both cases are positive, similar to findings in other studies (Amacher *et al.*, 1996; Cooke, 1998a; Amacher *et al.*, 1999). Higher female wages increase the amount of fuelwood collected off-farm, demonstrating the impact of female labor opportunity on fuelwood collection. Meanwhile, male wages have no significant effect on fuelwood quantity collected off-farm, consistent with our data showing men are not actively engaged in off-farm fuelwood collection in this area. The opposite relationship is found with respect to fuelwood produced on-farm, as men's work opportunity is correlated with fuelwood produced on-farm through their shadow wage, and the time women work has no significant effect on fuelwood produced on-farm. This again is consistent with our data and qualitative evidence from the area that women are less engaged in on-farm fuelwood production than men.

Our results also show that changes in the shadow price of off-farm collecting lead to significant increases in the amount of female labor spent gathering fuelwood off-farm. Using the same 2SLS regression as above (Equation 27) but with the number of hours spent collecting off-farm per month as the dependent variable, we find that a ten percent increase in the shadow price of off-farm collecting increases the hours spent collecting off-farm by 3.6 percent (table 2.A.5 in the Appendix. See table 2.1 for comparisons with other papers). The magnitude of the elasticity of labor for fuelwood collection is greater than the magnitudes of the cross-price elasticities of off-farm collecting with respect to either on-farm producing or buying fuelwood. This illustrates that households prefer to increase labor devoted to the off-farm collecting of fuelwood rather than substitute toward on-farm producing or purchasing fuelwood in the wake of shadow price increases.

Coefficients for additional estimation regressors, including household characteristics and charcoal and kerosene prices, are in Appendix table 2.A.6. We find few variables to be statistically significant. The coefficient on age of household head is one exception, which is positive and statistically significant with respect to fuelwood produced on-farm, and negative and significant for fuelwood collected off-farm. In addition, while substitution with alternative energy sources is not a focus of this study (due to data constraints), we include the price of charcoal and kerosene in the regressions. The price of charcoal lacks statistical significance for any fuelwood source group, while the price of kerosene is negatively correlated and marginally statistically significant with fuelwood produced on-farm, suggesting a complementary relationship. Due to the relatively small number of individuals using charcoal and kerosene in the sample, we are cautious to give weight to these findings.

2.5 Conclusion and policy implications

This paper examines households' energy use in the western Kenya highlands, focusing specifically on the substitution among different fuelwood sources – fuelwood produced on-farm, collected off-farm, and purchased – and the role of households' labor endowments in energy sourcing. We find that the median household shadow price of fuelwood collected off-farm (1.5 Kenyan shillings per kilogram (KES/kg)) is well below the median shadow price of fuelwood produced on-farm (5.2 KES/kg) and the median market price of purchased fuelwood (5.8 KES/kg). The most plausible explanations for this result, suggested by the patterns in our data and estimation, are the potential lack of off-farm employment opportunities for women (that depress the shadow price of fuelwood collected off-farm) and possible competition of agroforestry with on-farm crops, among other factors (that increase the shadow price of fuelwood produced on-farm). In our sample of households, women earn less than men: 3,000 KES ($36) per month compared to 4,000 KES ($47), 95 percent of fuelwood collectors are women, and most woodlots are managed by men (male household heads or male children or grandchildren in households headed by women). In line with the data, coefficients on the female shadow wages in the demand equations are statistically significant with respect to fuelwood collected off-farm and coefficients on the male shadow wages are likewise statistically significant with respect to fuelwood produced on-farm. These results echo earlier findings from qualitative studies in western Kenya that show strong social and cultural norms behind household division of labor (Chavangi and Adoyo, 1993).

Looking specifically at fuelwood collected off-farm, we also find that the own-price elasticity for non-purchased fuelwood is greater in absolute value terms than the labor supply elasticity. Households prefer to increase the labor dedicated to exploiting a fuelwood source with an increasing shadow price rather than substitute away from it. Given higher shadow prices for fuelwood collected off-farm, rural Kenyan households respond by increasing female labor, rather than substituting toward other fuelwood sources. This finding has important implications for efforts against forest degradation. It shows, for example, as fuelwood scarcity increases, households will expend more time and effort to collect fuelwood off-farm rather than switch to fuelwood production on-farm or fuelwood purchases, which may come from more renewable sources.

Our findings also imply that fuelwood is not a homogeneous resource. The very limited substitution as shown by our results suggests that increases in the shadow price of fuelwood from any particular source can potentially have significant effects on household labor. Increasing opportunities for female work off-farm increases the relative cost of fuelwood collected off-farm, which could lead to greater relative dependence on on-farm fuelwood production and lower overall use as fewer people are at home during the day (Burke and Dundas, 2015). Planting woodlots and producing fuelwood on-farm are arguably more sustainable practices than collecting off-farm, as the trees planted on-farm in agroforestry systems partially offset GHG emissions from biomass use and have other environmental and agricultural benefits; these include decreases in air and soil temperatures from greater tree coverage in the landscape, decreased soil erosion, and improved water retention, among others (Unruh *et al.*, 1993; Mbow *et al.*, 2014). In the short run, however, any increases in the shadow price of fuelwood collected off-farm are likely to increase the work burden for women. As Bluffstone (1995) finds, off-farm labor opportunities will stabilize tree coverage by increasing the opportunity cost of labor. However, these labor opportunities must exist for both genders, which implies increasing the labor substitutability between men and women in the labor force.

Another technology, improved cookstoves, also have the potential to decrease fuelwood use in SSA. Improved combustion or pyrolysis stoves are more efficient, can use substitutes for fuelwood (e.g., crop residues, grasses), require fewer units of biomass for cooking than traditional stoves, and can produce valuable soil amendments such as biochar (Torres-Rojas *et al.*, 2011). However, more efficient stoves can lead to a "rebound effect," wherein households continue to use similar quantities of biomass but increase cooking activities (Nepal *et al.*, 2011). Moreover, projects in Kenya seeking to increase the use of improved cookstoves have found adoption rates to be low, given many households' attachments to traditional cooking techniques (Tigabu, 2017). Further research is needed to identify strategies to increase the use of improved cookstoves in Kenya and to mitigate rebound effects.

In conclusion, reforestation efforts in western Kenya that include promotion of on-farm agroforestry may be ineffective in inducing households to collect less fuelwood off-farm unless there are changes to traditional norms regarding female participation in on-farm tree management and in the off-farm labor market. These norms indeed appear to be gradually changing. The new Kenyan constitution, approved by a significant majority of Kenyan citizens in 2010, codifies new rights for women in society (Kramon and Posner, 2011). Over the upcoming years, changing norms may lead to increasing substitution between male and female labor in rural labor markets, with consequent increases in agroforestry, tree coverage and the associated environmental benefits, as this paper suggests.

Table 2.1: Estimates of Fuelwood Elasticities from the Existing Literature

Source	Variable (Per-Unit)	Demand Elasticity (Own-Price)	N	Labor Elasticity (Total Collection Time)	N	Location
Amacher et al. (1993)	Collection Time	-0.157*	89			Nepal
Amacher et al. (1996)†	Market Price	-1.69***/-0.59*	286/240	0.82***/0.97*	286/240	Nepal
Cooke (1998a)	Shadow Cost	-0.25***	101	1.02***	101	Nepal
Mekonnen (1999)	Shadow Cost	-0.40***	419			Ethiopia
Amacher et al. (1999)†	Market Price	-0.21*/-1.47*	286/240			Nepal
Heltberg et al. (2000)	Collection Time	-0.11*	178	0.89*	176	India
Palmer and MacGregor (2009)	Collection Time	-0.05*	172	0.04***	172	Namibia
Baland et al. (2010)	Shadow Cost	-.134*	2190			Nepal

† These papers provide elasticity estimates for two distinct populations and do not provide a combined estimate.
*** $p<0.01$, ** $p<0.05$, * $p<0.1$, - not statistically significant.

Table 2.2: Summary Statistics

Variables	Collectors	Non-Collectors	Producers	Non-Producers	Buyers	Non-Buyers
Household Measures						
Asset Index	-0.23 (0.052)	0.12 (0.071)***	0.0021 (0.062)	-0.17 (0.12)	0.16 (0.096)	-0.13 (0.066)***
Off-Farm Income Ratio	0.63 (0.029)	0.54 (0.027)**	0.55 (0.023)	0.69 (0.046)***	0.58 (0.035)	0.58 (0.025)
Number of Children	5.98 (0.30)	5.99 (0.26)	6.01 (0.22)	5.86 (0.45)	6.40 (0.35)	5.77 (0.24)
Household Size	6.36 (0.19)	5.84 (0.19)*	6.03 (0.16)	6.21 (0.32)	6.75 (0.25)	5.71 (0.17)***
Adult Males	1.75 (0.10)	1.78 (0.10)	1.72 (0.08)	1.95 (0.16)	1.99 (0.13)	1.65 (0.09)**
Adult Females	1.75 (0.09)	1.77 (0.08)	1.79 (0.07)	1.71 (0.13)	2.02 (0.12)	1.65 (0.07)***
Female Children	1.33 (0.10)	1.15 (0.10)	1.24 (0.08)	1.16 (0.16)	1.41 (0.12)	1.13 (0.08)*
Number of Trees	114 (11.2)	144 (9.86)*	140 (8.61)	94.5 (16.8)**	131 (13.2)	131 (9.20)
Household Head Measures						
Age	48.6 (1.37)	54.0 (1.16)***	52.2 (0.94)	49.2 (2.01)	51.3 (1.51)	51.8 (1.09)
Gender (1=Male)	0.82(0.033)	0.79 (0.030)	0.82 (0.025)	0.74 (0.052)	0.83 (0.038)	0.80 (0.028)
Years of Education	6.69 (0.34)	6.54 (0.34)	6.77 (0.28)	5.91 (0.58)	7.23 (0.44)	6.28 (0.31)*
Financials (KES)						
Household Income	104,684 (8,237)	145,662 (11,249)**	123,861 (9,233)	144,446 (19,414)	148,905 (15,600)	116,862 (10,467)*
Per Cap. Household Income	16,662 (1,402)	24,161 (1,780)***	20,451 (1,470)	22,750 (3,083)	22,017 (2,254)	20,315 (1,669)
Share of Family (%)						
Off-farm Emp. (Women)	0.32 (0.041)	0.28 (0.035)	0.28 (0.029)	0.40 (0.060)*	0.30 (0.045)	0.30 (0.033)
Off-farm Emp. (Men)	0.56 (0.044)	0.48 (0.038)	0.48 (0.032)	0.66 (0.065)**	0.57 (0.049)	0.48 (0.036)
Share of Women in Household	0.498 (0.016)	0.52 (0.016)	0.51 (0.014)	0.47 (0.028)	0.51 (0.020)	0.50 (0.015)
Distance Measure (KM)						
Village Center	0.36 (0.022)	0.46 (0.024)***	0.43 (0.022)	0.36 (0.042)	0.39 (0.025)	0.43 (0.023)
Area Measure (Acres)						
Land	2.72 (0.25)	5.93 (0.75)***	5.08 (0.70)	2.27 (1.30)*	2.67 (0.22)	5.50 (0.70)**
Tropical Livestock Units (TLU)						
Herd Size	1.99 (0.21)	2.69 (0.21)**	2.58 (0.18)	1.54 (0.35)***	2.22 (0.23)	2.47 (0.19)
N	131	173	243	61	103	201

Standard deviations located next to respective means. Total sample is 304 observations. Categories (e.g. Collectors vs. Non-Collectors) are not mutually exclusive.
*** $p<0.01$, ** $p<0.05$, * $p<0.1$.

Table 2.3: Price and Quantity Statistics

Variable	N	Median	Mean	Std. Dev.	Min	Max
Price (KES/Kg)						
Collector Shadow Price	131	1.53	2.21	3.01	0.02	30.01
Producer Shadow Price	243	5.21	13.91	28.31	0.06	280.27
Market Price	103	5.85	9.02	10.64	0.59	73.1
Kerosene Market Price (KES/Liter)	234	100.00	105.95	33.65	50.0	250.0
Charcoal Market Price	126	16.67	22.33	18.54	3.33	150.0
Shadow Wages (KES/Month)						
Female (Heckman-adjusted)	301	2040.95	2428.88	1312.67	622.76	7506.82
Male (Heckman-adjusted)	301	1831.93	2073.66	1139.18	196.66	6380.68
Quantities (Kg/Month)						
Fuelwood Used	301	130.00	209.78	316.43	1.85	3000
Fuelwood Collected	131	66.67	96.11	129.03	0.25	800
Fuelwood Produced	243	99.75	184.27	280.24	0.285	2395
Fuelwood Bought	103	48.45	169.12	390.49	2.85	3000

Note: Statistics are for households who participate in the particular fuelwood source group/energy source group.

Table 2.4: Fuelwood Demand Elasticities

VARIABLES	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
	Fuelwood Collected (Kg/Month)			Fuelwood Produced (Kg/Month)			Fuelwood Bought (Kg/Month)		
	OLS	2SLS	2SLS+Olsen	OLS	2SLS	2SLS+Olsen	OLS	2SLS	2SLS+Olsen
Shadow Price Collecting	-0.480***	-0.605**	-0.609**						
	(0.156)	(0.245)	(0.263)						
Shadow Price Producing				-0.500***	-0.546***	-0.551***			
				(0.0544)	(0.132)	(0.137)			
Market Price							-0.863***	-0.920***	-0.925***
							(0.141)	(0.164)	(0.162)
Full Shadow Price Collecting				0.127***	0.126***	0.174**	0.0887**	0.0910**	0.0229
				(0.0220)	(0.0232)	(0.0725)	(0.0407)	(0.0424)	(0.130)
Full Shadow Price Producing	0.165**	0.183***	0.210				0.109	0.112	0.238
	(0.0656)	(0.0708)	(0.179)				(0.0794)	(0.0782)	(0.236)
Full Market Price	0.210	0.216	0.199	0.191**	0.193**	0.208***			
	(0.172)	(0.175)	(0.201)	(0.0778)	(0.0786)	(0.0779)			
Female Wage	0.724**	0.823**	0.811**	-0.105	-0.112	-0.0928	-0.174	-0.154	-0.527
	(0.368)	(0.372)	(0.394)	(0.199)	(0.205)	(0.215)	(0.575)	(0.611)	(0.986)
Male Wage	-0.499*	-0.511*	-0.550	0.788***	0.836***	0.961***	0.169	0.150	0.444
	(0.265)	(0.276)	(0.414)	(0.181)	(0.215)	(0.295)	(0.442)	(0.446)	(0.730)
HH Controls, Block Dummies	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Constant	-2.186	-3.053	-2.603	-0.268	-0.600	-1.702	0.173	0.288	-4.880
	(3.406)	(3.169)	(4.257)	(1.497)	(1.604)	(2.344)	(3.446)	(3.334)	(10.63)
Observations	131	131	131	243	243	243	103	103	103
R-squared	0.421	0.414	0.414	0.485	0.483	0.484	0.475	0.475	0.477
Kleibergen-Paap rk LM P-value		0.00	0.00		0.00	0.00		0.00	0.00
Kleibergen-Paap rk Wald F stat		28.162	29.320		22.102	21.833		28.682	30.617
Stock-Yogo 10% maximal IV size		16.38	16.38		16.38	16.38		16.38	16.38
Hausman Test P-value		0.451	0.448		0.681	0.681		0.428	0.362
Lambda coefficient			-0.323			1.317			-3.664
			(1.916)			(1.865)			(6.775)

HH Controls and additional variables included in Appendix A6. All continuous variables are in log form and in adjusted per capita units after Cavendish (2002). "Full" variables include imputed values for households not participating in respective fuelwood source groups. Instruments for the shadow prices of fuelwood collection and production are the shadow prices of amatched household located outside own-block (additional explanation provided in text). Instrument for market price is the village average of except-own market price. Bootstrapped standard errors (1000 replications) in parentheses. *** p<0.01, ** p<0.05, * p<0.1.

Table 2.A.1: Asset Index

Variable	Weight	Variable	Weight
Durables: number of		**Characteristics: indicator for**	
House	0.411	Brick/cement walls	0.700
Radio	0.389	Mabati (corrugated iron) roof	0.379
Telephone (mobile)	0.649	Cement/wood floor	0.666
Fridge/freezer	0.620	Private piped water	0.447
Television	0.688	Water from neighbor	-0.068
Electronic Equipment	0.559	Borehold water	0.036
Air conditioning	0.339	River/stream water	-0.184
Furniture	0.743	No toilet	-0.279
Kettle/iron	0.446	Traditional toilet	-0.293
Mosquito net	0.602	Improved toilet	-0.703
Computer	0.529	Kerosene light	-0.717
Internet access	0.351	Electricity light	-0.763
Electric/gas stove	0.526	Solar light	-0.112
Improved stove	0.217		
Bicycle	0.332		
Motorcylce	0.483		
Car/truck	0.568		
Bank account	0.699		
Generator	0.291		
Large battery	0.177		
Solar panel	0.338		
LPG	0.636		
Observations	309	Observations	309

Table 2.A.2: Wage Regressions

	Male			Female		
	OLS	MLE Heckman		OLS	MLE Heckman	
VARIABLES	Male Wages (KES/Month)	Male Wages (KES/Month)	Selection	Female Wages (KES/Month)	Female Wages (KES/Month)	Selection
Household Head Measures:						
Years of education	0.0772***	0.0596***	-0.00993	0.106***	0.102***	0.00807
	(0.0150)	(0.0177)	(0.0194)	(0.0234)	(0.0209)	(0.0220)
Age	0.0778***	0.173***	0.213***	-0.0148	0.0253	0.185***
	(0.0239)	(0.0311)	(0.0230)	(0.0225)	(0.0441)	(0.0296)
Age squared	-0.000807***	-0.00195***	-0.00244***	0.000368	-0.000126	-0.00225***
	(0.000296)	(0.000381)	(0.000280)	(0.000277)	(0.000554)	(0.000387)
TLU Herd size	0.0554***	0.0170	-0.0669**	0.0461	0.0414	-0.0178
	(0.0205)	(0.0252)	(0.0262)	(0.0332)	(0.0320)	(0.0319)
Dependency Ratio †			-0.196***			0.0986
			(0.189)			(0.321)
Distance from Village Center			-0.518***			-1.236***
			(0.205)			(0.322)
Village Dummies	√	√	√	√	√	√
Clustered SE at Household Level	√			√		
Constant	6.058***	3.687***	-3.565***	7.032***	5.837***	-4.091***
	(0.480)	(0.667)	(0.526)	(0.570)	(1.159)	(0.667)
Observations	202	536	536	105	540	540
R-squared	0.432			0.436		
LR Test P-Value			0.0030***			0.2091

Robust Standard errors in parentheses. †Dependency Ratio=number of household members less than 15 and over 65 years old divided by the number of household members between 15 and 65 years old. *** p<0.01, ** p<0.05, * p<0.1. All dependent variables are in log form.

Table 2.A.3: First Stage Joint Linear Regressions for Shadow Price Estimation

	(1)	(2)	(3)
VARIABLES	Collect Wood	Produce Wood	Buy Wood
Full Shadow Price Collecting		0.038***	-0.018**
		(0.007)	(0.007)
Full Shadow Price Producing	0.087***		0.033**
	(0.015)		(0.014)
Full Market Price	-0.044	0.011	
	(0.027)	(0.022)	
Female Wage	-0.014	0.012	-0.079
	(0.077)	(0.063)	(0.069)
Male Wage	-1.74***	0.090*	0.045
	(0.063)	(0.049)	(0.055)
Household Head Measures			
Age	-0.005**	0.004	-0.002
	(0.002)	(0.002)	(0.002)
Male	0.040	0.185***	0.011
	(0.080)	(0.064)	(0.071)
Land Area (Acres)	-0.024**	0.003	-0.010
	(0.010)	(0.008)	(0.009)
Asset Index	-0.330***	-0.119	0.236**
	(0.126)	(0.101)	(0.110)
TLU Herd Size	0.037	0.006	0.017
	(0.046)	(0.038)	(0.044)
Adult Men	0.069**	-0.039	0.047
	(0.030)	(0.024)	(0.026)
Adult Women	-0.010	0.040	0.060**
	(0.033)	(0.027)	(0.029)
Children Number	0.001	-0.008	0.017*
	(0.010)	(0.008)	(0.009)
Female Children Number	-0.026	0.021	0.041*
	(0.026)	(0.021)	(0.023)
Charcoal Price	-0.043	-0.013	0.014
	(0.058)	(0.047)	(0.054)
Kerosene Price	-0.200**	-0.104	0.213
	(0.090)	(0.074)	(0.079)
Number of Parcels	0.114***		
	(0.041)		
Slope		0.042*	
		(0.022)	
Distance to Town (Km)			-0.043**
			(0.017)
Block Dummies	Yes	Yes	Yes
Constant	2.435***	0.272	-0.157
	(0.511)	(0.426)	(0.501)
Observations	301	301	301

All shadow prices, market prices, and wages are in log form. Continuous variables in per-capita form after Cavendish (2002). “Full” variables include imputed values for households not participating in respective fuelwood source groups. Regressions are jointly estimated. *** p<0.01, ** p<0.05, * p<0.1.

Table 2.A.4: First Stage IV Regressions for Demand Estimation

VARIABLES	(1) Shadow Price of Collecting	(2) Shadow Price of Producing	(3) Market Price
Hours Collect (IV)	0.252***		
	(0.0553)		
Number of Trees (IV)		0.0225***	
		(0.00552)	
Except-own Mean Market Price (IV)			-57.64***
			(14.77)
Shadow price Collecting (Full)		-0.0277	-0.000
		(0.0278)	(0.0196)
Shadow Price Producing (Full)	0.157***		0.0434
	(0.0597)		(0.0362)
Market Price (Full)	0.0118	0.0206	
	(0.137)	(0.0959)	
Female Wage	0.413	-0.386*	0.227
	(0.261)	(0.225)	(0.231)
Male Wage	-0.115	0.839***	-0.159
	(0.219)	(0.219)	(0.156)
HH Controls, Block dummies	√	√	√
Constant	-3.950	-2.476	127.1***
	(2.565)	(1.756)	(32.34)
R-squared	0.533	0.463	0.742
Observations	131	243	103

All variables in log form. Continuous variables in per-capita form after Cavendish (2002). "Full" variables include imputed values for households not participating in respective fuelwood source groups. All regressions bootstrapped with 1000 repetitions. Standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1.

Table 2.A.5: Elasticity of Labor for Fuelwood Collection

VARIABLES	(1) OLS	(2) 2SLS	(3) 2SLS+Olsen
Shadow Price Collecting	0.417***	0.368*	0.363*
	(0.112)	(0.200)	(0.220)
Full Shadow Price Producing	0.0721	0.0789	0.114
	(0.0523)	(0.0524)	(0.138)
Full Market Price	0.171	0.173	0.152
	(0.124)	(0.131)	(0.151)
Female Wage	-0.283	-0.244	-0.259
	(0.269)	(0.315)	(0.318)
Male Wage	-0.133	-0.138	-0.188
	(0.190)	(0.200)	(0.297)
Household head age	-0.0231***	-0.0243**	-0.0266**
	(0.00854)	(0.00952)	(0.0135)
Household head sex (1=male)	0.330	0.320	0.344
	(0.291)	(0.301)	(0.311)
Land area (acres)	0.0817	0.0771	0.0706
	(0.135)	(0.142)	(0.152)
Asset index	-0.436	-0.528	-0.685
	(0.866)	(0.954)	(1.158)
TLU herd size	0.175	0.189	0.213
	(0.188)	(0.195)	(0.231)
Number adult males	-0.183	-0.183	-0.208
	(0.123)	(0.124)	(0.149)
Number adult females	0.0653	0.0721	0.0700
	(0.137)	(0.143)	(0.145)
Number of children	0.0151	0.0197	0.0212
	(0.0340)	(0.0365)	(0.0399)
Number of female children	-0.0369	-0.0359	-0.0440
	(0.0900)	(0.0942)	(0.0969)
Charcoal price	-0.119	-0.117	-0.132
	(0.213)	(0.229)	(0.231)
Kerosene price	0.303	0.319	0.233
	(0.354)	(0.341)	(0.470)
Lambda coefficient			-0.414
			(1.359)
Constant	2.703	2.366	2.942
	(2.506)	(2.465)	(3.007)
Observations	131	131	131
Kleibergen-Paap rk LM P-value		0.0000	0.0000
Kleibergen-Paap rk Wald F stat		35.657	31.271
Stock-Yogo 10Hausman Test P-value		0.801	0.784

All continuous variables are in log form. Explanations for “full” variables and IV provided in text. Bootstrapped standard errors (1000 replications) in parentheses. ***p¡0.01, **p¡0.05, *p¡0.1. ***p<0.01, **p<0.05, *p<0.1.

Table 2.A.6: Fuelwood Demand Elasticities - Additional Variables Used in Regression Analysis

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
VARIABLES	Fuelwood Collected (KG/Month)			Fuelwood Produced (KG/Month)			Fuelwood Bought (KG/Month)		
Household head age	-0.0250**	-0.0282**	-0.0300*	0.0206***	0.0207***	0.0262**	0.0133	0.0139	0.0075
	(0.0120)	(0.0128)	(0.0181)	(0.00786)	(0.00783)	(0.0108)	(0.0123)	(0.0132)	(0.0163)
Household head sex (male=1)	0.453	0.427	0.446	-0.0631	-0.0838	0.164	0.0154	0.00833	0.0379
	(0.361)	(0.358)	(0.366)	(0.246)	(0.258)	(0.435)	(0.400)	(0.415)	(0.429)
Land area (acres)	0.132	0.120	0.115	0.0161	0.0170	0.0177	0.599***	0.592***	0.559**
	(0.196)	(0.188)	(0.214)	(0.0380)	(0.0392)	(0.0422)	(0.224)	(0.229)	(0.237)
Asset index	-1.141	-1.377	-1.499	0.0307	0.0457	-0.0963	0.217	0.216	1.031
	(1.148)	(1.110)	(1.334)	(0.356)	(0.372)	(0.408)	(0.746)	(0.729)	(1.839)
TLU herd size	0.242	0.278	0.297	0.0434	0.0329	0.0519	-0.0206	-0.0181	0.0526
	(0.259)	(0.262)	(0.320)	(0.108)	(0.109)	(0.113)	(0.264)	(0.269)	(0.336)
Number adult men	-0.170	-0.170	-0.190	-0.120	-0.116	-0.171	-0.158	-0.178	-0.0127
	(0.172)	(0.161)	(0.207)	(0.0789)	(0.0815)	(0.112)	(0.135)	(0.145)	(0.332)
Number adult women	0.124	0.142	0.140	-0.0933	-0.0900	-0.0315	0.0634	0.0558	0.275
	(0.192)	(0.184)	(0.188)	(0.0908)	(0.0917)	(0.124)	(0.155)	(0.164)	(0.459)
Number of children	-0.0185	-0.00663	-0.00544	-0.0449	-0.0429	-0.0532	0.0279	0.0271	0.0928
	(0.0495)	(0.0523)	(0.0535)	(0.0353)	(0.0351)	(0.0389)	(0.0523)	(0.0539)	(0.130)
Number female children	-0.0710	-0.0686	-0.0748	-0.157**	-0.161**	-0.128	-0.0959	-0.0944	0.0493
	(0.114)	(0.117)	(0.130)	(0.0722)	(0.0765)	(0.0830)	(0.131)	(0.135)	(0.314)
Charcoal price	-0.175	-0.171	-0.182	-0.0750	-0.0597	-0.0777	0.119	0.137	0.169
	(0.315)	(0.301)	(0.309)	(0.185)	(0.185)	(0.187)	(0.278)	(0.282)	(0.297)
Kerosene price	0.649	0.690	0.623	-0.418*	-0.407*	-0.521*	0.229	0.197	0.966
	(0.462)	(0.461)	(0.623)	(0.249)	(0.245)	(0.286)	(0.555)	(0.591)	(1.635)

All continuous variables are in log form and in adjusted per capita units after Cavendish (2002). Bootstrapped standard errors (1000 replications) in parentheses. *** p¡0.01, ** p¡0.05, * p¡0.1. *** p<0.01, ** p<0.05, * p<0.1.

Chapter 3: Underground Knowledge: Estimating the Impacts of Soil Information Transfers through Experimental Auctions

Climate change, environmental degradation, land fragmentation, and rapid population growth represent some of the biggest challenges facing smallholder farmers in Sub-Saharan Africa (SSA). Together they contribute to widespread food insecurity and rural poverty (Sanchez, 2002; Frelat *et al.*, 2016; García-Ruiz *et al.*, 2017). One potential response to food insecurity is to increase crop productivity (yield per acre), though yields have remained stagnant over the past decade. Growth in agricultural production has more commonly been driven by cropland expansion and soil nutrient mining. Soil mining – the insufficient application of inputs into the soil to replenish those removed by harvest and erosion – frequently leads to soil degradation and a downward yield spiral (Drechsel *et al.*, 2001; Tully *et al.*, 2015). This "poverty trap," in which poor yields lead to still lower incomes, further limits the amounts of inputs that farmers can invest in their land (Dasgupta, 1997; Barrett and Bevis, 2015). Because of the highly degraded state of many African soils, the quantities of inorganic fertilizers required to achieve adequate crop yields are often not profitable in the short term (Antle *et al.*, 2006; Marenya and Barrett, 2009). To achieve sustainable agricultural intensification, therefore, it is of major importance that inorganic and organic fertilizers are used in quantities and locations where they are most efficient, profitable, and environmentally appropriate.

A major challenge arises, however, because crop growth response to soil amendments often varies significantly according to soil properties and fertility levels (Tittonell *et al.*, 2008a,b; Kihara *et al.*, 2016). Because of this, as well as the often limited ability of smallholder farmers to purchase fertilizers, plot-specific soil information and recommendations represent potentially significant improvements in fertilizer use and efficiency for staple crops. Little experimental evidence exists, however, as to whether farmers alter their behavior and input choices when presented with improved soil information and input recommendations. In this study, we test whether providing individualized soil test information and input recommendations from low-cost soil analysis kits to a sample of farmers in western Kenya influences their demands for a number of agricultural inputs, including both inorganic nitrogen fertilizers and organic inputs. We develop an economic model that posits that farmers develop perceptions about their farms' soil properties through exposure to information via a Bayesian process which affects their willingness to pay (WTP) for a particular input. The model shows that farmers' WTP for inputs will change as they learn about the optimal combinations of inputs for their particular soil nutrient levels and that they adjust their valuations accordingly.

To estimate WTP for this study, we use experimental auctions among a sample of 884 small-scale western Kenyan farmers randomly chosen from village rosters. We conducted auctions after Becker, DeGroot, and Marschak (BDM; 1964) for various organic and inorganic inputs, implemented in two rounds to measure changes in WTP before and after the information transfers to participants. We test the efficacy of different transfer strategies by randomly dividing the total sample into several treatment groups and a control group. Farmers in the primary treatment group received the individualized soil test results and recommendations for most effective combinations of inorganic and organic input usage.

Results from triple difference estimations show that the effects of the primary treatment are uniformly positive and large for inorganic fertilizers: a recommendation to use DAP (diammonium phosphate) increases average bids for that fertilizer by 61%

compared to the baseline.[16] However, the effects of recommendations for the use of organic inputs (e.g. animal manure, vermicompost, biochar) are mixed and vary significantly by gender. We conduct a cost benefit analysis using the willingness to pay results from our experimental auctions that shows that individual soil tests have positive net benefits under most scenarios, with the net benefits of the tests particularly high for those who had been using relatively expensive inorganic fertilizers on unresponsive soils.

Determining whether a farmer's increased knowledge of his/her soil type and quality alters one's behavior with regard to input demands can help address whether widespread soil testing and accompanying input recommendations can effectively contribute to moving farmers out of poverty by overcoming information constraints. At present, most farmers are unable to discern nutrient deficiencies since they lack access to standard diagnostic analysis. Significant spatial variation of soil properties and fertilizer requirements across individual farmer fields (Tittonell *et al.*, 2005, 2013) limits the ability of farmers to learn from their neighbors (Tjernström, 2016). High-resolution geo-referenced soil data that rely on interpolation also appear to be limited as a mechanism to estimate soil characteristics at the individual farm level (Berazneva *et al.*, 2018). Scaling up the use of low-cost soil analysis kits that provide comprehensive recommendations for plot-specific soil fertility management is one of the more viable and promising options to improve input use efficiency and crop productivity.

This study adds to several strands of literature. A significant amount of previous research has analyzed the effects of information transfers on farmer behavior in the developing world. This study, however, is among the first that uses individual soil tests and personalized input recommendations as the focus of the information trans-

[16] Or, by another measure, this recommendation increases willingness to pay by 24% of the average price compared to the counterfactual group.

fer. A major contribution of this research therefore is to demonstrate the favorable benefits relative to the costs of individualized soil testing in improving the well-being and food security of small-scale farmers in SSA. In addition, while several past studies looking at the effect of information transfers on individuals in the developing world have used data from experiments (Lybbert *et al.*, 2013; Steur *et al.*, 2013; De Groote *et al.*, 2016), as far as we know, this is the first study that uses a two-round BDM auction methodology to analyze changes in behavior after an information transfer in a developing country. A major contribution of this research therefore in using experimental data is to demonstrate the benefits of individualized soil test information, relative to the costs of the technology, in potentially improving farm productivity and economic well-being for smallholder farmers.

The effects on behavior that arise from peer comparisons have also been extensively explored in the behavioral economics literature (Goldstein *et al.*, 2008; Ayres *et al.*, 2009; Allcott, 2011). This study, which includes treatments that seek to discover how comparisons between peers in a village influence the behavior of farmers with respect to agricultural inputs, is the first to examine this research question in a developing country. Finally, this research contributes to the literature on the role of gender in agriculture in developing countries. We conduct separate auctions with men and women to test hypothesized gender differences in perceptions and impacts of soil information transfers. We find significant differences in organic input demands between men and women, which appear to be related to the lower access to organic resources among female farmers. This appears to be due both to the preferences given in intra-household organic input allocation to male-managed plots (Udry *et al.*, 1995) and to lower levels of livestock ownership among female-headed households (Ndiritu *et al.*, 2014).

This article is organized as follows: first, we review some of the most important literature on learning and information transfers, then explain the state of soil degradation and agriculture in western Kenya. Next, we present the theoretical model, followed by a description of the data and the experimental design. Following that, we discuss the empirical results from our difference-in-differences estimations, as well as the findings from our cost-benefit analysis. We conclude by summarizing the implications of this study and the policy recommendations stemming from this research.

3.1 Learning and information transfers

A substantial literature exists that analyzes technology adoption and learning in the developing world (see Feder *et al.* (1985) and Foster and Rosenzweig (2010) for extensive reviews). Several studies in this area have demonstrated that a major hurdle to the diffusion of information regarding productivity-enhancing technologies is the heterogeneity of growing conditions among farms at the village scale. Munshi (2004) finds that during the Green Revolution in India, the adoption of high- yielding varieties was more rapid for wheat than for rice, and concludes that this was due to the fact that wheat-growing regions have more homogeneous growing conditions than those growing rice. Therefore, a farmer can gain more reliable information from his neighbor's experience growing wheat than growing rice. In the U.S. context, because GM soybean seeds are sensitive to individual farm characteristics, Ma and Shi (2015) find that the information impact from peers is less than that from self-experimentation. Meanwhile, Magnan *et al.* (2015) find rather muted peer effects on the adoption of a newly introduced practice, laser land leveling in India, in part due to the heterogeneity in production characteristics that leads to varying levels of yield improvements from using the technology. Tjernström (2016) shows that information exchange about the use of new hybrid seed varieties within a peer network in Kenya can be severely constrained by heterogeneity in soil health characteristics of croplands. However, farmers can reduce uncertainty regarding soil fertility status and input usage by sharing or selling the results of their soil tests among their peers.

Information transfers to individuals can in fact significantly affect the demands for new products and technologies. The economics of advertising literature, in particular, has shown that information transfers can increase the ability of individuals to make optimal matches between product characteristics and their preferences (e.g., Anand and Shachar (2011)). Johnson and Myatt (2006) demonstrate that such "real information" aspects of information transfers increase the dispersion in product valuations by individuals across a sample and rotates the demand curve, as individuals are able to more accurately determine the suitability of the product relative to their wants or needs. A number of empirical papers such as Rickard *et al.* (2011) and Liaukonyte *et al.* (2015) have demonstrated strong effects of in-formation transfers on behavior, often using experimental auctions to elicit incentive-compatible WTP. Rickard *et al.* (2011) show, for instance, that commodity-specific (versus broad-based) information transfers regarding fruit and vegetable attributes lead to an increase in the dispersion of valuations, providing evidence of individuals matching their preferences to products as a result of the new information.

The behavior of individuals can also be affected through social norms or proscriptions (Akerlof, 1980). For example, studies show that farmers may not adopt a new technology or practice if it is antithetical to social traditions (Moser and Barrett, 2006). Many microfinance interventions and savings groups take advantage of social norms to ensure payments or savings (e.g, Dupas and Robinson, 2013). Information provided to an individual regarding the characteristics or behavior of their peers can also significantly affect behavior. An example of this in a different context is that of Goldstein *et al.* (2008), who show that informing hotel guests that the majority of their fellow guests reuse their towels decreases laundry costs more than using a normative message regarding environmental sustainability. Similarly, Ayres *et al.* (2009) and Allcott (2011) find that providing information to utility consumers regarding the energy usage of their neighbors decreases energy consumption by a significant margin. It remains to be seen, however, if such effects also exist in the context of this study, that is, with respect to soil management in developing countries.

In a parallel literature that informs the present study, there has also been an expansion of research that has looked specifically at the role of women in agriculture. In SSA, plots managed by women tend to have lower productivity than men (Croppenstedt *et al.*, 2013). In part, this appears to be due to inefficient allocation of resources within the household, where plots managed by men receive relatively larger amounts of inputs (Udry *et al.*, 1995; Udry, 1996). Women also tend to adopt agricultural technologies and practices at lower rates than men (Doss and Morris, 2001; Ndiritu *et al.*, 2014). This gender gap is likely due to several factors. Gilligan *et al.* (2014) suggest that this result is due to weaker household bargaining positions by women, although this would not necessarily account for the effect in female-headed households where males are absent. Doss and Morris (2001), Croppenstedt *et al.* (2013), and Ndiritu *et al.* (2014) suggest that at least part of the explanation lies with the lower levels of resources available to female farmers. For example, if female-headed households tend to have lower levels of livestock, this would limit the adoption of manure or compost as organic fertilizers since markets for these inputs are not highly developed.

As mentioned previously, farmers do not necessarily hold accurate beliefs regarding their own soil nutrient levels, which may further limit their use of the appropriate kinds and levels of agricultural inputs. Perceptions of soil health are primarily based on yield, which not only is a lagged indicator, but does not necessarily inform the farmer precisely of which nutrients, if any, are deficient (Marenya *et al.*, 2008; Berazneva *et al.*, 2018). Uncertainty therefore exists on the part of the farmer as to whether a particular input will be an optimal match for the soil characteristics of that particular farm plot. The model in this article suggests that this will likely decrease input usage if the farmer is risk averse. Information transfers in the form of results from soil tests may thus be an effective remedy to increase agricultural input adoption and usage, and in turn generate higher crop yields and potentially an escape from the poverty. Along with this study, several recent and ongoing projects are also looking at this or related questions. Fishman *et al.* (2016) find no evidence of an effect on fertilizer usage from an Indian government initiative to provide fertilizer recommendations to farmers. In Tanzania, early results from Michelson *et al.* (2017) show that fertilizer recommendations coupled with vouchers are effective in increasing fertilizer purchases compared to the practices of a control group and a group comprised of participants who only received recommendations. Ongoing research by Corral *et al.* (2017) analyzes the effects of both individualized and average village-level fertilizer recommendations coupled with subsidies for fertilizer purchases. The research reported in this article is thus highly complementary to ongoing research in this area.

3.2 Soil degradation and agricultural input use in western Kenya

Western Kenya, where this research was carried out, has some of the highest population densities in SSA. The majority of rural communities consist of smallholder farm households that depend on own agricultural production to meet most of their food needs. Limiting crop productivity in the humid tropics of Kenya, however, is the heavy weathering of the region's soils. This makes farmers' croplands more prone to nutrient depletion and acidification, significantly undercutting the productivity and fertilizer efficiency of staple crops like maize, sorghum, and beans (Sanchez, 2002; Tittonell *et al.*, 2008b; Tully *et al.*, 2015).Compounding the problems faced by farmers, soils with low fertility cause additional ecological problems, including a greater susceptibility to pests, diseases, and weeds, especially Striga (Witchweed), which has caused substantial yield losses in maize crops across East Africa (Mateete *et al.*, 2010). Over the long run, the well-being of smallholder households can be further jeopardized by downward spiraling crop yields and soil fertility, leading to a "poverty trap" outcome (Dasgupta, 1997; Barrett and Bevis, 2015). To reverse this situation, it is crucial that farmers' usage rates, as well as the efficiency of inorganic and organic inputs on farmlands, be increased to reverse these outcomes and generate sustainable agricultural intensification.

In our research area in western Kenya, however, we find that farmers are often investing, sometimes substantially, in fertilizers, especially inorganic fertilizers such as diammonium phosphate (DAP) which is widely used as a source of nitrogen and phosphorus in crop cultivation. Indeed, Sheahan *et al.* (2013) found that in many regions of Kenya, including the area of this study, farmers are often using nitrogen fertilizers *in excess of* profitable levels. This comes after years of promotion and subsidization of such fertilizers, especially DAP, by the government of Kenya.

DAP and other inorganic nitrogen and phosphorus fertilizers are, however, not necessarily universally effective across farms or even on plots within the same farm. Without comprehensive information regarding the soil nutrient levels on a farm, there exists significant risk of non-responsiveness - that is, that no satisfactory response in crop yields will be achieved with the input investments (Wopereis *et al.*, 2006). Meta-analyses of maize fertilizer responses in SSA show that a large proportion of farmer fields exhibit low yield responses when blanket input recommendations are followed (Vanlauwe *et al.*, 2011; Kihara *et al.*, 2016). On-farm trials by Roobroeck *et al.* (2017) find that rates of inorganic fertilizer applications in line with government recommendations did not increase maize yields by more than one ton per hectare for 19% to 30% of farmers' fields, thus leading to a frequent financial loss on this investment by farmers.

Croplands that demonstrate non-responsiveness to nitrogen fertilizers typically have low amounts of soil organic matter, extractable nutrients, and/or high acidity. For reference, table 3.1 shows the proportion farmers' fields in this experiment where soil nutrient levels were rated very low. Applications of nitrogen fertilizers often do not increase crop yields for soils that have low soil organic matter (Vanlauwe *et al.*, 2002), further decreasing profitability (Marenya and Barrett, 2009). The efficiency of inorganic nitrogen by crops grown on acidic soils is commonly lower, and prolonged use of DAP can acidify the soil (Bekunda *et al.*, 1997), further decreasing fertilizer profitability (Burke *et al.*, 2017). Farmers in western Kenya have little access to soil fertility information services to determine whether crop yields are suffering due to a lack of a soil nutrient (e.g. nitrogen, phosphorus), organic matter, whether their soils are acidic, or whether a combination of these factors is at play in limiting crop productivity.

In sum, field studies in western Kenya show that the use of both organic and inorganic fertilizers is of critical importance in sustainably intensifying agricultural production, particularly on highly degraded soils (Solomon *et al.*, 2007; Ngoze *et al.*, 2008; Güereña *et al.*, 2016). Organic fertilizers like animal manure and crop residues have been used in cropping systems for millennia to replenish soil nutrients. Organic fertilizers not only are a source of soil nutrients, but they also help increase soil organic matter (carbon), reduce soil acidity, enhance water retention, suppress pests and diseases, and stimulate beneficial soil biota (Ngetich *et al.*, 2012). Because of many factors including high transportation costs, most organic fertilizers are produced on-farm in the developing world (Place *et al.*, 2003). Yet, farmers frequently face tradeoffs in allocating household resources for use as fertilizer, since crop residues can serve as animal fodder, fuel for cooking, and for many other uses, reducing the availability of organic resources for agriculture. Gathering, processing, and applying organic inputs are also labor-intensive activities and are thus frequently limited by household labor availability. As a result, household shadow prices (opportunity costs) for these products are high relative to household incomes (Berazneva *et al.*, 2017). Thus, although many households produce and manage manure, compost, and crop residues, their high opportunity costs frequently limit the intensity of their use and adoption as fertilizers.

Organic inputs such as animal manure, traditional compost, crop residues, etc. accordingly can be extremely effective tools to aid in the recovery of degraded soil. However, they will not be equally effective on all types of soils and in many cases inorganic fertilizers should be used in combination with these inputs, especially given the high opportunity costs of their use. Studies have demonstrated that organic and inorganic inputs should be used together to replenish depleted soils, as inorganic fertilizers have higher concentrations of nutrients such as nitrogen and phosphorus, while organic inputs help to replenish soil organic matter and control acidity (Palm *et al.*, 1997; Mateete *et al.*, 2010). At the same time, the combination with inorganic fertilizers enables farmers to offset the high opportunity costs of organic inputs.

For farmers to assess whether they should expend valuable resources – cash, time, opportunity costs – to acquire these inputs, knowledge about the soil fertility of their croplands can be very advantageous. Here we investigate whether providing information to smallholder farmers about the nutrient status of their soils, as well as input recommendations for maize crops, will effectively alter their demands for agricultural inputs. In general, it can be expected that comprehensive decision support tools enable farmers to form more accurate and effective matches between soils, crops and inputs, leading to greater efficacy and profitability of inorganic and organic fertilizers. Understanding how the behavior of farmers is affected by soil fertility information, here determined by low-cost testing kits, and the associated input recommendations individualized to each farmer's circumstances, is an important contribution to research and policies that are undertaken to improve the use of fertilizer inputs not only in Kenya but across SSA.

3.3 A model of farmer information updating

The main objective of this model is to demonstrate the willingness to pay (WTP) of an individual farmer for a particular input combination and the effect on farmer WTP of having received new information about his/her soils. We examine the case of a farmer evaluating his/her expected profit using an input that differs from the input combination the farmer has traditionally used. It may be the case that for a particular farmer, the traditional input combination will be a null set (no inputs used). WTP in this context can be conceptualized as the difference in the utility to the farmer between the use of the novel input set (k), comprised of new input or input combinations used, and the traditional input set (r), consisting of prior inputs or input combinations used (Zapata and Carpio, 2014). For producers, this utility is derived from differences in expected profits, and is known as a variation function (since compensating and equivalent variation are identical for producers (Just *et al.*, 2004, p. 53)). We model this variation function for farmer i in period t as:

$$E(d^k_{it}) = E[\pi^{k,r}_{it} - \pi^r_{it}] \tag{29}$$

where $E(\pi^{k,r}_{it})$ is the expected profits in period t using novel input set k and traditional input set r, and π^r_{it} is the expected profits, all in period t, if only the traditional inputs are used. In other words, the willingness to pay for the new input set is the difference between the maximized expected profits of the farmer using both new and traditionally used input sets, and expected profits if s/he had not used new inputs. If farmer i has not yet adopted the new input set, this value will be zero.

A representation of a farmer's profit function optimizing across the new and traditionally used input sets is:

$$E(\pi^{k,r}_{it}) = E\left[P\left(A^k_{it}(x^k_{it})^{\frac{1}{2}}L^k_{it} + (x^k_{it})^{\frac{1}{2}}L^k_{it}\epsilon_{it}\right) - c^k_t x^k_{it} L^k_{it} + RL^r_{it} + \gamma L^k_{it} + \alpha_i\right] \tag{30}$$

where P is the per unit sale price for the crop, which is assumed to be fixed in the short run and known by all farmers in period t, and R is the net per acre returns using the traditional input set. We let A^k_{it} be the product, and plot-specific agronomic efficiency, that represents the farmer's soil suitability for that particular input set k. Agronomic efficiency is defined as the increase in maize yield per unit of fertilizer inputs applied on the soil of farmer i (Vanlauwe *et al.*, 2011). The agronomic efficiency for input set k is not known by farmers with accuracy, but their estimate is distributed $N(\bar{A}^k_{it}, \zeta^2_{it})$. We define the accuracy of the estimate of the agronomic efficiency A by farmer i for product k as $\theta^k = \frac{1}{\zeta^2}$ $(\theta^k > 0)$, or the inverse of the variance. Term x is the quantity of the new input set used by farmer i in period t, and L^k is the number of acres on which the input set is applied (including land on which traditionally used inputs are also used), while L^r represents land on which only the traditionally used input set is used. Term c is the per unit price of the input set x, which is assumed to be fixed in the short run and known by all farmers in period t. Stochastic variable ϵ represents weather or others sources of variability: $\epsilon \sim N(0, \eta_{it}^2)$. Term α_i is a measure of farmer expertise, ability, or other demographic characteristics such as education, which may affect profitability. We assume $L^k + L^r = L$, L is fixed in the short run, and for simplicity we assume that in estimating his/her expected profit, the farmer assumes s/he will continue to devote at least some land to the traditional set of inputs, thus $L^r > 0$. However, the farmer may not plan to devote any land to the new input set, in which case $L^k \geq 0$, thus $\gamma L^k = 0$.

If we assume that the farmer has CARA (constant absolute risk aversion) risk preferences on the profit function in Equation 2, we have:

$$E(\pi_{it}^{k,r}) = P\bar{A}_{it}^k (x_{it}^k)^{\frac{1}{2}} L_{it}^k - c_t^k x_{it}^k L_{it}^k + R(L_i - L_{it}^k) + \gamma L_{it}^k + \alpha_i - \frac{1}{2}\varrho P^2 (L_{it}^k)^2 x_{it}^k \left(\frac{1}{\theta_{it}^k} + \eta_{it}^2\right) \tag{31}$$

where ϱ is the measure of absolute risk aversion. Expected profit is clearly increasing in the accuracy of the belief of the agronomic efficiency of θ, and in the mean of the perceived agronomic efficiency for input set k, $\bar{A}^k$. We thus have the following FOCs:

$$\frac{\partial E(\pi^{k,r})}{\partial L^k} = P\bar{A}_{it}^k (x_{it}^k)^{\frac{1}{2}} - c_t^k x_{it}^k - R + \gamma - \varrho P^2 x_{it}^k L_{it}^k \left(\frac{1}{\theta_{it}^k} + \eta_{it}^2\right) = 0 \tag{32}$$

$$\frac{\partial E(\pi^{k,r})}{\partial x^k} = \frac{P_t \bar{A}_{it}^k L_{it}^k}{2(x_{it}^k)^{\frac{3}{2}}} - c_t^k L_{it}^k - \varrho P^2 (L_{it}^k)^2 \left(\frac{1}{\theta_{it}^k} + \eta_{it}^2\right) = 0 \tag{33}$$

Due to the complementary slackness condition, if $L^k = 0$, $\gamma \geq 0$, then

$$0 \leq \gamma = -P\bar{A}_{it}^k (x_{it}^k)^{\frac{1}{2}} + c_t^k x_{it}^k + R \tag{34}$$

which is equivalent to:

$$R \geq P_t \bar{A}^k_{it}(x^k_{it})^{\frac{1}{2}} - c^k_t x^k_{it} \tag{35}$$

From Equation 7, we see that the farmer will not use the new input set if the returns to the traditional input set (R) are greater or equal to the net returns of the new input set ($P_t \bar{A}^k_{it}(x^k_{it})^{\frac{1}{2}} - c^k_t x^k_{it}$) for farmer i. Therefore, if the expected return to the new input set k, $\bar{A}^k_{it}(x^k_{it})^{\frac{1}{2}}$, increases, it is more likely that the farmer would adopt the input set. As shown below, receiving a positive information signal related to the agronomic efficiency of input set k for farmer i's soil would be one potential pathway for this to occur.

The variation equation for input set k (producer's willingness to pay) from Equation 1 will thus equal Equation 3 minus the expected profits using the traditional input set, $E[\pi^r] = RL_i$:

$$\begin{aligned} d^k_{it} = E[\pi^{k,r}_{it} - \pi^r] &= P\bar{A}^k_{it}(x^k_{it})^{\frac{1}{2}} L^k_{it} - c^k_t x^k_{it} L^k_{it} + R(L_i - L^k_{it}) + \gamma L^k_{it} + \alpha_i - RL_i \\ &\quad - \frac{1}{2}\varrho P^2 (L^k_{it})^2 x^k_{it}\left(\frac{1}{\theta^k_{it}} + \eta^2_{it}\right) \\ = E[\pi^{k,r}_{it} - \pi^r] &= P_t\bar{A}^k_{it}(x^k_{it})^{\frac{1}{2}} L^k_{it} - c^k_t x^k_{it} L^k_{it} + \alpha_i - RL^k_{it} + \gamma L^k_{it} - \frac{1}{2}\varrho P^2 (L^k_{it})^2 x^k_{it}\left(\frac{1}{\theta^k_{it}} + \eta^2_{it}\right) \end{aligned}$$

Rearranging terms,

$$d^k_{it} = L^k_{it}\left[P\bar{A}^k_{it}(x^k_{it})^{\frac{1}{2}} - c^k_t x^k_{it} - R + \gamma + \alpha_i - \frac{1}{2}\varrho P^2 L^k_{it} x^k_{it}\left(\frac{1}{\theta^k_{it}} + \eta^2_{it}\right)\right] \tag{36}$$

which has the following FOCs,

$$\frac{\partial d_{it}^k}{\partial L^k} = P\bar{A}_{it}^k(x_{it}^k)^{\frac{1}{2}} - c_t^k x_{it}^k - R + \gamma - \frac{1}{2}\varrho P^2 L_{it}^k x_{it}^k\left(\frac{1}{\theta_{it}^k} + \eta_{it}^2\right) = 0 \tag{37}$$

$$\frac{\partial d_{it}^k}{\partial \bar{A}^k} = P(x_{it}^k)^{\frac{1}{2}} L_{it}^k = 0 \tag{38}$$

$$\frac{\partial d_{it}^k}{\partial \theta} = \frac{\varrho P^2 (L_{it}^k)^2 x_{it}^k}{2(\theta_{it}^k)^2} = 0 \tag{39}$$

We can see from Equation 8 that if the amount of land on which the new input set k is used is zero ($L^k = 0$), then the WTP for input set k is 0. We can also see that WTP is decreasing with the cost of the input set c, the net return of the traditional input set R, and the variance, η^2, of the stochastic variable ε. WTP is increasing in the net profit of the new input, in the level of accuracy of farmer i's estimate of his/her agronomic efficiency of k on the soil of i, θ^k, and in the mean of the perceived agronomic efficiency for input set k, $\bar{A}^k$. The closer that the individual farmer believes that the input set k matches his/her soil characteristics, the greater the perceived agronomic efficiency the input set will have, and the more s/he values input set k.

3.3.1 Updating WTP

On learning new information about the farm-specific agronomic efficiency of input set k, farmers will update their WTP. We assume that this information about the agronomic efficiency is accurate, and thus we can represent the updating of his/her WTP as Bayesian. We also assume that the incoming information signal υ regarding the agronomic efficiency of k for the soil of i is distributed $N(\mu_{it}, \sigma_{it}^2)$. The inverse of the variance of the signal, $\frac{1}{\sigma_{it}^2}$, can be thought of as the trust the individual places in the accuracy of the information signal. Intuitively, an individual would tend to have greater trust in information if the information is more precise (i.e. less variable). Thus, if we assume a standard Bayesian updating process (for example, as used by Foster and Rosenzweig (1995)), the accuracy of the farmer's belief in the agronomic efficiency of input set k on his/her soil, θ^k, in period $t+1$ after receiving υ will be:

$$\begin{aligned}
\zeta^2_{it+1} &= \frac{1}{\frac{1}{\zeta^2_{it}} + \frac{1}{\sigma^2_{it}}} \\
\theta^k_{it+1} &= \theta^k_{it} + \frac{1}{\sigma^2_{it}}
\end{aligned} \tag{40}$$

In other words, the accuracy of farmer i's information regarding his/her soil suitability in $t+1$ is equal to the accuracy in his/her belief in the previous period plus his/her trust in the accuracy of the new signal s/he has received regarding the agronomic efficiency of k on his/her soil. Thus, with each new information transfer, the farmer's beliefs in his/her soil characteristics will be nondecreasing in accuracy (it is possible that they may stay the same). Again using a Bayesian process, the updated mean of the farmer's estimate of the agronomic efficiency of k is:

$$\bar{A}^k_{it+1} = \frac{\zeta^2_{it}\mu_{it} + \sigma^2_{it}\bar{A}^k_{it}}{\zeta^2_{it} + \sigma^2_{it}} = \frac{\mu_{it} + \theta^k_{it}\sigma^2_{it}\bar{A}^k_{it}}{1 + \theta^k_{it}\sigma^2_{it}} \tag{41}$$

which states that the farmer's actual belief in the level of his/her soil characteristics and its compatibility with a particular input will increase if the information received indicates an agronomic efficiency greater than was believed previously, and will decrease if the information received indicates that the agronomic efficiency of a particular input is lower than the previous belief.

Assuming variables x, c, L, and η remain constant (as is the case in our empirical setting) and we are analyzing the effect of information on WTP as a result of the information transfer alone, we develop several propositions that we can empirically test:

3.3.1.1 Proposition 1

If $d_{it}^k \neq d_{it+1}^k$, then $\bar{A}_{it+1}^k \neq \bar{A}_{it}^k$ or $\theta_{it+1}^k > \theta_{it}^k$.

This means that if we observe a difference in the WTP for an individual who receives information about the agronomic efficiency of k, then at least one of two possibilities must be true: the information either increased/decreased his/her belief of A^k, or it increased his/her accuracy or certainty of θ^k.

Proof. See Appendix 3.A.1

If we assume a scenario where $|\theta_{it+1}^k - \theta_{it}^k| < \varepsilon$ (at or near zero), then we can conclude that:

3.3.1.2 Corollary 1

$d_{it+1}^k > d_{it}^k$ if and only if $\bar{A}_{it+1}^k > \bar{A}_{it}^k$

If there is no effect of an increase in accuracy on WTP, then the impact of an increase in the belief of the agronomic efficiency of k will increase the WTP for that input set.

Proof. See Appendix 3.A.1

3.3.1.3 Proposition 2

If $d_{it+1}^k < d_{it}^k$, then $\bar{A}_{it+1}^k < \bar{A}_{it}^k$.

Because θ^k is nondecreasing, if the WTP of a farmer for input set k decreases, it must follow that s/he received negative information related to (A^k), the agronomic efficiency of k.

Proof. See Appendix 3.A.1

Empirically testing these propositions can show whether farmers are updating their beliefs as a result of the soil information transfer through a change in their previously held perceptions about their soil fertility and fertilizer efficiencies, and whether these changes are significantly different compared to a counterfactual control group that did not receive the soil information.

3.4 Data and experiment

We collected data for this research in three counties of western Kenya: Bungoma, Busia, and Kakamega. The partner organization, the International Institute of Tropical Agriculture (IITA), selected eighteen villages based on familiarity with the area. The villages covered a wide area of western Kenya (see map in Appendix 3.A.2 for positions of the sampled villages). We obtained village rosters of household heads from village elders or regional chiefs, and randomly selected household heads using a random number generator. Staff from the project then visited each of the randomly chosen household heads, and after obtaining consent, took a sample of their soil. To analyze the acidity and levels of key nutrients in soil samples from farmers' plots, we used the SoilDoc wet chemistry system because it is mobile, relatively inexpensive, and easy to operate (Earth Institute, 2017). Additional information about SoilDoc is given in Appendix 3.A.3.

After two to three months, completed soil test information was returned to the farmers by trained research staff. For each household, we attempted to survey the husband and wife individually, although in many instances, the spouse was not present and could not be interviewed.[17] The survey included questions about household and individual demographic characteristics, household market activity, and agricultural

[17]This was often due to migration, where the husband had gone to work long-term in a larger city, or where the household head was a widow/er.

production, including input use practices, over the past two complete cropping seasons for each of their crops. The final sample consists of 884 individuals in 548 households. Table 4.1 shows key summary statistics for individuals and households.

The table demonstrates that we captured a wide variety of individuals through the random sample selection. The average age and years of education of the respondents was 48.3 and 8.0 years respectively, with large ranges on each variable. Because the number of years of education does not necessarily capture the quality of education for an individual, we also tested their math ability at the time of the survey by asking each respondent to perform a simple multiplication problem, which 56% could answer correctly. The final sample contained a majority of women (58%), due in part to migration by many men to work in cities, and due to the number of widows in the sample. While many individuals in the sample had more than one occupation, 88% identified farming as their primary occupation.

As mentioned above, farms in the sample were generally very small, 1.06 acres on average. This was in part due to design: our selection criteria excluded from the village rosters farms with significantly greater than average farm sizes; these were often commercial farms with absentee landlords. Household sizes of the farms varied significantly, with a mean of 5.29 people per household. Household expenditures also varied significantly across households. Only considering expenditures on food and beverages, average weekly household expenditures averaged 1,228 KSh (roughly $12.00 U.S.), with a range from 0 to 21,000 KSh and a median of 800 KSh. Although most of the individuals sampled were women, a majority of the household heads in the sample are male (55%). Most of the households use inputs on their agricultural plots, with 88% using at least one type of inorganic fertilizer in the past two cropping seasons prior to the survey. The number who used organic inputs (e.g. animal manure) was lower, amounting to 45% of households over the same period. Only 7% of households in the sample did not use any inputs, organic or inorganic, over the same period of time.

As shown in table 3.1, farmers in the sample had highly degraded soils, with high proportions of croplands demonstrating very low levels of N, P, K, S and/or active C. Yet, these soil characteristics were not reflected in farmers' perceptions of soil quality. We find that 85% of farmers ranked the overall soil quality on their farm as average ("3" on a 5-point scale), and only 11.5% ranked their soil as below average. This corresponds to research by Berazneva *et al.* (2018) in Kenya and Tanzania which shows that farmers' perceptions of soil quality are generally unrelated to soil nutrient levels. Therefore, providing soil information transfers to farmers should result in significant updating of their prior beliefs regarding input suitability to their soils.

3.4.1 Experimental auction design

In this research, we use experimental auctions after Becker *et al.* (1964) to estimate incentive-compatible WTP for agricultural inputs among individual respondents in our sampled households. We implement these auctions in two rounds, both before and after the information treatment to measure the effect of the information transfer on behavior.

Experimental auctions have been commonly used for decades in the industrialized world, but have only recently begun to be used in developing countries to elicit incentive-compatible estimates of WTP. Several recent studies have used experimental auctions to investigate the demand for a variety of products, such as insecticide-treated bednets in Uganda (Hoffmann *et al.*, 2009), biofortified maize in Ghana (De Groote *et al.*, 2011), and aflatoxin-free maize in Kenya (De Groote *et al.*, 2016). Many studies use a Becker-DeGroot-Marschak (BDM) methodology to obtain their WTP estimates, as they have been shown to be incentive-compatible and are convenient to implement at the individual level (Shogren, 2005). However, as discussed in Morawetz *et al.* (2011), precautions should be taken when transferring experimental auction methodologies used in the developed world to Africa, including the need for intensive training of auction experimenters and multiple practice auction rounds for each participant.

BDM auctions, as used in this research, are particularly suited to field experiments as participants make bids against a randomly generated price (Becker *et al.*, 1964). This makes it possible to conduct the auctions with individual participants, which limits the bias that might otherwise arise from the respondent's observations of the behavior of other participants. While this methodology is somewhat more complex than more conventional auctions, with practice auctions to familiarize the participants with the methodology in Ghana, Morawetz *et al.* (2011) find that the increased complexity does not lead to bias in the WTP results when compared to a first-price auction. One innovation in our study was using a two-stage BDM auction, where we conducted a baseline auction for the agricultural inputs prior to receiving the treatment, and a post-treatment auction was then conducted immediately afterward.

Prior to beginning the baseline auction for the inputs, the enumerators conducted practice auctions with the participants in which the auction methodology was explained in detail (all auction scripts are in Appendix 3.A.4). The enumerator explained to the participant that s/he would receive a cash endowment and make bids for several items, and afterward, one item would be chosen at random and a random price would be chosen for that item. If the respondent bid at least the amount of the random price, they would pay that random price and receive the item. Otherwise, they would keep the cash endowment. The enumerator gave each participant 70 KSh (about 0.69 USD) and the participant bid on different varieties of cookies and 50 KSh cash notes.[18] After this first practice auction, the enumerator gave each participant another 70 KSh and repeated the practice auction to assure participants' familiarity with the auction methodology.

For the actual experimental auction, each participant was given a cash endowment of 700 KSh (about 6.90 USD).[19] We found that in our experiment 95% of bids were below 500,[20] and the mean bid across all inputs and all quantities was about 200, thus we conclude that this was an effective choice for the cash endowment.

In this study, the auction was first conducted with the household head, and the spouse (if present) was asked to leave. After the household head participated in the auction, the spouse, if available, was asked to return for his/her auction. After receiving the cash endowment, the enumerator read a brief statement that described some of the inputs that might be new to the participant (the script used by enumerators included in Appendix 3.A.4). The participant then made bids for DAP, biochar,[21] a biochar DAP mix, vermicompost,[22] a biochar vermicompost mix, and cow manure

[18] The motivation for auctioning cash notes is to emphasize that the participant should be bidding what s/he perceives as the true value of the good. Thus, for a 50 KSh cash note, the participant should bid 50 KSh. If they did not, then the enumerator would explain to the participant why this is the optimal strategy.

[19] In choosing the amount of the cash endowment, we needed to ensure that the endowment was sufficient enough that the participant could bid his/her true WTP. If it was too little, the bids might be biased downward. On the other hand, some research has shown that a larger cash endowment can lead to overstated WTP estimates (Loureiro *et al.*, 2003). Adding to this concern is that the majority of the farmers in this area are poor, and average agricultural wages are about 300 KSh per day. Thus, 700 KSh is about twice the value of the more expensive quantities of goods that we were auctioning, and is therefore a standard cash endowment in line with the literature (Morawetz *et al.*, 2011; De Groote *et al.*, 2016).

[20] 98% of bids were below 700. If a respondent bid greater than 700 KSh, the enumerator was instructed to remind the participant that they would be responsible for payment above 700 if necessary and asked them to confirm their bid.

[21] Biochar results from the thermal decomposition of biomass in the absence of oxygen, generating a type of charcoal. It is produced from left-over plant material of field crops on farm like maize cobs and stovers, rice husks and haulms, sugarcane bagasse, coconut shells, and others. If applied to soil at the optimal rate, biochar helps to improve crop production by increasing the uptake of fertilizers, manure and water (Lehmann and Joseph, 2009).

[22] Vermicompost is the end-product of the breakdown of organic matter by an earthworm, also called worm castings. If applied to the soil at the optimal rate vermicompost will improve crop production because it contains substantial amounts of nutrients, has a large water holding capacity

and the enumerator presented them in a random order. After all of these bids were collected, a random number generator on the enumerator's tablet computer assigned the participant to one of four treatment groups:

Treatment 1 (Input Recommendation [IR]): Enumerators presented the participants with their soil test results, which included a binary indicator indicating whether a particular soil nutrient level was low, and provided fertilizer recommendations tailored to their individual farms using the SoilDoc system.

Treatment 2 (Village Comparison [VC]): Enumerators showed the participant a chart comparing their soil test results with other anonymized test results for all the participants in their village. The enumerator pointed out the average nutrient level in the village, and the results for each participating farmer relative to the village distribution, but did not provide specific fertilizer recommendations about nutrient levels.

Treatment 3 (Combined Treatment [IR&VC]): Participants in this treatment received both Treatments 1 and 2 together.

Control: Participants received no information transfer between auction rounds.

Participants assigned to Treatments 1 and 3 received a copy of their soil test results that included a binary (yes/no) indicator of whether to apply or not apply N, P, K, and organic C inputs based on critical thresholds of soil nutrient requirements for cereal production in SSA built into the SoilDoc system (an example is shown in Appendix 3.A.5). The enumerators, who had been extensively trained in making input recommendations based on the soil test results, then gave a detailed explanation of the meaning of the soil tests, and how farmers could optimize their input usage based on this information. For example, a farmer with low nitrogen or phosphorus was

and enriches the soil with micro-organisms (Jack and Thies, 2006).

advised to use DAP, CAN, or NPK fertilizer, unless their soil was acidic ($pH<5.5$), in which case the farmer was advised to avoid DAP because of its potential to further acidify the soil. For soils with low carbon levels, compost, animal manure, biochar, and crop residues were typically recommended. The enumerator also explained the benefits of the complementary use of inorganic and organic resources in improving soil health. After these recommendations, the enumerator answered any questions the respondent had.

If the participant was assigned to Treatments 2 or 3, they were presented with a chart that showed their soil nutrient levels compared to others in the village (Appendix 3.A.5). The enumerator showed the participant his/her placement on the chart, and also identified the average in the village. Because most soils in the sample were of poor quality, the enumerator explained to the participant that this was a relative position, and explained where the threshold was for adequate levels of the nutrient in the soil. However, for those in Treatment 2, no input recommendations were made to the participant.[23]

The second auction round proceeded exactly as the first. Immediately after the baseline auction (and before receiving any treatment) each participant played a five minute memory game on the enumerator's tablet computer. This was done primarily so that the control group would have an activity between the two rounds, and for consistency this was done for all participants. Afterward, the tablet computer randomly chose one auction round (the baseline or the second round), one product, and a random price. If the participant had bid at least the amount of that random price for that item in that round, they paid the random price and received the input, otherwise they kept the full cash endowment.

[23]For consistency, those participants in Treatment 3 always received the soil test results and recommendations before seeing the village charts.

3.4.2 Sample from experiment

By having the tablet computer randomly assign participants to one of the four experimental treatments between the auction rounds, we prevented any possibility of bias arising from prior knowledge of the participant's treatment group by the enumerator. The minor disadvantage to this method was that this created unevenly distributed participant numbers in treatment groups, as table 3 below illustrates. Because individuals were randomly assigned to the various treatment groups and a control group, we would assume that there would be no significant differences in the characteristics of individuals between the various groups, nor in the likelihood of a particular enumerator implementing that treatment. In Appendix 3.A.6, we include balance tables (tables 3.A.6-1 through 3.A.6-4) that show average differences between those in a particular treatment/control group and those who are not in that group. Overall, these tables demonstrate that the randomization was effective, as most variables are balanced. One notable exception is farmland area, which is the only variable that is unbalanced in more than one treatment group (Treatment 1 and Control). The magnitude of the difference, however, is not particularly large (0.92 and 1.17 acres for Treatment 1 and non-Treatment 1, respectively, and 1.29 and 1.04 acres for Control and non-Control, respectively). We include farmland in all regressions to help prevent potential bias from this imbalance affecting the estimation results.[24]

[24] Overall, we have 884 individuals making bids in two auction rounds for two quantities of six inputs, providing a total bid sample of 21,216. To avoid making any inferences based on extreme values, we decided to trim the sample in two ways. First, we took the difference between the second bid and first bid for a particular input-quantity for an individual, and dropped both bids if the difference was in the top or bottom 1% of the sample. These individuals were outliers who changed their bids by extreme amounts between auction rounds. This amounted to 358 total bids or 1.7% of the sample. We next dropped any remaining bids that were at least double the cash endowment (1,400 KSh), as these were not realistic bids by participants. This eliminated another 42 bids. The final sample size of all bids for both auction rounds used in the analysis is 20,816. By trimming the sample, there were no individuals who were completely eliminated, and the number of individuals remains at 884. Table 3.A.6-5 in the Appendix shows the average bids and standard deviations for each input by treatment and by auction round.

3.5 Empirical model

Using triple and quad difference estimations, we take a closer look at the precise input recommendations (for Treatments 1 and 3) and placement on the village charts (for Treatments 2 and 3) and analyze the direction and the degree to which these treatments affect farmer valuations of different types of inputs (testing Corollary 1 and Proposition 2). We find that the treatments do have an impact, but the impacts are heterogeneous across gender, treatment, and input type. Our primary results from Treatment 1 show that the soil test results and input recommendations had a large and significant effect for inorganic fertilizers (e.g., DAP): recommendations to use DAP led to 61% higher average bids compared to the baseline. For organic inputs, the results are statistically significant, though lower, and we find evidence of heterogeneity in the effects by gender (discussed below). In addition, we conducted a cost-benefit analysis of the program (described below) and find that under most conditions, an enlarged program of soil testing and fertilizer recommendations is likely cost-effective.

3.5.1 Difference-in-differences estimation

To empirically test Corollary 1 and Proposition 2 and analyze the directional effects of the information provided, we use a difference-in-differences estimation with three differences: the auction round (first and second), treatment status (particular treatment and control), and information type (treatment specific – in Treatment 1, for example, this is the difference between those who received a positive recommendation to use a particular type of input, and those who were recommended not to use that input).

A key assumption of the difference-in-differences estimation is the parallel paths or parallel trend assumption, that is, the two groups being compared would have the same trend in the absence of any treatment (Angrist and Pischke, 2008). Because we are using triple-differences (auction round, treatment status, and information type), we have two sets of parallel paths that should hold: the difference between the treatment and control, and the difference in the type of information received. The parallel paths assumption in the former should hold based on the random assignment: both the treatment and control were randomly assigned at the time of the auction and the characteristics among these groups are well balanced on the whole (see tables 3.A.6-1 through 3.A.6-4). In the latter case, the parallel paths assumption for the difference in the information treatment is less clear-cut. If parallel paths existed, this would mean, for example, that individuals in Treatment 1 who receive information advising them to use more organic inputs would, in the absence of treatment, behave in the same way as those who received advice that they do not need to use organic inputs. Because farmers have limited knowledge of the nutrient levels of their soils, especially with respect to specific nutrients, the parallel paths assumption seems plausible. Using the control group as a guide, we show estimation results in Appendix 3.A.7 (tables 3.A.7-1 and 3.A.7-2) that support this assumption. We conclude that there is sufficient evidence to support the existence of parallel paths in this sample.

We therefore perform triple difference estimations for each treatment, comparing the impact of each treatment compared to the control:

$$\begin{aligned} d_{ikt} = \alpha &+ \delta_1 round_t + \delta_2 treatment_i + \delta_3 info_i + \delta_4 (round \times treatment)_{it} \\ &+ \delta_5 (round \times info)_{it} + \delta_6 (treatment \times info)_i \\ &+ \delta_7 (round \times treatment \times info)_{it} + \sum_n \phi_{in}\beta_n + \sum_k \omega_k \iota_k + \xi + \varepsilon_{ikt} \end{aligned} \tag{42}$$

where $round$, $treatment$, and $info$ are binary variables. The variable $round$ indicates the auction round t, $treatment$ indicates whether the individual is in the treatment or the control, and $info$ is the information type that the individual receives (or would have received if in the control group). For example, in Treatment 1, $info$ would be equal to 1 if an individual either received a recommendation that s/he should use input k on his/her farm, or would have received this recommendation given his/her soil test results, but was in the control group. Interactions between these variables are included, and the coefficient of interest, δ_7, for Treatment 1 measures the impact of the information on those in the treatment group after the second auction compared to those in the control group in the baseline auction who would not have received a recommendation for input k. Exogenous characteristics of the individual, household, and farm are given by ϕ_n, ω controls for the input type k, ξ are fixed effects for enumerator, village, and survey month, and ε are i.i.d errors. Additionally, given the strong differences in results between men and women in the mixed model, the difference between men and women is also added to our estimations.[25]

3.6 Results

We first look at the results of the triple and quad differences for Treatment 1 (input recommendations). In this treatment, the individual received the results of his/her soil test and the enumerator also made targeted recommendations based on these results. For example, if the participant had a farm with low soil nitrogen levels, the enumerator would recommend using DAP (unless the soil was also acidic), CAN (calcium ammonium nitrate fertilizer), or NPK (nitrogen phosphorus potassium fertilizer). If nitrogen levels were high, the enumerator informed the farmer that the use of nitrogen inputs was not a priority. If active carbon was low, the enumerator would recommend the use of animal manure, crop residues, compost, etc. For each individual, we create a variable ("Use nitrogen input" recom.), which is "1" if a nitrogen

[25] We show in Appendix 3.A.7 that the parallel paths assumption still holds with this additional difference added.

input (DAP) was recommended (78.12% of individuals), and “0” otherwise. Likewise, we create a variable (“Use organic input” recom.) for organic inputs (e.g., animal manure), which is “1” if an organic input was recommended (66.21% of individuals), and “0” otherwise. Those in the control group also have observations for this variable, indicating the potential treatment that was never received during the experiment. For example, a farmer with low soil nitrogen levels in the control group would have a “1” value for “Use nitrogen input” variable, though being in the control group, would not receive the actual recommendation. Therefore, as a shorthand in the discussion of our triple difference results, we refer to the “counterfactual group” as those in the control group who, if had they been randomly chosen to be in the treatment group, would not have received a recommendation to use a particular input.

Table 3.4 shows the results of these estimations. Looking first at the results of the treatment on bids for DAP, we see that in the triple difference estimation (Column I), those in the treatment who received a positive recommendation for inorganic nitrogen fertilizers on average bid 61.97 KSh more than those in the counterfactual group. The average quantity of DAP auctioned was about 2.5kgs, and at the time of the survey, one *goro-goro* (about 2kgs) of DAP in the local market was about 200 KSh. Therefore the recommendations for DAP led to an economically meaningful effect on farmer valuation for the input.

In our estimations, we use clustered standard errors to control for possible within-village dependence between individuals, but the small number of village clusters (17) may lead to underestimates of the standard errors. Therefore, we also estimate adjusted p-values using the Wild bootstrap method after Cameron *et al.* (2008), which corrects for the small number of village clusters. For the triple difference estimations in Column I, both of the p-values from these methods are near zero, and the results are highly statistically significant.

Moving to Column II, which has results for the quad difference estimates that include gender, we see that including the difference between men and women lowered the average impact of the soil tests to 49.56 KSh, while the difference for women compared to men in the counterfactual are positive but not statistically significant (22.04 KSh). This indicates that men and women did not have statistically significant differences in their behavior in the auctions with respect to DAP fertilizer.

Marginal calculations from these experimental results for DAP on treated individuals show that after setting all regressors at their means, the average predicted WTP at the baseline for one kilogram of DAP is 56.6 KSh. After a positive recommendation to use DAP, the predicted WTP for DAP increases to 91.2 KSh, a 61% increase. After a negative recommendation, the average WTP for DAP decreases to 6.8 KSh. These values are all below the average market price for one kilogram of DAP in the project area (100 KSh), which we would expect: farmers should not be willing to pay more than the market price for a particular input. These values along with 95% confidence intervals are included in figure 1. This appears to support Corollary 1 and Proposition 2, which imply that farmers' WTP for a fertilizer generally move in the direction of fertilizer recommendations.

Columns III and IV on table 3.4 show estimation results for bids by participants on all organic inputs.[26] Looking at the triple-difference estimates, we find that those in the treatment group who received a recommendation to use organic inputs increased their bids by a modest 17.8 KSh (significant at p=0.05) compared to those in counterfactual group. When we include a gender difference in a quad-differences estimation as shown in Column IV, however, we find that the information treatment appears to have had no overall effect compared to the counterfactual group. For

[26]These included animal manure, biochar, compost (vermicompost), biochar-compost mix, and biochar-DAP mix.

women though, we see an increase in the average bid by 28.5 KSh compared to men in the counterfactual (though less statisically significant).

To understand these estimation results, it is helpful to go back to the different levels of access that men and women have to organic inputs. When we break down the data to the crop-plot-season level, we see that women use far lower levels of organic inputs on the crops they manage compared to men. Because women face greater constraints in obtaining organic inputs, when our auctions essentially created a temporary market at their home and recommended to women that their soil results indicated that they should use organic inputs, women strongly increased their bids in hopes of making a purchase. For men, slackness may exist in the supply of organic inputs, as they can reallocate organic inputs from the woman's plot to their own, or they may allocate fewer organic resources for cooking, as building material, etc., and instead apply these inputs to agricultural uses. Women in the household, however, are typically not able to make these input reallocations due to their underlying scarcity. We believe that the intrahousehold allocation of organic inputs is likely the major reason why there was no significant effect from the organic fertilizer recommendation for men compared to women.

These conclusions echo results from other studies, where men in SSA are found to be the primary household decision-makers in the allocation of agricultural inputs (Udry *et al.*, 1995; Udry, 1996). As decision-makers, men are more likely to allocate organic inputs to their own plots than to their spouse's. In addition, female-headed households are less likely to own livestock. Markets for organic inputs are nearly non-existent in rural Kenya, and most organic inputs are obtained from a household's own animals. Thus, lower levels of livestock ownership in female-headed households are a significant constraint to the use of organic inputs.

3.6.1 Additional treatments

In the remaining treatments, we test whether comparisons with peers can help incentivize farmers to invest in their soil quality and optimize their fertilizer usage. Project enumerators presented participants in Treatments 2 and 3 with five charts, each for a different soil nutrient (Nitrate, Phosphate, Phosphorus, Potassium, and Active Carbon). On each chart, soil test results for each household were plotted so that an individual could compare his/her soil results with his/her village peers' results (example in Appendix 3.A.5). The charts also show the village mean of the soil results for that particular nutrient.

To estimate the effects of this information on WTP, we create a variable that divides the placements in each chart into quintiles and then averages the quintile placements for each individual.[27] We then create a binary variable that is equal to "1" if their average placement is greater than the third quintile of village soil nutrient levels (432 individuals), and "0" otherwise (452 individuals). This variable represents the broad impression that the respondent received from viewing his/her placements on the charts. Similar to Treatment 1, those in the control group also had observations for this variable, indicating the potential treatment that was not received during the experiment.

The results from Treatment 2 (village comparison) are given in table 3.5. The first column shows the triple-difference estimation for all inputs that were auctioned. We see that for those who had an average soil nutrient quintile placement in their village greater than "3" (i.e., above-average levels of soil nutrients compared to their peers), they on average decreased their bids by 13.23 KSh (significant at the 5%

[27]For example, if an individual is in the first quintile in his/her village for nitrogen, in the third quintile for phosphorus, second quintile for potassium, fifth quintile for sulfur, and third quintile for carbon, his/her average would be 2.8, meaning their overall soil nutrient level was below average (3.0).

level) compared to those in the control who had an average soil nutrient quintile placement less than or equal to "3". This suggests that participants who had soil nutrient levels that were better than average decreased their bids after learning their relative position. Column II shows results including an additional gender difference. In this estimation, compared to those in the control group who have below-average soil quality, there is a stronger decrease in bids when participants learned they had better than average soils (-22.2 KSh).

The results for DAP from Treatment 2 in Columns III and IV of table 3.5 are insignificant and near zero, indicating that the treatment did not impact these individuals. However, the results for organic inputs in Columns V and VI are similar to the overall input results in Columns I and II: there is a strong overall decrease in bids when participants learned that their soils have better than average soil quality compared to the those in the control group with below average soil quality.

Farmers in general, on learning that they have below average soil quality, increase their demand for organic inputs. However, if they see their soil quality is above average, they decrease their bids, reverting towards the village mean. This decrease in average WTP for inputs among individuals who learn that their soil is above average represents what social psychologists describe as a "boomerang effect," or an unintended consequence of peer information. For example, studies have documented that those consuming below average amounts of electricity will increase their electricity consumption if they learn that others are consuming more (Clee and Wicklund, 1980; Schultz *et al.*, 2007). This is an example of the "destructive power of social norms," and in our context, is related to farm households' knowledge that their soils have higher soil quality than average in their village. These outcomes can be avoided, as discussed by Cialdini (2003), by including "injunctive norms," which suggest to the respondent what they should do as a result of new information. Participants in

Treatment 3, then, receive the village comparisons *in addition to* the input recommendations that may serve as injunctive norms.

The results for WTP for fertilizer inputs in the combined treatment in Treatment 3 are presented in Tables 3.6 and 3.7. In Table 3.6, which reports results for Treatment 3 using the same quintile variables as the estimations for Treatment 2, the estimated coefficients are generally smaller in magnitude than those for Treatment 2 and lack statistical significance. The decrease in magnitude compared to the coefficients from the Treatment 2 results indicate that the input recommendations are affecting behavior in the opposite direction as the social norms (from those who have above average soil quality), effectively eliminating the boomerang effect. In table 3.7, we estimate the results from Treatment 3, showing the effects of the input recommendations conditional on having also seen the village comparison charts. The direction of the results is similar to Treatment 1: the effect of the recommendation to use DAP is positive and statistically significant across the full sample. When the gender difference variable is added, however, the coefficients remain positive but lose statistical significance. Overall, the treatment impacts on WTP for those in Treatment 3 are smaller compared to Treatment 1.

This outcome is a bit surprising; we expected that Treatment 3 might have the strongest effect as the participants received the most information and the enumerator spent the most time discussing the results. However, there are several reasons why this may not be the case. First, the combination of input recommendations and comparisons with one's peers may have influenced the participants in opposite directions. For example, if a respondent had higher than average soil quality compared to his/her neighbors, the results from Treatment 2 (Column I) suggest that s/he would decrease their WTP for inputs. However, a recommendation to use an input would likely increase one's WTP for that input. This therefore could lead to the offsetting effects that we see on table 3.7 for Treatment 3. Also, the relatively long amount of time spent on this treatment compared to the others may have caused some participants to lose focus and their bids may have been less accurate. Finally, due to how the random assignment was implemented, Treatment 3 contained the fewest number of participants. As discussed earlier, the tablet computer randomly assigned a treatment to a participant between the auction rounds. As a result, the treatment groups had uneven numbers of participants. Treatment 3 had the fewest participants in total, especially among men, possibly leading to coefficients estimated with less precision.

3.6.2 Cost-benefit analysis

The results from the experimental auctions in this study show that personalized fertilizer recommendations exert an optimizing effect on the behavior of small-scale farmers with respect to their input allocations. This raises the question – relevant to SSA governments and NGO's – as to whether scaling up this intervention is a cost-effective method to improve crop productivity and enhance farmer well-being. A related question is whether the SoilDoc system (or potentially other forms of soil testing technology, such as soil spectroscopy) used to assess soil nutrient levels could be self-financing, wherein revenues generated from providing plot-specific soil fertility information to farmers could cover the expenses involved, rather than relying on subsidies or other sources of financing.

Table 3.8 presents the assumed component parameters and estimated results of cost-benefit analyses that assess the net benefits of testing soils under 12 different price-cost-yield scenarios that might be faced by farmers in western Kenya (the methodology and assumptions for the cost-benefit analysis are described more fully in Appendix 3.A.8). Because DAP fertilizer is widely available for purchase in the study area and farmers are very familiar with its price, we focus on its net benefits.

From the soil test information, we can identify two groups of farmers: those for whom DAP is recommended (Group A), and those for whom it is not recommended (Group B) (these groups constitute 78.1% and 21.9% of the sample, respectively). For those in Group A, the net benefit is therefore the average increase in crop yield minus the cost of the soil test and the cost of the additional DAP used. Based on the results from the soil testing, we assume for this analysis that use of DAP is ineffective on the soils of those in Group B and does not affect crop yield; thus the net benefit for Group B is the average change in the value of DAP used minus the cost of the soil test.

Although we include 12 scenarios in table 8, the discussion here focuses on Scenario 9, which uses average values for many of the parameters and a per sample cost of the soil test of 1,000 KSh (approximately $10 U.S), which includes program and infrastructure costs for a large-scale project (additional detail on cost estimates are available in Appendix 3.A.8). Under these assumptions, Scenario 9 shows that farmers in Group A would enjoy an average net benefit from the soil test information of 2,481 KSh per hectare of maize planted, while those in Group B would have a larger average net benefit of 13,894 KSh per hectare. The higher average net benefit for those in Group B is due to the large savings that accrue from not purchasing DAP, a relatively expensive input, for use on unresponsive soil, i.e., soils that scored above the threshold for N or had acidic soil. Despite the fact that most farms in this area are less than one hectare in size, the amount of fertilizer inputs used – and thus the potential savings from not purchasing DAP when it is unnecessary – are nonetheless significant; recall from Table 4.1 that the mean amount spent on food and drink per week for a household in the survey is 1,229 KSh. For farmers in Group B, extension agents would recommend that they allocate their savings toward fertilizers that are expected to be more effective on their soils than solely using DAP fertilizer.

For nearly all of the scenarios identified in table 8, the improved information about soil fertility and nutrient input requirements leads to major net benefits, calculated using the total cost of the soil analysis. This leads to the question whether farmers would be willing to pay for individualized soil information. Ongoing complementary research by Fabregas *et al.* (2014) in western Kenya suggests that farmers are willing to pay for soil information, indicating that inexpensive soil testing may well be an effective tool for increasing crop productivity and food security in SSA.

3.7 Discussion and implications

In Sub-Saharan Africa, smallholder farmers have access to very little information about the soil fertility of their croplands that can be used to facilitate profit-maximizing decisions regarding the use of fertilizer inputs. Although individuals know when soil fertility is lower than in the past or than that on nearby croplands, they are usually unable to discern what specific types of nutrient inputs are required for improving crop yields. This can lead to a downward cycle of agricultural productivity, farm incomes, and overall living standards, in which progressively fewer resources are available to spend on fertilizer input for croplands.

This article analyzes whether providing small-scale farmers with soil test results and associated information in the form of personalized agricultural input recommendations will affect their behavior and lead to improved optimization of their agricultural input choices. We took soil samples from randomly selected farms of 884 individuals in western Kenya, analyzed them, and returned to the households with farm specific test results. At the time of the survey, we divided the sample into three information treatment groups and a control group. To measure the impact of the information treatments, we used a two-round experimental auction methodology after Becker *et al.* (1964) and compared these treatment groups to a control group. While experimental auctions have been used in the recent past in SSA, especially to measure WTP for novel food items (for example, De Groote *et al.*, 2011, 2016), this is the first study that, to our knowledge, uses a two-stage BDM field auction in SSA to test the impacts of information transfers on behavior.

To test Corollary 1 and Proposition 2 (explained above) and acquire a more precise understanding of the impacts of the type of information received with each treatment, we use triple-difference estimations (differences in auction round, treatment-control, and information type). For Treatment 1, we specifically look at the difference in WTP between those who receive a particular input recommendation (such as those advised to use an organic input) and those who do not. Overall, we find a statistically significant and economically large effect in the change in WTP for DAP fertilizer between those who are recommended a nitrogen input in the treatment group and those in the counterfactual control group. Overall, the recommendation to use DAP fertilizer increased bids by 61% compared to the baseline. For organic inputs, however, we find a significant effect, though one smaller in magnitude. When we add a fourth difference to the estimation (between men and women), we find no overall effect of the recommendation to use organic inputs compared to the counterfactual control group, but for women, there is a statistically significant and positive effect: women in the treatment group who are recommended to use an organic input (e.g., animal manure) to recover soil health on average bid 28.5 KSh higher than men in the counterfactual control group. We believe that the differential impact between men and women from the organic input recommendation is connected to the general lack of access to resources among women in the household.

In Treatment 2, enumerators provided the participants with charts that enabled them to compare their own soil quality levels to that of their peers. We find that, in general, if their soil quality is above average, they tend to decrease their bids for agricultural inputs, suggesting a counterproductive effect of the peer information. Participants in Treatment 3 received both types of information transfers (input recommendations and village comparisons). We find that this treatment had fewer statistically significant effects than either Treatments 1 or 2. We posit that the primary reason for these results are the offsetting effects of the two types of information treatments: comparisons with peers appears to decrease bids for agricultural inputs among those who have higher than average soil quality levels, while recommendations to use inputs for the majority of individuals tend to increase bids.

The results from Treatment 2 inform us about the potential limitations of using peer comparisons to influence behavior. The very limited number of economic studies that have tested the effects of peer comparison on behavior (Goldstein *et al.*, 2008; Ayres *et al.*, 2009; Allcott, 2011) show that they lead to private and public benefits. However, it appears that showing the peer comparisons alone (Treatment 2) can lead to a "boomerang effect," as those with soil quality above-average decrease their demand for inputs. Treatment 3 seems to have reduced the boomerang effect, decreasing the magnitude of the effects to the point where they are not statistically significant. Cialdini (2003) and Schultz *et al.* (2007), among others, demonstrate that "injunctive norms" can be used to mitigate this negative impact of peer comparisons, and our study appears to show that input recommendations in this context can serve the same purpose.

Throughout this research, we attempted to reduce the information constraint that farmers face in optimizing their agricultural input choices. However, farmers face numerous other constraints, including the lack of resources available to invest in agricultural inputs, lack of financial liquidity, and high opportunity costs of domestically available inputs such as animal manure. By providing a cash endowment to participants through the experimental auction design, we eliminate any liquidity constraint and directly provide resources for the bidding. Moreover, during the explanation of the auction method, it was made clear to participants that they could use the cash endowment for other purposes outside of the auction. This assures us that our estimates are incentive-compatible since participants' actual money was used during the auction. These results therefore should be interpreted as those that would arise given access to sufficient liquidity. Other research, such as the well-known Duflo *et al.* (2011) study, analyzes methods to alleviate liquidity constraints among farmers that arise from the timing of fertilizer purchasing. This article therefore attempts to address one important constraint, accurate information, and while we also address the liquidity constraint, we cannot adequately address all constraints simultaneously.

While we advocate the above explanations of our results, other explanations for the participant's behavior during the auctions are possible. Specifically, it could be argued that in the auctions, the participants may have attempted to bid in a manner to appease the enumerators. This "experimenter demand effect" (Zizzo, 2010), it is argued, could bias the auction results. A "house money" effect also may have caused bias as experiment participants usually make more risky decisions with "windfall" cash that they do not fully consider their own (Thaler and Johnson, 1990). While not dismissing these effects, we believe that several aspects of the experiment mitigate the possibility of these sources of bias.

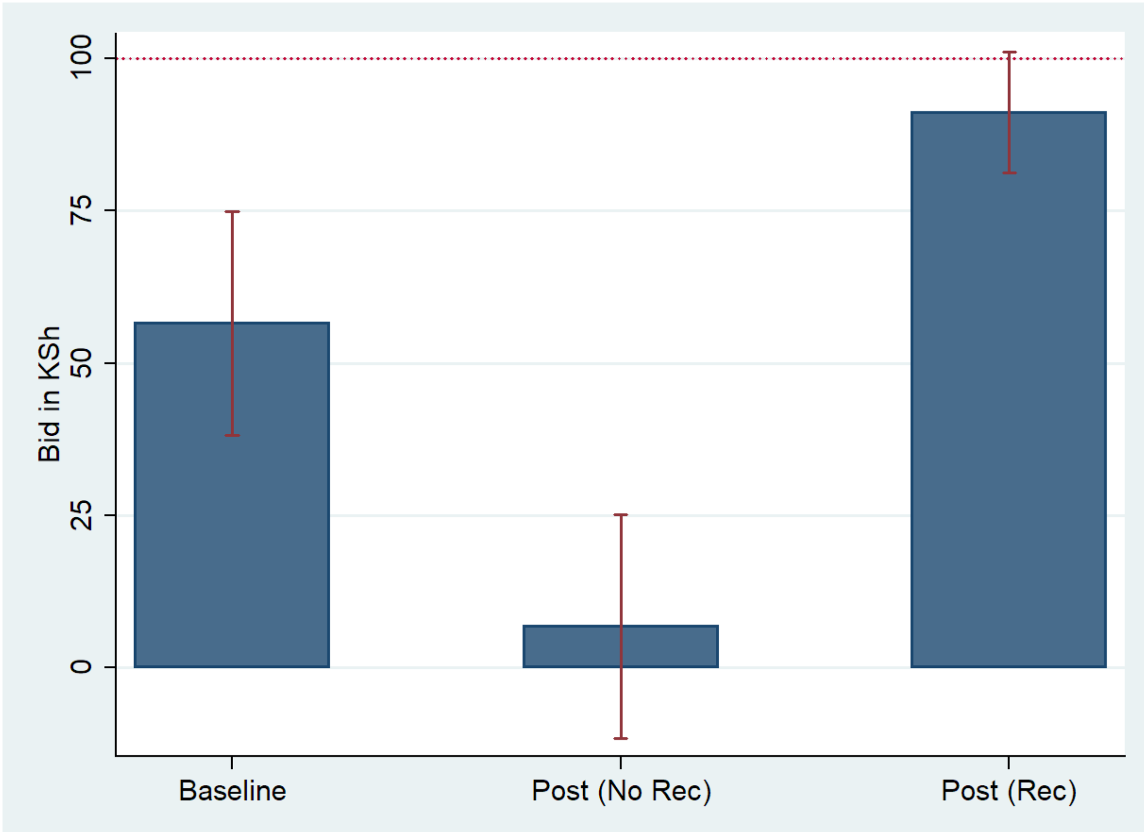
The first is that the cash endowment was made prior to the auction, which Davis *et al.* (2010) show mitigates the house money effect: by giving participants actual possession of the money, this transforms the house money into their own money in the minds of the participants. Second, enumerators strongly emphasized to respondents that the cash endowment was theirs to keep and could be used for any purpose. Third, the enumerators conducted practice auctions with the participants so that they knew that real money was at stake. These actions emphasized to participants that rational decision-making was in their best interest. Our results also serve to minimize concern over these potential sources of bias: we find heterogeneous effects of the information across treatments, inputs, and by gender. If experimenter demand bias was a significant problem, we would expect to see stronger results across the board in the direction of the recommendations. Finally, we also control for enumerator fixed effects in the regressions, neutralizing idiosyncratic impacts on auction results arising from the behavior of particular enumerators.

In addition to the experimental results above, our cost-benefit analysis shows the potential of plot-specific soil tests to significantly enhance the well-being of farm households in western Kenya: information about the prevalence of soil nutrient deficiencies that curb the productivity of crops suggests the importance of making further investments in fertilizer inputs and directs the allocation of scarce resources toward more profitable uses. The results in this study suggest that widespread soil testing may be an effective way of increasing agricultural input optimization among farmers. Subsidization is one vehicle that might be employed, but this is costly to governments, generates distortions in the market, and can cause farmers to over-allocate their resources towards ineffective or deleterious inputs. It appears from this study that testing soils, whether by SoilDoc (used in this research) or some other soil testing technology, has significant potential to pay for itself in terms of the additional revenues that accrue to farmers from improved crop productivity relative to the cost of the tool. More research is needed to gain a more detailed understanding of the costs versus benefits of soil testing programs when scaled up to reach farmers at a much wider scale.

These results also suggest that development programs should increase their targeting of women to increase access to both inorganic and organic agricultural inputs.

While many studies and development projects have focused on the former, there has been insufficient attention placed on organic input access given the importance of complementary usage of both inorganic and organic inputs for soil nutrient recovery. One potential example is livestock access. Because it appears that men tend to allocate their household's resources, including organic resources like animal manure, for use on their own plots, women do not have the same access to organic resources necessary for improved soil as do men. Programs that focus on increasing ownership of livestock among women will help to give women access to animal manure, as well as animals, and thus enhance soil health and crop yields on plots that they manage.

Extensive rural poverty and food insecurity has remained a persistent problem in SSA and much of the developing world. A major cause is soil and environmental degradation and the resulting low crop yields that prevent accumulation of assets. Farmer optimization of agricultural input use can improve soil health and move farmers out of a resource poverty trap. Organic agricultural inputs are generally underused in SSA yet have particular potential for improving crop yields, especially in areas with highly carbon-degraded soils. However, it is difficult for farmers to gain accurate measures of their soil nutrient levels in order to determine an optimal match with agricultural input levels. Coupled with liquidity constraints, uncertainty regarding appropriate input use for a particular farmer's soil nutrients and soil type limits the adoption and intensity of use of often necessary soil amendments. Soil testing may therefore be a key tool in optimizing farmers' agricultural input choices, reversing soil degradation, and improving the ecology and farmer livelihoods in rural SSA.



Note: Marginal values for treated individuals. Post (No Rec): Those who did not receive a recommendation to use DAP; Post (Rec): Those who received a recommendation to use DAP. Horizontal red dotted line is the market price for 1KG of DAP. Vertical red lines show 95% confidence intervals.

Figure 3.1: Predicted WTP for 1KG Diammonium Phosphate

Table 3.1: Soil fertility in western Kenya[a]

Indicator	Threshold Value[b]	Farms below threshold (%) (N=559[c])
Nitrate-N	20mg NO_3 per kg soil	71.7
Phosphate-P	0.5mg PO_4^{-3} per kg soil	84.3
Potassium-K	30mg K per kg soil	37.8
Sulfate-S	10mg SO_4^{-2} per kg soil	26.7
Active Carbon	350mg Active C per kg soil	33.3
pH	5.5	18.4

Note: [a]Kakamega, Bungoma, and Busia counties. See Appendix A1 for a map of sample village locations. [b]Thresholds developed by Weill and Palm. [c]The number of households in this sample greater than elsewhere in paper due to sample attrition between soil sample collection and household interviews.

Table 3.2: Summary statistics

Variable	Mean	Std. Dev.	Min	Max
Individual (n=884)				
Age	48.29	16.09	19.00	109.00[a]
Years of Education	7.95	3.80	0.00	26.00[b]
Yes=1:				
Basic math ability[c]	0.56	0.50	0.00	1.00
Female	0.58	0.49	0.00	1.00
Widow/er	0.14	0.35	0.00	1.00
Primary occupation is farmer	0.88	0.33	0.00	1.00
Anglican	0.29	0.45	0.00	1.00
Catholic	0.17	0.37	0.00	1.00
Pentecostal	0.42	0.49	0.00	1.00
Bukusu subtribe	0.37	0.48	0.00	1.00
Luhya tribe (except Bukusu)	0.31	0.46	0.00	1.00
Iteso tribe	0.29	0.45	0.00	1.00
Household (n=548)				
Household size[d]	5.29	3.27	0.00	40.00
Total farm area (excluding homestead)	1.06	1.06	0.02	8.87
Household spending on food per week (KSh)[e]	1228.52	1773.12	0.00	21000.00
Yes=1:				
Household head is male	0.55	0.50	0.00	1.00
Organic inputs (within past two seasons)	0.45	0.50	0.00	1.00
Inorganic inputs (within past two seasons)	0.88	0.33	0.00	1.00
No inputs (within past two seasons)	0.07	0.25	0.00	1.00
NGO contact	0.13	0.34	0.00	1.00
River as water source	0.43	0.50	0.00	1.00
Electricity (grid)	0.13	0.33	0.00	1.00
Solar panels	0.29	0.45	0.00	1.00
Metal roof	0.87	0.33	0.00	1.00
Mud walls	0.77	0.42	0.00	1.00
Earth/Mud floor	0.72	0.45	0.00	1.00
Polygamous household	0.10	0.30	0.00	1.00
Own cow(s)	0.37	0.48	0.00	1.00
Village (n=17)				
Individuals (interviewed per village)	52.00	13.91	38.00	97.00
Households (sampled per village)	32.24	8.39	21.00	57.00

Note: [a]There was one woman who claimed she was 109 years old. [b]The sample included a couple of individuals who were university professors and had PhDs. [c]Was able to do a basic multiplication problem. [d]Defined as the number of individuals who spent the night at that dwelling last night. [e]1 USD was approximately equal to 102 KSh at the time of the survey.

Table 3.3: Sample size by group[a]

Treatment	Women	Men	Total
T1: IR	137	96	233
T2: VC	117	101	218
T3: IR & VC	128	77	205
Control	129	99	228
Total	511	373	884

Note: IR: Input recommendation; VC: Village comparison; IR & VC: Input recommendation and village comparison.
[a]Uneven distribution among treatments due to random assignment by tablet computer at time of auction.

Table 3.4: Difference-in-differences results: Treatment 1 – Input recommendation

		DAP[a]		**Organic inputs**[b]	
		I	II	III	IV
"Use nitrogen input" recom. × treatment × time	Treatment effect	61.970***	49.562***		
	SE	10.015	14.711		
	Cluster p-value	0.000	0.004		
	WB p-value	0.000	0.000		
"Use nitrogen input" recom.× treatment × time × female	Treatment effect		22.040		
	SE		21.440		
	Cluster p-value		0.319		
	WB p-value		0.326		
"Use organic input" recom.× treatment × time	Treatment effect			17.770**	1.013
	SE			8.286	14.155
	Cluster p-value			0.048	0.944
	WB p-value			0.058	0.910
"Use organic input" recom. × treatment × time × female	Treatment effect				28.491*
	SE				15.123
	Cluster p-value				0.078
	WB p-value				0.110
Household & demographic controls		Yes	Yes	Yes	Yes
Fixed effects (Village, month, enumerator, input)		Yes	Yes	Yes	Yes
N[c]		1824	1824	9010	9010
R^2		0.659	0.659	0.496	0.496

Note: [a]Diammonium Phosphate, an inorganic fertilizer. [b]Includes several products including biochar, vermicompost, farmyard manure, and combinations of these products. [c]Differing N between estimations due to the larger number of organic products auctioned than inorganic. Additional difference-in-differences regressors included in estimation and are available upon request. Standard errors (SE) are clustered at the village level. Wild bootstrap (WB) p-values are included to correct for small sample size of clusters (17) with 1000 repetitions. ***p<0.01, **p<0.05, *p<0.1.

Table 3.5: Difference-in-differences results: Treatment 2 – Village comparison

		All inputs		**DAP**[a]		**Organic inputs**[b]	
		I	II	III	IV	V	VI
Avg. soil nut. quintile>3.0 × treatment × time	Treatment effect	-13.233**	-22.190***	1.656	-5.921	-16.229**	-25.523***
	SE	5.384	7.246	6.802	9.360	5.612	8.314
	Cluster p-value	0.026	0.007	0.811	0.536	0.011	0.008
	WB p-value	0.014	0.012	0.766	0.540	0.008	0.012
Avg. soil nut. quintile>3.0 × treatment × time × fem.	Treatment effect		16.438		15.539		16.667
	SE		10.476		17.170		11.491
	Cluster p-value		0.136		0.382		0.166
	WB p-value		0.146		0.376		0.192
Household & demographic controls		Yes	Yes	Yes	Yes	Yes	Yes
Fixed effects (Village, enumerator, month, input)		Yes	Yes	Yes	Yes	Yes	Yes
N[c]		10511	10511	1766	1766	8745	8745
R^2		0.509	0.510	0.656	0.658	0.475	0.475

Note: Average quintile: Designates whether individual had an average soil quality level above the 3.0 quintile level in his/her village. [a]Diammonium Phosphate, an inorganic fertilizer. [b]Includes several products including biochar, vermicompost, farmyard manure, and combinations of these products. [c]Differing N between estimations due to the larger number of organic products auctioned than inorganic. Additional difference-in-differences regressors included in estimation and are available upon request. Standard errors (SE) are clustered at the village level. Wild bootstrap (WB) p-values are included to correct for small sample size of clusters (17) with 1000 repetitions. ***p<0.01, **p<0.05, *p<0.1.

Table 3.6: Difference-in-differences results: Treatment 3 – Input recommendation & village comparison

		All inputs		**DAP**[a]		**Organic inputs**[b]	
		I	II	III	IV	V	VI
Avg. soil nut. quintile>3.0 × treatment × time	Treatment effect	-2.940	-5.371	-2.810	1.875	-2.952	-6.830
	SE	4.829	9.581	11.022	19.037	5.048	9.906
	Cluster p-value	0.551	0.583	0.802	0.923	0.567	0.500
	WB p-value	0.524	0.604	0.610	0.778	0.540	0.510
Avg. soil nut. quintile>3.0 × treatment × time × fem.	Treatment effect		5.386		-6.637		7,815
	SE		13.251		24.615		15.399
	Cluster p-value		0.690		0.791		0.619
	WB p-value		0.706		0.842		0.630
Household & demographic controls		Yes	Yes	Yes	Yes	Yes	Yes
Fixed effects (Village, enumerator, month, input)		Yes	Yes	Yes	Yes	Yes	Yes
N^c		10229	10229	1722	1722	8507	8507
R^2		0.507	0.510	0.634	0.639	0.477	0.480

Note: Average quintile: Designates whether individual had an average soil quality level above the 3.0 quintile level in his/her village. [a]Diammonium Phosphate, an inorganic fertilizer. [b]Includes several products including biochar, vermicompost, farmyard manure, and combinations of these products. [c]Differing N between estimations due to the larger number of organic products auctioned than inorganic. Additional difference-in-differences regressors included in estimation and are available upon request. Standard errors (SE) are clustered at the village level. Wild bootstrap (WB) p-values are included to correct for small sample size of clusters (17) with 1000 repetitions. ***p<0.01, **p<0.05, *p<0.1.

Table 3.7: Difference-in-differences results: Treatment 3 – Input recommendation & village comparison

		DAP[a]		**Organic inputs**[b]	
		I	II	III	IV
"Use nitrogen input" recom. × treatment × time	Treatment effect	44.624***	21.887		
	SE	8.431	17.356		
	Cluster p-value	0.000	0.225		
	WB p-value	0.000	0.284		
"Use nitrogen input" recom. × treatment × time × female	Treatment effect		37.769		
	SE		28.165		
	Cluster p-value		0.199		
	WB p-value		0.288		
"Use organic input" recom. × treatment × time	Treatment effect			10.656*	1.239
	SE			6.093	13.190
	Cluster p-value			0.099	0.926
	WB p-value			0.180	1.000
"Use organic input" recom. × treatment × time × female	Treatment effect				15.545
	SE				13.901
	Cluster p-value				0.280
	WB p-value				0.296
Household & demographic controls		Yes	Yes	Yes	Yes
Fixed effects (Village, enumerator, month, input)		Yes	Yes	Yes	Yes
N^c		1714	1714	8467	8467
R^2		0.659	0.661	0.492	0.493

[a]Diammonium Phosphate, an inorganic fertilizer. [b]Includes several products including biochar, vermicompost, farmyard manure, and combinations of these products. [c]Differing N between estimations due to the larger number of organic products auctioned than inorganic. Additional difference-in-differences regressors included in estimation and are available upon request. Standard errors (SE) are clustered at the village level. Wild bootstrap (WB) p-values are included to correct for small sample size of clusters (17) with 1000 repetitions. ***p<0.01, **p<0.05, *p<0.1.

Table 3.8: Cost-benefit analysis (Sc=Scenario)

	Sc 1	Sc 2	Sc 3	Sc 4	Sc 5	Sc 6	Sc 7	Sc 8	Sc 9	Sc 10	Sc 11	Sc 12
Soil test price	500	500	500	500	500	500	1000	1000	1000	1000	1000	1000
DAP price (KSh kg^{-1})	100	100	100	100	100	100	100	100	100	100	100	100
Maize value (KSh kg^{-1})	24.1	24.1	29.1	34.1	34.1	34.1	24.1	24.1	29.1	34.1	34.1	34.1
N - AE (kg (kg N^{-1}))	19	32	25.5	19	19	32	19	32	25.5	19	19	32
Price elasticity DAP	-1.02	-1.2	-1.02	-0.87	-1.2	-1.02	-1.02	-1.2	-1.02	-0.87	-1.2	-1.02
Q_0 (DAP kg ha^{-1})	165.94	165.94	165.94	165.94	165.94	165.94	165.94	165.94	165.94	165.94	165.94	165.94
DAP Prices (KSh kg^{-1})*												
P_0 (Baseline)	56.55	56.55	56.55	56.55	56.55	56.55	56.55	56.55	56.55	56.55	56.55	56.55
P_+ (Group A)	91.20	91.20	91.20	91.20	91.20	91.20	91.20	91.20	91.20	91.20	91.20	91.20
P_- (Group B)	6.79	6.79	6.79	6.79	6.79	6.79	6.79	6.79	6.79	6.79	6.79	6.79
Group A												
ΔQ (DAP kg ha^{-1})	103.71	122.01	103.71	88.46	122.01	103.71	103.71	122.01	103.71	88.46	122.01	103.71
DAP value (KSh)	10371.03	12201.21	10371.03	8845.88	12201.21	10371.03	10371.03	12201.21	10371.03	8845.88	12201.21	10371.03
ΔN (kg N ha^{-1})	18.67	21.96	18.67	15.92	21.96	18.67	18.67	21.96	18.67	15.92	21.96	18.67
ΔMaize (kg ha^{-1})	354.69	702.79	476.03	302.53	417.28	597.37	354.69	702.79	476.03	302.53	417.28	597.37
Maize value (KSh)	8548.01	16937.23	13852.48	10316.24	14229.30	20370.36	8548.01	16937.23	13852.48	10316.24	14229.30	20370.36
Net benefit (KSh ha^{-1})	-2323.02	4236.02	2981.45	970.36	1528.09	9499.33	-2823.02	3736.02	2481.45	470.36	1028.09	8999.33
Group B												
ΔQ (DAP ha^{-1})	148.94	175.22	148.94	127.03	175.22	148.94	148.94	175.22	148.94	127.03	175.22	148.94
DAP value (KSh)	14893.58	17521.86	14893.58	12703.35	17521.86	14893.58	14893.58	17521.86	14893.58	12703.35	17521.86	14893.58
Net benefit	14393.58	17021.86	14393.58	12203.35	17021.86	14393.58	13893.58	16521.86	13893.58	11703.35	16521.86	13893.58

Note: Group A are those who increase DAP use. Group B are those who decrease DAP use. Prices are margins at means calculated from estimations in Section 3.5.1.

Appendix 3.A.1: Proofs of propositions

Proposition 1 $d^k_{it} \neq d^k_{it+1}$ if and only if $\bar{A}^k_{it+1} \neq \bar{A}^k_{it}$ or $\theta^k_{it+1} > \theta^k_{it}$

Proof. We begin with Equation 36, and assume that variables L^k, x^k, c^k, and η^2 do not change between t and $t+1$, as would be the case between two auction rounds in the same sitting, and that θ^k in the first period t is greater than zero. Thus we have:

$$d^k_{it} = L^k_i \left[P\bar{A}^k_{it}(x^k_i)^{\frac{1}{2}} - c^k x^k_i - R + \gamma + \alpha_i - \frac{1}{2}\varrho P^2 L^k_i x^k_i \left(\frac{1}{\theta^k_{it}} + \eta^2_i \right) \right] \tag{43}$$

The proposition is that $d^k_{it} \neq d^k_{it+1}$, or

$$\begin{aligned} L^k_i & \left[P\bar{A}^k_{it}(x^k_i)^{\frac{1}{2}} - c^k x^k_i - R + \gamma + \alpha_i - \frac{1}{2}\varrho P^2 L^k_i x^k_i \left(\frac{1}{\theta^k_{it}} + \eta^2_i \right) \right] \\ & \neq L^k_i \left[P\bar{A}^k_{it+1}(x^k_i)^{\frac{1}{2}} - c^k x^k_i - R + \gamma + \alpha_i - \frac{1}{2}\varrho P^2 L^k_i x^k_i \left(\frac{1}{\theta^k_{it+1}} + \eta^2_i \right) \right] \end{aligned} \tag{44}$$

Canceling out and rearranging terms:

$$\bar{A}^k_{it} - \bar{A}^k_{it+1} \neq \left(\frac{\varrho P^2 L^k_i (x^k_i)^{\frac{1}{2}} (\theta^k_{it+1} - \theta^k_{it})}{2\theta^k_{it}\theta^k_{it+1}} \right) \tag{45}$$

First, let $\bar{A}^k_{it+1} \neq \bar{A}^k_{it}$ and $\theta^k_{it+1} = \theta^k_{it}$ (θ is nondecreasing so cannot take a negative value at any t). If $\theta^k_{it+1} = \theta^k_{it}$, the the right side of Equation 17 is zero. Since $\bar{A}^k_{it+1} \neq \bar{A}^k_{it}$, the left side of Equation 17 must be nonzero, and the inequality holds.

Next, let $\bar{A}^k_{it+1} = \bar{A}^k_{it}$ and $\theta^k_{it+1} > \theta^k_{it}$. If $\bar{A}^k_{it+1} = \bar{A}^k_{it}$, then the left side of Equation 17 is zero. If $\theta^k_{it+1} > \theta^k_{it}$, then the right side of Equation 17 must be greater than zero, and the inequality holds.

The other direction follows trivially and is not shown here. □

Corollary 1 If we assume that $\theta^k_{it+1} - \theta^k_{it}$ is zero, then $d^k_{it+1} > d^k_{it}$ if and only if $\bar{A}^k_{it+1} > \bar{A}^k_{it}$

Proof. We continue with the same assumption as Proposition 1 but add that $\theta^k_{it+1} - \theta^k_{it}$ is zero. We first show that $d^k_{it+1} > d^k_{it}$ implies $\bar{A}^k_{it+1} > \bar{A}^k_{it}$. Again using Equation 8, we have:

$$\begin{aligned} & L^k_i \left[P\bar{A}^k_{it}(x^k_i)^{\frac{1}{2}} - c^k x^k_i - R + \gamma + \alpha_i - \frac{1}{2}\varrho P^2 L^k_i x^k_i \left(\frac{1}{\theta^k_{it}} + \eta^2_i \right) \right] \\ & < L^k_i \left[P\bar{A}^k_{it+1}(x^k_i)^{\frac{1}{2}} - c^k x^k_i - R + \gamma + \alpha_i - \frac{1}{2}\varrho P^2 L^k_i x^k_i \left(\frac{1}{\theta^k_{it+1}} + \eta^2_i \right) \right] \end{aligned} \tag{46}$$

Again canceling terms and rearranging, we have:

$$\bar{A}^k_{it} - \bar{A}^k_{it+1} < \left(\frac{\varrho P^2 L^k_i (x^k_i)^{\frac{1}{2}} (\theta^k_{it+1} - \theta^k_{it})}{2\theta^k_{it}\theta^k_{it+1}} \right) \tag{47}$$

If $\theta^k_{it+1} - \theta^k_{it}$ is zero, then the right hand side of Equation 19 is zero, and $\bar{A}^k_{it+1} > \bar{A}^k_{it}$.

Showing the other direction is trivial and not shown here.

□

Proposition 2 If $d^k_{it+1} < d^k_{it}$, then $\bar{A}^k_{it+1} < \bar{A}^k_{it}$.

Proof. We assume again that variables L^k, x^k, c^k, and η^2 do not change between t and $t + 1$, and that θ^k at time t is greater than zero and nondecreasing. We thus have:

$$\begin{aligned} & L^k_i \left[P\bar{A}^k_{it}(x^k_i)^{\frac{1}{2}} - c^k x^k_i - R + \gamma + \alpha_i - \frac{1}{2}\varrho P^2 L^k_i x^k_i \left(\frac{1}{\theta^k_{it}} + \eta^2_i \right) \right] \\ & > L^k_i \left[P\bar{A}^k_{it+1}(x^k_i)^{\frac{1}{2}} - c^k x^k_i - R + \gamma + \alpha_i - \frac{1}{2}\varrho P^2 L^k_i x^k_i \left(\frac{1}{\theta^k_{it+1}} + \eta^2_i \right) \right] \end{aligned} \tag{48}$$

After canceling terms, we get the familiar equation below:

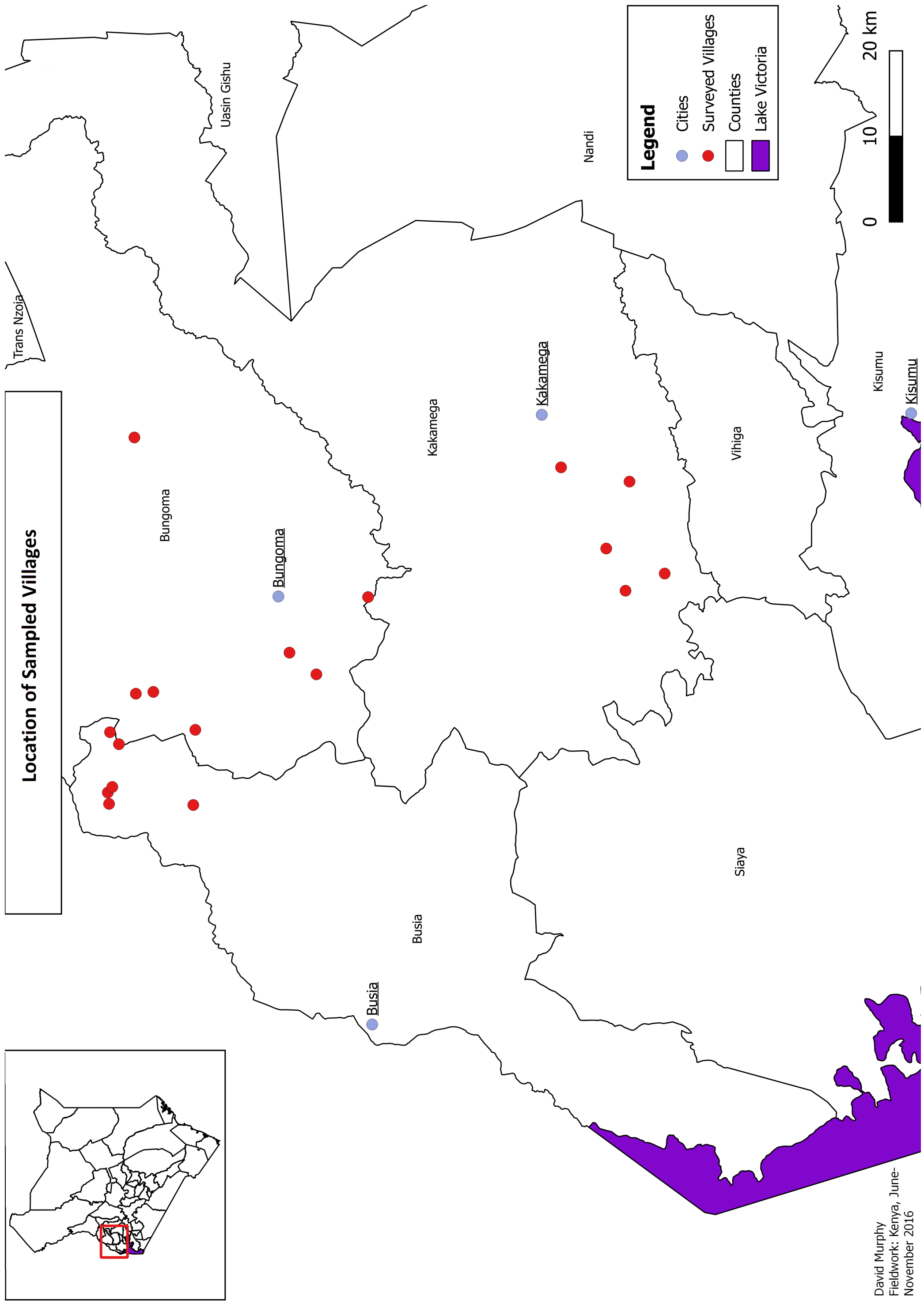
$$\bar{A}^k_{it} - \bar{A}^k_{it+1} > \left(\frac{\varrho P^2 L^k_i (x^k_i)^{\frac{1}{2}} (\theta^k_{it+1} - \theta^k_{it})}{2\theta^k_{it}\theta^k_{it+1}} \right) \tag{49}$$

First assume that $\bar{A}^k_{it+1} = \bar{A}^k_{it}$. Then the left hand side of the equation is zero. As we know, θ^k_{it} is nondecreasing, so $\min(\theta^k_{it+1} - \theta^k_{it})$ is zero. Thus, when we have $\min(\theta^k_{it+1} - \theta^k_{it})$, we get $0 > 0$ and a contradiction.

Next, assume $\bar{A}^k_{it+1} > \bar{A}^k_{it}$. We again take the minimum value of the right hand side, which is zero as θ^k_{it} is nondecreasing. This leads to $\bar{A}^k_{it} > \bar{A}^k_{it+1}$, a contradiction.

□

Appendix 3.A.2: Project area



Appendix 3.A.3: SoilDoc soil testing system information

SoilDoc offers a diagnostic kit and management system for the use in 'plot-specific' analysis of soil properties that are related to fertility and nutrient availability, and uses geo-referencing to generate soil maps. The analysis of N, P, active C, sulfur, and aggregate stability offered by the SoilDoc kit has comparable accuracy as standard laboratories. Other advantages of the SoilDoc system are that 1) it can be deployed anywhere, 2) it offers low cost per test, 3) it has high accuracy, and 4) gives fast diagnosis of the soil fertility status. One trained person can complete full SoilDoc analysis for 40 and 60 samples per week (2000 samples per year). Including the chemicals, other consumables, and replacement of all instruments over a three year period, the cost for one full suite of tests remains under 3 USD with less than 1,000 samples per year, and is close to 2 USD with 2,000 samples per year. Overall, soil analysis by the SoilDoc system costs significantly less than at any research or commercial lab. At the moment, research efforts are being undertaken to validate and improve the fertilizer recommendations for maize crops through meta-analysis of trials. So far, a single threshold value is being used to establish whether fertilizers should be used or not, but more comprehensive models of SoilDoc tests are being developed to give more detailed recommendations.

In this research project, soil samples were taken two to three months before the onset of rains by trained staff, making a composite from five evenly staggered positions in a field where farmers planned to grow maize in the subsequent growing season. The soil was thoroughly mixed by hand and a 250g subsample was carried to the field lab. Before analysis, the soil samples were placed in solar driers for 4 days and manually sieved over 2mm. The pH of soils (a measure of acidity) and electric conductivity (an indicator for the total amount of exchangeable nutrients in soils) were measured in a solution of 10g soil and 20mL water using electrodes. These soil solutions were further subjected to calcium chloride extraction (0.01 mol salt per liter) and then filtered for analysis. Nitrate and potassium in soil extracts were measured with ion-specific electrodes that were calibrated against two standard solutions each day of measurement. Phosphorus and sulfate-S in soil extracts were analyzed through reactions with molybdate and barium chloride that were measured with pocket photo-spectrometers, checked against one standard solution each day. Active C in soils was analyzed through permanganate digestion of 2.5g subsamples, after which supernatant was measured against a three point calibration curve using a pocket photo-spectrometer. Samples were processed in batches of 15 samples per day plus one reference soil that was measured in all batches. All of the analysis were carried out with drinking water from local sources, and blank water corrections were made in calculations of soil nutrient contents for each batch separately. The values of nutrient content were disaggregated in multiple ranges related to soil fertility, whereas fertilizer recommendations were binary, i.e. advising to apply when nutrient levels were below moderate values listed below in table 3.A.3.

Table 3.A.3: SoilDoc nutrient thresholds

Parameter	Extremely Low	Very Low	Low	Moderate (or Opt. pH)	High	Very High
Nitrate-N mg NO_3 per kg soil^{-1}	$x < 4.5$	$4.5 \leq x < 5$	$5 \leq x < 5.5$	$5.5 \leq x < 6$	$6 \leq x < 6.5$	$x \geq 6.5$
Phosphate-P mg PO_4^{-3} per kg soil^{-1}	$x < 10$	$10 \leq x < 20$	$20 \leq x < 40$	$40 \leq x < 60$	$60 \leq x < 80$	$x \geq 80$
Potassium˙K mg K per kg soil^{-1}	$x < 15$	$15 \leq x < 30$	$30 \leq x < 60$	$60 \leq x < 90$	$90 \leq x < 120$	$x \geq 120$
Sulfate-S mg SO_4^{-2} per kg soil^{-1}	$x < 5$	$5 \leq x < 10$	$10 \leq x < 20$	$20 \leq x < 30$	$30 \leq x < 40$	$x \geq 40$
Active C mg per kg soil^{-1}	$x < 200$	$200 \leq x < 350$	$350 \leq x < 500$	$500 \leq x < 650$	$650 \leq x < 800$	$x \geq 800$
pH	$x < 4.5$	$4.5 \leq x < 5$	$5 \leq x < 5.5$	$5.5 \leq x < 6$	$6 \leq x < 6.5$	$x \geq 6.5$

Appendix 3.A.4: Auction scripts

Practice Auction Script We will now play a market game. Here is 70 shillings. This 70 shillings is yours to keep and do as you wish. You can use the money in the game, but you are not required to. We are interested in finding out how much you would pay for several items. We have a vanilla cupcake, chocolate cupcake, some cookies, and a 50 shilling note. We will ask you to tell us the maximum price you are willing to pay for each of these items. After you have told how much you would pay for each item, one item will be selected at random by the computer. A price will then be randomly chosen for that item by the computer. If the price you tell me is higher than the random price, you will pay the random price that was chosen and I will give you the item. If the random price is lower than the maximum you are willing to pay, you will keep all the money I have given you and I will keep the item. Under this procedure, it is in your best interest to tell me exactly the maximum you are willing to pay for each item; no more and no less. If you tell me a price that is higher than the maximum you actually want to pay for an item and it is chosen, you will be required to pay this price if it is randomly chosen. If the price you tell me is lower than the maximum you would pay for an item, then if a good price is chosen by the computer but your price is lower, you will not be allowed to buy the item at the good price even if you want to. Do you understand how this game works?

Baseline Auction Script We will now play the same market game for agricultural inputs. Here is 700 shillings. This 700 shillings is yours to keep and do as you wish. You can use the money in the game, but are not required to. We are now interested in finding out how much you would pay for several agricultural items. The game procedure will be exactly the same as for the food items. Under this procedure, it is in your best interest to tell me exactly the maximum you are willing to pay for each item; no more and no less. If you tell me a price that is higher than the maximum you actually want to pay for an item and it is chosen, you will be required to pay this price if it is randomly chosen. If the price you tell me is lower than the maximum you would pay for an item, then if a good price is chosen by the computer but your price is lower, you will not be allowed to buy the item at the good price even if you want to. Do you understand how this game works?

Second Auction Script Now that you have heard the (soil test results), we will play the same market game again. After this round, either this round or your previous round will be chosen as the binding round. One item will be randomly selected from either this round or the previous round, and a random price will be chosen for it by the computer. It is in your best interest to tell me exactly the maximum you are willing to pay for each item; no more and no less. If you tell me a price that is higher than the maximum you actually want to pay for an item and it is chosen, you will be required to pay this price if it is randomly chosen. If the price you tell me is lower than the maximum you would pay for an item, then if a good price is chosen by the computer but your price is lower, you will not be allowed to buy the item at the good price even if you want to. **[You may leave your bid unchanged from the first round if you desire.]

**Note: As part of another study, this sentence was added to the instructions to about half of the participants (randomized by village).

Agricultural Input Explanation "Biochar" is a type of charcoal that is produced from left-over plant material of field crops on farm like maize cobs and stovers, rice husks and haulms, sugarcane bagasse, coconut shells, and others. If applied to soil at the correct rate, biochar helps to improve crop production by increasing the uptake of fertilizers, manure and water. “Vermicompost” is the end-product of the breakdown of organic matter by an earthworm, also called worm castings. It is compost produced using earthworms. If applied to the soil in the correct rate vermicompost will improve crop production because it contains substantial amounts of nutrients, has a large water holding capacity and enriches the soil with micro-organisms.

Kiswahili

Zoezi la mnada wa nakala Sasa tutacheza mchezo wa soko. Chukua hii shillingi 70. Hii shilling 70 ni yako na unaweza kufanya nalo kile unachotaka. Unaweza tumia pesa hii kwa mchezo huu, na hiyo pia sio lazima. Tungependa kujua ni kiasi gani ya thamani gani utalipia vitu mbali mbali. Tunalo (queen cakes) aina ya vanilla, chakoleti , kuki zingine , na shilling 50. Tutakuuliza utuambie kile bei ya juu zaidi unaweza lipa kununua kila mmoja ya hivi vitu. Baada ya kutuambia kile malipo unaweza lipa kwa kila bidhaa , bidhaa moja itachaguliwa ki nasibu kupitia njia ya tarakilishi . Bei ya bidhaa hiyo vile vile itachaguliwa ki nasibu kupitia njia ya tarakilishi. Kama bei ulichoniambia ni zaidi ya kile bei kilichochaguliwa ki nasibu na tarakilishi, utalipa kile bei kilichochaguliwa ki nasibu na tarakilishi na nitakupatia bidhaa hiyo. Kama bei ilichochaguliwa ki nasibu ni chini zaidi ya ile bei ya juu uliyosema unaweza lipa kununua bidhaa, utabaki na pesa zote nilichokupatia na mimi nitabaki na bidhaa zangu. Katika hii utaratibu ni kwa mvuto yako kuniambia bei ya juu kamili na halisi unaweza lipia kununua kila bidhaa; bila kuweka bei ya juu zaidi au ya chini sana. Ukiniambia hile bei ya juu zaidi ya hile wewe hasa ungependa kuweka kama bei yako ya juu ya kununua bidhaa na bidhaa ichaguliwe ki nasibu, itabidii ulipe hii bei kununua bidhaa hiyo. Kama bei ulichoniambia ni chini zaidi ya bei wewe hasa ungependa kulipa kama kiwango cha juu basi utalipia bidhaa, Kisha kama bei imechaguliwa ki nasibu na bei yako ni chini, hutakubaliwa kununua bidhaa kwa bei nafuu hata ukiwa unahitaji. Je' unaelewa jinsi huu mchezo unachezwa?

Msingi wa mnada wa nakala Sasa tutacheza huu mchezo ya soko tena. Kutumia hii ni shilingi 700. Hii shilingi 700 ni yako ya kuweka na kutumia utakavyo. Unaweza kumia hii pesa kati huu mchezo. Lakini sio lazima. Sasa tungependa kujua ni kwa thamani gani utalipia pembejeo kadhaa. Utaratibu ya mchezo huu utafanana kabisa na ule wa vitu vya kula hapo . Katika huu utaratibu ni kwa mvuto wako kuniambia bei ya juu kamili na halisi unaweza lipia kununua kila bidhaa; bila kuweka bei ya juu zaidi au ya chini sana. Ukiniambia hile bei ya juu zaidi ya ile wewe hasa ungependa kuweka kama bei yako ya juu ya kununua bidhaa na bidhaa mabei wewe hasa ungependa kulipa kama kiwango cha juu basi utalipia bidhaa, kisha ikiwa bei imechaguliwa ki nasibu na bei yako ni chini, hutakubaliwa kununua bidhaa kwa bei nafuu hata ikiwa unahitaji. Je' unaelewa jinsi huu mchezo unachezwa?

Udongo ya Mnada Kwa vile umesikia kuhusu (utafiti ya udongo) Tutacheza michezo ya hawali tena . Baada ya huu msururu, huu msuru au msururu ya hapo hawali itachaguliwa kuwa msururu wa mwisho. Bidhaa moja itachaguliwa ki nasibu kati huu msururu au musururu ya hapo hawali, na kwa njia ya kinasibu bei itachaguliwa na tarakilishi .Katika hii utaratibu ni kwa mvuto wako kuniambia bei ya juu kamili na halisi unaweza lipa kununua kila bidhaa; bila kuweka bei ya juu zaidi au ya chini sana. Ukiniambia hile bei ya juu zaidi ya hile wewe hasa ungependa kuweka kama bei yako ya juu ya kununua bidhaa na bidhaa ichaguliwe kinasibu, itabidii ulipe hii bei kununua bidhaa hiyo. Kama bei ulichoniambia ni chini zaidi ya bei wewe hasa ungependa kulipa kama kiwango cha juu basi utalipia bidhaa, Kisha ikiwa bei imechaguliwa kinasibu na bei yako ni chini, hutakubaliwa kununua bidhaa kwa bei nafuu hata ikiwa unahitaji.

Maelezo ya mbolea ya kilimo "Biochar" "Makaa ya shamba" ni aina ya makaa ambaye inatengenezwa kutoka kwenye mabaki ya mimeya kama msogoro, na vijiti za mahindi, bagasse ya miwa, mabakio ya nazi na zinginezo. Ikimwagwa kwenye udongo shambani kwa kiwango inayo faa, Makaa ya shamba (Biochar) usaidia kuwepo mazao mazuri kwa kuongeza uwepo wa madini, mbolea ya wanyama na maji. "Vermicompost" ni bidhaa inayo totakana na kinyesi ya earthworm (mniambo). Ikimwagwa kwenye udongo shambani kwa kiwango inayo faa mbolea ya vermicompost itaongeza mazao kwa sababu ina madini mingi sana, na pia inashikilia unyevu kwa kiwango kikubwa na vile vile inaoneza vihini bora kwenye udongo.

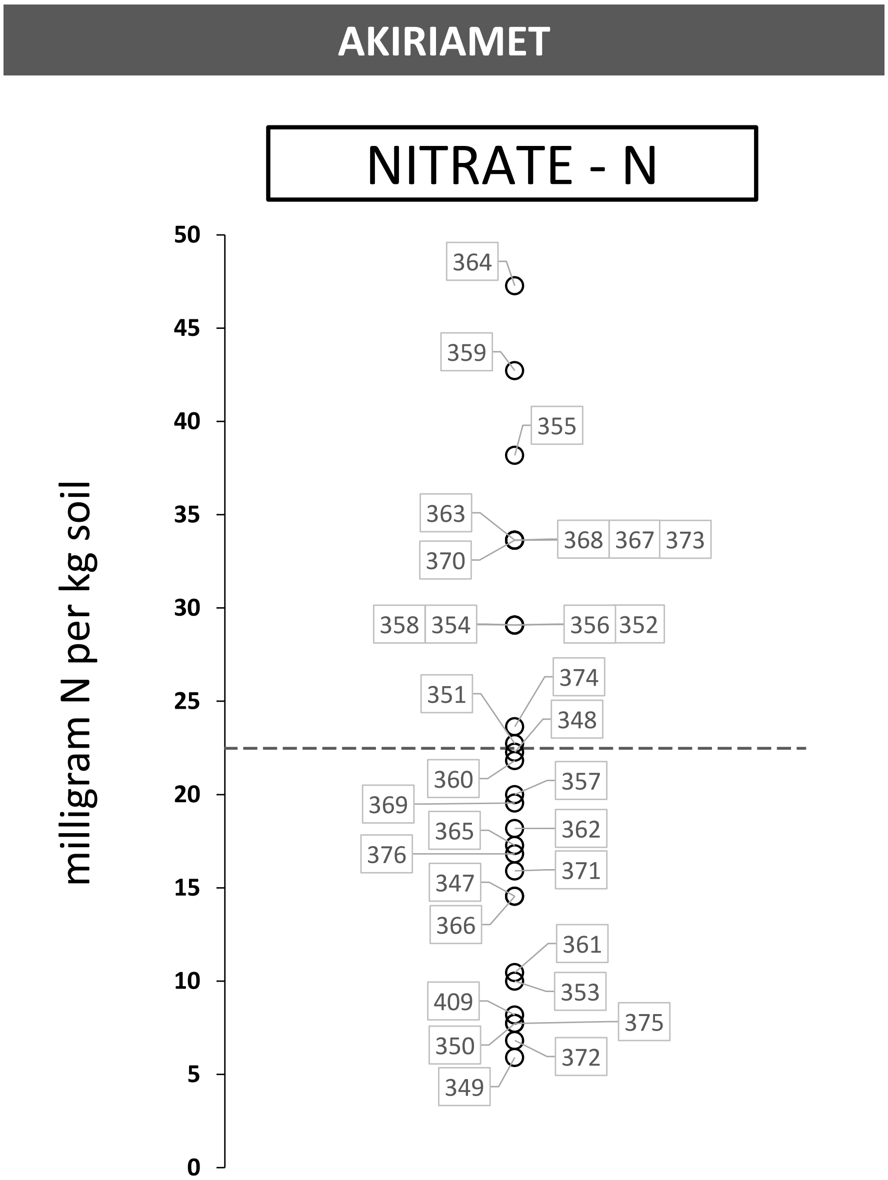
Appendix 3.A.5: Experimental auction supplements

Sample soil test report

SAMPLE CODE	Village	Farmer	Acidity (pH)	Acidity Level	Soil NO3-N (mg NO3-N kg soil-1)	Nitrate Level	Use N input	Soil P (mg P kg soil-1)	P Level	Use P input
Biv 347	Akiriamet	Benard Omusugu	6.78	Optimum	14.55	Very Low	Yes	0.50	Low	Yes

Soil K (mg K kg soil-1)	K Level	Use K input	Soil S (mg kg soil-1)	S Level	Use S input	Active C (mg kg soil-1)	C Level	Use organic input
20	Very Low	Yes	30.25	High	No	389.9	Low	Yes

Sample village comparison chart: Akiriamet village



Appendix 3.A.6: Sample tables

Table 3.A.6-1: Treatment 1 balance table

Variable	Treatment 1	Non-Treatment 1
Female	0.58 (0.49)	0.57 (0.49)
Age	47.75 (16.12)	48.48 (16.09)
Household size	5.52 (3.25)	5.39 (3.10)
Years education	7.76 (3.70)	8.02 (3.83)
Asset index[a]	0.01 (0.93)	0.03 (0.96)
TLU[b]	1.21 (2.38)	1.07 (1.58)
Uses inorganic inputs	0.87 (0.34)	0.89 (0.31)
Uses organic inputs	0.42 (0.49)	0.47 (0.50)
Math ability	0.55 (0.50)	0.56 (0.50)
Widow	0.15 (0.36)	0.14 (0.34)
Usually home	0.98 (0.13)	0.98 (0.14)
NGO contact	0.11 (0.32)	0.14 (0.35)
Total acres	0.92 (0.85)	1.17 (1.13)***
Anglican	0.26 (0.44)	0.30 (0.46)
Catholic	0.19 (0.39)	0.16 (0.36)
Pentecostal	0.43 (0.50)	0.42 (0.49)
Other Christian	0.10 (0.30)	0.11 (0.31)
Other religion	0.02 (0.13)	0.02 (0.13)
Bukusu	0.35 (0.48)	0.38 (0.49)
Other Luhya	0.33 (0.47)	0.30 (0.46)
Iteso	0.29 (0.46)	0.29 (0.45)
Other tribe	0.03 (0.17)	0.03 (0.17)
Enumerator 1	0.23 (0.42)	0.20 (0.40)
Enumerator 2	0.30 (0.46)	0.29 (0.46)
Enumerator 3	0.23 (0.42)	0.30 (0.46)**
Enumerator 4	0.18 (0.39)	0.17 (0.38)
Enumerator 5	0.06 (0.24)	0.04 (0.18)

Note: Standard deviations located next to respective means. [a] Asset index compiled through factor analysis after Sahn and Stifel (2003). [b] Tropical Livestock Units. Difference between means T-test statistical significance: ***p<0.01, **p<0.05, *p<0.1.

Table 3.A.6-2: Treatment 2 balance table

Variable	Treatment 2	Non-Treatment 2
Female	0.54 (0.50)	0.59 (0.49)
Age	49.04 (15.82)	48.04 (16.19)
Household size	5.20 (3.64)	5.49 (2.96)
Years education	8.30 (3.64)	7.84 (3.84)
Asset index†	0.12 (0.98)	-0.01 (0.94)*
TLU	1.24 (1.74)	1.07 (1.85)
Inorganic inputs	0.90 (0.30)	0.88 (0.32)
Organic inputs	0.47 (0.50)	0.45 (0.50)
Math	0.59 (0.49)	0.55 (0.50)
Widow	0.15 (0.36)	0.14 (0.34)
Usually home	0.97 (0.16)	0.98 (0.13)
NGO contact	0.14 (0.35)	0.13 (0.34)
Total acres	1.15 (0.97)	1.09 (1.10)
Anglican	0.31 (0.46)	0.28 (0.45)
Catholic	0.17 (0.38)	0.16 (0.37)
Pentecostal	0.38 (0.49)	0.44 (0.50)
Other Christian	0.11 (0.31)	0.11 (0.31)
Other religion	0.03 (0.16)	0.01 (0.12)
Bukusu	0.34 (0.48)	0.38 (0.49)
Other Luhya	0.34 (0.48)	0.30 (0.46)
Iteso	0.29 (0.45)	0.29 (0.45)
Other tribe	0.02 (0.15)	0.03 (0.17)
Enumerator 1	0.23 (0.42)	0.20 (0.40)
Enumerator 2	0.29 (0.45)	0.30 (0.46)
Enumerator 3	0.29 (0.46)	0.27 (0.45)
Enumerator 4	0.16 (0.36)	0.18 (0.38)
Enumerator 5	0.03 (0.18)	0.05 (0.21)

Note: Standard deviations located next to respective means. [a] Asset index compiled through factor analysis after Sahn and Stifel (2003). [b] Tropical Livestock Units. Difference between means T-test statistical significance: ***p<0.01, **p<0.05, *p<0.1.

Table 3.A.6-3: Treatment 3 balance table

Variable	Treatment 3	Non-Treatment 3
Female	0.62 (0.49)	0.57 (0.50)
Age	48.79 (16.32)	48.14 (16.03)
Household size	5.36 (2.37)	5.44 (3.34)
Years education	7.89 (4.08)	7.97 (3.71)
Asset index[a]	0.01 (0.96)	0.03 (0.95)
TLU[b]	1.07 (1.55)	1.12 (1.90)
Inorganic inputs	0.90 (0.30)	0.88 (0.32)
Organic inputs	0.48 (0.50)	0.45 (0.50)
Math	0.55 (0.50)	0.56 (0.50)
Widow	0.14 (0.35)	0.14 (0.35)
Usually home	0.99 (0.12)	0.98 (0.14)
NGO contact	0.16 (0.36)	0.13 (0.34)
Total acres	1.07 (1.04)	1.12 (1.08)
Anglican	0.27 (0.45)	0.29 (0.46)
Catholic	0.16 (0.37)	0.17 (0.37)
Pentecostal	0.45 (0.50)	0.42 (0.49)
Other Christian	0.11 (0.31)	0.11 (0.31)
Other religion	0.01 (0.10)	0.02 (0.14)
Bukusu	0.39 (0.49)	0.37 (0.48)
Other Luhya	0.27 (0.44)	0.32 (0.47)
Iteso	0.31 (0.46)	0.28 (0.45)
Other tribe	0.03 (0.18)	0.03 (0.17)
Enumerator 1	0.18 (0.38)	0.22 (0.42)
Enumerator 2	0.30 (0.46)	0.29 (0.46)
Enumerator 3	0.31 (0.46)	0.27 (0.44)
Enumerator 4	0.18 (0.39)	0.17 (0.38)
Enumerator 5	0.03 (0.18)	0.04 (0.21)

Note: Standard deviations located next to respective means. [a] Asset index compiled through factor analysis after Sahn and Stifel (2003). [b] Tropical Livestock Units. Difference between means T-test statistical significance: ***p<0.01, **p<0.05, *p<0.1.

Table 3.A.6-4: Control balance table

Variable	Control	Non-Control
Female	0.57 (0.50)	0.58 (0.49)
Age	47.67 (16.18)	48.50 (16.07)
Household size	5.60 (3.13)	5.36 (3.15)
Years education	7.88 (3.79)	7.98 (3.81)
Asset index[a]	-0.05 (0.94)	0.05 (0.96)
TLU[b]	0.92 (1.44)	1.18 (1.94)*
Inorganic inputs	0.88 (0.33)	0.89 (0.31)
Organic inputs	0.46 (0.50)	0.46 (0.50)
Math	0.54 (0.50)	0.57 (0.50)
Widow	0.11 (0.32)	0.15 (0.36)
Usually home	0.98 (0.13)	0.98 (0.14)
NGO contact	0.14 (0.34)	0.14 (0.34)
Total acres	1.29 (1.32)	1.04 (0.96)***
Anglican	0.31 (0.46)	0.28 (0.45)
Catholic	0.14 (0.34)	0.18 (0.38)
Pentecostal	0.43 (0.50)	0.42 (0.49)
Other Christian	0.11 (0.31)	0.11 (0.31)
Other religion	0.01 (0.11)	0.02 (0.13)
Bukusu	0.41 (0.49)	0.36 (0.48)
Other Luhya	0.29 (0.46)	0.32 (0.47)
Iteso	0.27 (0.44)	0.30 (0.46)
Other tribe	0.03 (0.17)	0.03 (0.17)
Enumerator 1	0.21 (0.41)	0.21 (0.41)
Enumerator 2	0.29 (0.45)	0.30 (0.46)
Enumerator 3	0.29 (0.45)	0.27 (0.45)
Enumerator 4	0.18 (0.38)	0.17 (0.38)
Enumerator 5	0.04 (0.20)	0.04 (0.20)

Note: Standard deviations located next to respective means. [a] Asset index compiled through factor analysis after Sahn and Stifel (2003). [b] Tropical Livestock Units. Difference between means T-test statistical significance: ***p<0.01, **p<0.05, *p<0.1.

Appendix 3.A.7: Parallel paths assumption

Identification based on difference-in-differences regressions relies on the assumption that the two groups being compared would have the same general trajectory over time in the absence of an intervention. We believe that because the treatment and control groups were randomly assigned at the auction and the groups are generally well-balanced, we can assume that the parallel paths trend applies to these groups. Here, we analyze whether we can make the same assumption for those who received different information. For example, in Treatment 1, we need to demonstrate that in the absence of information, those who received recommendations to use organic inputs would have on average changed their bids in the same manner as those who did not receive a recommendation to use these inputs. Because the recommendations are based on the soil nutrient levels of an individual's farm, it is possible (though seemingly unlikely), that they may behave in a fundamentally different way based on their own soil characteristics. We can easily test for this by looking at the control group, which did not receive any information treatment. However, we have information about what the recommendation would have been for each individual if they were in the treatment group. We can thus compare the differences between these two groups among those who did not receive a treatment.

In Columns I and III of table 3.A.7-1 we show results of a difference-in-differences estimation among those in the control group, looking at the difference between whether they would have received the input recommendation if they had been randomly assigned to the treatment group. For organic input recommendation, there are no statistically significant differences between the change in bids between the first and second auctions. For the nitrogen input recommendation, there is only very marginal statistical significance at the p=0.1 level. Similarly, in table 3.A.7-2, we do the same estimation but divide the control group by whether they would have seen their position on the village soil charts as above or below that of their peers. Columns I and III in table 3.A.7-2 also show no significant differences. This leads us to conclude that the parallel paths assumption holds for the triple difference estimation.

When we include gender however, we need to show parallel paths for men and women. We look at this difference among those in the control group conditional on whether they would have received the input recommendation for Treatment 1 or seen they had above average soil quality in Treatment 2. These results are in Columns II and IV in tables 3.A.7-1 and 3.A.7-2, and we find no significant differences between men and women for organic input recommendations. For nitrogen input recommendations, there is only a very marginal significance at the p=0.1 level. These results suggest that conditional on the potential information treatment, there are parallel paths between men and women.

Table 3.A.7-1: Difference-in-differences results: Parallel paths checks

		DAP[a]		**Organic inputs**[b]	
		I	II	III	IV
"Use nitrogen input" recom. × time	Treatment effect	-10.037*	2.022		
	SE	5.642	6.359		
	Cluster p-value	0.094	0.755		
"Use nitrogen input" recom. × time × female	Treatment effect		-21.233*		
	SE		11.354		
	Cluster p-value		0.080		
"Use organic input" recom. × time	Treatment effect			-2.804	0.832
	SE			4.927	7.353
	Cluster p-value			0.577	0.911
"Use organic input" recom. × time × female	Treatment effect				-6.904
	SE				9.803
	Cluster p-value				0.491
Household & demographic controls		Yes	Yes	Yes	Yes
Fixed effects (Village, enumerator, month, input)		Yes	Yes	Yes	Yes
N[c]		900	900	4454	4454
R^2		0.515	0.674	0.513	0.674

Note: [a]Diammonium Phosphate, an inorganic fertilizer. [b]Includes several products including biochar, vermicompost, farmyard manure, and combinations of these products. [§]Differing N between estimations due to the larger number of organic products auctioned than inorganic. Additional difference-in-differences regressors included in estimation and available upon request. Standard errors (SE) are clustered at the village level. ***p<0.01, **p<0.05, *p<0.1.

Table 3.A.7-2: Difference-in-Differences Results: Parallel Paths Checks

		DAP[a]		**Organic inputs**[b]	
		I	II	III	IV
Avg. soil nutrient quintile>3.0 × time	Treatment effect	0.147	-4.686	3.018	4.329
	SE	4.957	7.376	3.482	5.911
	Cluster p-value	0.977	0.534	0.399	0.475
Avg. soil nutrient quintile>3.0 × time × female	Treatment effect		8.567		-2.320
	SE		11.021		6.710
	Cluster p-value		0.734		0.448
Household & demographic controls		Yes	Yes	Yes	Yes
Fixed effects (Village, enumerator, month, input)		Yes	Yes	Yes	Yes
N[c]		900	900	4454	4454
R^2		0.670	0.675	0.512	0.514

Note: [a]Diammonium Phosphate, an inorganic fertilizer. [b]Includes several products including biochar, vermicompost, farmyard manure, and combinations of these products. [c]Differing N between estimations due to the larger number of organic products auctioned than inorganic. Additional difference-in-differences regressors included in estimation and available upon request. Standard errors (SE) are clustered at the village level. ***p<0.01, **p<0.05, *p<0.1.

Appendix 3.A.8: Cost-benefit analysis methodology

In this appendix, we describe additional assumptions and methods for the calculations of the cost-benefit analysis of table 3.8. We present these in the order of the parameters in table 3.8.

Soil price As described in Appendix A2, the cost of the soil test (including materials) is 3 USD with less than 1,000 samples per year, and about 2 USD with 2,000 samples per year. We add to this wages for the soil technician, field staff, transportation, lab space, and other associated costs. This increases the price to between 5 and 10 USD per sample (about 500 to 1000 KSh). The price depends greatly on the scale of the testing, with larger quantities of soil testing leading to a lower price per sample.

DAP price (KSh kg^{-1}) We use the average self-reported per kilogram DAP price paid by farmers for 2KG packs of DAP (1 goro-goro) in the sample.

Maize price (KSh kg^{-1}) The average self-reported price of one kilogram of maize in the sample is 29.1 KSh. The price is the amount that the farmer can sell his/her maize at the market. In some scenarios, we also use a lower per kilogram maize price (24.1 KSh) or a higher maize price (34.1 KSh) to illustrate the effect of different average maize prices on the net benefit.

N-Agronomic efficiency (kg (kg N^{-1}) N-AE measures the increase in kilograms of maize produced per hectare for each additional kilogram of nitrogen. We use values from a survey of agronomic trials in SSA by Vanlauwe *et al.* (2011), who find that the mean N-AE value of farmer managed maize plots is 19 (kg (kg N^{-1})) while those who use organic fertilizers as complements with inorganic fertilizers on research managed plots reach a N-AE of 32 (kg (kg N^{-1})). We use values in this range for the various scenarios presented in table 3.8.

Price elasticity of DAP To determine the price elasticity of DAP in the sample, we estimate a 2SLS regression of DAP quantity on the household specific price paid per kilogram with exogenous controls and fixed effects for village, enumerator, and survey month. As an instrument for the price of DAP, we use the average of all (except own) prices in each village (Hausman *et al.*, 1994). The results are not reproduced in full here but are available upon request. We find a demand elasticity of -1.03, with a 95% confidence interval between -0.87 and -1.20. We therefore use these three values in the scenarios present in table 3.8.

Q_0 (DAP kg ha^{-1}) The initial quantity of kilograms of DAP used per hectare is the average value used by respondents in our sample.

DAP prices (KSh kg^{-1}) The initial DAP price (P_0) is the average WTP for one kilogram of DAP from the experimental auctions, 63.36 KSh. The average post-information DAP price for those for whom DAP was recommended (P_+) is 91.74, and for those for whom DAP was not recommended (P_-) is 13.60 KSh. These post-information values are margins predicted by setting all regressors at their means and varying the binary variables for type of information received, treatment group, and auction round.

ΔQ_0 (DAP kg ha^{-1}) We assume that the supply of DAP fertilizer does not change as a result of the change in demand produced by the soil test information. Thus, the predicted change in quantity of DAP used is calculated as $(P_1 - P_0)(\frac{Q_0}{P_0})(-\epsilon_d)$ where P_1 is either P_+ or P_- and ϵ_d is the demand elasticity for DAP.

DAP value (KSh) This is the ΔQ_0 multiplied by the DAP price.

ΔN (DAP kg ha^{-1}) DAP is 18% nitrogen. Thus the change in nitrogen is $(.18)\Delta Q_0$.

ΔMaize (kg ha^{-1}) The change in the quantity of maize produced is ΔN multiplied by the agronomic efficiency (N-AE).

Maize value (KSh) We calculate the value of the maize by multiplying Δmaize by the maize price.

Net benefit (KSh) The net benefit is calculated by taking the maize value, subtracting the DAP value, and subtracting the soil test price. For those in Group B (not recommended DAP), the net benefit is the DAP value (amount of funds saved by not purchasing the DAP) and subtracting the soil test price.

Chapter 4: Social Capital and Gendered Peer Networks

In Sub-Saharan Africa (SSA), soil degradation is a pressing issue. Years of intensive farming with insufficient inputs and ineffective soil practices have led to low soil nutrient levels and decreasing yields per acre (Sanchez, 2002; Tully *et al.*, 2015). Information about improved seeds, land management practices, and optimal input combinations can enable farmers to recover soil nutrient levels (Place *et al.*, 2003) and break the cycle of rural poverty (Barrett and Bevis, 2015). Because of the advanced state of soil degradation in many areas of SSA, information about improved agricultural inputs and practices is particularly valuable. While media and government/NGO extension efforts are important for information diffusion, agricultural information in SSA continues to be primarily spread through peer-to-peer interactions (Conley and Udry, 2010; Krishnan and Patnam, 2014). In rural areas of SSA, peers exchange agricultural information that contribute towards increases in an individual's productivity and economic output. Additionally, an individual's peers often have a "safety credibility," or trustworthiness, that facilitates information exchange (Rogers, 1995). Thus, individuals with higher levels of social capital who are more central in the information network are likely to more rapidly acquire information on important topics such as information related to agricultural practices or inputs.

Social capital, though recognized as a crucial component of economic development, has been defined in numerous ways in the literature (Coleman, 1988; Putnam, 1993, 2000; Burt, 1998, 2000; Dasgupta and Serageldin, 1999; Woolcock and Narayan, 2000; Grootaert and van Bastelaer, 2002a,b; Fafchamps and Minten, 2001; Fafchamps, 2006). In this study, we use a structural definition after Burt (2000) and Fafchamps and Minten (2001), among others, in that social capital represents the competitive advantage accruing to those with favorable locations within a social structure: or alternatively, "better connected people enjoy higher returns" (Burt, 2000). This is shown clearly, for example, in Fafchamps and Minten (2001), who, in three different African countries, show that there is a causal impact among traders on the number of other traders one knows and his/her profits. In general, those centrally located within networks are better able to acquire obligations from others, building stocks of "credit" that can be exchanged for information or favors in the future (Coleman, 1988). Relatedly, the concept of "structural holes" has been described extensively by Burt (1992), who demonstrates that individuals located centrally within networks and who serve as a bridge between sub-networks have a particularly strong competitive advantage in information acquisition and control. Therefore, in our analysis, we use various measures of how central an individual is in his/her village peer network, or "network centrality" (measured in various ways and described in detail in forthcoming sections), as a proxy for social capital.

In developing countries, it has generally been found that men have, on average, higher levels of social capital than women (Katungi *et al.*, 2008; Meinzen-Dick *et al.*, 2014). Potential reasons for this include gender norms that prevent women from participating in certain social activities (Katungi *et al.*, 2008) and the higher opportunity costs of women's time leading to less time available to cultivate bridging social ties (Meinzen-Dick and Zwarteveen, 2003). Also, in many African communities, it is traditional for a husband's wife to be from another village (Luke and Munshi, 2006), causing a reduction, at least initially, in her proximate information links when joining the husband's household after marriage. If this difference in social capital between women and men in fact exists in SSA, it has important implications for development strategies and the need for gender targeting in development programs. However, there have thus far been few rigorous economic studies using data from SSA that demonstrate differences in social capital between men and women in terms of their productivity-enhancing effects.

In our study, we seek to fill this gap in the literature. Using household data collected in 2016-2017, we analyze the social networks of nearly one thousand men and women in four counties of Kenya to determine whether peer network structures in rural villages in SSA point to differential levels of social capital between men and women and, if so, the effects of these differences on agricultural productivity. Using our structural definition of social capital, we construct a theoretical framework after Cowan and Jonard (2004) that illustrates the impact of social capital on bargaining in asymmetrical relationships and its effect on information diffusion. Using several different measures of network centrality, we first show that women are less central in their networks and, as a result, generally have less influence in village networks compared to men. Then, using a linear-in-means empirical model, we find that among men, a one standard deviation increase in their share of female peers corresponds to an increase in per-hectare maize yields by 1.6 percent, controlling for numerous other characteristics. No productivity effect of the gender of peers is found among women in the sample. Our model and empirical analysis suggests that this is due to information bargaining that leads to asymmetrical directional ties between men and women, at least in the Kenyan context, with relatively greater levels of information flowing towards those with more social capital, i.e. men, and less information flowing in the opposite direction.

These findings make significant contributions to the economic literature related to network structure, gender, and agricultural productivity in SSA. While there exists a rich body of work related to social networks in developing countries, with important exceptions (Magnan *et al.*, 2015; Mekonnen *et al.*, 2017), few have focused on differences in social networks related to gender. In this study, we link the literature on gendered social networks with that on gender and social capital, and provide evidence of the impacts of differential levels of social capital on agricultural productivity. The results of this analysis suggest that if improvements in female well-being are the goals of a development project, given existing network structures, it may be insufficient to simply target women. Instead, projects that focus on improving potential social capital among women, such as by targeting the village network structure, may be more effective. In a recent study, Vasilaky and Leonard (2018) show that exogenous creation of new information network links in villages of Uganda increased farm yields for both men and women. Our study shows why structural network reform in villages of SSA can be beneficial for an equitable increase in information diffusion.

4.1 Agriculture, networks, and social capital in SSA

In rural Sub-Saharan Africa, women figure significantly in household agricultural production. Recent studies have shown that the contribution of women to agricultural labor in East Africa is high, ranging from forty to above fifty percent of total labor hours across the region (Food and Agriculture Organization, 2011; Palacios-Lopez *et al.*, 2017). However, numerous studies across SSA have also demonstrated that female-managed plots are less productive than male-managed plots (Peterman *et al.*, 2011; Slavchevska, 2015; Aguilar *et al.*, 2015; Oseni *et al.*, 2015; Kilic *et al.*, 2015; de Brauw, 2015). Yet, as many of these studies reveal, once characteristics such as education, input use, asset ownership, and other resources are included in the estimations, much of the gender gap in productivity disappears.

Why, then, do women suffer from a relative lack of resources that contribute to agricultural productivity? Early work by Udry *et al.* (1995) shows that in Burkina Faso, female-managed plots in a household are allocated fewer resources, such as fertilizer and farmyard manure, than male-managed plots, causing lower productivity. Women in general also suffer from a lack of access to information that leads to lower adoption of improved inputs and agricultural practices compared to men (Doss and Morris, 2001; Ndiritu *et al.*, 2014). In Kenya, women often marry into a village, leading to lower levels of proximate information contacts. Cultural restrictions further limit the speed at which women are able to expand the number of contacts and the strength of their information network. Thus, a less advantageous position within the village social network, corresponding to lower social capital, likely has a direct effect on the gender gap in agricultural productivity observed between men and women in SSA.

A large body of literature has examined the effects of social networks on information diffusion and the adoption of agricultural technologies in SSA (Munshi, 2004; Conley and Udry, 2010; Maertens, 2017; Crane-Droesch, 2017). While a household can obtain information regarding input use and agricultural practices through multiple adult household members, network models in this literature often only account for information flows through a single household head. Magnan *et al.* (2015), however, collected social network data from Uttar Pradesh, India that included information from both the household head and spouse. Using these gender-disaggregated data, they find that male and female networks had little overlap, and, therefore, that households obtain significantly more information through the utilization of both gendered networks. Mekonnen *et al.* (2017) in Ethiopia also collected gender-disaggregated social network data. Within a household, they find that the number of adopters in the female's social network impacts the household decision to adopt an improved agricultural practice more than the number of adopters in the male's network, though the authors do not hypothesize why this is the case. From our analysis however, we may be able to answer this question. If women are more peripheral in village networks and have fewer peers, the marginal impact on the adoption decision of each additional peer among women is likely greater than among men.

Most studies attempting to explicitly measure differential levels of social capital between men and women in developing countries use group membership as a proxy for social capital. They generally find that women belong to less production-oriented groups than men, decreasing their likelihood of agricultural information acquisition (see Meinzen-Dick *et al.* (2014) for a review). Kumar and Quisumbing (2011) find in Bangladesh, however, that if information on an agricultural technology is disseminated through female groups, asset ownership increases relative to men in the sample. Katungi *et al.* (2008) use various measures of social capital including number of strong/weak network links, civic engagement, and membership in social institutions to measure the likelihood of sending or receiving information in rural Uganda. They find that male household heads are more likely to receive information on agricultural technologies than female household heads, and men are more likely to participate in civic engagement or social institutions, which increases their likelihood of exchanging this type of information. Our study further develops this literature by using robust analysis of gender disaggregated network data to analyze levels of social capital and their likely effects on agricultural production.

4.2 A model of information bargaining

In rural villages in SSA, peers exchange information on various subjects that contribute towards increases in an individual's economic output. In our model, we assume that an individual does not share information altruistically, but does so in exchange for a peer's information.[28] To illustrate this, we construct a model after Cowan and

[28]In a more complex model, we could show information shared by one individual with another in time t can garner goodwill for future reciprocal favors from that individual. For simplicity, however,

Jonard (2004) where peers myopically barter information. In our model, however, lower levels of information across specific subjects decrease an individual's bargaining position, leading to unbalanced exchanges of information with his/her peer.

Let i, j be individuals in finite set I^v, which includes all individuals in particular village v. The adjacency matrix representing the network connections of village v is given as $\boldsymbol{G}^v$, where $G^v_{ij} = 1$ if i and j are connected to one another, and 0 otherwise. Each $i \in I$ has information on category $c \in C$, where C is a set of information subjects, where for simplicity, all information and information categories are of the same importance and quality. After Cowan and Jonard (2004), we designate a knowledge vector $o_i = (o_{i,c})$, where $o_{i,c}$ is the information level of i on topic c. For each peer dyad in $\boldsymbol{G}^v$, we denote $n_{ij} = \#\{c : o_{i,c} > o_{j,c}\}$, or the number of topics where i's information level strictly dominates j, and correspondingly $n_{ji} = \#\{c : o_{j,c} > o_{i,c}\}$, which corresponds to the number of categories in which j strictly dominates i. Provided $G^v_{ij} = 1$, i and j will interact and share information if both $n_{ij} > 0$ and $n_{ji} > 0$, or in other words, where both i and j have at least some information that the other does not yet have.[29]

At this point, we depart from Cowan and Jonard (2004), as we show that an individual who has a higher n_{xy} (where $n_{xy} \in \{n_{ij}, n_{ji}\}$) will have more bargaining power. Greater bargaining power will translate into increased information flowing to that individual, increasing his/her economic outcomes compared to those with lower bargaining power. If we assume both i and j have identical utility functions characterized by diminishing marginal utility of information and that all information categories, c, have equal value, then the marginal utility of new information to the individual with higher n_{xy} will be less than that for the individual with lower n_{xy}.

we do not include this scenario in this model.

[29]In this model, we do not account for differences in frequency of communication across peer dyads, though this can also be an important consideration.

For example, let i strictly dominate j in o_1 and o_2, and j strictly dominate i in o_3. Individual i will provide information to j regarding either o_1 or o_2 (assume o_2 in this example), and j will provide information to i regarding o_3. Assuming that i and j both are fully aware of who dominates whom in each category, individual i is willing to trade relatively less information to j about o_2 than i is willing to trade to i about o_3, given their differences in marginal utility of additional information. Parameter α represents the level of information received by a peer, which varies depending on which individual has higher n_{xy}:

$$\alpha = \begin{cases} \frac{1}{\delta} & \text{if } n_{xy} \geq n_{yx} \\ \frac{1}{\delta+\gamma} & \text{if } n_{xy} < n_{yx} \end{cases} \tag{50}$$

where $n_{yx} \in \{n_{ij}, n_{ji}\}, n_{yx} \neq n_{xy}$, $\alpha < 1$, and δ and γ are network level parameters. Under our assumptions, Equation 1 shows that those who have less information across categories will, as a result, obtain relatively less information in *quid pro quo* exchanges of information.[30] After bargaining and the exchange of information, individuals i and j would have the following information vectors:

$$\begin{aligned}
o_{i,1}(t+1) &= o_{i,1}(t) \\
o_{j,1}(t+1) &= o_{j,1}(t) \\
o_{i,2}(t+1) &= o_{i,2}(t) \\
o_{j,2}(t+1) &= o_{j,2}(t) + \frac{1}{\delta+\gamma}[o_{i,2}(t) - o_{j,2}(t)] \\
o_{i,3}(t+1) &= o_{i,3}(t) + \frac{1}{\delta}[o_{j,3}(t) - o_{i,3}(t)] \\
o_{j,3}(t+1) &= o_{j,3}(t)
\end{aligned}$$

[30] As an illustration, in this example, we could assume that $\gamma = (n_{ij} - n_{ji})^2$, such that α decreases as the number of information categories in which j is strictly dominated increases.

Individual j, who is strictly dominated by i in the number of information categories in which s/he has greater information levels, receives a smaller share of the information difference between i and j in information category o_2 compared to what i receives for information category o_3 (as $\frac{1}{\delta+\gamma} < \frac{1}{\delta}$).

In our sample, we hypothesize that men are more central in the network, and on average have higher levels of social capital and information across categories (c). If this is the case, then in our sample we would expect men to have increased productivity when connected to female peers compared to other male peers. An individual needs to give up a lower share of information when exchanging information with another individual whom they strictly dominate in more information categories. If men have information advantages over women in this sample, this provides a benefit to men in their bargaining with women compared to bargaining with other men. If this is the case, then there is an advantage for men to befriend as many women as possible given this favorable bargaining position. In Kenya, however, many women marry into a village, having been born outside of the area. Cultural norms restrict the number of men whom a woman can befriend, limiting the number of new linkages per time period and decreasing the frequency of communication along these links. However, we would expect that as time passes and women build expertise in additional information categories, they will become more central to their network, building their stock of social capital.

4.3 Data

In 2016-2017, in collaboration with the International Institute of Tropical Agriculture (IITA) and the World Agroforestry Centre (ICRAF), we collected data from 991 individuals in 612 rural households, randomly chosen from official rosters of 21 villages in four neighboring counties of Kenya (Bungoma, Busia, Kakamega, and Nandi).[31] This area of Kenya is primarily composed of small-scale farmers with holdings of land less than one hectare. Maize is the dominant food crop, with farmers usually able to harvest crops twice per year. Households are generally poor, and soils highly degraded due to high intensity farming on small plots with insufficient additions of inputs back into the soil through fertilizers and organic material. When visiting these households, enumerators asked the respondents questions related to agricultural production, household assets, demographic information, and other topics, while field technicians took GPS coordinates of the homestead and agricultural plots to provide precise area calculations. Enumerators interviewed both household heads and spouses (if available). Summary statistics of key variables are presented in table 4.1.

Many of the variables in the dataset are characterized by wide ranges (high standard deviations). The average age of the respondent was 48.5 years with an average of 7.9 years of education. In addition, when asked a simple math problem in the survey, 54 percent of respondents could answer correctly. The sample was a majority female, in part due to the large number of men in these areas who have migrated to larger cities for work. A large majority of sample individuals (87 percent) identified farming as their primary occupation. We also see that a majority of houses were not electrified, had a metal roof, and mud walls. Just under 50 percent of the households obtained their drinking water from rivers or lakes. Farms were generally very small in our sample, with the average farm area (excluding the homestead) at 1.04 acres.

In addition to questions related to household, agricultural, and demographic characteristics, our survey also contained a network module. The enumerator asked each individual (both household head and spouse) about 10 to 13 randomly selected indi-

[31]For our analysis, the networks of two villages were combined given their close proximity and historical linkages, giving us effectively twenty distinct villages.

viduals from rosters of his/her village: whether they knew the individual, had met the individual, level of friendship, etc. This strategy used to generate network data, known as "random matching within sample," is useful when time or resource constraints prevent sampling of the entire village peer network (Conley and Udry, 2010). Moreover, Santos and Barrett (2008) show that this strategy can produce results closer to the true population than other forms of network sampling. The effective sampling rate of household heads per village in our study varied: the minimum number of individuals sampled represented 31% of the household heads in the village, and in the most widely sampled village, we surveyed 100% of household heads (the mean effective sampling rate among villages was 54%)[32]. This method produces a dataset of 9,705 peer dyads. In tables 4.2 and 4.3, we include network statistics for these data divided between the gender of the dyad's peers: table 4.2 includes statistics for male$_i$-male$_j$ and male$_i$-female$_j$ dyads, and table 4.3 includes statistics for female$_i$-male$_j$ and female$_i$-female$_j$ dyads. It is striking to see the strong differences that exist in these statistics conditional on the gender of the dyad. For example, i farmers in female$_i$-female$_j$ dyads are much less likely than other gendered dyads to report knowing j. Also, conditional on having met j, individuals i of female$_i$-female$_j$ dyads rate the quality of information from individual j less highly, have known j for less time, are less likely to receive any kind of agricultural advice from j, and trust j less. As also seen in tables 4.2 and 4.3, the directions of many of these statistics are reversed for the other gendered peer dyads, indicating that there are strong connections between links other than female$_i$-female$_j$. While these summary statistics are pronounced, we

[32]We calculate the effective sampling rate in the following way. We received an official village roster of household heads, then randomly selected a percentage of the households to visit. In most villages, a subset of these households no longer existed, having moved or passed away. We then randomly selected replacement household heads from the list. The mean effective sampling rate of 54% represents the effective sampling rates of 18 of 21 villages (three were not able to be computed with the information available)

should keep in mind that these statistics are formulated in the absence of a control for family ties. The intimacy of spousal ties is likely one reason why, without controls, we see significant negative relationships between female$_i$-female$_j$ ties compared to male$_i$-female$_j$ and female$_i$-male$_j$ ties. As marriage in these societies only exists between those of opposite genders, this likely explains why these statistics show closer ties between those of opposite genders compared to same gender dyads (at least in the female$_i$-female$_j$ case). After controlling for same-household relationships, we would likely see a less stark difference.

For our analysis, we construct the full "induced subgraph" of the village network. The induced subgraph is constructed by imputing the outside-sample links between same-village individuals. This provides us with a complete set of peer dyads among those who were sampled in a particular village. Using logit estimation, we estimate the impact of various observable characteristics on the likelihood of linkages between any two same-village peers. After predicting the likelihood of the linkages, we generate binary variables for links between individuals based on a 0.5 likelihood threshold, giving us a full sample of 51,894 dyads. In other words, using our sample data, we estimate whether any two peers in a village are likely to be peers. Robustness checks detailed later demonstrate that these imputed links appear to be well-specified and our results are robust to any minor inaccuracies in the link generation.

We present some examples of village networks in figures 4.1-4.3. The generated graphs automatically place those who are more central within the networks more towards the center of the figure, and less central individuals are on the periphery.[33] Based on the Stata algorithm used, distances of lines and spacing have no literal meaning in these representations. Only the location of the individual (circles) is meaningful and represents who is more central within the village. In these network

[33]We use the *nwplot* function in Stata to construct these graphs.

graphs, men in the village are represented by blue circles, and women by red circles. Same household links are shown using a yellow line, while non-household links have a grey line. Based on these network graphs, in each village it appears that men tend to be located more in the center of the network, and women located more in the periphery. From these visualizations, it appears that women may have lower levels of social capital (using our definition after Burt (2000)) in these networks compared to men. While these graphs show an intuitive relationship, they are insufficient to draw any meaningful conclusions and we seek to more rigorously quantify the relationship between gender and network centrality in the following section.

4.4 Empirical strategy and results

To further explore the relationship between gender and the individual's position with the village peer network, we calculate various measures of network centrality (which we define below), and regress these measures on various demographic characteristics including gender. After this, we estimate the impact of the gender of an individual's peers, among other variables, on own agricultural productivity using a linear-in-means estimation after Manski (1993). The results support our hypothesis that women have lower levels of social capital in rural Kenya, with men, on average, benefiting more through cross-gender ties than do women.

4.4.1 Gender and network centrality

The hypothesis underlying this analysis is that women in villages of rural Kenya are located less centrally in the local social network. To test this, we construct various measures of network centrality using data on peer links within the sampled villages and use these centrality measures as dependent variables in fixed effects estimations.[34] We begin by summarizing the various measures of network centrality that we use in our estimations.

Degree centrality

The simplest measure of centrality, degree centrality, counts the number of links an individual has with other individuals in his/her network. For directed networks, which account for asymmetries in linkages between individuals (as in our data), the "in-degree" measure counts the number of links others have with an individual, while the "out-degree" measure counts the number of links between an individual and his/her network peers.

Closeness centrality

Closeness centrality, first described by Bavelas (1950), measures the centrality of an individual in a particular network by the social distance between each individual i and his/her peer j. For example, if i is connected to j, and j is connected to k, but i and k are not directly connected, then the social distance between i and k is two (two degrees of separation). Closeness centrality is therefore measured in the following way:

$$C_i = \frac{1}{\sum_j \psi_{ij}}$$

where ψ_{ij} is the social distance between the network nodes (individuals) of i and j.

[34] Chandrasekhar and Lewis (2011) show that centrality measures of induced subgraphs lead to biased coefficient estimates when used as independent regressors. However, we then use these variables as dependent variables. Because the individuals in our sampled network were randomly chosen, we assume that measurement error on the network centrality measures deriving from unsampled missing links is uncorrelated with the independent regressors leading to unbiased estimates.

By this measure, those with a higher closeness value in the network are more central.

Eigenvector centrality

As this term suggests, this centrality measure uses the eigenvectors of the network to create a measure that accounts for the importance of individual i's links (Bonacich, 1972, 2007). Using this measure, individuals who are closely connected with popular individuals have a higher centrality than those connected to less popular individuals. This measure can produce very different centrality measures for those who may have the same degree of closeness centrality, since the eigenvector centrality measure for individual i is defined using the following equation as follow:

$$\lambda e_i = \sum_{j=1}^{n} G_{ij} e_j$$

where G_{ij} has a value of 1 if i and j are connected with one another (e.g. are peers) and 0 otherwise, λ is the largest eigenvalue of the adjacency matrix $\boldsymbol{G}$, and e_i is the eigenvalue centrality of individual i. From this equation, we can see that the eigenvalue centrality of an individual is proportional to the total centralities of the peers with whom s/he is connected.

Katz centrality

While degree centrality defines centrality as the number of *direct* connections, Katz centrality (Katz, 1953) counts the number of peers that can be connected through a path beginning or ending at individual i. More distant links have an attenuated effect on Katz centrality, so individuals with few first-degree connections have a lower centrality score. This measure is constructed using the following equation:

$$\kappa_i = \sum_{n=1} \sum_{j=1} \alpha^n (\boldsymbol{G}^n)_{ij}$$

where $\boldsymbol{G}^n$ is the adjacency matrix raised to the n power, and shows whether individuals i and j are connected through $n-1$ intermediaries of i. Parameter α is an attenuation factor, which we set at 0.33 (the results are not meaningfully different with alternative choices for α). Katz centrality can be measured both in outgoing links and incoming links (In-Katz and Out-Katz) in a directed sample.

Betweenness centrality

Betweenness centrality measures the number of times individual i acts as the shortest path between two other peers, j and k (Freeman, 1977). This measure is often effective at measuring the influence of peers, though is especially sensitive to missing links in the network. The equation for betweenness centrality is given as:

$$B_i = \sum_{i \neq j \neq k} \frac{\rho_{ijk}}{\rho_{jk}}$$

where i, j, and k are are individuals in the network, variable ρ_{jk} represents the total number of paths between j and k in the network, and ρ_{ijk} is the total number of paths between j and k that pass through i. The higher the share of paths that pass through i as a share of total paths, the larger the betweenness centrality for individual i.

We include summary statistics of each of the centrality measures of our network sample in table 4.4. By construction, both In-degree and Out-degree centrality and In-Katz and Out-Katz have the same mean values, though the standard deviation of the distributions are slightly different. The Betweenness centrality measure has the largest range and standard deviation, indicating that some individuals serve as important bridges between others in the network, while other individuals do not serve as information bridges. The degree centrality measures have a minimum of zero, indicating no connections, and a maximum of 96, which is the total number of individuals in the largest village sampled.

Using these various measures of centrality, we estimate the following fixed effects regression:

$$\phi_i = \alpha + \beta_i f_i + \beta_2 t_i + \beta_3 (f_i \times t_i) + \sum_{m=1}^{M} X_{im}\gamma_m + \theta + \varepsilon_i \tag{51}$$

where ϕ_i is the network centrality measure of individual i, f_i is a binary vairable representing whether i is female, t_i is the number of years i has lived in the village, θ are fixed effects at the village level, and X are additional variables expected to influence an individual's network centrality including distance from village center, asset index, tropical livestock unit index, mathematical ability, years of education, years of education squared, age, age squared, household size, total number cultivated acres, and dummy variables for tribe. We cluster standard errors at the village level, but the relatively small number of clusters (20) may lead to underestimates of the standard errors. Because of this, we also compute standard errors using the Wild bootstrap method after Cameron *et al.* (2008).

In our sample, women, on average, have lived fewer years in their village of residence than men. This is primarily due to the widespread practice in western Kenya of women marrying and joining the household of the (often older) husband in another village. We include overlapping histograms in figure 4.4 demonstrating the differences in the distribution of years lived in the village by gender. For women, the distribution is skewed right (median 22), while for men, the distribution is more normally distributed (median 43).

Results for Equation 51 are in table 4.5. From these estimations, we find significant correlations between network centrality (ϕ) and female (f), years in the village

(t), and the interaction of these two variables ($f \times t$). The results suggest that network centrality increases along with years living in the village, as one would expect. The variable for female gender is generally negatively correlated with the various measures of network centrality, however we find positive correlations between $f \times t$ and ϕ. Because of this, it is easiest to interpret the results by analyzing the marginal effects. In figures 4.5-4.11, we show the marginal effects of gender on various centrality measures plotted over years in the village. Overall, the figures show that the marginal effects have a similar pattern for each measure of network centrality.[35] For In-degree, Closeness, Eigenvector, and Out-Katz measures of centrality, we find that, for low values of years in the village (roughly between zero and thirty years across network variables), women have significantly lower measures of centrality. Many of the figures show that the slope of the marginal effect line is greater at the lower range of years lived in the village for women, possibly indicating diminishing returns of years lived in the village on centrality. For many of the centrality variables in this range of years, 95% confidence intervals are not overlapping between genders, indicating a significant difference in centrality between men and women.

However, the results show that by the time a woman has lived in the village thirty or forty years, they have "caught up" in network centrality with men. That being said, there is the potential for within-village idiosyncratic omitted variables correlated with measures of both women and network centrality to bias these results, and therefore we do not seek to imply causation with these particular estimations. Instead, together with additional results presented later, these results provide evidence that women have lower levels of social capital then men.

[35]From this analysis, we drop individuals who report having lived in the village for more than 60 years, due to the sparseness of those data.

4.4.2 Linear-in-Means estimation of peer effects on agricultural productivity

As discussed earlier, a large body of literature has shown that farmers learn about agricultural technologies and improved practices from their peers, which increases their agricultural productivity. Our model shows, however, that those who have higher levels of social capital or leverage within a social network have greater bargaining power, providing advantages in acquiring information especially when connected with those who have relatively lower levels of social capital. In this section, we seek to determine the effects of one's peers on agricultural productivity. Since we have established that women, on average, have lower levels of social capital than men, we hypothesize that among men, connections with women increase men's agricultural productivity in our Kenyan sample, though not necessarily in the other direction. As the share of women increase in a man's peer network, he is able to take advantage of social capital asymmetries, as female farmers on-average must trade relatively more of their information to men to obtain their information in return.

We explore this hypothesis using a linear-in-means empirical model after Manski (1993) to analyze the effect of an individual's peers on his/her agricultural productivity. This model divides peer effects into three distinct effects: effects caused by traits and characteristics of peers (exogenous or contextual effects), the environment in which the peers live (correlated effects), and the influence of the behavior of peers on the behavior of other peers (endogenous effects). A key issue in identifying this model is to solve the "reflection problem", or the simultaneity bias that exists between the behavior of an individual and his/her peers. While j's behavior may influence i's, i's behavior may also influence j's. To identify this relationship, we use an instrumental variable strategy discussed further below.

As a measure of productivity in the linear-in-means estimation, we use the per-hectare maize yields over two cropping seasons (long rains and short rains) of the primary maize plot of each household. Along with other household and demographic characteristics, we include a variable indicating the gender of the primary maize plot manager (in our sample, 42% of these were women). We constructed the adjacency matrix (i.e. matrix of peer connections) based on data collected through a random matching within sample method, discussed earlier. Given that we are analyzing the primary maize plot of each household, for the following analysis, we restrict the peer samples to the primary plot manager in each household.

We present the basic linear-in-means model as:

$$y_i = \alpha + \frac{1}{n_i}\sum_{j\in I_i} y_j\beta + \sum_{m=1}^{M} x_{mi}\gamma_m + \frac{1}{n_i}\sum_{m=1}^{M}\sum_{j\in I_i} x_{mj}\delta_m + \theta + \eta_v + u_i \tag{52}$$

where y represents the per-hectare yield of the farmer's primary maize plot, the x are household, demographic, soil, and farm management variables, and θ are enumerator and survey month fixed effects. The coefficient β represents the average endogenous effect of i's peers' maize productivity on i's maize productivity, which could be caused by i learning improved agricultural practices from j. Coefficients δ_m are the average contextual or exogenous effects of the characteristics of i's peers on his/her maize productivity – for example – the average of effect of j's education on i's per-hectare maize yields. The coefficient γ is the effect of own characteristics on the farmer's own agricultural productivity, and η_v shows the correlated, or common, village-level effects. By convention and to ensure that the resulting adjacency matrix is well-defined, we row standardize the adjacency matrix, and divide by the sum of social connections of i in his/her peer set, n_i.

As Manski (1993) discusses, a key challenge to this estimation is solving the "reflection problem," or the simultaneity that exists between Y_i and Y_j. Most researchers use instruments to identify this relationship, using, for example, partially overlapping peer groups (i is friends with j, who is friends with k, who is not necessarily friends with i) (Bramoullé *et al.*, 2009; Giorgi *et al.*, 2010). However, because we are using a random sub-sample of village networks (which is itself endogenously determined), we cannot use this strategy (Chandrasekhar and Lewis, 2011). Instead, we find that homestead altitude, which was gathered at the time of the survey, has a strong correlation with agricultural productivity after controlling for village fixed effects. Maize yields have been shown to be especially sensitive to altitude in Kenya, with higher altitudes leading to significantly more grains per plant (Cooper, 1979). With altitude varying significantly within many villages in the sample and most farm locations exogenously determined through inheritance, this variable appears to be an optimal instrument.

We also incorporate a distance decay term to better account for negative correlation between physical distance and the impact of effects from one's peers. Physically closer peers, all else equal, will likely more strongly influence an individual, given the increased frequency of interaction, more similar farm characteristics, etc. If we were to neglect distance in this estimation, the bias may be compounded by our row-standardization. For example, each peer will affect an individual's outcomes more strongly in the case of isolated individuals with few, physically distant peers than for an individual with many, physically close peers. More than likely, however, the impact of the peers on the former, isolated individual will be less than those of the latter individual. Following Bell and Bockstael (2000), we take account of this in our estimation in two ways. First, we include a distance decay term to the empirical model that places a greater weight on physically closer individuals. Second, we estimate the model several times, each time using a different distance cut-off c for peer linkages: our first estimation has no cut-off, our second limits an individual's peers to those within one kilometer, and our third uses a cut-off of one third of a kilometer. This method should more accurately reflect the relatively greater impact of physically close peers compared to those physically distant.

Using this strategy, we estimate the following model:

$$y_i = \alpha + \left(\frac{1}{n_i}\right) \sum_{j \in I_i} w_j \bar{y}_j \beta + \sum_{m=1}^{M} x_{mi} \gamma_m + \left(\frac{1}{n_i}\right) \sum_{m=1}^{M} \sum_{j \in I_i} w_j x_{mj} \delta_m + \theta + \eta_v + u_i \quad (53)$$

$$w_j = \begin{cases} \left(\frac{d_{ij}}{\sum_{j \in I_i} d_{ij}}\right)^{-1} & \text{if } d_{ij} < c \\ 0 & \text{if } d_{ij} \geq c \end{cases} \quad (54)$$

where $\bar{y}_j$ is the agricultural productivity of y after instrumentation and d_{ij} is the physical distance between i and j. Term w_j becomes zero for an observation if the peer is further than c kilometers. All other variables are as defined previously. First stage regressions using our IV are in Appendix 4.A.1. and detailed results are in tables 4.6 through 4.8.

In table 4.6 we show the results from the 2SLS linear-in-means without a distance cutoff. We find that across specifications in table 4.6, the endogenous effect – the impact of an individual's peers' agricultural productivity on his/her own productivity (the first row of results) – is small and often not statistically significant. We should not be surprised by this: here we are measuring contemporaneous effects, while most likely a time lag is required to account for i learning from the productivity of j. It may take a season for an individual to update their farming practices, for example, after learning about improved practices from his/her peer. Past yields of j then would likely have a more significant impact on current yields of i (Manski, 2000; Mekonnen *et al.*, 2017). Given that our data are essentially a cross-section (one season each of long rains and short rains harvests for each individual), we cannot effectively analyze this lagged effect.

More interesting for our analysis regarding the correlation between social capital and agricultural productivity, however, is the impact of the exogenous effects – specifically, the genders of farmers i and j. Looking first at i's own gender (table 4.6), we see results that are broadly consistent with other literature in the field: the impact of male management on agricultural productivity (compared to female management) is large and positive: female plot management decreases maize yields per hectare by 26% compared to male plot management (Column I), but the magnitude of this impact decreases and loses statistical significance once we add in additional regressors to control for demographic and soil characteristics (Columns II and III, respectively). As in papers such as Slavchevska (2015), who uses a sample from Tanzania and Kilic *et al.* (2015), using data from Malawi, we conclude that the gender gap in agricultural productivity in Kenya is likely primarily a result of differences in education, assets, and resources between male and female farmers.

We believe our most notable results lie in analyzing the correlation between a peer's gender and i's agricultural productivity (table 4.6). Using the full sample, we find that after controlling for other exogenous variables and fixed effects, the binary variable on peers' gender ranges from 0.014 to 0.018 (Columns II through IV). However, this coefficient is only statistically significant at the 5 percent level in Column III. When we split the sample between men and women and estimate each sub-sample separately, stronger effects emerge. Among men in a split sample, we find the coefficients on peers' gender increase in magnitude and gain stronger statistical significance (Columns X to XII). Among women, however, we do not find an effect resulting from the gender of their peer link.

Because we believe that proximity matters in learning and technology adoption, we include a distance cut-off for peer linkages. As we would expect, we find that the magnitude of the gender effect increases further when we use the cut-off. Table 4.7 shows peer results for a 1KM cutoff, and table 4.8 shows these results for a 0.33KM cutoff. Using a 0.33KM cutoff for peer connections, the coefficient on the binary gender variable (female=1) of one's peers increases to around 0.07 in the male sub-sample (Columns X through XII, table 4.8 continued), with a high degree of statistical significance. We continue to see no relationship, however, among women from this binary variable after including the distance cutoffs for peer connections. Defining peers as those who know one another and are within 0.33KM of each other, we can interpret this binary variable as a one standard deviation increase in the female share of a male farmer's links increasing agricultural productivity by 1.6 percent.[36] As can be expected given our long list of controls, the magnitude of this effect is not particularly large. However, the difference between the magnitude and statistical significance of the coefficients on this variable between genders in the sample is striking, suggesting major differences between men and women in relationships with their opposite gender in peer relationships. While we cannot conclusively claim that differences in bargaining power for information are the primary cause, our earlier evidence from our measures of network centrality suggests that this may likely be the case.

Our estimations also reveal other interesting results. Years of own education has a consistent positive effect on agricultural productivity across specifications – likely because those who have more education are likely to be more informed about best agricultural practices. Own household size, on the other hand, has a surprising

[36] The binary variable represents the percent change in agricultural productivity between a (theoretical) man with no female peers, to a (theoretical) man with only female peers. Because, in our sample, 39% of the mean male farmer's links are with female farmers (with a 0.23 standard deviation), this corresponds to a one standard deviation increase in the female share of a male farmer's peers increasing per-hectare maize yields by 1.6 percent.

negative effect (table 4.6, Columns II - IV). Once we split the sample between men and women, however, we see that the effect of household size is only statistically significant with respect to female farm managers (Columns VI - VIII), which may be due to the increased attention needed from these women on household work for the larger household.

In Column IV in table 4.6, we include potentially endogenous variables and the results should therefore be interpreted with caution. These estimations include agricultural input use in log form. However, because we do not want to lose observations that have zero quantities of fertilizers, we follow a strategy after Battese (1997) and Slavchevska (2015) and create a dummy variable for each input, which equals one if that input is not used on that plot. We then transform the log variable of input use per hectare to be $Max(Inp_{ik}, D_{ik})$, where Inp_{ik} is the transformed log variable for fertilizer k usage by household i and D_{ik} is the dummy variable that takes a value of one if no fertilizer k is used by household i, and zero otherwise.

We find that variables such as hired labor, total hours worked per hectare, purchased seeds, and organic fertilizers and DAP used per hectare are all associated with higher maize yields per hectare, as expected, while maize monocrops have lower per-hectare yields compared to intercropping two or more crops. Column IV also reveals that increased levels of phosphorus and potassium in the soil are positively related with maize per-hectare yields. In table 4.6, which includes the effects from i's peers, we see that the number of adult men in a peer's household also has a positive relationship with i's productivity, likely because these adult men can aid in maize planting and harvesting on i's farm.

We can only claim that any of our above estimations are causal, however, if we assume that our network links were formed exogenously conditional on the included observable individual and peer characteristics. Because individuals form peer relationships based on shared characteristics ("homophily") (McPherson *et al.*, 2001), which are not necessarily observed, this assumption may be violated. However, as Patacchini and Venanzoni (2014) demonstrate, when networks are fairly small (as in our case), network (village-level) fixed effects can be an effective way to control for these unobservables, as these unobserved characteristics are likely common to individuals within the network. All of our estimations include village-level fixed effects, and together with our long list of control variables for individual and household characteristics, we believe that we have controlled for potential endogeneity in network formation.

4.5 Robustness checks

Because collecting data on all peer dyads in each village was not feasible, our analysis depends on the induced subgraph of village networks. As described earlier, our network data were collected through a random matching within sample methodology; enumerators asked respondents about 10-13 randomly selected individuals from their own village. We then mapped the induced subgraph by imputing the missing links using these data. This of course means that some of these links could be incorrectly specified. As a robustness check, we conduct an analysis after Liu *et al.* (2013) to determine whether our results are robust to misspecified links. First, we randomly replace links for 0.5 percent of the total peer dyads ten times, and increase this percentage in 0.5 percent increments. In each increment, we conduct our linear-in-means 2SLS estimates, repeating these ten times for each increment, and plot the results. We choose to use a 1KM distance cutoff as a reference point, and focus on the split sample results, specifically the effects of female peer links among men in the sample (corresponding to the third panel of results on table 4.7 (second page)). We present these results in Appendix figure 4.A.1.

Figure 4.A.1 corresponds to the results for gender of peer in Column XI of table 4.7 (second page) (the specification including household and demographic variables and soil characteristics). Vertical bars show the range of the t-statistic estimate over the ten iterations, and the dot corresponds to the mean value for that particular share of randomly replaced links. We can see that as the share of randomly replaced links in the sample increases, the variance of the t-statistic also increases, with increasing numbers of increments including negative values in the range. This indicates that, as we would expect, increasing the number of random links in the sample decreases the statistical significance of the coefficient estimate. Because of the increasing variance in the statistical significance of the coefficients as we add more randomly generated links, after about five percent of the peer links are randomly replaced (corresponding to 2,595 links), we cannot be confident of statistical significance.

4.6 Discussion and conclusion

Women in Sub-Saharan Africa face many economic hurdles compared to men, including lower levels of average education and resources, greater demands on their time stemming from responsibilities within the household, and cultural prohibitions on many social activities. In addition, in the areas of our sample in Kenya, men usually take wives from other villages, and as a result, these women lose their physically proximate network connections. In an extreme case, for the Teso tribe in far western Kenya, we found that men often took wives not only from another village, but another country – sometimes marrying women from the Teso tribe in Uganda. Having lost their close network connections, these women will likely have lower levels of social capital within their village peer network than men.

In our study, we use data collected from nearly one thousand individuals in twenty-one villages spread throughout four counties of Kenya to address two primary questions: 1) are women more peripheral in village-level social networks than men in SSA? and 2) is there evidence of differential economic effects between men and women stemming from their levels of social capital? Our model demonstrates that individuals with higher levels of social capital (proxied in our estimations by measures of network centrality) will have an advantage in bargaining with other members of the network. As a result, they will benefit from connections with individuals having lower levels of social capital. We included a network module within our surveys, and using a "random matching within sample" method, estimate linkages between peers in each village.

To answer the first question above, we use several measures of network centrality and regress these measures on various demographic and household characteristics including years lived in the village and the gender of the respondent, also incorporating village-level fixed effects. We analyze the marginal effects, and find that given fewer years in the village, women have significantly lower measures of centrality than men. They only "catch-up" in network centrality with men after living in the village thirty to forty years. Women, on average, have lived in a village fewer years compared to men: the median value for years lived in a sample village is 22 for women, but 43 for men. Together these results indicate that women have significantly lower levels of network centrality than men. Using the definition of social capital after Burt (2000), which states that social capital is the competitive advantage accruing to those with favorable locations within a social structure, this result presents initial evidence suggesting lower levels of social capital among women in SSA compared to men.

This finding would suggest, as our model predicts, that women have a disadvantage in information acquisition compared to men. In our second set of results, we find, all else equal, that male farmers are more advantaged in networks than their female peers. Using a linear-in-means regression and an instrumental variable strategy to solve the "reflection" problem highlighted by Manski (1993), we find that among male farmers, an increase of one standard deviation in the share of women in their peer set corresponds with a statistically significant increase in the male farmer's agricultural productivity. Women, on the other hand, receive no productivity benefit from particular gender connections. This difference between genders in these estimations suggests that men are able to leverage the knowledge and resources of their female peers to enhance their own agricultural productivity, while women are unable to do the same. Our results demonstrating that women are located more peripherally in their village networks indicate that lower levels of social capital are likely a major cause of this finding, though we cannot conclusively say whether this is the primary or sole mechanism. As a robustness check, we follow Liu *et al.* (2013) and substitute increasing numbers of random links into the network to test whether the results are robust to the likelihood of misspecified links. These results show that substitution of an increasing number of random network links up to about five percent of total links does not materially change our results.

The results suggest that unequal positions within social networks between men and women likely influence their agricultural productivity. Men, who on average have greater levels of social capital, can more effectively take advantage of information within the network, increasing their per-hectare yields compared to women. The implication is that traditional cultural practices, including the taking of wives from other villages, has detrimental economic outcomes for these women. While men tend to be the plot managers of maize crops in households having both a husband and wife, men in these areas tend to have a low life expectancy. When a woman takes over the management of plots after the death of her husband, her social network is less developed than a man of the same age. Moreover, the practice of polygamy, while decreasing in prevalence, still is common in these areas. In these polygamous households, wives often live on separate farms and manage their own plots. Often coming from different villages, they usually do not have their husband living with them full time, causing especially limited networks and low social capital.

There has been an increased emphasis in past and planned development projects focusing on the economic advancement of women (World Bank, 2015). Transmitting information to female-centric networks is a common practice, and has shown positive outcomes among women in many situations (Meinzen-Dick *et al.*, 2014). However, because men on average have higher levels of social capital, given the existing network structure in a village, they often serve as the information brokers for the community. As a result, the most effective policy tool may be to increase the network centrality of women within their social networks. In a recent experiment in Uganda, Vasilaky and Leonard (2018) trained women in agricultural techniques, then paired these trained women with other randomly chosen women in the village. The trained women were encouraged to share the agricultural information that they had acquired with their untrained peers, who they may not have known prior to the experiment. The results show that the creation of these information-sharing links produced positive agricultural outcomes not only for women in that village, but for men also when compared to a control group. Our research in this paper explains one potential mechanism for their finding. The intervention likely increased network centrality for these women, increasing their social capital and enabling more effective information diffusion.

The implication of this research is that a potential low-cost way to decrease the economic disparity between men and women in SSA is to increase peer connections of women. Because women generally begin adulthood with fewer network connections than men, it appears to take a significant number of years for women to reach the same level of network centrality in their village as men. Helping to increase the speed of this process could enable women to more effectively obtain important information from their networks. This, for example, can be done by helping women to form new, information sharing social connections with others in their village.

This study adds to the important discussion about the effects of one's peers found in the literature on gendered social networks and gender and social capital. The results suggest that asymmetry in social capital within peer networks enhances economic outcomes among those with higher social capital, i.e. men, compared to those with lower social capital. An increased focus on increasing the network centrality of women could have impacts on reducing the differential impacts we find within gendered peer linkages. However, unless cultural practices change among large numbers of people in this region, women will continue to disadvantaged by having lower social capital than men. Common practices such as polygamy are becoming less common, and as villages and towns increase in size, it may also become less common for men to take wives from outside their own village. If this social trajectory continues, it may alleviate some of the differences that we find between the social capital of men and women in rural villages of Kenya.

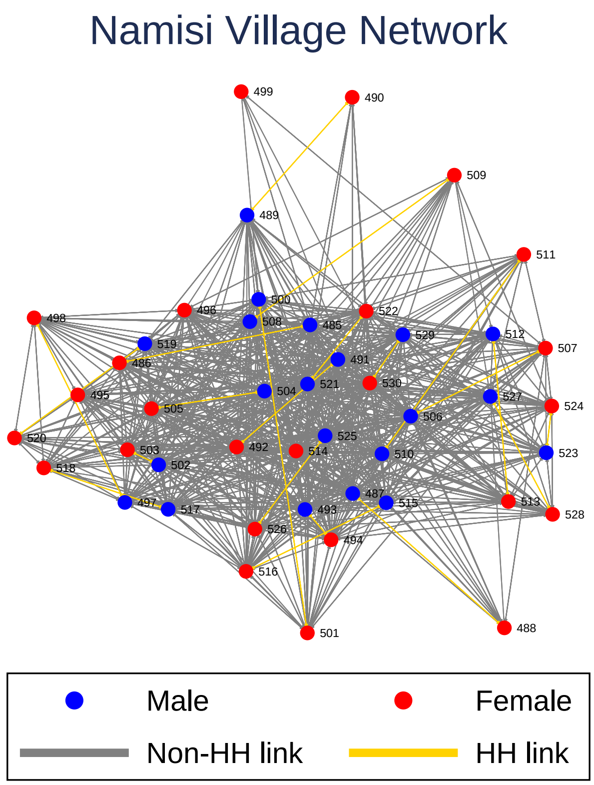


Figure 4.1: Network Example 1

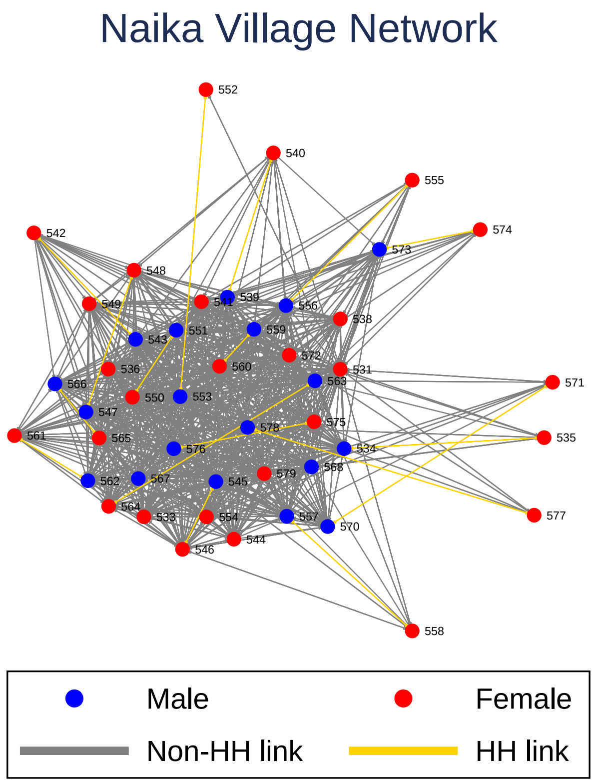


Figure 4.2: Network Example 2

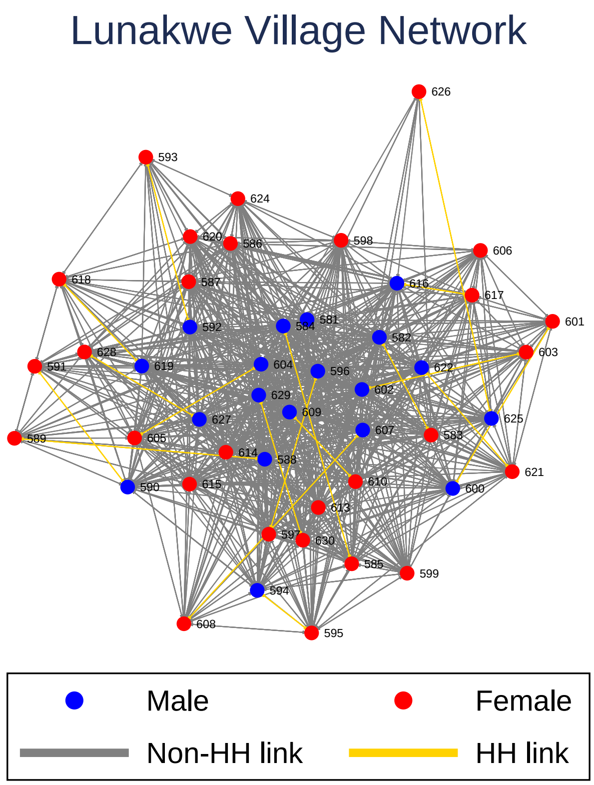


Figure 4.3: Network Example 3

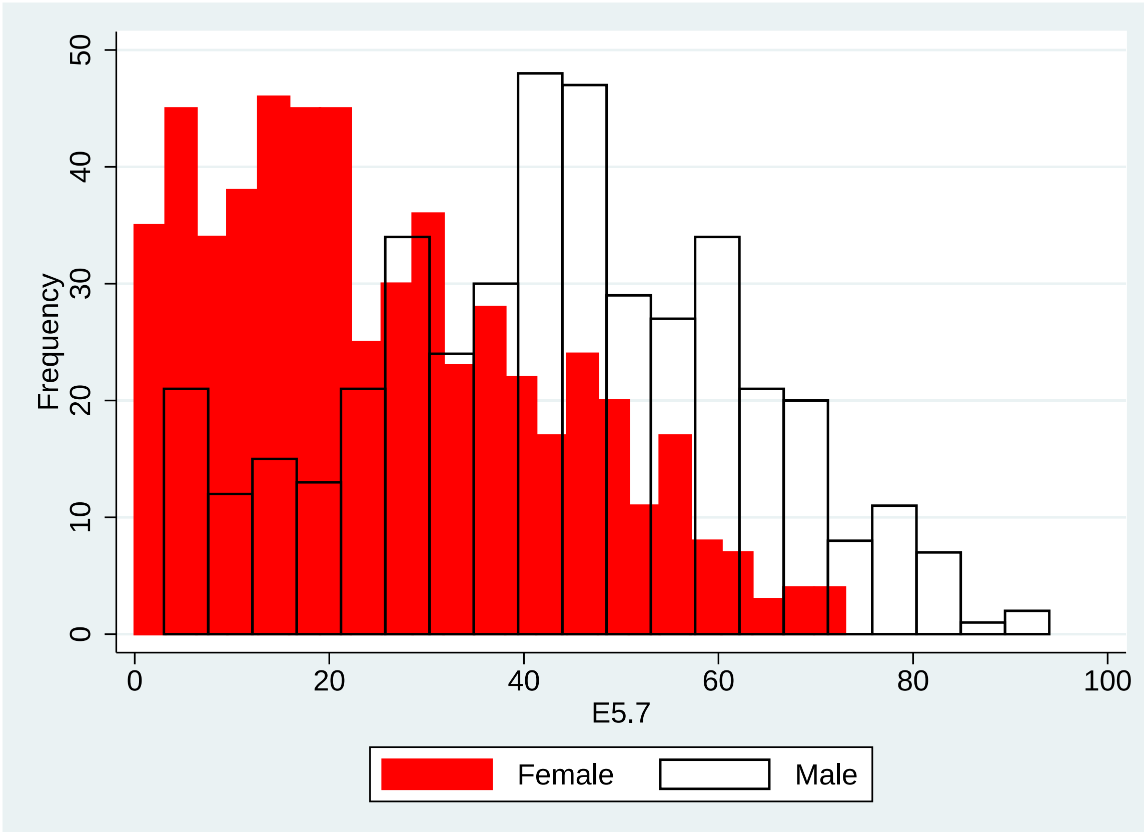
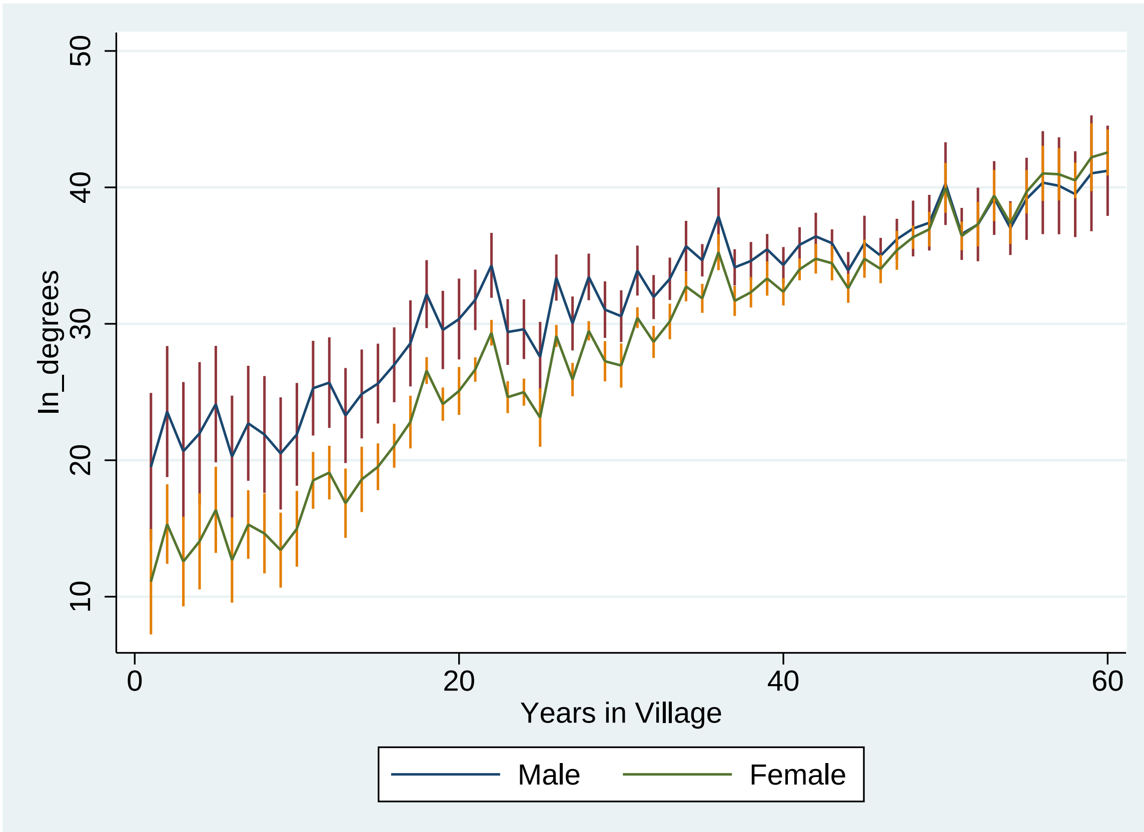
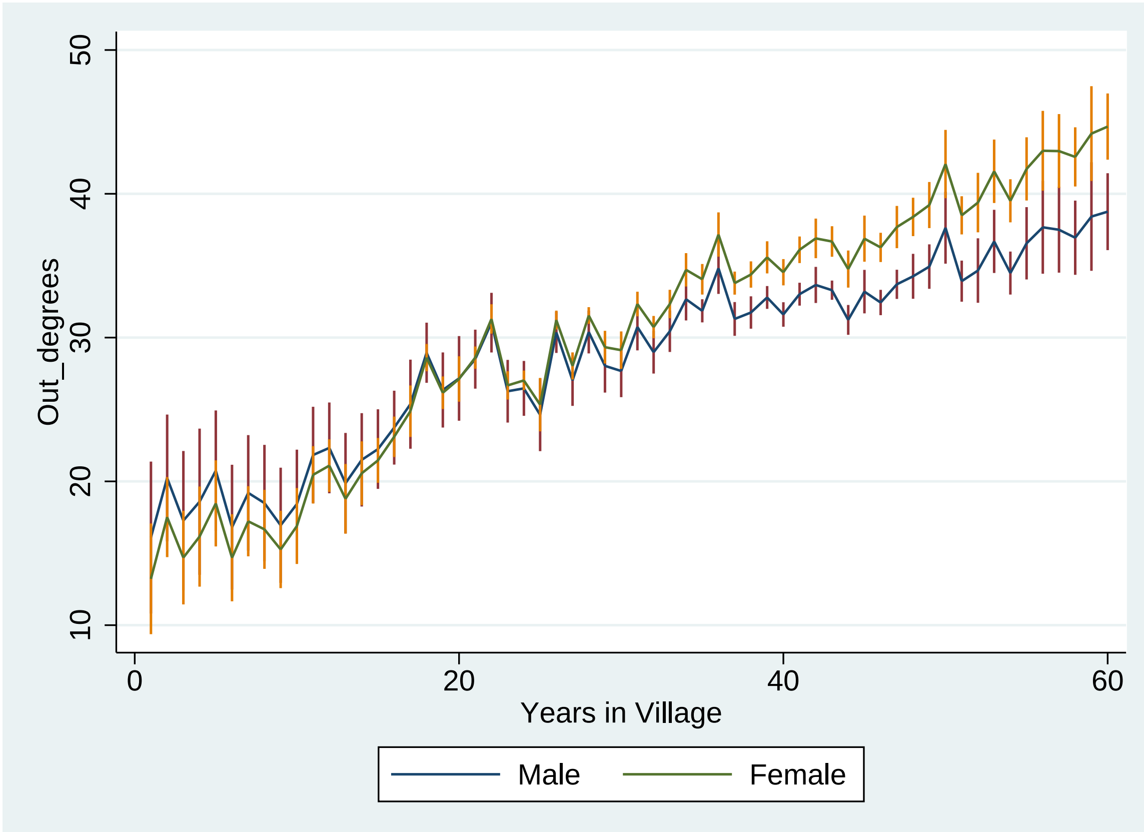


Figure 4.4: Years in Village by Gender



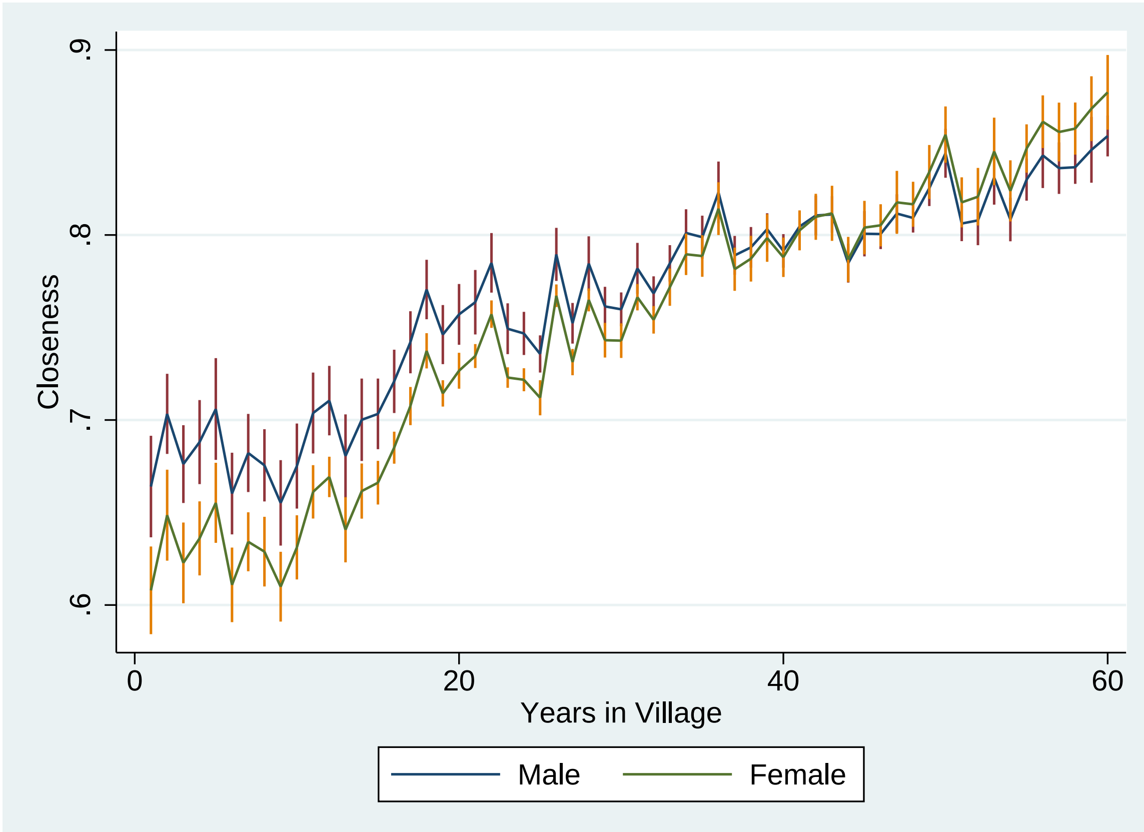
Note: Vertical bars represent 95% confidence intervals.

Figure 4.5: Marginal Effects: In-Degree over Years in Village



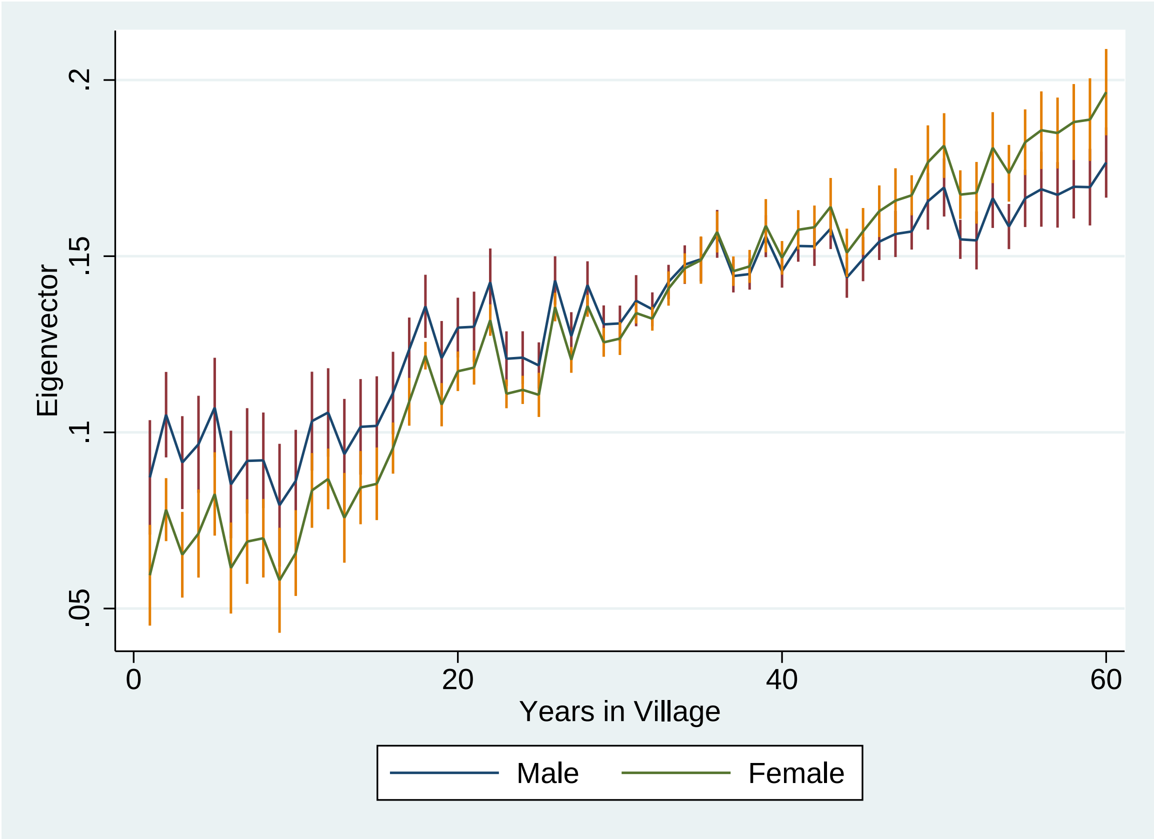
Note: Vertical bars represent 95% confidence intervals.

Figure 4.6: Marginal Effects: Out-Degree over Years in Village



Note: Vertical bars represent 95% confidence intervals.

Figure 4.7: Marginal Effects: Closeness over Years in Village

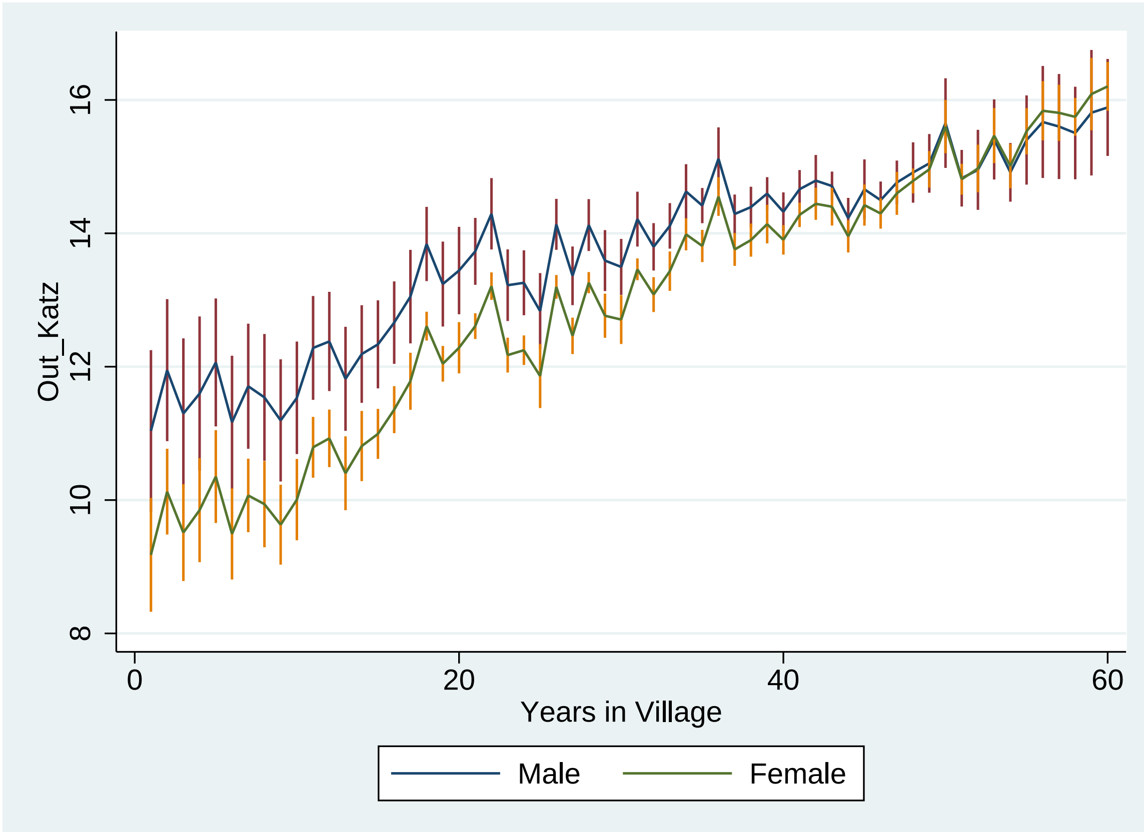


Note: Vertical bars represent 95% confidence intervals.

Figure 4.8: Marginal Effects: Eigenvector over Years in Village

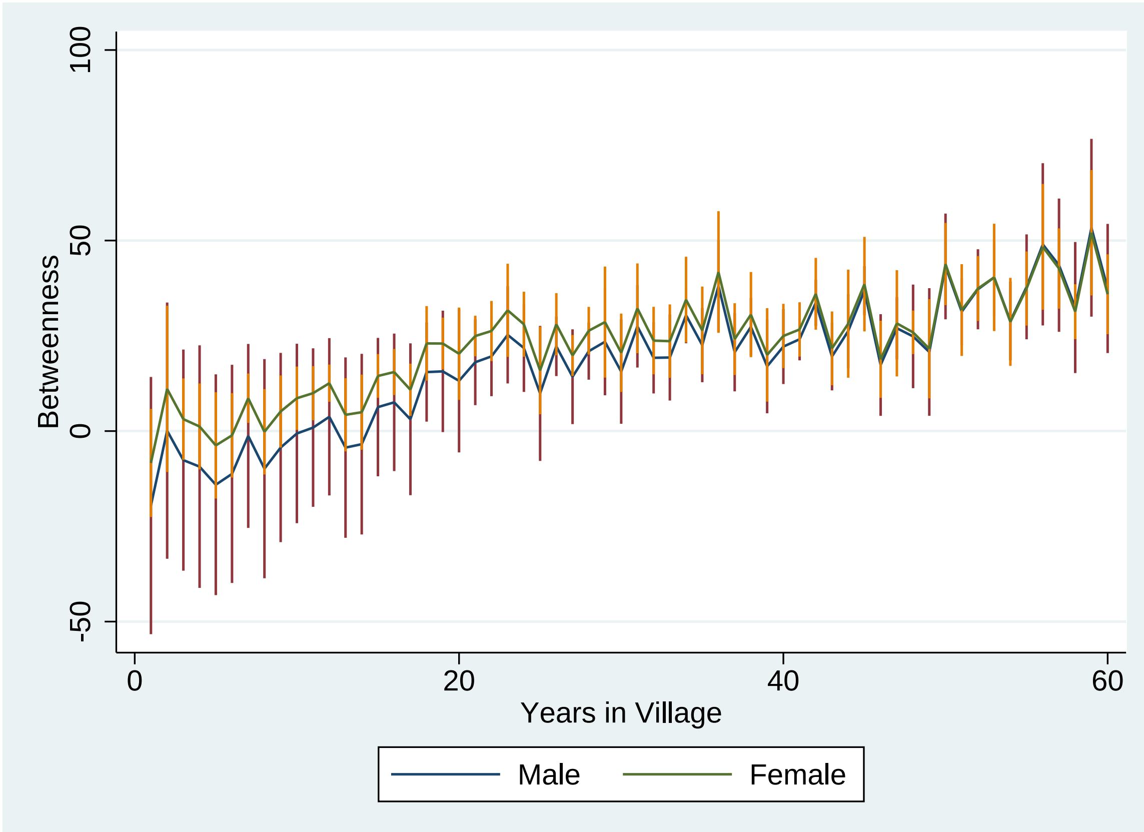
Note: Vertical bars represent 95% confidence intervals.

Figure 4.9: Marginal Effects: ln-Katz over Years in Village



Note: Vertical bars represent 95% confidence intervals.

Figure 4.10: Marginal Effects: Out-Katz over Years in Village



Note: Vertical bars represent 95% confidence intervals.

Figure 4.11: Marginal Effects: Betweenness over Years in Village

Table 4.1: Summary statistics

Variable	Mean	Std. Dev.	Min	Max
Individual (n=992)				
Age	48.51	15.82	19.00	109.00[a]
Years of Education	7.90	3.71	0.00	26.00[b]
Yes=1:				
Basic math ability[c]	0.54	0.50	0.00	1.00
Female	0.57	0.50	0.00	1.00
Widow/er	0.14	0.34	0.00	1.00
Primary occupation is farmer	0.87	0.34	0.00	1.00
Religion:				
Anglican	0.26	0.44	0.00	1.00
Catholic	0.16	0.37	0.00	1.00
Tribe:				
Pentecostal	0.43	0.50	0.00	1.00
Bukusu subtribe	0.33	0.47	0.00	1.00
Luhya tribe (except Bukusu)	0.33	0.47	0.00	1.00
Iteso tribe	0.26	0.44	0.00	1.00
Kalenjin tribe	0.06	0.24	0.00	1.00
Household (n=612)				
Household size[d]	5.24	3.21	0.00	40.00
Total farm area (acres)	1.04	1.05	0.02	8.87
Yes=1:				
Household head is male	0.57	0.50	0.00	1.00
Organic inputs (within past two seasons)	0.43	0.50	0.00	1.00
Inorganic inputs (within past two seasons)	0.88	0.32	0.00	1.00
No inputs (within past two seasons)	0.07	0.26	0.00	1.00
NGO contact	0.16	0.37	0.00	1.00
River as water source	0.45	0.50	0.00	1.00
Electricity access (grid)	0.11	0.32	0.00	1.00
Solar panels	0.32	0.47	0.00	1.00
Metal roof	0.89	0.32	0.00	1.00
Mud walls	0.78	0.41	0.00	1.00
Polygamous household	0.09	0.29	0.00	1.00
Own cow(s)	0.37	0.48	0.00	1.00

Note: [a]There was one woman who claimed she was 109 years old. [b]The sample included a couple of individuals who were university professors and had PhDs. [c]Was able to do a basic multiplication problem. [d]Defined as the number of individuals who spent the night at that dwelling last night. [e]1 after trimming outliers.

Table 4.2: Network Summary Statistics by Peer Dyad 1

	$\text{Male}_i\text{-Male}_j$		All Others		$\text{Male}_i\text{-Female}_j$		All Others	
Know individual j	0.78	(0.010)	0.67***	(0.0053)	0.66	(0.0098)	0.70***	(0.0054)
Have met individual j in person	0.76	(0.010)	0.64***	(0.0053)	0.63	(0.0100)	0.67***	(0.0055)
N (dyads)	1,830		7,874		2,337		7,367	
Friendship level (1=Not a friend, 5=Close friend)	4.11	(0.019)	3.98***	(0.011)	3.98	(0.022)	4.02*	(0.011)
Level of information quality from j (1=Poor, 5=Great)	3.65	(0.027)	3.55***	(0.014)	3.56	(0.028)	3.57	(0.014)
Yes=1								
j is immediate family	0.055	(0.0061)	0.054	(0.0032)	0.069	(0.0066)	0.05***	(0.0032)
j is extended family	0.34	(0.013)	0.34	(0.0067)	0.35	(0.012)	0.34	(0.0067)
Known j at least ten years	0.88	(0.0087)	0.77***	(0.0057)	0.81	(0.010)	0.79	(0.0057)
Speak at least weekly with j	0.83	(0.010)	0.80**	(0.0055)	0.8	(0.010)	0.81	(0.0056)
Speak daily with j†	0.29	(0.012)	0.27	(0.0063)	0.26	(0.011)	0.28	(0.0063)
Attend same RI as j	0.31	(0.012)	0.33	(0.0066)	0.35	(0.012)	0.32**	(0.0066)
Attend same non-religious organization as j	0.19	(0.010)	0.19	(0.0055)	0.18	(0.010)	0.19	(0.0056)
Fertilizer advice received from j	0.18	(0.010)	0.17	(0.0053)	0.19	(0.010)	0.17	(0.0054)
Planting advice received from j	0.16	(0.0099)	0.15	(0.0051)	0.17	(0.0097)	0.15	(0.0051)
Crop buyer advice received from j	0.072	(0.0069)	0.09**	(0.0040)	0.11	(0.0080)	0.08***	(0.0040)
At least one form of advice received from j	0.23	(0.011)	0.22	(0.0059)	0.24	(0.011)	0.22	(0.0059)
Worked together with j in past 12 months	0.32	(0.012)	0.32	(0.0066)	0.33	(0.012)	0.32	(0.0066)
Trust j to watch a valuable for you for one week	0.25	(0.012)	0.22**	(0.0059)	0.23	(0.011)	0.22	(0.0059)
N (dyads)	1,394		5,050		1,478		4,966	

† "Speak daily with j" is a subset of "Speak at least weekly with j." Standard deviations in parenthesize to right of mean.

Table 4.3: Network Summary Statistics by Peer Dyad 2

	Female$_i$-Male$_j$		All Others		Female$_i$-Female$_j$		All Others	
Know individual j	0.71	(0.0089)	0.68***	(0.0054)	0.63	(0.0088)	0.71***	(0.0056)
Have met individual j in person	0.68	(0.0092)	0.66**	(0.0054)	0.61	(0.0089)	0.69***	(0.0057)
N (dyads)	2,570		7,134		2,967		6,737	
Friendship level (1=Not a friend, 5=Close friend)	3.98	(0.020)	4.02*	(0.012)	3.99	(0.019)	4.02	(0.012)
Level of information quality from j (1=Poor, 5=Great)	3.56	(0.024)	3.57	(0.015)	3.53	(0.023)	3.59**	(0.015)
Yes=1								
j is immediate family	0.067	(0.0060)	0.049***	(0.0033)	0.029	(0.0039)	0.064***	(0.0033)
j is extended family	0.37	(0.011)	0.33***	(0.0069)	0.3	(0.011)	0.36***	(0.0070)
Known j at least ten years	0.77	(0.010)	0.8**	(0.0059)	0.73	(0.010)	0.82***	(0.0059)
Speak at least weekly with j	0.8	(0.0096)	0.81	(0.0057)	0.81	(0.0091)	0.81	(0.0058)
Speak daily with j†	0.29	(0.011)	0.27**	(0.0065)	0.25	(0.010)	0.28**	(0.0066)
Attend same RI as j	0.33	(0.011)	0.32	(0.0068)	0.31	(0.011)	0.33*	(0.0069)
Attend same non-religious organization as j	0.21	(0.0096)	0.19*	(0.0057)	0.19	(0.0091)	0.19	(0.0058)
Fertilizer advice received from j	0.18	(0.0092)	0.17	(0.0055)	0.15	(0.0084)	0.18***	(0.0056)
Planting advice received from j	0.16	(0.0088)	0.15	(0.0053)	0.13	(0.0079)	0.16***	(0.0053)
Crop buyer advice received from j	0.1	(0.0073)	0.08***	(0.0041)	0.063	(0.0057)	0.095***	(0.0041)
At least one form of advice received from j	0.23	(0.010)	0.22	(0.0061)	0.2	(0.0095)	0.23***	(0.0061)
Worked together with j in past 12 months	0.33	(0.011)	0.32	(0.0068)	0.31	(0.011)	0.33	(0.0069)
Trust j to watch a valuable for you for one week	0.22	(0.0099)	0.22	(0.0061)	0.2	(0.0093)	0.23***	(0.0061)
N (dyads)	1,757		4,687		1,815		4,629	

† "Speak daily with j" is a subset of "Speak at least weekly with j." Standard deviations in parenthesize to right of mean.

Table 4.4: Network Centrality Statistics (n=991)

Variable	Mean	Std. Dev.	Min	Max
In-degree	29.75	15.35	1.00	96.00
Out-degree	29.75	14.54	0.00	96.00
Closeness	0.76	0.16	0.32	1.00
Eigenvector	0.13	0.05	0.00	0.25
In-Katz	13.31	3.91	2.94	32.68
Out-Katz	13.31	3.95	2.36	32.68
Betweenness	22.84	84.91	0.00	1359.64

Table 4.5: Network Centrality Estimations

	Out-degree	In-degree	Closeness	Eigenvector	Out-Katz	In-Katz	Betweenness
Female	-2.875	-8.415***	-0.056***	-0.028***	-1.859***	-0.802*	11.179
SE	1.693	2.265	0.014	0.007	0.507	0.388	14.289
Cluster p-value	0.106	0.001	0.001	0.001	0.002	0.052	0.444
WB p-value	0.130	0.002	0.004	0.002	0.002	0.058	0.572
Years in village	0.269***	0.249***	0.002***	0.001***	0.056***	0.050***	0.810**
SE	0.059	0.064	0.000	0.000	0.014	0.016	0.301
Cluster p-value	0.000	0.001	0.000	0.000	0.001	0.006	0.014
WB p-value	0.000	0.000	0.000	0.004	0.000	0.002	0.016
Female × years in village	0.149***	0.165**	0.001***	0.001***	0.037**	0.033***	-0.215
SE	0.040	0.059	0.000	0.000	0.013	0.009	0.398
Cluster p-value	0.001	0.011	0.001	0.000	0.011	0.002	0.595
WB p-value	0.000	0.004	0.002	0.000	0.002	0.000	0.810
Distance from village center	-11.427***	-11.285***	-0.134***	-0.069***	-2.648***	-2.340***	4.290
SE	1.525	1.593	0.020	0.013	0.354	0.422	4.451
Cluster p-value	0.000	0.000	0.000	0.000	0.000	0.000	0.347
WB p-value	0.002	0.002	0.002	0.004	0.002	0.002	0.282
Asset index	1.757***	1.859***	0.015***	0.006***	0.404***	0.403***	7.188
SE	0.517	0.485	0.004	0.002	0.107	0.120	4.353
Cluster p-value	0.003	0.001	0.002	0.002	0.001	0.003	0.115
WB p-value	0.006	0.002	0.010	0.002	0.002	0.006	0.090
Years of education	1.092***	1.143***	0.010***	0.004***	0.256***	0.248***	0.773
SE	0.382	0.367	0.003	0.001	0.081	0.084	4.422
Cluster p-value	0.010	0.006	0.002	0.002	0.005	0.008	0.863
WB p-value	0.006	0.000	0.020	0.006	0.000	0.000	0.892
Years of education sq.	-0.045**	-0.048**	-0.000*	-0.000**	-0.011**	-0.011**	0.090
SE	0.020	0.019	0.000	0.000	0.004	0.004	0.293
Cluster p-value	0.034	0.022	0.055	0.031	0.022	0.020	0.761
WB p-value	0.080	0.040	0.240	0.174	0.042	0.012	0.810
Age	1.153***	1.142***	0.010***	0.005***	0.251***	0.271***	1.292**
SE	0.128	0.132	0.001	0.001	0.029	0.031	0.490
Cluster p-value	0.000	0.000	0.000	0.000	0.000	0.000	0.016
WB p-value	0.000	0.000	0.000	0.000	0.000	0.000	0.014
Age sq.	-0.011***	-0.011***	-0.000***	-0.000***	-0.002***	-0.003***	-0.015***
SE	0.001	0.001	0.000	0.000	0.000	0.000	0.003
Cluster p-value	0.000	0.000	0.000	0.000	0.000	0.000	0.000
WB p-value	0.002	0.002	0.002	0.002	0.002	0.002	0.002
Constant	-7.580	-3.508	0.462***	-0.009	5.985***	4.970***	-31.946
SE	6.327	6.350	0.034	0.019	1.398	1.419	38.573
Cluster p-value	0.246	0.587	0.000	0.666	0.000	0.002	0.418
WB p-value	0.902	0.582	0.000	0.704	0.000	0.000	0.008
Additional HH and demo. vars.	√	√	√	√	√	√	√
Village fixed effects	√	√	√	√	√	√	√
N	991	991	991	991	991	991	991
R sq.	0.528	0.555	0.580	0.633	0.558	0.430	0.078

Dependent variable is the measure of network centrality (e.g. Out-Degree, In-Degree). Standard errors clustered at the village level. Wild bootstrap (WB) standard errors to control for small number of clusters (20) bootstrapped with 1000 repetitions. ***p<0.01, **p<0.05, *p<0.1.

Table 4.6: Peer Effects - No Distance Cutoff

2SLS	**Full sample**				**Female$_i$**				**Male$_i$**			
	I	II	III	IV	V	VI	VII	VIII	IX	X	XI	XII
Value of per-hecatre maize yield (j)	0.0007	0.0080	0.0188	0.0155	0.0005	0.0094	0.0294	0.0262	0.0002	0.0098	0.0214	0.0112
	(0.0005)	(0.0058)	(0.0125)	(0.0187)	(0.0012)	(0.0077)	(0.0226)	(0.0289)	(0.0004)	(0.0071)	(0.0139)	(0.0166)
Long rains season	0.6876***	0.7777***	0.7710***	0.3449**	0.6242**	0.7819***	0.7744***	0.3426	0.7152***	0.7533***	0.7508***	0.3470**
	(0.1383)	(0.1367)	(0.1371)	(0.1389)	(0.2492)	(0.2295)	(0.2315)	(0.2466)	(0.1620)	(0.1629)	(0.1616)	(0.1397)
Individual (i)												
Female	-0.2643***	-0.0935	-0.0921	-0.1067								
	(0.0891)	(0.0964)	(0.0964)	(0.0815)								
Years in village		0.0409	0.0460	0.0177		0.1189	0.0974	0.0729		-0.0681	-0.1105	-0.0765
		(0.0883)	(0.0873)	(0.0719)		(0.1405)	(0.1430)	(0.1203)		(0.0962)	(0.0948)	(0.0876)
Distance to plot (meters)		0.0404	0.0430	-0.0124		0.0806	0.0822	0.0131		0.0012	-0.0051	-0.0695
		(0.0557)	(0.0563)	(0.0571)		(0.0916)	(0.0955)	(0.0965)		(0.0671)	(0.0672)	(0.0693)
Age		-0.0747	-0.0935	0.0083		-0.5910	-0.5340	-0.2382		0.2510	0.2455	0.1742
		(0.1548)	(0.1598)	(0.1369)		(0.3696)	(0.3669)	(0.3068)		(0.1849)	(0.1914)	(0.1838)
Years of education		0.2708***	0.2657***	0.1514*		0.3692**	0.3288**	0.2025		0.0522	0.0652	-0.0174
		(0.0965)	(0.0952)	(0.0789)		(0.1490)	(0.1513)	(0.1401)		(0.0962)	(0.0934)	(0.0760)
Crop area		-0.5201***	-0.5178***	0.0216		-0.5008***	-0.4590***	0.0909		-0.5494***	-0.5734***	0.0113
		(0.0523)	(0.0537)	(0.0893)		(0.0740)	(0.0801)	(0.1469)		(0.0728)	(0.0741)	(0.1277)
Farmer		0.0511	0.0670	0.1613		0.0292	0.0716	0.2160		0.0643	0.1053	0.1505
		(0.1619)	(0.1584)	(0.1414)		(0.2773)	(0.2848)	(0.2715)		(0.1931)	(0.1841)	(0.1499)
Own plot		-0.1398	-0.1409	-0.0134		-0.9632***	-0.8574***	-0.3428		0.1740	0.1142	0.0726
		(0.3849)	(0.3744)	(0.2627)		(0.2595)	(0.2971)	(0.3091)		(0.4797)	(0.4845)	(0.3516)
Math ability		0.0540	0.0469	-0.0127		0.0289	0.0561	0.0525		0.0429	0.0154	-0.0653
		(0.1156)	(0.1170)	(0.1068)		(0.2068)	(0.2175)	(0.2003)		(0.1151)	(0.1211)	(0.1112)
TLU		0.1204*	0.1116*	0.0389		0.0744	0.0561	-0.1420		0.2045**	0.1977**	0.1251*
		(0.0661)	(0.0673)	(0.0668)		(0.1017)	(0.0999)	(0.1149)		(0.0899)	(0.0902)	(0.0757)
Purchased plot		0.1209	0.1111	0.1260		0.2316	0.2483	0.1762		0.0154	-0.0399	0.0316
		(0.1148)	(0.1142)	(0.1041)		(0.1697)	(0.1849)	(0.1782)		(0.1528)	(0.1484)	(0.1207)
Asset index		0.1840***	0.1876***	0.0455		0.2497*	0.2706*	0.1020		0.1391**	0.1564**	0.0264
		(0.0687)	(0.0675)	(0.0611)		(0.1433)	(0.1452)	(0.1207)		(0.0625)	(0.0608)	(0.0596)
Adult men		0.0022	0.0026	0.0379		-0.0107	-0.0329	0.1200		0.2672	0.2837*	0.1642
		(0.0973)	(0.0971)	(0.0898)		(0.1253)	(0.1287)	(0.1305)		(0.1693)	(0.1707)	(0.1627)
Adult women		0.1849	0.1725	0.0266		-0.0819	-0.0810	-0.1513		0.0922	0.0983	-0.0709
		(0.1180)	(0.1173)	(0.1058)		(0.1770)	(0.1810)	(0.1567)		(0.1589)	(0.1594)	(0.1428)
Household size		-0.2297***	-0.2305***	-0.1778**		-0.4020***	-0.4011***	-0.4015***		-0.1150	-0.1118	0.0253
		(0.0769)	(0.0765)	(0.0694)		(0.1507)	(0.1440)	(0.1317)		(0.0852)	(0.0864)	(0.0785)
Soil nutrients (i)												
Nitrogen			-0.0694	-0.0626			-0.0651	-0.0199			-0.0584	-0.0125
			(0.0457)	(0.0382)			(0.0795)	(0.0743)			(0.0620)	(0.0384)
Phosphorus			0.0083	0.0930**			0.1047	0.1703**			-0.0812	-0.0044
			(0.0467)	(0.0445)			(0.0812)	(0.0776)			(0.0541)	(0.0476)
Potassium			0.0192	0.1251**			0.1294	0.2141**			-0.1243**	-0.0352
			(0.0604)	(0.0523)			(0.0956)	(0.1006)			(0.0607)	(0.0590)
Sulphur			0.1072	0.0775			0.2089	0.1228			0.0534	-0.0027
			(0.0817)	(0.0650)			(0.1658)	(0.1473)			(0.0938)	(0.0622)
Carbon			0.0776	0.0012			0.1108	0.1615			0.0100	-0.0781
			(0.0749)	(0.0596)			(0.1363)	(0.1229)			(0.0888)	(0.0641)
Acidity (pH)			0.2565	-0.2572			0.0795	-0.2990			0.6339	0.3693
			(0.6056)	(0.5464)			(1.1746)	(0.9780)			(0.6928)	(0.6253)
Farm management vars. (i)												
Ag. NGO participant				0.0583				0.2510				-0.0183
				(0.1076)				(0.1898)				(0.1449)
Hired labor				0.2023**				0.1557				0.2259**
				(0.0827)				(0.1345)				(0.0885)
Total hours worked per-hectare				0.5087***				0.5540***				0.5349***
				(0.0877)				(0.1506)				(0.1167)
Seed payment				0.0505***				0.0276				0.0747***
				(0.0146)				(0.0188)				(0.0226)
Monocrop (1=yes)				-0.3309***				-0.1393				-0.5257***
				(0.0934)				(0.1475)				(0.1099)
No organics used (1=yes)				0.6668**				1.0665*				0.0609
				(0.3179)				(0.5850)				(0.2971)
No DAP used (1=yes)				0.3701				-0.0084				0.9298**
				(0.3014)				(0.4193)				(0.4072)
No CAN used (1=yes)				0.0162				0.6232				-0.5491
				(0.2843)				(0.4408)				(0.4243)
Total organics used per-hectare				0.0977**				0.1729**				0.0090
				(0.0405)				(0.0794)				(0.0371)
Total DAP used per-hectare				0.1311**				0.0821				0.2003***
				(0.0551)				(0.0867)				(0.0691)
Total CAN used per-hectare				0.0381				0.1942**				-0.0858
				(0.0511)				(0.0953)				(0.0651)
FEs (village, enumerator, svy. month)	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
Observations	17,512	17,455	17,455	17,455	7,472	7,435	7,435	7,435	10,040	10,020	10,020	10,020
R-squared	0.126	0.276	0.279	0.411	0.113	0.288	0.303	0.439	0.186	0.341	0.351	0.515

2SLS regression results - OLS results available upon request. Standard errors clustered at the individual level. Non-binary variables in log form. ***p<0.01, **p<0.05, *p<0.1.

Table 4.6: Peer Effects Continued - No Distance Cutoff

2SLS	**Full sample**				**Female$_i$**				**Male$_i$**			
	I	II	III	IV	V	VI	VII	VIII	IX	X	XI	XII
Individual (j)												
Female	0.0042	0.0144*	0.0181**	0.0147*	-0.0018	0.0009	0.0124	-0.0005	0.0086	0.0208**	0.0193**	0.0280***
	(0.0067)	(0.0077)	(0.0081)	(0.0078)	(0.0106)	(0.0122)	(0.0126)	(0.0120)	(0.0066)	(0.0082) (0.0080)	(0.0077)	
Years in village		0.0049	0.0060	0.0030		0.0042	0.0035	-0.0085		-0.0001	-0.0015	0.0036
		(0.0068)	(0.0066)	(0.0055)		(0.0123)	(0.0124)	(0.0089)		(0.0073)	(0.0069)	(0.0061)
Distance to plot (meters)		-0.0044	-0.0067	-0.0066		-0.0042	-0.0091	-0.0131**		-0.0037	-0.0082	-0.0083*
		(0.0039)	(0.0044)	(0.0044)		(0.0074)	(0.0077)	(0.0065)		(0.0046)	(0.0051)	(0.0044)
Age		-0.0183	-0.0210	-0.0134		-0.0128	-0.0084	0.0093		-0.0228	-0.0256*	-0.0235
		(0.0124)	(0.0130)	(0.0128)		(0.0190)	(0.0241)	(0.0186)		(0.0150)	(0.0154)	(0.0150)
Years of education		-0.0017	-0.0042	-0.0030		-0.0120	-0.0218	-0.0095		0.0011	-0.0010	-0.0001
		(0.0063)	(0.0073)	(0.0071)		(0.0113)	(0.0137)	(0.0110)		(0.0063)	(0.0069)	(0.0061)
Crop area		0.0024	0.0093	0.0016		0.0013	0.0127	0.0077		0.0077	0.0159	0.0038
		(0.0062)	(0.0078)	(0.0069)		(0.0095)	(0.0109)	(0.0134)		(0.0066)	(0.0109)	(0.0062)
Farmer		0.0158	0.0204*	0.0089		0.0017	0.0075	0.0002		0.0173	0.0228	0.0121
		(0.0114)	(0.0116)	(0.0100)		(0.0246)	(0.0218)	(0.0235)		(0.0138)	(0.0147)	(0.0104)
Own plot		-0.0075	-0.0017	-0.0062		-0.0129	0.0175	-0.0048		0.0067	0.0048	0.0031
		(0.0192)	(0.0187)	(0.0175)		(0.0401)	(0.0434)	(0.0324)		(0.0149)	(0.0157)	(0.0148)
Math ability		0.0049	0.0046	-0.0031		0.0100	0.0088	-0.0096		-0.0014	-0.0012	-0.0005
		(0.0060)	(0.0063)	(0.0073)		(0.0113)	(0.0109)	(0.0115)		(0.0061)	(0.0062)	(0.0069)
TLU		-0.0066	-0.0073	-0.0018		-0.0168	-0.0172	-0.0119		0.0007	-0.0000	0.0072
		(0.0050)	(0.0051)	(0.0061)		(0.0105)	(0.0105)	(0.0112)		(0.0047)	(0.0047)	(0.0064)
Purchased plot		0.0123	0.0110	0.0131		0.0096	0.0088	0.0308		0.0091	0.0083	0.0103
		(0.0120)	(0.0120)	(0.0098)		(0.0259)	(0.0251)	(0.0194)		(0.0116)	(0.0114)	(0.0100)
Asset index		0.0071	0.0060	0.0014		0.0189**	0.0222**	0.0074		0.0048	0.0049	0.0013
		(0.0044)	(0.0042)	(0.0038)		(0.0093)	(0.0091)	(0.0082)		(0.0042)	(0.0043)	(0.0033)
Adult men		0.0215***	0.0237***	0.0164**		0.0217*	0.0306**	0.0149		0.0232**	0.0234**	0.0239***
		(0.0075)	(0.0084)	(0.0075)		(0.0117)	(0.0140)	(0.0123)		(0.0091)	(0.0091)	(0.0079)
Adult women		-0.0093	-0.0099	-0.0106		0.0050	0.0105	0.0010		-0.0234**	-0.0229*	-0.0129
		(0.0078)	(0.0081)	(0.0084)		(0.0136)	(0.0148)	(0.0127)		(0.0096)	(0.0117)	(0.0085)
Household size		-0.0138***	-0.0109**	-0.0042		-0.0187**	-0.0243***	-0.0078		-0.0087	-0.0029	-0.0050
		(0.0051)	(0.0053)	(0.0042)		(0.0076)	(0.0094)	(0.0068)		(0.0058)	(0.0078)	(0.0055)
Soil nutrients (j) Nitrogen			0.0002	0.0019			-0.0196***	-0.0085*			0.0063**	0.0046*
			(0.0033)	(0.0024)			(0.0048)	(0.0044)			(0.0030)	(0.0026)
Phosphorus			0.0022	-0.0016			0.0043	0.0006			0.0001	-0.0046
			(0.0032)	(0.0030)			(0.0057)	(0.0082)			(0.0035)	(0.0031)
Potassium			-0.0025	-0.0027			0.0001	-0.0022			-0.0028	-0.0023
			(0.0048)	(0.0056)			(0.0093)	(0.0111)			(0.0039)	(0.0038)
Sulphur			-0.0037	-0.0035			0.0121	0.0073			-0.0114*	-0.0070
			(0.0075)	(0.0069)			(0.0162)	(0.0149)			(0.0060)	(0.0049)
Carbon			0.0067	-0.0027			0.0027	-0.0061			0.0034	-0.0029
			(0.0056)	(0.0059)			(0.0096)	(0.0112)			(0.0061)	(0.0047)
Acidity (pH)			-0.0667	-0.0505			-0.1177	-0.1132			-0.0455	-0.0113
			(0.0683)	(0.0620)			(0.1528)	(0.1278)			(0.0509)	(0.0382)
Farm management vars. (j) Ag. NGO participant				0.0104				-0.0101				0.0196**
				(0.0099)				(0.0146)				(0.0098)
Hired labor				0.0135**				0.0423***				-0.0032
				(0.0067)				(0.0120)				(0.0058)
Total hours worked per-hectare				-0.0077				-0.0007				-0.0077
				(0.0093)				(0.0140)				(0.0079)
Seed payment				0.0045**				0.0046				0.0043
				(0.0022)				(0.0031)				(0.0027)
Monocrop (1=yes)				0.0102				0.0150				0.0053
				(0.0078)				(0.0113)				(0.0094)
No organics used (1=yes)				0.0274				0.0394				0.0146
				(0.0253)				(0.0458)				(0.0260)
No DAP used (1=yes)				0.0586*				0.0587				0.0784**
				(0.0311)				(0.0534)				(0.0356)
No CAN used (1=yes)				-0.0411				-0.0624				-0.0492**
				(0.0269)				(0.0492)				(0.0237)
Total organics used per-hectare				0.0030				0.0033				0.0009
				(0.0033)				(0.0060)				(0.0034)
Total DAP used per-hectare				0.0081				0.0052				0.0124*
				(0.0061)				(0.0103)				(0.0068)
Total CAN used per-hectare				-0.0093*				-0.0124				-0.0119**
				(0.0053)				(0.0088)				(0.0049)
FEs (village, enumerator, survey month)	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
Observations	17,512	17,455	17,455	17,455	7,472	7,435	7,435	7,435	10,040	10,020	10,020	10,020
R-squared	0.126	0.276	0.279	0.411	0.113	0.288	0.303	0.439	0.186	0.341	0.351	0.515

2SLS regression results - OLS results available upon request. Standard errors clustered at the individual level. Non-binary variables in log form. ***$p<0.01$, **$p<0.05$, *$p<0.1$.

Table 4.7: Peer Effects - 1KM Cutoff

2SLS	**Full sample**				**Female$_i$**				**Male$_i$**			
	I	II	III	IV	V	VI	VII	VIII	IX	X	XI	XII
Value of per-hecatre maize yield (j)	0.0006	0.0074	0.0150	0.0067	0.0003	0.0099	0.0301	0.0169	0.0003	0.0087	0.0150	0.0047
	(0.0005)	(0.0066)	(0.0132)	(0.0224)	(0.0013)	(0.0098)	(0.0242)	(0.0330)	(0.0004)	(0.0080)	(0.0143)	(0.0222)
Long rains season	0.6878***	0.7793***	0.7761***	0.3483**	0.6246**	0.7827***	0.7774***	0.3461	0.7153***	0.7553***	0.7579***	0.3511**
	(0.1384)	(0.1372)	(0.1383)	(0.1394)	(0.2493)	(0.2294)	(0.2311)	(0.2459)	(0.1622)	(0.1639)	(0.1649)	(0.1408)
Individual (i)												
Female	-0.2645***	-0.0934	-0.0920	-0.1064								
	(0.0891)	(0.0964)	(0.0964)	(0.0815)								
Years in village		0.0416	0.0465	0.0185		0.1197	0.0983	0.0735		-0.0678	-0.1103	-0.0759
		(0.0883)	(0.0872)	(0.0718)		(0.1402)	(0.1426)	(0.1202)		(0.0961)	(0.0947)	(0.0876)
Distance to plot (meters)		0.0406	0.0431	-0.0123		0.0809	0.0834	0.0139		0.0011	-0.0055	-0.0697
		(0.0557)	(0.0562)	(0.0570)		(0.0915)	(0.0954)	(0.0964)		(0.0669)	(0.0672)	(0.0693)
Age		-0.0744	-0.0927	0.0090		-0.5893	-0.5332	-0.2381		0.2507	0.2448	0.1748
		(0.1548)	(0.1597)	(0.1368)		(0.3693)	(0.3663)	(0.3065)		(0.1849)	(0.1914)	(0.1837)
Years of education		0.2707***	0.2656***	0.1512*		0.3695**	0.3291**	0.2024		0.0519	0.0649	-0.0176
		(0.0966)	(0.0953)	(0.0789)		(0.1491)	(0.1513)	(0.1400)		(0.0961)	(0.0933)	(0.0758)
Crop area		-0.5203***	-0.5180***	0.0210		-0.5011***	-0.4595***	0.0906		-0.5496***	-0.5736***	0.0108
		(0.0523)	(0.0537)	(0.0892)		(0.0740)	(0.0801)	(0.1469)		(0.0728)	(0.0741)	(0.1276)
Farmer		0.0516	0.0674	0.1616		0.0296	0.0714	0.2158		0.0647	0.1063	0.1508
		(0.1620)	(0.1585)	(0.1414)		(0.2775)	(0.2850)	(0.2712)		(0.1932)	(0.1842)	(0.1499)
Own plot		-0.1387	-0.1404	-0.0133		-0.9620***	-0.8543***	-0.3383		0.1750	0.1133	0.0725
		(0.3848)	(0.3746)	(0.2627)		(0.2594)	(0.2974)	(0.3092)		(0.4793)	(0.4845)	(0.3515)
Math ability		0.0549	0.0481	-0.0118		0.0300	0.0573	0.0528		0.0433	0.0161	-0.0646
		(0.1155)	(0.1169)	(0.1068)		(0.2067)	(0.2172)	(0.1998)		(0.1150)	(0.1208)	(0.1109)
TLU		0.1205*	0.1113*	0.0385		0.0746	0.0563	-0.1426		0.2047**	0.1975**	0.1250*
		(0.0661)	(0.0672)	(0.0667)		(0.1017)	(0.1000)	(0.1150)		(0.0898)	(0.0900)	(0.0754)
Purchased plot		0.1205	0.1105	0.1256		0.2303	0.2472	0.1764		0.0153	-0.0400	0.0314
		(0.1148)	(0.1142)	(0.1041)		(0.1696)	(0.1847)	(0.1779)		(0.1527)	(0.1482)	(0.1207)
Asset index		0.1838***	0.1878***	0.0454		0.2498*	0.2704*	0.1023		0.1389**	0.1565**	0.0262
		(0.0687)	(0.0675)	(0.0611)		(0.1431)	(0.1452)	(0.1208)		(0.0625)	(0.0608)	(0.0596)
Adult men		0.0021	0.0026	0.0382		-0.0108	-0.0322	0.1203		0.2668	0.2830*	0.1641
		(0.0973)	(0.0971)	(0.0898)		(0.1254)	(0.1288)	(0.1306)		(0.1692)	(0.1706)	(0.1627)
Adult women		0.1840	0.1714	0.0255		-0.0820	-0.0820	-0.1529		0.0906	0.0970	-0.0720
		(0.1180)	(0.1173)	(0.1058)		(0.1771)	(0.1812)	(0.1566)		(0.1589)	(0.1594)	(0.1428)
Household size		-0.2293***	-0.2302***	-0.1775**		-0.4018***	-0.4007***	-0.4015***		-0.1147	-0.1122	0.0254
		(0.0768)	(0.0765)	(0.0694)		(0.1507)	(0.1439)	(0.1317)		(0.0852)	(0.0864)	(0.0785)
Soil nutrients (i)												
Nitrogen			-0.0693	-0.0624			-0.0647	-0.0194			-0.0586	-0.0126
			(0.0457)	(0.0382)			(0.0796)	(0.0743)			(0.0620)	(0.0384)
Phosphorus			0.0082	0.0929**			0.1048	0.1705**			-0.0809	-0.0043
			(0.0467)	(0.0444)			(0.0812)	(0.0776)			(0.0540)	(0.0475)
Potassium			0.0193	0.1252**			0.1299	0.2138**			-0.1241**	-0.0349
			(0.0604)	(0.0523)			(0.0956)	(0.1006)			(0.0606)	(0.0590)
Sulphur			0.1065	0.0769			0.2074	0.1226			0.0531	-0.0033
			(0.0816)	(0.0650)			(0.1658)	(0.1474)			(0.0938)	(0.0622)
Carbon			0.0774	0.0010			0.1095	0.1607			0.0098	-0.0781
			(0.0749)	(0.0595)			(0.1363)	(0.1230)			(0.0887)	(0.0640)
Acidity (pH)			0.2549	-0.2576			0.0851	-0.2918			0.6326	0.3691
			(0.6053)	(0.5463)			(1.1742)	(0.9778)			(0.6919)	(0.6247)
Farm management vars. (i)												
Ag. NGO participant				0.0602				0.2550				-0.0170
				(0.1074)				(0.1897)				(0.1442)
Hired labor				0.2025**				0.1557				0.2259**
				(0.0827)				(0.1343)				(0.0884)
Total hours worked per-hectare				0.5083***				0.5545***				0.5341***
				(0.0877)				(0.1506)				(0.1166)
Seed payment				0.0507***				0.0277				0.0748***
				(0.0147)				(0.0188)				(0.0226)
Monocrop (1=yes)				-0.3304***				-0.1385				-0.5255***
				(0.0933)				(0.1475)				(0.1100)
No organics used (1=yes)				0.6678**				1.0661*				0.0611
				(0.3177)				(0.5832)				(0.2968)
No DAP used (1=yes)				0.3726				-0.0065				0.9322**
				(0.3019)				(0.4192)				(0.4079)
No CAN used (1=yes)				0.0152				0.6209				-0.5494
				(0.2843)				(0.4404)				(0.4243)
Total organics used per-hectare				0.0978**				0.1729**				0.0089
				(0.0404)				(0.0791)				(0.0370)
Total DAP used per-hectare				0.1316**				0.0825				0.2008***
				(0.0552)				(0.0866)				(0.0692)
Total CAN used per-hectare				0.0379				0.1937**				-0.0860
				(0.0511)				(0.0951)				(0.0651)
FEs (village, enumerator, (svy. month)	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
Observations	17,512	17,455	17,455	17,455	7,472	7,435	7,435	7,435	10,040	10,020	10,020	10,020
R-squared	0.126	0.276	0.279	0.411	0.113	0.288	0.303	0.439	0.186	0.341	0.352	0.515

2SLS regression results - OLS results available upon request. Standard errors clustered at the individual level. Non-binary variables in log form. ***p<0.01, **p<0.05, *p<0.1.

Table 4.7: Peer Effects Continued - 1KM Cutoff

2SLS	**Full sample**				**Female$_i$**				**Male$_i$**			
	I	II	III	IV	V	VI	VII	VIII	IX	X	XI	XII
Individual (j)												
Female	0.0041	0.0177**	0.0208**	0.0174*	-0.0025	0.0044	0.0175	0.0003	0.0085	0.0239***	0.0211**	0.0305***
	(0.0076)	(0.0088)	(0.0096)	(0.0095)	(0.0122)	(0.0137)	(0.0157)	(0.0150)	(0.0074)	(0.0093)	(0.0088)	(0.0088)
	(0.0001)	(0.0001)	(0.0001)									
Years in village		0.0026	0.0030	0.0006		0.0014	0.0014	-0.0079		-0.0018	-0.0045	-0.0004
		(0.0070)	(0.0069)	(0.0061)		(0.0144)	(0.0144)	(0.0100)		(0.0080)	(0.0077)	(0.0072)
Distance to plot (meters)		-0.0050	-0.0069	-0.0059		-0.0046	-0.0099	-0.0139*		-0.0046	-0.0081	-0.0068
		(0.0046)	(0.0050)	(0.0050)		(0.0082)	(0.0084)	(0.0074)		(0.0057)	(0.0060)	(0.0052)
Age		-0.0165	-0.0187	-0.0100		-0.0144	-0.0101	0.0095		-0.0194	-0.0206	-0.0183
		(0.0134)	(0.0135)	(0.0141)		(0.0212)	(0.0258)	(0.0199)		(0.0164)	(0.0164)	(0.0168)
Years of education		-0.0019	-0.0042	-0.0034		-0.0152	-0.0243*	-0.0079		0.0031	0.0020	-0.0001
		(0.0068)	(0.0077)	(0.0079)		(0.0127)	(0.0142)	(0.0115)		(0.0070)	(0.0075)	(0.0077)
Crop area		0.0021	0.0067	0.0020		0.0026	0.0131	0.0100		0.0077	0.0114	0.0036
		(0.0072)	(0.0083)	(0.0081)		(0.0113)	(0.0120)	(0.0151)		(0.0075)	(0.0114)	(0.0069)
Farmer		0.0193	0.0241*	0.0133		0.0070	0.0135	0.0066		0.0212	0.0266*	0.0144
		(0.0120)	(0.0123)	(0.0109)		(0.0253)	(0.0224)	(0.0263)		(0.0148)	(0.0160)	(0.0115)
Own plot		-0.0057	-0.0021	-0.0123		-0.0025	0.0287	-0.0140		0.0079	0.0024	0.0047
		(0.0199)	(0.0198)	(0.0217)		(0.0476)	(0.0482)	(0.0417)		(0.0166)	(0.0183)	(0.0207)
Math ability		0.0086	0.0092	0.0002		0.0159	0.0140	-0.0058		0.0013	0.0014	0.0024
		(0.0070)	(0.0074)	(0.0084)		(0.0136)	(0.0130)	(0.0129)		(0.0066)	(0.0068)	(0.0077)
TLU		-0.0056	-0.0068	-0.0021		-0.0187*	-0.0187	-0.0143		0.0026	0.0015	0.0080
		(0.0056)	(0.0055)	(0.0072)		(0.0114)	(0.0115)	(0.0138)		(0.0054)	(0.0054)	(0.0071)
Purchased plot		0.0081	0.0080	0.0089		0.0064	0.0068	0.0325		0.0053	0.0050	0.0048
		(0.0126)	(0.0129)	(0.0107)		(0.0285)	(0.0279)	(0.0201)		(0.0124)	(0.0122)	(0.0112)
Asset index		0.0059	0.0052	-0.0003		0.0213*	0.0238**	0.0057		0.0040	0.0057	0.0002
		(0.0049)	(0.0046)	(0.0041)		(0.0109)	(0.0100)	(0.0091)		(0.0048)	(0.0048)	(0.0039)
Adult men		0.0275***	0.0281***	0.0198**		0.0265*	0.0345**	0.0163		0.0295**	0.0279**	0.0284***
		(0.0091)	(0.0098)	(0.0088)		(0.0137)	(0.0160)	(0.0139)		(0.0115)	(0.0109)	(0.0097)
Adult women		-0.0103	-0.0110	-0.0146		0.0050	0.0086	-0.0042		-0.0267**	-0.0247*	-0.0143
		(0.0096)	(0.0096)	(0.0101)		(0.0156)	(0.0161)	(0.0137)		(0.0120)	(0.0136)	(0.0113)
Household size		-0.0153**	-0.0135**	-0.0069		-0.0223**	-0.0286***	-0.0105		-0.0106	-0.0068	-0.0078
		(0.0061)	(0.0061)	(0.0049)		(0.0091)	(0.0106)	(0.0074)		(0.0066)	(0.0085)	(0.0063)
Soil nutrients (j)												
Nitrogen			0.0004	0.0034			-0.0184***	-0.0041			0.0061	0.0063**
			(0.0038)	(0.0029)			(0.0057)	(0.0053)			(0.0038)	(0.0028)
Phosphorus			0.0029	-0.0021			0.0043	0.0023			0.0001	-0.0062*
			(0.0037)	(0.0036)			(0.0070)	(0.0095)			(0.0041)	(0.0037)
Potassium			-0.0072	-0.0068			-0.0081	-0.0062			-0.0050	-0.0052
			(0.0061)	(0.0070)			(0.0116)	(0.0136)			(0.0050)	(0.0045)
Sulphur			-0.0040	-0.0016			0.0160	0.0097			-0.0146**	-0.0089
			(0.0081)	(0.0076)			(0.0173)	(0.0161)			(0.0069)	(0.0058)
Carbon			0.0048	-0.0058			-0.0004	-0.0111			0.0038	-0.0029
			(0.0063)	(0.0067)			(0.0106)	(0.0125)			(0.0076)	(0.0055)
Acidity (pH)			-0.0316	-0.0208			-0.1003	-0.0789			-0.0103	0.0023
			(0.0692)	(0.0692)			(0.1563)	(0.1358)			(0.0537)	(0.0488)
Farm management vars. (j)												
Ag. NGO participant				0.0096				-0.0202				0.0226**
				(0.0113)				(0.0186)				(0.0106)
Hired labor				0.0131*				0.0457***				-0.0032
				(0.0074)				(0.0143)				(0.0062)
Total hours worked per-hectare				-0.0043				0.0044				-0.0049
				(0.0110)				(0.0158)				(0.0098)
Seed payment				0.0059**				0.0062*				0.0053
				(0.0029)				(0.0037)				(0.0038)
Monocrop (1=yes)				0.0105				0.0134				0.0065
				(0.0084)				(0.0125)				(0.0098)
No organics used (1=yes)				0.0475				0.0687				0.0233
				(0.0337)				(0.0531)				(0.0383)
No DAP used (1=yes)				0.0795**				0.0766				0.0965**
				(0.0387)				(0.0634)				(0.0480)
No CAN used (1=yes)				-0.0490				-0.0729				-0.0582*
				(0.0306)				(0.0505)				(0.0302)
Total organics used per-hectare				0.0058				0.0073				0.0022
				(0.0045)				(0.0072)				(0.0052)
Total DAP used per-hectare				0.0127				0.0093				0.0160*
				(0.0078)				(0.0127)				(0.0095)
Total CAN used per-hectare				-0.0110*				-0.0146				-0.0143**
				(0.0061)				(0.0091)				(0.0058)
FEs (village, enumerator, (svy. month)	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
Observations	17,512	17,455	17,455	17,455	7,472	7,435	7,435	7,435	10,040	10,020	10,020	10,020
R-squared	0.126	0.276	0.279	0.411	0.113	0.288	0.303	0.439	0.186	0.341	0.352	0.515

2SLS regression results - OLS results available upon request. Standard errors clustered at the individual level. Non-binary variables in log form. ***p<0.01, **p<0.05, *p<0.1.

Table 4.8: Peer Effects - 0.33KM Cutoff

2SLS	**Full sample**				**Female$_i$**				**Male$_i$**			
	I	II	III	IV	V	VI	VII	VIII	IX	X	XI	XII
Value of per-hecatre maize yield (j)	0.0020	0.0131	0.0464*	0.0247	0.0023	0.0036	0.0432	0.0179	0.0012	0.0239	0.0411	0.0236
	(0.0014)	(0.0147)	(0.0256)	(0.0421)	(0.0029)	(0.0208)	(0.0456)	(0.0671)	(0.0011)	(0.0179)	(0.0278)	(0.0385)
Long rains season	0.6874***	0.7802***	0.7746***	0.3512**	0.6238**	0.7891***	0.7896***	0.3539	0.7148***	0.7544***	0.7580***	0.3511**
	(0.1383)	(0.1367)	(0.1368)	(0.1391)	(0.2492)	(0.2299)	(0.2312)	(0.2423)	(0.1618)	(0.1629)	(0.1626)	(0.1413)
Individual (i)												
Female	-0.2641***	-0.0917	-0.0907	-0.1060								
	(0.0891)	(0.0965)	(0.0964)	(0.0814)								
Years in village		0.0426	0.0478	0.0196		0.1220	0.0998	0.0764		-0.0670	-0.1084	-0.0744
		(0.0885)	(0.0875)	(0.0720)		(0.1412)	(0.1432)	(0.1209)		(0.0959)	(0.0947)	(0.0876)
Distance to plot (meters)		0.0402	0.0427	-0.0129		0.0801	0.0826	0.0140		0.0007	-0.0055	-0.0704
		(0.0556)	(0.0562)	(0.0571)		(0.0915)	(0.0952)	(0.0963)		(0.0669)	(0.0670)	(0.0693)
Age		-0.0754	-0.0944	0.0077		-0.6003	-0.5457	-0.2558		0.2521	0.2439	0.1728
		(0.1549)	(0.1601)	(0.1371)		(0.3714)	(0.3687)	(0.3088)		(0.1847)	(0.1914)	(0.1839)
Years of education		0.2713***	0.2659***	0.1514*		0.3682**	0.3266**	0.2007		0.0529	0.0658	-0.0162
		(0.0964)	(0.0952)	(0.0789)		(0.1490)	(0.1514)	(0.1401)		(0.0959)	(0.0929)	(0.0757)
Crop area		-0.5200***	-0.5176***	0.0216		-0.5000***	-0.4590***	0.0914		-0.5496***	-0.5735***	0.0110
		(0.0523)	(0.0537)	(0.0893)		(0.0738)	(0.0799)	(0.1466)		(0.0727)	(0.0740)	(0.1277)
Farmer		0.0506	0.0658	0.1617		0.0239	0.0638	0.2069		0.0650	0.1060	0.1517
		(0.1616)	(0.1580)	(0.1413)		(0.2771)	(0.2847)	(0.2710)		(0.1926)	(0.1835)	(0.1496)
Own plot		-0.1401	-0.1420	-0.0162		-0.9627***	-0.8535***	-0.3312		0.1741	0.1138	0.0701
		(0.3846)	(0.3741)	(0.2624)		(0.2576)	(0.2954)	(0.3074)		(0.4789)	(0.4833)	(0.3510)
Math ability		0.0547	0.0474	-0.0119		0.0274	0.0538	0.0466		0.0431	0.0155	-0.0648
		(0.1153)	(0.1167)	(0.1064)		(0.2058)	(0.2156)	(0.1980)		(0.1150)	(0.1210)	(0.1111)
TLU		0.1213*	0.1120*	0.0388		0.0754	0.0558	-0.1410		0.2053**	0.1982**	0.1256*
		(0.0661)	(0.0671)	(0.0666)		(0.1015)	(0.0997)	(0.1145)		(0.0898)	(0.0899)	(0.0756)
Purchased plot		0.1207	0.1109	0.1268		0.2320	0.2470	0.1777		0.0156	-0.0393	0.0339
		(0.1149)	(0.1144)	(0.1043)		(0.1700)	(0.1847)	(0.1785)		(0.1526)	(0.1481)	(0.1204)
Asset index		0.1848***	0.1889***	0.0468		0.2503*	0.2716*	0.1025		0.1400**	0.1578***	0.0274
		(0.0688)	(0.0678)	(0.0614)		(0.1432)	(0.1452)	(0.1206)		(0.0623)	(0.0607)	(0.0596)
Adult men		0.0026	0.0029	0.0383		-0.0111	-0.0334	0.1202		0.2690	0.2849*	0.1651
		(0.0972)	(0.0970)	(0.0898)		(0.1255)	(0.1289)	(0.1306)		(0.1688)	(0.1700)	(0.1625)
Adult women		0.1838	0.1715	0.0251		-0.0819	-0.0837	-0.1550		0.0897	0.0951	-0.0737
		(0.1179)	(0.1172)	(0.1057)		(0.1769)	(0.1811)	(0.1562)		(0.1588)	(0.1593)	(0.1424)
Household size		-0.2296***	-0.2305***	-0.1774**		-0.4013***	-0.3998***	-0.4014***		-0.1147	-0.1114	0.0272
		(0.0768)	(0.0765)	(0.0694)		(0.1506)	(0.1437)	(0.1315)		(0.0851)	(0.0862)	(0.0787)
Soil nutrients (i)												
Nitrogen			-0.0688	-0.0618			-0.0643	-0.0189			-0.0577	-0.0117
			(0.0458)	(0.0383)			(0.0797)	(0.0742)			(0.0619)	(0.0384)
Phosphorus			0.0076	0.0926**			0.1049	0.1708**			-0.0814	-0.0050
			(0.0467)	(0.0444)			(0.0811)	(0.0772)			(0.0540)	(0.0476)
Potassium			0.0194	0.1260**			0.1308	0.2156**			-0.1244**	-0.0343
			(0.0604)	(0.0523)			(0.0955)	(0.1000)			(0.0606)	(0.0590)
Sulphur			0.1065	0.0767			0.2077	0.1221			0.0533	-0.0031
			(0.0815)	(0.0649)			(0.1652)	(0.1468)			(0.0937)	(0.0621)
Carbon			0.0794	0.0019			0.1091	0.1581			0.0115	-0.0781
			(0.0748)	(0.0596)			(0.1358)	(0.1222)			(0.0888)	(0.0641)
Acidity (pH)			0.2577	-0.2597			0.0825	-0.2937			0.6341	0.3676
			(0.6062)	(0.5471)			(1.1662)	(0.9707)			(0.6918)	(0.6250)
Farm management vars. (i)												
Ag. NGO participant				0.0616				0.2570				-0.0179
				(0.1077)				(0.1897)				(0.1454)
Hired labor				0.2019**				0.1580				0.2261**
				(0.0828)				(0.1345)				(0.0885)
Total hours worked per-hectare				0.5082***				0.5543***				0.5341***
				(0.0876)				(0.1501)				(0.1169)
Seed payment				0.0509***				0.0280				0.0752***
				(0.0147)				(0.0188)				(0.0227)
Monocrop (1=yes)				-0.3308***				-0.1370				-0.5255***
				(0.0934)				(0.1471)				(0.1099)
No organics used (1=yes)				0.6640**				1.0539*				0.0563
				(0.3174)				(0.5814)				(0.2978)
No DAP used (1=yes)				0.3744				-0.0058				0.9312**
				(0.3018)				(0.4185)				(0.4077)
No CAN used (1=yes)				0.0157				0.6221				-0.5490
				(0.2840)				(0.4378)				(0.4243)
Total organics used per-hectare				0.0972**				0.1714**				0.0083
				(0.0404)				(0.0786)				(0.0371)
Total DAP used per-hectare				0.1316**				0.0823				0.2006***
				(0.0551)				(0.0867)				(0.0692)
Total CAN used per-hectare				0.0381				0.1942**				-0.0862
				(0.0511)				(0.0946)				(0.0651)
FEs (village, enumerator, (svy. month)	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
Observations	17,512	17,455	17,455	17,455	7,472	7,435	7,435	7,435	10,040	10,020	10,020	10,020
R-squared	0.1260	0.2758	0.2792	0.4111	0.1132	0.2888	0.3032	0.4399	0.1863	0.3416	0.3525	0.5155

2SLS regression results - OLS results available upon request. Standard errors clustered at the individual level. Non-binary variables in log form. ***p<0.01, **p<0.05, *p<0.1.

Table 4.8: Peer Effects Continued - 0.33KM Cutoff

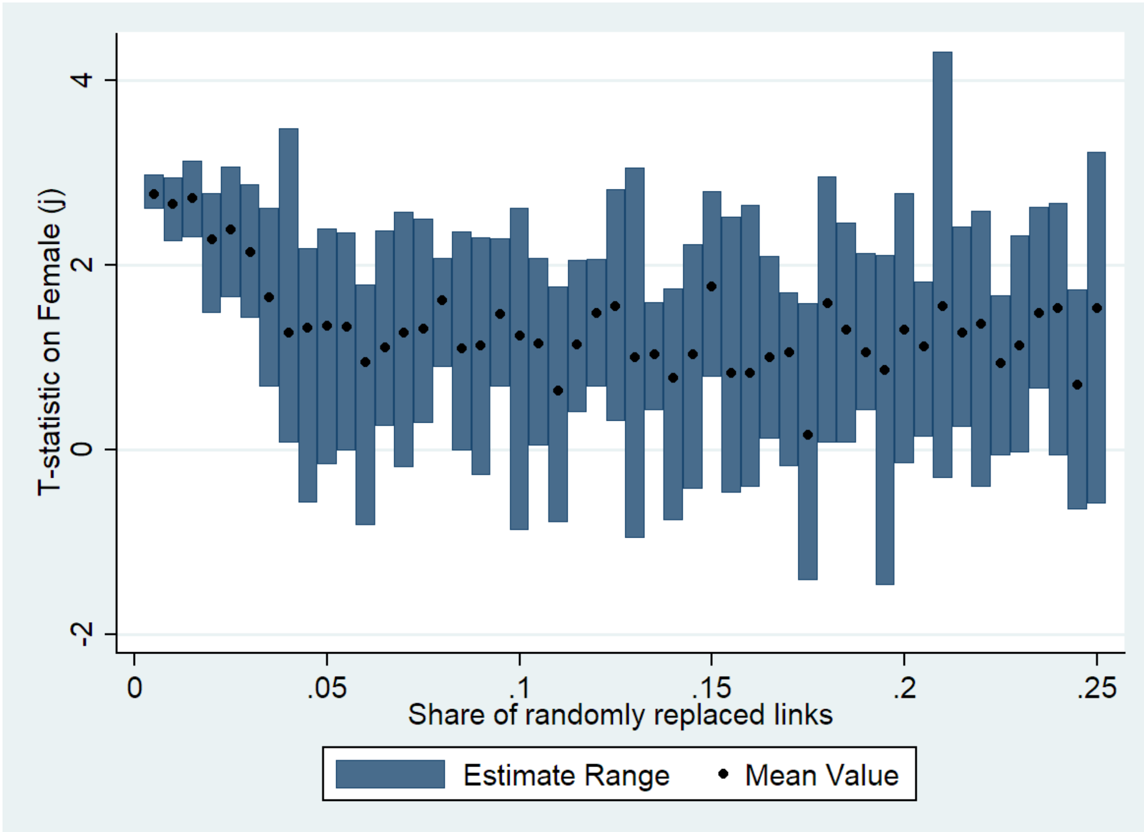
2SLS	**Full sample**				**Female$_i$**				**Male$_i$**			
	I	II	III	IV	V	VI	VII	VIII	IX	X	XI	XII
Individual (j)												
Female	0.0144	0.0527**	0.0625**	0.0483**	-0.0041	-0.0025	0.0324	-0.0027	0.0237	0.0731***	0.0630***	0.0698***
	(0.0163)	(0.0215)	(0.0247)	(0.0231)	(0.0256)	(0.0290)	(0.0423)	(0.0397)	(0.0168)	(0.0215)	(0.0217)	(0.0200)
Years in village		0.0203	0.0199	0.0080		0.0110	0.0089	-0.0179		0.0182	0.0131	0.0095
		(0.0151)	(0.0156)	(0.0133)		(0.0310)	(0.0306)	(0.0207)		(0.0168)	(0.0160)	(0.0146)
Distance to plot (meters)		-0.0153	-0.0229**	-0.0132		-0.0068	-0.0195	-0.0295**		-0.0159	-0.0219*	-0.0131
		(0.0106)	(0.0117)	(0.0115)		(0.0216)	(0.0189)	(0.0146)		(0.0109)	(0.0120)	(0.0098)
Age		-0.0319	-0.0285	-0.0000		0.0085	0.0372	0.0877*		-0.0630*	-0.0607	-0.0381
		(0.0299)	(0.0348)	(0.0336)		(0.0438)	(0.0633)	(0.0527)		(0.0371)	(0.0372)	(0.0358)
Years of education		-0.0145	-0.0221	-0.0124		-0.0460*	-0.0639**	-0.0150		-0.0012	-0.0017	-0.0049
		(0.0156)	(0.0170)	(0.0164)		(0.0267)	(0.0292)	(0.0249)		(0.0174)	(0.0187)	(0.0169)
Crop area		0.0007	0.0208	0.0101		-0.0095	0.0071	0.0165		0.0187	0.0293	0.0111
		(0.0199)	(0.0177)	(0.0251)		(0.0333)	(0.0229)	(0.0497)		(0.0198)	(0.0246)	(0.0154)
Farmer		0.0285	0.0440*	0.0298		0.0024	0.0123	0.0128		0.0421	0.0555*	0.0343
		(0.0254)	(0.0262)	(0.0267)		(0.0536)	(0.0535)	(0.0574)		(0.0281)	(0.0306)	(0.0230)
Own plot		-0.0134	0.0075	0.0131		0.0103	0.0820	0.0190		0.0009	0.0135	0.0337
		(0.0457)	(0.0429)	(0.0434)		(0.1287)	(0.1250)	(0.1126)		(0.0376)	(0.0405)	(0.0440)
Math ability		0.0285*	0.0297	0.0239		0.0620**	0.0570**	-0.0002		-0.0023	-0.0053	0.0179
		(0.0167)	(0.0182)	(0.0209)		(0.0264)	(0.0271)	(0.0319)		(0.0174)	(0.0183)	(0.0178)
TLU		-0.0188	-0.0201	-0.0098		-0.0532*	-0.0493	-0.0488		0.0005	0.0017	0.0215
		(0.0153)	(0.0157)	(0.0213)		(0.0306)	(0.0303)	(0.0396)		(0.0140)	(0.0141)	(0.0159)
Purchased plot		0.0131	0.0081	0.0032		-0.0214	-0.0229	0.0407		0.0271	0.0221	0.0126
		(0.0283)	(0.0294)	(0.0255)		(0.0614)	(0.0628)	(0.0488)		(0.0282)	(0.0279)	(0.0241)
Asset index		0.0200	0.0156	0.0004		0.0577**	0.0630**	0.0202		0.0108	0.0130	-0.0050
		(0.0137)	(0.0123)	(0.0103)		(0.0288)	(0.0272)	(0.0227)		(0.0122)	(0.0125)	(0.0083)
Adult men		0.0553**	0.0631**	0.0511**		0.0498	0.0712	0.0472		0.0644**	0.0607**	0.0632***
		(0.0234)	(0.0280)	(0.0258)		(0.0357)	(0.0479)	(0.0410)		(0.0263)	(0.0271)	(0.0237)
Adult women		-0.0024	-0.0067	-0.0115		0.0454	0.0527	0.0166		-0.0488**	-0.0374	-0.0117
		(0.0212)	(0.0218)	(0.0207)		(0.0357)	(0.0377)	(0.0277)		(0.0248)	(0.0268)	(0.0180)
Household size		-0.0363**	-0.0285*	-0.0184		-0.0540**	-0.0702***	-0.0391*		-0.0215	-0.0137	-0.0182
		(0.0143)	(0.0146)	(0.0127)		(0.0219)	(0.0252)	(0.0217)		(0.0150)	(0.0170)	(0.0138)
Soil nutrients (j)												
Nitrogen			-0.0059	0.0032			-0.0346***	-0.0113			0.0096	0.0157*
			(0.0102)	(0.0085)			(0.0131)	(0.0112)			(0.0112)	(0.0086)
Phosphorus			0.0083	-0.0045			0.0155	0.0069			-0.0001	-0.0167**
			(0.0090)	(0.0090)			(0.0175)	(0.0258)			(0.0104)	(0.0085)
Potassium			-0.0165	-0.0168			-0.0369	-0.0311			-0.0002	-0.0044
			(0.0160)	(0.0185)			(0.0303)	(0.0350)			(0.0105)	(0.0095)
Sulphur			0.0013	0.0020			0.0306	0.0066			-0.0233	-0.0146
			(0.0233)	(0.0223)			(0.0436)	(0.0395)			(0.0177)	(0.0149)
Carbon			0.0182	-0.0018			-0.0067	-0.0151			0.0241	0.0026
			(0.0160)	(0.0184)			(0.0220)	(0.0295)			(0.0174)	(0.0131)
Acidity (pH)			-0.2103	-0.1734			-0.1777	-0.2138			-0.1575	-0.0656
			(0.1716)	(0.1698)			(0.3557)	(0.3149)			(0.1202)	(0.1060)
Farm management vars. (j)												
Ag. NGO participant				0.0101				-0.0835*				0.0593**
				(0.0333)				(0.0487)				(0.0286)
Hired labor				0.0396**				0.0958***				0.0031
				(0.0191)				(0.0326)				(0.0133)
Total hours worked per-hectare				-0.0010				0.0196				-0.0141
				(0.0235)				(0.0380)				(0.0185)
Seed payment				0.0042				0.0123				0.0019
				(0.0053)				(0.0087)				(0.0043)
Monocrop (1=yes)				0.0037				0.0179				-0.0094
				(0.0168)				(0.0279)				(0.0161)
No organics used (1=yes)				0.0929				0.1864				-0.0319
				(0.0855)				(0.1573)				(0.0766)
No DAP used (1=yes)				0.1204				0.1582				0.1657**
				(0.0816)				(0.1336)				(0.0707)
No CAN used (1=yes)				-0.1193				-0.1878				-0.1156**
				(0.0760)				(0.1226)				(0.0539)
Total organics used per-hectare				0.0111				0.0198				-0.0047
				(0.0109)				(0.0215)				(0.0096)
Total DAP used per-hectare				0.0173				0.0201				0.0271*
				(0.0159)				(0.0261)				(0.0139)
Total CAN used per-hectare				-0.0236				-0.0349				-0.0271**
				(0.0145)				(0.0223)				(0.0107)
FEs (village, enumerator, (svy. month)	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
Observations	17,512	17,455	17,455	17,455	7,472	7,435	7,435	7,435	10,040	10,020	10,020	10,020
R-squared	0.1260	0.2758	0.2792	0.4111	0.1132	0.2888	0.3032	0.4399	0.1863	0.3416	0.3525	0.5155

2SLS regression results - OLS results available upon request. Standard errors clustered at the individual level. Non-binary variables in log form. ***p<0.01, **p<0.05, *p<0.1.

Table 4.A.1: First Stage Regressions

	Value of Per-hectare Maize Yield (j)		
Peer Cutoff	Full	1KM	0.33KM
Altitude (j)	1.3980***	1.3844***	1.467***
	(0.0862)	(0.0919)	(0.103)
Long rains season (j)	0.8048***	0.7048***	0.245***
	(0.0638)	(0.0570)	(0.028)
Years in village (j)	-0.0974*	-0.0946	-0.097*
	(0.0585)	(0.0581)	(0.053)
Distance to plot (j)	0.1306***	0.1509***	0.151***
	(0.0401)	(0.0373)	(0.038)
Female (j)	-0.2878***	-0.3781***	-0.333***
	(0.0817)	(0.0769)	(0.070)
Age (j)	0.0404	0.0738	-0.057
	(0.1346)	(0.1274)	(0.149)
Years of education (j)	0.2209***	0.1866***	0.150**
	(0.0637)	(0.0677)	(0.066)
Crop area (j)	-0.5387***	-0.5446***	-0.509***
	(0.0386)	(0.0338)	(0.045)
Farmer (j)	-0.1769**	-0.2315***	-0.120
	(0.0860)	(0.0864)	(0.137)
Own plot (j)	-0.4193*	-0.3982	-0.533*
	(0.2230)	(0.2568)	(0.286)
Math ability (j)	-0.0075	-0.0467	0.037
	(0.0967)	(0.0967)	(0.075)
TLU (j)	0.1145	0.1502**	0.066
	(0.0754)	(0.0743)	(0.055)
Plot purchased (j)	-0.0227	-0.0002	0.072
	(0.0918)	(0.0917)	(0.071)
Asset index (j)	0.1070**	0.1074**	0.093*
	(0.0427)	(0.0447)	(0.051)
Adult men (j)	-0.1978**	-0.2155***	-0.262***
	(0.0830)	(0.0742)	(0.076)
Adult women (j)	0.0990	0.1585	0.149
	(0.1156)	(0.1007)	(0.115)
Household size (j)	-0.0854	-0.1310	-0.090
	(0.1116)	(0.0962)	(0.119)
Constant	-1.2397***	-1.0671***	-0.333***
	(0.1480)	(0.1337)	(0.052)
F-stat	4454	6212	9485
Observations	17,483	17,483	17,483

Note: Non-binary variables in log form. Standard errors clustered at the individual level. ***p<0.01, **p<0.05, *p<0.1.



Note: Vertical bars represent t-stat range of ten estimates. Uses 1KM distance cutoff - corresponds to variable Female on Table 4.7, Column XI.

Figure 4.A.1: Robustness Check 1: Effect of Misspecified peer links on T-stat

Chapter 5: Summary and Conclusions

For most of those living in rural areas of Sub-Saharan Africa (SSA), agriculture is the predominant economic activity. Due to intensive agricultural activity, depleted soil nutrients, and soil erosion – the latter caused in part by unsustainable environmental practices and exacerbated by climate change – soils are generally highly degraded. This has often led to stagnant crop productivity, low rural incomes, and persistent underinvestment in effective agricultural inputs. The result is frequently a downward yield and income spiral, causing widespread rural poverty in many areas of the continent. Meanwhile, social inequities, often caused through customary peer relationships, leave those with less social capital with even lower levels of agricultural productivity. While there are many challenges for economic development in SSA, the foregoing chapters of this dissertation examine several obstacles confronting resource management and make policy recommendations that can potentially lead towards more sustainable agricultural practices, equitable social dynamics, and thereby enhance agricultural productivity. From the research reported in this dissertation, the lessons learned and resulting implications are discussed below. In addition, future areas for research are identified and are also discussed at the conclusion of the chapter.

5.1 Lessons and implications

In rural areas of western Kenya, there exists strong linkages among the environment, energy, and agricultural productivity. Farmers rely on fertile soils for crop production and on the broader environment, particularly trees and forests, as a primary source for household energy. Our data show that in the sample of rural households from western

Kenya surveyed in Chapter 2 of this dissertation, 98 percent use fuelwood as a primary energy source. Traditionally, fuelwood has often been collected off-farm from village commons or illegally from forest reserves. With increasing demands for fuelwood resulting from a growing population, this has led to substantial deforestation in many areas of western Kenya. Deforestation negatively affects agricultural productivity by altering biogeochemical processes, increasing erosion, and decreasing water retention in soils, all of which contribute to declining productivity. Agroforestry, or the planting of trees on the farm as a crop, is a potential solution to both rehabilitate the local environment, enhance soil health, and provide a source of household energy. Indeed, agroforestry has been adopted (to various degrees of intensity) by the majority of farmers in the study area. While only 15 percent of farmers in our sample have dedicated woodlots, 81 percent grow trees on their farm, primarily as a source of fuelwood.

As shown in Chapter 2, however, agroforestry may have limits in the extent to which it can substitute for nonrenewable fuelwood collected off-farm. In most studies on fuelwood collection in developing countries, fuelwood is considered an undifferentiated product. However, the persistence of off-farm fuelwood collection (practiced by 44 percent of households surveyed) in the study area here suggests continued demand for fuelwood collected off-farm, in addition to on-farm production and fuelwood purchased from the local market. In our analysis based on household survey data from western Kenya, we use maximum likelihood estimation of Heckman selection estimators to control for endogenous selection into fuelwood source groups. Then, using a two-stage least squares identification strategy, we calculate own-price and cross-price elasticities for fuelwood to measure the substitutability amongst the various fuelwood sources. As expected, the own-price elasticity of non-purchased fuelwood demand was negative and inelastic, with own-price elasticities ranging from -0.48 to -0.61. Cross-price elasticities were positive and remarkably low (0.02 to 0.24), indicating that increases in the opportunity cost of one fuelwood source does not translate into large increases in demand for fuelwood from another source.

Our analysis suggests that the most likely cause of this phenomenon is household adherence to traditional gender norms that limit the ability of men or women to engage in labor activities traditionally held by the opposite gender. In our data from western Kenya, 94 percent of primary fuelwood collectors off-farm are women, while 67 percent of on-farm woodlots are managed by men. As a result, we find that as the opportunity cost (shadow price) increases for off-farm fuelwood, women will tend to increase their time searching and collecting fuelwood rather than switching to another source (e.g., on-farm fuelwood production). Women tend to work both inside the home (cooking, caring for family members) and in on-farm agricultural work – known as the "double workday" of women (Kes and Swaminathan, 2006). By increasing the time burden rather than changing the source of fuelwood, these traditional gender norms have particularly negative ramifications for female well-being in the household.

In Chapter 4, we learn further that women are also disadvantaged in social capital compared to men. In this analysis, we define social capital as the competitive advantage that accrues to those with a more favorable location within the local social structure (Burt, 2000), proxied by centrality within a village social network. Using a network module in our household survey instrument, we mapped the social networks of nearly one thousand individuals in 21 villages. Using several measures of network centrality, and looking at the marginal results of female gender on network centrality over years lived in the village (using village fixed effects), we found that women are, on average, located more at the periphery of their village social networks. Our results show that it can take up to 30 years of living in a village for women to approach the level of network centrality that an average man has in a typical village. There are several likely causes of this finding. One is that social norms prevent women from participating in many types of social activity – for example, patronizing local bars – which inhibits the formation of social ties. Also, as already mentioned above, women are significantly busier than men and, as a result, have less time available to expand their social network. Finally, custom dictates that in many areas of western Kenya, men take brides from villages other than their own. Women join the household of their husband, which means that upon marriage, women effectively "reset," and reduce in number, their local network connections.

The relative periphery of women within social networks in these villages is a likely explanation for a subsequent finding of the analysis in Chapter 4: using a linear-in-means regression framework, we find that among men, increasing the share of female peers increases their own agricultural productivity, while among women, there is no gender-based effect coming from their peers. This result, robust across various specifications, is likely due to the information bargaining power differential between peers that have low levels of social capital (generally women), and those with higher levels (generally men). From this research, we learn that location within the social network, i.e., social capital, affects actual agricultural productivity in a measurable way. A focus on increasing the centrality of women in their networks, for example, by helping women forge new personal and business relationships in their individual localities, could lead to more equitable power relationships across peer networks and thus enhance information acquisition and improve agricultural productivity among women farmers.

Agricultural productivity in general, however, has stagnated in recent years in western Kenya due, in particular, to agricultural intensification and nutrient mining, leading to highly degraded soils. More effective fertilizer and input practices could help farmers enhance their soil fertility; however, farmers face uncertainties in applying the most effective input combinations due to heterogeneity in soil nutrient levels across farms and villages, and farmers' lack of information about their basic soil nutrient characteristics. Moreover, the government has traditionally emphasized inorganic fertilizers in its agricultural recommendations, while organic fertilizers are also necessary for effective soil nutrient recovery. In Chapter 3 of this dissertation, we report the results of experimental auctions we conducted to learn whether information provided to farmers in the form of soil test results and agricultural input recommendations affected farmer demand for these inputs. We used the SoilDoc kit, a relatively new product that can quickly and accurately measure the nutrients in a soil sample (Earth Institute, 2017). Using experimental auctions after Becker, DeGroot, and Marschak (1964) implemented in two rounds – both before and after providing farmers with information about their soils – we measured farmers' willingness to pay (WTP) for various agricultural inputs: DAP (diammonium phosphate), biochar, (vermi)compost, cow manure, and various combinations of these inputs (see Chapter 3 for details on these inputs). We divided participants into three different information treatment groups and a control group and used a triple difference estimation technique to identify the effects of alternative information transfers (see Chapter 3 for a detailed description of the auction methodology).

From this experiment, we learned that farmers are responsive to information derived from tests of their soils, but we discovered significant heterogeneity in the information effects by type of input and gender. First, we found that a recommendation to a farmer to use DAP (based on the results of his/her soil test) increased WTP for 2.5 Kg of DAP by about 62 KSh more than those in the counterfactual control group – a significant increase in WTP. However, we discovered the WTP results for organic inputs to be more nuanced: recommendations to use organic inputs increased WTP for those inputs by only about 18 KSh per unit auctioned compared to the counterfactual control group. However, when we include gender as an additional difference in the triple difference estimation – making the estimations quad differences – we find that the magnitude of the WTP increases for women compared to men in the counterfactual control. This gender difference we posit is likely due to the lower level of access that most women have to organic inputs compared to men in this area of Kenya. The relative lack of access to these inputs, we believe, causes women on average to bid higher for these inputs than men when given a recommendation to use them.

In addition, we analyzed in Treatment 2 whether comparisons of farmers' soils with the soil nutrient levels of anonymized peers influenced changes in WTP. Participants were shown charts that graphically illustrated their soil nutrient levels compared to other soils tested in the same village. As a result of this information, we found that farmers tended to shift their bids towards the mean: farmers who had below average soil nutrient levels increased their bids for agricultural inputs, while those who had above-average soil nutrient levels decreased their bids. This reversion to the mean, or "boomerang effect," has been observed by social psychologists and is an example of the negative effects that can arise from social norms (Clee and Wicklund, 1980). Interestingly, our Treatment 3, which included both comparisons with peers and input recommendations based on individualized soil tests, showed no significant boomerang effect, illustrating that information regarding optimal input choices can negate this deleterious social effect. Lessons from this experiment therefore inform researchers of the negative potential effects of using peer comparisons to influence policy outcomes.

To learn whether the SoilDoc soil testing system has the potential to be cost-effective in larger-scale settings, we conduct cost-benefit tests under various scenarios in Chapter 3. We find that in most scenarios, the SoilDoc tests lead to net benefits accruing to farmers – even when assuming that the farmers bear the full cost of the test. This is because, on average, the results from the SoilDoc test enable farmers to learn more optimal input combinations, which, if used, increase their crop yields. The largest benefits we found to be among those who were already using relatively expensive DAP fertilizer, but for whom the test showed that using this input on their soil was ineffective. This chapter thus shows that soil testing, whether through SoilDoc or other types of soil analysis (e.g. soil spectroscopy) can be an effective method to enable farmers to better optimize their input use, increase crop yields, and, presumably, increase incomes and decrease poverty levels.

Overall, foregoing Chapters 3 and 4 demonstrate both the promise and limitations of better information in stimulating economic development in rural SSA. As just noted, the results of Chapter 3 shows that information regarding farmers' soil nutrient levels influences their demands for agricultural inputs, leading to more effective optimization of their input choices. When farmers receive plot-specific soil nutrient information, for example, informing them that nitrogen inputs (e.g., DAP) should be used to increase crop productivity, this leads to significant increases in farmers' WTP for these inputs. Calculations further show that these benefits exceed the costs under most circumstances. The general implication is that soil testing can be an effective method to increase farmers' ability to rehabilitate their soils by increasing their optimization of agricultural inputs and reducing inefficiencies caused by the use of ineffective inputs for the farmers soil profile.

Chapter 4 illustrates that the structure of the social network is critically important for the diffusion of information in western Kenya. Those at the periphery of the village-level social networks have fewer contacts, less access to information, and as a result, have lower levels of bargaining power to obtain additional information. Our research in Chapter 4 demonstrates that women, on average, tend to be located more peripherally in village social networks. The result is that men are able to acquire information more inexpensively when their social network contains relatively more women. With more inexpensive information, agricultural productivity thus increases for men relative to women. These findings imply that economic development programs that target women may have limited success unless they address the structure of the network itself. Rather, measures that help women to establish additional social ties in their villages, thus improving their social capital and network centrality, may be one of the most effective means of increasing economic development among women in western Kenya.

These results echo, in a different context, the findings of Chapter 2, that information regarding agroforestry practices will not be fully effective unless household labor roles become more flexible. Cultural barriers exist that hinder the involvement of women in on-farm fuelwood production, i.e., managing woodlots, which effectively places a cap on the amount of on-farm fuelwood produced, thus limiting the renewable sources of fuelwood. Given that off-farm collection of fuelwood degrades the environment and places a time burden on women who are the primary fuelwood collectors off-farm in western Kenya, increasing the on-farm production of fuelwood is an important policy goal. Increasing labor substitutability in the workforce through actions encouraging female participation may therefore aid in establishing norms within the household for greater substitution of labor between genders.

A key overall finding in this dissertation is that cultural norms remain strong in SSA, often leading to inequities between genders and retarding economic growth. In Chapter 4, for example, we find that the practice of men taking brides from villages other than their own resets women's local social connections, leading to lower levels of social capital overall for women. We also find that women are the primary collectors of fuelwood off-farm, and rather than switching to on-farm fuelwood production in the wake of higher opportunity costs, the household increases the collection times for off-farm fuelwood, thus increasing women's time burden. In Chapter 3, we found that women behave differently in experimental auctions for agricultural inputs, bidding higher for organic agricultural inputs than men. Anecdotal evidence suggests that these traditional cultural norms may slowly be changing, indicating the promise for more equitable economic outcomes for men and women in the future.

5.2 Future research

The findings from the essays contained in this dissertation suggest several areas for future research. We learned in Chapter 2 that fuelwood is a differentiated product and that households are generally unwilling to substitute among sources. However, our data were insufficient to rigorously analyze the substitution between other sources of energy, such as charcoal and kerosene. The consumption of charcoal, in particular, has similar deleterious effects on the environment as off-farm fuelwood collection, as it also involves the use of trees as fuel. Yet, the use of kerosene, as a fossil fuel, carries its own negative environmental consequences. Discovering whether and how substitution readily occurs between fuelwood (in its various sources) and these other energy products could contribute to a more robust analysis of a household's overall energy use profile and the resulting trajectories for environmental outcomes.

Given the findings in Chapter 3, several additional avenues for research are possible. While we measured WTP for several agricultural inputs (DAP, biochar, vermicompost, and cow manure), many other inputs are available to farmers and can potentially improve soil fertility in western Kenya. For example, due to the number of farmers with acidic soil, agricultural lime would be very useful in improving crop productivity. However, few farmers use this product in the sample area. Projects that introduce farmers to agricultural lime and that conduct experimental auctions to identify farmers' demands in the study area could inform businesses, NGOs, and government agencies as to whether this input could be effectively introduced into the area, and whether subsidization would be a realistic policy.

The results from the cost-benefit analysis in Chapter 3 also point to the possibility of using soil testing on a wider scale. Further research is necessary to determine whether farmers are willing to pay for soil tests, given the high return that appears to result from their use. Projects such as that reported by Fabregas *et al.* (2014) have shown preliminary evidence that farmers are willing to pay for soil tests, but more research is needed prior to implementation of widespread soil testing. If further research finds that farmers will pay fully, or partially, for soil testing, then implementation of more extensive testing, accompanied by input recommendations, would likely contribute towards increases in agricultural yields in the area as farmers better optimize their input combinations.

A limitation of our experimental results in Chapter 3 is that they measured short-term changes in farmer behavior after a soil information transfer. We are also interested in learning whether this information leads to any long-term changes in a farmer's agricultural input usage and how these changes affect his/her crop productivity. To measure these effects, a panel dataset could be constructed to re-test farmers' soils. These data would enable testing as to how much information farmers retain from the first project visit and then measure any improvements in their soil nutrient levels. This would enable measurements of the long-term effects of the information transfer and more effectively gauge whether this type of program is effective, both agronomically and economically, in the long run. Collecting these data would also enable a deeper analysis of the peer network effects reported in Chapter 4. In this dissertation, only contemporaneous peer effects are measured. However, with panel data, an analysis of the lagged effects of a peer's agricultural productivity on a farmer's own productivity could strengthen the empirical results.

In addition, following the results from Chapter 4, further research is needed to determine whether interventions to increase female social linkages in these villages in fact lead to increases in agricultural productivity. Vasilaky and Leonard (2018) show in their study from Uganda that random pairings of women in a village increase crop yields. However, we need additional research of this nature in Kenya that specifically analyzes whether this kind of intervention improves the network centrality of women in the village. If so, this could demonstrate that this intervention strategy is useful in addressing the inequities in social capital discussed in Chapter 4 of this dissertation.

Overall, the chapters of this dissertation illustrate the complexities of economic development in Sub-Saharan Africa. Because the primary occupation in these areas is farming, the physical characteristics of the environment and the soils play a significant role in the economic condition of the majority of the rural population. Meanwhile, cultural practices and relationships among the members of village social networks have impacts on the diffusion of information that can increase agricultural productivity. While this dissertation addresses some of these issues, additional research on the intersection of agriculture, the environment, and village social networks is important if substantial economic gains are to be made.

References

Acharya, G. and E. Barbier (2002), 'Using domestic water analysis to value groundwater recharge in the Hadejia-Jama'are Floodplain, Northern Nigeria', *American Journal of Agricultural Economics* **84**(2): 415–426.

Aguilar, A., E. Carranza, M. Goldstein, T. Kilic, and G. Oseni (2015), 'Decomposition of gender differentials in agricultural productivity in Ethiopia', *Agricultural Economics* **46**(3): 311–334.

Aker, J. C. (2011), 'Dial "A" for Agriculture: A Review of Information and Communication Technologies for Agricultural Extension in Developing Countries', *Agricultural Economics* **42**(6): 631–647.

Akerlof, G. A. (1980), 'A Theory of Social Custom, of which Unemployment May be One Consequence', *The Quarterly Journal of Economics* **94**(4): 749–775.

Aleman, J. C., M. A. Jarzyna, and A. C. Staver (2018), 'Forest extent and deforestation in tropical Africa since', *Nature Ecology and Evolution* **2**(1): 26–33.

Allcott, H. (2011), 'Social Norms and Energy Conservation', *Journal of Public Economics* **95**(9-10): 1082–1095.

Amacher, G. S., W. F. Hyde, and B. R. Joshee (1993), 'Joint production and consumption in traditional households: Fuelwood and crop residues in two districts in Nepal', *Journal of Development Studies* **30**(1): 206–225.

Amacher, G. S., W. F. Hyde, and K. R. Kanel (1996), 'Household fuelwood demand and supply in Nepal's tarai and mid-hills: Choice between cash outlays and labor opportunity', *World Development* **24**(11): 1725–1736.

Amacher, G. S., W. F. Hyde, and K. R. Kanel (1999), 'Nepali fuelwood production and consumption: Regional and household distinctions, substitution and successful intervention', *Journal of Development Studies* **35**(4): 138–163.

Anand, B. N. and R. Shachar (2011), 'Advertising, the Matchmaker', *RAND Journal of Economics* **42**(2): 205–245.

Anderson, J. R. and G. Feder (2004), 'Agricultural Extension: Good Intentions and Hard Realities', *World Bank Research Observer* **19**(1): 41–60.

Angrist, J. D. and J.-S. Pischke (2008), *Mostly Harmless Econometrics: An Empiricist's Companion*, Princeton, NJ: Princeton University Press.

Antle, J. M., J. J. Stoorvogel, and R. O. Valdivia (2006), 'Multiple Equilibria, Soil Conservation Investments, and the Resilience of Agricultural Systems', *Environment and Development Economics* **11**(04): 477–491.

Ayres, I., S. Raseman, and A. Shih (2009), 'Evidence from Two Large Field Experiments that Peer Comparison Feedback Can Reduce Residential Energy Usage', Unpublished manuscript, National Bureau of Economic Research.

Bailis, R., M. Ezzati, and D. M. Kammen (2005), 'Mortality and greenhouse gas impacts of biomass and petroleum energy futures in Africa', *Science* **308**(April): 98–103.

Baland, J.-M., P. Bardhan, S. Das, D. Mookherjee, and R. Sarkar (2010), 'The environmental impact of poverty: evidence from firewood collection in rural Nepal.', *Economic Development and Cultural Change* **59**(1): 23–61.

Barrett, C. B. and L. E. M. Bevis (2015), 'The Self-Reinforcing Feedback between Low Soil Fertility and Chronic Poverty', *Nature Geoscience* **8**(12): 907–912.

Battese, G. E. (1997), 'Frontier Production-Functions and technical efficiency - A Survey of empirical applications in agricultural-economics', *Agricultural Economics* **7**: 185–208.

Bavelas, A. (1950), 'Communication Patterns in Task-Oriented Groups', *The Journal of the Acoustical Society of America* **22**(6): 725–730.

Becker, G., M. DeGroot, and J. Marschak (1964), 'Measuring Utility by a Single Response Sequential Method', *Behavioral Science* **9**(3): 226–232.

Bekunda, M. A., A. Bationo, and H. Ssali (1997), 'Soil fertility management in Africa: A review of selected research trials', in R. J. Buresh, P. A. Sanchez, and F. Calhoun (eds.), 'Replenishing Soil Fertility in Africa', Madison, WI: American Society of Agronomy, 63–80.

Bell, K. P. and N. E. Bockstael (2000), 'Applying the Generalized-Moments Estimation Approach To Spatial Problems Involving Microlevel Data', *The Review of Economics and Statistics* **82**(1): 72–82.

Berazneva, J., D. Lee, F. Place, and G. Jakubson (2017), 'Allocation and Valuation of Non-Marketed Crop Residues in Smallholder Agriculture: The Case of Maize Residues in Western Kenya', Unpublished manuscript, Dept. of Economics, Middlebury College.

Berazneva, J., L. McBride, M. Sheahan, and D. Güereña (2018), 'Empirical assessment of subjective and objective soil fertility metrics in east Africa: Implications for researchers and policy makers', *World Development* **105**: 367–382.

Binder, M. and D. Scrogin (1999), 'Labor force participation and household work of urban schoolchildren in Mexico: Characteristics and consequences', *Economic Development and Cultural Change* **48**(1): 123–154.

Bluffstone, R. A. (1995), 'The effect of labor market performance on deforestation in developing countries under open access: An example from rural Nepal', *Journal of Environmental Economics and Management* **29**(1): 42–63.

Bonacich, P. (1972), 'Factoring and weighting approaches to status scores and clique identification', *The Journal of Mathematical Sociology* **2**(1): 113–120.

Bonacich, P. (2007), 'Some unique properties of eigenvector centrality', *Social Networks* **29**(4): 555–564.

Bradley, P. and M. Huby (1993), *Woodfuel, Women and Woodlots*, Washington DC: The World Bank.

Bramoullé, Y., H. Djebbari, and B. Fortin (2009), 'Identification of peer effects through social networks', *Journal of Econometrics* **150**(1): 41–55.

Buck, L., J. Lassoie, and E. Fernandes (1999), *Agroforestry in Sustainable Agricultural Systems*, Washington DC: CRC Press.

Burke, P. J. and G. Dundas (2015), 'Female labor force participation and household dependence on biomass energy: Evidence from national longitudinal data', *World Development* **67**: 424–437.

Burke, W. J., T. S. Jayne, and J. R. Black (2017), 'Factors Explaining the Low and Variable Profitability of Fertilizer Application to Maize in Zambia', *Agricultural Economics* **48**: 115–126.

Burt, R. S. (1992), *Structural Holes*, Cambridge, MA: Harvard University Press.

Burt, R. S. (1998), 'The Gender of Social Capital', *Rationality and Society* **10**(1): 5–46.

Burt, R. S. (2000), 'The Network Structure of Social Capital', *Research in Organizational Behavior* **22**: 345–423.

Cameron, A. C., J. B. Gelbach, and D. L. Miller (2008), 'Bootstrap-Based Improvements for Inference with Clustered Errors', *Review of Economics and Statistics* **90**(3): 414–427.

Carletto, C., S. Gourlay, and P. Winters (2015), 'From guesstimates to GPStimates: Land area measurement and implications for agricultural analysis', *Journal of African Economies* **24**(5): 593–628.

Casaburi, L., M. Kremer, S. Mullainathan, and R. Ramrattan (2014), 'Harnessing ICT to Increase Agricultural Production: Evidence From Kenya', Unpublished manuscript, Stanford University.

Cavendish, W. (2002), 'Quantitative methods for estimating the economic value of resource use to rural households', in B. M. Campbell and M. K. Luckert (eds.), 'Uncovering the hidden harvest: Valuation methods for woodland and forest resources', London: Earthscan.

Chandrasekhar, A. G. and R. Lewis (2011), 'Econometrics of sampled networks', Unpublished manuscript, Stanford University.

Chavangi, N. and F. Adoyo (1993), 'Towards an Extension Strategy', in P. Bradley and M. Huby (eds.), 'Woodfuel, Women and Woodlots, Volume 2: The Kenya Woodfuel Development Programme', London and Basingstoke: Stockholm Environment Institute, 58–73.

Cialdini, R. B. (2003), 'Crafting Normative Messages to Protect the Environment', *Current Directions in Psychological Science* **12**(4): 105–109.

Clee, M. A. and R. A. Wicklund (1980), 'Consumer Behavior and Psychological Reactance', *Journal of Consumer Research* **6**(4): 389–405.

Cole, S. A. and A. N. Fernando (2012), 'The Value of Advice: Evidence from Mobile Phone-Based Agricultural Extension', Unpublished manuscript, Harvard Business School.

Coleman, J. S. (1988), 'Social Capital in the Creation of Human Capital', *American Journal of Sociology* **94**: 95–120.

Conley, T. G. and C. R. Udry (2010), 'Learning about a New Technology: Pineapple in Ghana', *The American Economic Review* **100**(1): 35–69.

Cooke, P. (2014), 'Where to Collect? Community Forest Management and Disadvantaged Households in Nepal', in E. J. Z. Robinson and R. Bluffstone (eds.), 'Forest Tenure Reform in Asia and Africa: Local Control for Improved Livelihoods, Forest Management, and Carbon Sequestration', EBSCO Publishing, chap. 10, 149–183.

Cooke, P., G. Köhlin, and W. F. Hyde (2008), 'Fuelwood, forests and community management – evidence from household studies', *Environment and Development Economics* **13**(01): 103–135.

Cooke, P. A. (1998a), 'Intrahousehold labor allocation responses to environmental good scarcity: A case study from the hills of Nepal', *Economic Development and Cultural Change* **46**(4): 807–830.

Cooke, P. A. (1998b), 'The effect of environmental good scarcity on own-farm labor allocation: The case of agricultural households in rural Nepal', *Environment and Development Economics* **3**(4): 443–469.

Cooper, P. (1979), 'The Association between Altitude, Environmental Variables, Maize Growth and Yields in Kenya', *The Journal of Agricultural Science* **93**(3): 635–649.

Corral, B. Y. C., X. Giné, A. Mahajan, and E. Seira (2017), 'Fertilizer Recommendations and In-Kind Subsidies Increase Update and Yields among Maize Farmers in Mexico', *Tech. rep.*, Feed the Future, USAID.

Cowan, R. and N. Jonard (2004), 'Network Structure and the Diffusion of Knowledge', *Journal of Economic Dynamics and Control* **28**(8): 1557–1575.

Crafford, J., R. Strohmaier, P. Muñoz, T. D. Oliveira, C. Lambrechts, M. Wilkinson, A. Burger, J. Bosch, A. Burger, and J. Bosch (2012), 'The role and contribution of montane forests and related ecosystem services to the Kenyan economy', *Tech. rep.*, UNEP, Nairobi, Kenya.

Crane-Droesch, A. (2017), 'Technology Diffusion, Outcome Variability, and Social Learning: Evidence from a Field Experiment in Kenya', *American Journal of Agricultural Economics* (January): 1–20, [Available at] http://academic.oup.com/ajae/advance-article/doi/10.1093/ajae/aax090/4781930.

Croppenstedt, A., M. Goldstein, and N. Rosas (2013), 'Gender and Agriculture: Inefficiencies, Segregation, and Low Productivity Traps', *World Bank Research Observer* **28**(1): 79–109.

Damte, A., S. F. Koch, and A. Mekonnen (2012), 'Coping with fuelwood scarcity: Household responses in rural Ethiopia', Unpublished manuscript, Environment for Development.

Dasgupta, P. (1997), 'Nutritional Status, the Capacity for Work, and Poverty Traps', *Journal of Econometrics* **77**: 5–37.

Dasgupta, P. and I. Serageldin (eds.) (1999), *Social Capital: A Multifaceted Perspective*, September, Washington D.C.: The World Bank.

Davis, K., E. Nkonya, E. Kato, D. A. Mekonnen, M. Odendo, R. Miiro, and J. Nkuba (2012), 'Impact of Farmer Field Schools on Agricultural Productivity and Poverty in East Africa', *World Development* **40**(2): 402–413.

Davis, L. R., B. P. Joyce, and M. R. Roelofs (2010), 'My money or yours: House money payment effects', *Experimental Economics* **13**(2): 189–205.

de Brauw, A. (2015), 'Gender, control, and crop choice in northern Mozambique', *Agricultural Economics* **46**(3): 435–448.

De Groote, H., S. C. Kimenju, and U. B. Morawetz (2011), 'Estimating Consumer Willingness to Pay for Food Quality with Experimental Auctions: the Case of Yellow versus Fortified Maize Meal in Kenya', *Agricultural Economics* **42**(1): 1–16.

De Groote, H., C. Narrod, S. C. Kimenju, C. Bett, R. P. B. Scott, M. M. Tiongco, and Z. M. Gitonga (2016), 'Measuring Rural Consumers' Willingness to Pay for Quality Labels Using Experimental Auctions: The Case of Aflatoxin-Free Maize in Kenya', *Agricultural Economics* **47**(1): 33–45.

DeGraff, D. S. and D. Levison (2009), 'Children's work and mothers' work—what is the connection?', *World Development* **37**(9): 1569–1587.

Dillon, B. and C. B. Barrett (2017), 'Agricultural factor markets in Sub-Saharan Africa: An updated view with formal tests for market failure', *Food Policy* **67**(November): 64–77.

Doss, C. R. and M. L. Morris (2001), 'How does Gender affect the Adoption of Agricultural Innovations?: The Case of Improved Maize Technology in Ghana', *Agricultural Economics* **25**(1): 27–39.

Drechsel, P., L. Gyiele, D. Kunze, and O. Cofie (2001), 'Population density, soil nutrient depletion, and economic growth in sub-Saharan Africa', *Ecological Economics* **38**(2): 251–258.

Duflo, E., M. Kremer, and J. Robinson (2011), 'Nudging Farmers to Use Fertilizer: Theory and Experimental Evidence from Kenya', *American Economic Review* **101**(October): 2350–2390.

Dupas, P. and J. Robinson (2013), 'Why Don't the Poor Save More? Evidence from Health Savings Experiments', *American Economic Review* **103**(4): 1138–1171.

Earth Institute (2017), 'SoilDoc: Bringing the Lab to the Field', Last Modified May 24, 2017. http://agriculture.columbia.edu/projects/agriculture/soildoc/.

Fabregas, R., M. Kremer, J. Robinson, and F. Schilbach (2014), 'What Institutions are Appropriate for Generating and Disseminating Local Agricultural Information?', Unpublished manuscript, Harvard Kennedy School.

Fafchamps, M. (2006), 'Development and social capital', *Journal of Development Studies* **42**(7): 1180–1198, [Available at] http://www.tandfonline.com/doi/abs/10.1080/00220380600884126.

Fafchamps, M. and B. Minten (2001), 'Social Capital and Agricultural Trade', *American Journal of Agricultural Economics* **83**(3): 680–685.

Feder, G., R. Just, and D. Zilberman (1985), 'Adoption of Agricultural Innovations in Developing Countries: A Survey', *Economic Development and Cultural Change* **33**(2): 255–298.

Feder, G., R. Murgai, and J. B. Quizon (2004a), 'Sending farmers back to school: The impact of farmer field schools in Indonesia', *Review of Agricultural Economics* **26**(1): 45–62.

Feder, G., R. Murgai, and J. B. Quizon (2004b), 'The Acquisition and Diffusion of Knowledge: The Case of Pest Management Training in Farmer Field Schools, Indonesia', *Journal of Agricultural Economics* **55**(2): 221–243.

Fishman, R., A. Kishore, Y. Rothler, P. S. Ward, S. Jha, and R. K. P. Singh (2016), 'Can Information Help Reduce Imbalanced Application of Fertilizers in India? Experimental Evidence from Bihar', Unpublished manuscript, IFPRI.

Food and Agriculture Organization (2011), 'Women in Agriculture: Closing the Gender Gap for Development', *Tech. rep.*, Food and Agriculture Organization, Rome, Italy, [Available at] http://www.fao.org/docrep/013/i2050e/i2050e.pdf.

Food and Agriculture Organization: UN (2010), 'Deforestation and Net Forest Area Change', [Available at] http://www.fao.org/forestry/30515/en.

Foster, A. D. and M. R. Rosenzweig (1995), 'Learning by Doing and Learning from Others: Human Capital and Technical Change in Agriculture', *Journal of Political Economy* **103**(6): 1176–1209.

Foster, A. D. and M. R. Rosenzweig (2010), 'Microeconomics of Technology Adoption', *Annual Review of Economics* **2**(1): 395–424.

Freeman, L. C. (1977), 'A Set of Measures of Centrality Based on Betweenness', *Sociometry* **40**(1): 35–41.

Frelat, R., S. Lopez-Ridaura, K. E. Giller, M. Herrero, S. Douxchamps, A. A. Djurfeldt, O. Erenstein, B. Henderson, M. Kassie, B. K. Paul, C. Rigolot, R. S. Ritzema, D. Rodriguez, P. J. A. van Asten, and M. T. van Wijk (2016), 'Drivers of Household Food Availability in sub-Saharan Africa Based on Big Data from Small Farms', *Proceedings of the National Academy of Sciences* **113**(2): 458–463.

García-Ruiz, J. M., S. Beguería, N. Lana-Renault, E. Nadal-Romero, and A. Cerdà (2017), 'Ongoing and Emerging Questions in Water Erosion Studies', *Land Degradation and Development* **28**(1): 5–21.

Gilligan, D. O., N. Kumar, S. McNiven, J. V. Meenakshi, and A. R. Quisumbing (2014), 'Bargaining Power and Biofortification: The Role of Gender in Adoption of Orange Sweet Potato in Uganda', Unpublished manuscript, IFPRI.

Giorgi, G. D., M. Pellizzari, and S. Redaelli (2010), 'Identification of Social Interactions through Partially Overlapping Peer Groups', *American Economic Journal: Applied Economics* **2**(2): 241–275.

Goldstein, N. J., R. B. Cialdini, and V. Griskevicius (2008), 'A Room with a Viewpoint: Using Social Norms to Motivate Environmental Conservation in Hotels', *Journal of Consumer Research* **35**(3): 472–482.

Grootaert, C. and T. van Bastelaer (eds.) (2002a), *The Role of Social Capital in Development Processes: An Empirical Assessment*, Cambridge, UK: Cambridge University Press.

Grootaert, C. and T. van Bastelaer (eds.) (2002b), *Understanding and Measuring Social Capital: A Multidisciplinary Tool for Practitioners*, Washington D.C.: The World Bank.

Güereña, D. T., J. Kimetu, S. Riha, H. Neufeldt, and J. Lehmann (2016), 'Maize productivity dynamics in response to mineral nutrient additions and legacy organic soil inputs of contrasting quality', *Field Crops Research* **188**: 113–120.

Hausman, J., G. Leonard, and J. D. Zona (1994), 'Competitive Analysis with Differenciated Products', *Annales d'Economie et de Statistique* **34**(261): 159–180.

Heckman, J. (1979), 'Sample selection bias as a specification error', *Econometrica* **47**(1): 153–161.

Heltberg, R., T. C. Arndt, and N. U. Sekhar (2000), 'Fuelwood consumption and forest eegradation: A household model for domestic energy substitution in rural India', *Land Economics* **76**(2): 213–232.

Henry, M., P. Tittonell, R. J. Manlay, M. Bernoux, A. Albrecht, and B. Vanlauwe (2009), 'Biodiversity, carbon stocks and sequestration potential in aboveground biomass in smallholder farming systems of western Kenya', *Agriculture, Ecosystems and Environment* **129**(1-3): 238–252.

Hoffmann, V., C. B. Barrett, and D. R. Just (2009), 'Do Free Goods Stick to Poor Households? Experimental Evidence on Insecticide Treated Bednets', *World Development* **37**(3): 607–617.

Hosonuma, N., M. Herold, V. De Sy, R. S. De Fries, M. Brockhaus, L. Verchot, A. Angelsen, and E. Romijn (2012), 'An assessment of deforestation and forest degradation drivers in developing countries', *Environmental Research Letters* **7**(4): 1–12.

International Energy Agency (2014), 'Africa Energy Outlook: Executive Summary', [Available at] http://www.worldenergyoutlook.org/africa.

Jack, A. L. and J. E. Thies (2006), 'Compost and Vermicompost as Amendments Promoting Soil Health', in N. Uphoff (ed.), 'Biological Approaches to Sustainable Soil Systems', CRC Press, 453–467.

Jacoby, H. G. (1993), 'Shadow wages and peasant family labor supply: An econometric application to the Peruvian Sierra', *The Review of Economic Studies* **60**(4): 903–921.

Johnson, J. P. and D. P. Myatt (2006), 'On the Simple Economics of Advertising, Marketing, and Product Design', *The American Economic Review* **96**(3): 756–784.

Joint Research Center: The European Commission (2013), 'Deforestation rates in Africa slowed down by 50%', [Available at] https://ec.europa.eu/jrc/en/news/deforestation-rates-africa-slowed-down-50-9929.

Just, R. E., D. Hueth, and A. Schmitz (2004), *The Welfare Effects of Public Policy: A Practical Approach to Project and Policy Evaluation*, North Hamptom, MA: Edward Elgar Publishing Company.

Karltun, E., M. Lemenih, and M. Tolera (2013), 'Comparing farmers' perception of soil fertility change with soil properties and crop performance in beseku, ethiopia', *Land Degradation and Development* **24**(3): 228–235.

Katungi, E., S. Edmeades, and M. Smale (2008), 'Gender, Social Capital and Information Exchange in Rural Uganda', *Journal of International Development* **20**(1): 35–52.

Katz, L. (1953), 'A New Status Index Derived from Sociometric Analysis', *Psychometrika* **18**(1).

Kenya Forest Service (2009), 'A guide to on-farm Eucalyptus growing in Kenya', [Available at] http://www.kenyaforestservice.org/documents/Eucalyptus%20guidelines%20%20Final%202.pdf.

Kes, A. and H. Swaminathan (2006), 'Gender and time poverty in Sub-Saharan Africa', in C. Blackden and Q. Wodon (eds.), 'Gender, Time Use, and Poverty in Sub-Saharan Africa', Washington DC: The World Bank, 13–38.

Key, N., E. Sadoulet, and A. D. Janvry (2000), 'Transactions costs and agricultural household supply Response', *American Journal of Agricultural Economics* **82**(May): 245–259.

Kihara, J., G. Nziguheba, S. Zingore, A. Coulibaly, A. Esilaba, V. Kabambe, S. Njoroge, C. Palm, and J. Huising (2016), 'Understanding Variability in Crop Response to Fertilizer and Amendments in sub-Saharan Africa', *Agriculture, Ecosystems and Environment* **229**: 1–12.

Kilic, T., A. Palacios-López, and M. Goldstein (2015), 'Caught in a Productivity Trap: A Distributional Perspective on Gender Differences in Malawian Agriculture', *World Development* **70**: 416–463.

Kiptot, E. and S. Franzel (2012), 'Gender and agroforestry in Africa: A review of women's participation', *Agroforestry Systems* **84**(1): 35–58.

Kohlin, G. and G. S. Amacher (2005), 'Welfare implications of community forest plantations in developing countries: The Orissa social forestry project', *American Journal of Agricultural Economics* **87**(4): 855–869.

Kondylis, F., V. Mueller, and J. Zhu (2017), 'Seeing is believing? Evidence from an extension network experiment', *Journal of Development Economics* **125**: 1–20.

Kramon, E. and D. N. Posner (2011), 'Kenya's new constitution', *Journal of Democracy* **22**(2): 0–21.

Krishnan, P. and M. Patnam (2014), 'Neighbors and Extension Agents in Ethiopia: Who Matters More for Technology Adoption?', *American Journal of Agricultural Economics* **96**(1): 308–327.

Kumar, N. and A. R. Quisumbing (2011), 'Does Social Capital Build Women' s Assets? The Long-Term Impacts of Group-Based and Individual Dissemination of Agricultural Technology in Bangladesh', *Journal of Development Effectiveness* **3**(2).

Lehmann, J. and S. Joseph (eds.) (2009), *Biochar for Environmental Management*, London; Sterling, VA: Earthscan.

Levison, D., K. S. Moe, and F. Knaul (2008), 'Marking time: An analysis of youth hours of work and study in urban Mexico', *Review of Development Economics* **12**(4): 751–763.

Liaukonyte, J., N. A. Streletskaya, and H. M. Kaiser (2015), 'The Long-Term Impact of Positive and Negative Information on Food Demand', *Canadian Journal of Agricultural Economics* **63**: 539–562.

Lim, S. S., T. Vos, A. D. Flaxman, G. Danaei, K. Shibuya, and H. Adair-Rohani (2012), 'A comparative risk assessment of burden of disease and injury attributable to 67 risk factors and risk factor clusters in 21 regions, 1990-2010: A systematic analysis for the Global Burden of Disease Study 2010', *The Lancet* **380**(9859): 2224–2260.

Liu, X., E. Patacchini, and Y. Zenou (2013), 'Peer Effects: Social Multiplier or Social Norms?', Unpublished manuscript, University of Colorado at Boulder.

Loureiro, M. L., W. J. Umberger, and S. Hine (2003), 'Testing the Initial Endowment Effect in Experimental Auctions', *Applied Economics Letters* **10**(5): 271–275.

Luke, N. and K. Munshi (2006), 'New Roles for Marriage in Urban Africa: Kinship Networks and the Labor Market in Kenya', *The Review of Economics and Statistics* **88**(2): 264–282, [Available at] http://www.jstor.org/stable/40042994.

Lybbert, T. J., N. Magnan, A. K. Bhargava, K. Gulati, and D. J. Spielman (2013), 'Farmers' Heterogeneous Valuation of Laser Land Leveling in Eastern Uttar Pradesh: An Experimental Auction to Inform Segmentation and Subsidy Strategies', *American Journal of Agricultural Economics* **95**(2): 339–345.

Ma, X. and G. Shi (2015), 'A Dynamic Adoption Model with Bayesian Learning: An Application to U.S. Soybean Farmers', *Agricultural Economics* **46**(1): 25–38.

Maathai, W. (1993), *The Green Belt Movement: Sharing the Approach and the Experience*, Lantern Books.

Maertens, A. (2017), 'Who cares what others think (or do)? Social learning and social pressures in cotton farming in India', *American Journal of Agricultural Economics* **99**(4): 988–1007.

Magnan, N., D. J. Spielman, K. Gulati, and T. J. Lybbert (2015), 'Information Networks among Women and Men and the Demand for an Agricultural Technology in India', Unpublished manuscript, IFPRI.

Manski, C. (2000), 'Economic Analysis of Social Interactions', *The Journal of Economic Perspectives* **14**(3): 115–136, [Available at] http://www.nber.org/papers/w7580.pdf.

Manski, C. F. (1993), 'Identification of Social Endogenous Effects: The Reflection Problem', *The Review of Economic Studies* **60**(3): 531–542.

Marenya, P. P. and C. B. Barrett (2009), 'State-conditional Fertilizer Yield Response on Western Kenyan Farms', *American Journal of Agricultural Economics* **91**(4): 991–1006.

Marenya, P. P., C. B. Barrett, and T. Gulick (2008), 'Farmers' Perceptions of Soil Fertility and Fertilizer Yield Response in Kenya', Unpublished manuscript, Dyson School of Applied Econ, Cornell University, [Available at] http://ssrn.com/abstract=1845546orhttp://dx.doi.org/10.2139/ssrn.1845546.

Masera, O. R., B. D. Saatkamp, and D. M. Kammen (2000), 'From linear fuel switching to multiple cooking strategies: A critique and alternative to the energy ladder model', *World Development* **28**(12): 2083–2103.

Mateete, B., S. Nteranya, and P. L. Woomer (2010), 'Restoring Soil Fertility in Sub-Sahara Africa', *Advances in Agronomy* **108**(10): 183–236.

Mathu, W. (2011), 'Forest platation and woodlots in Kenya', *African Forest Forum Working paper series* **1**(13): 48.

Mbow, C., P. Smith, D. Skole, L. Duguma, and M. Bustamante (2014), 'Achieving mitigation and adaptation to climate change through sustainable agroforestry practices in Africa', *Current Opinion in Environmental Sustainability* **6**(1): 8–14.

McPherson, M., L. Smith-Lovin, and J. M. Cook (2001), 'Birds of a Feather: Homophily in Social Networks', *Annual Review of Sociology* **27**(1): 415–444.

Meinzen-Dick, R., J. A. Behrman, L. Pandolfelli, A. Peterman, and A. R. Quisumbing (2014), 'Gender and Social Capital for Agricultural Development', in A. R. Quisumbing, R. Meinzen-Dick, T. L. Raney, A. Croppenstedt, J. A. Behrman, and A. Peterman (eds.), 'Gender in Agriculture: Closing the Knowledge Gap', Dordrecht: FAO, 235–266.

Meinzen-Dick, R. and M. Zwarteveen (2003), 'Gendered Participation in Water Management: Issues from Water Users' Associations in South Asia', in A. R. Quisumbing (ed.), 'Household Decisions, Gender, and Development', Washington D.C.: IFPRI, 153–158.

Mekonnen, A. (1999), 'Rural household biomass fuel production and consumption in Ethiopia: A case study', *Journal of Forest Economics* **5**(1): 69–97.

Mekonnen, D. A., N. Gerber, and J. A. Matz (2017), 'Gendered Social Networks, Agricultural Innovations, and Farm Productivity in Ethiopia', *World Development* **In Press**, [Available at] http://dx.doi.org/10.1016/j.worlddev.2017.04.020.

Michelson, H., N. Amuri, A. Harou, C. Magomba, A. Rose, J. Semoka, K. Tschirhart, and C. Palm (2017), 'Targeted Soil Recommendations May Improve Smallholder Yields in Tanzania', Unpublished manuscript, Feed the Future, USAID.

Miteva, D. A., R. A. Kramer, Z. S. Brown, and M. D. Smith (2017), 'Spatial patterns of market participation and resource extraction: Fuelwood collection in Northern Uganda', *American Journal of Agricultural Economics* **99**(4): 1008–1026.

Morawetz, U. B., H. De Groóte, and S. C. Kimenju (2011), 'Improving the Use of Experimental Auctions in Africa: Theory and Evidence', *Journal of Agricultural and Resource Economics* **36**(2): 263–279.

Moser, C. M. and C. B. Barrett (2006), 'The Complex Dynamics of Smallholder Technology Adoption: the Case of SRI in Madagascar', *Agricultural Economics* **35**(607): 373–388.

Mugure, A. and P. Oino (2013), 'Benefits of Agroforestry Farming Practices among Rural Households in Kenya: Experiences among Residents of Busia County', *International Journal of Science and Research* **2**(4): 442–449.

Munshi, K. (2004), 'Social Learning in a Heterogeneous Population: Technology Diffusion in the Indian Green Revolution', *Journal of Development Economics* **73**(1): 185–213.

Nakasone, E., M. Torero, and B. Minten (2014), 'The Power of Information: The ICT Revolution in Agricultural Development', *Annual Review of Resource Economics* **6**(1): 533–550.

Ndiritu, S. W., M. Kassie, and B. Shiferaw (2014), 'Are There Systematic Gender Differences in the Adoption of Sustainable Agricultural Intensification Practices? Evidence from Kenya', *Food Policy* **49**: 117–127.

Nepal, M., A. Nepal, and K. Grimsrud (2011), 'Unbelievable but Improved Cookstoves are Not Helpful in Reducing Firewood Demand in Nepal', *Environment and Development Economics* **16**(1): 1–23.

Ngetich, F. K., C. A. Shisanya, J. Mugwe, M. Mucheru-muna, and D. Mugendi (2012), 'The Potential of Organic and Inorganic Nutrient Sources in Sub-Saharan African Crop Farming Systems', in J. Whalen (ed.), 'Soil Fertility Improvement and Integrated Nutrient Management - A Global Prospective', 135–156.

Ngoze, S., S. Riha, J. Lehmann, L. Verchot, J. Kinyangi, D. Mbugua, and A. Pell (2008), 'Nutrient constraints to tropical agroecosystem productivity in long-term degrading soils', *Global Change Biology* **14**(12): 2810–2822.

Nkonya, E., T. Johnson, H. Y. Kwon, and E. Kato (2016), *Economics of Land Degradation and Improvement – A Global Assessment for Sustainable Development*, [Available at] http://link.springer.com/10.1007/978-3-319-19168-3.

Oballa, P., P. K. A. Konuche, M. N. Muchiri, and B. N. Kigomo (2010), 'Facts on growing and use of eucalyptus in Kenya', *Tech. Rep. September*, Kenya Forestry Research Institute, Nairobi.

Olsen, R. J. (1980), 'A least squares correction for selectivity bias', *Econometrica* **48**(7): 1815–1820.

Oseni, G., P. Corral, M. Goldstein, and P. Winters (2015), 'Explaining Gender Differentials in Agricultural Production in Nigeria', *Agricultural Economics* **46**(3): 285–310.

Palacios-Lopez, A., L. Christiaensen, and T. Kilic (2017), 'How much of the labor in African agriculture is provided by women?', *Food Policy* **67**: 52–63, [Available at] http://dx.doi.org/10.1016/j.foodpol.2016.09.017.

Palm, C. A., R. J. K. Myers, and S. M. Nandwa (1997), 'Combined use of organic and inorganic nutrient sources for soil fertility maintenance and replenishment', in R. J. Buresh, P. A. Sanchez, and F. Calhoun (eds.), 'Replenishing Soil Fertility in Africa', Madison, WI: American Society of Agronomy, 193–218.

Palmer, C. and J. MacGregor (2009), 'Fuelwood scarcity, energy substitution, and rural livelihoods in Namibia', *Environment and Development Economics* **14**(06): 693–715.

Patacchini, E. and G. Venanzoni (2014), 'Peer effects in the demand for housing quality', *Journal of Urban Economics* **83**: 6–17.

Pattanayak, S. K., E. O. Sills, and R. A. Kramer (2004), 'Seeing the forest for the fuel', *Environment and Development Economics* **9**(2): 155–179.

Peterman, A., A. Quisumbing, J. Behrman, and E. Nkonya (2011), 'Understanding the Complexities Surrounding Gender Differences in Agricultural Productivity in Nigeria and Uganda', *Journal of Development Studies* **47**(10): 1482–1509, [Available at] http://www.tandfonline.com/doi/abs/10.1080/00220388.2010.536222.

Place, F., C. B. Barrett, H. Freeman, J. J. Ramisch, and B. Vanlauwe (2003), 'Prospects for Integrated Soil Fertility Management using Organic and Inorganic Inputs: Evidence from Smallholder African Agricultural Systems', *Food Policy* **28**(4): 365–378.

Putnam, R. D. (1993), *Making Democracy Work: Civic Institutions in Modern Italy*, Princeton: Princeton University Press.

Putnam, R. D. (2000), *Bowling Alone*, New York: Simon and Schuster.

Republic of Kenya: Ministry of Environment (2014), 'Forest Policy 2014', .

Rickard, B. J., J. Liaukonyte, H. M. Kaiser, and T. J. Richards (2011), 'Consumer Response to Commodity-Specific and Broad-Based Promotion Programs for Fruits and Vegetables', *American Journal of Agricultural Economics* **93**(5): 1312–1327.

Rogers, E. M. (1995), *Diffusion of Innovations*, New York: Free Press, 4th edn.

Roobroeck, D., G. Nziguheba, B. Vanlauwe, and C. Palm (2017), 'Fertilizer use in maize and soybean production across African agro-ecosystems: Assessing prevalence of non-responsiveness and understanding constraints to input use.', Unpublished manuscript, International Institute of Tropical Agriculture.

Sahn, D. E. and D. Stifel (2003), 'Exploring alternative measures of welfare in the absence of expenditure data', *Review of Income and Wealth* **49**(4): 463–489.

Sanchez, P. A. (2002), 'Soil Fertility and Hunger in Africa.', *Science* **295**(5562): 2019–2020.

Scherr, S. J. (1995), 'Economic factors in farmer adoption of agroforestry: patterns observed in Western Kenya', *World Development* **23**(5): 787–804.

Schultz, P. W., J. M. Nolan, R. B. Cialdini, N. J. Goldstein, and V. Griskevicius (2007), 'The Constructive, Destructive, and Reconstructive Power of Social Norms', *Psychological Science* **18**(5): 429–434.

Sheahan, M., R. Black, and T. Jayne (2013), 'Are Kenyan Farmers Under-utilizing Fertilizer? Implications for Input Intensification Strategies and Research', *Food Policy* **41**: 39–52.

Shogren, J. F. (2005), 'Experimental Methods and Valuation', in K.-G. Maler and J. Vincent (eds.), 'Handbook of Environmental Economics', North Holland, vol. 2, 969–1027.

Simonit, S. and C. Perrings (2011), 'Sustainability and the value of the 'regulating' services: Wetlands and water quality in Lake Victoria', *Ecological Economics* **70**(6): 1189–1199.

Skoufias, E. (1994), 'Using shadow wages to estimate labor supply of agricultural households', *American Journal of Agricultural Economics* **76**(2): 215–227.

Slavchevska, V. (2015), 'Gender Differences in Agricultural Productivity: The Case of Tanzania', *Agricultural Economics* **46**(3): 335–355.

Solomon, D., J. Lehmann, J. Kinyangi, W. Amelung, I. Lobe, A. Pell, S. Riha, S. Ngoze, L. Verchot, D. Mbugua, J. Skjemstad, and T. Schäfer (2007), 'Long-term impacts of anthropogenic perturbations on dynamics and speciation of organic carbon in tropical forest and subtropical grassland ecosystems', *Global Change Biology* **13**(2): 511–530.

St. Clair, P. C. (2016), 'Community forest management, gender and fuelwood collection in rural Nepal', *Journal of Forest Economics* **24**: 52–71.

Steur, H. D., J. Buysse, S. Feng, and X. Gellynck (2013), 'Role of Information on Consumers' Willingness-to-pay for Genetically-modified Rice with Health Benefits: An Application to China', *Asian Economic Journal* **27**(4): 391–408.

Strauss, J. (1986), 'The Theory of Comparative Statics', in I. Singh, L. Squire, and J. Strauss (eds.), 'Agricultural Household Models', Baltimore: Johns Hopkins University Press, 71–91.

Thaler, R. H. and E. J. Johnson (1990), ‘Gambling with the House Money and Trying to Break Even: The Effects of Prior Outcomes on Risky Choice’, *Management Science* **36**(February 2015): 643–660.

The World Bank (2018), ‘PovcalNet’, [Available at] http://iresearch.worldbank.org/PovcalNet/povDuplicateWB.aspx.

Tigabu, A. (2017), ‘Factors Associated with Sustained use of Improved Solid Fuel Cookstoves: A Case Study from Kenya’, *Energy for Sustainable Development* **41**: 81–87.

Tittonell, P., a. Muriuki, C. J. Klapwijk, K. D. Shepherd, R. Coe, and B. Vanlauwe (2013), ‘Soil Heterogeneity and Soil Fertility Gradients in Smallholder Farms of the East African Highlands’, *Soil Science Society of America Journal* **77**(2): 525.

Tittonell, P., K. D. Shepherd, B. Vanlauwe, and K. E. Giller (2008a), ‘Unravelling the Effects of Soil and Crop Management on Maize Productivity in Smallholder Agricultural Systems of Western Kenya-An Application of Classification and Regression Tree Analysis’, *Agriculture, Ecosystems and Environment* **123**(1-3): 137–150.

Tittonell, P., B. Vanlauwe, M. Corbeels, and K. E. Giller (2008b), ‘Yield gaps, Nutrient use Efficiencies and Response to Fertilisers by Maize across Heterogeneous Smallholder Farms of Western Kenya’, *Plant and Soil* **313**(1-2): 19–37.

Tittonell, P., B. Vanlauwe, P. A. Leffelaar, E. C. Rowe, and K. E. Giller (2005), ‘Exploring Diversity in Soil Fertility Management of Smallholder Farms in Western Kenya: I. Heterogeneity at Region and Farm Scale’, *Agriculture, Ecosystems and Environment* **110**(3-4): 149–165.

Tjernström, E. (2016), ‘Signals, Similarity and Seeds: Social Learning in the Presence of Imperfect Information and Heterogeneity’, Unpublished manuscript, University of Wisconsin, Madison.

Torres-Rojas, D., J. Lehmann, P. Hobbs, S. Joseph, and H. Neufeldt (2011), ‘Biomass Availability, Energy Consumption and Biochar Production in Rural Households of Western Kenya’, *Biomass and Bioenergy* **35**(8): 3537–3546.

Tully, K., C. Sullivan, R. Weil, and P. Sanchez (2015), ‘The State of Soil Degradation in sub-Saharan Africa: Baselines, Trajectories, and Solutions’, *Sustainability* **7**(6): 6523–6552.

Udry, C. (1996), ‘Gender, Agricultural Production, and the Theory of the Household’, *Journal of Political Economy* **104**(5): 1010–1046.

Udry, C., J. Hoddinott, H. Alderman, and L. Haddad (1995), ‘Gender Differentials in Farm Productivity: Implications for Household Efficiency and Agricultural Policy’, *Food Policy* **20**(5): 407–423.

Unruh, J., R. Houghton, and P. Lefebvre (1993), 'Carbon storage in agroforestry: Am estimate for Sub-Saharan Africa', *Climate Research* **3**: 39–52.

Van Der Kroon, B., R. Brouwer, and P. J. H. Van Beukering (2013), 'The energy ladder: Theoretical myth or empirical truth? Results from a meta-analysis', *Renewable and Sustainable Energy Reviews* **20**: 504–513.

Vanlauwe, B., K. Aihou, S. Aman, E. N. O. Iwuafor, B. K. Tossah, J. Diels, N. Sanginga, O. Lyasse, R. Merckx, and J. Deckers (2001), 'Maize yield as affected by organic inputs and urea in the West African moist savanna', *Agronomy Journal* **93**(6): 1191–99.

Vanlauwe, B., J. Diels, K. Aihou, E. Iwuafor, O. Lyasse, N. Sanginga, and R. Merckx (2002), 'Direct Interactions between N Fertilizer and Organic Matter: Evidence from Trials with N-labelled Fertilizer', in B. Vanlauwe, J. Diels, N. Sanginga, and R. Merckx (eds.), 'Integrated Plant Nutrient Management in Sub-Saharan Africa: From Concept to Practice', Wallingford, Oxon: CABI, 173–184.

Vanlauwe, B., J. Kihara, P. Chivenge, P. Pypers, R. Coe, and J. Six (2011), 'Agronomic Use Efficiency of N Fertilizer in Maize-Based Systems in sub-Saharan Africa within the Context of Integrated Soil Fertility Management', *Plant Soil* **339**(1): 35–50.

Vasilaky, K. and K. L. Leonard (2018), 'As Good as the Networks They Keep?: Improving Farmers' Social Networks via Randomized Information Exchange in Rural Uganda', *Economic Development and Cultural Change* **Forthcomin**.

Woolcock, M. and D. Narayan (2000), 'Social Capital: Implications for Development Theory', *World Bank Research Observer* **15**(2): 225–249.

Wooldridge, J. (2002), *Econometric Analysis of Cross Section and Panel Data*, Cambridge, MA: The MIT Press.

Wopereis, M. C. S., A. Tamélokpo, K. Ezui, D. Gnakpénou, B. Fofana, and H. Breman (2006), 'Mineral Fertilizer Management of Maize on Farmer Fields Differing in Organic Inputs in the West African Savanna', *Field Crops Research* **96**(2-3): 355–362.

World Bank (2015), 'World Bank Group Gender Strategy (FY16-23): Gender Equality, Poverty Reduction and Inclusive Growth', *Tech. rep.*, The World Bank.

World Meteorological Organization (2010), 'Climate and Land Degradation', [Available at] http://www.wamis.org/agm/pubs/brochures/wmo989e.pdf.

World Wide Fund for Nature (2007), 'Dirty thirty: Ranking of the most polluting power plants in Europe', [Available at] http://assets.panda.org/downloads/dirty30rankingfinal260905.pdf.

Yasuoka, J. and R. Levins (2007), 'Impact of deforestation and agricultural development on anopheline ecology and malaria epidemiology', *American Journal of Tropical Medicine and Hygene* **76**(3): 450–460.

Zapata, S. D. and C. E. Carpio (2014), 'The Theoretical Structure of Producer Willingness to Pay Estimates', *Agricultural Economics* **45**(5): 613–623.

Zingore, S., J. Mutegi, B. Agesa, L. Tamene, and J. Kihara (2015), 'Soil Degradation in sub-Saharan Africa and Crop Production Options for Soil Rehabilitation', *Better Crops with Plant Food* **99**(1): 65–67.

Zizzo, D. J. (2010), 'Experimenter Demand Effects in Economic Experiments', *Experimental Economics* **13**(1): 75–98.