Environment and Development: Essays on the Link Between Household Welfare and the Environment in Developing Countries

A DISSERTATION
SUBMITTED TO THE FACULTY OF THE GRADUATE SCHOOL
OF THE UNIVERSITY OF MINNESOTA
BY

Martha H. Rogers

IN PARTIAL FULFILLMENT OF THE REQUIREMENTS
FOR THE DEGREE OF
Doctor of Philosophy

Paul Glewwe, Stephen Polasky

July, 2014

Acknowledgements

Five years ago, I could not have imagined the amount of personal and professional growth that would take place over the course of my graduate work. I could not have gotten through this long and winding road without the invaluable help and support of my advisors, professors, friends, and family. Thank you all for being willing to be a part of this journey.

First and foremost, I owe my deepest gratitude to my advisors, Paul Glewwe and Stephen Polasky. Their seemingly endless willingness to listen to my good (and bad) ideas and constant encouragement has motivated me to write a dissertation that I am truly proud of. I am also indebted to the rest of my committee members, Jay Coggins, Joe Knight, and Deborah Levison, as well as Gerard McCullough for their thoughtful comments and belief in me as an academic researcher.

I would also like to acknowledge the financial support that I received from the University of Minnesota's Interdisciplinary Center for the Study of Global Change, William W. Stout Fellowship, Consortium on Law and Values, and Healthy Food, Healthy Lives Institute. This financial support has been invaluable in providing me the opportunities to pursue research questions in my thesis that are of the most interest to me.

Finally, my life over the last five years has been greatly enriched by my community of family and friends. Helen Markelova, Kari Heerman, and Alison Sexton all added much needed sanity to thesis writing and were an endless source of entertainment, encouragement, and support. Thank you to my brother and sister, Matt and Maggie Rogers, and the Pacas family for being constant sources of motivation. And, last but not least, I am forever thankful to my parents, Fred Rogers and Jenny Hartley, and husband, José Pacas. Words cannot express how grateful I am for your support, patience, and nourishment these last five years.

Dedication

To my husband, José Pacas, who has encouraged and supported me every step of the way. We did it.

Abstract

In this dissertation, I present three methods of evaluating local populations' interactions with their natural environments using household-level data from Tanzania. To date, little effort has been made to evaluate the non-market benefits of natural resources for local populations and this dissertation makes important contributions to this budding research area. First, I apply a travel cost model and estimate that households in Kagera, Tanzania are willing to pay approximately $200 per year (2012 U.S. dollars) for local community forests access, a value equal to roughly 25 percent of annual total household expenditures. Second, using a long-term panel data set I estimate that an additional hour required to collect firewood when a child is young translates into $475 (2010 USD) in lost earnings over 30 years, roughly 1.7 percent of income. Finally, I show evidence of significant interdependencies between a household's agricultural production and food consumption decisions. This inter-dependency implies that programs aimed at environmental conservation through agricultural intensification may have important unintended consequences on a household's food consumption and subsequent micronutrient levels. In sum, the results in this dissertation indicate that households in Tanzania interact with their environments in complex ways and receive significant non-market benefits from natural resources.

Contents

List of Tables

List of Figures

Chapter 1

Introduction

The welfare of rural households in developing countries is dependent upon both the local natural environment and local markets. Not only does the local natural environment and local markets affect the welfare of these households but, these households also affect the environment, and global public goods, with their actions. For example, forests absorb 25 percent of annual global carbon dioxide emissions (Gillis, 2011). But, in Africa, 90 percent of the continent's population uses firewood for cooking and, in contrast to other parts of the world, human activity is the leading cause of deforestation (Agyei, 1998). Effective management of the natural environment requires that we measure benefits and understand how the ways in which these rural households interact and rely on their local environments. These benefit estimates are key inputs into cost-benefit analyses of conservation and climate mitigation policies, such as payment for ecosystem services programs.

Despite many policymakers' understanding of these different interests, very few economic studies have used household socio-economic data to quantify the local non-market benefits that households in developing countries derive from natural resources. Indeed, the need for more economic papers addressing this research topic was recently recognized in the working paper by Greenstone and Jack (2013). In this dissertation, I use household-level data to estimate the ways in which local populations interact with their natural environments. In Chapters 2 and 3, I estimate two non-market dollar value benefits of forest access: (1) the consumptive value of forests and (2) the loss in human capital that results from reduced forest access and increased firewood collection time. In Chapter 4, I show that food market imperfections affect households' food consumption decisions. Ultimately, an understanding of the varying ways in which local populations interact and use their natural environment is important for promoting sustainable development policies that manage both the current and future stock of natural resources.

In the first essay, I estimate the benefit that households in Kagera, Tanzania, derive from access to local community forests. Specifically, I combine an environmental economics travel cost framework with a traditional agricultural household model to derive household demand for weekly firewood collection trips in the presence of constrained labor markets. I then estimate household demand for weekly firewood collection trips as a function of the time costs associated with firewood collection. My results show that households would be willing to pay approximately \$200 per year for access to their local community forests, or roughly 25 percent of their total annual expenditures. This empirically-derived benefit is, on average, twice as large as the household-reported annual value of firewood consumption, which indicates that direct cost estimates may undervalue the full opportunity costs associated with forest conservation programs like United Nations' Reducing Emissions from Deforestation and Forest Degradation program. That is, the welfare loss to households when conservation programs restrict their use of local ecosystem services may be greatly underestimated.

In the second essay, I test for the effects of forest access, measured as firewood collection trip time, on human capital formation in both the short- and long-run. Tanzania's heavy reliance on forests (95 percent of households report using firewood as their primary source of cooking fuel) coupled with Tanzania's high rate of deforestation (the country has lost 81,000 km^2 of forests over the last 20 years (World Bank, 2010)) implies that children could be forced to compromise their school attendance as access to local forests becomes more limited. In the short-term, I find that a one hour increase in firewood collection trip time results in a child spending 25 minutes less in school a week, regardless of whether the child is collecting firewood him or herself. Over the long-term, this reduction in weekly school attendance due to a one hour increase in firewood collection time when the child is young translates into a child completing one-fifth fewer grades 19 years later. Previous studies have estimated an 8 percent annual return to education in Tanzania (Psacharopoulos and Patrinos, 2004) and, using this rate, a one hour increase in firewood collection time implies a 1.7 percent reduction in annual income when the child is older, or a net present value of $475 in 2010 USD over the course of 30 years. Even though the human capital costs of forest conservation are not as large in magnitude as the consumptive values estimated in the first essay, they are still non-trivial amounts when aggregated to the population level– 3 million cumulative years of lost education if all 15 million rural children in Tanzania were affected.

Finally, in the third essay, I change focus from forest access and forest use to a household's food consumption and agricultural production. Agriculture accounts for 10-12 percent of total greenhouse gas emissions (Stocker et al., 2013) and scientists estimate that agricultural production will need to double in order to meet the needs of a growing population (Foley et al., 2011). The most popular method of increasing agricultural production is through agricultural intensification, particularly in Africa where agricultural production is below its estimated potential (Foley et al., 2011). Indeed, agricultural development programs emphasize agricultural intensification, the commercialization of crops and, consequently, the income generated from agricultural sales (Barrett, 2008). But, if food markets are imperfect then any change in a household's agricultural production will also affect its food, and subsequent nutrient, consumption. Consequently, the recognized productivity gap in African agriculture may not be so much a gap as much as a rational response by households to local markets and a need to meet household nutrient needs.

To study this relationship between household food consumption and production, I test whether household demand for food is affected by cash crop prices. If food markets work perfectly, an increase in the price of cash crops indirectly affects household food consumption through a positive effect on household income; as the price of cash crops rises, farmers will produce more cash crops and less food crops. But, because food markets work perfectly, home produced food and purchased food are perfect substitutes and households can supplement any drop in home-consumed food with purchased food. With imperfect food markets, however, home-produced food and purchased food are no longer perfect substitutes. In this scenario, an increase in the price of cash crops affects food consumption both indirectly, through an increase in household agricultural profits, and directly, through a decrease in home-produced food. In this essay, I provide evidence for the presence of food market constraints in Tanzania and their effects on a household's food and nutrient consumption.

Together, these three essays shed light on the ways in which households in Tanzania affect and are effected by their local environments and local markets. All three essays provide insight into how rural households value their natural environment and have implications for how policy-makers can effectively manage the natural environment. While the effects of climate change are global, mitigating climate change comes from altering local decisions. And though much attention has been paid to the global need to conserve natural resources and mitigate climate change, less effort has been made to understanding local populations' interactions with their environments. Importantly, the previous literature on this topic was scarce and thus each essay makes an important and timely contribution. I discuss more of the specific contributions and policy implications of each study in their respective chapters.

Chapter 2

Measuring the Welfare Effects of Forests in Tanzania: An Application of the Travel Cost Model

2.1 Introduction

Societies have long relied on forests to meet their daily needs – in the form of either firewood for fuel or timber for building construction. Currently over 350 million people, the majority of them poor, live in or near forests (Klugman, 2011), and 57 countries experience firewood shortages (Perlin, 2005). The value of forests is also reflected in the vast resources currently being spent to reduce climate emissions through forest conservation; Denmark, Japan, Norway, and Spain have pledged over US$170 million in funds to support the United Nations' Reducing Emissions from Deforestation and Degradation (UN-REDD) programme that has been implemented in 47 different countries (UN–REDD Programme, 2009). Similarly, the Nature Conservancy has its own suite of forest conservation-based climate mitigation programs in Brazil and Indonesia (Nature Conservancy, 2013).

These two benefits of forests unavoidably conflict with each other; conserving forests to reduce greenhouse gas emissions reduces the amount of firewood and timber that can be obtained from forests. Ultimately, an accurate monetary estimate of the costs of forest conservation programs from reduced procurement of firewood and timber depends not only on the market benefits associated with these uses of forest resources but also on the non-market uses. Household well-being derived from firewood collection, however, is difficult to measure because this service is rarely exchanged in markets. To date, very few economic studies have attempted to use household socio-economic data to quantify the benefits that households in developing countries derive from local community forests. Indeed, the need for more economic papers addressing this research topic was recently recognized in the working paper by Greenstone and Jack (2013). In this paper, I attempt to fill this gap in the literature and contribute to the growing micro-level literature on the impact of climate change programs on households in developing countries. I estimate the demand for household firewood collection trips and derive a measure of household-level willingness to pay for local community forest access in the Kagera region of Tanzania.[1]

Tanzania has lost approximately 80,760 square kilometers of forest over the last 20 years (World Bank, 2010), an area the size of South Carolina, and, over the next 40 years, deforestation in sub-Saharan Africa is predicted to occur at a faster rate than anywhere else in the world (Millennium Ecosystem Assessment, 2005). Tanzania joined the UN-REDD programme in 2008 and, to date, international donors have pledged over four million dollars to support, in part, forest-conservation programs (UN–REDD Programme, 2009). Tanzania is still developing and assessing the costs of possible forest conservation programs, including participatory forest management and payment for ecosystem services (UN–REDD Programme, 2009), and the costs evaluated include opportunity costs, implementation costs, transaction costs, and institutional costs (UN–REDD Programme, 2012). Most notably, opportunity costs are defined as the difference in household net earnings from agriculture, timber, or charcoal production between a forest conservation regime, where forest access and use is limited, and the status quo, where forests can be freely used by local populations. In both of the previous cost analyses (Fisher et al., 2011; UN–REDD Programme, 2012), however, opportunity costs are defined only in areas where markets for forest products exist and thus neglect households' opportunity costs associated with reduced firewood collection.

[1]The correct welfare measure is compensating variation but throughout this paper I rely on the results in Willig (1976) and use willingness to pay as a proxy for compensating variation.

These non-market opportunity costs associated with forest conservation are especially important in Tanzania, where recent estimates show that approximately 40 million people (out of a total population of 43 million) rely on biomass for cooking fuel (Mushi, 2012), with wood being the most common type of biomass (Mwampamba, 2007). I contend that commonly used opportunity cost measurements fall short of measuring the full costs borne by local households associated with these forest conservation programs.

In this paper, I estimate the household-level welfare, measured as willingness to pay (WTP), associated with firewood collection in Kagera, Tanzania. I adapt traditional environmental economics techniques, in particular travel cost models, to a developing country setting and use these revealed preference estimates to derive WTP welfare estimates. I estimate that households in Kagera, Tanzania are willing to pay, on average, approximately \$200 per year (2012 U.S. dollars) for access to local community forests. This WTP estimate is twice as large as household-reported values of firewood consumption, which suggests that households significantly underestimate the value of firewood. This paper is one of the first papers to derive estimates of households' willingness to pay for forest access in sub-Saharan Africa and one of only a few studies using a revealed preference approach to natural resource valuation in a developing country.

A second, more methodological contribution of this paper is that I estimate WTP values in the presence of constrained labor markets and limited wage data, an important feature of many developing countries (Greenstone and Jack, 2013). Travel cost models estimate the demand for local community forests by estimating the number of firewood collection trips households make as a function of households' opportunity costs of the time devoted to firewood collection trips. If household members are unable to freely choose how many hours they work outside of the home, then local wage rates are unlikely to represent these households' true opportunity costs of time. To account for this constraint, I use a household profit function to estimate household-specific shadow wages. In addition, I construct a household-specific travel cost index that accounts for intra-household differences in shadow wages and firewood collection participation levels. Finally, in contrast to previous research that relied on cross-sectional data (Pattanayak et al., 2004; Baland et al., 2010), I employ a household fixed effects estimation strategy that controls for unobserved household characteristics that affect household travel costs and that would otherwise lead to inconsistent coefficient estimates.

This paper proceeds as follows: In Section 2.2, I review the literature and further highlight the contributions made by this paper. In Section 2.3, I describe the data source used in this study, the Kagera Health and Development Survey, and present descriptive statistics of the sample population. In Section 2.4, I combine a travel cost model with an agricultural household production model and derive the relevant demand functions to be estimated. In Section 2.5, I describe the estimation strategy and the construction of the travel cost variable. I present the estimation results in Section 2.6, as well as robustness checks and a discussion of the resulting household welfare estimates associated with community forest access. Finally, I discuss drawbacks and caveats of my results in Section 2.7, and provide concluding remarks in Section 2.8.

2.2 Literature Review

This paper extends the small but growing subset of literature that considers the degree to which agricultural households in sub-Saharan Africa rely on forest resources, and the effects that proposed reductions in forest access would have on these households. The "other energy crisis" (Eckholm, 1975), firewood shortages, has long been a focus of research, but the majority of previous studies have focused on the crisis in the context of southeast Asia.[2] As noted by Cooke, Köhlin and Hyde (2008), more empirical studies from Africa are needed to compare results from Southeast Asian studies and thus provide more reliable generalizations. In particular, the two recent studies that provide monetary welfare estimates for household forest access are both estimated for Southeast Asian countries (Pattanayak et al., 2004; Baland et al., 2010). To the best of my knowledge, no comparable studies exist with similar estimates for households in sub-Saharan Africa.

[2]See Amacher, Hyde and Joshee (1993); Bluffstone (1995); Amacher, Hyde and Kanel (1996); Cooke (1998*a*,*b*); Amacher, Hyde and Kanel (1999); Adhikari, Falco and Lovett (2004); Baland et al. (2010); Baland, Libois and Mookherjee (2013) for studies relating to Nepal, Heltberg, Arndt and Sekhar (2000); Foster and Rosenzweig (2003); Kohlin and Amacher (2005); Gupta and Köhlin (2006); Gundimeda and Kohlin (2008) for studies relating to India, Shively and Fisher (2004) for studies relating to the Philippines, and Pitt (1985); Pattanayak et al. (2004) for studies relating to Indonesia.

To date the majority of relevant studies have focused on observable consumption- and labor-based household responses to the economic scarcity of firewood (Dewees, 1989; Heltberg, Arndt and Sekhar, 2000). Consumption-based measures have focused on estimating household firewood consumption, measured through either firewood expenditures or firewood collected, as a function of household, village, and environmental characteristics (Amacher, Hyde and Joshee, 1993; Amacher, Hyde and Kanel, 1996; Chen, Heerink and Van Den Berg, 2006), or in relation to household consumption of substitutes such as coal and kerosene in a demand systems approach (Pitt, 1985; Gupta and Köhlin, 2006; Gundimeda and Kohlin, 2008). The majority of these studies find that own-price elasticities for firewood fall between negative one and zero. Income elasticities are more ambiguous: some studies find a negative income elasticity (Amacher, Hyde and Joshee, 1993; Heltberg, Arndt and Sekhar, 2000) and others a positive income elasticity (Amacher, Hyde and Joshee, 1993; Cooke, 1998*b*; Amacher, Hyde and Kanel, 1999).[3] In contrast with this paper, none of these studies makes the link between household demand for firewood and the welfare loss that may result from lost forest access for households.

Labor-based outcome measures have focused on the effects of forest scarcity on two dimensions of labor supply: First, who in the household collects firewood (men, women, or children), and, second, firewood collection time relative to other household activities such as agriculture. Most results indicate that scarcity measures (such as firewood price, firewood collection trip time, and firewood distance) are not correlated with increased total collection time for specific household members (Amacher, Hyde and Joshee, 1993; Cooke, 1998*b*; Cooke et al., 2000; Amacher et al., 2004; Palmer and MacGregor, 2009) though Amacher et al. (2004) did show an increased burden for children.[4] Finally, both Amacher et al. (2004) and Cooke (1998*b*) find little evidence of any effects of firewood scarcity on agricultural labor but, Kumar and Hotchkiss (1988) find the opposite results. For a complete review of these studies see Cooke, Köhlin and Hyde (2008).

The studies listed above are important and provide insights into the ways in which

[3]For a more complete, albeit somewhat dated, review of studies on firewood consumption see Hyde, Kohlin and Amacher (2000). The authors review the income elasticities across seven different studies.

[4]It is worth pointing out that four of the five studies cited here relied on less than 200 observations (Amacher, Hyde and Joshee, 1993; Cooke, 1998*b*; Cooke et al., 2000; Palmer and MacGregor, 2009) so their power to detect a significant change was likely very low.

households physically respond to changes in forest scarcity but they all fall short of linking observed household behavior to an estimate of the welfare benefits derived from local community forests. Only five studies explicitly discuss the effects of a change in forest access on household welfare, and only two of these studies, neither of which examines an African country, measure welfare in monetary terms. In Southeast Asia, Kohlin and Amacher (2005) estimate annual time saved from access to a community forest (250 hours per household) while Baland et al. (2010) estimate the welfare loss to a household associated with a one hour increase in collection time per trip (503.2 rupees/year). In Africa, Fisher (2004) examined the reduction of observed inequality if firewood income (a proxy for forest benefits) is explicitly included in account measures (12 percent reduction in the Gini coefficient) and MacDonald, Adamowicz and Luckert (2001) estimate the caloric value associated with a change in the forest resource base. To date, only Pattanayak et al. (2004) has used travel cost estimation to measure WTP for forest access, doing so for Manggarai households in Indonesia.

This paper extends the previous wood fuel literature by providing an estimate of agricultural households' willingness to pay for forest access in Tanzania. This paper is the first to estimate households' willingness to pay for forest access in sub-Saharan Africa. Previous studies on African countries have examined firewood consumption in Ethiopia (Amacher et al., 2004), Malawi (Fisher, 2004; Fisher, Shively and Buccola, 2005; Fisher, Chaudhury and McCusker, 2010), Namibia (Palmer and MacGregor, 2009), and Uganda (Khundi et al., 2011), but none has estimated local households' willingness to pay for forest access.

This paper extends the travel cost model employed by Pattanayak et al. (2004) by developing a household model with different types of household workers. More specifically, I use a generalization of Pattanayak et al. (2004) household production model of firewood collection that is placed in the context of the household agricultural model developed by Jacoby (1993) in which the household has two distinct labor types. I also allow for the case of constrained labor markets, where household shadow wages are derived from estimates of a household profit function (Jacoby, 1993; Baland et al., 2010) as well as corner labor markets (Bockstael, Strand and Hanemann, 1987).

This paper is one of the few in the literature that employs panel data (others include

Cooke, 1998*a*,*b*; Shively and Fisher, 2004; Baland, Libois and Mookherjee, 2013). Most previous studies rely on cross-sectional data to estimate the impact of forest scarcity; they are potentially biased because they do not account for unobservable variables such as household collection efficiency. I employ a household fixed effects estimation approach to control for these unobservable household characteristics.

2.3 Data

The ideal data for travel cost estimation would include information on household firewood collection trips for all household members over the past year for a representative sample. The most recent Tanzanian household survey is the National Panel Survey (NPS), collected in both 2008 and 2010, but the survey collects information only on firewood collection trips made the previous day. A one-day recall period loses much of the variability in the distribution of weekly firewood collection trips across households; households that normally collect firewood but that did not collect firewood yesterday are recorded with zero firewood collection trips and all households that collected firewood yesterday are recorded as collecting firewood at least seven times in the last week. In contrast, the four rounds of the Kagera Health and Development Survey (KHDS), a regionally representative longitudinal household survey, collected information on firewood collection trips made over the previous week. This more refined firewood collection trip question allows for more variability in the frequency of household firewood collection trips. Indeed, 47.5% of households report making between one and six weekly firewood collection trips in the KHDS - responses that would be recorded as either zero or one in the NPS.

Although the KHDS data used in this paper are about 20 years old, household firewood use in Tanzania has been largely unchanged over the last twenty years. According to the 2010 NPS, approximately 93% of rural households in Kagera report firewood as their cooking fuel compared to 99% of rural households in the KHDS. Recent estimates suggest that in 2011, natural forests in Tanzania could produce an annual supply of about 18 million cubic meters of wood products but that the annual wood harvest is over 50 million cubic meters (Mushi, 2012). This relatively small adoption of firewood substitutes in Kagera and small change in deforestation patterns in Tanzania over the last 20 years means that the observed firewood collection patterns in the KHDS are still relevant today. Consequently, I rely on the KHDS for this paper to take advantage of the data set's more refined firewood collection data.

The KHDS surveyed over 800 households in the Kagera Region of Tanzania four times between September 1991 and January 1994.[5] The Kagera region (40,838 km^2) lies in the northwest corner of Tanzania on the western shore of Lake Victoria and borders Uganda, Rwanda, and Burundi. Kagera is one of the farthest regions in Tanzania from the country's capita, Dar es Salaam (see Figure 2.1 for a map showing the geographical placement of Kagera in Tanzania). This study uses an unbalanced panel with 3,375 observations (840 in round 1, 849 in round 2, 858 in round 3, and 828 in round 4). The household attrition rate is low: between the 1991 and 1994 survey rounds the annual household attrition rate was 0.88 percent and only 0.70 percent after excluding deaths (Outes-Leon and Dercon, 2008). Survey rounds were conducted six to seven months apart and households were surveyed from 50 different villages across all five districts of Kagera. The KHDS used a two-stage stratified random sample based on the 1988 Tanzanian Census. Communities were stratified based on agroclimatic zone and adult mortality rates. The KHDS questionnaires were adapted from the World Bank's Living Standard Measurement Study questionnaire, with questionnaires administered to households, communities, local markets, local medical centers, and local schools.

In Tanzania, most households have access to community forests and rely on them as a major source of firewood. Specifically, all land in Tanzania is owned by the government but is divided into three categories: general land, reserved land, and village land (i.e. community land) (Carpano, 2010; Rurai, 2014). General land is either privately managed and used at the discretion of the owner or held by the government. Reserved land includes game reserves and forest reserves, is used at the discretion of the national government, and is not available for private household use. Finally, community land is managed by the village council, includes grazing land and village forest reserves, and is available for use by all households in the village; a village may even allow specific individuals to manage a piece of the community

[5] A household was defined as "a person or group of persons who live in the same dwelling and eat together for at least three of the twelve months preceding the date of the survey" (Ainsworth, 2004).

land. Each village also creates a Forest Management Plan that designates different areas of the village forest for different uses. Usually the plan forbids the cutting of fresh wood but may allow collection of dead woods (Rurai, 2014). Liversage (2004) estimates that as of 2004 general land comprised 29 percent of all land, reserved land 2 percent, and community land 69 percent.

An additional benefit of using data collected in the early 1990s is that there is little self-selection into households that do and do not have tree nurseries and do and do not sell firewood. To accommodate declining firewood availability on community forest land, more households in recent years have allocated some of their personal land to tree nurseries and have begun selling firewood. Thus, without data on firewood collection sites, using the 1991–1994 data allows me to credibly assume that the majority of households collected firewood on community forest land. Table 2.1 presents summary statistics from the KHDS on the proportion of households that collect firewood, own tree nurseries, and sell firewood. Approximately 95 percent of households report using firewood in all four survey rounds with over 85 percent of households collecting firewood in each round. Additionally, virtually no households have income from firewood sales in rounds 1 and 2 and only two percent of households in round 3. Very few households have firewood in stock but approximately 23 percent of households do report having firewood crops, most likely in the form of a tree nursery.

In the KHDS all household members seven years and older were asked "How many hours did you spend collecting firewood in the last seven days?" I use this question to construct a measure of the number of firewood collection trips that each individual made in the last week. The number of individual collection trips last week is the number of days that a certain household member reported non-zero collection time. This trip frequency is accurate under the assumption that a certain household member is not making more than one firewood collection trip on any given day. I also use this question to construct a measure of the average time per trip across all household members. An individual's trip time is measured as the total number of hours that she or he spent collecting firewood in the last week divided by the number of trips she or he made. Exact construction of the travel cost variable is explained in more detail in Section 2.5.

The number of household firewood collection trips is used as a proxy variable for a household's annual number of firewood collection trips. Weekly household firewood collection trips are a good proxy for annual firewood collection trips as long as firewood collection trips do not vary significantly with the season. Few households store firewood (see Table 2.1) mostly because firewood collection is predominately done by hand, which implies that a significant amount of time would be required for a household to build up a firewood stock. In addition, households use firewood to cook throughout the year so they have a continuous need to collect firewood. Figure 2.2 displays weekly household firewood collection trips by the month of interview. The plot shows that the median number of weekly firewood collection trips varies little across the 12 months, corroborating my statement that seasonality is not a major factor in household firewood collection decisions. The upper extreme of firewood collection trips, however, is smaller during the middle of the two rainy seasons (April and December) but, between the upper and lower quartile (where most of the data lie), there is little variation in weekly trips across the twelve months.

For the analysis in this paper I drop households that report any firewood sales in the last six months (38 households) because these households are assumed to use local community forests for commercial purposes and the interest here is in private non-market forest use. This restriction leaves a total sample of 3,337 observations. Only one household in round 1 and 17 in round 4 report any firewood sales so dropping these observations should have minimal effects on the results. If households with firewood sales are assumed to have a more inelastic demand with respect to own-price for firewood collection trips then any resulting bias would be downward which implies that WTP estimates are underestimated.

Table 2.2 displays summary statistics for the sample of interest. Prices and monetary variables are all presented in real terms with round 1 as the base year. These variables are deflated using a Laspeyre's price index that is measured at the village level. Households, on average, cultivate approximately four acres of land and tend to have more valued assets in the form of livestock than they do in the form of non-farm business assets. The average household has between five and six individuals and a little more than a quarter of households are female-headed. The price of charcoal is relatively constant across the four survey rounds while the price of kerosene increases across the four rounds. Finally, Table 2.2 shows that, on average, households make between six and seven firewood collection trips per week. To further investigate the distribution of weekly firewood collection trips I present both histogram and kernel density estimates by round in Figure 2.3. Weekly firewood collection trips appear highly skewed with spikes in the observed frequency at zero, two, four, and seven weekly firewood collection trips. In the next section I describe the theoretical foundations of the travel cost model before empirically estimating the model in sections 2.5 and 2.6.

2.4 Travel Cost Model

In traditional single-site travel cost demand models (Phaneuf and Smith, 2005) household utility, u, is a function of household trips to the forest, y, and household consumption of some aggregate good, z.[6] I start with this traditional single site model but adapt it to allow trips to the forest to be an input into a household firewood production process, an approach similar to that taken by Pattanayak et al. (2004) and Baland et al. (2010). I also assume that households earn income from both marketed labor and household agricultural production, similar to the household model laid out in Jacoby (1993).

In this adapted model, household utility, u, is a function of firewood $f \geq 0$, an aggregate good $z \geq 0$, leisure $\ell \geq 0$, and a vector of household taste-shifters $\boldsymbol{\zeta}$. I assume that the utility function is concave in all of its arguments, that $\partial^2 u/\partial z \partial f \geq 0$, and that there are diminishing returns of income on the maximized utility level. Both Pattanayak et al. (2004) and Baland et al. (2010) assume identical household workers with a single household measure of leisure. More realistically, household workers choose different amounts of leisure, firewood collection trips, home and market labor hours, and different firewood collection and agriculture productivity levels. I incorporate this intra-household differentiation into my model and allow for two types of household members.[7]

Household firewood production, $f = g(y_1 + \theta y_2, \delta)$, is a function of the total number of

[6]Throughout this theoretical section, I rely on a unitary household model that is a derivation of the traditional agricultural household model (Bardhan and Udry, 1999). The unitary household model relies on the assumption that either utility is transferable (Alderman et al., 1995) or that household resource allocation is driven by one altruistic household member (Becker, 1993). I maintain this assumption because firewood is consumed jointly within the household.

[7]For notational simplicity I only consider the case of two distinct household members but the model can easily be extended to include n different household members.

firewood collection trips that a household makes, $y = y_1 + \theta y_2$ $(y, y_1, y_2 \geq 0)$, where $\theta \in [0,1]$ captures differences in firewood collection productivity and $\delta \in [0,1]$ denotes environmental quality. The firewood production function, $g(\cdot,\cdot)$, is assumed to have increasing but diminishing marginal returns to firewood collection trips which implies that $\partial g/\partial y_j > 0$ and $\partial^2 g/\partial y_j^2 < 0$ for $j = 1, 2$. Given the same environmental quality, household members of the same type are assumed to collect the same amount of firewood per trip across households.[8] Finally, time per trip, t^f, (including travel and collection time) is assumed to be the same for both household member types but varies across households.

Household revenue (agriculture, home business, livestock, and fishing profits), measured in local currency units, are a function of household fixed assets, A_{fi}, household variable assets, A_{vi}, household labor, L_j^a, and hired labor, L^h: $F(A_{fi}, A_{vi}, L_1^a, L_2^a, L^h)$. Household marketed labor is given by L_j^m, and household members are endowed with E_j^L total units of labor.

In full, the household problem can be written as:

$$\max_{y_1,y_2,z,\ell_1,\ell_2,L_1^m,L_2^m,L_1^a,L_2^a,L^h} u\big(g(y_1+\theta y_2,\delta), z, \ell_1, \ell_2; \boldsymbol{\zeta}\big)$$

subject to:

$$pz = F(A_{fi}, A_{vi}, L_1^a, L_2^a, L^h) - wL^h + w_1 L_1^m + w_2 L_2^m$$

$$E_1^L = \ell_1 + L_1^m + L_1^a + y_1 t^f$$

$$E_2^L = \ell_2 + L_2^m + L_2^a + y_2 t^f$$

where the aggregate good's price is p, w_j is the market wage rate for household member of type j, and w is the market wage rate for hired labor. The household problem can be

[8] This assumption implies that some households do not use carts or animals in order to help individuals carry more firewood per trip. This assumption seems plausible since carrying firewood on one's back or shoulders remains the dominant form of firewood transport in the region.

rewritten as (Singh, Squire and Strauss, 1986):

$$\max_{y_1,y_2,z,\ell_1,\ell_2,L_1^m,L_2^m,L_1^a,L_2^a,L^h} u\big(g(y_1+\theta y_2,\delta),z,\ell_1,\ell_2;\boldsymbol{\zeta}\big)$$

subject to:

$$pz+\sum_{j=1,2} y_j w_j t^f+\sum_{j=1,2} w_j\ell_j = F(A_{fi},A_{vi},L_1^a,L_2^a,L^h)-wL^h-\sum_{j=1,2} w_j L_j^a+\sum_{j=1,2} w_j E_j^L.$$

Household member type j's travel cost associated with one firewood collection trip is denoted as $w_j t^f$.

2.4.1 Perfectly Functioning Labor Markets

If labor markets are perfectly functioning then households are indifferent between hiring labor, providing labor at home, and working in the market. In this case, household consumption and agricultural production decisions are separable and household characteristics, such as household size and gender of household head, do not affect the level of household production. Household home labor is determined by the first order condition:

$$\frac{\partial F(A_{fi},A_{vi},L_1^a,L_2^a,L^h)}{\partial L_j^a}=w_j \text{ for } j=1,2.$$

The optimal number of household firewood collection trips for person j, y_j^*, is a function of travel costs and other exogenous variables:

$$y_j^* = y_j(w_1t^f,w_2t^f,\theta,\delta,p,\pi,E_1^L,E_2^L,\boldsymbol{\zeta}) \geq 0 \text{ for } j=1,2 \tag{2.1}$$

where $\pi = F(A_{fi},A_{vi}^*,L_1^{a*},L_2^{a*},L^{h*})-w_1L_1^{a*}-w_2L_2^{a*}-wL^{h*}$ denotes household profit. It is straightforward to show that $\partial y_j/\partial(w_jt^f)<0$ so that as household member type j's travel cost, w_jt^f, increases, that member type j will make fewer firewood collection trips. In contrast, $\partial y_j/\partial(w_kt^f)>0$ for $j\neq k$ so that as the travel cost increases for household member type k then type j will make more trips, ceteris paribus. Additionally, firewood collection trips will increase as income rises, $\partial y_j/\partial\pi>0$ for $j=1,2$, because fuel is assumed to be a normal good.

Estimates for this case of perfectly functioning labor markets rely on observed household wages. These estimates are the least robust because they do not allow for any imperfections in the local labor market, a common feature in many developing countries.[9] In the next subsection I relax the assumption of perfectly functioning labor markets.

2.4.2 Constrained Labor Markets

If hired and own-farm agricultural labor are not perfect substitutes, or if local labor markets have transaction costs such as limiting the amount of hours a household member can work outside the household, then household production and consumption decisions are no longer separable (Jacoby, 1993). For example, if both household members work in home-production and are unable to find jobs in the local labor market so that $L_1^m = L_2^m = 0$ then household production is affected by household characteristics. In this case, each household member's individual shadow wage, $\hat{w}_j$, is given by:

$$\frac{\partial u/\partial \ell_j}{\partial u/\partial z} \cdot p = \frac{\partial F(A_{fi}, A_{vi}, L_1^a, L_2^a, L^h)}{\partial L_j^a} = \hat{w}_j \text{ for } j = 1, 2$$

where the shadow wage is equivalent to the price of leisure. These shadow wages can be estimated from a household profit function and, in the presence of constrained labor markets or missing or poor wage data, are a good measure of household members' opportunity costs of time (Jacoby, 1993).

The optimal number of firewood collection trips for household member j is now written as a function of shadow wages and not market wages:

$$y_j^* = y_j(\hat{w}_1 t^f, \hat{w}_2 t^f, \theta, \delta, p, \pi, E_1^L, E_2^L, \boldsymbol{\zeta}) \text{ for } j = 1, 2 \tag{2.2}$$

where profit is $\pi = F(A_{fi}, A_{vi}^*, L_1^{a*}, L_2^{a*}, L^{h*}) - \hat{w}_1 L_1^{a*} - \hat{w}_2 L_2^{a*} - wL^{h*}$. Household member type j's travel costs still negatively affect the number of firewood collection trips that type j makes and positively affect the number of firewood collection trips that type $k \neq j$ makes. With constrained labor markets, however, an increase in household profits now positively affects type j's travel costs and the effect of household profits on the number of type j's firewood collection trips is ambiguous. In particular, increased profits raise the demand for

[9] See Jacoby (1993); Skoufias (1994); Bhattacharyya and Kumbhakar (1997); Abdulai and Regmi (2000); Grimard (1997); Le (2010) for studies that all reject the separation hypothesis.

firewood but also raise the opportunity cost of collecting firewood.

In the case of constrained labor markets, shadow wages and household profit both become choice variables and are endogenous in any cross-sectional estimation. Both Jacoby (1993) and Baland et al. (2010) instrument for shadow wage using the number of household workers and for household profit using household potential income, the predicted profit from a household profit estimation. In my analysis I do not use instrumental variables but instead include household fixed-effects, thereby controlling for all unobserved time-invariant household characteristics, such as ability and location, that affect both shadow wages and profits.

Travel cost estimates from this case of constrained labor markets are more robust to unobserved constraints in local labor markets and to misreporting in the observed wage variable. In the discussion of the empirical results, estimates from this case are taken as the most robust and realistic travel cost estimates. In the next section I consider a third possible labor market condition.

2.4.3 Corner Solutions in the Labor Markets

Finally, it may be possible that household labor market decisions are at a corner solution. At a corner solution household member $j = 1, 2$ cannot work additional hours in the labor market for wage w_j and there are little or no productivity gains from working at home, i.e. $\hat{w}_j \approx 0$ (Bockstael, Strand and Hanemann, 1987). In this case L_j^a and L_j^m for $j = 1, 2$ are fixed and the household only chooses z, ℓ_1, ℓ_2 , y_1, and y_2. Consequently, the optimal number of firewood collection trips for type j is given by:

$$y_j^* = y_j(t^f, \theta, \delta, p, \pi, E_1^L, E_2^L, \boldsymbol{\zeta}). \tag{2.3}$$

This case eliminates the need to consider wage measurement or construct shadow wages. And, as in the case of perfectly functioning labor markets, household income positively affects the number of firewood collection trips made by person j, as long as there is decreasing marginal utility of income.

This case of corner solutions in the labor market is predominately estimated as a robustness check. Under the assumption that home production exhibits increasing returns to home labor (i.e. $\partial F/\partial L_j^a > 0$ for $j = 1, 2$) then $\hat{w}_j > 0$ for all $L_j^a > 0$. Estimates from this case are straightforward, do not rely on any wage construction, and travel cost coefficients are easily interpreted in terms of firewood collection travel time.

2.4.4 Aggregate Household Demand model

Equations (2.1), (2.2), and (2.3) express household member type j's demand for firewood collection trips, but household WTP estimates for forest access are derived from total household demand for firewood collection trips. To obtain household level firewood collection trip demand I aggregate up the member-based demands solved for in equations (2.1), (2.2), or (2.3). Household members' firewood collection travel costs are assumed to be multicollinear over time. Because travel time is assumed to be the same for all household members this assumption holds as long as household members' observed wages and shadow wages are positively correlated. Under this assumption, Hicks' generalized composite commodity theorem applies to household members' firewood collection trips and demand for firewood collection can be aggregated to the household level (Hicks, 1946; Lewbel, 1996).

The construction of aggregate household demand for firewood collection trips is difficult for both the case of perfect labor markets and constrained labor markets because in each case trip demand for household member type j is a function of both the travel costs for type j and the travel costs for type $k \neq j$. To see why, consider the case of perfect labor markets (the case of constrained labor markets is analogous, simply replace w with $\hat{w}$). Total household firewood collection trips is an aggregate household good composed of the firewood collection trips of household members 1 and 2:

$$y^* = y_1^* + y_2^* = \sum_{j=1}^{2} y_j(w_1 t^f, w_2 t^f, \theta, \delta, p, \pi, E_1^L, E_2^L, \boldsymbol{\zeta}).$$

In order to write this aggregate model as a function of a single travel cost variable a price index is needed. I create a price index that weights both member 1 and 2's travel costs:

$$\bar{w} t^f = \sum_{j=1}^{2} \frac{y_j}{y_1 + y_2} \times w_j t^f = \left(\sum_{j=1}^{2} \frac{y_j}{y_1 + y_2} \times w_j \right) t^f. \tag{2.4}$$

The travel cost index is a weighted average of the travel costs for all household member types where the weight for type j is the share of firewood collection trips made by type j. Total household demand for firewood collection trips can now be written as:

$$y^* = y(\bar{w}t^f, \theta, \delta, p, \pi, E_1^L, E_2^L, \boldsymbol{\zeta}). \tag{2.5}$$

Similarly, in the case of constrained labor markets aggregate household demand for firewood collection trips is:

$$y^* = y(\bar{\hat{w}}t^f, \theta, \delta, p, \pi, E_1^L, E_2^L, \boldsymbol{\zeta}). \tag{2.6}$$

In contrast to other demand models, the price index, equation (2.4), used in equations (2.5) and (2.6) is now a function of the endogenous variables y_1^* and y_2^*. In empirical estimation this endogeneity will generally lead to inconsistent coefficient estimates. A fixed effects estimation approach will control for endogeneity that results from correlation of y_1^* and y_2^* with any unobserved time-invariant household characteristics, such as bargaining structure.

Aggregation of equation (2.3) is trivial because trip demand for any given household member is a function of a single travel time measure. This case does not require the construction of a price index and total household demand for firewood collection trips is given by:

$$y^* = \sum_{j=1,2} y_j(t^f, \theta, \delta, p, \pi, E_1^L, E_2^L, \boldsymbol{\zeta}) = y(t^f, \theta, \delta, p, \pi, E_1^L, E_2^L, \boldsymbol{\zeta}). \tag{2.7}$$

Equations (2.5), (2.6), and (2.7) all represent demand equations and, consequently, welfare valuations can also be applied to each equation. Specifically, household-level WTP for forest access is measured as the area under the Marshallian demand curve (Bockstael et al., 1990; Haab and McConnell, 2002):

$$\text{WTP}(forest\ access) = \int_{(wt^f)^0}^{(wt^f)^1} y(wt^f, \theta, \delta, p, \pi, E_1^L, E_2^L, \boldsymbol{\zeta})\, d(wt^f) \tag{2.8}$$

where $(wt^f)^0$ is the current household-level travel cost and $(wt^f)^1$ is the choke price (the travel cost at which households make zero firewood collection trips). For the cases of perfect and constrained labor markets WTP is measured in local currency units. For the case of corner labor markets WTP is measured in hours and the conversion to currency units is more difficult because wages and shadow wages are not good representations of opportunity costs of time.

2.5 Estimation Strategy

In this section, I estimate household demand for firewood collection trips under the three labor market scenarios presented in equations (2.5), (2.6), and (2.7). I use the number of household collection trips as the outcome variable and the indexed travel cost as the main explanatory variable of interest for the purpose of welfare analysis. In the case of equation (2.5), the estimated demand equation is of the form:

$$\begin{aligned} y^*_{kvt} &= y(\bar{w}t^f, \theta, \delta, p, \pi, E^L_1, E^L_2, \boldsymbol{\zeta}) = y(\boldsymbol{x}'_{kvt}\boldsymbol{\beta}, \eta_k, \varepsilon_{kvt}) \\ &= y(\beta_1(\bar{w}t^f)_{kvt} + \beta_\theta\theta_{kvt} + \beta_\delta\delta_{kvt} + \boldsymbol{\beta}_p\boldsymbol{p}_{vt} + \beta_2\pi_{kvt} + \boldsymbol{\beta}_h\boldsymbol{\zeta}_{kvt}, \eta_k, \varepsilon_{kvt}) \end{aligned} \tag{2.9}$$

for household k in village v at time t. The term η_k represents unobservable time-invariant household characteristics and ε_{kvt} is an independent and identically distributed error term. I allow for the possibility that η_k is correlated with some of the regressors in $\boldsymbol{x}_{kvt}$. Specifically, unobservable household characteristics, such as women's bargaining power, may affect which household members collect firewood, which, in turn, affects a household's travel cost measure and the number of collection trips taken. In the estimation, I use a fixed effects model and avoid bias due to correlation of η_k with $\boldsymbol{x}_{kvt}$ by differencing it out of the regression. A description of all the variables used in the estimation is provided in Table 2.3.

The vector of prices, $\boldsymbol{p}$, includes the prices of firewood substitutes – charcoal and kerosene – and a food price index where prices are collected at the village level in each survey round. Observed household taste-shifters, $\boldsymbol{\zeta}$, include household size, a dummy equal to one if the reported household head is female, average household adult education years, and the number of household members with restricted activity in the last seven days. Unfortunately, data limitations prevent any precise measurement of firewood collection trip substitutability (θ) and environmental quality (δ). These two variables are partially controlled for in both the travel cost measure and the household fixed-effects term.

Household net income, π, is measured as reported household expenditures.[10] Household income data, especially data on household business income, may have a large amount of measurement error. Household expenditures, on the other hand, suffer from less misreporting and are strongly correlated with household income, making that variable a good proxy for household income (Deaton, 1997).

Household travel cost values are created by indexing adult male, adult female, teenager, and child travel costs. Adults are categorized as individuals 16 years or older, teenagers are 12 to 15 years, and children are 7 to 11 years. Consequently, the travel cost for household k in village v at time t is indexed as:

$$travel\ cost_{kvt} = \sum_{j=male,female,teen,child} \frac{y_{jkvt}}{\sum_{i=male,female,teen,child} y_{ikvt}} \left(w_{jkvt} \times t^{f}_{kvt} \right) \tag{2.10}$$

The weight used in this index varies across households and, consequently, is potentially correlated with unobservable household characteristics. The fixed effects estimation strategy, however, will remove all bias resulting from correlation between a household's indexed travel cost and time-invariant unobservable household characteristics.

Travel cost values are created by using survey information on both firewood collection travel time and wages. For travel time construction, I continue to assume that firewood collection travel time varies across, but not within, the household. I calculate travel time as the average time per trip across all household members (i.e. total number of collection hours divided by the total number of trips). Travel time has substantial within and between household variation (travel time has a mean of 1.5 hours with a between standard deviation of 0.77 and within standard deviation of 0.67) indicating that there are changes in forest access within households over time. I control for any changes in forest access that may be related to technological changes and not environmental changes by including separate

[10]Household expenditures include food expenditures, consumption of home production, non-food consumption expenditure, remittances sent, and wage income in kind (Ainsworth, 2004).

asset ownership dummies for bicycles, cars, and motorcycles. Households with no reported firewood collection trips are assigned their village average travel time at time t (402 of the 3,295 observations). This assignment is reasonable as long as households are not overly dispersed within a village.

I use three wage constructions to account for the three different labor market scenarios evaluated in Subsections 2.4.1 to 2.4.3: the observed wage (equation (2.5)), the estimated shadow wage (equation (2.6)), and travel time (equation (2.7)). For the case of perfectly functioning labor markets, I rely on a survey question asked to each community leader about how much an agricultural laborer earns for a day's work. This question is asked separately for men, women, and children. Unfortunately, this variable has a large number of missing observations (as many as 35 of the 50 villages in some of the survey rounds have missing child wages), and it also does not allow for any intra-village variation in wages. Second, in the case of constrained labor markets, I use an estimate of the marginal product of labor as a proxy for household shadow wage rates, as proposed by Jacoby (1993). This shadow wage construction has the additional benefit that it allows for sample variation in opportunity costs of time both within and across households. Finally, in the case of corner solutions I set the wage equal to one and estimate travel cost as travel time. In the next subsection I discuss estimation and construction of the shadow wage estimates.

2.5.1 Estimating Shadow Wages

Household-specific male, female, teenager, and child shadow wages are derived from their corresponding labor hour coefficients in a household profit function. Estimation follows the approach laid out in Jacoby (1993) and adapted by Baland et al. (2010). Initially, profits are assumed to follow a Cobb-Douglas functional form:[11]

$$\begin{aligned}\ln profit_{kvt} = {} & \alpha_0 + \alpha_1 \ln land_{kvt} + \alpha_2 \ln livestock_{kvt} + \alpha_3 \ln hh\ education_{kvt} \\ & + \alpha_4 \ln business_{kvt} + \alpha_5 \ln variable\ inputs_{kvt} + \alpha_m \ln male\ labor_{kvt} \\ & + \alpha_f \ln female\ labor_{kvt} + \alpha_t \ln teen\ labor_{kvt} + \alpha_c \ln child\ labor_{kvt} + \mu_{kv} + \epsilon_{kvt}.\end{aligned} \tag{2.11}$$

[11] A log transformation is used on all outcome and explanatory variables in the shadow wage estimates. All observations, except for adult male and adult female labor hours, are replaced with one if they are equal to zero. Results are robust to a hyperbolic sine transformation (results not reported).

Variable descriptions are provided in Table 2.3 and the inclusion of these variables is based on the theoretical household profit function in Section 2.4. The error term μ_{kv} denotes unobservable time-invariant household characteristics that are correlated with household labor hours and ϵ_{kvt} is a serially independent, identically distributed error term that is assumed to be uncorrelated with all regressors.

A Cobb-Douglas functional form is advantageous because of its intuitive interpretation and simple calculation of shadow wages. On the other hand, such a profit function may not be appropriate because it assumes strong separability and requires that the marginal rates of transformation between any two household labor inputs do not depend on any other input (Jacoby, 1993). To test these assumptions, I estimate both a Cobb-Douglas and translog profit function that includes squared labor terms and interaction terms between adult female and child labor hours.

Both Jacoby (1993) and Baland et al. (2010) use cross-sectional data and instrument for the household labor hour variables using the number of working age household members as instruments. Specifically, Jacoby (1993) claims that variable inputs (i.e. labor hours) must be instrumented in order to remove correlation between labor inputs and household fixed effects such as management ability. I use panel data with household fixed effects to control for unobserved household managerial ability and so avoid the need for any instrumental variables.

Household profit estimates are reported in Table 2.4. Column (1) reports results from the Cobb-Douglas profit function presented in equation (2.11) and shows a positive and significant effect of both male and female annual labor hours on household profit, a positive but insignificant effect of teenage annual labor hours, and an unexpected significantly negative effect of child annual labor hours. A Wald Test fails to reject the null hypothesis that the elasticity of male and female labor hours are equal in column (1) (p-value of 0.680) and rejects the null that the elasticity of teenage and child labor are equal at the 5 percent level (p-value of 0.016). Consequently, I re-run regression (1) with a single coefficient for adult labor hours; these results are presented in column (3). As expected, annual adult labor hours is positive and statistically significant in column (3). The corresponding translog profit function estimates are presented in columns (2) and (4). With the exception of business assets in column (2) all coefficients on fixed assets are insignificant; limited within household variation in these variables could increase their standard errors and reduce the power to detect an effect.

The coefficient estimates on adult, teenage, and child labor hours are used to estimate group-level shadow wages (mean shadow wage estimates are reported at the bottom of Table 2.4). The translog profit function produces shadow wage estimates that are notably different from those of the Cobb-Douglas profit function, with the largest difference in teenage shadow wage estimates. The translog profit function also has the disadvantage that it yields negative shadow wage estimates for a significant portion of the sample whereas the Cobb-Douglas profit function will yield negative estimates only if coefficient estimates are negative. Consequently, I use the estimated shadow wages from column (3) in estimates of household demand for firewood collection trips; these estimates are also more in line with the observed wages (see Table 2.5). Finally, all four profit function estimates yield a child shadow wage that is less than zero. This negative coefficient could be related to a surplus labor story (Sen, 1966)[12] but could also be the result of sample selection into which households do and do not use child labor, of the 3,375 total observations only 909 have non-zero child labor hours. I interpret these coefficients as saying that households earn no additional profit from children working an additional hour at home. Thus, children are assumed to be at a corner solution in the labor market, which is equivalent to having a wage of one.

Table 2.5 reports the summary statistics for community wage rates, shadow wages (from column (3) in Table 2.4), and all other variables used in the construction of the indexed household-level travel cost. Adult males make the largest proportion of firewood collection trips in rounds 1 and 2, while there is no significant difference in the proportion of trips made by adult males and adult females in rounds 3 and 4. Adult shadow and sample wages are similar in all four rounds and, in contrast, teenage shadow wage estimates are, on average, much lower than the observed sample average wage rate. Finally, travel cost estimates differ substantially depending on the method of wage measurement used; estimates based on the sample average wage are, on average, higher than travel cost estimates based on the shadow

[12]Sen (1966) defines surplus labor as as labor that can be removed from, in this case, agricultural production, without reducing household profits. Surplus labor requires labor costs (i.e. wages) are not sensitive to the omission of children from the labor force.

wage. This discrepancy could be largely accounted for by the much lower estimate of the opportunity cost of time for teenagers based on shadow wages compared to reported wages.

2.5.2 Estimating Household Demand for Firewood Collection Trips

Because firewood collection trips are always discrete non-negative values, count data models are used for estimation (Haab and McConnell, 2002). I estimate both Poisson and negative binomial models with household fixed effects.[13] The Poisson model is preferred to the negative binomial model because Poisson fixed effects estimates are consistent under much weaker distributional assumptions than those required for consistent estimation of the negative binomial model with fixed effects (Cameron and Trivedi, 2005). Consistency of the coefficient estimates with the Poisson model requires only that the conditional mean, given by:

$$\begin{aligned} \mathrm{E}[y_{kvt} \mid \boldsymbol{x}_{kvt}, \eta_k] = \eta_k \exp(&\beta_1 \mathit{travel\ cost}_{kvt} + \beta_2 \mathit{income}_{kvt} \\ &+ \boldsymbol{\beta}_p \boldsymbol{prices}_{vt} + \boldsymbol{\beta}_h \boldsymbol{hh\ characteristics}_{kvt} \\ &+ \beta_3 \mathit{bike}_{kvt} + \beta_4 \mathit{car}_{kvt} + \beta_5 \mathit{motorcycle}_{kvt}), \end{aligned} \tag{2.13}$$

be correctly specified; it does not require that any full distributional assumption holds.[14]

As discussed in Section 2.4, the coefficients on travel cost are expected to be negative and, for the case of perfectly functioning labor markets and corner solutions, the coefficient on income, measured as household expenditures, to be positive. Ex ante, the effect of income on household firewood collection trips is ambiguous in the case of constrained labor markets. In particular, an increase in income increases household demand for fuel and

[13] For a Poisson model with household fixed effects the probability that household k in village v at time t takes y_{kvt} firewood collection trips is:

$$\Pr[Y = y_{kvt} \mid \boldsymbol{x}_{kvt}, \boldsymbol{\beta}, \eta_k] = \frac{\exp(-\eta_k \lambda_{kvt})(\eta_k \lambda_{kvt})^{y_{kvt}}}{y_{kvt}!}. \tag{2.12}$$

where $\lambda_{kvt} = \exp(\boldsymbol{x}'_{kvt}\boldsymbol{\beta} + \varepsilon_{kvt})$.

[14] This conditional mean requirement contrasts with the conditional mean requirement in the negative binomial model:

$$\mathrm{E}[y_{kvt} \mid \boldsymbol{x}_{kvt}, \eta_k] = \frac{\eta_k \exp(\boldsymbol{x}'_{kvt}\boldsymbol{\beta})}{\phi_{kv}}$$

where ϕ_{kv} represents the overdispersion parameter. If the overdispersion parameter is misspecified then estimates from the fixed effects negative binomial model will be inconsistent (Cameron and Trivedi, 1998).

household members' opportunity costs of time. As mentioned earlier, previous literature remains inconclusive on the sign of firewood income elasticities (Amacher, Hyde and Joshee, 1993; Amacher, Hyde and Kanel, 1999; Cooke, 1998*b*; Heltberg, Arndt and Sekhar, 2000).

Conditional maximum likelihood is used to obtain estimates of $\boldsymbol{\beta}$ (conditional on household-specific totals, $T\bar{y}_{kv} = \sum_{t=1}^{T} y_{kvt}$). For count data panel models there is no incidental parameters problem, so the coefficient estimates will be consistent as long as the conditional mean assumption, equation (2.13), holds (Cameron and Trivedi, 2005). Estimates for the perfect labor market scenario and for the corner labor market scenario are displayed with clustered standard errors at the village level to account for heteroskedasticity in the error term. Estimates for the constrained labor market scenario are displayed with block-bootstrapped standard errors where I block-bootstrap (500 replications) over both the first-stage shadow wage estimates and second-stage demand estimates.

2.6 Estimation Results

I now proceed to estimation of equations (2.5), (2.6), and (2.7). The perfect labor market scenario, equation (2.5), is estimated using the observed sample wage to construct travel cost, the constrained labor market scenario, equation (2.6), using estimated household-level shadow wages, and the corner labor market scenario, equation (2.7), using firewood collection trip time. All three estimates include household fixed-effects and use observed household expenditure as a proxy indicator of household income.

Travel cost estimates are presented in Table 2.6. The travel cost, household size, and household expenditure coefficients are significant in all three regressions. Most notably, the sign on travel cost is negative in all three regressions, indicating that households in Kagera behave rationally in deciding how many firewood collection trips to make; an increase in household travel costs reduces the number of weekly firewood collection trips made by a household. Additionally, the coefficient on household expenditures is positive providing support for a positive firewood income elasticity. Surprisingly, the coefficients on price of kerosene, price of charcoal, and the food price index are insignificant in all three regressions, indicating that households may not have a substitute for firewood.

Coefficient estimates can be interpreted as semi-elasticities. Using the average shadow wages and weights in Table 2.5, a one hour increase in travel time is correlated with, on average, a 14 shilling increase in travel cost.[15] From the estimates in column (1), holding all else constant, a one hour increase in travel time corresponds to, on average, a 25 percent decrease in the weekly number of household firewood collection trips. In column (2), a one hour increase in travel time, holding all else constant, corresponds to, on average, a 17 percent decrease in household firewood collection trips over the previous week. Finally, in column (3), a one hour increase in travel time corresponds to a 39 percent decrease in total household trips holding all else constant. Unfortunately, due to the nature of the Poisson fixed effects estimates, it is impossible to estimate the marginal effects of travel cost on weekly household firewood collection trips because the marginal effects are a function of the unobserved household characteristics captured by η_{kv}.[16]

2.6.1 Robustness Checks

In this sub-section I focus on alternative specifications of the aggregate household travel cost model. I look at the effects of a change in the construction of the indexed travel cost, a change in sample, and changes in the distributional assumptions of weekly firewood collection trips.

First, I compare my results to Pattanayak et al. (2004) travel cost estimates for Indonesia. Their analysis relied on a simpler household travel cost aggregation method. The authors had information only on a household's "typical trip to collect firewood" so they could not estimate an indexed travel cost that accounts for the distribution of collection trips across household members, as done in equation (2.10). I replicate a similar analysis. Households where adults make the largest share of firewood collection trips are assigned a travel cost equivalent to an adult's travel cost (65.4 percent of observations), households where teenagers make the largest share of firewood collection trips are assigned a travel cost equivalent to

[15]The average increase in travel cost for a one hour increase in travel time was calculated as: $19.24 \times 0.34 + 19.24 \times 0.29 + 7.19 \times 0.21 + 1.00 \times 0.11 = 13.74$.

[16]For example, the marginal effects for adult travel costs, $travel\ cost_{akvt}$, are given by:

$$ME_{travel\ cost_{akvt}} = \frac{\partial \mathrm{E}[y_{kvt}]}{\partial travel\ cost_{akvt}} = \eta_{kv} \exp(\boldsymbol{x}_{kvt}\boldsymbol{\beta})\beta_a$$

and η_{kv} is unobservable.

a teenager's travel cost (15.6 percent of observations), and households where children make the largest share of firewood collection trips are assigned a travel cost equivalent to travel time (8.1 percent of observations). Households where two groups make the largest share of firewood collection trips are randomly assigned one of the two group's travel costs (10.8 percent of observations).

Estimation results are presented in column (1) of Table 2.7. The coefficients on travel cost, household expenditures, and household size remain statistically significant and in the same directions as the coefficients reported in Table 2.6. Travel cost estimates are smaller in absolute value when this new household majority travel cost measure is used. The difference in the coefficient estimates between column (2) in Table 2.6 and column (1) in Table 2.7 is not large but the corresponding per trip willingness to pay estimates are almost 20 percent higher when the household majority travel cost is used. These results do not show strong bias when a non-indexed travel cost is used but they do show evidence of potentially different WTP estimates when household travel costs do not account for variations in the opportunity costs of time across household members. More specifically, WTP estimates increase when this simpler travel cost aggregation method is used suggesting that the indexed method used in this paper produces more conservative WTP estimates.

In column (2) of Table 2.7 I estimate a Poisson fixed effects count model dropping households that report having a firewood crop. These households are dropped in order to see whether there is a differential impact of travel cost on firewood collection trips for households that have fewer choices in where to go to collect firewood. Again, travel cost coefficient estimates are smaller in absolute value compared to column (2) in Table 2.6. The difference between these two estimates is in the expected direction because households without a firewood crop have no alternative firewood sources and are likely to be less responsive to travel costs. For example, a household that has a firewood crop may choose to collect firewood in the local community forest if travel costs are low but when travel costs increase the household will switch to collecting firewood from its own crop.

In columns (3) through (7) of Table 2.7, I re-run the original model assuming different distributional and demand form assumptions. In column (3), I estimate a negative binomial count model with household fixed effects; travel cost coefficients differ substantially from the original Poisson fixed-effects estimates. Because the Poisson fixed effects estimates are more robust to distributional assumptions, the negative binomial results may be inconsistent due to a misspecified variance.

Finally, in columns (4) through (7) I estimate household demand for firewood collection trips as a linear function using both ordinary least squares (OLS) with household fixed effects and ordered multinomial choice models. The ordered probit and logit models allow weekly firewood collection trips to be a proxy for unobserved household use of forests and model firewood collection trips on an ordinal scale, as compared to the cardinal scale modeled with the count models (Cameron and Trivedi, 1986). The ordered probit and logit models are both run as cross-sectional estimates with additional household elevation, urban dummy, and round and village fixed-effects. The coefficient on travel cost remains significant and negative in columns (5) and (6) but is insignificant in column (4). In the OLS fixed effects model a 14 shilling increase in travel cost, holding all else constant, corresponds to a 0.28 decrease in weekly household firewood collection trips. In the ordered probit model a 14 shilling increase in travel cost, holding all else constant, corresponds to, on average, a 0.35 decrease in weekly household firewood collection trips, equivalent to a five percent decrease. For the ordered logit model, the same increase in travel cost translates to a 0.71 decrease in weekly household firewood collection trips, or an 11 percent decrease.[17] Even though the percent changes associated with these three regressions are all smaller than the estimates in Table 2.6 they do provide further evidence that households respond rationally to firewood collection travel costs and reduce the number of weekly firewood collection trips they make as their travel cost increases.

2.6.2 Community Forest Valuation

One of the main advantages of travel cost models is that they provide a revealed preference approach to measuring the value of a natural resource. The value of the natural resource

[17]The average change in weekly household firewood collection trips is calculated as:

$$\mathrm{E}[\Delta trips] = \sum_{n=0}^{N} n \times \Delta \mathrm{Prob}[trips = n]$$

where N denotes the maximum number of weekly firewood collection trips made by a household in the sample (50 trips). I hold all explanatory variables constant except travel cost, which is increased by 14 shillings.

is measured as the consumer surplus (willingness to pay) associated with a household's demand for the natural resource, measured in frequency of visits. The general equation for consumer surplus is given in equation (2.8). The exponential demand function used in estimation can be plugged into the general household demand function $y(wt^f, \cdot)$ in equation (2.8). The choke price associated with the exponential demand function is infinity and the corresponding willingness to pay estimates (per trip) are calculated as $-1/\hat{\beta}_1$ where $\hat{\beta}_1$ is the estimated coefficient on travel cost from the Poisson fixed effects count data model (Yen and Adamowicz, 1993; Haab and McConnell, 2002).

Expected willingness to pay estimates (per trip) derived from the travel cost coefficient estimates are given at the bottom of Tables 2.6 and 2.7 along with the 95% confidence interval. The 95% confidence intervals are calculated from a Monte Carlo simulation using 10,000 draws from the multivariate normal distribution.[18] For the main estimation results in Table 2.6, expected WTP estimates range from 55 shillings per trip in column (1) to 84 shillings in column (2). WTP for annual access to the community forest is estimated as the WTP per trip multiplied by the sample mean number of household firewood collection trips.[19] In this section, I discuss the welfare benefits of community forest access using the travel cost estimates given in column (2) of Table 2.6. These results are the most robust because they allow for the possibility of constrained labor markets and variability in household opportunity costs of time. A test for separability (see Appendix) also rejects the null hypothesis that observed sample wages and estimated shadow wages are equivalent at the one percent level. The confidence interval on the WTP per trip measure also suggests that the two estimates are different with the WTP measure from the constrained labor market scenario almost certainly being larger than the WTP measure from the perfect labor market scenario.

Estimates for annual household WTP (measured in 1991 shillings) for local community

[18] I follow Creel and Loomis (1991) and Yen and Adamowicz (1993) and use a Monte Carlo simulation that relies on the asymptotic normality of the coefficient estimates under the assumption that the estimation given model is the true model. I make 10,000 draws from the multivariate normal distribution with a mean vector given by the coefficient estimates and a covariance matrix given by the estimated covariance matrix. The Monte Carlo simulation is advantageous because the estimated confidence intervals are accurate even if the distribution of $\hat{\beta}_1$ is asymmetric.

[19] Because the unobserved household fixed effect is not estimated I cannot estimate annual WTP estimates using the predicted number of household collection trips.

forest access are given in Table 2.8. The value of annual forest access to households ranges from 24,400 shillings in round 3 to 30,100 shillings in round 1. On average, these annual benefits translate to approximately \$200 a year (2012 U.S. dollars) with a 95% confidence interval of \$140 to \$356, suggesting that the benefits of forest access are always positive and significant.[20]

I compare these values against household-reported values of firewood consumption. All households in the survey were asked to value the amount of firewood that they used over the last two weeks, including both purchased and household collected firewood. These values were then aggregated to an annual number. These two values will be equivalent if firewood is correctly valued in the household and households are predominately using local forests for firewood collection. Table 2.8 displays summary statistics for the annual household value of firewood consumption across all four rounds. Household-reported values of firewood consumption are consistently lower than estimated WTP values and these differences are significant at the one percent level across all five columns. More importantly, estimated WTP values, on average, are over twice as large as household-reported firewood consumption values. Anecdotal evidence suggests that community forests are predominately used by households for firewood collection – little hunting or foraging takes place on these lands – suggesting that the majority of the benefits of local community forest access can be attributed to its effects on firewood collection. Thus, I hypothesize that most of the difference between the estimated benefits of forest access and firewood consumption value can be attributed to an undervaluation of firewood on the part of households and in the local markets. The WTP estimates associated with local community forest access appear to provide additional information on a household's use of local community forests that is not captured directly in the firewood consumption survey question. This comparison also suggests that using household-reported firewood consumption values as an estimate of the benefits of local forests to households will lead to an undervaluation of the benefits derived by households from forest access.

Finally, Table 2.8 includes the estimated travel cost elasticity of household firewood

[20]To convert 1991 Tz shillings to 2012 U.S. dollars I first use the Tanzanian consumer price index from the International Financial Statistics to convert 1991 Tz shillings to 2012 Tz shillings (CPI index of 11.94 with a base year of 1991). I then use the average annual exchange rate for 2012 (1,583 shillings per U.S. dollar) from the World Development indicators to convert this value to 2012 U.S. dollars.

collection trips.[21] The corresponding elasticity estimates are approximately -0.20 and add to the growing body of literature showing that firewood consumption in many rural areas is inelastic with an own-price elasticity between negative one and zero.

2.7 Caveats

There are two main caveats in my estimation that merit attention. First, I have no direct measure of the relative firewood collection productivity across household members, nor of environmental quality, so these variables are excluded from the estimation. Most obviously, environmental quality, measured as biomass or total number of trees, is time-variant which means that it is not fully captured by the household fixed-effect, and failure to include it in my analysis may lead to biased coefficient estimates. Future analyses would benefit from direct and reliable estimates of environmental quality but no such data exist for Tanzania in the early 1990s.

Second, I estimate a regional travel cost model where the dependent variable is the total number of household firewood collection trips independent of the site visited. Consequently, I could not measure the benefits of any particular forest site or the effects of changing forest quality on firewood collection trips. As mentioned earlier, community forests in Tanzania are becoming much less dense, and even disappearing, because of poor management and over use due to increased agricultural pressures from a growing population. These changes have potentially profound effects on the local ecosystem and household welfare. A more detailed multi-site travel cost estimation would allow for the analysis of different environmental quality characteristics on household firewood collection decisions and provide more information on which forest qualities households value most.

[21]The coefficient on travel cost represents the semi-elasticity, or:

$$\mathrm{E}[y_{kvt}|\boldsymbol{x}_{kvt}]\frac{\partial \mathrm{E}[y_{kvt}|\boldsymbol{x}_{kvt}]}{\partial travel\ cost_{kvt}} = \beta_1$$

so an elasticity estimate is obtained by multiplying travel cost coefficient estimates by the average household travel cost.

2.8 Policy Implications and Conclusion

My findings show that households derive significant welfare benefits from local community forest access for firewood collection, a predominately non-market activity. Households in Kagera, Tanzania are willing to pay approximately 25 percent of their annual household expenditure, or \$200 (2012 U.S. dollars) a year, for access to local forests. In addition, this WTP estimate is over twice as large as household-reported values of firewood consumption. Most importantly, these estimates are the first set of forest access welfare estimates to be derived for agricultural households in Tanzania. Greenstone and Jack (2013) recently recognized the need for more econometric papers of this type that focus on household willingness to pay estimates for environmental quality in developing countries.

For policy makers, these values suggest that recent opportunity cost estimates associated with the UN-REDD programs in Tanzania may be significantly underestimated because they do not incorporate the costs of forest conservation associated with non-market forest use. With data on households' firewood collection sites, including area and density of forest, which unfortunately are not in the Kagera survey, one could estimate the average household's WTP for a hectare of local community forest and, consequently, per Mg of CO_2. For example, suppose, for the sake of illustration, that the average forest area in Tanzania is 100 hectares. Using the estimates from this paper, a local household would be willing to pay approximately \$200 per year to maintain access to this forest, or \$2 per year per hectare of forest. Assume that with forest conservation the 100 hectares of forests store 207 Mg of CO_2 per hectare and that without forest conservation 100 Mg of CO_2 are lost per hectare.[22] Then, the average local household would be willing to pay approximately \$0.02 per Mg of CO_2. If there are 100 households in the village, this village would be willing to pay a total of \$2 per Mg of CO_2, an opportunity cost that represents a 16% increase from current UN-REDD opportunity costs estimates (Fisher et al., 2011). Additionally, this same forest conservation program would need to make carbon storage payments of \$200 per household per year (i.e. payments for ecosystems services) to all 100 local households to fully compensate the village for its loss

[22] This carbon storage figure accounts for the above ground carbon stored in an open woodlands land cover class (Fisher et al., 2011).

of forest access, which amounts to an increase in program costs of $20,000 per year.

To further evaluate the range of costs associated with forest conservation programs, future research and data collection should also focus on the effects of changing community forest landscapes. The disappearance and change in community forest landscapes is motivating many households to begin to develop local tree nurseries. The most common forest nursery is planted with Eucalyptus trees, an exotic tree species in Tanzania. Eucalyptus trees are preferred for tree nurseries because they are one of the fastest growing hardwood trees in the world. Eucalyptus trees, however, have a considerably higher moisture extraction rate compared to other crops and have been shown to lower neighboring crop yields (Malik and Sharma, 1990). Consequently, future research should investigate any potential negative effects of Eucalyptus plantations on local crop yields. Attention should also be paid to the distributional consequences of tree plantations on the landless households and households that lack adequate land to establish a tree plantation.

Second, the decision of who collects firewood in a given household is one that deserves further exploration. Specifically, empirical tests can be done to see if the allocation of labor within the household is Pareto efficient. Firewood collection labor is allocated efficiently if the relative marginal products of firewood collection are equal for all collecting household members. In addition, future analyses should focus on the effects of increased collection time or increased collection trips on children's labor hours and schooling. Previous literature has focused on the potentially adverse effects of increased collection time on agriculture output or children's collection burden (see Cooke, Köhlin and Hyde, 2008, for a list of relevant studies), but very little research has focused on the effects on children's time in school.

The travel cost model employed in this paper estimates the effects of a change in a household's travel cost to reach a local community forest on the number of weekly household firewood collection trips; this demand model is a revealed preference estimation approach that allows for the estimation of the consumer surplus associated with local community forest access for households in Kagera, Tanzania. The travel cost construction used in my analysis is based on an agricultural household model for which collection trips are an input into firewood production, a household produced good. Unlike previous studies, I allow for three distinct types of household labor, adult, teenager, and child, and show how a household level travel cost index can be created to estimate household level collection trips. Additionally, I allow for the possibility of constrained labor markets and estimate household wages using a household profit function, similar to the approach laid out in Jacoby (1993). My analysis also takes advantage of four rounds of panel data and the ability to control for time-invariant unobserved household characteristics. Ultimately, this paper highlights the importance of revealed preference tools for measuring the benefits of non-market environmental services to households in developing countries.

Table 2.1: Household Firewood Use Summary Statistics

	Survey Round				
	1	2	3	4	Total
Household uses firewood	0.95	0.95	0.95	0.95	0.95
	(0.22)	(0.21)	(0.22)	(0.21)	(0.22)
Household collects firewood	0.90	0.88	0.86	0.85	0.87
	(0.30)	(0.33)	(0.34)	(0.36)	(0.33)
Household has firewood crops	0.16	0.26	0.23	0.25	0.23
	(0.37)	(0.44)	(0.42)	(0.43)	(0.42)
Household has income from firewood sales	0.00	0.00	0.02	0.02	0.01
	(0.05)	(0.03)	(0.14)	(0.14)	(0.11)
Household has firewood in stock	0.02	0.04	0.03	0.03	0.03
	(0.14)	(0.19)	(0.18)	(0.17)	(0.17)
Observations	840	849	858	828	3375

Standard deviation in parentheses.

Table 2.2: Sample Summary Statistics

	Survey Round				
	1	2	3	4	Total
Household profit, TSh (log)[a]	11.76	10.87	10.72	10.52	10.97
	(2.03)	(1.62)	(1.51)	(1.59)	(1.76)
Land acres (log)	0.94	0.99	1.03	1.03	1.00
	(0.91)	(0.89)	(0.86)	(0.93)	(0.90)
Value of livestock, TSh (log)[a]	5.77	6.25	5.96	6.15	6.03
	(4.63)	(4.49)	(4.55)	(4.40)	(4.52)
Average household education (log)	1.31	1.33	1.35	1.33	1.33
	(0.71)	(0.72)	(0.71)	(0.73)	(0.71)
Business assets, TSh (log)[a]	2.01	2.38	2.73	2.58	2.42
	(3.87)	(3.98)	(4.00)	(4.04)	(3.98)
Value of variable inputs, TSh (log)[a]	8.95	8.26	7.61	7.15	8.00
	(2.72)	(2.12)	(2.21)	(2.68)	(2.54)
Adult male annual labor hours (log)	7.24	6.98	6.96	6.86	7.01
	(0.95)	(0.96)	(0.93)	(0.99)	(0.96)
Adult female annual labor hours (log)	7.27	6.97	6.98	6.98	7.05
	(0.83)	(0.90)	(0.87)	(0.84)	(0.87)
Teenager annual labor hours (log)	2.62	2.79	2.99	2.80	2.80
	(3.26)	(3.20)	(3.23)	(3.16)	(3.21)
Child annual labor hours (log)	1.25	1.55	1.87	1.69	1.59
	(2.54)	(2.68)	(2.86)	(2.73)	(2.71)
Household firewood collection trips	6.90	6.06	5.60	6.15	6.18
	(6.31)	(5.84)	(5.38)	(5.44)	(5.77)
Household expenditure, TSh (log)[a]	12.51	11.66	11.60	11.53	11.83
	(0.77)	(0.83)	(0.77)	(0.82)	(0.89)
Price of kerosene, TSh[a]	79.59	75.01	81.40	101.83	84.29
	(15.85)	(15.57)	(17.42)	(22.64)	(20.73)
Price of charcoal, TSh[a]	0.06	0.04	0.04	0.05	0.05
	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)
Food price index, TSh[a]	0.21	0.17	0.14	0.12	0.16
	(0.13)	(0.13)	(0.13)	(0.13)	(0.14)
Household size	5.79	5.69	5.71	5.68	5.72
	(3.08)	(3.01)	(3.12)	(3.22)	(3.10)
Female household head	0.27	0.27	0.28	0.27	0.27
	(0.44)	(0.44)	(0.45)	(0.44)	(0.44)
Average household education	4.47	4.57	4.61	4.60	4.56
	(2.82)	(2.85)	(2.78)	(3.03)	(2.87)
Members with restricted activity	0.58	0.67	0.69	0.60	0.64
	(0.78)	(0.81)	(0.88)	(0.77)	(0.81)
Household owns bicycle	0.28	0.31	0.35	0.39	0.33

Continued on next page

Table 2.2 – continued from previous page

	Survey Round				
	1	2	3	4	Total
	(0.50)	(0.54)	(0.60)	(0.66)	(0.58)
Household owns car	0.02	0.02	0.02	0.02	0.02
	(0.20)	(0.21)	(0.22)	(0.23)	(0.21)
Household owns motorcycle	0.02	0.02	0.02	0.02	0.02
	(0.14)	(0.15)	(0.15)	(0.16)	(0.15)
Observations	838	848	840	811	3337

Standard deviation in parentheses. Sample does not include households that report selling wood, 38 observations.

[a] Values normalized to 1991 Tanzanian shillings.

Table 2.3: Description of Variables

Variable	Variable name	Variable description
	Household profit estimates	
Profit (π)	*profit*	Sum of profits from agriculture and non-farm self-employment.
Fixed assets (A_f)	*land*	Number of acres of household land in cultivation.
	livestock	Value of household-owned livestock (TSh)[a].
	hh education	Average years of education for adults 15 years and older.
	business	Non-farm business assets (buildings, vehicles, and tools).
Variable inputs (A_v and L^h)	*variable*	Total cost of purchased inputs for crop production – includes land, seed, hired labor, fertilizer, pesticide, transportation, and processing costs.
Household labor (L^a_j)	j'*s labor*	Member j's (child, teenager, adult female, or adult male) annually aggregated hours on agriculture, livestock, fishing, and household business activities.
	Household demand for firewood collection trips	
Household firewood collection trips (y)	*trips*	Sum of firewood collection trips in last week across all household groups.
Travel cost ($\bar{w}t^f$)	*sample wage*$_j$	Community reported agricultural laborer day rate for j for each round (day rate is divided by eight to obtain hourly rate).
	shadow wage$_j$	Member j's marginal product of labor estimated from household self-income function.
	travel time	Average time (in hours) per trip across all household members.
Income (π)	*hh expenditure*	Annual household expenditure (TSh) - proxy for π.
Prices ($\boldsymbol{p}$)	*price kerosene*	Price per beer bottle of kerosene (TSh) in village v - varies across community and time.
	price charcoal	Price per small quantity of charcoal (TSh) in village v - varies across community and time.

Continued on next page

Table 2.3 – continued from previous page

Variable	Variable name	Variable description
	price food	Food price index using Stone price index and mean budget share in village v - varies across community and time.
Taste-shifters ($\boldsymbol{\zeta}$)[b]	*hh size*	Household size (unweighted)
	hh head	Dummy variable equal to one if household head is female, zero otherwise.
	restrict	Number of household members with restricted activity in the last 7 days due to own illness
Additional controls	*bike* *car* *motorcycle*	Dummy variable equal to one if household owns bicycle, car, or motorcyle.
Trip substitutability (θ) and environmental quality (δ)	–	No direct measurement. Partially accounted for in shadow wage estimates (θ), travel time values (δ), and the time-invariant components of both variables are absorbed into the household fixed effects.
Total labor (E_j^L)	–	Accounted for in household size and household fixed effects.

[a] Household livestock includes cattle, sheep, goats, chickens, pigs, other poultry, and rabbits.

[b] Household education is also included as a taste-shifter variable in the firewood collection estimates.

[c] TSh denotes Tanzanian shillings. All prices are expressed in terms of 1991 prices (survey round 1).

Table 2.4: Household Profit: Ordinary Least Squares with Fixed Effects
Dependent variable: Log of profit from self-employment (business and agriculture)

	Cobb-Douglas	Translog	Cobb-Douglas	Translog
	(1)	(2)	(3)	(4)
Land acres (log)	−0.101	−0.048	−0.100	−0.046
	(0.072)	(0.124)	(0.073)	(0.121)
Value of livestock, TSh (log)	0.009	−0.001	0.008	0.001
	(0.014)	(0.047)	(0.014)	(0.047)
Average household education (log)	−0.043	−0.058	−0.051	−0.045
	(0.116)	(0.166)	(0.119)	(0.162)
Business assets, TSh (log)	0.010	−0.032*	0.010	−0.026
	(0.016)	(0.019)	(0.016)	(0.020)
Variable inputs, TSh (log)	0.224***	−0.083	0.223***	−0.082
	(0.026)	(0.086)	(0.026)	(0.086)
Adult male annual labor hours (log)	0.080**	−0.219		
	(0.039)	(0.359)		
Adult female annual labor hours (log)	0.110**	−0.877***		
	(0.054)	(0.290)		
Adult annual labor hours (log)			0.217***	−1.475**
			(0.067)	(0.729)
Teenager annual labor hours (log)	0.024	−0.061	0.023	−0.063
	(0.018)	(0.074)	(0.018)	(0.073)
Child annual labor hours (log)	−0.028**	0.180	−0.028**	0.209
	(0.011)	(0.137)	(0.011)	(0.182)
Adult male labor (log) squared		0.021		
		(0.027)		
Adult female labor (log) squared		0.075***		
		(0.021)		
Adult labor (log) squared				0.110**
				(0.049)
Adult female × Child 7-11 labor		−0.017		
		(0.012)		
Adult × Child 7-11 labor				−0.020
				(0.018)
Teenage labor (log) squared		0.012		0.013
		(0.010)		(0.010)
Child labor (log) squared		−0.013		−0.012
		(0.012)		(0.012)
Land acres squared (log)		−0.017		−0.016
		(0.045)		(0.044)
Livestock value squared, TSh (log)		0.001		0.001
		(0.005)		(0.005)

Continued on next page

Table 2.4 – continued from previous page

	Cobb-Douglas	Translog	Cobb-Douglas	Translog
	(1)	(2)	(3)	(4)
Household education (log) squared		0.016		0.006
		(0.065)		(0.065)
Business assets squared, TSh (log)		0.004		0.004
		(0.003)		(0.003)
Variable inputs squared, TSh (log)		0.023***		0.023***
		(0.006)		(0.006)
Observations	1,904	1,904	1,904	1,904
R^2 within	0.183	0.222	0.185	0.222
P-value for Wald test $\hat{\alpha}_f = \hat{\alpha}_m$	0.680			
P-value for Wald test $\hat{\alpha}_c = \hat{\alpha}_t$	0.016		0.019	
Adult male shadow wage[a]	15.014	7.792		
Adult female shadow wage[a]	16.590	15.317		
Adult shadow wage[a]			17.243	9.944
Teenage shadow wage[a]	8.944	29.070	7.793	26.910

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Standard errors clustered at the village level.

Note: Both estimates include additional dummies for interview month (coefficients not reported). Regressions include households with reported firewood sales. Sample used for this regression is based on direct estimates of net revenue for home-business and excludes households with negative reported incomes (141 observations total). A one is added to reported-zeros for household income, teenage labor hours, child labor hours, non-farm business assets, household education, livestock value, land acres, and variable inputs. Only 1,904 observations have non-zero female *and* male labor hours.

[a] Shadow wages are calculated as $\hat{w}_j = \frac{\partial \mathrm{E}[\ln profit]}{\partial \ln L^a_j} \times \frac{\hat{profit}}{L^a_j}$ for $j = m, f, t$. Thus, for the case of the Cobb-Douglas production function the shadow wage estimates are equivalent to $\hat{w}_j = \hat{\beta}_j \times \frac{\hat{profit}}{L^a_j}$ where $\hat{\beta}_j$ is the estimated coefficient on group j's annual labor hours variable. Sample means reported over all 3,337 observations. In columns (2) 516 adult male, 631 adult female, and 97 teenage shadow wages are dropped because they are estimated to be negative and, similarly, in column (4) 1,026 adult and 15 teenage shadow wages were dropped. Child shadow wages are omitted because they are always estimated to be negative.

Table 2.5: Travel Cost Summary Statistics Using Reported and Estimated Wages

	Survey Round				
	1	2	3	4	Total
Adult male weight	0.37	0.37	0.33	0.32	0.34
	(0.40)	(0.40)	(0.38)	(0.37)	(0.39)
Adult female weight	0.29	0.27	0.29	0.30	0.29
	(0.37)	(0.36)	(0.37)	(0.37)	(0.37)
Child weight	0.10	0.11	0.12	0.12	0.11
	(0.21)	(0.21)	(0.23)	(0.23)	(0.22)
Teenage weight	0.20	0.21	0.22	0.21	0.21
	(0.29)	(0.29)	(0.29)	(0.30)	(0.29)
Travel time (hours)	1.70	1.56	1.45	1.37	1.52
	(1.11)	(1.08)	(0.90)	(0.91)	(1.01)
Adult shadow wage	20.11	16.86	17.31	14.61	17.24
	(20.25)	(21.55)	(16.58)	(20.38)	(19.86)
Adult sample wage[a]	17.88	20.24	22.27	23.20	20.88
	(5.72)	(6.53)	(8.19)	(8.26)	(7.52)
Teenage shadow wage	9.98	7.68	6.69	6.72	7.77
	(10.83)	(6.57)	(5.79)	(6.85)	(7.88)
Teenage sample wage[a]	13.63	19.38	15.22	21.12	17.31
	(3.76)	(5.20)	(4.70)	(9.03)	(6.70)
Travel cost (shadow wage)	23.93	16.86	15.53	13.48	17.48
	(30.26)	(25.81)	(20.09)	(48.11)	(32.83)
Travel cost (sample wage)	25.05	28.21	25.27	27.04	26.39
	(20.74)	(24.32)	(22.52)	(27.09)	(23.78)
Observations	838	848	840	811	3337

Standard deviation in parentheses. [a] Values normalized with round 1 as base prices.

Note: Predicted profits were obtained for the entire sample population by replacing the 257 households with zero reported adult labor hours with a value of one (64 households in round 1, 68 in round 2, 62 in round 3, and 63 in round 4). Finally, shadow wages could not be predicted for households with reported adult or teen labor hours of zero.; these households were assigned values equal to their village average shadow wage in a given round. This assignment resulted in the same 257 adult shadow wages as above being replaced and 1,868 teenage shadow wages (503 households in round 1, 473 in round 2, 445 in round 3, and 447 in round 4) being replaced. In addition, there are 418 observations (82 households in round 1, 102 in round 2, 113 in round 3, and 121 in round 4) with zero reported weekly firewood collection trips. These households receive weight, trip time, and travel cost values equal to the village average in a given round. Sample does not include households that report positive firewood sales (38 observations).

[a] Villages with missing reported wage data are given a wage equal to the sample mean for the relevant survey round.

Table 2.6: Household Travel Cost Estimates: Poisson Fixed Effects
Dependent variable: Weekly household fire collection trips

	Perfect labor markets	Constrained labor markets	Corner markets
	(1)	(2)	(3)
Travel cost (sample wage)	−0.018***		
	(0.001)		
Travel cost (shadow wage)		−0.012***	
		(0.003)	
Travel time (hours)			−0.392***
			(0.022)
Household expenditure, TSh (log)	0.136***	0.162***	0.180***
	(0.037)	(0.041)	(0.035)
Price of kerosene	0.005	0.004	0.007
	(0.004)	(0.005)	(0.005)
Price of charcoal	−2.267	−0.758	−4.423
	(5.944)	(6.804)	(6.001)
Food price index	1.366	1.651	2.836
	(1.936)	(2.132)	(2.027)
Household size	0.072***	0.063***	0.070***
	(0.014)	(0.014)	(0.013)
Female household head	−0.151	−0.155	−0.167
	(0.131)	(0.155)	(0.120)
Average household education	0.013	0.001	0.008
	(0.010)	(0.012)	(0.010)
Members with restricted activity	−0.006	−0.008	0.001
	(0.019)	(0.021)	(0.020)
Household owns bicycles	−0.072	−0.025	−0.060
	(0.052)	(0.062)	(0.051)
Household owns car	−0.259	−0.226	−0.339
	(0.308)	(0.435)	(0.344)
Household owns motorcycle	0.881***	0.982	0.920***
	(0.308)	(1.926)	(0.342)
Observations	3170	3337	3170
Log-likelihood	−6200.12		−6181.72
Chi-squared	411.23		534.29
WTP per trip (Tz shillings)	55.38	83.87	
WTP per trip confidence interval (95%)	50.68 – 61.30	58.89 – 147.40	

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Cluster robust standard errors in parentheses for the perfect and corner labor market scenarios. For constrained labor markets, standard errors are block-bootstrapped over both the first stage (household profit function) and second stage (travel cost).

Note: Willingness to pay measures reported in local currency units (Tanzanian shillings). Households with zero firewood collection trips in *all* periods were dropped from the estimation (35 households and 120 observations total) and households with only one observation (42 households) were dropped.

Table 2.7: Household Travel Cost Estimates: Robustness Checks

Dependent variable: Weekly household firewood collection trips

	Poisson FE		Negative binomial FE	OLS FE	Ordered Multinomial	
	(1)	(2)	(3)	(4)	(5)	(6)
	HH majority	No wood crop			Probit[a]	Logit[a]
Travel cost (shadow wage)		−0.010**	−0.007***	−0.020	−0.005**	−0.019***
		(0.004)	(0.002)	(0.014)	(0.003)	(0.005)
Travel cost (HH majority)	−0.009***					
	(0.003)					
Value of HH expenditure (log)	0.154***	0.116**	0.097**	0.701***	0.024	0.018
	(0.041)	(0.051)	(0.045)	(0.245)	(0.056)	(0.098)
Price of kerosene	0.002	0.000	−0.001	0.017	0.013	0.023
	(0.005)	(0.006)	(0.003)	(0.031)	(0.010)	(0.018)
Price of charcoal	1.742	3.661	9.104*	18.847	−15.175	−31.259
	(6.235)	(7.518)	(4.961)	(42.310)	(12.289)	(22.349)
Food price index	0.867	0.896	−0.848	3.968	−0.260	−0.842
	(2.017)	(2.458)	(1.159)	(12.721)	(4.502)	(8.160)
Household size	0.060***	0.063***	0.017	0.459***	0.114***	0.196***
	(0.014)	(0.020)	(0.014)	(0.097)	(0.013)	(0.023)
Female household head	−0.175	−0.018	−0.116	−1.110	0.099	0.205*
	(0.141)	(0.182)	(0.110)	(0.945)	(0.065)	(0.107)
Average household education	−0.000	−0.001	−0.001	−0.030	−0.035***	−0.060***
	(0.011)	(0.014)	(0.010)	(0.060)	(0.011)	(0.019)
Members with restricted activity	−0.007	−0.011	−0.003	−0.129	0.020	0.049
	(0.021)	(0.026)	(0.021)	(0.170)	(0.027)	(0.046)
Household owns bicycle	−0.019	0.005	−0.062	−0.337	−0.169***	−0.237***

Continued on next page

Table 2.7 – continued from previous page

	Poisson FE		Negative binomial FE	OLS FE	Ordered Multinomial	
	(1)	(2)	(3)	(4)	(5)	(6)
	HH majority	No wood crop			Probit[a]	Logit[a]
	(0.060)	(0.073)	(0.060)	(0.355)	(0.048)	(0.081)
Household owns car	−0.206	−0.018	−0.545	−1.288	−0.318**	−0.411
	(0.462)	(0.469)	(0.451)	(1.342)	(0.131)	(0.251)
Household owns motorcycle	0.983	0.790***	−0.318	2.813*	−0.231	−0.620
	(2.367)	(0.212)	(0.626)	(1.461)	(0.242)	(0.468)
Observations	3175	2390	3175	3333	2379	2390
WTP per trip (Tz shillings)	112.64	100.89	134.47	152.58		
WTP per trip confidence interval (95%)	73.88 – 237.76	61.50 – 273.84	86.87 – 300.68	72.29 – 1317.53		

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Two stage block-bootstrap standard errors displayed for columns (1) through (4). Cluster robust standard errors displayed for columns (5) and (6).

Note: Willingness to pay measures reported in local currency units (Tanzanian shillings).

[a] Columns (5) and (6) include household elevation, urban dummy, and round and village fixed effects. Coefficients not reported.

Table 2.8: Estimated and Observed Wages and Incomes

	Round				
	1	2	3	4	Total
WTP for annual forest access (1991 TSh)[a]	30,125.35	26,424.70	24,408.35	26,835.52	26,925.55
	(27513.65)	(25477.18)	(23472.51)	(23700.74)	(25172.24)
Annual firewood consumption (1991 TSh)[b]	15,055.69	8,611.59	7,241.44	6,882.98	9,460.59
	(13567.99)	(7801.93)	(5526.08)	(5348.78)	(9331.24)
Elasticity of trips with respect to travel cost	-0.29	-0.20	-0.19	-0.16	-0.21
	(0.36)	(0.31)	(0.24)	(0.57)	(0.39)
WTP as proportion of household expenditure (%)	13.62	28.86	26.14	32.70	25.27
	(17.28)	(37.29)	(30.78)	(46.67)	(35.29)
Observations	835	848	840	810	3337

Standard deviation in parentheses.

[a] Statistic calculated as $-y_{kvt}/\hat{\beta}_1$. Standard deviations come from variations in y_{kvt} in the sample population.

[b] Values normalized with round 1 as base prices.

Figure 2.1: Kagera Region, Tanzania

Figure 2.2: Weekly Household Firewood Collection Trips by Month of Interview

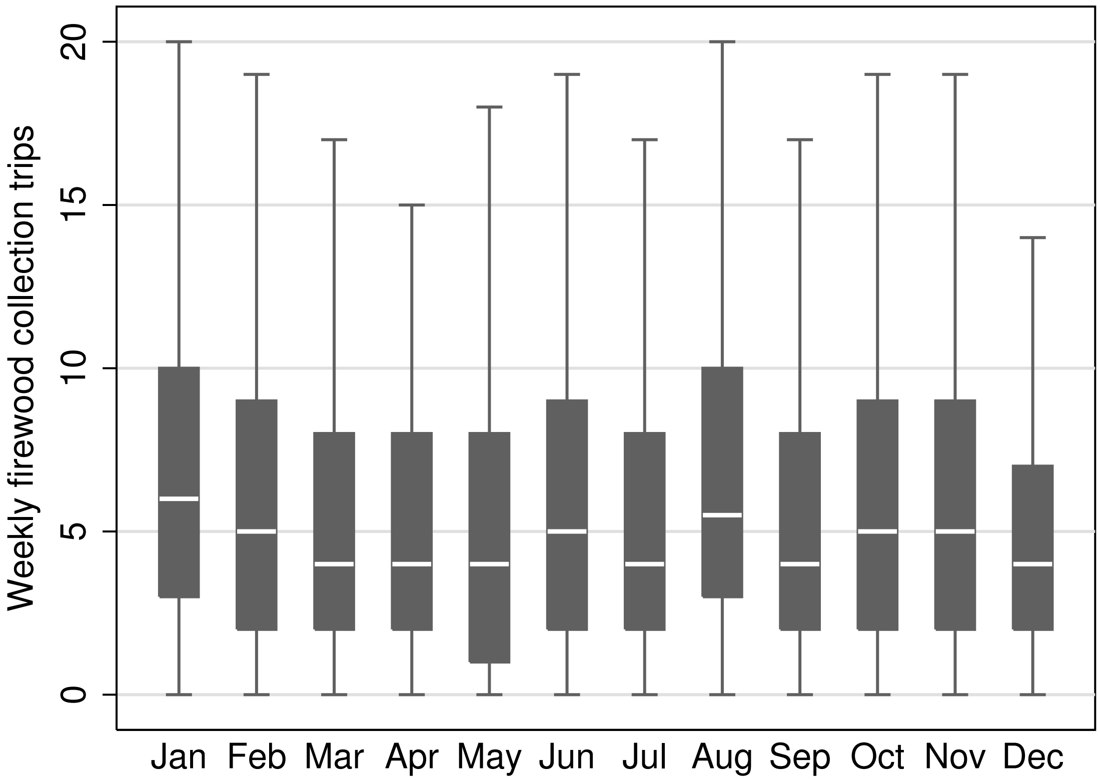
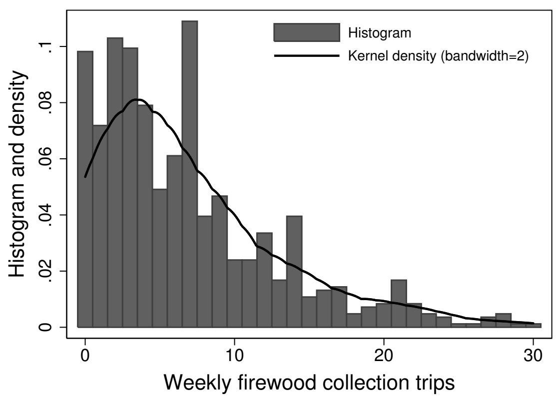
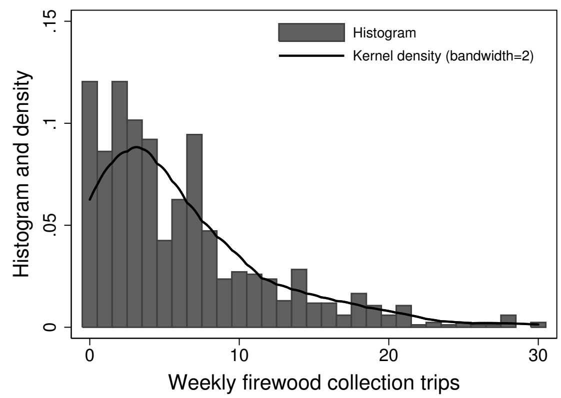


Figure 2.3: Weekly Household Firewood Collection Histogram and Density

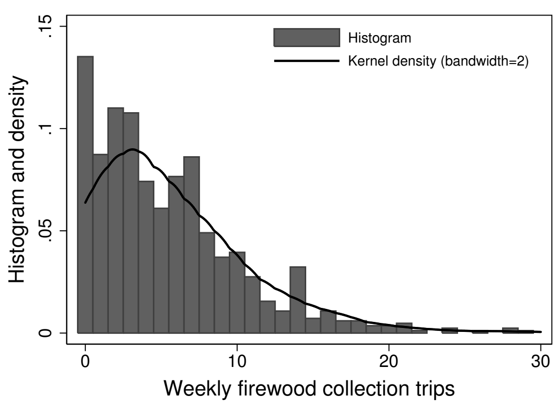
(a) Survey Round 1



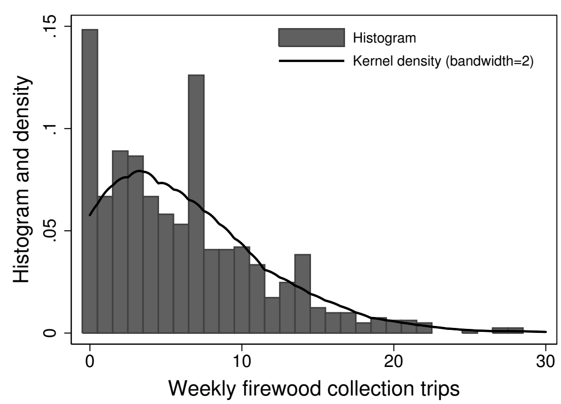
(b) Survey Round 2



(c) Survey Round 3



(d) Survey Round 4



Note: Histogram and kernel density plots drop the 10 households that report more than 30 firewood collection trips per week (3 households in round 1, 1 in round 2, 4 in round 3, and 2 in round 4).

Chapter 3

Forest Access and Human Capital Accumulation

3.1 Introduction

Forests provide important ecosystem services to many people in Africa. Approximately 90 percent of the continent's population uses firewood for cooking (Agyei, 1998) and 90 percent of wood used is for firewood or charcoal production (Hassan, Scholes and Ash, 2005). Firewood is predominately collected and used outside of the formal market system with households collecting firewood by either chopping down entire trees or particular branches. Despite this high level of forest use in Africa, the continent is currently experiencing a decline in its forest cover. The most recent Millennium Ecosystem Assessment estimates that Africa accounts for over 50 percent of current global deforestation (Hassan, Scholes and Ash, 2005); in particular, Tanzania has lost 81,000 square kilometers of forests over the last 20 years, a 19 percent decline in forests that is equivalent to 9 percent of the country's total area (World Bank, 2010). Deforestation and forest degradation reduce a household's access to firewood and force households to devote more of their time to collecting firewood. But, time is a limited resource, more time devoted to firewood collection makes less time available for other household activities.

In this paper, I test for the effects of a reduction in access to natural capital on human capital investment, human capital accumulation, and an individual's long-term earning potential. Specifically, I test for the effects of forest access on human capital formation in Kagera Region, Tanzania using the Kagera Health and Development Survey (KHDS). Forest access may affect school attendance and human capital formation through the influence it has on time needed for firewood collection. As in the rest of the country, firewood collection is an important component of household welfare in Kagera. Between 1991 and 1994, 95 percent of households report using firewood as the primary source of cooking fuel, while 90 percent of households collect firewood themselves. On average, households spend approximately eight hours per week collecting firewood and children ages 7 to 15 account for roughly two-thirds of all household firewood collection trips. This heavy reliance on the natural resource coupled with Tanzania's high rate of deforestation implies that children could be forced to compromise their school attendance as access to local forests becomes more limited.

I use a sample size of over one thousand children ages 7 to 15 to empirically link forest access to both school attendance and long-term human capital formation. In the short-run, I find that a one hour increase in firewood collection trip time results in a child spending 25 minutes less in school per week, regardless of whether the child is collecting firewood him or herself. Over the long-term, this reduction in weekly school attendance due to a one hour increase in firewood collection time when the child is young translates into a child completing one-fifth fewer grades 19 years later. Previous studies have estimated an 8 percent annual return to education in Tanzania (Psacharopoulos and Patrinos, 2004), and using this rate a one hour increase in firewood collection time implies a 1.7 percent reduction in annual income when the child is older, or a net present value of \$475 in 2010 USD over the course of 30 years. These figures are the first to attempt to quantify the impact of forest access on human capital formation. Together they imply that there is a significant effect of reduced forest access on educational achievement with large aggregate costs to the Tanzanian economy – 3 million cumulative years of lost education if all 15 million rural children in Tanzania were affected.

By using a long-term panel survey, this paper also benefits from data that explicitly links children's access to forests when they are young with their long-term education levels. Most retrospective studies that look at school attendance and school achievement use cross-sectional data sets and fail to control for unobserved individual characteristics that are likely to lead to omitted variable bias (Glewwe, 2002). In addition, ascertaining the long-term effects of reduced forest access on human capital accumulation is much more difficult with cross-sectional data because the researcher only knows an individual's forest access level (for children 7 to 15) or the number of years of education completed by an individual (for adults). Thus, making a causal statement about the effects of forest access when children are young and attending school on long-term education levels is difficult empirically. I use a panel data set covering 19 years with detailed information on children both when they are young (7 to 15 years) and when they are older (26 to 34 years) and have completed their education. For the analysis, I track the same individuals across their lifetimes and observe how educational outcomes differ solely based on variation in early childhood firewood collection time.

Ultimately, this paper is one of the first papers to empirically link reduced access to natural capital with lower human capital accumulation and to quantify the magnitude of the effects. This paper also takes advantage of a rare long-term panel survey to show that reduced forest access at a young age has long-term effects on an individual's education level. A reliable derivation of this dollar-value estimate can then be used in policy development and provide more information on the non-market costs of deforestation and forest degradation.

This paper proceeds as follows: In Section 3.2, I review the literature and further highlight the contributions made by this paper. In Section 3.3, I describe the data before proceeding to the short-term analysis in Section 3.4 and the long-term analysis in Section 3.5. Finally, in Section 3.6, I summarize my results and provide concluding remarks.

3.2 Literature Review

This paper draws on the firewood collection and the child labor literature; there is almost no literature that directly links forest access with school attendance and human capital formation. The one exception is a recent paper by DeGraff, Levison and Dungumaro (2014) that estimates the link between a child's times fetching water or firewood and their school attendance. I extend on this paper by also analyzing the effects of increased firewood collection time on long-term school completion levels.

There is a relatively large set of literature on the relationship between household labor, firewood collection, and the environment. While much of the early literature focused on firewood collection in South Asia, specifically Nepal, where there is a formal market for firewood sales, more recent literature examines firewood collection labor in sub-Saharan Africa. In what follows, I explain how this broad set of literature motivates and substantiates two important claims that I make in this paper: First, that firewood collection time is negatively correlated with forest access and, second, that time spent collecting firewood affects intra-household time allocation.

In this paper, I am interested in measuring the extent to which forest access, and changes in forest access, affect human capital formation. Forest access can be measured in two ways: one, by directly observing the distance from a household to the nearest community forest and the quality of that forest and, two, by observing a household's relative use of each forest. I measure forest access using the latter of these two methods; for each household, I measure the amount of time a single firewood collection trip takes. Previous studies provide evidence that better access to forests reduces firewood collection time (Cooke, 1998*b*; Amacher et al., 2004; Kohlin and Amacher, 2005). But, the magnitude of a household's response to a change in forest access depends on whether substitute fuels are available in the local market (Cooke, 1998*b*; Heltberg, Arndt and Sekhar, 2000; Chen, Heerink and Van Den Berg, 2006) and whether there are strong off-farm labor markets (Bluffstone, 1995; Shively and Fisher, 2004; Fisher, Shively and Buccola, 2005). Thus, by measuring forest access using firewood collection trip time I not only capture information on the quality of the forests but also capture critical information on local markets and the availability of cooking fuel substitutes.

Second, this paper tests whether forest access, measured as firewood collection trip time, negatively impacts school attendance and human capital formation. This hypothesis is motivated by previous literature that has estimated the effects of increased firewood collection time on household agricultural labor. If an increase in firewood collection time negatively affects household agricultural labor then there is reason to believe that an increase in firewood collection time will also negatively affect school attendance. Unfortunately, the evidence on the direction of the relationship between firewood collection and agricultural labor is mixed: both Cooke (1998*a*,*b*) and Amacher et al. (2004) show that there is no effect, while Kumar and Hotchkiss (1988) show a negative effect. A second subset of literature, however, has shown that children are more likely to leave school as local labor market conditions improve (Duryea, 2003; Edmonds, 2007; Kruger, 2007).

An increase in firewood collection trip time may affect a children's school attendance regardless of whether or not the child collects firewood. There is substantial evidence that many households in developing countries employ children in home labor activities, including firewood collection but also fetching water, tending to livestock, and cleaning (Amacher, Hyde and Joshee, 1993). Even though education has large social gains, it may still be individually optimal for parents to employ a child at home in order to free up parents' time to work in other income-earning activities (Basu and Van, 1998; Baland and Robinson, 2000). Consequently, even if children do not collect firewood themselves an increase in firewood collection time may trigger an increase in firewood collection time for their parents. As parents devote more time to firewood collection they may shift additional home work onto the child so that the parents can maintain the time they devote to income-earning activities. Thus, this study contributes to this current debate by providing an additional piece of empirical evidence on the effects of firewood collection labor on intra-household time allocation and school attendance.

3.3 Data

The data used in this paper come from the 1991-1994, 2004, and 2010 survey rounds of the Kagera Health and Development Survey (KHDS). The Kagera region (40,838 km^2) lies in the northwest corner of Tanzania on the western shore of Lake Victoria and borders Uganda, Rwanda, and Burundi (see Figure 3.1). The first four survey rounds of the KHDS were conducted once every six or seven months, between October of 1991 and January of 1994. The 2004 survey round was conducted between January and August of 2004 and the 2010 survey round between April and December of 2010. In both 2004 and 2010, the KHDS tracked as many initial respondents as they could that had moved both within Tanzania and to neighboring Uganda.

For the analysis, I need a sample of children who were of school age in the initial four survey rounds, so I focus on children who were between 7 and 15 years old during the first survey round (October 1991 to May 1992). These children were in primary school (ages 7 to 13, grades 1 to 7) or had just started secondary school (ages 14 to 17, forms 1 to 4). Summary statistics for this sample are presented in Table 3.3. Roughly half the sample is female with the majority of children living in rural areas. On average, children spent approximately 15 hours per week in school in the first four survey rounds and had completed seven years of education by 2010. By 2010, roughly 40 percent of the observed sample had moved to another village within Kagera and roughly 20 percent have moved outside of the Kagera region, either within Tanzania or to another country.

Firewood collection trip time is approximately one and a half hours in all of the first four survey rounds. Trip time is measured as the average time (in hours) spent collecting firewood per day across all household members with non-zero firewood collection trips.[1] I use average household trip time to minimize any potential measurement error in an individual's trip time. Children that do not collect firewood themselves but reside in a household that does collect firewood are then assigned this average household trip time. Only 50 percent of the children in the sample report non-zero firewood collection trips, so without this imputation I would be forced to drop roughly 700 individuals from the analysis. In addition, households that did not make any firewood collection trips are assigned a hypothetical trip time equal to their village average trip time.[2] As long as households that collect and do not collect firewood are randomly located throughout the village this village average replacement will adequately represent forest access for non-collecting households. Consequently, variation of firewood collection trip time is at the household and time period level.

In total, there are 1,442 individuals from 621 different households in the sample, but with

[1]The exact survey question used to derive household trip is "How many hours did you spend collecting firewood for each of the last seven days?" that was asked to all household members 7 years and older. Trip time is then measured as the total number of hours an individual reports collecting firewood divided by the number of days an individual collects firewood. Assuming individuals make only one firewood collection trip per day, the number of trips made is equal to the number of days an individual spent collecting firewood.

[2]In total, 90 percent of children in each survey round have a trip time equal to their household average trip time and only 10 percent of children live in households that make no firewood collection trips and have a trip time equal to the village mean (9.6 percent in round 1, 5.5 percent in round 2, 11.1 percent in round 3, and 10.4 percent in round 4).

attrition this number decreases throughout the survey rounds to a low of 881 individuals in 2004. Attrition in the first four survey rounds is low with only seven percent of individuals leaving the survey by round four. By the 2004 survey round, however, 39 percent of the sample was not interviewed and 30 percent by 2010, both attrition rates are relative to the initial sample size of 1,442.[3] The reduction in attrition between the 2004 and 2010 survey rounds is largely a result of higher re-interview rates among children initially 10 to 19 years old; only 77 percent (1,453 individuals) of surviving 10 to 19 year olds were re-interviewed in 2004 but 84 percent (1,523 individuals) were re-interviewed in 2010 (Weerdt et al., 2012). Attrition in the first four survey rounds is unlikely to bias coefficient estimates, but attrition could lead to bias in the last two survey rounds, by which time over a third of individuals were not re-interviewed had attrited. In the long-run, attrition will cause bias from self-selection if it is non-random and correlated with long-term education levels (Baulch and Quisumbing, 2011). In the empirical analysis, I adjust for attrition with selection on observables by using inverse probability weights (Fitzgerald, Gottschalk and Moffitt, 1998).

3.3.1 Putting Kagera in Context

As described above, the data used in this paper only contain individuals that initially lived in the Kagera region of Tanzania. Kagera is a remote area of Tanzania but households in Kagera are comparable to the rest of the country. As shown in Table 3.1, roughly 80 percent of households in the survey live in rural areas compared to national rural population estimates of 81 percent for the same time period (World Bank, 2010). The vast majority of households participate in agriculture and livestock herding; over 90 percent of households report spending time on agriculture activities in the last week and almost two-thirds of households report spending time on livestock herding. These rates are only slightly higher than the 1991 World Development Indicators' estimate that 84 percent of individuals were employed in agriculture. Most pertinent to this analysis, over 95 percent of households report firewood as their primary source of cooking fuel in all four survey rounds and almost 90 percent of households collect firewood themselves.

Second, I compare education levels and school attendance levels in Kagera during the

[3] Of the 460 individuals that were not interviewed in round 4 and 2004, 51 of the cases were due to death. There were an additional 8 deaths between the 2004 and 2010 survey rounds.

initial four survey rounds to national education levels. As shown in Table 3.2, approximately 60 percent of sample children were enrolled in school in the first four survey rounds and enrollment rates were almost identical for girls and boys. These estimates are slightly higher than the primary school net enrollment rate of 50 percent estimated by the United Nations Educational, Scientific, and Cultural Organization but are not implausibly far off. In the initial survey round, the sample of 1,442 children had completed, on average, 2.6 years of education and this number rises to 2.9 by the fourth survey round. Across all four survey rounds children are spending approximately 2 days a week in school totaling 14 weekly hours (or about 6.5 hours a day in school on average). There is also little difference between males and females in terms of the number of grades completed, hours spent in school a week, and days spent in school a week.

These statistics demonstrate that Kagera is only slightly more agricultural than the rest of Tanzania but school attendance is similar to the rest of country. In the initial four survey rounds, education levels were relatively low with less than three completed grades but this figure does rise substantially by the final survey round in 2010 (specifically 7 years, see Table 3.3). These statistics provide evidence that the results found in this paper may be generalizable to all of Tanzania.

In what follows, I first investigate the short-term effects of firewood collection trip time on weekly school attendance using the 1991 to 1994 survey rounds. The empirical strategy and results for this short-term estimation are presented in Section 3.4. I then investigate the long-term effects of firewood collection trip time in 1991 on school completion levels using the 2004 and 2010 survey rounds. The empirical strategy and results for this long-term estimation are presented in Section 3.5. In both estimations, I exploit spatial and temporal variation in firewood collection trip time (in hours) across households. Trip time is positively correlated with a household's distance to the forest and the greater a household's distance to the forest the more scarce the forest resource. Trip time could be affected by a household's mode of transportation to collect firewood. In Kagera, however, walking is the dominant method of collecting firewood – less than two percent of households own cars or motorbikes and only thirty percent of households own bicycles.

3.4 Short-Term School Attendance Effects in 1990s

In the short-term, I investigate the extent to which an increase in firewood collection trip time affects the number of hours each week that a child spends in school using data from the initial four survey rounds, 1991 to 1994. An increase in time spent collecting firewood could negatively affect weekly school attendance for both children that are and are not currently collecting firewood. For children that are currently collecting firewood, increased collection time forces the child to spend more time each week collecting firewood and less time each week in school. For children that are not currently collecting firewood, increased collection time increases the time that other household members have to spend collecting firewood which, consequently, could cause these children to participate in additional household chores and reduce the amount of time each week that they spend in school. As shown in Table 3.4 it is quite common for children that do not collect firewood to participate in other household chores; over 30 percent of children do not collect firewood but do participate in other household chores and the majority of children that do collect firewood participate in other household chores.

3.4.1 Empirical Strategy

In the short-term estimation, I specify the number of hours in the last week that a child was in school as the outcome variable. I also use a set of explanatory variables to control for individual, village, and household characteristics that affect weekly hours a child spends in school. Individual-level variables include current age and gender of the child. I include village-level fixed effects to control for all time-invariant observed and unobserved village characteristics, including whether or not the village is urban, the number of primary schools in the village, and unobserved environmental and school access characteristics.[4] I also include survey round dummies and interview month dummies for the lean season (February), harvest season (June and July), post-harvest season (August), and school break (December).[5]

[4] I also ran the analysis with controls for the price of kerosene and the price of charcoal but the coefficient on trip time was not affected by the inclusion of these variables.

[5] I also ran the analysis using interview month dummies for all twelve months but the coefficient on trip time was also not affected by the inclusion of these variables.

Household-level variables include firewood collection trip time, gender of household head, mother's education, father's education, household size, and annual household expenditures (log). Both mother's and father's education levels are measured using two mutually-exclusive dummy variables. First, there is a dummy variable equal to one if the parent has some years of primary education, but did not finish primary schooling. Second, there is a dummy variable equal to one if the parent has either completed primary school or been educated beyond primary school. The base category for each of the parents' education variables is parents who have not completed any years of schooling or have missing data. Household size is measured as the (unweighted) number of individuals living in the same dwelling and eating together three-quarters of the time in the last year (Ainsworth, 2004).[6] Lastly, I use annual household expenditures in the analysis as a proxy for household income because annual expenditures are strongly correlated with annual income but have less measurement error (Deaton, 1997).[7]

The general short-term model that I estimate for individual i in village v at time t has a conditional mean given by:

$$\mathrm{E}[school\ hours_{ivt}|trip\ time_{ivt}, \boldsymbol{x}_{ivt}] = g\Big(\beta_0 + \beta_1 trip\ time_{ivt} + \boldsymbol{x}_{ivt}\boldsymbol{\alpha} + \gamma_v + \delta_t\Big). \tag{3.1}$$

where $\boldsymbol{x}_{ivt}$ is vector of the individual characteristics, household characteristics, and interview month dummies, γ_v denotes village fixed-effects, and δ_t denotes survey round fixed-effects. The function $g(\cdot)$ is a general function and the form it takes will vary by the type of estimation method. In this short-term analysis, I am interested in the coefficient estimate on *trip time*. Specifically, if environmental access has a negative effect on school attendance then β_1 will be negative.

In choosing the appropriate estimation strategy, I need to account for not only the censored nature of the data (no child reports spending less than zero hours per week in

[6]The analysis was run with separate coefficients for the number of female children, male children, female adults, male adults, female elders, and male elders in the household but the coefficient on trip time was not affected by the inclusion of these more expansive household size variables.

[7]Annual household expenditures includes food expenditures, consumption from home production, non-food consumption, remittances sent, and wage income in kind (Ainsworth, 2004).

school) but also for the fact that the data are highly skewed; almost 50 percent of children report not attending school in each of the four survey rounds. Thus, the distribution of weekly school hours is not normal but the log of weekly school hours better approximates a normal distribution. In general, there are two ways to estimate a lognormally distributed outcome variable: estimating the relationship $\log \textit{school hours} = \boldsymbol{x}'\boldsymbol{\beta}$ or estimating $\textit{school hours} = \exp(\boldsymbol{x}'\boldsymbol{\beta})$, where $\boldsymbol{x}$ is the set of of explanatory variables described above. I focus my estimation on a generalized linear model with log-linear relationship using the Poisson distribution (GLM-LL) (Cameron and Trivedi, 2005; Nichols, 2010) that estimates the relationship $\textit{school hours} = \exp(\boldsymbol{x}'\boldsymbol{\beta})$. Robustness checks are done using Tobit, trimmed least absolute deviations (LAD), and ordinary least squares (OLS) model estimates that model the relationship $\log \textit{school hours} = \boldsymbol{x}'\boldsymbol{\beta}$.

With the GLM-LL Poisson estimation I am estimating a special case of (3.1) where the conditional mean follows an exponential function:

$$\mathrm{E}[\textit{school hours}_{ivt} | \textit{trip time}_{ivt}, \boldsymbol{x}_{ivt}] = \exp\Big(\beta_0 + \beta_1 \textit{trip time}_{ivt} + \boldsymbol{x}_{ivt}\boldsymbol{\alpha} + \gamma_v + \delta_t\Big).$$

The GLM-LL Poisson estimation strategy is advantageous for a number of reasons. First, by estimating an exponential function I allow the actual realized value of weekly hours in school to be zero ($1 = e^0$). This result does not hold when I estimate $\log \textit{school hours} = \boldsymbol{x}'\boldsymbol{\beta}$ using an OLS or tobit model. Because the natural log of zero is undefined for these models, I must replace zero observations with small positive numbers, η, so that they are not dropped from the estimation. The GLM-LL Poisson estimation avoids any sensitivity of the results to the choice of η because no transformation of the outcome variable is needed (Nichols, 2010).

Second, the GLM-LL Poisson estimation strategy produces consistent coefficient estimates under a wide variety of realized distributions. Although the Poisson distributional assumption must hold for estimates to be efficient, consistency of coefficient estimates requires only that the exponential conditional mean assumption holds. A Poisson distribution assumes that $\mathrm{E}[\textit{school hours}_{ivt}] = \mathrm{Var}[\textit{school hours}_{ivt}]$, and even though the failure of this assumption will not lead to inconsistency, it will affect standard errors (Gourieroux, Monfort and Trognon, 1984; Cameron and Trivedi, 2005). Consequently, the final estimation strategy that I employ in this paper is a quasi-maximum likelihood estimate (QMLE). I estimate the model by maximizing the log-likelihood function:

$$\log \mathcal{L}(\boldsymbol{\beta}) = \sum_{t=1}^{T}\sum_{v=1}^{V}\sum_{i=1}^{N}\Big(school\ hours_{ivt} - \exp(\beta_0 + \beta_1 trip\ time_{ivt} + \boldsymbol{x}_{ivt}\boldsymbol{\alpha} + \gamma_v + \delta_t + \epsilon_{ivt}) \\ - \ln school\ hours_{ivt}!\Big)$$

with the assumption that the mean and variance of weekly school hours are not equal (Cameron and Trivedi, 2005). Practically, this estimation is done by first estimating the Poisson model assuming a Poisson density and then calculating Eicker-White robust standard errors (Cameron and Trivedi, 2010). In contrast, the tobit model requires *school hours* to be lognormally distributed in order for coefficient estimates to be consistent (Cameron and Trivedi, 2005).

Finally, because the GLM-LL Poisson estimation strategy estimates the conditional mean, I can calculate the magnitude of the partial effect of firewood collection trip time on the mean. While the trimmed least absolute deviation model also does not require *school hours* to be lognormally distributed for consistent coefficient estimates, it only estimates the conditional median (Powell, 1984).[8] Because the median and mean number of hours spent in school a week are not the same, the marginal effects of these two coefficient estimates will not be comparable. In addition, the median number of weekly school hours is zero for two out of the four initial survey rounds so it is unclear how to interpret the effects of an increase in firewood collection trip time on the median when the median is zero. For these three reasons listed above, I rely on GLM-LL Poisson estimates of equation (3.1) as my preferred estimates but also estimate tobit, trimmed LAD, and OLS models as robustness checks; all three of these alternative models use $\log school\ hours_{ivt}$ as the outcome variable.[9]

[8] As mentioned above, the tobit model requires school hours to follow a lognormal distribution in order for consistent estimates. In addition, if school hours follows a lognormal distribution then tobit and trimmed LAD coefficient estimates will be equal. I assume that school hours follows a log-normal distribution and transform the variable by taking its natural log. The trimmed LAD estimates will not be affected by this transformation because the trimmed LAD is not sensitive to outliers but the tobit and OLS estimates will be affected by how observations of zero hours are replaced.

[9] Children that report spending no time in school are given a value equal to one ten-millionth less than the minimum non-zero value.

3.4.2 Results

Results from estimating the short-term school attendance indicate that an increase in travel time for firewood collection leads to a reduction in the number of hours a week a child spends in school. Results from the Poisson, tobit, trimmed LAD, and OLS estimation are all reported in Table 3.5. In all four regressions the coefficient on firewood collection trip time is statistically significant and negative, indicating that children spend less time per week in school as the time it takes to collect firewood increases. The marginal effects are reported in the second column for each model estimated in Table 3.5. These marginal effects can be interpreted directly: using the Poisson model, a one hour increase in trip time is associated with, on average, a child spending 24.6 (60×0.41) minutes less per week in school, holding all else constant. The tobit and OLS model show the same direction of effect although their marginal effects are larger (46.2 (60×0.77) and 36 (60×0.60) minutes respectively).

The other coefficients in the model indicate that children spend more hours in school as they get older (although this effect is non-linear) and as their annual household expenditure increases. Conversely, children spend fewer hours if they are female, if they were interviewed during the harvest season (June and July) or during school break (December). Household size is also positively correlated with school attendance. One possible explanation is that, after controlling for annual household expenditures, households with more children are more easily able to distribute housework among the children throughout the week so that each child spends more time in school.

In comparing the models, although the sign of the coefficients is the same across all four estimation approaches, the magnitude of the effect differs. First, as expected with censored data, the coefficient estimates from the tobit model are larger than the coefficient estimates from the OLS estimate. The tobit estimate is also different from the trimmed LAD model, which suggests that the normality distributional assumption in the tobit model does not hold and thus those estimates are biased. From the trimmed LAD model, a one hour increase in firewood collection trip time is associated with a 30 minute decrease in the conditional median of the time a child spends in school a week. Both for ease of interpretation and for robustness to distributional assumptions, I focus on the coefficient estimates from the Poisson model. In the next subsection, I test the sensitivity of this model to different samples and sources of variation.

3.4.3 Robustness checks

In order to investigate the extent to which the above results are sensitive to the estimation sample and the source of variation, I run an additional suite of robustness checks. To test whether it matters if the child collects firewood or just that his or her household collects firewood, I run the the GLM-LL Poisson model for the sample of children that live in households that collect firewood ("households that collect") and for the sample of children that collect firewood themselves ("children that collect"). I also test whether it matters if the household collect firewood by assigning households that do not collect firewood a trip time of zero hours instead of the village mean that was used in the initial estimates ("zero hours"). The results from these three robustness check are displayed in Table 3.6. The coefficient estimates in all three new estimates are negative and significant, showing that the results are not sensitive to whether the household or child collects firewood. While the magnitude of the effect is slightly larger for children that collect firewood themselves, the difference only translates into an additional 5 minutes of school time lost per week. The main factor driving the results appears to be children simply living in a household that collects firewood. These results indicate that increased firewood collection time impacts children similarly regardless of whether they collect firewood themselves or they simply live in households where other members collect firewood.

Next, to test whether spatial or temporal variation in forest access is more important, I run three different GLM-LL Poisson estimations where identification is based on intra-village, inter-village, or temporal variation in forest access. These results are displayed in Table 3.7. Note that in the first column identification is based on variation within villages and this specification is equivalent to the original GLM-LL Poisson model displayed in Table 3.5. In the two additional columns, identification is based on between-village variation (column 2) and on temporal variation (column 3). To identify the coefficient on firewood collection trip time on between-village variation, I estimate the model holding time constant, i.e. survey round fixed effects, and transform the outcome and explanatory variables to be village averages so that there is no within-village variation. To identify the coefficient on firewood collection trip time on temporal variation, I estimate the model holding all individual-level variables constant, i.e child fixed effects.

Although the coefficient on firewood collection time is negative in all three columns it is only significant in the first column (p-values of 0.99 and 0.16, respectively, for inter-village and temporal variation) using intra-village variation. These results indicate that the biggest driver of a child's school attendance is the variation in trip time within a village. Changes in trip time both across villages and across time do not lead to children spending less time in school. I cannot, however, definitively rule out the possible significant effects of changes in firewood collection both across villages and across time because both columns 3 and 4 suffer from omitted variable bias. The intra-village variation estimates in column 1 of Table 3.7 include both village and survey round fixed-effects and, consequently, also control for school quality characteristics. If school quality characteristics are significant drivers of the number of hours a child spends in school per week then the omission of these controls in columns 2 and 3 could explain why firewood collection trip time becomes insignificant.

3.4.4 Caveats

Estimating the determinants of weekly school attendance using retrospective data in developing countries is notoriously difficult to do because of endogeneity, measurement error, and unobservable child, household, and school characteristics (Glewwe, 2002; Glewwe and Kremer, 2006). In this section, I discuss these three estimation problems and the potential effects that each of them could have on my coefficient estimates.

The school that a child attends is potentially an endogenous decision that can lead to selection bias in coefficient estimates. For example, children that are more motivated may be able to receive scholarships to attend better schools in nearby cities or other regions. If these children choose to leave their home village then the children remaining in the village will be less motivated than the average student and more likely to attend school for fewer hours per week. If the children remaining in the village are not only less motivated but are also more likely to forgo school to collect firewood then the coefficient estimates could be biased downward, or larger in magnitude than if the estimates include the full sample of children. In the KHDS, only 10 percent of households report having children ages 7 to 18 living away in each of the first four survey rounds and the majority of children that live away from home do so because of custody issues. Indeed, less than 1.5 percent of children are reported as migrating for schooling in each of the first four survey rounds (less than 20 total children). Because I do not have information on the weekly school hours attendance of children living away from home, it is not possible to test for selection bias, but with such a small proportion of the sample living away from home for schooling selection, bias due to endogenous school selection is unlikely to be a problem.

Second, many household surveys contain a substantial amount of measurement error (Glewwe and Kremer, 2006). Of concern is measurement error in the explanatory variables (random or nonrandom) which will bias coefficient estimates. Most notably, in my analysis there may be measurement error around observed annual household expenditures and firewood collection trip time. By using the household average firewood collection trip time, I remove some of the measurement error associated with this variable but coefficient estimates should still be interpreted with caution.

Finally, unobserved child, household or school characteristics could cause omitted variable bias in my coefficient estimates. Examples of such unobserved omitted variables include child's motivation, parents' emotional support for their child's education, and a teacher's ability. I use data from a panel survey and run a child fixed effects model that controls for time invariant unobserved child characteristics (column 3 of Table 3.7). The coefficient on firewood collection trip time is still negative and of similar magnitude to the original estimates showing that these unobserved characteristics are not biasing my coefficient estimates.

The three estimation problems noted above are important empirical issues that are difficult to address with retrospective data. Furthermore, the bias that results from these issues could lead to both over- or under-estimation of the coefficients (Glewwe, 2002). Although my results should be interpreted with these caveats in mind, the robustness of the coefficient estimate on firewood collection trip time across all four model specifications, varying sample specifications, and varying levels of variation provides strong evidence that an increase in firewood collection trip time results in a reduction in children's weekly school attendance.

3.5 Long-Term School Completion Effects

The short-term analysis tells us the immediate effects that firewood collection trip time has on a child's weekly school attendance. In order to quantify the effects of collection time on human capital formation, it is important to know how firewood collection trip time as a child affects an individual's completed level of education. With a 40 week school year the short-term effect of 24 minutes less per week amounts to a child spending almost 17 fewer hours in school per year. And the effect could be even larger for long-term outcomes such as grades completed if children that spend less time in school when they are young are then more likely to get held back a grade or to drop out of school altogether. In this section, I investigate the extent to which firewood collection trip time when a child is young affects his or her completed years of education 13 and 19 years after survey round 1.

3.5.1 Empirical Strategy

This estimation relies on firewood collection trip time data and individual, village, and household characteristics from the initial survey round (round 1) and on completed education levels from the 2004 and 2010 survey rounds. I use firewood collection trip time in round 1 in order to obtain the largest possible sample of individuals and because trip time in round 1 is positively correlated with trip time in all three of the other initial survey rounds (Table 3.8). I specify the number of grades completed as the outcome variable. I use the same set of individual and household characteristics that are included in the short-term analysis.[10] I also include village-level fixed effects corresponding to an individual's village in the first four survey rounds. In addition, I include two indicators relating to migration decisions for individuals who no longer live in the same village as they lived in during rounds one through four. I include a dummy variable that equals one for individuals who moved within the Kagera region and another dummy variable for individuals who moved outside the Kagera region, either to another region in Tanzania or to Uganda.[11]

[10] These variables are age, gender, firewood collection trip time, mother's education, father's education, gender of household head, household size, and annual household expenditures.

[11] Because of the possibility that these variables are endogenous and chosen simultaneously with school enrollment, I also ran the regression without these two migration variables. The coefficient on firewood collection trip time, however, is not affected by the inclusion or exclusion of these two variables.

The first regression that I run to test for long-run school completion effects is a simple linear regression. For individual i initially from village v the regression is:

$$school\ years_{ivt} = \beta_0 + \beta_1 trip\ time_{iv,1} + \boldsymbol{x}_{iv,1}\boldsymbol{\alpha} + \gamma_v + \varepsilon_{ivt} \tag{3.2}$$

with t equal to one of the two later survey periods, 2004 or 2010. Again, $\boldsymbol{x}_{ivt}$ denotes the vector of explanatory variables, and γ_v indicates village fixed-effects corresponding to the child's reported village from the first four survey rounds. The term ε_{ivt} is assumed to be independent and identically distributed. Similar to the short-term analysis, in the long-term analysis I am interested in the coefficient estimate on *trip time*. If forest access has a negative effect on long-term human capital formation then β_1 will be negative. The magnitude of the coefficient can be used to derive the estimated loss in education years as a result of increased firewood collection time. This loss in education years can then be assigned a dollar value using estimates of the return to education. I estimate equation (3.2) twice, using both the 2004 and 2010 survey rounds.

As mentioned in Section 3.3, the 2004 and 2010 survey rounds have 39 and 30 percent attrition rates, respectively. If an individual's attrition is non-random and correlated with long-term education levels then these attrition rates will cause bias in the coefficient estimates (Baulch and Quisumbing, 2011). In the Appendix, I test for non-random attrition using a pooling test that tests whether the coefficients in equation (3.1) are equal for attritors and non-attritors (Becketti et al., 1988). I perform the test for both the 2004 and 2010 survey rounds, and for the Poisson, tobit, and log-linear models. The p-value for the F-test that is reported at the bottom of tables B.1 and B.2 has a null hypothesis that attrition is random. This null hypothesis is rejected at the one percent level in all six estimates providing strong evidence that attrition is non-random.

Very generally, survey attrition that is correlated with the outcome variable of interest results in a sample selection problem. For the case of selection on unobservables, one can use James Heckman's two-stage selection model (Heckman, 1979). Heckman's model, however, requires the researcher to identify a set of exclusionary variables that are correlated with attrition but not with ε_{ivt} in equation (3.2). However, if no exclusionary variable exists or if it has weak explanatory power in terms of predicting attrition then the two-stage selection model will not adequately adjust for sample selection from attrition (Wooldridge, 2010). An appropriate exclusionary variable in the case of attrition is often hard to come by because few variables affect attrition without also affecting education years (Fitzgerald, Gottschalk and Moffitt, 1998). In contrast, if selection is based on observables then inverse probability weights can be used to correct for sample selection. Inverse probability weights rely on auxiliary variables to explain attrition and the auxiliary variables may be correlated to both attrition and long-term education (Fitzgerald, Gottschalk and Moffitt, 1998; Baulch and Quisumbing, 2011).

The use of inverse probability weights results in a weighted least squares estimation strategy. More specifically, assume a linear relationship as given in equation (3.2) and that attrition is governed by the relationship:

$$A^*_{ikt} = \theta_0 + \boldsymbol{x}'_{ikt}\theta_1 + \theta_2 z_{ikt} + \nu_{ikt} \text{ where } A = \begin{cases} 1 & \text{if } A^* \geq 0 \\ 0 & \text{if } A^* < 0 \end{cases} \tag{3.3}$$

and if $A^* \geq 0$ the individual is no longer in the sample at time t. The vector $\boldsymbol{x}_{ikt}$ corresponds to the explanatory variables given in equation (3.2) and z_{ikt} is a new auxiliary variable. The only requirement for this selection model is that ν_{ikt} and ε_{ikit} are independent given the explanatory variables $\boldsymbol{x}$ – the auxiliary variable may be correlated with long-term education. Under these assumptions, Fitzgerald, Gottschalk and Moffitt (1998) show that a weighted least squares approach can be used to estimate equation (3.2) where the weights are given by:

$$w(z, \boldsymbol{x}) = \frac{\Pr[A = 0|\boldsymbol{x}]}{\Pr[A = 0|z, \boldsymbol{x}]}. \tag{3.4}$$

In practice, these weights are estimated by running a probit model with and without the auxiliary variable and using an individual's predicted probabilities from the regressions to estimate her weight that is then used in the weighted least squares estimation.

I follow the approach laid out in both Fitzgerald, Gottschalk and Moffitt (1998) and Baulch and Quisumbing (2011) and rely on lagged variables as my auxiliary variables.

Specifically, I include all explanatory variables, except interview month and the outcome variable, hours in school, from the baseline round 1 survey as my auxiliary variables. I construct weights separately for both the 2004 and 2010 survey rounds by running a restricted and unrestricted probit model for each round. These probit estimates are reported in Table 3.9 for both the restricted (without auxiliary variables) and the unrestricted (with auxiliary variables) models for survey rounds 2004 and 2010. The long-term school completion effects regression, equation (3.2), is then run using both ordinary least squares and weighted least squares.

3.5.2 Results

Results from this long-term model are displayed in Table 3.10. This table shows the results from both the ordinary least squares and weighted least squares approach. Table 3.10 also reports the ordinary least squares regression results for the sample of individuals that are present in both the 2004 and 2010 survey rounds (columns 5 and 6).

The coefficient of interest for this paper is the coefficient on firewood collection trip time in round 1. The coefficient is negative with a similar marginal effect in all of the estimates and is statistically significant in four of the six estimates. In all of the results using the 2010 survey round the coefficient on firewood collection trip time is negative and statistically significant. In these 2010 results, a one hour increase in firewood collection trip time in round 1 corresponds to an individual completing 0.21 fewer grades when she is older. When I look at individuals that were in both the 2004 and 2010 survey rounds the coefficient is still significant but slightly smaller in magnitude than the estimates in columns 3 and 4.

Overall, these results show evidence that increased firewood collection trip time when a child is younger leads to a reduction in total grades completed 13 to 19 years older. The significance of these results implies that an increase in firewood collection trip time has two important effects on human capital: one, an immediate effect on a child's school attendance and, two, a long-term aggregated effect on the education completed by that child. This aggregated effect could result from children that miss school at a young age due to increased collection time having more difficulty successfully completing grades or being more likely to eventually drop out of school.

In all six of the estimates reported in Table 3.10, annual household expenditure has a positive and significant effect on the years of education completed by an individual. In three of the six estimates, household size has a negative and significant effect. The weighted least squares estimation approach has little effect on the coefficient estimates; this result is partly due to the small difference between the unrestricted and restricted probabilities, as shown at the bottom of Table 3.10.

3.5.3 Caveats

The outcome variable in my long-term school completion analysis is the number of grades completed by an individual. This variable does not take into account how many years an individual has been in school which may be larger than the number of grades completed if she had to repeat grades. The United Nations Educational, Scientific, and Cultural Organization estimates that primary school repetition rates in Tanzania averaged 3.3 percent between 1991 and 1994, with similar rates for boys and girls. Therefore, the effect of environmental access on human capital formation may be larger than my estimates were it to take into account the 3 percent of children repeating grades. But given that 3 percent is a relatively low repetition rate, the change in effect size is likely to be small.

In addition to the above-mentioned measurement error in grades completed, all of the estimation problems described in subsection 3.4.4 continue to hold. Because this survey tracked individuals that moved either to Uganda or another region in Tanzania, I have information on the years of education completed for children that left their home villages between 1994 and 2004 or 2010 in order to attend school. In contrast, in the first four initial survey rounds no attempt was made to track individuals that moved abroad or to other regions in Tanzania. Thus, there is likely to be less bias due to sample selection in the long-term analysis relative to the short-term analysis since I have information on both the children that stayed in their home villages and children that left. Measurement error, however, may still be present in the explanatory variables and unobservable individual characteristics could cause omitted variable bias.

3.5.4 Quantifying the Effects

In the previous subsection I show that individuals from households with shorter firewood collection trip times in round 1 had completed more years of education by 2010. Specifically, individuals that took one less hour to make a firewood collection trip in round 1 had, on average, almost a fifth of a year more of education. Combining return to education values with the results in Table 3.10 produces a monetary value of the costs of more restricted forest access on human capital.[12] These values are displayed in Table 3.11 for ten different return to education values ranging from one percent to ten percent. Psacharopoulos and Patrinos (2004) show an estimated rate of return of 7.9 and 8.8 percent to primary and secondary schooling, respectively, in Tanzania. These numbers are in line with Duflo's (2001) estimates of annual returns to education in Indonesia between 6.8 and 10 percent. Thus, I focus on the estimated value of the cost of increased firewood collection trip time with an assumed eight percent return to education. With an assumed annual return of eight percent to education, a one hour increase in firewood collection trip time translates into a 1.7 percent reduction in income per year. Average household expenditure for the 2010 survey round is approximately 2.5 million Tanzania shillings so this one hour increase in trip time corresponds to household expenditures being 42,400 Tanzanian shillings per year lower in 2010 ($30 in 2010 USD). Assuming that individuals work, on average, 30 years and using a five percent interest rate this loss in earnings has a net present value of $475 2010 USD.[13]

3.6 Conclusion

My findings show evidence that more restricted access to forests diminishes human capital formation. I measure forest access as the average time a household member spends making one firewood collection trip, which is positively correlated with distance to forest (Amacher et al., 2004; Kohlin and Amacher, 2005). For the analysis, I focus on children who are 7 to

[12] For all of these quantifications I rely on the firewood collection trip time coefficient estimate in column 3 of Table 3.10.

[13] A 2010 exchange rate of 1,409.27 Tanzanian shillings to the dollar is used for all conversions. An interest rate of 5.33 percent is used to calculate the net present value. This interest rate is the average real interest rate in Tanzania for the years 2010, 2011, and 2012. All figures come from the 2010 World Development Indicators (World Bank, 2010).

15 years old in 1991 and 26 to 34 years old in 2010. In the short-run, an additional hour required to collect firewood is associated with a child spending 25 minutes less in school per week. In the long-run, an additional hour required to collect firewood when the child is young is associated with the child completing 0.21 fewer grades 19 years later. Assuming an average return to education of eight percent a year, this long-term result translates into $475 2010 USD in lost earnings over 30 years, equal to roughly 1.7 percent of income.

This paper highlights one of the non-market benefits of forests, an important component of natural capital, to households in Tanzania. Deforestation and forest degradation that limit a household's access to forests may negatively impact human capital formation and, consequently, an individual's potential earnings when she is older. In 2010, Tanzania had approximately 15 million children ages 0 to 14 years living in rural areas and approximately 334,000 square kilometers of forest (World Bank, 2010). If, for the sake of illustration, all households in Tanzania suffered from the same loss in access to these forests then Tanzania would incur approximately 3 million cumulative years of lost education and $7 billion in lost lifetime earnings. Consequently, even though a 1.7 percent loss income is not necessarily economically significant at an individual level, the effect is large when aggregated to the country-level.

In the future, researchers should continue to examine how forest access, and access to other natural resources, such as water, affects human capital formation. Future studies can provide additional evidence on both the direction and magnitude of the relationship between environmental access and human capital formation. In addition, future studies should move beyond self-reported forest access data and instead rely on land use and land cover geospatial data. Firewood collection trip time is an endogenous variable to the extent that households have a choice as to where they collect firewood. In contrast, observed forest land cover data is exogenous and provides a clearer picture of how households respond to their natural environment.

Both the short- and long-term estimates presented in this paper are some of the first empirical analyses to link forest access and human capital formation in developing countries (see DeGraff, Levison and Dungumaro (2014) for another recent paper on the topic). Papers of this kind are timely given the current high levels of forest use and deforestation in sub-Saharan Africa and the current international focus on forest conservation policies. In the region, an estimated 90 percent of the population is reliant on firewood for cooking and human activity is one of the leading causes of deforestation (Agyei, 1998). An understanding of the link between forest access and human capital formation is helpful for understanding not only the full magnitude of program costs but also how best to design forest conservation programs to prevent increases in firewood collection trip times that may then lower school completion levels.

Table 3.1: Kagera Summary Statistics[a]

Households with children ages 7 to 15 in round 1.

	1991–1994			
	1	2	3	4
Household farms	0.93	0.91	0.94	0.93
	(0.26)	(0.28)	(0.23)	(0.25)
Household herds livestock	0.63	0.68	0.68	0.66
	(0.48)	(0.47)	(0.47)	(0.47)
Firewood is primary source of cooking fuel	0.96	0.96	0.97	0.98
	(0.19)	(0.19)	(0.18)	(0.15)
Household collects firewood	0.95	0.91	0.91	0.89
	(0.22)	(0.29)	(0.29)	(0.31)
Household is urban	0.20	0.19	0.19	0.19
	(0.40)	(0.39)	(0.40)	(0.39)
Observations	620	602	594	577

Standard deviation in parentheses.

[a]All statistics reported are proportions with a 1 denoting a "yes" response and 0 a "no" response.

Table 3.2: Education Summary Statistics
Sample of children ages 7 to 15 in round 1.

	1991–1994			
	1	2	3	4
Enrollment rate	0.59	0.64	0.64	0.61
Males	0.60	0.66	0.66	0.64
Females	0.59	0.62	0.62	0.57
Completed years of education	2.60	2.48	2.85	2.91
	(2.04)	(2.12)	(2.29)	(2.33)
Males	2.42	2.28	2.67	2.71
	(2.02)	(2.05)	(2.20)	(2.21)
Females	2.78	2.70	3.04	3.13
	(2.05)	(2.17)	(2.36)	(2.44)
Weekly hours spent in school	14.33	14.23	15.04	15.51
	(16.60)	(16.26)	(16.42)	(16.71)
Males	14.12	14.65	15.01	16.20
	(16.10)	(15.96)	(16.09)	(16.51)
Females	14.54	13.81	15.07	14.79
	(17.12)	(16.56)	(16.75)	(16.89)
Days attended school in last week	2.22	2.22	2.29	2.30
	(2.40)	(2.35)	(2.32)	(2.33)
Males	2.28	2.34	2.34	2.43
	(2.40)	(2.34)	(2.32)	(2.32)
Females	2.17	2.11	2.24	2.16
	(2.40)	(2.35)	(2.33)	(2.32)
Observations	1,442	1,396	1,378	1,341

Standard deviation in parentheses.

Table 3.3: Sample Summary Statistics, KHDS, Tanzania

Sample of children ages 7 to 15 in round 1.

	1991–1994				2004	2010
	1	2	3	4	5	6
Weekly hours spent in school	14.33	14.23	15.04	15.51		
	(16.60)	(16.26)	(16.42)	(16.71)		
Highest grade completed					6.90	7.26
					(2.12)	(2.52)
Age (years)	11.04	11.47	12.10	12.56	23.40	29.59
	(2.57)	(2.61)	(2.58)	(2.61)	(2.63)	(2.58)
Gender (1=female)	0.49	0.49	0.49	0.49	0.48	0.49
	(0.50)	(0.50)	(0.50)	(0.50)	(0.50)	(0.50)
Firewood collection trip time (hours)	1.76	1.63	1.54	1.45		
	(1.15)	(1.04)	(0.93)	(0.96)		
Mother: some primary education	0.12	0.12	0.12	0.12	0.27	0.33
	(0.32)	(0.32)	(0.32)	(0.32)	(0.45)	(0.47)
Mother: completed primary education or above	0.17	0.18	0.18	0.17	0.39	0.39
	(0.38)	(0.38)	(0.38)	(0.38)	(0.49)	(0.49)
Father: some primary education	0.11	0.11	0.11	0.11	0.42	0.43
	(0.32)	(0.32)	(0.32)	(0.32)	(0.49)	(0.49)
Father: completed primary education or above	0.26	0.25	0.25	0.25	0.48	0.49
	(0.44)	(0.43)	(0.44)	(0.43)	(0.50)	(0.50)
Gender of household head (1=female)	0.30	0.31	0.31	0.31	0.51	0.13
	(0.46)	(0.46)	(0.46)	(0.46)	(0.50)	(0.34)
Household size	7.84	7.53	7.48	7.39	4.80	4.57
	(3.47)	(3.39)	(3.56)	(3.56)	(2.76)	(2.18)
Annual household expenditure (TSh, log)[a]	12.75	11.87	11.82	11.79	14.37	14.53
	(0.71)	(0.77)	(0.71)	(0.75)	(0.65)	(0.65)

Continued on next page

Table 3.3 – continued from previous page

	1991–1994				2004	2010
	1	2	3	4	5	6
Moved within Kagera region					0.38	0.40
					(0.49)	(0.49)
Moved outside Kagera region					0.11	0.21
					(0.32)	(0.41)
Number of household deaths since last interview					1.20	0.38
					(1.19)	(0.65)
Number of very good years in household since last interview					0.17	0.09
					(0.29)	(0.21)
Number of very bad years in household since last interview					0.64	0.28
					(0.57)	(0.35)
Observations	1,442	1,396	1,378	1,341	881	1,008

Standard deviation in parentheses.

[a]Annual household expenditure measured in 1991 Tanzanian shillings (TSh) for survey rounds 1 through 4 and 2010 TSh for the 2004 and 2010 survey rounds.

Table 3.4: Two-Way Tabulation: Firewood Collection and Other Household Work
Sample of children ages 7 to 15 in round 1.

		Other household work	
		No	Yes
		Round 1 (N=1,442)	
Collects Firewood	No	220	547
		(15.32)	(38.09)
	Yes	51	618
		(3.55)	(43.04)
		Round 2 (N=1,396)	
Collects Firewood	No	143	489
		(11.24)	(38.44)
	Yes	37	603
		(2.91)	(47.41)
		Round 3 (N=1,378)	
Collects Firewood	No	102	454
		(8.55)	(38.06)
	Yes	36	601
		(3.02)	(50.38)
		Round 4 (N=1,341)	
Collects Firewood	No	96	369
		(8.82)	(33.88)
	Yes	36	588
		(3.31)	(53.99)

Relative frequencies in parentheses.

Note: Other household work includes: fetching water, preparing meals, cleaning the house, doing laundry, caring for other household members, other paid work, or helping neighbors with work.

Table 3.5: Short-Term School Attendance Effects: Hours

Dependent variable: Number of hours spent in school last week

	Poisson		Tobit (log)		Trimmed LAD (log)		OLS (log)	
	$\hat{\beta}$	$\frac{\partial \mathrm{E}[school\|\boldsymbol{x}]}{\partial x}$	$\hat{\beta}$	$\frac{\partial \mathrm{E}[school\|\boldsymbol{x}]}{\partial x}$	$\hat{\beta}$	$\frac{\partial \mathrm{E}[school\|\boldsymbol{x}]}{\partial x}$	$\hat{\beta}$	$\frac{\partial \mathrm{E}[school\|\boldsymbol{x}]}{\partial x}$
Age (years)	1.148***	13.85***	2.202***	23.08***	1.250***	18.45***	1.025***	15.13***
	(0.095)		(0.191)		(0.135)		(0.083)	
Age squared (years)	−0.044***	−0.52***	−0.087***	−0.91***	−0.050***	−0.73***	−0.040***	−0.59***
	(0.004)		(0.008)		(0.006)		(0.004)	
Gender (1=female)	−0.041	−0.50	−0.188**	−1.97**	−0.096**	−1.42**	−0.087*	−1.28*
	(0.037)		(0.086)		(0.048)		(0.045)	
Firewood collection trip time (hours)	−0.034*	−0.41*	−0.074*	−0.77*	−0.047**	−0.69**	−0.041**	−0.60**
	(0.017)		(0.042)		(0.024)		(0.020)	
Household size	0.014*	0.17*	0.045***	0.47***	0.016*	0.24*	0.021***	0.31***
	(0.007)		(0.016)		(0.009)		(0.008)	
Annual household expenditures (TSh, log)	0.123***	1.48***	0.274***	2.87***	0.199***	2.93***	0.151***	2.22***
	(0.036)		(0.087)		(0.042)		(0.044)	
Interviewed in February (1=yes)	0.245***	2.96***	0.371***	3.88***	0.279***	4.12***	0.210***	3.10***
	(0.076)		(0.130)		(0.097)		(0.077)	
Interviewed in June (1=yes)	−0.439**	−5.30**	−0.929**	−9.73**	−0.951***	−14.05***	−0.443**	−6.55**
	(0.190)		(0.426)		(0.232)		(0.206)	
Interviewed in July (1=yes)	−0.568**	−6.85**	−0.981**	−10.28**	−1.101***	−16.25***	−0.512***	−7.56***
	(0.224)		(0.393)		(0.175)		(0.177)	
Interviewed in August (1=yes)	0.220***	2.65***	0.593***	6.21***	0.391**	5.77**	0.257***	3.79***
	(0.062)		(0.139)		(0.159)		(0.069)	
Interviewed in December (1=yes)	−1.245***	−15.02***	−2.566***	−26.89***	−1.630***	−24.06***	−1.059***	−15.64***
	(0.337)		(0.546)		(0.147)		(0.174)	
Village fixed-effects	Yes		Yes		Yes		Yes	
Survey round fixed-effects	Yes		Yes		Yes		Yes	

Continued on next page

Table 3.5 – continued from previous page

	Poisson		Tobit (log)		Trimmed LAD (log)		OLS (log)	
	$\hat{\beta}$	$\frac{\partial \mathrm{E}[school\vert\boldsymbol{x}]}{\partial x}$	$\hat{\beta}$	$\frac{\partial \mathrm{E}[school\vert\boldsymbol{x}]}{\partial x}$	$\hat{\beta}$	$\frac{\partial \mathrm{E}[school\vert\boldsymbol{x}]}{\partial x}$	$\hat{\beta}$	$\frac{\partial \mathrm{E}[school\vert\boldsymbol{x}]}{\partial x}$
Observations	5,557		5,557		5,557		5,557	
Log-likelihood	-55,753.38		-7,819.91				-8,779.05	

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Standard errors in paranthesis. Cluster robust standard errors used for poisson, tobit and OLS models. Bootstrapped standard errors used for trimmed LAD model (500 replications).

Marginal effects are calculated at the mean for all three analyses. For the Poisson model the marginal effects are calculated as $\partial \mathrm{E}[school|\boldsymbol{x}]/\partial x_k = \beta_k \exp(\bar{\boldsymbol{x}}'\boldsymbol{\beta})$. For the tobit model the marginal effects are calculated as $\partial \mathrm{E}[school|\boldsymbol{x}]/\partial x_k = \Phi\big(\bar{\boldsymbol{x}}'\boldsymbol{\beta}/\sigma\big)\beta_k s\bar{chool}$ where $\Phi(\cdot)$ is the standard normal cumulative distribution function. And, for the trimmed LAD and OLS models the marginal effects are calculated as $\partial \mathrm{E}[school|\boldsymbol{x}]/\partial x_k = \beta_k s\bar{chool}$.

Additional controls include: Mother has some primary education, mother completed primary education, father has some primary education, father completed primary education, and female is head of household.

Table 3.6: Short-Term School Attendance Effects: Sample Size Robustness Checks
Dependent variable: Number of hours spent in school last week
GLM-LL Poisson estimation

	Original model	Households that collect	Children that collect	Zero hours
Age (years)	1.148***	1.152***	0.918***	1.148***
	(0.095)	(0.102)	(0.091)	(0.095)
Age squared (years)	−0.044***	−0.043***	−0.034***	−0.043***
	(0.004)	(0.004)	(0.004)	(0.004)
Gender (1=female)	−0.041	−0.051	−0.029	−0.042
	(0.037)	(0.040)	(0.045)	(0.037)
Firewood collection trip time (hours)	−0.034*	−0.030*	−0.044*	−0.024*
	(0.017)	(0.018)	(0.026)	(0.014)
Household size	0.014*	0.010	−0.006	0.014*
	(0.007)	(0.007)	(0.008)	(0.008)
Annual household expenditures (TSh, log)	0.123***	0.126***	0.128***	0.119***
	(0.036)	(0.034)	(0.042)	(0.036)
Interviewed in February (1=yes)	0.245***	0.266***	0.330***	0.244***
	(0.076)	(0.079)	(0.068)	(0.077)
Interviewed in June (1=yes)	−0.439**	−0.452**	−0.440**	−0.438**
	(0.190)	(0.191)	(0.204)	(0.191)
Interviewed in July (1=yes)	−0.568**	−0.640***	−0.565**	−0.567**
	(0.224)	(0.231)	(0.268)	(0.225)
Interviewed in August (1=yes)	0.220***	0.206***	0.211***	0.220***
	(0.062)	(0.060)	(0.066)	(0.062)
Interviewed in December (1=yes)	−1.245***	−1.290***	−1.152***	−1.247***
	(0.337)	(0.354)	(0.350)	(0.337)
Village fixed-effects	Yes	Yes	Yes	Yes
Survey round fixed-effects	Yes	Yes	Yes	Yes
Observations	5,557	5,098	2,568	5,557
Log-likelihood	-55,753.38	-50,963.79	-24,047.31	-55,769.08

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Cluster robust standard errors in parantheses.

Note: All coefficient estimates rely on a Poisson estimation approach. Additional controls include: Mother has some primary education, mother completed primary education, father has some primary education, father completed primary education, and female is head of household.

Table 3.7: Short-Term School Attendance Effects: Source of Variation Robustness Checks
Dependent variable: Number of hours spent in school last week
GLM-LL Poisson estimation

	Intra-village variation	Inter-village variation	Temporal variation
Age (years)	1.148***	1.726**	1.335***
	(0.095)	(0.722)	(0.112)
Age squared (years)	−0.044***	−0.064**	−0.049***
	(0.004)	(0.030)	(0.004)
Gender (1=female)	−0.041	0.133	−0.038
	(0.037)	(0.399)	(0.051)
Firewood collection trip time (hours)	−0.034*	−0.001	−0.022
	(0.017)	(0.061)	(0.016)
Household size	0.014*	−0.026	0.051***
	(0.007)	(0.033)	(0.010)
Annual household expenditures (TSh, log)	0.123***	0.271***	−0.004
	(0.036)	(0.096)	(0.026)
Interviewed in February (1=yes)	0.245***	0.249***	0.226***
	(0.076)	(0.080)	(0.035)
Interviewed in June (1=yes)	−0.439**	−0.305	−0.416***
	(0.190)	(0.187)	(0.060)
Interviewed in July (1=yes)	−0.568**	−0.541**	−0.555***
	(0.224)	(0.240)	(0.074)
Interviewed in August (1=yes)	0.220***	−0.036	0.185***
	(0.062)	(0.085)	(0.036)
Interviewed in December (1=yes)	−1.245***	−1.329***	−1.251***
	(0.337)	(0.363)	(0.099)
Child fixed-effects	No	No	Yes
Village fixed-effects	Yes	No	No
Survey round fixed-effects	Yes	Yes	No
Village averages	No	Yes	No
Observations	5,557	5,557	5,557
Log-likelihood	-55,753.38	-21,399.32	-34,289.35

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Cluster robust standard errors in parantheses.

Note: All coefficient estimates rely on a Poisson estimation approach. Additional controls include: Mother has some primary education, mother completed primary education, father has some primary education, father completed primary education, and female is head of household..

Table 3.8: Cross-Correlation: Firewood Collection Trip Time

	Round 1	Round 2	Round 3	Round 4
Round 1	1.00			
Round 2	0.20	1.00		
Round 3	0.19	0.18	1.00	
Round 4	0.18	0.27	0.15	1.00

Table 3.9: Attrition Probit Model

Dependent variable: Individual interviewed in survey round 2004/2010 (1=yes)

	2004 Round		2010 Round	
	Unrestricted	Restricted	Unrestricted	Restricted
Age in round 1 (years)	−0.016	0.026	0.363**	
	(0.158)	(0.155)	(0.162)	
Age squared in round 1	0.001	−0.001	−0.015**	
	(0.007)	(0.007)	(0.007)	
Gender (1=female)	−0.046	−0.016	0.067	0.068
	(0.081)	(0.080)	(0.084)	(0.084)
Weekly hours spent in school in round 1	−0.001	−0.002	0.002	0.004
	(0.003)	(0.003)	(0.003)	(0.003)
Firewood collection trip time in round 1 (hours)	−0.008	−0.041	−0.006	−0.001
	(0.035)	(0.035)	(0.037)	(0.037)
Gender of household head in round 1 (1=female)	0.071	0.005	0.170	0.166
	(0.101)	(0.099)	(0.110)	(0.110)
Age of household head in round 1 (years)	−0.001	−0.002	−0.001	−0.002
	(0.003)	(0.003)	(0.003)	(0.003)
Household size in round 1	0.062***		0.010	0.013
	(0.020)		(0.016)	(0.017)
Annual household expenditures in round 1 (TSh, log)	−0.226**		0.010	−0.006
	(0.089)		(0.088)	(0.087)
Mother: some primary education in round 1	−0.147	−0.248**	−0.196	−0.099
	(0.129)	(0.125)	(0.136)	(0.131)
Mother: completed primary education or above in round 1	0.044	−0.005	−0.229*	
	(0.120)	(0.116)	(0.119)	
Father: some primary education in round 1	0.061	0.087	−0.182	−0.214
	(0.145)	(0.142)	(0.139)	(0.140)

Continued on next page

Table 3.9 – continued from previous page

	2004 Round		2010 Round	
	Unrestricted	Restricted	Unrestricted	Restricted
Father: completed primary education or above in round 1	−0.179	−0.136	−0.006	−0.093
	(0.111)	(0.109)	(0.121)	(0.116)
Number of household deaths since last interview	−0.013	−0.001	−0.056	−0.074
	(0.040)	(0.036)	(0.063)	(0.062)
Average village attrition rate	−2.107***		−1.256**	
	(0.579)		(0.565)	
Urban village (1=yes)	0.298**		−0.105	−0.261**
	(0.134)		(0.124)	(0.111)
Number of primary schools in village in round 1	0.043	0.091	−0.079	−0.076
	(0.080)	(0.077)	(0.080)	(0.080)
Number of very good years in household since last interview	0.247	0.344*	0.320	0.375
	(0.177)	(0.184)	(0.256)	(0.250)
Number of very bad years in household since last interview	0.013	0.003	−0.114	−0.105
	(0.069)	(0.069)	(0.123)	(0.122)
Observations	1,334	1,346	1,319	1,319
Pseudo-R^2	0.035	0.013	0.032	0.018
Wald test, p-value	0.000		0.094	
Baseline predicted probability	0.818	0.816	0.840	0.840

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Cluster robust standard errors in parentheses.

Table 3.10: Long-Term School Completion Effects

Dependent variable: Years of education completed

	2004 Round		2010 Round		Individuals in both 2004 and 2010 (OLS)	
	OLS	WLS	OLS	WLS	2004	2010
Age in round 1 (years)	−0.152	−0.197	−0.048	−0.053	−0.092	−0.232
	(0.298)	(0.295)	(0.248)	(0.250)	(0.307)	(0.305)
Age squared in round 1	0.007	0.009	0.002	0.002	0.005	0.010
	(0.014)	(0.014)	(0.011)	(0.011)	(0.014)	(0.014)
Gender (1=female)	0.041	0.034	−0.335**	−0.342**	0.009	−0.215
	(0.113)	(0.115)	(0.153)	(0.152)	(0.123)	(0.144)
Firewood collection trip time in round 1	−0.157	−0.158	−0.212**	−0.215***	−0.162*	−0.184**
	(0.095)	(0.095)	(0.079)	(0.080)	(0.084)	(0.071)
Mother: some primary education in round 1	−0.760***	−0.733***	−0.563**	−0.573***	−0.656***	−0.529***
	(0.199)	(0.203)	(0.216)	(0.214)	(0.200)	(0.192)
Mother: completed primary education or above in round 1	−0.161	−0.175	0.098	0.104	0.033	0.142
	(0.203)	(0.203)	(0.226)	(0.229)	(0.205)	(0.204)
Father: some primary education in round 1	−0.437*	−0.452*	−0.614*	−0.596*	−0.375	−0.234
	(0.235)	(0.234)	(0.317)	(0.318)	(0.265)	(0.245)
Father: completed primary education or above in round 1	0.418**	0.371*	0.133	0.150	0.605***	0.505**
	(0.198)	(0.207)	(0.209)	(0.209)	(0.203)	(0.233)
Gender of household head in round 1 (1=female)	0.192	0.227	0.168	0.161	0.205	0.113
	(0.197)	(0.203)	(0.194)	(0.196)	(0.186)	(0.192)
Household size in round 1	−0.065*	−0.068*	−0.059	−0.057	−0.055	−0.078*
	(0.037)	(0.037)	(0.038)	(0.038)	(0.039)	(0.042)
Annual household expenditures in round 1 (TSh, log)	0.996***	1.019***	0.791***	0.787***	1.107***	0.969***
	(0.187)	(0.189)	(0.186)	(0.186)	(0.168)	(0.169)
Moved within Kagera region	−0.043	−0.040	0.204	0.218	−0.007	0.252

Continued on next page

Table 3.10 – continued from previous page

	2004 Round		2010 Round		Individuals in both 2004 and 2010 (OLS)	
	OLS	WLS	OLS	WLS	2004	2010
	(0.149)	(0.152)	(0.198)	(0.199)	(0.162)	(0.172)
Moved outside Kagera region	1.218***	1.243***	2.201***	2.218***	1.200***	1.933***
	(0.171)	(0.173)	(0.283)	(0.283)	(0.261)	(0.306)
Village fixed-effects	Yes	Yes	Yes	Yes	Yes	Yes
Observations	949	949	1011	1011	820	820
Log-likelihood	-1,959.76	-1,960.712	-2,203.451	-2,206.361	-1,647.967	-1,654.863
Inverse probability weight (mean)		1.00		1.00		

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Cluster robust standard errors in parentheses.

Table 3.11: Quantifying the Effects

	Return to education (annual)									
	1%	2%	3%	4%	5%	6%	7%	8%	9%	10%
Lost school-days	42.41	42.41	42.41	42.41	42.41	42.41	42.41	42.41	42.41	42.41
Lower bound	10.52	10.52	10.52	10.52	10.52	10.52	10.52	10.52	10.52	10.52
Upper bound	74.30	74.30	74.30	74.30	74.30	74.30	74.30	74.30	74.30	74.30
Lost income (percent)	0.21	0.42	0.64	0.85	1.06	1.27	1.48	1.70	1.91	2.12
Lower bound	0.05	0.11	0.16	0.21	0.26	0.32	0.37	0.42	0.47	0.53
Upper bound	0.37	0.74	1.11	1.49	1.86	2.23	2.60	2.97	3.34	3.72
Lost income, annual (100's TSh)	53.01	106.03	159.04	212.05	265.07	318.08	371.09	424.10	477.12	530.13
Lower bound	13.15	26.30	39.44	52.59	65.74	78.89	92.04	105.18	118.33	131.48
Upper bound	92.88	185.76	278.63	371.51	464.39	557.27	650.15	743.03	835.90	928.78
Lost income, annual (USD)	3.76	7.52	11.29	15.05	18.81	22.57	26.33	30.09	33.86	37.62
Lower bound	0.93	1.87	2.80	3.73	4.66	5.60	6.53	7.46	8.40	9.33
Upper bound	6.59	13.18	19.77	26.36	32.95	39.54	46.13	52.72	59.31	65.91
Net present value (30 yrs, USD)	59.48	118.95	178.43	237.90	297.38	356.85	416.33	475.81	535.28	594.76
Upper bound	104.20	208.40	312.60	416.80	521.00	625.20	729.41	833.61	937.81	1,042.01
Lower bound	14.75	29.50	44.25	59.00	73.75	88.51	103.26	118.01	132.76	147.51

Note: Calculations are done assuming an income of 2.5 million shillings. The lower bound estimates use a β estimate of -0.053 and the upper bound estimates use a β estimate of -0.372, i.e. the 95% confidence interval from the OLS estimates in column 3 of Table 3.10.

Figure 3.1: Kagera Region, Tanzania

Chapter 4

Food Market Constraints and Households' Food and Nutrient Consumption: Evidence from Tanzania*

4.1 Introduction

Households have access to foods both through the local marketplace and through home production. If food markets work perfectly, then households are able to produce crops to maximize agricultural profits and then purchase food in the local marketplace to maximize household utility. An increase in the price of cash crops indirectly affects household food consumption through a positive effect on household income; as the price of cash crops rises, farmers will produce more cash crops and less food crops. But, because food markets work perfectly, home produced food and purchased food are perfect substitutes and households can supplement any drop in home-consumed food with purchased food. If food markets do not work perfectly, however, then an increase in the price of cash crops may negatively impact

*This paper was written in collaboration with Helen Markelova, Ph.D. Candidate, Department of Applied Economics, University of Minnesota household welfare. With imperfect food markets, home-produced food and purchased food are no longer perfect substitutes. If an increase in the price of cash crops causes farmers to produce more cash crops and less food crops then household food consumption may decline because they can no longer supplement a decrease in home-produced food consumption with purchased food consumption. In this scenario, an increase in the price of cash crops affects food consumption both indirectly, through an increase in household agricultural profits, and directly, through a decrease in home-produced food. In this paper, we focus on empirically testing for the presence of food market constraints and non-separability in household food demand.

We test for the presence of constraints in the market for food using a household dataset from Tanzania and applying it in the context of the traditional agricultural household model. Specifically, we test whether a household's demand for food and a household's demand for nutrients are affected by the price of cash crops.[1] We construct regional-level prices for five food groups (staples, pulses, fruits/vegetables, animal products, and meal complements), cash crops, and non-food market goods to be able to estimate a food demand system for the five food groups, measured in kilocalories, and for eight different nutrients (protein, iron, zinc, vitamin A, riboflavin, folate vitamin B_{12}, and vitamin C). The latter is an innovative approach of this paper–examining the demand for food using not just the available weight measures, but decomposing the food groups into measures of key nutrients. Earlier papers focused on household energy demand (Subramanian and Deaton, 1996), but few papers have extended the analysis to include household nutrient consumption (see Abdulai and Aubert (2004) and Ecker and Qaim (2011) from some exceptions). Ultimately, we provide evidence for the presence of food market constraints in Tanzania and their effects on a household's food consumption.

This paper makes two main contributions to the literature. It is one of the first studies to empirically test for the inter-dependencies between household food consumption and household agricultural production by estimating households' demand for food. Previous

[1]By "cash crops" we refer to non-food agricultural crops that are traditionally only sold in the market place (not consumed). These markets may include local market (sales at the farm gate), national markets (for example, supermarkets), and export-oriented production. All three types of markets can be accessed by farming households in developing countries (Barrett, 2008).

studies that have tested for separability have focused on households' agricultural production (Benjamin, 1992; Jacoby, 1993; Skoufias, 1994; Bhattacharyya and Kumbhakar, 1997; Grimard, 1997; Abdulai and Regmi, 2000; Le, 2010). With the exception of Tekgüç (2012), little effort has been made to testing for separability using household food demand. Thus this paper is important not only for corroborating the findings of previous papers (the majority of which find evidence of nonseparability) but also for providing an alternative test that can be used to test for nonseparability. In contrast to the previous tests that rely on extensive household agricultural production data, this test requires only limited agricultural data (input and output prices) and data on household food consumption.

Second, this paper provides evidence on the effects of nonseparability on household food demand. Previous literature widely acknowledges that nonseparability causes household food demand to depend directly on agricultural prices (Pitt and Rosenzweig, 1985; de Janvry, Fafchamps and Sadoulet, 1991; Taylor and Adelman, 2003; Duflo and Udry, 2004). But, previous food demand estimates in developing countries do not include any agricultural prices as explanatory variables (Subramanian and Deaton, 1996; Abdulai and Aubert, 2004; Ecker and Qaim, 2011). We show that in the presence of food market constraints, agricultural prices significantly affect household food demand and consequently should be included as explanatory variables in future food demand analyses.

The results of this paper are relevant to two ongoing policy discussions. First, understanding the effects of food market constraints on households' food consumption is useful for understanding how households obtain key nutrients. Micronutrient malnutrition is now being labeled as "the hidden hunger," with consequences lasting for decades for children as it leads to delayed cognitive development, increased susceptibility to diseases, and overall poor physical development (Asare-Marfo et al., 2013). These health issues caused by micronutrient deficiencies affect individual welfare and national economic growth – improved nutrition has been linked to higher levels of human capital accumulation and increased productivity on the labor market (Deolalikar, 1988; Strauss and Thomas, 1998; Welch and Graham, 1999; Alderman, Hoddinott and Kinsey, 2006). This paper is one of the few empirical studies to emphasize that not only does health affect agricultural productivity but agricultural production also potentially affects health (Grossman, 1972).

Second, the traditional emphasis of development programs remains on agricultural intensification and commercializations (Barrett, 2008; Von Braun and Kennedy, 1994). Beginning in the 1950s with the onset of the Green Revolution, development programs have predominately involved the introduction and promotion of new agricultural technologies, such as high-yield crops, irrigation systems, fertilizers and pesticides. The primary focus of these technologies and programs was on increasing both the production and sale of staples (i.e. maize and rice) and cash crops (i.e. cotton and tobacco) in the market (Borlaug, 2007). More recently, Feed the Future, the U.S. Government's Global Food Security and Hunger Initiative, focuses primarily on increasing maize and rice production in Tanzania for the market (Bernard, Taffesse and Gabre-Madhin, 2008; U.S. Government's Global Hunger and Food Security Initiative, 2014). If food markets are constrained, however, then agricultural intensification programs that generate incentives for households to produce more crops for the market, which improve households' income but do not address agricultural production for home consumption, could lower household consumption of high-nutrient low-profit agricultural crops and increase household micronutrient deficiency levels. To date, none of these programs have considered the potential adverse effects that market-oriented agricultural production may have on household health and nutrition. As shown in this paper, in the presence of food market constraints, household food consumption, and subsequent nutrient consumption, is affected by a household's agricultural production decisions. Thus, agriculture plays a dual role in farming households in Tanzania; agricultural production not only contributes to households' profits but also their nutritional status (Welch and Graham, 1999).

The paper proceeds as follows. The next section reviews the literature and further highlights the contributions made by this paper. In Section 4.3 we present the theoretical set-up for the scenarios of fully-functioning and constrained markets for food. The empirical approach is discussed in the next section, followed by sections that describe the data and the methodology used for the analyses. Then, in Section 4.6 we present the results from estimating the demand systems and interpret both the statistical and economic significance of the coefficient estimates. We then discuss in more detail the agricultural policy implications of our results and provide some concluding remarks in Section 4.7.

4.2 Literature Review

The conceptual foundations of this paper are rooted in the traditional agricultural household model that describes how agricultural households in developing countries make production and consumption decisions. The main underlying principle in this model is that, under certain assumptions, a household's production and consumption decisions are made independently (see Bardhan and Udry (1999) for more details on this model). In the presence of separability agricultural output and input prices will affect a household's consumption decisions only via their effect on farm prices (Pitt and Rosenzweig, 1985; de Janvry, Fafchamps and Sadoulet, 1991; Taylor and Adelman, 2003; Duflo and Udry, 2004). The separability hypothesis, however has been widely tested (and rejected) by estimating a household's agricultural production behavior and testing the null hypothesis that household characteristics do not affect farm profits (Benjamin, 1992; Jacoby, 1993; Skoufias, 1994; Bhattacharyya and Kumbhakar, 1997; Grimard, 1997; Abdulai and Regmi, 2000; Le, 2010). With nonseparability agricultural output and input prices have a direct effect on households' consumption decisions. Despite the evidence for nonseparability, almost no food demand models include agricultural prices as explanatory variables (for example, see Subramanian and Deaton (1996), Abdulai and Aubert (2004), and Ecker and Qaim (2011)).

In this paper, we investigate whether cash crop prices affect consumption decisions and, subsequently, reveal market constraints in purchasing food (and thus nutrients) on the local market. To date, only Tekgüç (2012) has tested for nonseparability by estimating a household food demand model. In his paper, Tekgüç (2012) estimates an almost ideal demand system for household food demand for eleven food groups.[2] He then tests for separability by including the home-produced food budget share as an additional explanatory variable; if home-produced food is significant then there is evidence of nonseparability. In this paper, we estimate household food demand using both household consumption and agricultural production data. Unlike Tekgüç (2012), we use a direct measure of agricultural output prices and do not proxy for agricultural data using self-produced food, an endogenously chosen

[2]The eleven food groups are: bread, cereals, meat and meat products, vegetable oils, vegetables, fruits, dairy products and eggs, sugar, confectionary and jams, tea and coffee, non-alcoholic beverages, and other food products.

variable. In addition, instead of estimating food budget shares we estimate household demand for energy and key micronutrients and thus provide additional information on the implications of nonseparability for household health and nutrition.

The separability property holds only in the context of complete markets (Bardhan and Udry, 1999). However, due to the institutional, economic, legal, and political environments in many developing countries, households face limitations or even failures in local markets. An earlier study by de Janvry, Fafchamps and Sadoulet (1991) provided several theoretical arguments for the existence of pervasive food market imperfections in developing countries as a result of poor infrastructure, high transportation costs, and price volatility; de Janvry, Fafchamps and Sadoulet's (1991) theoretical results have been elaborated on and empirically confirmed by findings in more recent papers as well (Fafchamps, 1992; Omamo, 1998; Key, Sadoulet and de Janvry, 2000; Barrett, 2008). In particular, through a series of simulations de Janvry, Fafchamps and Sadoulet (1991) show that in the presence of food market constraints an increase in the price of cash crops reduces household food consumption; a 10 percent increase in the price of cash crops reduces household food consumption by 0.8 percent. In this paper, following de Janvry, Fafchamps and Sadoulet (1991), we investigate the effects of food market constraints on household food consumption but, in contrast to de Janvry, Fafchamps and Sadoulet (1991), we rely on household survey data for the analysis.

In Tanzania, there are several reasons to believe that food markets are constrained. First, households may face food market constraints because of poor infrastructure or incomplete market access; poor infrastructure not only makes it more costly to transport goods but also reduces the amount of information that a farmer has on the market for a specific crop, making transactions more risky. In Tanzania, there is substantial subjective evidence supporting the existence of poor infrastructure; Tanzania ranked 104 out of 153 ranked countries in 2012 for the "logistics performance index" which ranks the quality of its trade and transportation. Additionally, Tanzania has only 9 km of road per 100 km^2 of land relative to the United States' 66 km of road per 100 km^2 (Tanzania ranked 80 out of 92 ranked countries in 2011) (World Bank, 2010).

Second, households may also face constraints in their local food markets because of their limited ability to purchase food storage technology, such as refrigerators, combined with limited household access to food markets. Without refrigerators, households cannot store perishable food items, including meats and leftovers, and will need to make daily trips to the market in order to consume these food items (The Economist, 2014). According to Tanzania's most recent National Panel Survey (2010-2011) only three percent of farming households own a refrigerator and only 25 percent have access to a daily market within their village; households without a market in their village must travel, on average, 22 km to reach the nearest daily market. Either one of these two situations could create a situation where agricultural production decisions and household consumption decisions are made jointly, leading to a breakdown in the separability property (Bardhan and Udry, 1999).

Consequently, this paper follows in the footsteps of Pitt and Rosenzweig (1985) by not assuming separability between production and consumption decisions and then empirically tests the validity of that assumption. Our main proposition is that in developing countries, one of the missing or imperfect markets is the market for food, which ultimately affects a household's ability to purchase nutritious food.

4.3 Household Food Demand

Our conceptual framework is based on the classic agricultural household model used in development economics (see Bardhan and Udry, 1999). Our model, however, emphasizes household demand for food in the presence of food market imperfections, similar to the model laid out in de Janvry, Fafchamps and Sadoulet (1991). To begin, we consider the simple case of household demand for food when all markets are perfectly functioning. Households receive utility, u, from food consumption, denoted by $z \geq 0$, and an aggregate non-food good $c \geq 0$. We assume that the utility function is concave in both of its arguments. Households purchase both z and c in the local market for prices $p_z > 0$ and $p_c > 0$, respectively. Households choose to produce two types of crops: a food crop, $F(\alpha_z L_z, \beta_z E_z, \delta_z f_z)$, and a cash crop, $F(\alpha_a L_a, \beta_a E_a, \delta_a f_a)$, both of which are functions of labor, L_j, land, E_j, and fertilizer, f_j, for $j = z, a$. Households can choose to hire labor L_j^h or work on the farm themselves L_j^f so that $L_j^h + L_j^f = L_j$ for $j = z, a$. Hired labor is paid at the rate w and households have a fixed amount of labor given by $\bar{L} = L_z^f + L_a^f + L^m$ where L^m is household marketed labor.

The terms α_j, β_j, δ_j for $j = z, a$ are scalars that reflect differences in input efficiency across food and cash crops.[3] Aside from these two efficiency scalars, the agricultural production process governing both food and cash crop production is the same and is assumed to be strictly concave in labor, land, and fertilizer.[4]

For simplicity, we assume that hired and home labor are perfectly substitutable and that labor markets are fully functioning so that the households do not face any labor constraints in agricultural production decisions. We assume that there are no land rental markets and that households also have a fixed amount of agricultural land, $\bar{E}$, that they can choose to plant with either food crops (for example, orange flesh sweet potatoes), or cash crops (such as tobacco).[5] Finally, cash crop agricultural production can be sold in the market for the price $p_a > 0$, food crops can be sold for $p_z > 0$, and fertilizer is bought at the price $p_f > 0$. Thus, the full household problem is given by:

$$\max_{z,c,L_z^h,L_z^f,L_a^h,L_a^f,E_z,E_a,f,L^m} \quad u(z,c)$$

$$\text{subject to:}$$

$$p_z z + p_c c = p_z F(\alpha_z L_z, \beta_z E_z, \delta_z f_z) + p_a F(\alpha_a L_a, \beta_a E_a, \delta_a f_a) - \sum_{j=z,a} \left(w L_j^h + p_f f_j \right) + w L^m + y$$

$$\bar{E} = E_z + E_a$$

$$L_j = L_j^h + L_j^f \text{ for } j = z, a$$

$$\bar{L} = L_z^f + L_a^f + L^m$$

$$z, c, L_z^h, L_z^f, L_a^h, L_a^f, E_z, E_a \geq 0$$

where y is exogenously-determined non-labor household income. With perfectly functioning labor markets, if the household buys labor it does not sell labor (i.e. $\sum_{j=z,a} L_j^h > 0$ then $L^m = 0$) and if the household sells labor it will not buy labor (i.e. $L^m > 0$ then $\sum_{j=z,a} L_j^h = 0$). As long as $L_j^f > 0$ for $j = z, a$ the theoretical implications of the effects of imperfect

[3] The distinction between food and cash crop production is important because it allows households to grow crops that they can't consume but that have high market value (e.g. tobacco or cotton).

[4] In addition, we assume that $\partial F / \partial L_j(0, E_j, \cdot) = \infty$ and $\partial F / \partial E_j(L_j, 0, \cdot) = \infty$ for $j = z, a$ to avoid the case of a corner solution where the household will choose to produce only one of the two goods.

[5] This assumption is in line with observed land use in Tanzania. Only 1 percent of farming households (30 households) report renting out land and only 6 percent of households (122) report renting in land.

food markets are unchanged. Therefore, we consider the case of the household selling labor, $L^m > 0$, and not buying labor, $\sum_{j=z,a} L_j^h = 0$ and let $L_j = L_j^f$ for $j = z, a$ so that $\bar{L} = L_z + L_a + L^m$. The alternative case of the household buying labor is derived in Appendix C.

For the case of the household selling labor, the associated Lagrangian is:

$$\mathcal{L} = u(z,c) - \lambda\Big(p_z z + p_c c - p_z F(\alpha_z L_z, \beta_z E_z, \delta_z f_z) - p_a F(\alpha_a L_a, \beta_a E_a, \delta_a f_a) + \sum_{j=z,a} p_f f_j$$
$$- wL^m - y\Big) - \phi(\bar{E} - E_z - E_a) - \eta(\bar{L} - L_z - L_a - L^m)$$

Assuming an interior solution for all choice variables, the first order conditions and the constraints are a system of twelve equations (including the budget constraint, the land constraint, and the home labor constraint) and twelve variables (including the three Lagrange multipliers λ, ϕ, and η), and so they can be used to solve for optimal household consumption and production levels. Specifically, the first order conditions determining optimal household consumption and agricultural production levels are given by:

$$\frac{\partial u}{\partial i} = \lambda p_i \text{ for } i = z, c \tag{4.1}$$
$$p_j \alpha_j \frac{\partial F}{\partial L_j} = -\frac{\eta}{\lambda} \text{ for } j = z, a \tag{4.2}$$
$$p_j \beta_j \frac{\partial F}{\partial E_j} = -\frac{\phi}{\lambda} \text{ for } j = z, a \tag{4.3}$$
$$p_j \delta_j \frac{\partial F}{\partial f_j} = p_f \text{ for } j = z, a. \tag{4.4}$$
$$w = -\frac{\eta}{\lambda} \tag{4.5}$$

From equation (4.1) households' consumption decisions are determined by the prices of the consumed goods p_z and p_c up until their marginal utilities are equal:

$$\frac{1}{p_c}\frac{\partial u}{\partial c} = \frac{1}{p_z}\frac{\partial u}{\partial z}. \tag{4.6}$$

And first order conditions (4.2), (4.3), and (4.4) imply that land and labor are allocated between household-produced food and cash crop production up until the point where the value of their marginal products are equal:

$$p_z\alpha_z\frac{\partial F}{\partial L_z} = p_a\alpha_a\frac{\partial F}{\partial L_a},\ p_z\beta_z\frac{\partial F}{\partial E_z} = p_a\beta_a\frac{\partial F}{\partial E_a},\ \text{and}\ p_z\delta_z\frac{\partial F}{\partial f_z} = p_a\delta_a\frac{\partial F}{\partial f_a}. \tag{4.7}$$

As long as p_z, p_c, and p_a are exogenously determined prices, these first order conditions imply that a household's agricultural production and consumption decisions are not linked and that the decisions can be made sequentially. Specifically, households first choose agricultural land and labor inputs to satisfy (4.7) and then choose food and non-market good consumption levels to maximize utility subject to the income constraint $y + \Pi^* = p_zF(\alpha_zL_z^*, \beta_zE_z^*, \delta_zf_z^*) + p_aF(\alpha_aL_a^*, \beta_aE_a^*, \delta_af_a^*) - p_f(f_z^* + f_a^*) + wL^{m*}$, an exogenous variable in the household consumption problem. Total household food consumption is given by:

$$z^* = z(p_z, p_c, \Pi^* + y).$$

Most importantly, in this separable household model the price of cash crops has no direct effect on household food consumption decisions. The sequential nature of this separable model implies that a household's food consumption decisions is equivalent to solving the simplified problem:

$$\max_{z,c} u(z,c)$$
$$\text{subject to:}$$
$$p_zz + p_cc = \Pi^*(p_z, p_a, p_f, w, \bar{E}) + y.$$

Thus, an increase in the price of cash crops will affect household food consumption decisions only through it's affect on Π^* but will have no direct effect on food consumption.

4.3.1 Food Market Failures

Following de Janvry, Fafchamps and Sadoulet (1991), we now consider household food demand with a food market failure. We assume that there is no food market but that there are complete markets for cash crops and the non-food good. Although a total failure in the food market is unlikely to be present, this exaggerated model best highlights a household's adjustments in food demand that would result even from a less severe constraint in the food market. With a food market failure food is no longer traded in the market and the household must meet all of its food needs through the production of food crops. Again, we assume that the household sells labor and does not hire labor – the alternative case of the household hiring labor and marketing no labor is derived in Appendix C. The adapted household model is:

$$\max_{z,c,L_z,L_a,E_z,E_a,f,L^m} u(z,c)$$

subject to:

$$p_c c = p_a F(\alpha_a L_a, \beta_a E_a, \delta_a f_a) - \sum_{j=z,a} p_f f_j + wL^m + y$$

$$\bar{E} = E_z + E_a$$

$$L_j = L_j^f \text{ for } j = z, a$$

$$\bar{L} = L_z + L_a + L^m$$

$$z \leq F(\alpha_z L_z, \beta_z E_z, \delta_z f_z)$$

$$z, c, L_z, L_a, E_z, E_a \geq 0.$$

The adapted Lagrangian for this constrained household problem is:

$$\begin{aligned}\mathcal{L} = u(z,c) - \lambda\Big(p_c c - p_a F(\alpha_a L_a, \beta_a E_a, \delta_a f_a) + \sum_{j=z,a} p_f f_j - wL^m - y\Big) \\ - \phi(\bar{E} - E_z - E_a) - \eta(\bar{L} - L_z - L_a - L^m) - \mu\Big(z - F(\alpha_z L_z, \beta_z E_z, \delta_z f_z)\Big).\end{aligned}$$

At an interior solution, the first order conditions governing food consumption and agricultural production decisions are now:

$$\frac{\partial u}{\partial z} = \mu \tag{4.8}$$

$$\frac{\partial u}{\partial c} = \lambda p_c \tag{4.9}$$

$$\alpha_z \frac{\partial F}{\partial L_z} = -\frac{\eta}{\mu} \tag{4.10}$$

$$p_a \alpha_a \frac{\partial F}{\partial L_a} = -\frac{\eta}{\lambda} \tag{4.11}$$

$$\beta_z \frac{\partial F}{\partial E_z} = -\frac{\phi}{\mu} \tag{4.12}$$

$$p_a \beta_a \frac{\partial F}{\partial E_a} = -\frac{\phi}{\lambda} \tag{4.13}$$

$$\delta_z \frac{\partial F}{\partial f_z} = \frac{\lambda}{\mu} p_f \tag{4.14}$$

$$p_a \delta_a \frac{\partial F}{\partial f_a} = p_f \tag{4.15}$$

$$w = -\frac{\eta}{\lambda} \tag{4.16}$$

A household's decision of how much food crops to produce is now directly affected by the household's food consumption preferences through first order conditions (4.8), (4.9), (4.10), and (4.11):

$$\frac{\partial u}{\partial z} = \frac{p_a}{p_c} \frac{\partial u}{\partial c} \frac{\beta_a \partial F/\partial E_a}{\beta_z \partial F/\partial E_z}. \tag{4.17}$$

Consequently, in this constrained problem the price of cash crops now affects household food consumption both through it's direct effect on z and it's indirect effect on household agricultural profits. Total household food consumption is given by:

$$z^* = z\big(p_c, p_a, p_f, \Pi + y\big)$$

where $\Pi(p_z, p_a, p_f, w, \bar{E}) = p_a F(\alpha_a L_a, \beta_a E_a, \delta_a f_a) - p_f(f_z + f_z) + wL^m$ is still household agricultural profits.

If the food market constraint does not cause a complete food market failure then some trade in food at the price p_z is still possible but food demand will continue to also be a function of the price of cash crops. In the case of an incomplete food market with partial food trade household food demand is given by:

$$z^* = z\big(p_z, p_c, p_a, p_f, \Pi + y\big). \tag{4.18}$$

In this paper, we proceed by not assuming the complete failure of food markets and instead estimate equation (4.18).

Since one of the objectives of the paper is to test for the existence of constrained food markets, we take advantage of the fact that, controlling for income, household demand for food is a function of the price of cash crops in the presence of constrained food markets but not with perfectly functioning markets. Empirically, we estimate household demand for food and test for the significance of the price of cash crops. In our main analysis we do not include the price of fertilizer in our estimates because we have limited and unreliable data on the price of fertilizers, as explained in Section 4.4.

Ex-ante, the effects of the price of cash crops on food consumption in the presence of constrained markets is ambiguous. An increase in the price of cash crops raises income which increases household demand for food (under the standard assumption that food is a normal good), we call this positive effect the income effect. But, at the same time, an increase in p_a also shifts household agricultural production away from home-production which, with constrained markets, reduces household food consumption, we call this negative effect the agricultural substitution effect. This paper is one of the first to explicitly test for the existence of constraints in the market for food. We believe that an understanding of these market constraints is extremely important for informing both agricultural and health policies.

In the following section we describe the data used for this demand analysis. In Section 4.5 we describe the estimation and in Section 4.6 we present the estimation results and interpret both the statistical and economic significance of the coefficient estimates. We then discuss in more detail the policy implications of our results and provide some concluding remarks in Section 4.7.

4.4 Data

The socio-economic data for this paper come from the 2010-2011 Tanzania National Panel Survey. The survey contains individual, household, and community level information on 3,924 households and produces nationally representative estimates of agricultural production and poverty for Tanzania. Because we are concerned with the relationship between household agricultural production and food consumption, we use only households that participate in farming. Additionally, we drop households that have unreasonable daily caloric per capita estimates. These two restrictions leave us with a final sample size of 2,337 households.[6]

Most notably for the purposes of this paper, the survey contains a comprehensive food consumption questionnaire. This questionnaire asks households for the total amount consumed in the last seven days for 46 food items, including cereals, pulses, nuts, vegetables, fruits, meats and fish, as well as the source of food consumed, including purchased and home-produced food. In order to understand the degree to which households are nutrient poor, we use the food consumption questionnaire to derive nutrient consumption values for each household. In addition, we use the food consumption questionnaire to derive weekly household food consumption values, measured in kilocalories, for five different aggregate food groups – staples, pulses, vegetables and fruits, animal products, and meal complements – and we construct price indices for each of the five food groups. In what follows, we describe both the nutrient decomposition methodology and the construction of the price indices.

4.4.1 Nutrient Decomposition

In motivating this paper, we claim that identifying the presence of food market constraints that affect household food demand is important because food is the primary source of nutrients. In order to better understand household nutrient consumption and the prevalence of nutrient deficiencies in Tanzania we decompose the household food consumption data

[6]Roughly 38 percent of households do not participate in farming (1,416 households). We define unreasonable daily caloric per capita consumption as being less than 500 kilocalories per day (29 households). Finally, another 142 households have no observed consumption data and are dropped from the analysis. In total, 1,587 households are dropped and the observations that could bias our results (171) only account for seven percent farming households.

into nutrient consumption levels. For the decomposition, we use both the FAO food decomposition table for Kenya, Senegal, and Mexico and the United States Department of Agriculture's National Nutrient Database for Standard Reference.[7] All of the 46 food items were measured in grams.[8] For food items in the consumption survey that contain multiple specific foods, such as "onions, tomatoes, carrots, and green pepper," the average nutrient value per gram of those foods was used as the nutrient conversion rate.

To obtain population-level nutrient deficiency rates, we use data from multiple sources to create an individual-specific data set of both recommended and required nutrient levels. Recommended and required daily nutrient levels are created for energy (kcal), protein (g), iron (mg), zinc (mg), vitamin A (μg RE), riboflavin (mg), folate (μg), vitamin B_{12} (μg) and vitamin C (mg).[9] An individual is labeled as deficient in a particular nutrient if his or her derived nutrient consumption level is below his or her required level.

Recommended and required energy levels come from the FAO, WHO, and United Nations University technical report "Human energy requirements" (2001). Recommended energy levels are the amount of kilocalories necessary to maintain a normal lifestyle with a physical activity level (PAL) of 1.75 and a body mass index (BMI) of 21. Required energy levels are the amount of kilocalories necessary to maintain a normal lifestyle with a PAL of 1.45 and a BMI of 18.5. Ultimately, food and nutrient consumption levels are reported as daily per capita consumption values where individual specific weights are calculated based on kilocalorie requirement levels relative to adult males.[10]

Recommended and required protein levels come from the FAO and WHO technical report "Energy and Protein Requirements" (1985). Recommended protein levels are defined as the "safe level of protein intake" which is the protein level recommended to meet protein needs for 97.5% of individuals in the associated age- and gender-specific population group. Required protein levels are the "average protein intake" which is equivalent to subtracting two standard deviations of the average protein intake from the "safe level of protein intake."[11]

---

[7] The authors gratefully acknowledge Olivier Ecker for help in obtaining the FAO conversion tables. The USDA National Nutrient Database was used only to obtain nutrient levels of standard foods, e.g. spaghetti and canned milk.

[8] Each egg consumed was given a weight of 45 grams.

[9] The symbol μg denotes microgram, RE retinol equivalence, and mg milligram.

[10] The average individual weight in the survey is 0.785 adult males.

[11] The formula used to relate average and safe levels of protein intake is: $safe = average + (2 \times SD)$.

The FAO (1985) reported both recommended and required protein levels in terms of grams per kilogram of body mass. To convert these levels to grams of protein per day this reported rate was multiplied by an individual's weight.

For the remaining seven vitamins and minerals the recommended nutrient intake (RNI) was reported as the recommended level of these nutrients. The RNI is the daily intake which meets the nutrient requirements of 97.5% of individuals in the age- and gender-specific population group. All RNI levels were taken directly from the "Vitamin and Mineral Requirements" report published jointly by WHO and FAO (2004). The required nutrient level was reported as the estimated average requirements (EAR), which is equivalent to two standard deviations below the RNI. The conversion rates for converting the RNI to the EAR were found in the "Guidelines on Food Fortification with Micronutrients" report published by WHO and FAO (2006).

Per capita daily food and nutrient consumption levels for the 2,337 households in the 2010-2011 Tanzania survey are reported in Table 4.1.[12] This table shows that many Tanzanians are deficient in a number of different nutrients, most notably iron, zinc, and vitamin B_{12}. Our nutrient decomposition reveals that 37 percent of Tanzanians are undernourished – a statistic that is in line with the FAO (2008) estimate of undernourishment in Tanzania at 44 percent for 2001-2003. In addition, the table shows that, on average, almost 75 percent of an individual's daily energy needs are met through consumption of staple foods, with maize and rice being the two largest sources of nutrients – both of which are poor sources of vitamins A, B_{12}, and C.

4.4.2 Food Group Price Construction

In order to perform a demand analysis for household food demand we also need to construct food price indices. We estimate household demand for five mutually exclusive food groups – staples, pulses, fruits/vegetables, animal products, and meal complements – and, therefore we require a group-specific price index for each of these five food groups measured in Tanzanian shillings (TSh) per kilocalorie. In addition, we also calculate an overall food price

[12]Households that reported consuming more than 5,000 kilocalories per day were assigned nutrient consumption levels equal to the 99th percentile for any food group where reported nutrient consumption was in excess of the 99th percentile level.

index measuring the price of kilocalories. In this subsection, we explain the steps taken to construct these six price indices.

To begin, each household in the survey is assigned a price per gram of food for each of the 46 food items in the consumption questionnaire. The price per gram is equal to the regional average of the observed district price from the price questionnaire. For the 17 food items where no price data were collected, the price is measured as the regional average of the household-reported purchase cost and purchase amount values from the consumption questionnaire. Any remaining households with missing price data were assigned the sample average price.

A food-specific price per kilocalorie was then created by dividing the price per gram by the amount of kilocalories per gram for each food item, obtained from the FAO food decomposition tables described above. The price per kilocalorie is measured in terms of Tanzanian shillings per kilocalorie for food item k:

$$p_{k,kcal} = \frac{p_k}{kcal_k} = \frac{TSh}{gram} \times \frac{gram}{kcal} = \frac{TSh}{kcal}.$$

From the nutrition decomposition, we know the amount of kilocalories that household n in village v consumed through food k, $kcal_{knv}$. This amount is then multiplied by the price per kilocalorie of food k, $p_{k,kcal}$ to obtain a measure of household expenditures on kilocalories from food k, $\pi_{knv} = kcal_{knv} p_{k,kcal}$.

After having obtained food-specific prices measured in Tanzanian shillings per kilocalorie we construct a price index for each of the five food groups using a Stone index (Deaton and Muellbauer, 1980). We use the Stone price index instead of the Paasche, Laspeyres, or Fisher index because we have prices from only one point in time and it is the only one of these four indices to not require a set of base period prices in its construction. With the Stone index, the price of food group i is a weighted average of the price per kilocalorie for each of the K food items in food group i:

$$p_i = \exp\left(\sum_{k=1}^{K} \frac{\bar{w}_{jk}}{\sum_j \bar{w}_{jk}} p_{k,kcal}\right) \tag{4.19}$$

where $\bar{w}_{jk}$ is the mean household budget share spent on food j, measured in kilocalorie expenditures.[13] Ultimately, this Stone index price construction results in six food related price indices: price per kilocalorie of food, price per kilocalorie of staples, price per kilocalorie of pulses, price per kilocalorie fruits and vegetables, price per kilocalorie of animal products, and price per kilocalorie of meal complements.

4.4.3 Market and Agricultural Prices

In addition to the food price indices, household food demand estimation also requires price indices for non-food market goods and for cash crops. In this subsection, we describe the construction of these two final price indices and the derivation of the price of three agricultural inputs – organic fertilizer, inorganic fertilizer, and pesticides.

Non-Food Market Goods Price Index

Non-food market goods consist of all goods in the community price questionnaire that are not food items. This restriction results in the inclusion of three goods into this price index: kerosene, charcoal, and maize grinding costs. All costs are expressed in Tanzanian shillings per gram before being indexed together. We obtain the prices for all three of these goods from the price questionnaire and again use the regional average price as a household's price. Household expenditures on each of the three goods were obtained directly from the household non-food expenditure questionnaire. We then construct the price index for non-food market goods using a Stone price index identical to the index formula in equation (4.19).

Cash Crops Price Index

To construct the cash crop price index, we focus on agricultural crops that households produce solely for sale in the market and that are never consumed at home. Ultimately, only three crops meet these criteria and are used in the construction of the cash crop price index: cotton, sesame, and tobacco. Overall, nine percent of households (214) report selling at least one of these three crops and they account for 13 percent of total household agricultural sales.

---

[13] Using the notation from above and, assuming the survey has N different households, we have:

$$\bar{w}_{jk} = \frac{1}{N}\sum_{v}\sum_{n}\frac{\pi_{knv}}{\Pi_{nv}} = \frac{1}{N}\sum_{v}\sum_{n}\frac{kcal_{knv}p_{k,kcal}}{\Pi_{nv}}$$

where Π_{nv} denotes total household expenditure for household n in village v.

The selling price per gram harvested for each crop is defined as the household-reported value of sales divided by the household-reported quantity (in grams) of sales. All households receive the regional average of these derived sale prices. We again use a Stone index to construct the price index but for this construction the weights are no longer household budget shares. For the cash crop price index, the weight used for crop k in the Stone index construction is the share of crop k sold, in Tanzanian shillings, out of the total value of all crops sold.

Agricultural Input Prices

Finally, the theoretical model displayed in equation (4.18) shows that in the presence of food market constraints household food demand is a function of both the price of cash crops and the price of agricultural inputs, including the price of fertilizer. Although the survey collects information on the amount of organic fertilizer, inorganic fertilizer, and pesticide used there is no direct question on the price of these purchased inputs. We calculate the price of an input as the amount spent on the input divided by the total amount used. Households are then assigned the regional average of these constructed prices. The survey, however, asks households for how much of each input was used and the total amount spent to purchase each input but not on the amount of each input that was received at a subsidized price or with a voucher. Because almost 50 percent of households report receiving vouchers these prices may be significantly overstated. Consequently, we include the price of fertilizers and pesticides as a sensitivity check but not as explanatory variables in the main estimation results.

The prices and weights (budget and harvest share weights) for all eight price indices are displayed in Table 4.2 along with the prices of the three agricultural inputs. Overall, food costs, on average, 0.8 Tanzanian shillings per kilocalorie. The most expensive food group per kilocalorie is vegetables and fruits, with an average price of 3.5 Tanzanian shillings per kilocalorie. Staple foods are the cheapest source of energy, with a price of 0.3 Tanzanian shillings per kilocalorie. The average price (per gram) of non-food marketed goods is 0.5 and the average price of cash crops is 1.7 Tanzanian shillings per gram.

Data Limitations

The nutrient decomposition method outlined above has certain drawbacks that could lead to measurement error in our observed food and nutrient consumption levels. In this subsection, we highlight some of those drawbacks and explain in what ways they may or may not affect our results.

All household surveys must deal with the issue of identifying the appropriate recall period. A long recall period increases the chance of an error in response by the household while a short recall period does not provide an accurate picture of a household's average food consumption (Clarke, Fiebig and Gerdtham, 2008). The Tanzania National Panel Survey uses a seven-day recall period for the food consumption questionnaire, which should limit measurement error from recall bias. Finally, households in the survey are questioned over all 12 months of the year so even with the seven-day recall period the survey will still provide an accurate picture of annual household nutrient consumption levels (Ecker and Qaim, 2011).

Second, the food decomposition process involves assumptions about the amount of food being consumed and the type of food. While we adjust food consumption for non-household members who ate inside the home, we still assume that all food reported on the consumption questionnaire is consumed by the household. If households store food or give food to neighbors then this assumption will lead to an overestimation of total food consumption (Bouis, 1994; Ecker and Qaim, 2011). We also have information only on total household food consumption, so in order to construct per capita consumption levels we assume that food is distributed equally among household members based on caloric needs. This intra-household food distribution method could lead to an over or underestimation of per capita consumption depending on the true intra-household distribution of food (Haddad and Kanbur, 1990).

Finally, the 46 food items reported in the food consumption questionnaire contain groups of foods whose nutrients need to be aggregated together for the nutrient decomposition. For example, the food consumption questionnaire asks households for the amount of "onions, tomatoes, carrots and green pepper" that the household consumed in the last seven days. In order to decompose this response into nutrients we assume that the household consumed equal amounts of onions, tomatoes, carrots, and green pepper and average together the nutritional makeup of each of the four foods. If, in reality, the households diet was unbalanced then this methodology will lead to erroneous nutrient consumption levels. A more detailed food consumption questionnaire could decompose these food item groups further but is also likely to suffer from more measurement error due to recall bias.

Despite these limitations, we believe that the nutrient decomposition methodology employed in this paper, and previously in Ecker and Qaim (2011), is an important step in gathering more detailed information on household nutrient deficiencies and food consumption patterns. From this decomposition method we can assess not only which nutrients household members are over- or under-consuming but also the food sources on which households rely to obtain different nutrients.

4.5 Empirical Setup

To empirically test for the presence of constrained food markets we estimate households' demand for food. Food demand estimates are typically performed in one of two ways. First, food demand is estimated based on the share of a household's budget spent on food using an almost ideal demand system (Deaton and Muellbauer, 1980; Ecker and Qaim, 2011; Tekgüç, 2012). While this approach follows a standardized demand analysis, estimating a household's food budget share is less accurate when households consume food that is grown at home and never purchased in the market. As shown in Figure 4.1, households in Tanzania consume a large amount of home-produced food: close to 60% of staple food consumption is obtained from home production. With large amounts of home-produced food a more accurate way to measure household food demand is by estimating household demand for energy, measured in kilocalories directly. In this paper, we estimate household demand for energy using an approach similar to the one laid out by Subramanian and Deaton (1996); household kilocalorie demand is a function of household income, household characteristics, market prices and, unlike Subramanian and Deaton (1996), agricultural output prices. In Section 4.6.3, we also estimate household demand for nine micronutrients to test for the differential effects that food market constraints may have on micronutrient consumption.

We divide food into five categories - staples, pulses, fruits and vegetables, animal products, and meal complements - and estimate the demand for each food group. We estimate demand for household daily per capita food consumption measured in kilocalories for each of the five groups. We chose this unit of measurement because kilocalorie consumption is highly correlated with total food consumption, measured in grams (see Table 4.3), is positively correlated to consumption of all other household nutrients (Table 4.3), and is one of the components of food that a household can most easily judge when making food consumption decisions.[14] The coefficient estimates are also more comparable across food groups when using kilocalorie consumption as the outcome variable because a kilocalorie of energy from staple foods is equivalent to a kilocalorie of energy from pulses. In Section 4.6.3, we explore the the effect of food market constraints on a household's nutrient consumption by estimating households' demand for nutrients.

In our main estimates, the outcome variable is measured as the daily per capita kilocalorie consumption from food group i by household n in village v, and the demand for kilocalories is given by:

$$\begin{aligned} cal_{inv} = \beta_{i0} + \boldsymbol{p}'_{cal,v}\boldsymbol{\beta}_1 + \beta_{i2} p_{nonfood,v} + \beta_{i3} p_{cash\ crops,v} + \beta_{i4} income \\ + \beta_{i5} household\ characteristics_{nv} + \beta_{i6} village\ characteristics_v + \epsilon_{inv} \end{aligned} \tag{4.20}$$

As shown in Section 4.3, constrained food markets exist if, after controlling for household income, $\beta_{i3} \neq 0$. As noted in Section 4.3, the effect could be positive or negative depending on whether the agricultural substitution effect or income effect dominates. If there are no constraints in the local food market then after controlling for household income the price of cash crops will have no effect on per capita household food consumption.

The three price variables included in equation (4.20) correspond to the notation from Section 4.3, with the omission of the price of agricultural inputs. The price of food, p_{cal}, is a five unit vector containing the five food group price indices: staples, pulses, fruits and vegetables, animal products, and meal complements. The price of aggregate non-food market goods, $p_{nonfood}$, is the price index of the three most commonly purchased non-food items: kerosene, charcoal, and maize grinding costs. Finally, the price of cash crops, $p_{cash\ crops}$, is a price index of the three most commonly sold non-food crops: cotton, sesame, and tobacco.

[14] For example, households might also make food consumption decisions based on personal taste preferences but this component of food choice is unobserved.

Exact construction of these price indices is given in Section 4.4.2.

For the entire demand system, we use the same set of explanatory variables to control for household, village, and market characteristics. We use annual household expenditures as a proxy for both a household's agricultural profits and non-labor income because annual expenditures is highly correlated to annual income but suffers from less measurement error (Deaton, 1997). Additional household controls include household size (kilocalorie adjusted based on an individual's kilocalorie requirement relative to adult males), proportion of children in the household, and a dummy equal to one if the household head is female. We also include month of interview dummies to control for seasonality in prices. Unfortunately, because the variation in the prices is at the regional level we cannot include any village fixed-effects. In order to control for village-level variables that will also affect prices we include a dummy equal to one if the household lives in an urban area, household distance to the nearest trunk road, and household distance to the nearest town of 20,000 or more people.[15]

Summary statistics for all household, village, and market characteristics are displayed in Table 4.4. These summary statistics show that the average household size is 5 individuals, with a calorie-adjusted size of 4. On average, 40 percent of household members are children and roughly 25 percent of household heads are female. In terms of geographic location, only 15 percent of households are classified as living in urban areas, the average distance to the nearest town of 20,000 or more people is 51 kilometers and, on average, households live 22 kilometers away from the nearest trunk road.

To estimate this demand system we use a system feasible generalized least squares approach (Zellner, 1962; Cameron and Trivedi, 2005). The major advantage of using an SFGLS approach is not only that it estimates the full variance-covariance matrix associated with our demand system but that we can use this matrix to test cross-equation restrictions. This feature is especially beneficial for this paper because we wish to test for the joint significance of the price of cash crops for all five food groups, i.e. $\hat{\beta}_{1,3} > 0, \hat{\beta}_{2,3} > 0, \ldots, \hat{\beta}_{5,3} > 0$. In the analysis, we run this test for joint significance of the home-production variables using a

[15]We also tested all estimation results without these village-level variables in case they are significantly correlated with the price of cash crops. The coefficient on price of cash crops is not significantly different with the inclusion or exclusion of these controls.

Wald test (Wooldridge, 2010, p. 172).

Ultimately, we rely on a seemingly unrelated regression but estimate it using an SFGLS approach so that we can obtain the full variance-covariance matrix. System feasible generalized least squares (SFGLS) is more efficient than equation by equation ordinary least squares if components of the error vector, $[\epsilon_{1nv}, \epsilon_{2nv}, \ldots, \epsilon_{nv9}]$, are correlated. In a system of demand equations, like the ones estimated in this paper, it is likely that any random shock to household k in village v that affects the demand for staples will also affect the demand for pulses, leading to correlation of the error terms. If, however, the same set of explanatory variables are used in all estimated equations, as in this paper, then SFGLS reduces to equation by equation ordinary least squares estimates.

With the SFGLS approach, the vector of coefficients is estimated as:

$$\hat{\boldsymbol{\beta}} = \left(\sum_v \sum_n \boldsymbol{x}'_{nv} \hat{\boldsymbol{\Omega}}^{-1}_{nv} \boldsymbol{x}_{nv} \right)^{-1} \sum_v \sum_n \boldsymbol{x}'_{nv} \hat{\boldsymbol{\Omega}}^{-1}_{nv} \boldsymbol{cal}_{nv}$$

For our analysis we use a two-step SFGLS approach, where $\hat{\boldsymbol{\Omega}} = 1/N \sum_n \sum_v \hat{\boldsymbol{\epsilon}}_{nv} \hat{\boldsymbol{\epsilon}}'_{nv}$, with cluster robust standard errors:

$$\mathrm{Var}[\hat{\boldsymbol{\beta}}] = \left(\sum_v \sum_n \boldsymbol{x}'_{nv} \hat{\boldsymbol{\Omega}}^{-1}_{nv} \boldsymbol{x}_{nv} \right)^{-1} \sum_v \sum_n \boldsymbol{x}'_{nv} \hat{\boldsymbol{\Omega}}^{-1}_{nv} \hat{\boldsymbol{\epsilon}}_{nv} \hat{\boldsymbol{\epsilon}}'_{nv} \hat{\boldsymbol{\Omega}}^{-1}_{nv} \boldsymbol{x}'_{nv} \left(\sum_v \sum_n \boldsymbol{x}'_{nv} \hat{\boldsymbol{\Omega}}^{-1}_{nv} \boldsymbol{x}_{nv} \right)^{-1}.$$

In the analysis, we report standard errors from a 500 replication block bootstrap. We block bootstrap the robust standard errors at the cluster level to account for any additional heteroskedasticity in the data and the sampling procedure where first-stage sample selection was done at the cluster level.

4.6 Results

Overall, these results provide evidence for the presence of constraints in the market for staple foods and little evidence of constraints in the other four food markets. Results from the main SFGLS estimation are presented in Table 4.5. The price of cash crops is significant in household demand for staple foods only, and the Wald test fails to reject the null hypothesis that the price of cash crops is not significant across all five equations. This significance provides some evidence for the non-separability of food consumption and agricultural production decisions, but stronger evidence for the non-separability of staple consumption and agricultural production decisions. Constraints in the market for staple foods and not other food groups are plausible because staple foods are the food group that accounts for the largest share of households' energy consumption (75 percent) and the largest share of households' agricultural harvest (30 percent), measured in value harvested.

Because the coefficient on the price of cash crops is negative for staples, these results provide evidence that for staples the agricultural substitution effect dominates the income effect. As the price of cash crops rises households produce more cash crops, less home-produced food, and, because of food market constraints, total consumption of staples falls. A direct interpretation of the coefficient estimates shows us that a doubling in the price of cash crops from 1.771 to 3.542 Tanzanian shillings per kilocalorie results in a reduction in staple food consumption by 177 kilocalories per day.

Examining the additional control variables in Table 4.5 we see that an increase in household size is associated with a reduction in daily per capita food consumption across all five food groups. This result supports the findings of Deaton and Paxson (1998), who showed evidence for no economies of scale in households and that larger households do not consume more per capita. Urban households also consume less food, but households with higher annual expenditures consume more food. Finally, household distance to the nearest town is positively associated with consumption of staples and pulses – as households live farther away from town they obtain, on average, more energy from the consumption of staples and pulses.

Finally, we see that, in contrast to standard demand theory, the own price coefficients for staples, pulses, and vegetables and fruits are positive and none of the five own price coefficients are statistically significant. Both de Janvry, Fafchamps and Sadoulet (1991) and Taylor and Adelman (2003) show that in the presence of food market constraints the impact of food prices on food demand is muted. Specifically, for staples and pulses a large share of households are net sellers of staples and pulses. Table 4.6 displays the proportion of households that buy and sell, and that are net buyers and net sellers, for each of the food groups that are also agricultural crops (staples, pulses, vegetables/fruits). Almost 30 percent of households are net sellers of staple foods and pulses. For these households, an increase in the price of staples or pulses could have a positive effect on household food demand because the price increase raises their incomes. For vegetables and fruits, however, over 95 percent of households are net buyers and the own-price coefficient is also positive but insignificant. The positive price for vegetables and fruits can partially be explained by the difference in consumption patterns for the food group; urban households consume more vegetables and fruits than rural households (75 kilocalories per day versus 66, significant at the five percent level) and there is little difference in the price of the food group between the two areas. In addition, urban households purchase 70 percent of their vegetables and fruits while rural households purchase only 50 percent. Therefore, one explanation for the positive own-price coefficient on vegetables and fruits in the presence of food market constraints is that it is driven by urban households that not only consume more of the good but that also purchase more of the good.

4.6.1 Agricultural Input Prices

In Table 4.5 we tested for non-separability by examining whether the price of cash crops significantly affects household food demand. As shown in Section 4.3, if household consumption and production decisions are non-separable then agricultural input prices will also affect household food demand. We did not include the price of fertilizers and pesticides in our initial estimation results because we believe these price data are not reliable. In this subsection, however, we run an additional test for non-separability (see Table 4.7) where food demand is a function of the price of cash crops and the price of three agricultural inputs: organic fertilizer, inorganic fertilizer, and pesticides.

Table 4.7 corroborates the findings in Table 4.5. The coefficient estimates for the price of cash crops are almost identical to those reported in Table 4.5 and estimates for the other explanatory variables in Table 4.5 are also unaffected. Of the three input price variables, the price of organic fertilizers is the largest in magnitude and is also significant in the demand for staple foods. Again, we reject the null hypothesis that these agricultural prices are jointly significant across all five demand estimates. Even though these results tell a similar market constraint story to the one told by Table 4.5, we rely on Table 4.5 as our main results. Not only are the input prices overestimated but also only 31 percent of households report using either organic or inorganic fertilizer and only 13 percent of households use pesticides.

4.6.2 Household Expenditures and Measurement Error

In Section 4.5 we mention that we use household expenditures to proxy for household annual income because household expenditures suffers from less measurement error. In reality, annual household expenditures may still contain substantial measurement error. It is difficult for any individual or household to accurately report the correct amount of all education, non-food, household, and food costs that they have incurred over the last week, month, or year. With this classical measurement error in one explanatory variable coefficient estimates from the initial estimates in Table 4.5 will be biased towards zero (Bound, Brown and Mathiowetz, 2001). In order to account for this measurement error in household expenditures we run an additional estimate instrumenting for annual household expenditures.

To attempt to remove the measurement, we estimate the system of five equations with instrumental variables using a generalized method of moments three-stage least squares estimator (GMM 3SLS) (Wooldridge, 2010). In order for our instruments to be plausible we need them to explain some of the variation in household expenditures but to be uncorrelated with the error term, ϵ_{inv}. For our instruments, we focus on vectors of fixed assets that are highly correlated with household expenditures but that do not affect household food demand. Specifically, we use instruments that measure a household's resource endowment for agricultural production, such as the number of adult males, adult females, girls, boys, acres of land cultivated, number of livestock owned, home ownership, water source, light source, cooking fuel, and sewage disposal (Jacoby, 1993).[16] These resource endowment variables are fixed endowments that are not correlated with short-term agricultural production shocks but are correlated with expected agricultural profits. The second set of instruments are asset holdings which take time to accumulate and are not easily sold in the short-run (Pattanayak

[16]The exact set of of these resource endowment instruments are: number of males 16 and older, number of females 16 and older, number of males 15 and younger, number of females 16 and younger, number of acres of land cultivated, number of livestock owned, home ownership dummy (1 if own, 0 otherwise), water source dummy (0 if from river lake, or rainwater, 1 otherwise), light source dummy (1 if electric, 0 otherwise), cooking fuel dummy (0 if firewood, 1 otherwise), and sewage disposal dummy (0 if none, 1 otherwise).

et al., 2004). Specifically, we include the number of phones and the number of televisions owned by a household. These two assets are not expected to directly affect food demand but are the first two assets that households buy once they have more expendable income (The Economist, 2014).

The GMM 3SLS is a GMM estimator with the weighting matrix:

$$\hat{\boldsymbol{w}} = \left(\frac{1}{N} \sum_{v} \sum_{n} \boldsymbol{z}'_{nv} \left(\frac{1}{N} \sum_{v} \sum_{n} \hat{\boldsymbol{u}}_{nv} \hat{\boldsymbol{u}}'_{nv} \right) \boldsymbol{z}_{nv} \right)^{-1}$$

where $\hat{\boldsymbol{u}}_{nv}$ are the residuals from the first stage equation (see Table D.1 in the Appendix D). And, the coefficient estimates are given by:

$$\hat{\boldsymbol{\beta}} = \left(\left(\sum_{v} \sum_{n} \boldsymbol{x}'_{nv} \boldsymbol{z}_{nv} \right) \hat{\boldsymbol{w}} \left(\sum_{v} \sum_{n} \boldsymbol{z}'_{nv} \boldsymbol{x}_{nv} \right) \right)^{-1} \left(\sum_{v} \sum_{n} \boldsymbol{x}'_{nv} \boldsymbol{z}_{nv} \right) \hat{\boldsymbol{w}} \left(\sum_{v} \sum_{n} \boldsymbol{z}'_{nv} \boldsymbol{cal}_{nv} \right).$$

This GMM 3SLS is asymptotically efficient as it places no restrictions on the unconditional and conditional variance matrix of $\boldsymbol{u}_{nv}$, i.e E$[\boldsymbol{u}_{nv}]$ and E$[\boldsymbol{u}_{nv}|\boldsymbol{z}_{nv}]$, respectively (Wooldridge, 2010). To account for heteroskedasticity in the error terms, we estimate the demand system using block bootstrapped standard errors (500 replications). In the case that our instruments contain no measurement error, or that any measurement error in the instruments is uncorrelated with the measurement error in annual household expenditures, estimates from this three-stage least squares estimate will be unbiased.

Estimation results for the instrumental variable model are displayed in Table 4.8. The cash crop price index is still significant and negative in household demand for staple foods but, as expected, is now larger in relative magnitude. In this new set of estimates a doubling in the price of cash crops results in a reduction in staple food consumption by 240 kilocalories per day.

These results show that even after controlling for the endogeneity of income the price of cash crops significantly affects household staple food demand. Thus, market constraints appear to play a role in a household's agricultural and consumption patterns. In the following subsection we test for the presence of market constraints in household nutrient consumption and discuss the health implications of our results.

4.6.3 Household Nutrient Consumption

As noted in the Introduction, one of the primary reasons why policy-makers will care about the presence of food market constraints is that they may have negative health and development repercussions. In the literature review, we cite evidence for the link between poor health, specifically nutrient deficiencies, and delayed or stunted cognitive and physical development. In this subsection we test for non-separability between a household's nutrient consumption and agricultural production decisions. Specifically, we run a series of nine demand estimates for kilocalorie, protein, iron, zinc, vitamin A, riboflavin, folate, vitamin B_{12}, and vitamin C. All explanatory variables are identical to the variables used in Table 4.5.

Results, displayed in Table 4.9, show that the price of cash crops significantly affects a household's daily per capita kilocalorie and vitamin A consumption. For these estimates the price of cash crops is also jointly significant at the five percent level. These results provide a final piece of evidence for the presence of food market constraints in Tanzania. Even after controlling for household income, an increase in the price of cash crops results in households consuming less total kilocalories per day. These results are especially relevant to current health policies because over 80% of households in Tanzania are estimated to be iron deficient and anemia affects over 25% of the global population (1.6 billion people) (HarvestPlus, 2014). Although understanding the exact mechanisms of this food substitution is beyond the scope of this paper these results provide an important first-step into understanding the link between food consumption, the types of foods consumed, and nutrient consumption in Tanzania, and developing countries more broadly, where an estimated 40 percent of the population does not meet their daily caloric needs and even more individuals suffer from micro-nutrient deficiencies (see Table 4.1).

4.7 Conclusion

Globally, Africa is frequently identified as a region with an agricultural production "yield-gap." To address this yield-gap, the majority of agricultural development programs currently operating in Tanzania, such as Feed the Future, focus on intensification through increased production of maize, rice, or tobacco or through commercialization of agricultural crops (see a 2006 special issue of *European Journal of Development Research*). The overall focus of these programs has been on increasing agricultural production for the market with little focus on health and nutrition (see Hawkes and Ruel (2012) for the one exception). Climate change further complicates this issue; scientists have recently provided substantial evidence that wheat, rice, field peas, soybeans, and maize grown in areas with higher levels of CO_2 have lower concentrations of iron and zinc (Myers et al., 2014). As shown in this paper, however, a household's agricultural production decisions are intimately linked to its food consumption levels, and subsequently nutrient consumption. Indeed, households in Tanzania make agricultural production decisions based not only on maximizing farm profits but also on household food needs. Consequently, this "yield-gap" in African agricultural production may not be so much a productivity gap but rather a rational response by households to food market constraints.

To test for the presence of constraints in local food markets, we estimate the demand of Tanzanian households for five food groups—staples, pulses, vegetables and fruits, animal products, and meal complements—as a function of cash crop prices. It is one of only two papers to test for food market constraints using household food demand estimates. Unlike Tekgüç (2012), we incorporate both food consumption and agricultural production data. Our results show that there is a significant effect of the price of cash crops on demand for staples, which points to the presence of imperfections in the market for food. These results are robust to both incorporating the price of fertilizers and pesticides and to instrumenting for household income. In all of these additional estimates, an increase in the price of cash crops significantly reduces household consumption of staple foods. Overall, we show evidence for the presence of market constraints in Tanzania that create inter-dependencies between household food consumption and agricultural production. This result is in contrast to previous estimates that have not included agricultural prices as explanatory variables in a household's demand for food (Subramanian and Deaton, 1996; Abdulai and Aubert, 2004; Ecker and Qaim, 2011). Future analyses that estimate food demand in Tanzania, and perhaps sub-Saharan Africa more broadly, should account for this nonseparbaility in their demand estimates with the inclusion of agricultural prices.

Table 4.1: Daily Per Capita Food and Nutrient Consumption and Prevalence of Nutritional Deficiencies (N=2,337)

	Quantity (g)	Calories (kcal)	Protein (g)	Iron (mg)	Zinc (mg)
Staple foods	785.3	2026.0	39.8	14.2	8.0
Maize	343.4	1026.2	23.6	9.7	5.1
Rice	111.1	403.2	7.4	0.7	1.2
Other cereals	110.5	171.3	3.2	1.4	0.6
Cereal products	23.7	64.4	1.6	0.2	0.1
Cassava	124.5	291.6	2.4	1.8	0.7
Potatoes	72.1	69.3	1.5	0.5	0.3
Pulses	72.8	202.0	12.0	3.0	1.6
Regular beans	58.6	120.2	8.6	2.4	1.2
Groundnuts	11.4	64.7	2.9	0.5	0.4
Other nuts	2.7	17.1	0.5	0.1	0.1
Vegetables and fruits	212.5	79.3	2.5	1.3	0.4
Onions and tomatoes	82.5	24.8	0.9	0.4	0.2
Green leafy vegetables	61.5	15.5	1.1	0.7	0.1
Bananas	14.5	13.4	0.1	0.0	0.0
Fruits	47.2	25.6	0.4	0.1	0.1
Animal products	126.7	274.9	25.3	1.3	2.2
Eggs	1.6	2.4	0.2	0.0	0.0
Fish	23.3	26.1	5.0	0.2	0.1
Red meat	17.7	50.5	3.8	0.4	0.6
White meat	40.6	115.6	10.9	0.6	0.7
Other meat	8.6	15.5	1.2	0.1	0.2
Milk and dairy products	35.1	64.9	4.3	0.1	0.5
Meal complements	120.1	311.3	0.2	0.1	0.0
Fat and oil	20.9	184.7	0.0	0.0	0.0
Sugar and sweets	48.9	109.9	0.0	0.1	0.0
Condiments	11.1	0.0	0.0	0.0	0.0
Beverages	39.2	16.7	0.2	0.0	0.0
Total	1239.7	2686.5	79.8	19.9	12.2
Recommendations		2677.4	42.8	40.6	15.1
Requirements		2216.9	34.3	27.2	12.6
Prevelance of deficiency (%)		37.1	11.4	79.7	64.4

Continued on next page

Table 4.1 – continued from previous page

	Vit. A (μg RE)	Riboflavin (mg)	Folate (μg DFE)	Vit. B12 (μg)	Vit. C (mg)
Staple foods	60.6	0.8	173.2	0.0	88.3
Maize	0.0	0.6	96.1	0.0	4.0
Rice	0.0	0.1	6.7	0.0	0.0
Other cereals	46.7	0.1	23.2	0.0	8.6
Cereal products	0.8	0.0	3.5	0.0	0.0
Cassava	13.1	0.0	33.4	0.0	66.8
Potatoes	0.0	0.0	10.4	0.0	8.8
Pulses	12.0	0.1	170.4	0.0	5.3
Regular beans	11.8	0.1	154.3	0.0	5.2
Groundnuts	0.0	0.0	14.4	0.0	0.0
Other nuts	0.2	0.0	1.7	0.0	0.1
Vegetables and fruits	587.5	0.1	68.8	0.0	60.3
Onions and tomatoes	289.2	0.0	13.8	0.0	19.5
Green leafy vegetables	256.7	0.1	41.4	0.0	16.8
Bananas	1.2	0.0	2.8	0.0	1.3
Fruits	40.3	0.0	10.7	0.0	22.6
Animal products	551.6	0.4	20.5	3.8	1.9
Eggs	3.0	0.0	0.7	0.0	0.0
Fish	10.0	0.0	2.6	0.3	0.0
Red meat	486.7	0.1	7.7	2.8	0.4
White meat	15.8	0.1	2.0	0.1	0.0
Other meat	0.2	0.0	0.8	0.1	0.2
Milk and dairy products	35.8	0.2	6.8	0.5	1.3
Meal complements	0.2	0.0	2.7	0.0	2.0
Fat and oil	0.0	0.0	0.0	0.0	0.0
Sugar and sweets	0.0	0.0	0.0	0.0	0.0
Condiments	0.0	0.0	0.0	0.0	0.0
Beverages	0.2	0.0	2.6	0.0	2.0
Total	1211.9	1.5	435.5	3.8	157.8
Recommendations	677.9	1.3	435.0	2.6	51.1
Requirements	484.2	1.1	348.0	2.2	41.7
Prevelance of deficiency (%)	36.9	41.6	47.7	65.4	17.7

Note: All amounts are sample means of daily per capita food and nutrient consumption levels.

Table 4.2: Price Index Summary Statistics

	Price	Budget/Harvest Share
Food group price indices (TSh/kcal)		
All foods	0.820	
	(0.093)	
Staple foods	0.325	
	(0.031)	
Maize	0.258	0.195
	(0.060)	(0.199)
Rice	0.309	0.066
	(0.026)	(0.087)
Other cereals	0.510	0.051
	(0.122)	(0.124)
Cereal products	0.697	0.019
	(0.155)	(0.040)
Cassava	0.248	0.062
	(0.041)	(0.144)
Potatoes	0.631	0.025
	(0.102)	(0.059)
Pulses	0.546	
	(0.091)	
Regular beans	0.610	0.052
	(0.132)	(0.075)
Groundnuts	0.350	0.015
	(0.043)	(0.042)
Other nuts	0.608	0.006
	(0.181)	(0.044)
Vegetables and fruits	3.519	
	(0.833)	
Onions and tomatoes	4.750	0.069
	(1.258)	(0.898)
Green leafy vegetables	2.652	0.034
	(0.957)	(0.056)
Bananas	0.927	0.006
	(0.264)	(0.021)
Fruits	1.427	0.018
	(0.490)	(0.047)
Animal products	1.961	
	(0.343)	
Eggs	3.674	0.004
	(0.334)	(0.012)
Fish	3.508	0.051
	(0.990)	(0.112)
Red meat	1.449	0.031
	(0.265)	(0.060)

Continued on next page

Table 4.2 – continued from previous page

	Price (TSh/g)	Budget/Harvest Share
White meat	1.843	0.092
	(0.454)	(3.258)
Other meat	2.367	0.017
	(0.303)	(0.066)
Milk and dairy products	0.507	0.016
	(0.127)	(0.047)
Meal complements	1.305	
	(0.152)	
Fat and oil	0.336	0.036
	(0.039)	(0.325)
Sugar and sweets	0.565	0.032
	(0.058)	(0.040)
Beverages	3.948	0.020
	(0.666)	(0.060)
Non-food market goods price index (TSh/g)		
Non-food market goods	0.486	
	(0.056)	
Kerosene	1.595	0.070
	(0.235)	(0.075)
Charcoal	0.386	0.022
	(0.102)	(0.075)
Maize grinding	0.057	0.044
	(0.016)	(0.053)
Cash crops price index (TSh/g)		
Cash crops	1.771	
	(0.632)	
Cotton	0.570	0.061
	(0.040)	(0.227)
Sesame	1.172	0.036
	(0.382)	(0.175)
Tobacco	54.413	0.032
	(81.153)	(0.173)
Agricultural input prices (TSh/g)		
Organic fertilizer	0.090	
	(0.092)	
Inorganic fertilizer	1.078	
	(0.362)	
Pesticides	228.949	
	(545.856)	
Observations	2.337	2.337

Note: Standard deviations in paranthesis.

Table 4.3: Cross-Correlation: Daily Per Capita Nutrient Consumption (N=2,337)

	Grams	Calories	Protein	Iron	Zinc	Vitamin A	Riboflavin	Folate	Vitamin B_{12}	Vitamin C
Grams	1.000									
Calories	0.814	1.000								
Protein	0.127	0.126	1.000							
Iron	0.324	0.370	0.914	1.000						
Zinc	0.212	0.234	0.989	0.958	1.000					
Vitamin A	0.187	0.162	0.173	0.358	0.253	1.000				
Riboflavin	0.223	0.222	0.981	0.941	0.992	0.282	1.000			
Folate	0.597	0.521	0.299	0.562	0.396	0.625	0.400	1.000		
Vitamin B_{12}	0.311	0.317	0.476	0.454	0.512	0.310	0.552	0.233	1.000	
Vitamin C	0.283	0.216	0.033	0.251	0.104	0.842	0.116	0.672	0.017	1.000

Table 4.4: Sample Summary Statistics
Household controls, village controls, and instruments

	Mean	Standard deviation
Household characteristics		
Household size	5.45	3.16
Household size (calorie adjusted)	4.19	2.18
Children in household (fraction)	0.42	0.24
Household head (1 if female)	0.23	0.42
Household expenditures (log)	14.50	0.70
Village characteristics		
Urban household (1 if urban)	0.15	0.36
Household distance to nearest trunk road	21.83	24.09
Household distance to nearest town ($>$ 20,000)	51.49	40.47
Instrumental variables		
Number of adult males (16 and older)	1.35	1.03
Number of adult females (16 and older)	1.48	0.91
Number of boys (15 and younger)	1.30	1.35
Number of girls (15 and younger)	1.33	1.36
Agricultural land (acres)	6.66	12.42
Number of livestock owned	5.69	20.06
Home ownership (1 if own, 0 otherwise)	0.89	0.32
Water source (0 if river, lake, or rainwater, 1 otherwise)	0.56	0.50
Light source (1 electric, 0 otherwise)	0.07	0.25
Cooking fuel (0 if firewood, 1 otherwise)	0.10	0.30
Sewage disposal (0 if none, 1 otherwise)	0.77	0.42
Number of mobile phones owned	0.82	1.09
Number of televisions owned	0.09	0.33
Observations	2,337	

Standard deviation in parentheses.

Table 4.5: Food Demand Models: Kilocalories

Dependent variable: Daily per capita kilocalorie consumption

	Staples	Pulses	Vegetables & Fruit	Animal Products	Complements
Household size (calorie adjusted)	−165.55***	−25.10***	−10.38***	−72.61*	−62.65***
	(19.23)	(4.30)	(2.85)	(37.53)	(17.39)
Children in household (fraction)	−413.99***	−98.59***	−56.44**	41.55	−72.17
	(107.75)	(28.04)	(24.44)	(164.27)	(83.11)
Household head (1 if female)	20.89	25.48*	40.07	−161.02	−35.58
	(53.72)	(15.41)	(36.96)	(165.35)	(44.66)
Household expenditures (log)	600.23***	89.51***	69.71**	267.63***	206.57***
	(43.59)	(13.12)	(28.29)	(70.82)	(20.21)
Urban household (1 if urban)	−333.95***	−70.75***	−38.19	−222.67	−8.24
	(71.17)	(19.55)	(30.14)	(176.26)	(43.96)
Household distance to nearest trunk road	−1.58	0.41	−0.34	−3.46	2.68
	(1.09)	(0.48)	(0.31)	(3.66)	(2.98)
Household distance to nearest town (> 20,000)	1.74***	1.10***	−0.01	2.36	−1.61
	(0.68)	(0.26)	(0.17)	(2.55)	(1.36)
Price of staples (TSh/kcal)	117.47	−703.58**	−307.16	−2612.46	−1157.15
	(835.81)	(326.54)	(246.62)	(2754.88)	(1069.92)
Price of pulses (TSh/kcal)	−213.67	50.49	−9.41	−626.82	−33.32
	(327.07)	(105.90)	(57.33)	(664.92)	(159.57)
Price of fruits/vegetables (TSh/kcal)	80.65**	−21.17*	23.07	−73.81	16.28
	(34.94)	(11.83)	(15.07)	(54.87)	(65.97)
Price of animal products (TSh/kcal)	−1.34	95.27***	35.82	−86.98	−34.58
	(83.31)	(32.75)	(45.22)	(96.52)	(71.01)
Price of meal complements (TSh/kcal)	175.53	232.89***	3.20	298.84	−222.67
	(198.97)	(64.41)	(26.41)	(193.12)	(413.70)
Cash crops price index (TSh/g)	−99.83**	−6.20	15.96	213.69	7.63
	(50.07)	(16.15)	(11.16)	(200.89)	(20.01)

Continued on next page

Table 4.5 – continued from previous page

	Staples	Pulses	Vegetables & Fruits	Animal Products	Complements
Non-food market goods price index (TSh/g)	1062.34**	−312.46**	41.38	−1741.23	−273.88
	(480.16)	(136.11)	(59.71)	(1938.86)	(346.35)
Observations	2,337				
Wald test (p-value)	0.176				

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Block bootstrapped standard errors (500 replications).

Note: Regression also contains month of interview dummies. Coefficients not reported.

Table 4.6: Net buyers vs Net Sellers

	Count	Buyer	Seller	Net buyer	Net seller
Staple foods	2,029	0.850	0.450	0.684	0.316
		(0.357)	(0.498)	(0.465)	(0.465)
Maize	1,299	0.645	0.444	0.610	0.390
		(0.479)	(0.497)	(0.488)	(0.488)
Rice	1,123	0.877	0.167	0.850	0.150
		(0.328)	(0.374)	(0.358)	(0.358)
Other cereals	394	0.919	0.112	0.901	0.099
		(0.274)	(0.315)	(0.299)	(0.299)
Cassava	349	0.980	0.023	0.977	0.023
		(0.140)	(0.150)	(0.150)	(0.150)
Potatoes	609	0.920	0.122	0.890	0.110
		(0.272)	(0.327)	(0.313)	(0.313)
Pulses	1,352	0.803	0.331	0.714	0.286
		(0.398)	(0.471)	(0.452)	(0.452)
Regular beans	1,217	0.811	0.277	0.759	0.241
		(0.392)	(0.448)	(0.428)	(0.428)
Groundnuts	343	0.752	0.283	0.726	0.274
		(0.432)	(0.451)	(0.447)	(0.447)
Vegetables and fruits	1,978	0.994	0.036	0.970	0.030
		(0.074)	(0.187)	(0.170)	(0.170)
Onions and tomatoes	1,887	0.992	0.031	0.971	0.029
		(0.092)	(0.174)	(0.168)	(0.168)
Green leafy vegetables	776	0.992	0.010	0.990	0.010
		(0.088)	(0.101)	(0.101)	(0.101)
Bananas	240	0.996	0.008	0.992	0.008
		(0.065)	(0.091)	(0.091)	(0.091)
Fruits	424	1.000	0.002	0.998	0.002
		(0.000)	(0.049)	(0.049)	(0.049)

Standard deviation in parentheses.

Note: Count refers to the number of households that report consuming that food item and report either purchasing the food or selling it. Net buyer refers to households that report buying more of the food than they sell while a net seller is households that report selling more of the food than they buy.

Table 4.7: Food Demand Models: Kilocalories with Agricultural Input Prices
Dependent variable: Daily per capita kilocalorie consumption

	Staples	Pulses	Vegetables & Fruit	Animal Products	Complements
Household size (calorie adjusted)	−168.55***	−25.74***	−10.90***	−69.55**	−65.01***
	(19.87)	(4.42)	(2.50)	(34.87)	(20.77)
Children in household (fraction)	−405.30***	−97.93***	−55.51**	35.54	−78.15
	(107.95)	(28.26)	(23.76)	(160.13)	(83.00)
Household head (1 if female)	22.57	22.93	36.88	−156.18	−39.61
	(53.34)	(15.12)	(35.48)	(161.71)	(53.00)
Household expenditures (log)	606.34***	90.25***	69.50**	262.61***	217.08***
	(44.28)	(13.16)	(28.86)	(66.39)	(28.66)
Urban household (1 if urban)	−308.06***	−66.61***	−34.35	−245.69	−4.03
	(69.89)	(19.81)	(27.46)	(196.48)	(36.74)
Household distance to nearest trunk road	−0.81	0.31	−0.51	−3.62	2.49
	(1.19)	(0.52)	(0.36)	(3.95)	(2.76)
Household distance to nearest town (> 20,000)	1.53**	1.11***	0.03	2.45	−1.81
	(0.67)	(0.26)	(0.16)	(2.68)	(1.47)
Price of staples (TSh/kcal)	1095.81	−668.73**	−308.10	−3209.10	−1208.74
	(912.01)	(339.41)	(206.70)	(3230.51)	(922.06)
Price of pulses (TSh/kcal)	−201.83	39.00	−17.19	−606.33	−132.66
	(326.30)	(105.76)	(62.49)	(645.89)	(207.19)
Price of fruits/vegetables (TSh/kcal)	105.57***	−18.76	21.02	−93.00	63.66
	(39.61)	(12.62)	(16.04)	(75.07)	(105.88)
Price of animal products (TSh/kcal)	−60.50	87.64***	26.27	−38.95	−5.48
	(92.10)	(32.94)	(38.47)	(89.63)	(85.90)
Price of meal complements (TSh/kcal)	13.64	246.97***	38.11	354.62	−310.11
	(222.01)	(72.82)	(31.41)	(255.36)	(447.14)
Price of organic fertilizer (TSh/g)	−815.00**	−57.39	28.81	569.96	−748.18
	(326.34)	(82.50)	(57.30)	(692.60)	(858.82)

Continued on next page

Table 4.7 – continued from previous page

	Staples	Pulses	Vegetables & Fruits	Animal Products	Complements
Price of inorganic fertilizer (TSh/g)	−47.02	41.58	50.06*	−68.75	100.75
	(79.23)	(27.95)	(26.63)	(120.40)	(163.45)
Price of pesticides (TSh/g)	0.10**	−0.00	−0.00	−0.04	−0.12
	(0.05)	(0.02)	(0.01)	(0.05)	(0.08)
Cash crops price index (TSh/g)	−104.31**	−3.64	17.88	210.12	28.59
	(49.66)	(16.51)	(12.16)	(197.34)	(30.01)
Non-food market goods price index (TSh/g)	1104.30**	−337.72**	31.03	−1703.49	−583.42
	(492.15)	(144.16)	(63.24)	(1893.03)	(576.22)
Observations	2,337				
Wald test (p-value)					
Wald test (organic fertilizer)	0.140				
Wald test (inorganic fertilizer)	0.366				
Wald test (pesticides)	0.166				
Wald test (cash crops)	0.114				

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Block bootstrapped standard errors (500 replications).

Note: Regression also contains month of interview dummies. Coefficients not reported.

Table 4.8: Food Demand Models: Kilocalories with Household Expenditures Instrumented

Dependent variable: Daily per capita kilocalorie consumption

	Staples	Pulses	Vegetables & Fruit	Animal Products	Complements
Household size (calorie adjusted)	−303.44***	−45.66***	−26.40***	−134.09**	−110.11***
	(37.77)	(7.51)	(4.73)	(56.02)	(25.62)
Children in household (fraction)	−337.38**	−87.17***	−47.54**	75.70	−45.81
	(132.21)	(29.54)	(22.58)	(178.38)	(90.18)
Household head (1 if female)	209.13***	53.55***	61.94	−77.09	29.20
	(65.39)	(17.48)	(44.39)	(145.72)	(39.45)
Household expenditures (log)	1471.18***	219.40***	170.87***	655.96***	506.30***
	(163.53)	(33.29)	(64.74)	(187.72)	(74.48)
Urban household (1 if urban)	−702.93***	−125.78***	−81.05*	−387.18*	−135.22**
	(107.83)	(24.54)	(45.14)	(219.98)	(60.62)
Household distance to nearest trunk road	−1.83	0.37	−0.37	−3.58	2.59
	(1.31)	(0.51)	(0.34)	(3.70)	(2.99)
Household distance to nearest town (> 20,000)	2.69***	1.24***	0.10	2.78	−1.29
	(0.81)	(0.29)	(0.15)	(2.68)	(1.31)
Price of staples (TSh/kcal)	25.22	−717.34**	−317.88	−2653.59	−1188.90
	(939.83)	(338.85)	(252.68)	(2769.70)	(1077.47)
Price of pulses (TSh/kcal)	−634.85*	−12.33	−58.33	−814.61	−178.27
	(384.50)	(111.56)	(78.92)	(726.84)	(163.35)
Price of fruits/vegetables (TSh/kcal)	169.71***	−7.89	33.41*	−34.10	46.93
	(46.03)	(13.46)	(18.98)	(51.19)	(69.50)
Price of animal products (TSh/kcal)	124.37	114.02***	50.42	−30.93	8.69
	(103.36)	(35.19)	(50.27)	(95.81)	(67.94)
Price of meal complements (TSh/kcal)	132.17	226.42***	−1.84	279.51	−237.59
	(231.97)	(70.77)	(30.95)	(196.77)	(417.25)
Cash crops price index (TSh/g)	−135.25**	−11.48	11.84	197.89	−4.56
	(60.12)	(17.53)	(10.63)	(198.05)	(21.39)

Continued on next page

Table 4.8 – continued from previous page

	Staples	Pulses	Vegetables & Fruits	Animal Products	Complements
Non-food market goods price index (TSh/g)	969.82*	−326.26**	30.63	−1782.48	−305.72
	(543.95)	(149.89)	(66.38)	(1951.40)	(351.13)
Observations	2,337				
Wald test (p-value)	0.115				

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Block bootstrapped standard errors (500 replications).

Note: Regression also contains month of interview dummies. Coefficients not reported. Household expenditure instruments: number of males 16 and older, number of females 16 and older, number of males 15 and younger, number of females 16 and younger, number of acres cultivated, number of livestock, home ownership dummy (1 if own, 0 otherwise), water source dummy (0 if from river lake, or rainwater, 1 otherwise), light source dummy (1 if electric, 0 otherwise), cooking fuel dummy (0 if firewood, 1 otherwise), sewage disposal dummy (0 if none, 1 otherwise), number of cell phones, and number of televisions.

Table 4.9: Household Nutrient Consumption: Ordinary Least Squares
Dependent variable: Daily per capita nutrient consumption

	Calories	Protein	Iron	Zinc
Household size (calorie adjusted)	−248.98***	−11.56***	−1.57***	−1.32***
	(26.31)	(3.69)	(0.28)	(0.28)
Children in household (fraction)	−542.27***	−11.66	−4.37***	−2.70**
	(106.09)	(16.02)	(1.33)	(1.21)
Household head (1 if female)	23.22	−12.83	0.35	−0.52
	(48.64)	(15.81)	(1.16)	(1.14)
Household expenditures (log)	986.03***	41.67***	5.00***	5.20***
	(41.40)	(6.98)	(0.71)	(0.57)
Urban household (1 if urban)	−406.71***	−31.25*	−5.20***	−3.46***
	(62.42)	(16.95)	(1.26)	(1.23)
Household distance to nearest trunk road	−1.55	−0.39	−0.04*	−0.04
	(1.05)	(0.35)	(0.02)	(0.02)
Household distance to nearest town (> 20,000)	2.39***	0.30	0.04**	0.03*
	(0.62)	(0.24)	(0.02)	(0.02)
Price of staples (TSh/kcal)	−600.23	−279.34	−27.35*	−21.20
	(791.21)	(262.93)	(16.35)	(18.06)
Price of pulses (TSh/kcal)	−142.18	−63.83	−4.18	−6.04
	(310.69)	(64.18)	(5.32)	(4.84)
Price of fruits/vegetables (TSh/kcal)	−10.68	−5.66	−0.78	−0.72*
	(32.93)	(5.13)	(0.53)	(0.40)
Price of animal products (TSh/kcal)	36.86	−7.47	−1.61	−1.17
	(82.62)	(9.44)	(1.29)	(0.85)
Price of meal complements (TSh/kcal)	557.37***	28.06	15.79***	9.09***
	(165.60)	(19.49)	(2.26)	(1.66)
Cash crops price index (TSh/g)	−84.77*	18.63	0.35	1.03
	(47.57)	(19.23)	(1.16)	(1.32)
Non-food market goods price index (TSh/g)	883.16**	−168.41	−4.94	−7.10
	(398.39)	(183.38)	(10.61)	(12.48)
Observations	2,337			
Wald test (p-value)	0.033			

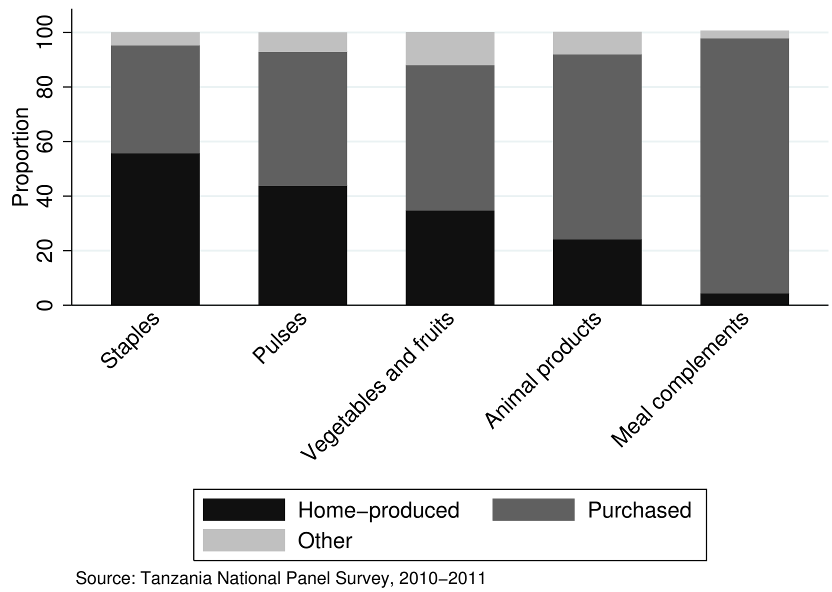
Table 4.9 – continued from previous page

	Vit. A	Riboflavin	Folate	Vit. B_{12}	Vit. C
Household size	−177.34***	−0.18***	−44.46***	−0.82***	−12.81***
	(35.76)	(0.04)	(5.34)	(0.12)	(2.77)
Children in household	−1150.47***	−0.42**	−162.11***	−3.40***	−64.32***
	(288.30)	(0.17)	(28.60)	(0.67)	(24.66)
Household head	474.47	−0.06	55.26**	0.07	50.41
	(435.46)	(0.16)	(27.39)	(0.37)	(31.21)
Household expenditures	1179.73***	0.78***	161.42***	4.52***	63.66***
	(338.30)	(0.08)	(19.95)	(0.37)	(22.66)
Urban household	−324.47	−0.46***	−123.39***	0.13	−70.93***
	(353.15)	(0.16)	(28.01)	(0.48)	(26.38)
Distance to nearest trunk road	−3.66	−0.00	−0.77	−0.00	0.06
	(3.60)	(0.00)	(0.51)	(0.01)	(0.38)
Distance to nearest town	−1.31	0.00	1.16***	−0.00	0.16
	(1.76)	(0.00)	(0.29)	(0.00)	(0.21)
Price of staples	−5281.89*	−2.45	−484.23	−7.80	−370.88
	(2899.04)	(2.38)	(319.33)	(5.57)	(275.66)
Price of pulses	−127.39	−0.39	376.19***	1.13	124.75
	(705.16)	(0.61)	(135.96)	(2.35)	(93.66)
Price of fruits/vegetables	23.29	−0.16***	−11.75	−1.03***	33.56**
	(179.78)	(0.06)	(16.80)	(0.22)	(16.95)
Price of animal products	313.88	−0.24**	41.67	−1.33***	113.76**
	(524.98)	(0.11)	(46.03)	(0.49)	(45.19)
Price of meal complements	1444.10***	1.50***	220.76***	5.61***	−38.88
	(319.36)	(0.22)	(69.14)	(1.19)	(48.04)
Cash crops price index	216.12*	0.20	9.10	0.29	−9.27
	(130.35)	(0.17)	(20.25)	(0.36)	(14.59)
Non-food market goods price index	−144.09	−0.50	−59.19	1.35	57.01
	(728.44)	(1.67)	(142.72)	(3.67)	(97.07)
Observations	2,337				
Wald test (p-value)	0.033				

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Block bootstrapped standard errors (500 replications).

Note: Regression also contains household distance to nearest trunk road, household distance to nearest town, and month of interview dummies. Coefficients not reported.

Figure 4.1: Food Group Kilocalorie Consumption by Source of Food (N=2,337)



Chapter 5

Conclusions

In this dissertation, I have used household-level data from Tanzania to empirically investigate the ways in which local populations interact with their natural environments. In all three essays, my findings show that households in Tanzania intimately rely on and use their local natural environments to meet important daily needs. Overall, this dissertation provides one of the first sets of comprehensive analyses aimed at monetizing the benefits of natural resources in developing countries.

In the first two essays, I demonstrate that households derive significant non-market benefits from forest access in Kagera, Tanzania. In Chapter 2, I estimate that households are willing to pay approximately 25 percent of their annual household expenditure, or $200 (2012 USD) a year for access to local community forests for firewood collection. In Chapter 3, I show evidence that restricted access to forests diminishes human capital formation; an additional hour required to collect firewood when a child is young results in the child completing 0.21 fewer grades of school 19 years later. Assuming an average return to education of eight percent a year, this result translates into $475 2010 USD in lost earnings over 30 years, equal to roughly 1.7 percent of annual income – a cost that aggregates to $7 billion in lost income if the program affects all 15 million rural children in Tanzania.

In the third essay, I show evidence of a strong inter-dependency between households' food consumption and agricultural production. I find that the price of cash crops significantly affects households' demand for food, which points to the present of imperfections in the market for food. This chapter highlights that in the presence of market constraints a household's agricultural production decisions are intimately linked to its food consumption levels, and subsequent nutrient consumption. If an agricultural household must produce food not only to maximize profits but also to meet nutritional needs then a perceived "yield-gap" may be the household's rational response to local market conditions.

The results in this dissertation have many potential policy implications for natural resource management. Effective management of the natural environment requires that all of the benefits of the a given resource are measured and monetized in a common unit. While many benefits can be easily measured and quantified, such as agricultural profits from cleared land, some benefits are experienced outside of the formal marketplace and are much harder to measure. In Chapters 2 and 3 I have measured two such benefits and in Chapter 4 I showed that local market imperfections can affect households' food consumption decisions. The results presented in this dissertation, however are only a small subset of the possible benefits that households in Tanzania derive from their local environments. To date, the literature in this area is sparse but it is increasingly important given the current state of our environment and and current policy focus on the effective management of local natural environments.

References

Abdulai, Awudu, and Dominique Aubert. 2004. "A Cross-Section Analysis of Household Demand for Food and Nutrients in Tanzania." *Agricultural Economics*, 31(1): 67–79. July.

Abdulai, Awudu, and Punya Prasad Regmi. 2000. "Estimating Labor Supply of Farm Households Under Nonseparability: Empirical Evidence from Nepal." *Agricultural Economics*, 22(3): 309–320.

Adhikari, Bhim, Salvatore Di Falco, and Jon C. Lovett. 2004. "Household Characteristics and Forest Dependency: Evidence from Common Property Forest Management in Nepal." *Ecological Economics*, 48(2): 245–257.

Agyei, Yvonne. 1998. "Deforestation in Sub-Saharan Africa." *African Technology Forum*, 8(1).

Ainsworth, Martha. 2004. "User's Guide to the Kagera Health and Development Survey Datasets." Development Research Group, The World Bank, Washington, D.C.

Alderman, H, J Hoddinott, and B Kinsey. 2006. "Long Term Consequences of Early Childhood Malnutrition." *Oxford Economic Papers*, 58(3): 450–474.

Alderman, Harold, Pierre-Andre Chiappori, Lawrence Haddad, John Hoddinott, and Ravi Kanbur. 1995. "Unitary versus Collective Models of the Household: Is it Time to Shift the Burden of Proof." *The World Bank Research Observer*, 10(1): 1–19.

Allen, L., B. de Benoist, O. Dary, and R. Hurrel, ed. 2006. *Guidelines on Food Fortification with Micronutrients.* Geneva, Rome:World Health Organization, Food and Agricultural Organization of the United Nations.

Amacher, Gregory S., Lire Ersado, Donald Leo Grebner, and William F. Hyde. 2004. "Disease, Microdams, and Natural Resources in Tigray, Ethiopia: Impacts on Productivity and Labour Supplies." *Journal of Development Studies*, 40(6): 122–145.

Amacher, Gregory S., William F. Hyde, and Bharat 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.” *The Journal of Development Studies*, 35(4): 138–163.

Asare-Marfo, D, E Birol, C Gonzalez, M Moursi, and S Perez. 2013. “Prioritizing Countries for Biofortification Interventions using Country-Level Data.” *HarvestPlus Working paper*, International Food Policy Research Institute: Washington, DC.

Baland, Jean-Marie, and James A. Robinson. 2000. “Is Child Labor Inefficient?” *Journal of Political Economy*, 108(4): 663–679.

Baland, Jean-Marie, Francois Libois, and Dilip Mookherjee. 2013. “Firewood Collections and Economic Growth in Rural Nepal 1995-2010: Evidence from a Household Panel.” *CEPR Discussion Paper no. 9394. London, Centre for Economic Policy Research.*

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.

Bardhan, Pranab, and Christopher Udry. 1999. *Development Microeconomics.* Oxford, UK:Oxford University Press.

Barrett, Christopher. 2008. “Smallholder Market Participation: Concepts and Evidence from Eastern and Southern Africa.” *Food Policy*, 33(4): 299–317.

Basu, Kaushik, and Pham Hoang Van. 1998. “The Economics of Child Labor.” *The American Economic Review*, 88(3): 412–427.

Baulch, Bob, and Agnes Quisumbing. 2011. “Testing and Adjusting for Attrition in Household Panel Data.” CPRC Toolkit Note.

Becker, Gary S. 1993. *A Treatise on the Family: Enlarged Edition.* Cambridge, MA:Harvard University Press.

Becketti, Sean, William Gould, Lee Lillard, and Finis Welch. 1988. “The Panel Study of Income Dynamics after Fourteen Years: An Evaluation.” *Journal of Labor Ecoomics*, 6(4): 472–492.

Benjamin, Dwayne. 1992. “Household Composition, Labor Markets, and Labor Demand: Testing for Separation in Agricultural Household Models.” *Econometrica: Journal of the Econometric Society*, 60(2): 287–322. March.

Bernard, Tanguy, Alemayehu Seyoum Taffesse, and Eleni Gabre-Madhin. 2008. “Impact of Cooperatives on Smallholders’ Commercialization Behavior: Evidence from Ethiopia.” *Agricultural Economics*, 39(2): 147–161. September.

Bhattacharyya, Anjana, and Subal C. Kumbhakar. 1997. “Market Imperfections and Output Loss in the Presence of Expenditure Constraint: A Generalized Shadow Price Approach.” *American Journal of Agricultural Economics*, 79(3): 860–871.

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.

Bockstael, Nancy E., Ivar E. Strand, and W. Michael Hanemann. 1987. "Time and the Recreational Demand Model." *American Journal of Agricultural Economics*, 69(2): 293–302.

Bockstael, Nancy E., Jr. Ivar E. Strand, Kenneth E. McConnell, and Firuzeh Arsanjani. 1990. "Sample Selection Bias in the Estimation of Recreation Demand Functions: An Application to Sportfishing." *Land Economics*, 66(1): 40–49. February.

Borlaug, N. 2007. "Feeding a Hungry World." *Science*, 318(5849): 359. October.

Bouis, Howard E. 1994. "The Effect of Income on Demand for Food in Poor Countries: Are our Food Consumption Databases Giving us Reliable Estimates?" *Journal of Development Economics*, 44(1): 199–226.

Bound, John, Charles Brown, and Nancy Mathiowetz. 2001. "Measurement Error in Survey Data." *Handbook of Econometrics*, 5: 3705–3843.

Cameron, A. Colin, and Pravin K. Trivedi. 1986. "Econometric Models Based on Count Data: Comparisons and Applications of Some Estimators and Tests." *Journal of Applied Econometrics*, 1(1): 29–53.

Cameron, A. Colin, and Pravin K. Trivedi. 1998. *Regression Analysis of Count Data. Econometric Society Monographs*, Cambridge, United Kingdom:Cambridge University Press.

Cameron, A. Colin, and Pravin K. Trivedi. 2005. *Microeconometrics: Methods and Applications.* Cambridge, United Kingdom:Cambridge University Press.

Cameron, A. Colin, and Pravin K. Trivedi. 2010. *Microeconometrics Using Stata, Revised Edition.* . 2 ed., College Station, TX:Stata Press.

Carpano, F. 2010. "Strengthening Women's Access to Land: The Tanzanian Experience of the Sustainable Rangeland Management Project." International Fund for Agricultural Development.

Chen, L., N. Heerink, and M. Van Den Berg. 2006. "Rural China: A Household Model for Three Villages in Jiangxi Province." *Ecological Economics*, 58(2): 407–420.

Clarke, Philip M., Denzil G. Fiebig, and Ulf-G. Gerdtham. 2008. "Optimal Recall Length in Survey Design." *Journal of Health Economics*, 27(5): 1275–1284.

Cooke, P., G. Köhlin, and W.F. Hyde. 2008. "Fuelwood, Forests and Community Management-Evidence from Household Studies." *Environment and Development Economics*, 13(1): 103.

Cooke, P.A. 1998*a*. "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.

Cooke, P.A. 1998*b*. "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., et al. 2000. "Changes in Intrahousehold Labor Allocation to Environmental Goods Collection: A Case Study from Rural Nepal, 1982 and 1997." *Food Consumption and Nutrition Division Discussion Paper series, Discussion Paper*, , (87).

Creel, Michael D., and John B. Loomis. 1991. "Confidence Intervals for Welfare Measures with Application to a Problem of Truncated Counts." *The Review of Economics and Statistics*, 73(2): 370–373.

de Janvry, Alain, Marcel Fafchamps, and Elisabeth Sadoulet. 1991. "Peasant Household Behavior with Missing Markets: Some Paradoxes Explained." *Economic Journal*, 101(409): 1400–1417.

Deaton, Angus. 1997. *The Analysis of Household Surveys: A Microeconomic Approach to Development Policy.* Johns Hopkins University Press.

Deaton, Angus, and Christina Paxson. 1998. "Economies of Scale, Household Size, and the Demand for Food." *Journal of Political Economy*, 106(5): 897–930. October.

Deaton, Angus, and John Muellbauer. 1980. "An Almost Ideal Demand System." *The American Economic Review*, 70(3): 312–326.

DeGraff, Deborah S., Deborah Levison, and Esther Dungumaro. 2014. "The Implications of Environmental Degradation for Human Capital: Children's TIme Fetching Water and Firewood in Tanzania." manuscript.

Deolalikar, Anil. 1988. "Nutrition and Labor Productivity in Agriculture: Estimates for Rural South India." *Review of Economics and Statistics*, 70(3): 406–413.

Dewees, Peter A. 1989. "The Woodfuel Crisis Reconsidered: Observations on the Dynamics of Abundance and Scarcity." *World Development*, 17(8): 1159–1172.

Duflo, Esther. 2001. "Schooling and Labor Market Consequences of School Construction in Indonesia: Evidence from an Unusual Policy Experiment." *The American Economic Review*, 91(4): 795–813.

Duflo, Esther, and Christopher Udry. 2004. "Intrahousehold Resource Allocation in Côte D'Ivoire: Social Norms, Separate Accounts, and Consumption Choices." National Bureau of Economic Research Working Paper 10498, Cambridge, MA.

Duryea, Suzanne. 2003. "School Attendance, Child Labor and Local Labor Market Fluctuations in Urban Brazil." *World Development*, 31(7): 1165–1178. July.

Ecker, Olivier, and Matin Qaim. 2011. "Analyzing Nutritional Impacts of Policies: An Empirical Study for Malawi." *World Development*, 39(3): 412–428.

Eckholm, Erik. 1975. *The Other Energy Crisis: Firewood.* Washington, D.C:World Watch Insitute.

Edmonds, Eric V. 2007. "Child Labor." *Handbook of Development Economics*, 4: 3607–3709.

Fafchamps, Marcel. 1992. "Cash Crop Production, Food Price Volatility, and Rural Market Integration in the Third World." *American Journal of Agricultural Economics*, 74(1): 90–99.

FAO. 2008. "Tanzania Nutrition Profile." Food and Agriculture Organization of the United Nations.

FAO, WHO, and UNU. 1985. "Energy and Protein Requirements." World Health Organization 724, Geneva, Switzerland.

FAO, WHO, and UNU. 2001. "Human Energy Requirement." Food and Agriculture Organization of the United Nations 1, Rome, Italy.

Fisher, Brendan, Simon L. Lewis, Neil D. Burgess, Rogers E. Malimbwi, Panteleo K. Munishi, Ruth D. Swetnam, R. Kerry Turner, Simon Willcock, and Andrew Balmford. 2011. "Implementation and Opportunity Costs of Reducing Deforestation and Forest Degradation in Tanzania." *Nature Climate Change*, 1: 161–164. June.

Fisher, M., G.E. Shively, and S. Buccola. 2005. "Activity Choice, Labor Allocation, and Forest Use in Malawi." *Land Economics*, 81(4): 503–517.

Fisher, Monica. 2004. "Household Welfare and Forest Dependence in Southern Malawi." *Environment and Development Economicsn*, 9(2): 135–154.

Fisher, Monica, Moushumi Chaudhury, and Brent McCusker. 2010. "Do Forests Help Rural Households Adapt to Climate Variability? Evidence from Southern Malawi." *World Development*, 38(9): 1241–1250.

Fitzgerald, John, Peter Gottschalk, and Robert Moffitt. 1998. "An Analysis of Sample Attrition in Panel Date: The Michigan Panel Study of Income Dynamics." *Journal of Human Resources*, 33(2): 251–299.

Foley, Jonathan A., Navin Ramankutty, Kate A. Brauman, Emily S. Cassidy, James S. Gerber, Matt Johnston, and Nathaniel D. Muellerand et al. 2011. "Solutions for a Cultivated Planet." *Nature*, 478(7369): 337–342.

Foster, Andrew D., and Mark R. Rosenzweig. 2003. "Economic Growth and the Rise of Forests." *Quarterly Journal of Economics*, 118(2): 601–637.

Gillis, Justin. 2011. "With Deaths of Forests, a Loss of Key Climate Protectors." *The New York Times*. October.

Glewwe, Paul. 2002. "Schools and Skills in Developing Countries: Education Policies and Socioeconomic Outcomes." *Journal of Economic Literature*, 40(2): 436–482.

Glewwe, Paul, and Michael Kremer. 2006. "Schools, Teachers, and Education Outcomes in Developing Countries." *Handbook of the Economics of Education*, 2: 945–1017.

Gourieroux, C., A. Monfort, and A. Trognon. 1984. "Pseudo Maximum Likelihood Methods: Theory." *Econometrica: Journal of the Econometric Society*, 52(3): 681–700.

Greenstone, Michael, and B. Kelsey Jack. 2013. "Envirodevonomics: A Research Agenda for a Young Field." *MIT Department of Economics Working Paper*, , (13-19).

Grimard, Franque. 1997. "Household Consumption Smoothing through Ethnic Ties: Evidence from Cote d'Ivoire." *Journal of Development Economics*, 53(2): 391–422.

Grossman, M. 1972. "On the Concept of Health Capital and the Demand for Health." *The Journal of Political Economy*, 80(2).

Gundimeda, Haripriya, and Gunnar Kohlin. 2008. "Fuel Demand Elasticities for Energy and Environmental Policies: Indian Sample Survey Evidence." *Energy Economics*, 30(2): 517–546.

Gupta, G., and G. Köhlin. 2006. "Preferences for Domestic Fuel: Analysis with Socio-Economic Factors and Rankings in Kolkata, India." *Ecological Economics*, 57(1): 107–121.

Haab, Timothy C., and Kenneth E. McConnell. 2002. *Valuing Environmental and Natural Resources: The Econometrics of Non-Market Valuation.* Northhampton, MA:Edward Elgar.

Haddad, Lawrence James, and SM Ravi Kanbur. 1990. "How Serious is the Neglect of Intra-Household Inequality?" *Economic Journal*, 100(402): 866–881. September.

HarvestPlus. 2014. "Iron." *Webpage*, `http://www.harvestplus.org/content/iron`.

Hassan, Rashid, Robert Scholes, and Neville Ash, ed. 2005. *Ecosystems and Human Well-Being: Current State and Trends.* Vol. 1, Washington, DC:Island Press.

Hawkes, C, and M T Ruel. 2012. "Value Chains for Nutrition." *Reshaping Agriculture for Nutrition and Health*, , ed. Shenggen Fan and Rajul Pandya-Lorch, Chapter 9, 73–81. Washington, DC: International Food Policy Research Institute.

Heckman, James J. 1979. "Sample Selection Basis as a Specification Error." *Econometrica: Journal of the Econometric Society*, 47: 153–161.

Heltberg, R., T.C. Arndt, and N.U. Sekhar. 2000. "Fuelwood Consumption and Forest Degradation: A Household Model for Domestic Energy Substitution in Rural India." *Land Economics*, 213–232.

Hicks, John R. 1946. *Value and Capital: An Inquiry Into Some Fundamental Principles of Economic Theory.* Vol. 2, Oxford, UK:Clarendon Press.

Hyde, W.F., G. Kohlin, and G.S. Amacher. 2000. "Social Forestry Reconsidered." *Silva Fennica*, 34(3): 285–314.

Jacoby, H.G. 1993. "Shadow Wages and Peasant Family Labour Supply: An Econometric Application to the Peruvian Sierra." *The Review of Economic Studies*, 60(4): 903–921.

Key, Nigel, Elisabeth Sadoulet, and Alain de Janvry. 2000. "Transaction Costs and Agricultural Household Supply Response." *American Journal of Agricultural Economics*, 82(2): 245–259.

Khundi, Fydess, Pamela Jagger, Gerald Shively, and Dick Sserunkuuma. 2011. "Income, Poverty, and Charcoal Production in Uganda." *Forest Policy and Economics*, 13(3): 199–205.

Klugman, Jeni. 2011. "Sustainability and Equity: A Better Future for All." United Nationals Development Programme Human Development Report, New York, NY.

Kohlin, Gunnar, and Gregory S. Amacher. 2005. "Welfare Implications of Community Forest Plantations in Developing Countries: The Orissa Social Foresty Project." *American Journal of Agricultural Economics*, 87(4): 855–869.

Kruger, Diana I. 2007. "Coffee Production Effects on Child Labor and Schooling in Rural Brazil." *Journal of Development Economics*, 82(2): 448–463. March.

Kumar, S., and D. Hotchkiss. 1988. *Consequences of Deforestation for Women's Time Allocation, Agricultural Production, and Nutrition in Hill Areas of Nepal. Research Report 69*, Washington, DC:International Food Policy Research Institute.

Le, Kien T. 2010. "Separation Hypothesis Tests in the Agricultural Household Model." *American Journal of Agricultural Economics*, 92(5): 1420–1431.

Lewbel, Arthur. 1996. "Aggregation Without Separability: A Generalized Composite Commodity Theorem." *The American Economic Review*, 86(3): 524–543.

Liversage, Harold. 2004. "Rangeland Management and Land Issues." International Fund for Agricultural Development Working Paper 3.

MacDonald, D., W. Adamowicz, and M. Luckert. 2001. "Fuelwood Collection in Northeastern Zimbabwe: Valuation and Calorific Expenditures." *Journal of Forest Economics*, 7(1).

Malik, R. S., and S. K. Sharma. 1990. "Moisture Extracation and Crop YIeld as a Function of Distance from a Row of Eucalyptus Tereticornis." *Agroforestry Systems*, 12(2): 187–195.

Millennium Ecosystem Assessment. 2005. *Ecosystems and Human Well-Being: Current State and Trends.* Vol. 1, Washington, D.C:Island Press.

Mushi, Deogratias. 2012. "Tanzania: How Biomass Use Fuels Rapid Deforestation." *Tanzania Daily News*, May 18.

Mwampamba, Tuyeni Heita. 2007. "Has the Woodfuel Crisis Returned? Urban Charcoal Consumption in Tanzania and its Implications to Present and Future Forest Availability." *Energy Policy*, 35: 4221–4234. May.

Myers, Samuel S, Antonella Zanobetti, Itai Kloog, Peter Huybers, Andrew D B Leakey, Arnold J Bloom, Eli Carlisle, Lee H Dietterich, Glenn Fitzgerald, Toshihiro Hasegawa, N Michele Holbrook, Randall L Nelson, Michael J Ottman, Victor Raboy, Hidemitsu Sakai, Karla A Sartor, Joel Schwartz, Saman Seneweera, Michael Tausz, and Yasuhiro Usui. 2014. "Increasing CO_2 Threatens Human Nutrition." *Nature*, 1–11. May.

Nature Conservancy. 2013. "Creating Incentives to Stop Deforestation." *Webpage*, `http://www.nature.org/ourinitiatives/urgentissues/global-warming-climate-change/how-we-work/creating-incentives-to-stop-deforestation.xml`.

Nichols, Austin. 2010. "Regression for Nonnegative Skewed Dependent Variables." Stata Conference. Boston, MA.

Omamo, Steven Were. 1998. "Transport Costs and Smallholder Cropping Choices: An Application to Siaya District, Kenya." *American Journal of Agricultural Economics*, 80(1): 116–123.

Outes-Leon, Ingo, and Stefan Dercon. 2008. "Survey Attrition and Attrition Bias in Young Lives." *Young Lives Technical Note*, 5.

Palmer, C., and J. MacGregor. 2009. "Fuelwood Scarcity, Energy Substitution, and Rural Livelihoods in Namibia." *Environment and Development Economics*, 14(6): 693–715.

Pattanayak, S.K., E.O. Sills, R.A. Kramer, et al. 2004. "Seeing the Forest for the Fuel." *Environment and Development Economics*, 9(2): 155–179.

Perlin, John. 2005. *A Forest Journey: The Story of Wood and Civilization.* Woodstock, VT:The Countryman Press.

Phaneuf, Daniel J, and V Kerry Smith. 2005. "Recreation Demand Models." *Handbook of Environmental Economics*, 2: 671–761.

Pitt, Mark M. 1985. "Equity, Externalities, and Energy Subsidies: The Case of Kerosense in Indonesia." *Journal of Development Economics*, 17(3): 201–217.

Pitt, Mark M, and Mark R Rosenzweig. 1985. "Health and Nutrient Consumption Across and Within Farm Households." *The Review of Economics and Statistics*, 67(2).

Powell, James L. 1984. "Least Absolute Deviations Estimation for the Censored Regression Model." *Journal of Econometrics*, 25(3): 303–325.

Psacharopoulos, George, and Harry Anthony Patrinos. 2004. "Returns to Investment in Education: A Further Update." *Education Economics*, 12(2): 111–134.

Rurai, Masageri Tumbuya. 2014. June 4, WMA Technical Advisor, Frankfurt Zoological Society. Tanzania.

Sen, Amartya K. 1966. "Peasant and Dualism with or without Surplus Labor." *Journal of Political Economy*, 74(5): 425–450.

Shively, Gerald E., and Monica Fisher. 2004. “Smallholder Labor and Deforestation: A Systems Approach.” *American Journal of Agricultural Economics*, 86(5): 1361–1366.

Singh, Inderjit, Lyn Squire, and John Strauss, ed. 1986. *Agricultural Household Models: Extensions, Applications, and Policy.* Baltimore, MD:Johns Hopkins University Press.

Skoufias, Emmanuel. 1994. “Using Shadow Wages to Estimate Labor Supply of Agricultural Households.” *American Journal of Agricultural Economics*, 76(2): 215–227.

Stocker, Thomas F., Dahe Qin, Gian-Kasper Plattner, Melinda M.B. Tignor, Simon K. Allen, Judith Boschung, Alexander Nauels, Yu Xia, Vincent Bex, and Pauline M. Midgley, ed. 2013. *Climate Change 2013: The Physical Science Basis.* Cambridge, UK:Cambridge University Press.

Strauss, John. 1986. “Does Better Nutrition Raise Farm Productivity?” *The Journal of Political Economy*, 94(2): 297–320.

Strauss, John, and Duncan Thomas. 1998. “Health, Nutrition, and Economic Development.” *Journal of economic literature*, 766–817.

Subramanian, Shankar, and Angus Deaton. 1996. “The Demand for Food and Calories.” *Journal of Political Economy*, 104(1): 133–162.

Taylor, J. Edward, and Irma Adelman. 2003. “Agricultural Household Models: Genesis, Evolution, and Extensions.” *Review of Economics of the Household*, 1(1–2): 33–58. January.

Tekgüç, Hasan. 2012. “Separability Between Own Food Production and Consumption in Turkey.” *Review of Economics and the Household*, 10(3): 423–439. September.

The Economist. 2014. “Cool Developments.” *The Economist*, May 31.

UN–REDD Programme. 2009. “About REDD+.” `http://www.un-redd.org/AboutREDD/tabid/102614/Default.aspx`.

UN–REDD Programme. 2012. “Estimating the Costs of REDD+ in Tanzania.” *Policy Brief*, by Tom Blomley and Timm Tennigkeit. April 2012.

U.S. Government’s Global Hunger and Food Security Initiative. 2014. “Feed the Future.” *Webpage*, `http://www.feedthefuture.gov/country/tanzania`.

Von Braun, J, and E Kennedy. 1994. *Agricultural Commercialization, Economic Development, and Nutrition.* Johns Hopkins University Press.

Weerdt, Joachim De, Kathleen Beegle, Helene Bie Lilleør, Stefan Dercon, Kalle Hirvonen, Martina Kirchberger, and Sofya Krutikova. 2012. “Kagera Health and Development Survey 2010: Basic Information Document.” The Rockwool Foundation Research Unit Study Paper 46, Copenhagen.

Welch, Ross M, and Robin D Graham. 1999. “A New Paradigm for World Agriculture: Meeting Human Needs: Productive, Sustainable, Nutritious.” *Field Crops Research*, 60(1): 1–10.

WHO, and FAO. 2004. *Vitamin and Mineral Requirements.* . 2nd ed., Geneva, Switzerland:World Health Organization.

Willig, Robert D. 1976. "Consumer's Surplus Without Apology." *The American Economic Review*, 66(4): 589–597.

Wooldridge, Jeffrey M. 2010. *Econometric Analysis of Cross Section and Panel Data.* . Second ed., Cambridge, MA:The MIT Press.

World Bank. 2010. *World Development Indicators.* Washington, DC:The World Bank.

Yen, Steven T., and Wiktor L. Adamowicz. 1993. "Statistical Properties of Welfare Measures from Count-Data Models of Recreation Demand." *Review of Agricultural Economics*, 15(2): 203–215.

Zellner, Arnold. 1962. "An Efficient Method of Estimating Seemingly Unrelated Regressions and Tests for Aggregation Bias." *Journal of the American Statistical Association*, 57(298): 348–368.

Appendix A

Perfect Labor Market Separability Test

In the case of perfect labor markets in Subsection 2.4.1 observed market wages represent for both household member types' true opportunity costs of time. Following Jacoby (1993) I test for perfect labor markets by running the regression:

$$shadow\ wage_{j,kvt} = a_0 + a_1 sample\ wage_{j,vt} + e_{kv} + e_t$$

where e_{kv} is a household fixed effect and e_t is an independent and identically distributed error term. The null hypothesis of efficient labor markets is the joint test that $a_0 = 0$ and $a_1 = 1$. I run this regression using OLS fixed effects and then run a joint Wald test to test the null hypothesis. The test is run to test separability both for adults and teenagers. Test results are displayed below in Table A.1. The perfect labor market scenario is soundly rejected both in the estimates for adult sand teenagers providing more evidence that the estimated shadow wage is the appropriate opportunity cost of time measure.

Table A.1: Perfect Labor Market Test: OLS with Household Fixed Effects
Test for equality of observed sample wage and estimated shadow wage
Dependent variable: Shadow wage estimates

	Adults	Teenagers
Sample wage	−0.252***	−0.052
	(0.082)	(0.060)
Constant	22.509***	8.666***
	(1.721)	(1.042)
Observations	3337	3337
F-test	171.00	69.11
Two-sided p-value	0.00	0.00

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Cluster robust standard errors in parentheses.
Note: Estimates are from an ordinary least squares model with household fixed effects.

Appendix B

Becketti et al. (1988) Pooling Tests

Table B.1: Becketti et al. (1988) Pooling Test (2004)

Dependent variable: Number of hours spent in school last week

	Poisson	Tobit	OLS (log)
Age (years)	1.13***	26.71***	1.04***
	(0.07)	(1.62)	(0.09)
Age squared (years)	−0.04***	−1.03***	−0.04***
	(0.00)	(0.07)	(0.00)
Gender (1=female)	−0.06**	−2.19***	−0.11**
	(0.03)	(0.84)	(0.05)
Firewood collection trip time (hours)	−0.05***	−1.25***	−0.05**
	(0.02)	(0.44)	(0.02)
Mother: some primary education	−0.10*	−2.51	−0.13
	(0.06)	(1.55)	(0.12)
Mother: completed primary education or above	−0.05	−1.83	−0.09
	(0.04)	(1.21)	(0.08)
Father: some primary education	−0.21***	−4.99***	−0.22*
	(0.06)	(1.56)	(0.11)
Father: completed primary education or above	0.01	0.35	0.02
	(0.04)	(1.21)	(0.09)
Gender of household head (1=female)	0.00	0.14	0.02
	(0.04)	(1.16)	(0.09)
Household size	0.00	0.19	0.01
	(0.01)	(0.16)	(0.01)
Annual household expenditures (TSh, log)	0.13***	3.52***	0.15***
	(0.02)	(0.71)	(0.05)

Continued on next page

Table B.1 – continued from previous page

	Poisson	Tobit	OLS (log)
Interviewed in February (1=yes)	0.22***	4.59***	0.16*
	(0.05)	(1.61)	(0.09)
Interviewed in June (1=yes)	−0.41***	−11.05***	−0.44**
	(0.07)	(1.98)	(0.22)
Interviewed in July (1=yes)	−0.48***	−10.98***	−0.47**
	(0.08)	(1.95)	(0.19)
Interviewed in August (1=yes)	0.24***	6.97***	0.27***
	(0.06)	(1.74)	(0.08)
Interviewed in December (1=yes)	−1.24***	−30.46***	−1.10***
	(0.12)	(2.18)	(0.18)
Individual attrited (1=yes)	−1.34	−26.58	−0.50
	(1.08)	(25.27)	(1.03)
Age x Attrition	0.08	0.53	−0.07
	(0.17)	(3.70)	(0.17)
Age squared x Attrition	−0.01	−0.06	0.00
	(0.01)	(0.16)	(0.01)
Gender x Attrition	0.11	2.58	0.08
	(0.07)	(1.86)	(0.12)
Trip time x Attrition	0.05	1.27	0.05
	(0.04)	(0.93)	(0.05)
Mother: some primary education x Attrition	0.38***	10.84***	0.48***
	(0.11)	(2.92)	(0.17)
Mother: primary education or above x Attrition	0.20**	6.47***	0.30**
	(0.09)	(2.48)	(0.13)
Father: some primary education x Attrition	−0.19	−6.33*	−0.21
	(0.14)	(3.63)	(0.20)
Father: primary education or above x Attrition	−0.07	−2.02	−0.12
	(0.09)	(2.43)	(0.11)
Gender of household head x Attrition	0.05	1.43	0.05
	(0.09)	(2.37)	(0.12)
Household size x Attrition	0.04***	1.09***	0.04***
	(0.01)	(0.30)	(0.01)
Annual household expenditures x Attrition	0.04	0.98	0.04
	(0.04)	(1.26)	(0.05)
Interviewed in February x Attrition	0.12	5.15*	0.28**
	(0.10)	(3.09)	(0.11)
Interviewed in June x Attrition	−0.12	−1.28	−0.01
	(0.15)	(3.63)	(0.14)
Interviewed in July x Attrition	−0.58***	−8.88**	−0.20
	(0.21)	(4.05)	(0.16)
Interviewed in August x Attrition	−0.12	−2.12	−0.07

Continued on next page

Table B.1 – continued from previous page

	Poisson	Tobit	OLS (log)
	(0.13)	(3.46)	(0.14)
Interviewed in December x Attrition	0.00	0.92	0.17
	(0.25)	(4.52)	(0.14)
Village fixed-effects	Yes	Yes	Yes
Survey round fixed-effects	Yes	Yes	Yes
Observations	5,557	5,557	5,557
Log-likelihood	-55,179	-14,782	-8,740
F-test, p-value	0.000	0.000	0.001

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Cluster robust standard errors in parentheses.

Note: F-test tests for the joint significance of the attrition dummy and all 13 attrition interaction terms. All regressions include village and survey round fixed-effects.

Table B.2: Becketti et al. (1988) Pooling Test (2010)
Dependent variable: Number of hours spent in school last week

	Poisson	Tobit	OLS (log)
Age (years)	1.20***	28.30***	1.10***
	(0.07)	(1.64)	(0.08)
Age squared (years)	−0.05***	−1.10***	−0.04***
	(0.00)	(0.07)	(0.00)
Gender (1=female)	−0.07**	−2.43***	−0.13**
	(0.03)	(0.83)	(0.05)
Firewood collection trip time (hours)	−0.04**	−1.05**	−0.05**
	(0.02)	(0.44)	(0.02)
Mother: some primary education	−0.02	−0.61	−0.04
	(0.05)	(1.47)	(0.10)
Mother: completed primary education or above	−0.05	−1.66	−0.08
	(0.04)	(1.24)	(0.08)
Father: some primary education	−0.21***	−5.37***	−0.24**
	(0.06)	(1.62)	(0.12)
Father: completed primary education or above	−0.03	−0.85	−0.03
	(0.04)	(1.20)	(0.10)
Gender of household head (1=female)	0.01	0.32	0.02
	(0.04)	(1.14)	(0.09)
Household size	0.01	0.27*	0.01
	(0.01)	(0.15)	(0.01)
Annual household expenditures (TSh, log)	0.12***	3.43***	0.15***
	(0.02)	(0.69)	(0.05)
Interviewed in February (1=yes)	0.21***	4.45***	0.16*
	(0.06)	(1.62)	(0.08)
Interviewed in June (1=yes)	−0.42***	−11.12***	−0.43*
	(0.07)	(1.91)	(0.22)
Interviewed in July (1=yes)	−0.51***	−11.64***	−0.48**
	(0.08)	(1.93)	(0.19)
Interviewed in August (1=yes)	0.22***	6.43***	0.25***
	(0.06)	(1.74)	(0.08)
Interviewed in December (1=yes)	−1.26***	−31.08***	−1.10***
	(0.12)	(2.16)	(0.18)
Individual attrited (1=yes)	1.23	36.27	1.70*
	(1.05)	(25.70)	(1.00)
Age x Attrition	−0.31**	−8.42**	−0.36**
	(0.16)	(3.64)	(0.17)
Age squared x Attrition	0.01**	0.35**	0.01**
	(0.01)	(0.15)	(0.01)
Gender x Attrition	0.17**	3.83*	0.18
	(0.07)	(1.97)	(0.11)

Continued on next page

Table B.2 – continued from previous page

	Poisson	Tobit	OLS (log)
Trip time x Attrition	0.00	0.05	0.01
	(0.04)	(0.93)	(0.05)
Mother: some primary education x Attrition	0.05	2.07	0.09
	(0.12)	(3.14)	(0.18)
Mother: primary education or above x Attrition	0.13	4.37*	0.23*
	(0.09)	(2.40)	(0.13)
Father: some primary education x Attrition	−0.19	−3.59	−0.08
	(0.14)	(3.46)	(0.17)
Father: primary education or above x Attrition	0.12	2.98	0.07
	(0.09)	(2.52)	(0.14)
Gender of household head x Attrition	0.03	0.64	0.01
	(0.09)	(2.49)	(0.16)
Household size x Attrition	0.04***	1.04***	0.04*
	(0.01)	(0.33)	(0.02)
Annual household expenditures x Attrition	−0.00	−0.25	−0.02
	(0.05)	(1.36)	(0.06)
Interviewed in February x Attrition	0.18*	5.36*	0.23**
	(0.10)	(3.08)	(0.10)
Interviewed in June x Attrition	−0.15	−1.95	−0.09
	(0.18)	(4.12)	(0.16)
Interviewed in July x Attrition	−0.38*	−5.59	−0.16
	(0.22)	(4.22)	(0.14)
Interviewed in August x Attrition	−0.01	0.87	0.04
	(0.13)	(3.57)	(0.16)
Interviewed in December x Attrition	0.10	3.73	0.22*
	(0.25)	(4.52)	(0.12)
Village fixed-effects	Yes	Yes	Yes
Survey round fixed-effects	Yes	Yes	Yes
Observations	5,557	5,557	5,557
Log-likelihood	-55,366	-14,796	-8,754
F-test, p-value	0.000	0.000	0.015

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Cluster robust standard errors in parentheses.

Note: F-test tests for the joint significance of the attrition dummy and all 13 attrition interaction terms. All regressions include village and survey round fixed-effects.

Appendix C

Household Food Demand with Hired Labor

If households are buying labor then they are not selling labor, $\sum_{j=z,a} L_j^h > 0$ and $L^m = 0$. In this case, agricultural labor is made up of both household labor and hired labor. The associated Lagrangian is:

$$\mathcal{L} = u(z,c) - \lambda\Big(p_z z + p_c c - p_z F\big(\alpha_z(L_z^h + L_z^f), \beta_z E_z, \delta_z f_z\big) - p_a F\big(\alpha_a(L_a^h + L_a^f), \beta_a E_a, \delta_a f_a\big)$$
$$+ \sum_{j=z,a} \big(wL_j^h + p_f f_j\big) - y\Big) - \phi(\bar{E} - E_z - E_a) - \eta(\bar{L} - L_z^f - L_a^f - L^m)$$

This problem now has thirteen equations (including the budget constraint, land, constraint, and home labor constraint) and thirteen variables (including the three Lagrange multipliers λ, ϕ, and η), and so the first order conditions can be used to solve for the optimal levels of household food consumption and agricultural production:

$$
\begin{aligned}
\frac{\partial u}{\partial i} &= \lambda p_i \text{ for } i = z, c \\
p_j \alpha_j \frac{\partial F}{\partial L_j} &= -\frac{\eta}{\lambda} \text{ for } j = z, a \\
p_j \alpha_j \frac{\partial F}{\partial L_j} &= w \text{ for } j = z, a \\
p_j \beta_j \frac{\partial F}{\partial E_j} &= -\frac{\phi}{\lambda} \text{ for } j = z, a \\
p_j \delta_j \frac{\partial F}{\partial f_j} &= p_f \text{ for } j = z, a.
\end{aligned}
$$

Similar to the case of household's selling labor, household consumption decisions are again determined by the prices of consumed goods up until the point where their marginal utilities are equal:

$$
\frac{1}{p_c}\frac{\partial u}{\partial c} = \frac{1}{p_z}\frac{\partial u}{\partial z}.
$$

Land and labor are allocated between household-produced food and cash crop production up until the point where the value of their marginal products are equal:

$$
p_z \alpha_z \frac{\partial F}{\partial L_z} = p_a \alpha_a \frac{\partial F}{\partial L_a},\ p_z \beta_z \frac{\partial F}{\partial E_z} = p_a \beta_a \frac{\partial F}{\partial E_a},\text{ and } p_z \delta_z \frac{\partial F}{\partial f_z} = p_a \delta_a \frac{\partial F}{\partial f_a}
$$

These first order conditions are identical to those given in equations (4.6) and (4.7) and imply that household consumption and production decisions can be made sequentially. Total household food consumption is still given by:

$$
z^* = z(p_z, p_c, \Pi^* + y)
$$

where now $y + \Pi^* = p_z F(\alpha_z L_z^{h*} + L_z^{f*}, \beta_z E_z^*, \delta_z f_z^*) + p_a F(\alpha_a L_a^{h*} + L_a^{f*}, \beta_a E_a^*, \delta_a f_a^*) - w(L_z^{h*} + L_a^{h*}) - p_f(f_z^* + f_a^*)$.

In the case of food market failures with hired labor, the associated Lagrangian is:

$$\mathcal{L} = u(z,c) - \lambda\Big(p_c c - p_a F\big(\alpha_a(L_a^h + L_a^f), \beta_a E_a, \delta_a f_a\big) + \sum_{j=z,a}\big(wL_j^h + p_f f_j\big) - y\Big)$$
$$- \phi(\bar{E} - E_z - E_a) - \eta(\bar{L} - L_z^f - L_a^f - L^m) - \mu\Big(z - F\big(\alpha_z(L_z^h + L_z^f), \beta_z E_z, \delta_z f_z\big)\Big).$$

And the first order conditions governing household food consumption and agricultural production decisions are:

$$\begin{aligned}
\frac{\partial u}{\partial z} &= \mu & \frac{\partial u}{\partial c} &= \lambda p_c \\
\alpha_z \frac{\partial F}{\partial L_z} &= \frac{\lambda}{\mu} w & \alpha_z \frac{\partial F}{\partial L_z} &= -\frac{\eta}{\mu} \\
p_a \alpha_a \frac{\partial F}{\partial L_a} &= w & p_a \alpha_a \frac{\partial F}{\partial L_a} &= -\frac{\eta}{\lambda} \\
\beta_z \frac{\partial F}{\partial E_z} &= -\frac{\phi}{\mu} & p_a \beta_a \frac{\partial F}{\partial E_a} &= -\frac{\phi}{\lambda} \\
\delta_z \frac{\partial F}{\partial f_z} &= \frac{\lambda}{\mu} p_f & p_a \delta_a \frac{\partial F}{\partial f_z} &= p_f.
\end{aligned}$$

These first order conditions produce the same relationship governing household food consumption decisions as displayed in equation (4.17):

$$\frac{\partial u}{\partial z} = \frac{p_a}{p_c} \frac{\partial u}{\partial c} \frac{\beta_a}{\beta_z} \frac{\partial F / \partial E_a}{\partial F / \partial E_z}.$$

Thus, the result that the price of cash crops only directly affects household food consumption decisions in the case of food market imperfections does not depend on whether or not households are buyers or sellers of labor.

Appendix D

GMM 3SLS: First Stage Results

Table D.1: Household Expenditures Instrument: First Stage Results
Dependent variable: Annual household expenditures, TSh (log)

	Household expenditure (log)
Number of adult males (16 and older)	0.046**
	(0.022)
Number of adult females (16 and older)	0.035*
	(0.021)
Number of boys (15 and younger)	−0.062***
	(0.017)
Number of girls (15 and younger)	−0.074***
	(0.019)
Agricultural land (acres)	−0.000
	(0.002)
Number of livestock owned	0.000
	(0.002)
Home ownership (1 if own, 0 otherwise)	0.053*
	(0.031)
Water source (0 if river, lake, or rainwater, 1 otherwise)	0.021
	(0.019)
Light source (1 electric, 0 otherwise)	0.209***
	(0.061)
Cooking fuel (0 if firewood, 1 otherwise)	0.305***
	(0.044)
Sewage disposal (0 if none, 1 otherwise)	0.029
	(0.027)
Number of mobile phones owned	0.200***
	(0.025)

Continued on next page

Table D.1 – continued from previous page

Number of televisions owned	0.112**
	(0.054)
Household size (calorie adjusted)	0.150***
	(0.024)
Children in household (fraction)	0.682***
	(0.138)
Household head (1 if female)	−0.130***
	(0.027)
Urban household (1 if urban)	0.110***
	(0.036)
Household distance to nearest trunk road	0.001
	(0.001)
Household distance to nearest town ($>$ 20,000)	−0.001
	(0.000)
Price of staples (TSh/kcal)	0.297
	(0.401)
Price of pulses (TSh/kcal)	0.326**
	(0.144)
Price of fruits/vegetables (TSh/kcal)	−0.039
	(0.026)
Price of animal products (TSh/kcal)	−0.125***
	(0.045)
Price of meal complements (TSh/kcal)	−0.042
	(0.114)
Cash crops price index (TSh/g)	0.073***
	(0.026)
Non-food market goods price index (TSh/g)	0.151
	(0.260)
Daily per capita staple consumption (kcal)	0.000***
	(0.000)
Daily per capita pulse consumption (kcal)	0.000***
	(0.000)
Daily per capita vegetable & fruit consumption (kcal)	0.000
	(0.001)
Daily per capita animal product consumption (kcal)	0.000
	(0.000)
Daily per capita meal complement consumption (kcal)	0.000
	(0.000)
Observations	2,337

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Block bootstrapped standard errors.

Note: Regression also contains month of interview dummies. Coefficients not reported.