AGRICULTURAL FINANCE, NON-FARM EMPLOYMENT, AND RURAL POVERTY: EVIDENCE FROM SUB-SAHARAN AFRICA

By

Serge Guigonan Adjognon

A DISSERTATION

Submitted to
Michigan State University
in partial fulfillment of the requirements
for the degree of

Agricultural, Food, and Resource Economics – Doctor of Philosophy

2016

ABSTRACT

AGRICULTURAL FINANCE, NON-FARM EMPLOYMENT, AND RURAL POVERTY: EVIDENCE FROM SUB-SAHARAN AFRICA

By

Serge Guigonan Adjognon

Efforts to eradicate poverty in the world require a particular focus on agricultural households of SSA, where poverty remains ubiquitous. This dissertation, titled *Agricultural Finance, Non-Farm Employment, and Rural Poverty,* uses evidence from SSA to explore some of the constraints faced by farming households in Sub-Saharan Africa (SSA). First, while agriculture remains central in the economy of most SSA, yields are still relatively low compared to other parts of the world. Financial restrictions are cited amongst the main constraints to inputs use in SSA. Thus, the first essay of this dissertation titled *Updating the Landscape on Farm Input Credit in SSA* explores input financing and the role of credit therein. Our results consistently show that traditional credit use is extremely low (across credit type, country, crop and farm size categories) and farmers primarily finance modern input purchases with cash from nonfarm activities and crop sales. Second, the consistent lack of credit for agricultural inputs observed in the first essay, motivated the second essay titled *Sustaining Input on Credit through Dynamic Incentives and Information Sharing*. This essay uses a framed field experiment to explore conditions that can minimize strategic default, a key source of market failure in input credit markets in developing countries, where institutions for contract enforcement are weak or nonexistent. The results show that the existence of an information exchange system, amongst input sellers, which mimics the role of a "credit score" (with potential benefits from its informal nature), can effectively deter default behavior by farmers receiving inputs on credit. Moreover, productivity shocks that affect the return to the use of inputs also affect the opportunity cost of

repayment, and thus farmer's decision to repay. Third, the importance of non-farm activities (revealed in Essay 1), in addition to recent evidence from the literature, indicate increasing contribution of non-farm activities to households' income in SSA. Therefore, the third essay of this dissertation titled the *Heterogeneous Welfare Effects of Non-Farm Employment* explores the effects of rural non-farm activities (wage and self-employment) on rural household welfare. It also explores the heterogeneous effects of participating in non-farm activities across the welfare distribution. The results confirm that participation in non-farm activities is generally welfare improving and poverty reducing. However, households at the lower tail of the welfare distribution benefit significantly less from participation than the wealthiest. Low education, assets, and access to credit are important barriers that limit the participation of the poorest in lucrative non-farm employment opportunities.

Together, these essays shed light on important policy considerations for improving the livelihood of poor households in developing countries. While access to credit remains extremely limited for farming households to finance agricultural intensification, the expansion of Information and Communication Technologies in SSA offers new hope for financial inclusion of those marginalized groups. Similarly, improving access to the non-farm sector by households will likely improve modern input use and agricultural productivity. Beyond just the effect on inputs use, participation in the non-farm sector significantly improves welfare and reduces poverty amongst rural households. However, it is important to address barriers that limit participation in more lucrative non-farm sectors by the poorest, who currently benefit less from participation compared to the non-poor.

I dedicate this dissertation to my mother, Elisabeth Katary, and my late father, Innocent Adjognon. For all the sacrifice you have made to provide for your children, may this accomplishment be an expression of my eternal gratitude. Your boundless support, encouragement, and prayers have kept me going. Merci beaucoup.
Je vous aime du fond de mon coeur.

ACKNOWLEDGEMENTS

I would like to express my profound gratitude for the incredible mentorship, guidance, inspiration, and motivation I received from my supervisor and major advisor Prof. Lenis Saweda Liverpool-Tasie who also funded my PhD study at Michigan State. Also, sincere thanks to the rest of my advisory committee: Prof. Mywish Maredia, Prof. Robert Shupp, Prof. Robert Myers, and Prof. Jeffrey Wooldridge, for precious comments on my research, teachings in class, as well as technical and personal advice throughout my PhD career.

I also extend my gratitude to the faculty and staff of the Department of Agricultural, Food and Resource Economics (AFRE), Michigan State University, especially the Graduate Secretary Ms. Debbie Conway, and the Graduate Chair Prof. Scott Swinton for consistent support and assistance while navigating the PhD study.

I am forever grateful for the support I received from my fellow student cohorts Jason Snyder, Joey Goeb, Sarah Kopper, Andrea Leschewski, Mukesh Ray, Piyayut Chitchimnong, Craig Carpenter, among others. I truly value the common experience and all the memories that we share.

To all my East Lansing friends, many thanks for having supported me and given me the opportunity to keep a balanced lifestyle throughout my PhD program. Special thanks to my best friend, and soon-to-be-wife, Marie Steele, who has been truly supportive through this whole journey.

Last but not least, I would like to recognize my family who encouraged me to accept my offer of admission into the PhD program and supported me throughout the whole program. I

want to thank especially my mother Elisabeth Katary for her countless hours of prayers for my success, and for all the sacrifices made to ensure my education since my young age. I thank and pray for my late dad who instilled in me the values of education and good work ethic since my childhood and encouraged me to work hard and follow fearlessly my dreams in life. Thank you very much for all the sacrifices.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

KEY TO ABBREVIATIONS

SSA	Sub-Saharan Africa
FE	Fixed effects
FRM	Fractional response Model
ICT	Information and Communication Technology
LSMS	Living Standards Measurement Surveys
MKW	Malawi Kwacha
SD	Standard Deviation
SE	Standard Error

1. INTRODUCTION

Poverty continues to affect millions of people globally, the majority of whom live in rural areas of Sub-Saharan Africa (SSA). The recent World Development Indicators[1] published by the World Bank indicate that the share of total world poor living in SSA has practically tripled from 14.7 percent in 1990 to 43.4 percent in 2012. These households in poverty are exposed to a host of interrelated socio-economic constraints including hunger and malnutrition, low education, low assets, low levels of infrastructure, amongst others; which limit their productivity and trap them into a vicious cycle of poverty. This implies that efforts to eradicate poverty in the world require a particular focus on rural households of SSA, with interventions aiming to improve their productivity and income from the activities they are involved in. Consequently, this dissertation titled *Agricultural Finance, Non-Farm Employment, and Rural Poverty* uses evidence from SSA to explore some of the constraints faced by farming households in SSA[2].

Agriculture is the main activity for most people in rural areas of SSA, yet agricultural yields in SSA are considerably lower than in other parts of the world. There is widespread agreement that this is partly due to the significantly lower use of modern inputs in SSA compared to the rest of the world. Financial constraints have been mentioned amongst the main demand side barriers to agricultural input use in developing countries, including SSA. Therefore, improving households' access to financial solutions for inputs purchases is an important requirement for agricultural development and poverty reduction. While recent evidence shows that many Sub-Saharan African farmers use modern inputs, there is limited current and

[1] http://data.worldbank.org/data-catalog/world-development-indicators

[2] Though most farm activities occur in rural areas, there are some farm households in urban areas. Therefore farm households, in this essay, include households involved in agricultural activities in both urban and rural areas, unless otherwise specified.

comparable information on how these input purchases are financed. The first essay of this dissertation titled *Updating the Landscape on Farm Input Credit in SSA* uses recently available nationally representative data from four countries to explore input financing and the role of credit therein. Our results consistently show that traditional credit use is extremely low (across credit type, country, crop and farm size categories) and farmers primarily finance modern input purchases with cash from nonfarm activities and crop sales. Tied output-factor market arrangements (largely ignored in the literature) appear to be the only form of credit relatively widely used but it is mostly for labor and not for external inputs. These results motivated the second and third essays of this dissertation.

The rational for the second essay comes from the consistent lack of credit for agricultural inputs observed in the first essay. Generally, rural credit markets in developing countries are characterized by market failures associated with imperfect information and risk. These failures persist due to weak contract enforcement institutions, thus increasing the potential for high strategic default rates. Knowing this, input suppliers are reluctant to provide inputs to farmers on credit. Therefore, the second essay of this dissertation, titled *Sustaining Input on Credit through Dynamic Incentives and Information Sharing*, uses a framed field experiment to explore conditions that can minimize default by farmers receiving input on credit from input sellers in developing countries where institutions for contract enforcement are weak or nonexistent. Using data collected through a framed field experiment that simulates a market for input on credit, the paper shows that the existence of an information exchange system, amongst input sellers, which mimics the role of a "credit score" (with potential benefits from its informal nature), can effectively deter default behavior by farmers receiving inputs on credit. Moreover, productivity shocks that affect the return to the use of inputs also affect the opportunity cost of repayment, and thus farmer's decision to repay.

The findings in Essay 1 also indicate that non-farm activities play an important role in agricultural finance in SSA. Recent evidence from the literature also shows that the contribution of non-farm activities (such as non-farm wage employment and non-farm enterprises) to household income in SSA is substantial and has increased over time. However, there are considerable entry barriers that restrict the poorest from participating in such activities, or limit them to low returns categories of employment with low potential for lifting them out if poverty. Therefore, the third essay of this dissertation titled the *Heterogeneous Welfare Effects of Non-Farm Employment* uses nationally representative panel data from Malawi and a combination of econometric approaches, to quantify the effects of rural non-farm activities (wage and self-employment) on rural household welfare. It also explores the heterogeneous effects of participating in non-farm activities across the welfare distribution. The results confirm that participation in non-farm activities is generally welfare improving and poverty reducing. However, households at the lower tail of the welfare distribution benefit significantly less from participation than the wealthiest. Low education, assets, and access to credit are important barriers that limit the participation of the poorest in lucrative non-farm employment opportunities.

Together, these essays shed light on important policy considerations for improving the livelihood of poor households in developing countries. While access to credit remains extremely limited for farming households to finance agricultural intensification, the expansion of Information and Communication Technologies in SSA offers new hope for financial inclusion of those marginalized groups. Indeed, the establishment of information sharing systems amongst input sellers, suggested in Essay 2, to enforce credit repayment and improve access to input credit, can be made easier and less costly, by leveraging on the opportunities presented by ICT technologies. Even in areas with limited access to digital technologies, the remarkable penetration of cell phones can supplement local social networks and facilitate the collection and sharing of credit repayment behavior, especially if community level retailers are integrated in the distribution of input credit. Meanwhile, farmers seem to rely increasingly on income from non-farm sources to address their farm inputs cash needs. This implies that, alongside strategies to provide better access to input credits, improving access to the non-farm sector by households will likely improve modern input use and agricultural productivity. More generally, the non-farm sector is shown in Essay 3 to significantly improve welfare and reduce poverty amongst rural households. This confirms that the benefits of developing the non-farm sector extend beyond just the financing of input purchases. However, Essay 3 also reveals an unequal distribution of welfare benefits from participation in non-farm employment which indicates that it is important to address barriers that limit participation in more lucrative non-farm sectors by the poorest, which currently limits their benefit from participation compared to the non-poor.

2. ESSAY 1: UPDATING THE LANDSCAPE ON FARM INPUT CREDIT IN SUB-SAHARAN AFRICA

2.1. Introduction

It is generally accepted that Sub-Saharan Africa (SSA) farmers often suffer low yields, and if they bought more modern inputs (e.g. fertilizer, pesticides, and seeds) they could increase yields, all else equal. It is a common hypothesis in the literature (e.g., Croppenstadt et al. (2003)) that one reason farmers in SSA do not buy enough of these inputs is that they are credit-constrained.

Starting however from the knowledge that many SSA farmers in fact do purchase farm inputs (Sheahan and Barrett 2014), we pose a basic question concerning how they finance those inputs – that is, does credit play any role and if so which kinds of credit are important? Do the empirical facts from a systematic analysis of farm household surveys match with conventional wisdom about these issues? The three research questions we address are thus: (1) how do farmers finance input purchases? (2) Are there correlations with farm size and thus "inclusiveness" of the financial arrangement used and (3) is there a relation with crop type and thus relation to cash crop versus food crops?

We derive the hypotheses to test from the literature, which we consider to be feeding and reflecting common wisdom about these questions. The key points are organized according to the set of potential finance arrangements available to smallholders. These include formal and informal tied and untied credit sources and then own retained earnings.

First, in principal, formal credit from formal-sector banks could be a source of finance for inputs. But this source is in general depicted in the literature to be scant in SSA rural areas for two reasons. On one hand, parastatal agrarian banks and government credit for farmers input purchases have been largely dismantled during Structural Adjustment programs in the past several decades (Kherallah et al. 2002). On the other hand, it is generally held that private-sector banks lend little (and then mainly to larger farmers) to nothing to farmers for inputs or otherwise (Poulton et al., 2006). The reasons are diverse: as Dorward et al. (2009: 11) state, “Credit markets fail because of the inability to insure borrowers, lack of collateral, difficulties of recovering loans, and limited diversification of local economies, all of which impede the development of a sustainable model of rural financial services.” This evaluation of formal credit markets is echoed in other developing regions, such as South Asia (Binswanger and Khandker, 1995 and more generally (Besley, 1994, Conning and Udry, 2007). Our hypothesis to test is thus that farmers source little of their finance from banks, and only large farmers would get this.

Second, informal credit from friends and family and local moneylenders appears to be in general deemed a major source of funds for farmers that do buy inputs and consumption items (Poulton et al., 2006, Binswanger and Khandker, 1995). Sometimes in the literature noting that farmers usually cannot get bank loans, it is implicit or explicitly noted that they must consequently resort to these local informal loans, or to trader credit noted below. Our hypothesis is thus that informal credit are important to all strata of farmers and all kinds of crop farmers would get them.

Third, (informal sector) trader credit in “tied output-credit” arrangements are deemed to be important and widespread (Bardhan, 1980, Dorward et al., 2009, Chao-Béroff, 2014). These can be crop traders or input traders. Traders and farmers enter these arrangements because local formal credit markets idiosyncratically fail for them, so in economics terms these are “second best” arrangements (Binswanger and Rosenzweig, 1986). We thus hypothesize that among farmers who buy inputs, obtaining these advances from output or input traders is important and not farm size or crop biased.

Fourth, another kind of tied output-factor market arrangement is via "labor-output" arrangements (Bardhan, 1984) where local farm workers advance labor in exchange for payment (typically in kind but can be in cash) at harvest. While this was researched in South Asia in the 1970s/1980s, it has not been examined empirically in SSA to our knowledge, and is an important gap in the literature we try to fill.

Fifth, credit advances from (formal sector) processors in contract farming schemes (such as cotton or tobacco) are deemed important for the minority of farmers who participate in these arrangements (Fafchamps, 1999; Tschirley 2009; Key and Runsten 1999; Swinnen and Maertens, 2014). We thus hypothesize that for that subset of crops and farmers this will be found to serve many farmers.

Sixth, household retained earnings are very rarely compared with credit sourcing by farmers. That seems to be because the credit literature focuses on patterns and determinants of sourcing credit rather than studying what are all the sources of funding inputs, per se. Non-credit related literature for example on rural nonfarm employment (RNFE) shows that source of cash to be the lead source in rural households in SSA (Haggblade et al. 2010). A few cases studies have compared RNFE with cash from credit and transfers and sales of crops and found the credit share (from any source) to be tiny, crops sales moderate, and RNFE as a cash source very important, predominant (for a case in the Sahel, see Reardon and Mercado-Peters 1994). We thus hypothesize that own cash sources will be important but possibly skewed toward larger farmers (due to skewed distribution of RNFE, see Reardon et al. 2000), with the smaller farmers relegated to funding inputs from the informal credit sources.

This paper undertakes what to our knowledge is a unique cross-country empirical examination over various types of credit and different crops. We analyze recent LSMS farm household data sets with about 10,000 households from Malawi, Nigeria, Tanzania, and Uganda. We examine the purchase of "external inputs" by which we mean non-labor variable inputs (fertilizer, pesticides, and seeds) and labor. We stratify by country, and also by farm size and crop type (primarily grains versus horticulture versus traditional "cash crops" like cotton. The determinants of fertilizer purchases are also explored in the specific case of Nigeria, to understand the importance of various sources of cash for agricultural inputs finances.

The paper proceeds as follows. The next section discusses the data sources. Then we present descriptive results testing the hypotheses, and in the section thereafter, our econometrics results. Finally, we conclude with food and agricultural policy implications.

2.2. Data

We use survey data on farm household use of inputs and cash and in-kind arrangements to pay for them. The analysis is done by crop, household, and plot. The data also have characteristics of the farm households such as nonfarm income and farm size. The data come from four Living Standard Measurement Study (LSMS) surveys. The country surveys differ somewhat in the specific questions they use to elicit information on the variables of interest. We treat the survey datasets as uniformly as possible to ensure that the information is comparable over the sets.

The details of the survey design and stratification methods are included in the documentation of the LSMS survey for each country. In general, the surveys used a two-stage sample design. In the first stage, enumeration areas were selected in each district of the country. Then, within each enumeration area a listing of households was conducted to provide the sample frame for the second stage selection of households. Then a random systematic sampling was used to select the households.

In all countries, we extend the analysis to both urban and rural areas but select only households doing any farming. This is because we are interested in how households invested in agriculture generally finance input purchases. Also in all countries, there are a very tiny proportion of farm households in urban areas, and the separate analysis for urban and rural areas does not add any particular insight. Besides, in order to explore input finance arrangement, we focus on input purchases instead of input use, as some of the input use, though a very small part, come from government subsidies or from friends and relatives. Below we summarize the country data sets.

First, we use the Malawi Integrated Household Panel Survey (IHPS) of 2012/2013, with 3219 households and 7705 plots. The data includes agriculture input credit use on fertilizer, pesticide and seed by season (rainy and dry). [3]The seed data identify the purchase, crop type, whether credit was used and the type of credit. The fertilizer and pesticide data are by type of fertilizer and pesticide. There are also data on the use of harvest (in kind) to reimburse for inputs purchased on credit. The dataset also has information on the use of loans for purchase of farm inputs (among other uses).

Second, we use the second wave of the Nigeria Living Standard Measurement Study – Integrated Survey on Agriculture (LSMS-ISA) Panel for 2012/2013, covering 3000 farm households and 5819 plots. The data include agriculture input use/purchases and credit information for seeds, pesticides/herbicides, and fertilizer. The seed and fertilizer data indicate the crop type of the plot. There are also crop-level data on household use of in-kind or cash payments from the harvest to pay for labor, seed, and fertilizer received on credit. There are also data on households receiving and use of loans.

---

[3] In Malawi we focus only on the rainy season information as agricultural activity is far less intense during dry season and inputs use and purchase decisions are considerably different.

Third, we use the United Republic of Tanzania National Panel Survey 2012/2013, covering 3047 households and 6165 plots. The data include (beside household and plot characteristics) input use by crop, use, purchase and finance method of organic fertilizer, inorganic fertilizer, and herbicides/pesticides. Credit use for these inputs was determined by reported purchase of input via credit. There are also data on loans and their uses (including for farming).

Fourth, we use the Uganda National Panel Survey 2010/2011 covering 2109 farm households and 6003 plots across first and second season[4]. The data show per plot and by crop, for two seasons, the use and source of purchase for organic fertilizer, inorganic fertilizer, and pesticides/herbicides. We combine both types of fertilizers, and aggregate all the information to the household level. The data also show reimbursement for non-household labor used on the household plots and then the laborer paid with part of the household's harvest. There is no loans (as contrasted with transaction related credit) section in the Uganda survey.

2.3. Farm input purchases – abstracting for the moment from input finance

In this section we examine farmers' purchase of what we call "external inputs," which are variable inputs apart from labor, and include inorganic fertilizer, seeds, and pesticides/herbicides. In this section we highlight the key patterns on how farmers financed these input purchases. In all the descriptive statistics we use sampling weights available in the dataset to account for the survey design and construct nationally representative statistics. The weight for each household is

[4] Contrary to Malawi, we use information from first and second seasons in Uganda, as farmers are active in agricultural activities in both seasons.

the inverse of the probability of being selected based on the sample frame structure discussed above.

Table 1 presents the shares of farmers purchasing what we call here “external inputs” (variable inputs apart from labor). For the overall share of households buying external inputs, there is a marked contrast between Nigeria and Malawi, with a high share of farmers buying external inputs (71 and 70% respectively), compared to Uganda and Tanzania (16% and 18% respectively). The Malawi-Nigeria results are at odds with the traditional notion that very few farmers in SSA use external inputs but consistent with recent literature (Sheahan and Barrett, 2014).

One might say that the Nigeria and Malawi results are driven by the fertilizer subsidy program. While this might be true in Malawi where about 60% of households receive subsidized fertilizer (IHS, 2013), this is not likely the case for Nigeria. In the Nigeria survey data, only about 5% of the households who purchase fertilizer bought it from government sources.[5] The great majority of the households’ external input purchases were with from the local market.

---

[5] Note that there is no explicit question in the Nigeria LSMS that captures if a household participated in the government fertilizer subsidy program. However, until recently, subsidized fertilizer was distributed by the government. Since the government is not typically involved in the sale of non-subsidized fertilizer, we use the source of fertilizer purchase being government as a proxy for receiving subsidized fertilizer (Takeshima and Nkonya, 2014, Takeshima and Liverpool-Tasie, 2015). While this might be an underestimate in 2012 (since it was possible starting in 2012 for farmers to purchase subsidized fertilizer from dealers in the market with a coupon) we find the very low numbers in 2012 to be similar to those in 2010 (when the government was the sole distributor of fertilizer).

Table 2-1: Share of households who purchase external inputs by country

Countries	Share of households who purchase external inputs (%)	Share of households (%) by type of inputs purchased		
		Fertilizers	Pesticides/ Herbicides	Seeds
Malawi	70	49	4	51
Nigeria	71	42	38	29
Tanzania	18	8	13	NA
Uganda	16	5	14	NA

Source: Generated by authors using LSMS data
Note: NA implies that information is unavailable in the dataset used
External inputs refer to fertilizer, seeds and agrochemicals (pesticides/herbicides)
The shares are calculated amongst all households who farmed a plot in the year of the survey. Households with no agricultural plot farmed in that year are thus excluded.

Moreover, for both Malawi and Nigeria, the data show the relative importance of fertilizer and seeds in terms of shares of households buying these. For Malawi, the incidence of purchase of pesticides/herbicides is much less. Interestingly, only about a half and a third of the farmers buying external inputs in Tanzania and Uganda buy fertilizer, yet a larger share buy pesticides and herbicides; this appears surprising, but is also consistent with Sheahan and Barrett (2014) for Uganda.

Table 2 disaggregates input purchases by farm size. In each country, we stratified the farms by farm size strata: very small farmers (with less than 0.5 hectares) to larger (more than 5 hectares). Several points to note.

First, farmland is very concentrated in the medium and large farm strata, while small farmers dominate in numbers of farms. Roughly 65-75% of the land is farmed by medium/large farmers, but 75-80% of the farms are small farmers and this is consistent across countries. Small farmers (less than 2 hectares, per Hazell and Rahman 2013) predominate in terms of shares of total numbers of farmers in the study countries: 96% in Malawi, 88% in Nigeria, 63% in

Tanzania, and 76% in Uganda, giving a simple average of 81% for all four countries, or 76% if one excludes Malawi. By contrast, the medium stratum (2 to 4.99 hectare) and larger farm stratum (5 and more hectares) have a total share of 38% (11 and 27% of farmland respectively) in Malawi, 65% (22 and 43% respectively) in Nigeria and 79% (32 and 47% respectively) in Tanzania, and 67% (30 and 37% respectively) in Uganda. For these two strata, the simple average over the study countries is 67%; excluding Malawi, it is 80%.

Second, the shares of farmers buying external inputs are in a surprisingly tight distribution over farm size strata in all the countries. Again roughly grouping small farmers (up to 2 ha) and comparing with medium and larger farmers, one sees in Malawi that the shares (of farmers buying these inputs) is 71% for small farmers versus 88% for medium/large; for Nigeria, 78 versus 83%, for Tanzania, 15% versus 23%, and for Uganda, 14% versus 24%.

However, despite farmland concentration in which the medium/large farmers have 67% of the land, they constitute a disproportionately lower share (36%) of the external input purchase "pie". Thus the medium/large farmers are farming less intensively (in terms of applications of external inputs) than the small farmers and thus engaging less in the input market, and conversely, the small farmers are producing intensively with more external inputs per hectare, as farm technology intensification theory would predict (Binswanger-Mkhize and Ruttan, 1978, Boserup, 2005). This does not vary much over input types. But it does vary a lot over the countries: for Malawi, the medium/large group undertook 20% of the purchases in volume terms of all external inputs (versus its share in farmland of 38%); for Nigeria, the share of medium/large in external input purchases is 24%, versus its land share of 65%. For Tanzania, its share in inputs is 68%, near its share in land of 79%. In Uganda, its share is 42%, versus its share in land of 67%.

Table 2-2: Household purchase of external inputs by farm size strata

Countries	Farm size strata (ha)	Share of farmers in this stratum (%)	Share of national farmland in this stratum (%)	Share of farmers who purchase external inputs (%)	Share of total fertilizer demand bought by households in each stratum (%)	Share of total pesticide/ herbicide demand bought by households in each stratum (%)	Share of total seed demand bought by households in each stratum (%)	Share of total inputs demand bought by households in each stratum (%)
Malawi								
	0 - 0.49	45	13	65	30	12	28	30
	0.5 – 0.99	33	24	69	21	11	34	22
	1 – 1.99	18	24	79	29	40	23	29
	2 – 4.99	4	11	91	19	30	13	19
	5+	0	27	84	1	7	2	1
Nigeria								
	0 - 0.49	53	8	62	30	19	55	30
	0.5 – 0.99	20	12	78	25	20	17	23
	1 – 1.99	15	16	83	23	24	13	22
	2 – 4.99	9	22	82	16	21	8	16
	5+	3	43	85	5	16	7	8
Tanzania								
	0 - 0.49	20	2	13	5	5	NA	5
	0.5 – 0.99	19	5	14	9	7	NA	9
	1 – 1.99	24	14	17	20	13	NA	19
	2 – 4.99	26	32	22	41	46	NA	42
	5+	11	47	24	25	29	NA	26

Table 2-2 (cont'd)

Uganda								
	0 - 0.49	26	4	6	6	5	NA	5
	0.5 – 0.99	24	10	16	9	10	NA	10
	1 – 1.99	26	20	20	35	48	NA	44
	2 – 4.99	19	30	20	34	25	NA	28
	5+	6	37	28	16	12	NA	14

Source: Generated by authors using LSMS data

Note: NA implies that information is unavailable in the dataset used
External inputs refer to fertilizer, seeds and agrochemicals (pesticides/herbicides)

2.4. Farm input finance for the overall sample and by farm size

In this section, we explore the extent to which households use any credit arrangements for external input purchases and how that varies by farm size. We find consistent evidence across countries of very low use of any form of credit to buy these inputs (table 3). While Table 1 shows strong variation across countries in terms of shares of farmers buying external inputs, Table 3 shows only modest differences with respect to the very low shares (on average about 6%) of households that buy these inputs, using *any* form of credit.

Table 2-3: Share of households purchasing external inputs that finance the input purchase on credit

Countries	Of those who bought external inputs, share of farmers buying on credit (%)	Of those who bought seeds, fertilizers or pesticides/herbicides, share of farmers who bought on credit by input type		
		Fertilizers	Pesticides/ Herbicides	Seeds
Malawi	5	5	7	3
Nigeria	3	2	NA	3
Tanzania	11	14	7	3
Uganda	6	14	4	NA

Source: Generated by authors using LSMS data
Note: NA implies that information is unavailable in the dataset used
External inputs refers to fertilizer, seeds and agrochemicals (pesticides/herbicides)
Column 2 is the share amongst households who purchased at least one external input
Column 3, 4, and 5 are shares amongst households who purchased fertilizers, agrochemicals, or seeds, respectively.

The converse is that 94% only use their own cash to buy external inputs; this can be from non-credit resources such as cash sales of crops, and employment earnings (farm wage labor, migration, and rural nonfarm employment). In general for SSA, survey evidence has shown that on average, among the employment earnings, rural nonfarm employment is a far greater source of income than migration or farm labor wage income (Haggblade et al. 2010).

As noted in the introduction, there has been a presumption in the literature that to the extent farmers buy external inputs, they do it at least with informal credit or trader credit. But the analysis here shows that conventional wisdom is not supported empirically, and it is not just a lack of formal credit, but a near absence of the use of any credit, formal or informal, tied with input or output traders, in kind or in cash. Some other points to note.

First, of the very small shares of farmers buying external inputs on credit, there is within those sets sharp variation over input types. There tends to be 2-3 times more households getting some kind of credit for fertilizer compared to seeds or pesticides/herbicides.

Second, across all inputs, the limited credit based expenditures are for larger farmers. Table 4.a shows the shares of the landholding strata in all credit-based input outlays, so a sort of "pie" of strata shares in all credit transactions. The table shows that in Malawi, Tanzania, and Uganda, input credit is roughly farm size correlated - the great majority of the credit-based external input expenditures are concentrated outside the below-one-hectare group. Nigeria is a sharp outlier, with the great majority of the input credit taken by the "under 1 ha" group. These results do not differ much over input type.

This is largely confirmed by the distribution of shares, by stratum, of credit-based input outlay for each input i in total outlay for input i (table 4b). The importance of input credit tends to be concentrated in the middle to higher end of farm sizes, and be mainly in fertilizer and not very much in pesticides and seeds. In Malawi, Tanzania and Uganda, input credit is important only for fertilizer, averaging around 9% of fertilizer input outlay in Malawi but concentrated in the upper-small and medium farmers (1-5ha) where it averages a fifth of input expenditure. In

Tanzania, the share of input expenditure done on credit is correlated with land size, with about 10% for smaller farmers and about a quarter and a half for medium and larger farmers. For Uganda, it is only highly important for the 1-5ha group, where it reaches 40-50% of fertilizer expenditure. In Nigeria, the share is low for all, with about 3% on average and does not differ much over strata (except for small spikes to 11-12% among large farmers for fertilizer and seeds for 1-2 ha farmers).

Table 2-4a: Shares of Strata in all credit-based expenditure on external inputs

Countries	Farm size strata	Share of stratum buying on credit (%)	Share of stratum in all credit-based fertilizer outlay (%)	Share of stratum in all credit-based pesticide/ herbicide outlay (%)	Share of stratum in all credit-based seed outlay (%)	Share of stratum in all credit-based input outlay (%)
Malawi						
	0 -0.49	3	4	11	13	4
	0.5 – 0.99	3	4	15	16	4
	1 – 1.99	10	61	38	44	60
	2 – 4.99	10	32	36	27	32
	5+	14	0	0	0	0
Nigeria						
	0 -0.49	3	49	NA	13	45
	0.5 – 0.99	5	22	NA	22	22
	1 – 1.99	4	11	NA	62	16
	2 – 4.99	1	2	NA	0	2
	5+	6	16	NA	3	14
Tanzania						
	0 -0.49	2	0	0	NA	0
	0.5 – 0.99	6	4	3	NA	4
	1 – 1.99	8	10	15	NA	10
	2 – 4.99	20	36	69	NA	38
	5+	24	50	12	NA	48
Uganda						
	0 -0.49	0	0	0	NA	0
	0.5 – 0.99	2	3	17	NA	5
	1 – 1.99	11	57	54	NA	56
	2 – 4.99	11	40	28	NA	39
	5+	0	0	0	NA	0

Source: Generated by authors using LSMS data
Note: NA implies that information is unavailable in the dataset used
External inputs refers to fertilizer, seeds and agrochemicals (pesticides/herbicides)
In Column 3 the share is amongst households who purchased at least one external input

Table 2-4b: Share of credit-based input outlay for input i in total outlay for input i per stratum

Countries	Farm size strata	Share of credit-based input outlay for fertilizer in total outlay per stratum (%)	Share of credit-based input outlay for pesticides/ herbicides in total outlay per stratum (%)	Share of credit-based input outlay for seeds in total outlay for per stratum (%)	Share of credit-based input outlay for all inputs in total outlay for per stratum (%)
Malawi					
	0 -0.49	1	3	2	1
	0.5 – 0.99	2	5	2	2
	1 – 1.99	22	4	8	21
	2 – 4.99	18	4	8	17
	5+	0	0	0	0
Nigeria					
	0 -0.49	6	NA	1	4
	0.5 – 0.99	3	NA	3	3
	1 – 1.99	2	NA	12	2
	2 – 4.99	1	NA	0	0
	5+	11	NA	1	5
Tanzania					
	0 -0.49	2	0	NA	2
	0.5 – 0.99	12	4	NA	11
	1 – 1.99	15	10	NA	14
	2 – 4.99	26	12	NA	23
	5+	58	3	NA	48
Uganda					
	0 -0.49	0	0	NA	0
	0.5 – 0.99	12	3	NA	6
	1 – 1.99	53	2	NA	17
	2 – 4.99	40	2	NA	19
	5+	0	0	NA	0

Source: Generated by authors using LSMS data

Note: NA implies that information is unavailable in the dataset used

External inputs refers to fertilizer, seeds and agrochemicals (pesticides/herbicides)

In Column 3 the share is amongst households who purchased at least one external input

2.5. Crop type and input purchase financing

In this section we explore the correlation between the type of crop grown by farmers and financing "external inputs" on credit. We aggregate the purchase and the financing of external inputs over plots (to the household level) by crop. We classify crops into a set of what are traditionally called "food crops" (although they are often also sold for cash), including grains, horticulture, legumes, and tubers (grown as a staple), and what are traditionally called "cash crops", including tobacco, cotton, tea/coffee, and edible oil crops.

Three points stand out from an analysis of the distribution of household producing at least some of the crop types of cash crop versus food crops, by country and by farm size strata (Table 5).

First, as expected, grains dominate, but interestingly are not ubiquitous, reaching only about three-quarters of the farms in Nigeria, Tanzania, and Uganda, being near 100% only in Malawi. While there is a lot of variation, over the countries on average nearly a third of the farms grow horticultural crops, and on average a half grow beans/pulses, and a third grow tubers. The food cropping is thus fairly diversified on average.

Second, by contrast, production of cash crops is much more concentrated over farms in every country. Overall, on average over countries only a fifth of farmers grow cash crops, and that is but a tenth if one excludes Uganda. The crop focus differs over countries, with tea/coffee and oil crops standing out in Uganda, cotton and oil crops in Tanzania, oil crops in Nigeria, and tobacco and cotton in Malawi.

Third, there is little farm size bias in participation in all the food crops. The exception is that the smallest farms (below a half hectare) have a modestly lower participation rate than the other strata in all food crops but tubers, where they have higher or similar participation compared with the other strata. By contrast, for cash crops, there is a marked correlation of the share of farms producing any cash crop and farm size.

Table 2-5: Share of households producing key cash and food crops across farm size strata

Crop types	Farm size strata (hectares)	Share of farmers producing each crop type (%)			
		Malawi	Nigeria	Tanzania	Uganda
Cash crops					
Tobacco					
	0 -0.49	2	0	0	0
	0.5 – 0.99	11	0	0	0
	1 – 1.99	24	0	1	3
	2 – 4.99	32	0	2	2
	5+	28	0	2	0
	All	10	0	1	1
Cotton					
	0 -0.49	2	1	1	2
	0.5 – 0.99	6	3	2	4
	1 – 1.99	13	3	4	4
	2 – 4.99	19	1	5	6
	5+	0	0	10	8
	All	6	1	4	4
Tea/coffee					
	0 -0.49	0	0	2	22
	0.5 – 0.99	0	0	2	21
	1 – 1.99	0	0	2	22
	2 – 4.99	0	0	3	22
	5+	0	0	0	36
	All	0	0	2	23
Oil crops					
	0 -0.49	0	10	2	3
	0.5 – 0.99	1	7	4	10
	1 – 1.99	3	8	5	12
	2 – 4.99	2	19	5	19
	5+	0	14	5	16
	All	1	10	4	11
All cash crops					
	0 -0.49	4	10	4	26
	0.5 – 0.99	18	10	8	34
	1 – 1.99	39	11	11	39
	2 – 4.99	49	20	14	46
	5+	28	14	18	54
	All	17	11	11	37

Table 2-5 (cont'd)

Food crops					
Grains					
	0 -0.49	98	69	61	70
	0.5 – 0.99	99	87	74	83
	1 – 1.99	99	86	79	86
	2 – 4.99	99	84	83	81
	5+	100	88	85	81
	All	99	77	76	80
Horticulture					
	0 -0.49	29	33	22	55
	0.5 – 0.99	31	21	13	50
	1 – 1.99	37	22	12	48
	2 – 4.99	32	23	9	46
	5+	43	17	7	63
	All	31	28	13	51
Legumes					
	0 -0.49	62	29	12	76
	0.5 – 0.99	76	56	10	75
	1 – 1.99	79	60	12	77
	2 – 4.99	77	53	16	82
	5+	93	54	16	82
	All	71	42	13	78
Tubers					
	0 -0.49	8	61	16	74
	0.5 – 0.99	9	30	19	79
	1 – 1.99	14	34	19	74
	2 – 4.99	16	39	18	76
	5+	0	49	20	71
	All	10	48	18	75
All food crops					
	0 -0.49	100	98	95	100
	0.5 – 0.99	100	98	97	99
	1 – 1.99	100	99	96	100
	2 – 4.99	100	98	95	100
	5+	100	99	97	99
	All	100	98	96	100

Source: Generated by authors using LSMS data

Conventional wisdom suggests that farmers growing cash crops would commonly access external inputs on credit, in particular from processors, while food crop producers may not. To test this, we explore the shares (by crop type) of farm plots on which inputs purchased on credit are used (Table 6). While there is a lot of variation over countries, the average over all cash crops is 13%, compared with 6% for food crops. First, this is surprising because the shares do not differ greatly, as we had expected. Second, this average figure masks higher ratios in two pairs of countries; Malawi and Tanzania who average 20% for cash crops and 7% for food, while Nigeria and Uganda average 6% for cash crops and 5% for food which is extremely similar. A closer look indicates that the main difference between cash and food crops in this respect is driven by tobacco in Tanzania and Uganda, where four-fifths of the plots are grown with inputs bought on credit from the processors. Removing the tobacco outlier (for just Tanzania and Uganda) puts the overall credit share for cash crops close to that of food; as observed in the other study countries. Also, that outlier is composed of a tiny group of tobacco farmers in the sample for each country, about 1% of the total sample. The very limited and "enclave" nature of tobacco farming and its correlation with farm size (see Table 5) in those countries could explain why these are the main cases where the conventional image of contract-farming related credit is manifest.

Table 2-6: Cash crops versus food crops on which purchased external inputs are used that were financed on cash-credit for key cash and food crops (% of plots)

	Malawi	Nigeria	Tanzania	Uganda
Cash crops				
Tobacco	16	NA	87	81
Cotton	11	8	11	0
Tea/coffee	NA	NA	22	1
Oil crops	6	3	4	11
All cash crops	14	4	26	8
Food crops				
Grains	5	3	11	7
Horticulture	4	3	0	4
Legumes	5	2	11	6
Tubers	7	3	4	5
All food crops	5	3	10	6

Source: Generated by authors using LSMS data
Note: NA implies that information is unavailable in the dataset used
The numbers are percentages amongst of households producing each type of crop in each country.

2.6. Tied output input credit arrangements

Tied output/input credit arrangements occur when repayment for inputs on credit (received at planting) are made at harvest time. The LSMS data for our four study countries include a section about the management of crop harvests. We use farmers' responses concerning use of part of their harvests to repay advances for inputs from input or output traders and processors (especially for cash crops) for external inputs, and labor from workers.

Table 7 shows the share of farmers, overall and per stratum, using part of their harvests for these ends. The main finding is that such "tied credit" is very rare for external inputs (fewer than 2% of the farmers) across all study countries. By contrast, and reported for the first time in the Sub-Saharan African literature using cross-country surveys for comparison, we find that labor-output tying is much more common, with as many as 42% of the Malawi, 26% of Nigerian, and 68% of Tanzanian farmers doing this practice. (The dataset for Uganda did not allow this calculation.) By contrast, and not reported in the table, tying the land and output markets was not found to be common; the land tenure section of the surveys showed that sharecropping was extremely limited.

Moreover, the patterns of differentiation over strata differ by country so no single story emerges. For harvest payment to labor, in Uganda, the share rises with farm size, in Nigeria it slightly declines, and in Malawi it is in an inverted-U shape relation with farm size. Thus one cannot say that this traditional-tying of labor and harvest is more a phenomenon of the smallest farmers holding on to an old practice, as one might expect, given our hypothesis that larger farms are more apt to use monetized labor relations only. For harvest payment for external inputs, the shares are so small that there are no interesting inter-strata differences.

Table 2-7: Share of farmers using harvest to reimburse for inputs received on credit by farm size

Countries	Farm size strata	Share of farmers using their harvest to repay labor received on credit (%)	Share of farmers using their harvest to repay external inputs received on credit (%)
Malawi			
	0 -0.49	37	1
	0.5 – 0.99	45	3
	1 – 1.99	50	2
	2 – 4.99	47	1
	5+	24	0
	All	42	1.8
Nigeria			
	0 -0.49	26	1
	0.5 – 0.99	29	1
	1 – 1.99	26	3
	2 – 4.99	21	2
	5+	22	3
	All	26	1.4
Tanzania			
	0 -0.49	NA	0
	0.5 – 0.99	NA	1
	1 – 1.99	NA	1
	2 – 4.99	NA	4
	5+	NA	5
	All	NA	1.9
Uganda			
	0 -0.49	54	NA
	0.5 – 0.99	63	NA
	1 – 1.99	74	NA
	2 – 4.99	78	NA
	5+	81	NA
	All	68	NA

Source: Generated by authors using LSMS data

Notes: NA implies that information is unavailable in the dataset used

When we consider the “reimbursement of credit with the harvest” by type of crop, it is very minor or zero for the other cash crops (except tobacco in Tanzania, discussed below), and all of the food crops (Table 8). By contrast, use of harvest repayment for labor is very minor for cash crops (except for oil crops in Uganda where it is a quarter of farmers using it), but is significant in food crops across the countries, such as about a third in horticulture and a quarter in grains.

Interestingly, there is only a single situation (crop plus country) where this arrangement is important for external inputs, and that is for tobacco in Tanzania. We conjecture that this high prevalence of the use of harvest to reimburse for external inputs received on credit to produce tobacco in Tanzania is related to a widespread use of contract farming arrangement over tobacco production in Tanzania. If our conjecture is true, we should then expect to see a lot more contract farming arrangement over tobacco compared to cotton, tea/coffee, and oil crops.

Therefore we investigated outgrower schemes in the Tanzania data set and found that a very small proportion of farmers (1.8%) are involved in outgrower schemes. Moreover, tobacco represents 78.1% of the crops grown as part of an outgrower scheme or contract farming system, followed by cotton (18.8%). Though this does not say how much of the tobacco produced is grown as part of outgrower scheme, at least it gives an indication of the dominance of tobacco amongst the crop grown as part of outgrower schemes, and therefore confirms our conjecture about the prevalence of tied output-input arrangements for tobacco in Tanzania. There is no information about contract farming or outgrower schemes in the other countries to allow us to compare this pattern across countries.

Table 2-8: Financing inputs on credit with harvest across key cash and food crops.

Crops types	Share of plots where harvest is used to repay (advanced) labor (%)				Share of plots where harvest is used to repay external inputs (%)			
	Nigeria	Malawi	Uganda	Tanzania	Nigeria	Malawi	Uganda	Tanzania
Cash crops								
Tobacco	0	2	0	NA	0	2	NA	79
Cotton	10	0	0	NA	0	1	NA	6
Tea/coffee	NA	NA	1	NA	NA	NA	NA	3
Oil crops	8	0	25	NA	0	0	NA	0
Food crops								
Grains	17	22	27	NA	1	1	NA	1
Horticulture	18	32	36	NA	1	0	NA	0
Legumes	9	21	25	NA	1	1	NA	0
Tubers	5	29	30	NA	1	1	NA	0

Source: Generated by authors using LSMS data

Notes: NA implies that information is unavailable in the dataset used

Overall our results indicate that there is much less tied credit arrangement to finance external input than expected. Even though those arrangements appear to be more formal (from contract farming arrangements) and more likely for cash crops, we still see far less than expected (except for tobacco), even though the literature indicates the contrary.

2.7. Households' use of loans not specifically linked to input transactions

We use the term "loans" for credit unconnected directly and specifically to transactions of outputs or inputs. Such loans can come from formal (banks), semi-formal (micro-finance), and informal sources (friends, relatives, cooperatives, etc.). The LSMS data show loan data for Malawi, Nigeria, and Tanzania, and Malawi and Tanzania show what the loans were used for.

We find evidence that households in SSA do take loans however, this is rarely used for agricultural purposes. Nigeria had as much as 38% of farmers taking loans (Table 9). In Malawi, 23% of the households took a loan, but only 5% of them did so for farming; in Tanzania, it was but 11% taking loans of which 2% for farming purposes, hence a 5 to 1 ratio of overall loans to farm-destined loans in both (Table 9).

Table 2-9: Share of households with a member taking a financial loan

Country	Share of HHs taking a loan (%) [a]	Of those who took loan, share of HHs taking loans for farming (%)
Malawi	23	5
Nigeria	38	NA
Tanzania	11	2
Uganda	NA	NA

Source: Generated by authors using LSMS data
Notes: NA implies no data
[a] captures whether any household member received a loan in the last 12 months

Table 2-10: Purpose for which loans are taken (by source of loan) in Malawi

Source of loan	Share of loans taken for land purchase (%)	Share of loans taken for inputs purchase for food crops (%)	Share of loans taken for input purchase for tobacco (%)	Share of loans taken for input purchase for other cash crops (%)	Share of loans taken for business start-up capital (%)	Share of loans taken for purchase of non-farm inputs (%)	Share of loans taken for consumption (%)	Share of loans taken for other purposes (%)
Relative	0	13	1	4	24	4	36	19
Neighbors	1	7	1	2	24	7	44	14
Grocery/local merchant	0	4	0	0	7	7	83	0
Money lender (katapila)	0	11	1	6	19	6	40	19
Employer	10	19	0	0	0	10	52	10
Religious institution	0	23	8	0	15	8	39	8
Mardef	0	0	0	0	71	0	29	0
Mrfc	0	0	14	0	43	29	14	0
Sacco	7	4	0	4	30	11	19	26
Loans from all sources	1	9	2	4	31	8	31	14

Source: Generated by authors using LSMS data

Notes: MArdef and Mrfc are leading microfinance institutions in Malawi

Table 2-11: Purpose for which loans are taken (by source of loan) in Tanzania

Source of loan	Share of loans for subsistence needs	Share of loans for medical cost	Share of loans for school fees	Share of loans for ceremony/ wedding	Share of loans for land purchase	Share of loans for purchase inputs	Share of loans for other business inputs	Share of loans for purchase Machinery	Share of loans for buy/build dwelling	Share of loans for other purposes
Commercial banks	12	4	20	0	4	2	33	0	19	7
Microfinance institutions	8	0	15	0	5	8	42	0	15	8
Building society./mortgage	0	0	0	0	0	100	0	0	0	0
Other financial institutions	8	0	25	4	4	0	29	0	25	4
Neighbors / friends	48	12	4	3	2	5	14	1	3	10
Grocery/local merchant	62	6	2	0	0	5	19	0	2	4
Money lender	15	15	5	0	0	5	30	0	5	25
Employer	14	0	21	7	14	0	7	0	7	29
Religious institutions	0	0	0	0	17	17	33	0	0	33
NGO	0	0	13	0	0	25	25	13	13	13
Self-help groups	32	0	11	2	1	6	30	0	9	9
Others	23	7	5	2	5	23	11	5	7	14
Loans from all sources	31	6	10	2	3	6	24	1	9	9

Source: Generated by authors using LSMS data

Table 2-12: Sources of Cash Income in Nigeria, North and South, 2010 and 2012

	HOUSEHOLD CASH SOURCES (000 Naira)						SHARE OF CASH FROM EACH SOURCE (%)					
	NIGERIA		SOUTH		NORTH		NIGERIA		SOUTH		NORTH	
INCOME SOURCES	2010	2012	2010	2012	2010	2012	2010	2012	2010	2012	2010	2012
CASH INCOME												
Net profit from household enterprise	48.3	113.1	41.2	102.2	53.8	121.5	17	26	12	18	22	38
Wage income	193.2	261	249.7	406	149.9	151	67	60	75	70	60	47
Crop sales	42	56.7	38.2	66.2	44.9	49.4	15	13	11	11	18	15
Livestock net sales	0.7	0.9	0.4	1	0.9	0.9	0	0	0	0	0	0
Remittances	2.2	1	5	1.8	0.1	0.4	1	0	1	0	0	0
Total cash	**286.4**	**432.7**	**334.5**	**577.2**	**249.6**	**323.1**	**100**	**100**	**100**	**100**	**100**	**100**
Inputs credit transactions	0.2	0.3	0.1	0.1	0.3	0.5	0	0	0	0	0	0
Inputs non credit transactions	7.2	10.6	3.8	3.9	9.7	15.7	3	2	1	1	4	5
Total input purchase	7.3	11	3.8	4	10	16.3	3	3	1	1	4	5
Hired labor value for harvest only	18.9	12.4	8	7	27.2	16.5	7	3	2	1	11	5
Imputed value of own crop output	112.9	137.4	72.5	83.1	143.8	178.6	39	32	22	14	58	55

Source: Generated by authors using LSMS data

Note: The numbers in the left panel are zero-in averages. The shares on the right are based on ration of number on the left to the total cash value. Inputs include fertilizer, seeds, and pesticides, except in Input credit transactions where it does not include pesticides credit because the information was not available in the data. For each value in the table, instead of deleting outliers we winsorized them i.e. replace top 10% values by the highest value within 90% of the distributions, thus creating a pile up at the top without changing the distribution (Cox, 2006).

For Imputation of value of own crop output method, we estimate unit prices of crops for crops that were sold, and then we use the median price in the local governments and multiply by harvest quantities to get the value of crop sales.

The harvest labor for planting activities is missing in the 2010 dataset, and therefore we focus on the harvest labor only in both years.

Instead, the loans were taken for nonfarm business startup and for consumption (Tables 10 and 11). This is striking because one would expect credit-constrained farmers to use these loans to finance farm input purchases. As is shown below, a key factor that determines external input purchase is engaging in nonfarm enterprises and wage labor. Thus it appears that farmers prefer to use financial credit to finance the set up/expansion of their nonfarm enterprises but use the generated cash from these nonfarm enterprises to finance external input purchases.

2.8. Determinants of inputs purchase in Nigeria

Thus far, we find consistent evidence that the use of any form of credit to finance external input purchase is extremely low. This begs to question how households finance these purchases. While one might expect credit-constrained farmers to use loans to finance input purchase, we find that farmers tend to use loans to finance business startup and consumption. This could be driven by the risky nature of agricultural investments and/or low expected returns on investments in modern inputs relative to the cost of credit.[6] In this section, we explore how farmers finance their input purchase by estimating the determinants of fertilizer purchases by Nigerian farmers. Our analysis lays emphasis on the role of non-farm employment (wage and self-employment) and agricultural productivity risks (captured by rainfall variability), as well as regional differences (north versus south) in fertilizer purchases decisions and intensity amongst Nigerian farm households.

---

[6] We do not test this hypothesis as it is beyond the goal of this paper. However, we recognize that it is an important question that deserves further investigation.

2.8.1. Conceptual and empirical framework

The fertilizer purchase decision follows a standard input demand function derived from a constrained household utility maximization problem and presented in Sadoulet and de Janvry (1996). Fertilizer demand can be expressed as a function of a vector of prices, risk proxies, a vector of complementary and substitute farm capital, and relevant shifter variables such as crop type. We first consider the decision to purchase a particular input or not and then the extent of input purchase.

We model the farmers fertilizer purchase decisions using the following unobserved effect binary dependent variable model (Green William, 2000, Wooldridge, 2010):

$$\boldsymbol{Y_{it}^{*} = 1[X_{it}'\beta + v_{it} + c_i > 0]} \; ; \quad \textbf{i=1, 2, …, N ; t=1, 2,…,T} \qquad \textbf{(1)}$$

In the model above, Y_{it}^{*} is the underlying latent variable which characterize farmer i's net benefit (or utility) from purchasing fertilizers, in period t. While this latent variable is unobserved, it determines the observed binary outcome variable Y_{it} (fertilizer purchase) which takes value 1 if $Y_{it}^{*} > 0$ and a fertilizer is purchased, and 0 otherwise. X_{it} is the vector of explanatory variables included in the model. β is the vector of parameters of interest. v_{it} is the error term assumed to follow a standard normal distribution, leading to an unobservable effects Probit model for the fertilizer purchase equation (Green William, 2000, Wooldridge, 2010).

$$Prob(Y_{it} = 1|X_{it}, c_i) = \Phi(X_{it}'\beta + c_i) \qquad (2)$$

For the intensity of fertilizer purchase, we model it using the following unobserved effects Tobit model to account for zeros dues to corner solution in the dependent variable (Wooldridge, 2010):

$$Y_{1_{it}} = \max\,(0, Z_{it}'\beta + c_i + u_{it}) \; ; \quad \text{i=1, 2, …, N ; t=1, 2,…,T} \qquad (3)$$

$$D(u_{it}|Z_{it}, c_i) = Normal\;(0, \sigma_u^2)$$

In this model, Y_1 is the dependent variable representing the number of kilograms of fertilizer purchased per unit of land cultivated. Z_{it} is the vector of explanatory variables that potentially affect quantity of fertilizer purchased. u_{it} is the error term assumed to follow a normal distribution with mean 0 and standard deviation σ_u.

In both models, c_i represents the unobserved effect parameter, modeled using the Mundlak (1978) special case of (Chamberlain, 1982) approach called correlated random effect (CRE):

$c_i = \psi + \overline{X}_\iota \xi + a_i, \quad a_i | X_i \sim Normal(0, \sigma_a^2)$

where $\overline{X}_\iota$ represents time averages of the explanatory variables. Assuming conditional independence, the full model becomes:

$$Prob(Y_{it} = 1 | X_{it}, \overline{X}_\iota) = \Phi(X_{it}'\beta_a + \psi_a + \overline{X}_\iota \xi_a), \tag{4}$$

with $\theta_a = \theta/(1 + \sigma_a^2)^{1/2}$, $\quad\quad \theta = (\beta, \psi, \xi\ )$

Average partial effects are identified and can be derived from the above model as:

$PE_{x_j} = \frac{\partial Prob(Y_{it}=1|X_{it},\overline{X}_\iota)}{\partial x_{jt}} = \beta_{aj}.\phi(X_{it}'\beta_a + \psi_a + \overline{X}_\iota \xi_a)$, with $\theta_a = \theta/(1 + \sigma_a^2)^{1/2}$, $\theta = (\beta, \psi, \xi\ )$, for the Probit model; and $PE_{x_j} = \beta_{aj}.\Phi(Z_{it}'\beta_a + \psi_a + \overline{Z}_\iota \xi_a)$ with $\gamma_a = \gamma/(\sigma_u^2 + \sigma_a^2)^{1/2}$, $\gamma = (\beta, \psi, \xi\ )$, for the Tobit model.

The use of the CRE model is preferred over alternative methods such as fixed effects (FE) and random effects (RE) models in the case of non linear models (Wooldridge, 2010). CRE models deal effectively with time invariant unobserved heterogeneity in both linear and non linear panel data, while FE with non linear models are known to produce inconsistent estimates as they treat the unobserved effects c_i as N parameters to estimate, leading to incidental parameters problem (for fixed T). CRE models includes the more efficient RE model as a special case, when $\xi_a = 0$ (Green William, 2000, Wooldridge, 2010). It is simple to compute a wald statistic to test the sensibility of the CRE model against the RE model (*Ho:* $\xi_a = 0$).

Consistent with the CRE model described above, the determinants of fertilizer purchase decisions and quantity purchased are estimated using pooled Probit [7] and pooled Tobit approaches respectively. Partial (or pooled) maximum likelihood estimation approach yield consistent estimates that can be made robust to serial correlation of the errors Wooldridge (2010).

Each regression equation includes a set of explanatory variables as well as the time averages of the explanatory variables. A Wald test of joint significance of those time average variables is performed to test whether a traditional random effects model would be appropriate or not. A dummy variable for time period is included taking value 1 for year 2010 and 0 for year 2012.

Though the use of CRE model addresses potential biases due to time invariant unobserved heterogeneity similar to FE models, the strict exogeneity assumption implies that there is no remaining endogeneity after controlling for time invariant unobservable. This assumes that, after conditioning on the heterogeneity parameter c_i, the explanatory variables X_{it} included in the Probit model or the Z_{it} included in the Tobit model are truly exogenous variables. If this assumption fails, and any explanatory variables are correlated with the time varying idiosyncratic shocks, our estimates might be biased. To minimize any remaining bias from time

[7] We also use Generalized Estimating Equation (GEE) approach, which generally produces more efficient estimation results for the Probit model. We find very similar results with the pooled Probit model, both in terms of margins estimates and standard errors. The GEE results are available from the authors upon request.

varying unobserved heterogeneity, we include a rich set of observable characteristics that can proxy for a lot of unobservable factors.

For this analysis we hone in on one study country, Nigeria using the two available waves (for 2010 and 2012) of The Nigeria Living Standard Measurement Study-Integrated Survey on Agriculture (LSMS-ISA); the panel version of the nationally representative dataset used in previous sections. [8]

The explanatory variables used in the model (see tables 13) capture the socio economic profile of the household and geographic factors likely to affect the decision to use modern inputs (Feder et al. 1985). Socio economic variables include the gender and age of the household head to proxy for systematic differences in resource access and use and the number of years of experience in agricultural activity respectively. We capture education of the household head by a dummy variable taking value 1 when the head of the household has received any formal education, and 0 if he has never been to school. Education matters for fertilizer purchase as it can improve access and understanding of information related to inputs use as well as market information such as prices. However it can also push people out of farming as they get access to better off-farm opportunities. We also include the household dependency ratio (measured as the number of household members aged less than 14 or older than 65 years old divided by the non-dependent members between 14 and 65 years) to capture households' productive structure. We also control for the size of total land holdings (in hectares) and households' ownership of

[8] We focus on Nigeria because it has one of the highest prevalence of inputs purchases amongst our 4 study countries, offering a more interesting analysis of the determinants of inputs purchases. We did not choose Malawi, which also has high prevalence of fertilizer purchases, because of high rates of input subsidies, which might affect input purchases and confound our estimations. Indeed, estimates from recent data indicate that up to 60% of farm households in Malawi receive subsidized fertilizer (IHS, 2013).

agricultural assets (farm machines), captured using an asset index computed using the principal component analysis approach[9]. Household ownership of productive agricultural assets increases the marginal productivity of input use, and thus is expected to have positive effects on fertilizer purchases. To account for the role that region specific effects such as infrastructure and soil conditions play in fertilizer purchase decision, we include geographical dummy variables for each of the 6 geographical zones of the country (North East, North West, South East, South Central, South West). For similar reasons, urban versus rural area dummy are also included as significant socio economic and cultural factors, as well as infrastructures and prices, differentiate urban and rural areas, at the same time as they are likely to affect inputs purchase decisions. Given bio-physical differences across crop types, the need for and the yield response to input use vary by crop types (De Geus and de Geus, 1967) and thus we control for types of crops produced by the household. We use fractional variables representing shares of grains, legumes, tubers, horticulture crops, oils crops, etc. in total land cultivated by the household. To capture crops grown for commercialization versus subsistence purposes, which likely informs investment in modern inputs, we control for the total value of crop sales per hectare of land cultivated.

For the potential sources of input finance, we first include a binary variable to capture farmer's access to loan. This variable takes value 1 when any member of the household took a loan the year prior to the survey period. Dummies for participation in off-farm activities such as household non-farm enterprises and off-farm wage employment are also included as they are

---

[9] Following Filmer and Pritchett (2001), principal component analysis is used to generate a productive asset index based on the household ownership of farm machinery (eg. tractors, ploughs and irrigation pumps) in the sample.

alternative sources of income for the household[10]. While participation in off farm employment could divert attention from own farm (Smale et al., 2016), it could also reduce the cash constraint farmers' face thus enabling the purchase of modern inputs (Oseni and Winters, 2009).

Agricultural risk such as rain variability plays an important role in input purchase decisions. In case of negative weather shocks, returns to input use may even fall below the returns from not using inputs. Therefore, the likelihood of a negative rainfall shock reduces farmers' incentives to invest in agricultural production, especially in absence of ex-post risk mitigation opportunities and lack of credit and insurance mechanisms (Dercon and Christiaensen, 2008). Agricultural productivity risk is captured by the coefficient of variation of rainfall in the local government (lga). Following standard law of demand, price of fertilizer in the lga (in Naira) is included to capture the slope of the demand for fertilizer. Household distance from nearest major road is also included to capture transaction costs of accessing fertilizer.

In addition to the country level analysis, we estimate regional level parameters for the Northern and Southern regions. As mentioned by Oseni and Winters (2009), there are important cultural and socio economic differences between the two regions which can affect the way farmers' in those regions respond to changes in determinants of inputs use and purchase. Compared to the South (See tables 13b and 13c), the north of Nigeria is more rural and traditional, with larger household sizes (7 vs. 5.3 on average), greater poverty and less education (Only about 53% education rate of household heads in north compared to 70% in the south). There are also fewer female-headed households (less than 4% in north vs. 26% in the South) with young household heads (age of 48 on average in north vs. 56 in south). In part because of

---

[10] Off-farm wage employment relates to all activities performed for a wage salary outside of one's own farm. These include both farm and non-farm wage employment as these are all potential sources of income household could use to finance input purchases.

higher urbanization with major cities such as Lagos in the south, rate of participation in non-farm wage employment is higher in the south (about 80%) than the north (about 45%). As for non-farm self-employment participation, households in the north seem to have a slightly higher participation rate (about 65%) compared to the south (about 55%).

Table 2-13: Summary statistics of variables used in the regression analysis by region

	Wave 2 (2012)				Wave 1 (2010)			
	N	Mean	sd	CV	N	Mean	sd	CV
PART A: NIGERIA								
Use fertilizer	2,951.00	45.2	49.8	110.1	3,036.00	45.3	49.8	110
Purchase Fertilizer	2,951.00	41.9	49.4	117.8	3,036.00	41.4	49.3	119
Household received loan from formal source (0/1)	2,930.00	4.8	21.5	443.3	3,035.00	2.9	16.8	578.6
Household received loan from informal source (0/1)	2,930.00	19.2	39.4	205	3,035.00	18	38.4	213.6
Household received loan from friends or relatives (0/1)	2,930.00	28.5	45.2	158.3	3,035.00	28.4	45.1	158.9
Loan (0/1)	2,930.00	39.9	49	122.7	3,035.00	39.3	48.9	124.3
Land holding size (hectares)	2,951.00	0.8	0.8	98.9	3,036.00	0.9	1	109
Household head is Male (0/1)	2,870.00	87.4	33.2	37.9	3,034.00	88.1	32.4	36.8
Age of the household head (years)	2,846.00	52.3	14.9	28.5	3,026.00	50.7	15.1	29.8
Household dependency ratio	2,870.00	1	0.8	85.4	3,034.00	1.1	0.9	82.5
Household head has formal education (0/1)	2,870.00	59.8	49	82	3,034.00	59.7	49.1	82.2
Household resides in an urban area (0/1)	2,870.00	11.7	32.2	274.6	3,034.00	13.3	33.9	255.9
Agricultural assets index	2,951.00	0.2	3.4	1576.2	3,036.00	0.3	4.3	1391.3
Fertilizer price in Naira per Kg	2,870.00	102.5	26.8	26.1	3,034.00	85.4	23.6	27.7
Distance to Nearest Major Road (Km)	2,951.00	7.3	7.8	107.7	3,036.00	17	18.5	109.4
Distance to Nearest Market (Km)	2,951.00	70.4	39.1	55.5	3,036.00	71.3	39.7	55.7
A household member is engaged in Non-Farm self-employment (0/1)	2,951.00	64.3	47.9	74.6	3,036.00	56.2	49.6	88.4
A household member is engaged in off Farm wage employment (0/1)	2,951.00	57.3	49.5	86.3	3,036.00	60.3	48.9	81.1
Value of sales per ha of land cultivated	2,633.00	43,078.00	59,615.00	138.4	2,689.00	42,684.00	61,066.00	143.1
Share of total land cultivated allocated to grains crops	2,633.00	44.2	33.7	76.2	2,689.00	43	35.3	81.9
Share of total land cultivated allocated to legumes crops	2,633.00	17.1	21.9	128.4	2,689.00	16	22.5	141.1
Share of total land cultivated allocated to tubers crops	2,633.00	24.5	35	143.1	2,689.00	27.7	37.7	136.2
Share of total land cultivated allocated to oil crops	2,633.00	2.9	11.4	386.7	2,689.00	3	11.4	382.3
Share of total land cultivated allocated to horticulture crops	2,633.00	8.3	19.1	230.7	2,689.00	7.4	17.7	238.1
Share of total land cultivated allocated to cotton	2,633.00	0.3	3.3	1133.4	2,689.00	0.1	2.1	1688.7

Table 2-13 (cont'd)

Share of total land cultivated allocated to tobacco	2,633.00	0	0	-	2,689.00	0	0	-
Share of total land cultivated allocated to tea/coffee	2,633.00	0	0	-	2,689.00	0	0	-
Share of total land cultivated allocated to other crops	2,633.00	2.7	14.1	513.3	2,689.00	2.8	14	504.3
zone==north central	2,951.00	16.5	37.1	225.3	3,036.00	16.9	37.5	221.6
zone==north east	2,951.00	20.4	40.3	197.5	3,036.00	20.3	40.2	198.4
zone==north west	2,951.00	24.1	42.8	177.7	3,036.00	21.7	41.2	189.9
zone==south east	2,951.00	19	39.3	206.5	3,036.00	19.9	40	200.5
zone==south south	2,951.00	12.7	33.3	262.6	3,036.00	12.7	33.3	262.5
zone==south west	2,951.00	7.4	26.2	354.1	3,036.00	8.5	27.9	328.2
Coefficient of variation of rainfall	2,878.00	94.8	27.4	28.9	2,956.00	93.5	27.3	29.2

PART B: SOUTHERN NIGERIA								
Use fertilizer	1,153.00	20.6	40.5	196.2	1,248.00	25.2	43.5	172.2
Purchase Fertilizer	1,153.00	19.5	39.7	203.2	1,248.00	22.6	41.8	185.1
Household received loan from formal source (0/1)	1,150.00	8.1	27.3	337.3	1,247.00	3.4	18.1	535.9
Household received loan from informal source (0/1)	1,150.00	24.1	42.8	177.6	1,247.00	18.4	38.7	210.9
Household received loan from friends or relatives (0/1)	1,150.00	26.3	44	167.6	1,247.00	22.3	41.6	186.8
Loan (0/1)	1,150.00	42.1	49.4	117.3	1,247.00	36.2	48.1	132.9
Land holding size (hectares)	1,153.00	0.4	0.6	147.6	1,248.00	0.5	0.7	160.4
Household head is Male (0/1)	1,125.00	73.6	44.1	59.9	1,247.00	76.4	42.5	55.6
Age of the household head (years)	1,109.00	57.4	14.6	25.5	1,241.00	55.5	14.7	26.5
Household dependency ratio	1,125.00	0.8	0.8	102.7	1,247.00	0.9	0.9	101.2
Household head has formal education (0/1)	1,125.00	70.1	45.8	65.3	1,247.00	70.5	45.6	64.7
Household resides in an urban area (0/1)	1,125.00	17.1	37.6	220.5	1,247.00	18.1	38.5	212.7
Agricultural assets index	1,153.00	0.1	2.7	2312.7	1,248.00	0.4	5	1237.4
Fertilizer price in Naira per Kg	1,125.00	105.8	28	26.4	1,247.00	92.9	24.6	26.4
Distance to Nearest Major Road (Km)	1,153.00	5.3	6.4	122	1,248.00	11	9.4	85.6
Distance to Nearest Market (Km)	1,153.00	65.9	36.7	55.7	1,248.00	66.2	37.1	56.1
A household member is engaged in Non-Farm self-employment (0/1)	1,153.00	61.1	48.8	79.8	1,248.00	50.7	50	98.6

Table 2-13 (cont'd)

A household member is engaged in off Farm wage employment (0/1)	1,153.00	84.4	36.3	43	1,248.00	74	43.9	59.4
Value of sales per ha of land cultivated	899	68,503.00	67,746.00	98.9	977	65,315.00	70,634.00	108.1
Share of total land cultivated allocated to grains crops	899	15.7	25.5	162.5	977	14.6	26.6	182.4
Share of total land cultivated allocated to legumes crops	899	1.2	7.9	663	977	0.7	6.9	947.3
Share of total land cultivated allocated to tubers crops	899	52.8	37.9	71.7	977	58.4	38.8	66.4
Share of total land cultivated allocated to oil crops	899	5.5	16.2	297.8	977	5	15.4	304.9
Share of total land cultivated allocated to horticulture crops	899	17.6	26.3	149.3	977	14.6	23.7	162.9
Share of total land cultivated allocated to cotton	899	0	0	-	977	0	0	-
Share of total land cultivated allocated to tobacco	899	0	0	-	977	0	0	-
Share of total land cultivated allocated to tea/coffee	899	0	0	-	977	0	0	-
Share of total land cultivated allocated to other crops	899	7.2	22.5	311.9	977	6.7	21.7	323.2
zone==north central	1,153.00	0	0	-	1,248.00	0	0	-
zone==north east	1,153.00	0	0	-	1,248.00	0	0	-
zone==north west	1,153.00	0	0	-	1,248.00	0	0	-
zone==south east	1,153.00	48.7	50	102.8	1,248.00	48.5	50	103.1
zone==south south	1,153.00	32.4	46.8	144.4	1,248.00	30.9	46.2	149.8
zone==south west	1,153.00	18.9	39.2	207.1	1,248.00	20.7	40.5	196
Coefficient of variation of rainfall	1,113.00	67.7	3.9	5.8	1,207.00	67.7	3.9	5.7

PART C: NORTHERN NIGERIA								
Use fertilizer	1,798.00	61	48.8	80.1	1,788.00	59.2	49.2	83
Purchase Fertilizer	1,798.00	56.2	49.6	88.2	1,788.00	54.5	49.8	91.3
Household received loan from formal source (0/1)	1,780.00	2.8	16.4	594.6	1,788.00	2.6	15.8	615.6
Household received loan from informal source (0/1)	1,780.00	16.1	36.7	228.6	1,788.00	17.7	38.2	215.5
Household received loan from friends or relatives (0/1)	1,780.00	30	45.8	152.8	1,788.00	32.6	46.9	143.8
Loan (0/1)	1,780.00	38.5	48.7	126.3	1,788.00	41.5	49.3	118.8
Land holding size (hectares)	1,798.00	1	0.8	76.2	1,788.00	1.2	1	84.3
Household head is Male (0/1)	1,745.00	96.3	18.8	19.5	1,787.00	96.2	19.1	19.9
Age of the household head (years)	1,737.00	49.1	14.2	29	1,785.00	47.3	14.4	30.5
Household dependency ratio	1,745.00	1.1	0.8	74.9	1,787.00	1.2	0.9	70.9

Table 2-13 (cont'd)

Household head has formal education (0/1)	1,745.00	53.1	49.9	94	1,787.00	52.1	50	95.9
Household resides in an urban area (0/1)	1,745.00	8.3	27.5	333.5	1,787.00	9.8	29.8	302.7
Agricultural assets index	1,798.00	0.3	3.7	1345.5	1,788.00	0.2	3.8	1538.7
Fertilizer price in Naira per Kg	1,745.00	100.4	25.8	25.7	1,787.00	80.1	21.4	26.8
Distance to Nearest Major Road (Km)	1,798.00	8.5	8.4	97.8	1,788.00	21.1	21.9	103.9
Distance to Nearest Market (Km)	1,798.00	73.3	40.2	54.9	1,788.00	75	41.1	54.8
A household member is engaged in Non Farm self employment (0/1)	1,798.00	66.2	47.3	71.4	1,788.00	60	49	81.7
A household member is engaged in off Farm wage employment (0/1)	1,798.00	39.9	49	122.7	1,788.00	50.8	50	98.5
Value of sales per ha of land cultivated	1,734.00	29,896.00	50,102.00	167.6	1,712.00	29,769.00	50,529.00	169.7
Share of total land cultivated allocated to grains crops	1,734.00	59	27.3	46.3	1,712.00	59.3	28.7	48.4
Share of total land cultivated allocated to legumes crops	1,734.00	25.3	22.3	88.3	1,712.00	24.7	23.7	96.1
Share of total land cultivated allocated to tubers crops	1,734.00	9.8	22.1	225.6	1,712.00	10.1	23	227
Share of total land cultivated allocated to oil crops	1,734.00	1.6	7.4	452.8	1,712.00	1.8	8.1	448.4
Share of total land cultivated allocated to horticulture crops	1,734.00	3.4	11.3	328.1	1,712.00	3.4	11.2	332.8
Share of total land cultivated allocated to cotton	1,734.00	0.4	4	916.6	1,712.00	0.2	2.6	1351.5
Share of total land cultivated allocated to tobacco	1,734.00	0	0	-	1,712.00	0	0	-
Share of total land cultivated allocated to tea/coffee	1,734.00	0	0	-	1,712.00	0	0	-
Share of total land cultivated allocated to other crops	1,734.00	0.4	4.5	1107.1	1,712.00	0.5	5.2	963.3
zone==north central	1,798.00	27	44.4	164.3	1,788.00	28.8	45.3	157.5
zone==north east	1,798.00	33.5	47.2	141	1,788.00	34.4	47.5	138.1
zone==north west	1,798.00	39.5	48.9	123.8	1,788.00	36.9	48.3	130.9
zone==south east	1,798.00	0	0	-	1,788.00	0	0	-
zone==south south	1,798.00	0	0	-	1,788.00	0	0	-
zone==south west	1,798.00	0	0	-	1,788.00	0	0	-
Coefficient of variation of rainfall	1,765.00	111.9	21.3	19.1	1,749.00	111.3	21.7	19.5

Source: Generated by authors using LSMS data
Note: N=Number of observations with non-missing values
sd=standard deviation, CV=Coefficient of variation of the variable

2.8.2. Regression results

Table 14 presents the average partial effects of the determinants of fertilizer purchase overall in Nigeria, and by region from the pooled Probit and pooled Tobit estimates. The results are generally consistent with the literature on modern input demand but reveal substantial differences between northern and southern Nigeria. Most relevant determinants of fertilizer purchase seem to show higher significance in the Northern region compared to the Southern region, possibly reflecting the fact that the Northern part of Nigeria uses and thus purchases more fertilizer in general and therefore is more responsive to various determinants than the South.

We find that participation in non-farm self-employment has positive and significant effects on fertilizer purchases; though, wage employment does not seem to matter significantly. The estimated APE indicates that participation in non-farm self-employment increases likelihood of purchasing fertilizer by about 7.3 percentage points in Nigeria; and this result is consistent across both southern Nigeria (10.5 percentage point increase) and northern Nigeria (5.4 percentage points). These findings coincide generally with the descriptive findings presented above, and echoes somewhat literature such as Adesina (1996) for Ivory Coast and Oseni and Winters (2009) for Nigeria. However, contrary to Oseni and Winters (2009), we find that wage employment did not appear as a significant determinant of fertilizer purchase and even has a negative coefficient. Neither wage employment, nor self-employment are significant determinants of the amounts of fertilizer purchased, according to the Tobit results. Smale et al. (2016), suggest that while we may expect participation in non-farm employment to relax farmers' financial constraints and allow increased purchase of fertilizer, ceteris paribus, potential competition in resource commitment between farm and non-farm sector may shrink the positive effect. It may also be that as households earn income from non-farm sector, they depend less on agriculture and therefore reduce investment in agricultural inputs. The balance between farm and off-farm competition for resources on one hand, and the relaxation of cash constraints to allow financing of agriculture inputs on the other hand, determine the observed effects of non-farm employment. In our case, the positive effect of non-farm self-employment in relaxing cash constraints has proven dominant compared to the negative effect, especially in the Northern part of Nigeria. In the South, they seem to balance each other out, leading to a non-significant effect. As for wage employment, it seems to induce a stronger competition with agriculture in terms of resources such as labor. The reduced labor in agriculture due to off-farm wage employment participation probably reduces the marginal profitability of investment in agricultural input, leading to a tendency to reduce input purchases, though this effect is not statistically significant.

While access to loan affect positively fertilizer purchase, the effect is significant only in the northern part of Nigeria. A closer investigation of the types of loans taken by farmers shows that loans from friends and relatives (rather than loan from formal and semi formal institutions) seem to drive most of these results[11]. This could illustrate the fact that loans, and in particular loan from formal and semi formal institutions, are limited for agricultural investment. Given the risks related to agricultural activities, formal and semi-formal credit suppliers are reluctant to provide loan for agricultural purposes, as they fear higher risk of default. Though we could not test specifically this hypothesis in Nigeria due to data limitation, table 10 and 11 show for Malawi and Tanzania respectively that food and non-food consumption and investment in business start-up are by far the primary purposes of the loan taken by households. Besides, the

[11] The regression results that use the various types of loans are available from the authors upon request.

fact that the effect of loan is significant only in the Northern part of the country could be explained by the dominant sources of loans in each region.

Friends and relatives seem be a more dominant source of loans taken by households in the North than in the South. Table 13 provides some evidence for this. Between 22 and 26 percent of households reported receiving loan from friends and relatives in the South, compared to 30 to 33 percent in the North. In the meantime, the same tables show that access to semi formal and formal sources of loan are higher in the South than the North.

The regression analyses seem to reflect a complementarity between loan and non-farm self-employment. Access to loan affects significantly fertilizer purchases only in North where the effect of non-farm self-employment seems weaker. Conversely, in the South, where access to loan does not have a significant effect on fertilizer purchase, we see a strong effect of non-farm self-employment.

The coefficient of variation of rainfall, which captures an important dimension of the risks related to agricultural production activity has, as expected, a strongly negative effect on fertilizer purchase, but this is only significant in the north. This result is very important as investments in modern input purchases though generally profitable, are costly and can yield very low (or even negative) returns in case of negative weather shocks. For that reason, poor farmers without access to risk mitigation opportunities can be very sensitive to rain variability (Dercon and Christiaensen, 2008).

Table 2-14: Estimation results of determinants of fertilizer purchase and quantity purchased by farmers in Nigeria

	APE on Fertilizer purchase decision (0/1)			APE on fertilizer purchase in Kg per ha of land cultivated		
VARIABLES	NIGERIA	SOUTH	NORTH	NIGERIA	SOUTH	NORTH
Household head is Male (0/1)	0.051**	0.022	0.123***	64.197***	14.944	94.837**
	[0.038]	[0.384]	[0.006]	[0.002]	[0.433]	[0.023]
Age of the household head (years)	0.000	0.000	0.000	0.087	0.342	0.298
	[0.819]	[0.563]	[0.861]	[0.786]	[0.541]	[0.467]
Household dependency ratio	-0.006	0.032+	-0.031**	-0.067	25.690	-13.470
	[0.628]	[0.119]	[0.048]	[0.995]	[0.157]	[0.357]
Household head has formal education (0/1)	0.081***	0.048*	0.099***	34.851***	42.246**	34.301*
	[0.000]	[0.073]	[0.000]	[0.009]	[0.044]	[0.050]
Land holding size (hectares)	0.018	-0.033	0.032**	-47.110***	-44.217*	-58.193***
	[0.184]	[0.269]	[0.036]	[0.000]	[0.062]	[0.000]
Agricultural asset index	0.002	0.003	0.002	1.713*	2.507	1.630+
	[0.275]	[0.410]	[0.389]	[0.078]	[0.403]	[0.123]
LOG of crop sales in naira per ha of harvested land	0.001*	0.000	0.001*	0.314	0.084	0.531
	[0.095]	[0.578]	[0.080]	[0.465]	[0.887]	[0.383]
A household member is engaged in Non-Farm Self-employment (1/0)	0.073***	0.105**	0.054*	18.508	61.324+	2.569
	[0.005]	[0.012]	[0.097]	[0.402]	[0.109]	[0.927]
A household member is engaged in wage employment (1/0)	-0.018	-0.011	-0.014	-7.715	2.286	5.372
	[0.389]	[0.772]	[0.591]	[0.603]	[0.937]	[0.792]
A household member took a loan (0/1)	0.056***	0.039	0.071***	16.752	29.803	16.979
	[0.008]	[0.241]	[0.007]	[0.306]	[0.253]	[0.405]
Coefficient of variation of rainfall	-0.003***	-0.003	-0.004***	-2.168***	-4.332	-2.576***
	[0.000]	[0.662]	[0.000]	[0.000]	[0.333]	[0.000]
LOG of fertilizer price in Naira per Kg	-0.019	-0.017	-0.018	-29.802	-38.036	-25.033
	[0.588]	[0.674]	[0.712]	[0.238]	[0.262]	[0.511]
Share of total land cultivated allocated to grains crops	0.002	0.000	0.004**	1.089	-0.048	2.597*
	[0.156]	[0.829]	[0.040]	[0.231]	[0.963]	[0.069]
Share of total land cultivated allocated to legumes crops	0.001	0.001	0.003+	1.046	0.880	2.372*
	[0.457]	[0.505]	[0.132]	[0.294]	[0.526]	[0.099]
Share of total land cultivated allocated to tubers crops	0.001	-0.000	0.003*	0.847	-0.074	1.960+
	[0.507]	[0.771]	[0.064]	[0.363]	[0.943]	[0.149]
Share of total land cultivated allocated to oil crops	0.001	0.000	0.001	1.269	0.581	1.462
	[0.559]	[0.849]	[0.624]	[0.241]	[0.596]	[0.411]
Share of total land cultivated allocated to horticulture crops	0.002*	0.001	0.004*	1.686*	0.922	2.298
	[0.078]	[0.374]	[0.086]	[0.073]	[0.356]	[0.153]
Share of total land cultivated allocated to cotton	-0.001		0.001	-6.547**		-6.974*
	[0.546]		[0.768]	[0.023]		[0.062]

Table 2-14 (cont'd)

Urban dummy variable (0/1)	0.085***	0.091***	0.052	45.380**	60.573**	16.485
	[0.004]	[0.007]	[0.225]	[0.015]	[0.013]	[0.487]
Household Distance in (KMs) to Nearest Market	-0.002***	-0.000	-0.003***	-1.058***	-0.399	-1.191***
	[0.000]	[0.467]	[0.000]	[0.000]	[0.332]	[0.000]
Year 2010 (0/1)	0.010	0.023	-0.000	15.427	36.245**	-3.478
	[0.469]	[0.279]	[0.994]	[0.181]	[0.028]	[0.837]
Zone dummies						
North east	0.089*		0.064	71.353**		55.221*
	[0.077]		[0.243]	[0.011]		[0.073]
North west	0.317***		0.302***	175.648***		167.472***
	[0.000]		[0.000]	[0.000]		[0.000]
South east	-0.182***	0.280***		-48.238	236.303***	
	[0.002]	[0.000]		[0.231]	[0.000]	
South south	-0.251***	0.180***		-178.055***	118.116***	
	[0.000]	[0.004]		[0.000]	[0.008]	
South west	-0.438***			-277.213***		
	[0.000]			[0.000]		
Observations	5,083	1,785	3,298	5,083	1,785	3,298

Source: Generated by authors using LSMS data

Note: ***, **, *, and + indicate that the corresponding regression coefficients are statistically significant at the 1%, 5%, 10%, and 15% levels, respectively. Model estimated using partial MLE estimation method. Pvalues based on clustered standard errors between brackets.

Other factors that significantly affect fertilizer purchase are as expected such as education of the household head with a positive and significant effect in both north and south of Nigeria. Landholding size effect is significant and positive only in the north of Nigeria, while it is negative but not significant in the south. Crop sales affect, positively, but not significantly fertilizer purchase decision. Similarly household distance to nearest market, which proxy for transaction costs of acquiring fertilizer has negative effect on fertilizer purchase decision but is only significant in the north.

2.9. Conclusion

Many believe that Sub-Saharan African farmers' increasing their purchase of external inputs such as fertilizer, seed, and pesticides/herbicides can bring a welcome increase in yields. It has also been observed (Sheahan and Barrett 2014), and echoed in our paper, that the purchase of these external inputs is widespread in SSA. There had not been a systematic exploration of how farmers are paying for these inputs – in particular, what were the relative roles of two sources of cash to pay for inputs (inter alia) - credit (informal and formal) and own cash income. This paper systematically delved into nationally representative datasets for four countries in SSA with widely varying characteristics (Malawi, Nigeria, Tanzania, Uganda) and examined the roles of these sources.

While the literature emphasized that with the reduction or elimination of parastatal agrarian banks formal bank credit is seldom or never available to Sub-Saharan African farmers for inputs, there was explicitly or implicitly in the literature the working hypothesis that farmers used traditional tied credit with output and input traders, and other sources of informal credit to finance their purchase of external inputs for non-contract farming situations. For cash contract-farming situations and cash cropping in general, the working hypothesis in much of the literature is that processors front inputs or cash for inputs to farmers.

By and large, our paper contradicted these "common wisdoms" concerning the use and role of credit in input purchase. First, we found that very few farmers use ***any*** form of credit, formal or informal to finance external input purchase. Second, we found that "tied" credit-output relations are very rare and very minor in external inputs, but especially among smaller farmers in poorer places. What is still significant is tied labor-output markets where local workers advance labor and are paid at the harvest, largely ignored in the literature. Third, we found that generally "traditional cash crop farmers" rarely receive credit from processors, except in a few enclaves like larger tobacco farmers in Tanzania.

Furthermore, we found econometrically that nonfarm self-employment (but not wage employment) plays a significant and positive role in inputs purchase decision, especially given the limited availability of credit for agricultural purposes. Farmers seem to use loans to start nonfarm enterprises (and finance consumption) and plow the cash partly back into their farm input needs.

These findings do not reflect on or test whether farmers face credit constraints; the fact that farmers use very little credit, informal or formal, for farm inputs, does not inform researchers or policymakers whether the farmers have too little access to credit. What we can say from the data is that nonfarm employment is providing a major source of cash that currently far eclipses use of credit for inputs purchases. When farmers take loans, they mainly use the funds to start nonfarm enterprises. They then often use nonfarm income cash to buy farm inputs. That appears to imply that farmers see that employment as a crucial cash source to meet their farm needs. This implies that rural development policies and programs that spur broad development of the rural nonfarm sector, in manufacture and services, would benefit farm input purchase and thus productivity and food security, and certainly be an important complement to credit policies and programs.

REFERENCES

REFERENCES

ADESINA, A. A. 1996. Factors affecting the adoption of fertilizers by rice farmers in Cote d'Ivoire. *Nutrient Cycling in Agroecosystems,* 46**,** 29-39.

BARDHAN, P. K. 1980. Interlocking factor markets and agrarian development: A review of issues. *Oxford Economic Papers,* 32**,** 82-98.

BARDHAN, P. K. 1984. *Land, labor, and rural poverty: Essays in development economics*, Columbia University Press.

BESLEY, T. 1994. How do market failures justify interventions in rural credit markets? *The World Bank Research Observer,* 9**,** 27.

BINSWANGER, H. P. & KHANDKER, S. R. 1995. The impact of formal finance on the rural economy of India. *The Journal of Development Studies,* 32**,** 234-262.

BINSWANGER, H. P. & ROSENZWEIG, M. R. 1986. Behavioural and material determinants of production relations in agriculture. *The Journal of Development Studies,* 22**,** 503-539.

BINSWANGER-MKHIZE, H. P. & RUTTAN, V. W. 1978. *Induced Innovation: Technology, Institutions, and Development*, Johns Hopkins University Press.

BOSERUP, E. 2005. *The conditions of agricultural growth: The economics of agrarian change under population pressure*, Transaction Publishers.

CHAMBERLAIN, G. 1982. Panel data. National Bureau of Economic Research Cambridge, Mass., USA.

CHAO-BÉROFF, R. 2014. Global Dynamics in Agricultural and Rural Economy, and its Effects on Rural Finance. *Finance for Food.* Springer.

CONNING, J. & UDRY, C. 2007. Rural financial markets in developing countries. *Handbook of agricultural economics,* 3**,** 2857-2908.

COX, N. J. 2006. WINSOR: Stata module to Winsorize a variable. *Statistical Software Components*.

CROPPENSTEDT, A., DEMEKE, M. & MESCHI, M. M. 2003. Technology Adoption in the Presence of Constraints: the Case of Fertilizer Demand in Ethiopia. *Review of Development Economics,* 7**,** 58-70.

DE GEUS, J. G. & DE GEUS, J. G. 1967. Fertilizer guide for tropical and subtropical farming. Cambridge Univ Press.

DERCON, S. & CHRISTIAENSEN, L. 2008. Consumption risk, technology adoption and poverty traps: evidence from Ethiopia. *World Economy & Finance Research Programme Working Paper*.

DORWARD, A. R., KIRSTEN, J. F., OMAMO, S. W., POULTON, C. & VINK, N. 2009. Institutions and the agricultural development challenge in Africa. *Institutional economics perspectives on African agricultural development*, 1.

FAFCHAMPS, M. 1999. Market institutions in sub-Saharan Africa: Theory and evidence. *MIT Press Books,* 1.

GREEN WILLIAM, H. 2000. Econometric analysis. *Forth Edition, Prentice Hall International, Inc, New York University*.

HAGGBLADE, S., HAZELL, P. & REARDON, T. 2010. The rural non-farm economy: Prospects for growth and poverty reduction. *World Development,* 38**,** 1429-1441.

HAZELL, P. B. & RAHMAN, A. 2014. *New directions for smallholder agriculture*, OUP Oxford.

KEY, N. & RUNSTEN, D. 1999. Contract farming, smallholders, and rural development in Latin America: the organization of agroprocessing firms and the scale of outgrower production. *World development,* 27**,** 381-401.

KHERALLAH, M., DELGADO, C. L., GABRE-MADHIN, E. Z., MINOT, N. & JOHNSON, M. 2002. *Reforming agricultural markets in Africa: Achievements and challenges*, Intl Food Policy Res Inst.

MATSUMOTO, T. & YAMANO, T. 2011. The Impacts of Fertilizer Credit on Crop Production and Income in Ethiopia. *In:* YAMANO, T., OTSUKA, K. & PLACE, F. (eds.) *Emerging Development of Agriculture in East Africa.* Springer Netherlands.

OSENI, G. & WINTERS, P. 2009. Rural nonfarm activities and agricultural crop production in Nigeria. *Agricultural Economics,* 40**,** 189-201.

POULTON, C., KYDD, J. & DORWARD, A. 2006. Overcoming Market Constraints on Pro‐Poor Agricultural Growth in Sub‐Saharan Africa. *Development Policy Review,* 24**,** 243-277.

REARDON, T. & MERCADO-PETERS, M. Self-financing of rural household cash expenditures in Burkina Faso: The case of net cereal buyers. Finance and Rural Development in West Africa. Proceedings of the 12th rural economy seminar (CIRAD, Ohio State University, CNCA, CEDRES, INERA) in Ouagadougou, 1993.

REARDON, T., TAYLOR, J. E., STAMOULIS, K., LANJOUW, P. & BALISACAN, A. 2000. Effects of non‐farm employment on rural income inequality in developing countries: an investment perspective. *Journal of agricultural economics,* 51**,** 266-288.

SADOULET, E. & DE JANVRY, A. 1995. *Quantitative development policy analysis*, Johns Hopkins University Press Baltimore.

SHEAHAN, M. & BARRETT, C. B. 2014. Understanding the agricultural input landscape in sub-Saharan Africa: Recent plot, household, and community-level evidence. *World Bank Policy Research Working Paper*.

SMALE, M., KUSUNOSE, Y., MATHENGE, M. K. & ALIA, D. 2016. Destination or Distraction? Querying the Linkage Between Off-Farm Work and Food Crop Investments in Kenya. *Journal of African Economies*.

SWINNEN, J. F. & MAERTENS, M. 2014. Finance through food and commodity value chains in a globalized economy. *Finance for Food.* Springer.

TAKESHIMA, H. & LIVERPOOL-TASIE, L. S. O. 2015. Fertilizer subsidies, political influence and local food prices in sub-Saharan Africa: Evidence from Nigeria. *Food Policy,* 54**,** 11-24.

TAKESHIMA, H. & NKONYA, E. 2014. Government fertilizer subsidy and commercial sector fertilizer demand: Evidence from the Federal Market Stabilization Program (FMSP) in Nigeria. *Food Policy,* 47**,** 1-12.

TSCHIRLEY, D. 2009. Input credit and extension. *Organization and Performance of Cotton Sectors in Africa: Learning from Reform Experience. D. Tschirley, C. Poulton and P. Labaste (ed). Washington DC, World Bank***,** 73-85.

WOOLDRIDGE, J. M. 2010. *Econometric analysis of cross section and panel data*, MIT press.

3. ESSAY 2: SUSTAINING INPUT ON CREDIT THROUGH DYNAMIC INCENTIVES AND INFORMATION SHARING: LESSONS FROM A FRAMED FIELD EXPERIMENT

3.1. Introduction

Increasing agricultural productivity is key for the structural transformation of societies and for poverty reduction (Johnston and Mellor, 1961). One potential mechanism to increase agricultural productivity is the increased use of modern technologies, including fertilizer. While there are signs of an increase in fertilizer use in countries with subsidy programs or other concerted input support strategies, fertilizer use in Sub Saharan Africa (SSA) generally remains low (Sheahan and Barrett, 2014).

Severe capital and credit constraints are one key reason for the low fertilizer use rates among smallholder farmers in many developing countries. Even when farmers believe that fertilizer use is profitable, they may be unable to purchase fertilizer because they lack cash, cannot obtain credit (e.g. due to lack of collateral) or cannot obtain fertilizer locally (Kelly et al., 2007). Thus, input on credit has been identified as a potential way to increase farmers' access to and use of modern inputs by solving both the credit and accessibility or availability constraints.

Despite the potential benefits of providing inputs on credit, market conditions often do not encourage the private sector to provide such credit to smallholder farmers (Kelly et al., 2003). Generally, credit markets in rural SSA are characterized by market failures associated with imperfect information in the presence of risk (Dorward et al., 1998, Poulton et al., 1998, Sadoulet, 2005, Tedeschi, 2006). These failures persist because institutions for contract enforcement are weak, increasing the potential for high default rates among farmers. Knowing this, input suppliers are reluctant to provide inputs on credit to farmers. This leads to the missing market problem as both the input provider and the farmer loose the potential gain from trade by not completing the transaction. However, input provision on credit can potentially be facilitated if it is commonly known that failure to repay implies future inability to get input on credit. Essentially, when the interaction is repeated over an indefinite period of time, input on credit arrangements can be sustained as long as farmers value gains from future access to fertilizer more than the temporary gain from reneging on current debt contracts, and if the threat of being prevented from accessing future input on credit is credible.[12] However, when multiple input sellers exist in the market, how can one ensure credibility of the threat since farmers can potentially approach another provider after defaulting? This paper adapts a game theoretic model drawn from the microfinance literature to answer this question, and then tests the model predictions using data from a framed field experiment conducted with farmers in rural Nigeria.

This paper makes several important contributions to the literature on agricultural input loan provision (by private input suppliers) in developing countries. First, there is no study (the authors are aware of) that has focused explicitly on strategic default in cases in which private input suppliers would sell inputs on credit to farmers and collect payment after harvest. While such input on credit arrangements share some characteristics of microfinance, they also have their peculiarities such as being in-kind, less prone to moral hazard, and mostly threatened by strategic default. Consequently, we build on the microfinance literature and develop ideas about additional measures that can help alleviate strategic default problems in input on credit arrangements. Specifically, we extend previous work on the role of dynamic incentives in addressing strategic default by exploring the importance of information sharing among credit suppliers for the effectiveness of dynamic incentives in a rural developing country setting. This

[12] This is a direct implication of the Folk theorem.

paper is timely given the recent focus by policy makers and development practitioners on private sector led approaches to input market development in developing countries. It informs the likely strategies that are necessary to encourage the development of private sector led input on credit provision. This paper also adds to the limited number of studies that use framed field experiments, and is also one among very few examples of an empirical application of the concepts of credit information sharing and dynamic incentives mechanisms. With a framed experiment with multiple rounds, we are able to explore the dynamics of the relationship between credit takers and suppliers over a longer-term period than is generally possible in most actual experiment.

The rest of the paper is organized as follow. In section 2, we provide a summary of the relevant literature on strategic default. Section 3 presents the theoretical framework from which empirically testable hypotheses are drawn. Section 4 describes the experimental design used to gather data for the empirical analysis and section 5 presents and discusses the results of the empirical analysis. We conclude with a summary of the key findings and policy implications in section 6.

3.2. Dealing with strategic default

Strategies to overcome moral hazard and strategic default issues inherent to offering uncollateralized loans to poor people in developing countries is a longstanding problem in the microfinance literature. One strand of the literature focuses on the use of group lending and joint liability as a mechanism to overcome those issues. This approach requires borrowers to sort themselves in groups. Loans are made to individuals, but the group as a whole is held jointly liable in case of default. The mechanism effectively transfers screening and monitoring costs from the bank to borrowers, providing an effective way for banks to reduce adverse selection, moral hazard and enforcement problems. However, the success of group lending becomes limited when we care about the poorest (Armendáriz de Aghion and Morduch, 2000), or when the group is either non-existent or too large to have the necessary information to ensure repayment (Tedeschi, 2006). Therefore it has become a subject of interest to find mechanisms through which individual non-collateralized lending to the poorest could be sustained.[13]

There is a relatively large literature, with an early contribution from Besley (1995), which has discussed dynamic mechanisms through repeated interaction and reputation mechanisms as alternative ways to overcome strategic default without relying on group lending based on joint liability. The fundamental idea is that when a borrower depends on successive loans to keep his business functional, the threat of being denied future loans can provide incentives to avoid default in current period (Hulme and Mosley, 1996, Armendáriz de Aghion and Morduch, 2000, Tedeschi, 2006).

Tedeschi (2006) focused on strategic default and default due to negative economic shocks and showed how dynamic incentives, in the form of additional or future loans, can reduce strategic default without relying on the group incentives used in the microfinance literature. Using a model based on a single microfinance institution ("lender") and a group of micro entrepreneurs ("borrowers") who may well be farmers, he models the repeated lender-borrower relationship by endogenizing the amount of time that a borrower who defaults must remain without a loan. He shows that the optimal length of the punishment phase can be less than infinity, especially when an individual has much to gain from the lending relationship. He notes

[13] Details about the mechanism and limitations of group lending are provided in Stiglitz, 1990, Besley and Coate, 1995, Armendáriz de Aghion and Morduch, 2000, etc.

that punishment should instead only be sufficiently long to prevent a borrower from strategic default, but not so long as to unduly punish the borrower that experiences a negative economic shock. An important aspect of this model is that it assumes the presence of a single lender or perfect sharing of default information if multiple lenders are present. But in reality there are usually several lenders and information is rarely perfectly shared amongst them. Tedeschi's paper does not discuss explicitly how this potential exchange of information between lenders may affect repayment behavior, nor does it empirically test the predictions.

As competition between lenders increases, the effectiveness of the dynamic incentive is weakened because the borrowers can take advantage of this competition and get loans from various sources. In such a case, coordination between lenders, in terms of credit information exchange can be an effective discipline device to mitigate various forms of moral hazard, and reduce strategic default (Padilla and Pagano, 1997, Padilla and Pagano, 2000). For example, communication and exchange of information was essential for the functioning of the *merchant guilds* that facilitated trade during the late medieval period (Greif et al., 1994), and the *Coalition* that enabled 11th century Maghribi traders' to benefit from employing overseas agents despite the commitment problem inherent in these relations (Greif, 1993). Ghosh and Ray (1999) also show the importance of communication between lenders in solving the issue of strategic default in individual lending. Moreover, there is a growing number of recent studies that provide theoretical and empirical evidence on the effect of credit information systems for mitigating problems of adverse selection and moral hazard in credit markets (McIntosh and Wydick, 2009, Padilla and Pagano, 1997, Padilla and Pagano, 2000, Vercammen, 1995). The general conclusion is that credit information sharing substantially increases lending, and decreases borrowers' default (Djankov et al., 2007, Jappelli and Pagano, 2002, Luoto et al., 2007, de Janvry et al., 2010).

In particular, Luoto et al. (2007)) and de Janvry et al. (2010) use field experiment data from a microfinance lender, *Génesis Empresarial*, one of the lending institutions participating in a credit bureau that was implemented across Guatemala in 2001. The credit bureau (CREDIREF) was established to solve the problem of multiple loan contracting and hidden debt exacerbated in the late 1990s by the growth in the number of microfinance institutions (MFIs) in Guatemala. By allowing for positive and negative information sharing between participating lenders, CREDIREF was proved to have positive screening and incentive effects. Essentially, the 39 branches of *Génesis Empresarial*, received the hardware and software necessary for the credit bureau in nine different waves between August 2001 and January 2003, providing a natural experiment to test the effects of the credit bureau on the lending portfolio of Génesis. Luoto et al. (2007) took advantage of this to identify the branch-level impacts from the screening effect of the bureau on loan delinquency rates. Their results indicate a reduction in default of approximately two percentage points after the bureau was implemented in branch offices. de Janvry et al. (2010) exploited the lack of awareness about the credit bureau among borrowers to isolate the incentive effects of bureaus via a field experiment. In the experiment, 573 Génesis borrowing groups were randomly selected from within 7 branches (the branches themselves randomly selected through stratified sampling) to receive a course that highlighted the existence and workings of the bureau.[14] The training course focused both on the positive repercussions of a

[14] A preliminary field survey with 184 borrowers in six branch offices of Génesis found that borrowers were remarkably poorly informed as to the presence of the credit bureau. This lack of awareness of the bureau at the time of its implementation was helpful in trying to decompose the different effects of a credit bureau empirically.

bureau (increased access to outside credit for those with good borrowing records) as well as the negative (heightened adverse consequences of failing to repay), and provided specific information about lenders using the bureau, when information was checked, and on whom. The results of their empirical analysis indicate that while new clients recruited after the bureau have better repayment rates, this improvement in default was counteracted by a doubling in the probability of serious delinquency among ongoing borrowers whose loan sizes grew sharply subsequent to the use of the bureau. However, de Janvry et al. (2010) are not able to explore the dynamics of the screening effect of the bureau de change, which our framed field experiment allows.

In this paper, we develop a theoretical model that characterizes ex-post moral hazard, or strategic default in the context of individual input loans made by private input suppliers to farmers in developing countries. Drawing insight from the models in Padilla and Pagano (2000) and McIntosh and Wydick (2009) the paper features a simple repeated game model of input credit and stresses the importance of information sharing amongst lenders, for farmers' repayment decision. The model also embeds the presence of a productivity shock that may affect farmers' repayment abilities or incentives. We then test the model predictions in the field using lab-in-the-field experimental methods referred to as a framed field experiment by Harrison and List (2004). The experimental design allows us to explore not only the contemporaneous effect of information exchange but also the dynamics of the interactions between farmers and input providers, which (as far as we are aware) has not been explored yet within the context of a field experiment in a developing country.

3.3. Theoretical framework and experimental hypotheses

3.3.1. A simple model of input on credit

Our model considers a repeated matching game between a set of firms $n_s = \{1, \dots, N_s\}$ and a set of farmers $n_b = \{1, \dots, N_b\}$. By assumption, the farmers need to buy inputs for agricultural production but they do not have the capital to pay upfront and thus must buy it on credit. The firms are agricultural input dealers or brokers who have inputs that they seek to sell. They consider selling on credit in addition to cash sales in order to maximize the volume of sales.[15] We fix the price of the input and assume all farmers receive the same input bundle so that brokers maximize profits by selling more inputs bundles to farmers with a higher likelihood of repaying. In each stage of the game, each broker is matched with every farmer and they play a 2-player sequential stage game. For each game the firm decides, at the beginning of the agricultural season, whether or not he should make an offer of input on credit to the farmer. After harvest, the farmer decides whether to repay or not. We assume the use of the agricultural input is always profitable i.e. that agricultural return is always higher with the use of the input than without. Net return without using the input is denoted R_{none} but there is a random productivity shock $\eta = \{Good, Bad\}$ that is realized after the input has been acquired and used.[16] The net return to the use of the input, $R_\eta = \{R_{good}, R_{bad}\}$, is assumed to be lower when the shock is bad and higher when it is good (i.e., $R_{good} > R_{bad} > R_{none}$)[17].

---

[15] Note that this model can be generalized to any relation between demanders and suppliers of credit.

[16] This can be thought of as a weather shock. Good weather implies higher productivity *ceteris paribus*.

[17] Though it can be seen as restrictive, the assumption that returns to the use of inputs is higher than returns from not using the inputs (even when the weather shock is bad) is made to focus on farmers' repayment decision without getting into input use decisions. This can be seen as a sort

In every period of the game, the firm's strategy can be described by a function $\sigma_i^S\colon \mathrm{H}^t \rightarrow \{Offer, Not\ offer\}$ for all farmers i $\in$ n_b, where $\mathrm{H}^t = \{\mathrm{H}^t_{Public} \cup \mathrm{H}^t_{Private}\}$ is the set of information available to firm s, and which contains, up to time t-1, the repayment history of the all the farmers including farmer i. Notice that we distinguish between public and private information. The public information set for a firm j $\in$ n_s contains the repayment history for all farmers that firm j has not made an offer to in past periods and therefore does not know privately how they behaved in those periods. The private information set contains repayment information about those farmers firm j has made offers to in past periods and therefore has observed farmer repayment behavior. The farmer's strategy in each period (given he receives an offer) is a mapping σ_j^B from the realization of productivity shock η to the set of possible actions $\{Renege, Not\ renege\}$, for all firms j from which the farmer took an offer. When he does not receive any offer, his set of possible actions is the empty set.[18]

Finally, for each initiated transaction with a farmer, the firm gets a payoff of P-c > 0 if the farmer does not renege, and –c < 0 if the farmer does renege. P is the price at which the input is being sold to the farmer, and c is the cost of the input to the firm. The firm's reservation payoff in case of no transaction with a farmer is 0. We assume that the firm's payoff function in the stage game is additively separable over all the transactions made with farmers in that stage. As for the farmers, they receive a reservation payoff R_{none} if they do not receive an offer in that stage and thus do not use any of the input. If a farmer receives an offer, their payoff function is

---

of incentive payment to ensure enough farmers will take input loans so we can observe their repayment behavior and test the predictions of the model.

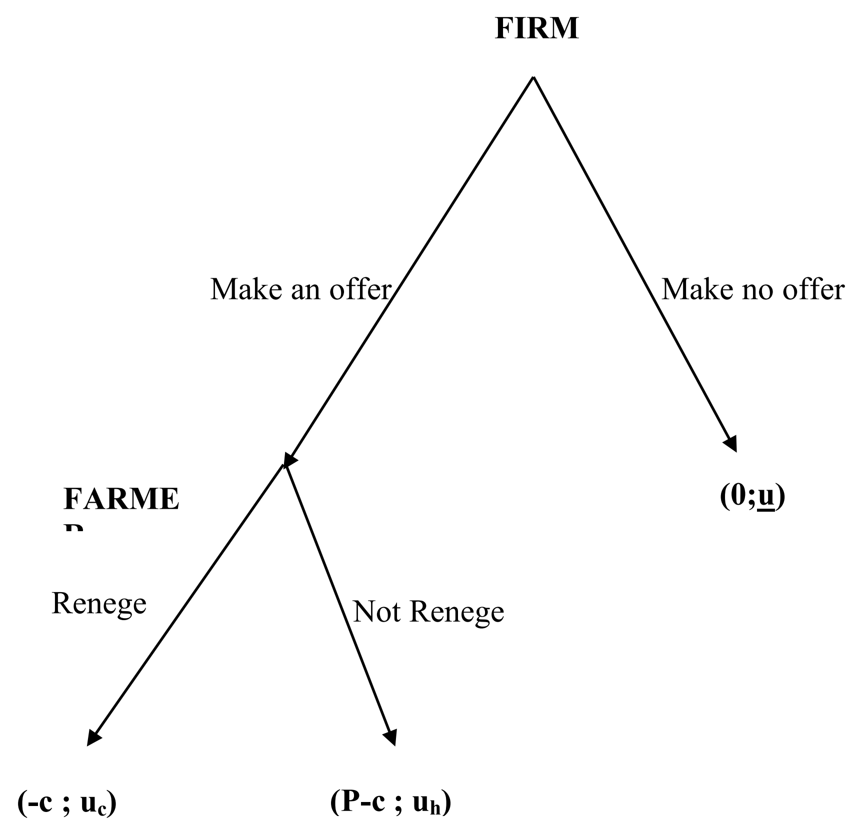
[18] Later, in our experimental design, we impose the constraints that farmers can only accept one offer in each period, and firms can only make offers to a limited number of farmers. These assumptions only simplify the game for the participants without fundamentally changing the implications of the model and the consequent empirical hypotheses.

described by a mapping g: $\{R_{good}, R_{bad}\} \times \{Renege, Not\ renege\} \rightarrow \mathbb{R}$. Their payoff depends on their repayment decision and the realization of the productivity shock. We define $u_{h\eta},$ $u_{c\eta},$ and $\underline{u}$ to be the farmer's state contingent utilities from not reneging, reneging, and not using the agricultural input, respectively.

3.3.2. The missing market problem in a single period game

In the single period case, each matching between farmer and firm in the game described above can be represented by the extensive form game in figure 1. As depicted, the farmers' dominant strategy is to take the loan from any firm that makes him an offer, and then renege. In anticipation of this, the firm's dominant strategy is to not lend in the first place and thus the market collapses (Conning and Udry, 2007). Figure 1 shows that the Subgame Perfect Equilibrium for this game is (no offer, renege) which gives a payoff profile of (0, $\underline{u}$). This is clearly pareto inferior to the (offer, not renege) option which results in a payoff of (P-c ; u_h). This happens irrespective of the realization of the productivity shock. Also, since there is no previous stage, the information available to the firm at the beginning of the game is the empty set. Typically, the loan might be secured or the firm could enforce the contract through the legal system causing the farmer's renege payoff to be greater than u_c. If this is high enough then the farmer has an incentive not to renege and the firm would make the offer and we would get to the pareto superior outcome. However, in our context, there is a high potential for default due mostly to the fact that the legal procedures for enforcing contracts are critically weak in most developing countries (Kelly et al., 2003). Thus, the input provider and the farmer both loose the potential gain from trade.

Figure 3.1: Extensive form representation of the farmer-trader theoretical game



3.3.3. Enforcement of the input-on-credit contract using dynamic incentives

As noted by Conning and Udry (2007), if the above interaction is repeated, it may be possible to generate incentives for the farmer to repay in every period, provided that the threat of no further loan activity is credible and sufficiently punishing. To illustrate, consider an infinitely repeated game where each round is the above stage game. Recall that, while in reality famers and firms do not enjoy an infinite lifespan, an infinitely repeated game is equivalent to a finite horizon model with a constant probability of terminating the relationship every period (Greif, 1993, Mas-Colell et al., 1995). In each period of the game, the threat of non-renewal implies that each firm is playing the following strategy with each farmer i they are matched with:

$$\sigma_i^S : \begin{cases} if\ \mathrm{H}_i^t\ indicates\ No\ past\ default\ behavior\ by\ farmer\ i, make\ him\ an\ offer; \\ \qquad\qquad Otherwise, make\ no\ offer\ to\ farmer\ i \end{cases}$$

Recall that $\mathrm{H}^t = \mathrm{H}^t_{Public} \cup \mathrm{H}^t_{Private}$ contains both public and private information about the farmer. In this model, where the market is competitive (several firms and several farmers), the public information aspect is crucial for sustaining cooperation, unless the firm and the farmer have an exclusive relationship. When farmers have the possibility to take input credit offers from other firms in subsequent periods, the expected punishment from default is less severe and may not be able to deter default. However, if default information is shared publicly amongst firms and all firms agree to collectively punish a defaulter, the farmer is forced to behave as if in an exclusive relationship with the firm.[19]

The farmers' response to the collective punishment is described as follows. At any period t, the present value of the lifetime expected utility to the farmer from never defaulting ($\mathrm{V_h}$) given the realization of the productivity shock $\eta = \{Good, Bad\}$ is:

$$\mathrm{V}_{\mathrm{h}\eta} = \mathrm{u}_{\mathrm{h}\eta} + \frac{\delta}{1-\delta} \qquad (1)$$

where δ and $\mathrm{u_h}$ are, respectively, the discount factor and payoff from not reneging as defined earlier. $E_\eta u_{\mathrm{h}}$ is the expected utility of the farmer for periods when he does not renege.

The present value of the lifetime expected utility from a one-time default is:

$$\mathrm{V}_{c\eta} = \mathrm{u}_{c\eta} + \frac{\delta}{1-\delta}\left[\theta E_\eta \underline{u} + (1-\theta) E_\eta u_{\mathrm{h}}\right], \quad \eta = \{Good, Bad\} \qquad (2)$$

where θ is the probability that a defaulting farmer gets punished. θ is affected by the number of input sellers in the market and the efficiency with which information about defaulters

---

[19] Note that the collective punishment assumes that firms in competition have incentive to punish farmers who defaulted any of the firms even if they have not been cheated on personally. Greif, 1993, and Kandori, 1992 describe reasons and institutions that can guarantee this.

flows between firms so that they can exclude the farmer from consideration. If θ=1, that implies information flows perfectly between firms and it is guaranteed that a defaulter will never get input on credit from any other firm in subsequent periods. Likewise, if θ=0, information does not flow between private firms and farmers can default and still get inputs on credit from other firms, depending on how many input firms there are. Eventually, the private information set alone will translate into a value of $\theta_{private}$ that is lower than when the firms have access to both the public and private information history.

According to the Nash Folk Theorem (Fudenberg and Tirole, 1991, Mas-Colell et al., 1995), cooperation between farmers and input suppliers can be achieved under the assumptions described above, as long as farmers are patient enough (δ is high enough).

The sustainability condition requires that:

$$\mathbf{V}_{\mathbf{h}\eta} \geq \mathbf{V}_{\mathbf{c}\eta}, \qquad \eta = \{Good, Bad\} \tag{3}$$

$$\mathbf{u}_{\mathbf{h}\eta} + \frac{\delta}{1-\delta} E_\eta u_{\mathrm{h}} \geq \mathbf{u}_{\mathbf{c}\eta} + \frac{\delta}{1-\delta}\left[\theta E_\eta \underline{u} + (1-\theta) E_\eta u_{\mathrm{h}}\right] \tag{4}$$

This is equivalent to:

$$\delta \geq \frac{1}{1+\theta\frac{\left(E_\eta u_{\mathrm{h}} - E_\eta \underline{u}\right)}{\mathrm{u}_{c\eta} - \mathrm{u}_{\mathrm{h}\eta}}} = \delta_\eta^* \tag{5}$$

Equations 5 demonstrates that in any period, only farmers with a discount factor greater than δ_η^* will not default and trade is sustainable only with those farmers. Assuming that the productivity shock is independently and identically determined in each round, the per-period forgone benefit from continuing to get inputs on credit ($E_\eta u_{\mathrm{h}} - E_\eta \underline{u}$) is fixed in each future period. Therefore, the minimum discount rate required to sustain trade depends mostly on how big the farmers' immediate gain from defaulting ($\mathrm{u}_{c\eta} - \mathrm{u}_{\mathrm{h}\eta}$) is in the current period. In particular, $\mathrm{u}_{c\eta} - \mathrm{u}_{\mathrm{h}\eta}$ can be interpreted as the opportunity cost of repaying for the input received on credit in the current period, and is a function of the realization of the productivity shock in that period. For risk averse farmers, $\mathrm{u}_{\mathrm{c},Good} - \mathrm{u}_{\mathrm{h},Good} < \mathrm{u}_{\mathrm{c},Bad} - \mathrm{u}_{\mathrm{h},Bad}$ and therefore, in the good state of the nature, δ_η^* is lower than in bad state of nature, *ceteris paribus*.

Many empirical hypotheses can be derived from equation 5. We focus on 2 main ones in this study:

Hypothesis 1: Equation 5 indicates that as θ increases, δ_η^* decreases for all η. That is, as the probability of being recognized as a defaulter by other firms increases, the minimum discount rate required for the farmer not to default decreases. This probability is related to the credibility and sufficiency of the punishment threat, and is determined by many factors such as the number of input suppliers and the degree of communication between them. This leads to the following testable hypothesis: "**As communication and exchange of information is facilitated amongst input suppliers, the probability of the farmer being caught and ostracized increases, and therefore, the probability of default by farmers decreases**."

Hypothesis 2: Equation 5 also indicates that as $(\mathrm{u}_{\mathrm{c}\eta} - \mathrm{u}_{\mathrm{h}\eta})$ increases, δ_η^* increases for all η. That is, as the opportunity cost of repaying increases, the minimum discount rate required for the farmer not to default increases. This leads to the second testable hypothesis: "**In the bad state of the nature (when productivity is lower due to some productivity shock), the probability of default by farmers receiving inputs on credit increases**."

3.4. Experimental design and procedures

Given that input-on-credit arrangements are not commonly observed in the setting of interest, it is difficult, if not impossible, to collect observational data to test our hypotheses.

Therefore, we conduct a lab-based field experiment using randomly selected farmers in 10 different villages in Kwara State, Nigeria (see Table 1). The experiment is designed to simulate a multiple round market for inputs-on-credit and test the above hypothesized communication and profitability shock effects.

Table 3-1: Experiment Villages in Kwara State, Nigeria

Local Government (LGA)	Village Name	Communication	Number of rounds
PATIGI	AGBOORO	Yes	10
PATIGI	CHAKYAGI	No	10
EDU	CHEWURU	Yes	11
EDU	CHIKANGI	No	10
EDU	CHIKANGI TIFIN	Yes	11
EDU	EFFAGI	No	10
EDU	GBARIGI	Yes	11
EDU	KPANGULU	No	10
PATIGI	KUSOGI GANA TSWALU	Yes	10
PATIGI	SHESHI TASHA	No	10

To test the communication and exchange of information effect, five out of the ten study villages were randomly selected to receive a communication treatment. In those five villages information regarding individual farmer default behavior was relayed to all creditors resulting in increasing the probability that a farmer is identified as a potential future defaulter. In the five non-communication treatment villages, creditors only knew the default behavior of the farmers to whom they made loans. Comparing farmers' behavior in the communication treatment to that in the non-communication treatment tests for the hypothesized communication effect.

To test hypothesis 2 – the impact of productivity and profitability on default behavior – a round-level treatment was implemented. Specifically, in each round the weather could take on one of two states – good or bad. If the weather was good, productivity and profitability of farmers were high, and if the weather was bad productivity and profitability of farmers were low.

Recall that the profitability shock hypothesis assumes that a higher net profitability reduces farmers' incentives to default. Given this we expect lower levels of farmer default in rounds with good weather than in rounds with bad. In each round, the weather state was determined by the flip of a coin after credit decisions were made, but before repayment.

Each experimental session (one per village) involved 20 participants. Participants were randomly assigned to be either a farmer (who might receive inputs on credit), or a paid broker of an agro dealer (henceforth, agro broker).[20] Allowing the brokers to have full discretionary power over input credit allocation to farmers, and making them incur a given proportion of the cost of the inputs in case of default by farmers, we can see brokers behavior in the game as representative of the input sellers. Therefore, we use the terms input dealers, inputs sellers, and brokers interchangeably in the remaining of this article.[21] Each session had 4 agro brokers and 16 farmers and participants remained in the same role for the entire experiment. Each experimental session consisted of 10 or 11 rounds. After the 9th round in each village, a coin was flipped at the end of each round to determine whether to continue an additional round of the game or not. This is to establish a random stopping point of the game and reduce farmers' incentive to behave opportunistically in the last rounds. Interestingly, we never had more than 11 rounds in any village, and all the 3 villages for which the experiment went for an 11th round were communication treatment villages. Each round represents an agricultural season and the

[20] We use the term broker to make the hypothetical situation more realistic to farmers. While they have not had any real experience as agro dealers, they are familiar with the concept of agro broker because a fertilizer company in the area has based its distribution system on locally based brokers who act as local retailers of the inputs in those villages.
[21] We recognize that this ignores the potential coordination problem between the brokers and the real input sellers. But we argue that by making the brokers residual claimant and by transferring all the credit allocation decision to them, incentives are aligned enough for us to interpret the brokers behavior as representing input sellers' behavior.

decisions made by participants were based on simulating the important aspects of actual input credit markets. As such, each round consisted of two periods – a pre-planting period and a post-harvest period. In the pre-planting period, the agro brokers offered inputs on credit to the farmers and the farmers decided which (if any) agro broker offer to accept. In the post-harvest period, farmers' harvest returns were determined (based on weather and input use) and farmers choose whether to repay the agro broker for the input or not. The possible decisions and their payoff implications for agro brokers and farmers are described in the following sections.

3.4.1. Decisions and Payoffs for Agro Brokers

Each of the four agro brokers in each village began each round with 300 kg of fertilizer to potentially be sold on credit to farmers. In the pre-planting period, the broker decided **for each farmer** whether to offer input on credit or not. To simplify the decisions, we assumed that the input comes in bags of 100kg and each farmer only needs 100kg. Therefore, an offer made to a farmer implied 100kg of input offered to the farmer by the broker. This means that the broker could make offers to at most 3 farmers in each round.[22] Once offered, each farmer could accept or decline the offer. In the post-harvest period, agro brokers received payments from the farmers to whom they made input loans. The value of the input loaned was set to N100 per kg. Thus a farmer who borrowed 100kg of fertilizer from an agro broker would be expected to repay N10,000. However, the actual amount received and the agro broker's commission/penalty depends on the farmers' repayment decision. The farmers had the option to: not repay at all (0% of amount owed), partially repay (50% of amount owed), or repay in full (100% of amount

[22] Note that agro brokers did not have to make any offers, but if they did not they would not receive the base salary.

owed). The possible outcomes for an agro broker, from any given farmer who received inputs on credit, are summarized in the Table 2.

Table 3-2: Brokers' Commission/Penalty Schedule

	Description	Amount/Value	
	Amount of fertilizer loaned	0	100kg/N10000
Repay in full	Amount collected	0	N10000
	Broker's commission/penalty	0	N2000
50% repayment	Amount collected	0	N5000
	Broker's commission/penalty	0	-N1500
0% repayment	Amount collected	0	N0
	Broker's commission/penalty	0	-N3000

Overall the agro broker's earnings from input sales during a round consist of two parts. First, a base salary of N3000 – paid if at least one farmer accepted an offer. This base salary was designed to incentivize agro brokers to make offers.[23] Second, the commissions/penalties from the repayment of loans made to farmers (3 or less per broker). As shown in Table 2, the broker receives a N2000 commission for every sale where repayment is complete but a penalty is imposed every time he offers inputs to farmers who do not repay fully. If a repayment is partial the agro broker has to pay a penalty of N1500 to the input dealer. Similarly, if the farmer repays nothing, the agro broker has to pay a penalty of N3000 to the input dealer. Note that, given the penalties, it is possible for the agro broker to lose money in a round. For example, assume that an agro broker makes offers to 3 different farmers and they all accept. The broker thus gets the base salary of N3,000. If all the farmers decide to fully default, the broker loses N3,000 per farmer or

[23] In fact, without this incentive (and because of the N50,000 payment given to ensure non-negative earnings discussed below) agro brokers might choose to sit out the game by not making offers once they made a single loan.

N9,000 total. Overall, the broker has a net loss of N6,000. In order to avoid the possibility that the broker owed us money at the end of the experimental session, every broker was promised N50,000, to be paid at the end of the session, provided that he had made at least one loan in any round. Net payments to agro dealers per round could vary from a loss of N6,000 as illustrated above to a net gain of N9,000 if three offers are accepted and fully repaid.

3.4.2. Decisions and Payoffs for Farmers

As described above, in each round farmers received offers from the agro brokers in the pre-planting period and, given that they received more than one offer, chose which one to accept. Note that, to simplify the game, farmers could only accept fertilizer on credit from one agro broker (100kg). Furthermore, fertilizer was assumed to always be advantageous for farmers in that using it always increased yields and thus payoffs. There was also no mechanism for farmers to get fertilizer in another way. This was done to ensure that all the farmers had the same resources available to them at the beginning of a round/season. In the post-harvest period, the weather for the season was determined via a coin-flip (a single coin flip applied to all farmers and individual farmers were invited to flip the coin) and this, along with whether they received fertilizer, determined harvest yields. As shown in Table 4, harvest yields were represented in terms of monetary returns to investment. Specifically, if the farmer used fertilizer and weather was bad they earned N13,000, while if the weather was good they earned N16,000. If they did not use fertilizer, the returns were much lower (N1,000) and were not dependent on the weather. After learning about the weather and resulting earnings, farmers that had received fertilizer chose a level of repayment (0%, 50%, or 100%). Recall that the fertilizer on credit was worth N10,000 or N100/kg. The possible round earnings for a farmer are shown in Table 3.

Table 3-3: Farmers' payoff structure

Description		Amount/Value	
Amount of fertilizer received (kg)		0	100kg
Low Return to investment (Bad Weather state)		N1000	N13000
if full repayment	Amount paid	0	N10000
	Farmer's net payoff	N1000	N3000
if partial (50%) repayment	Amount paid	0	N5000
	Farmer's net payoff	N1000	N8000
if no repayment	Amount paid	0	0
	Farmer's net payoff	N1000	N13000
High return to investment (Good Weather state)		N1000	N16000
if full repayment	Amount paid	0	N10000
	Farmer's net payoff	N1000	N6000
if partial (50%) repayment	Amount paid	0	N5000
	Farmer's net payoff	N1000	N11000
if no repayment	Amount paid	0	0
	Farmer's net payoff	N1000	N16000

3.4.3. Information Treatment Variation and General Implementation

The communication treatment sessions differed from the non-communication sessions in that the agro brokers were given complete information about all farmers' past repayment behavior in the game. This was done through a record kept publicly on a board in front of all the participants (see table 2). The repayment record board was updated after each round, thus showing each farmer's repayment decision in previous rounds. This implies that when a farmer does not repay the credit taken from a specific broker in a specific round, all other brokers will know about it before they make credit offers in the following round. Farmers in these sessions were informed prior to the start of the game that their repayment behavior would be made public. The default record was presented to participants as shown in Table 4.

Table 3-4: Public repayment records used in treatment villages

Farmers id	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16
Round 1																
Round 2																
Round 3																
Round 4																
Round 5																
Round 6																
Round 7																
Round 8																
Round 9																
Round 10																

The experiment was paper-based in that agro brokers and farmers made decisions using decision sheets (see appendix), but the data was recorded and payment amounts calculated using a computer. A team of six experimenters ran each session. Once all participants were present, the instructions were presented and questions answered. Participants were then separated into farmer and agro broker groups and received the appropriate decision sheets (broker sheet and farmer sheet). To give participants a chance to see the game in action and to ask questions an unpaid practice round was performed. During the experiment all decisions were anonymous in that brokers and farmers were assigned participant numbers and all decisions were entered on paper and communicated to other relevant participants via collection and transcription of decision sheets by the experimenters.

3.5. Results and discussion

3.5.1. General description of the data

As noted above, the experiment involved 16 farmers and 4 brokers per village, in 10 villages for 10 to 11 rounds. Overall, the 40 brokers that participated in the experiment made a total of 1205 input loan offers to farmers (see table 5).

Table 3-5: Statistics about the offers made and received through the game

	Communication	Non-Communication	Total
Total number of offers	614	591	1205
Average number of offers per farmer (amongst farmers who received at least one offer)	1.31	1.36	1.34
Total number of offers actually accepted throughout the game	466 (76%)	426 (72%)	892 (74%)

Source: Generated by authors using survey results

In the communication villages we observed more offers (614) than in the non-communication villages (591). Given that multiple brokers may make offers to the same farmer and farmers can only accept one offer, some offers are necessarily rejected. Farmers, when they received offers during a round, got on average 1.34 offers. This indicates that brokers did not necessarily spread out the offers across all farmers in each round. Consequently, while 1205 offers were made, the total number of offers, actually accepted, was 892 or 74% of the total number of offers made by brokers. Breaking this down by communication treatment, in villages with communication, 76% of offers were accepted whereas in the non-communication villages, only 72% were accepted. Note that, in the communication villages, more offers were made and a higher proportion were accepted resulting in more transactions relative to the non-communication villages. This is an initial, though still weak, indication that communication and exchange of information can allow the market to perform better due to the reduced information asymmetry problem.

In the following sections, we focus on farmers and brokers' behaviors and analyze the role of productivity shocks and communication treatments.

3.5.2. Farmers' behavior

3.5.2.1. Description of farmers' repayment behavior during the game

A key goal of this experiment was to evaluate how communication and exchange of information between brokers, as well as productivity shocks (weather), affect repayment decisions when farmers receive input on credit. Figures 2 and 3 describe the relationship between repayment behavior and our treatment variables. The pooled data contains 892 observations at the farmer level, with 47.3% observations with the good weather state, and 52.2% observations in the communication treatment villages.

Figure 2 (repayment behavior by communication treatment) indicates that the default rate – defined as the proportion of farmers who repay less than 100% – is higher in the no-communication treatment. More precisely, with no communication, 50.23% of farmers repaid half, while 11.27% did not repay anything, making the total default rate 61.5%. In contrast, with communication, the default rate, similarly defined, is 57.3%. This lower default rate in the communication treatment suggests that communication amongst input suppliers likely has a positive effect on farmers' repayment of input loans. However this difference is not statistically different from zero overall ($Pr(|T| > |t|) = 0.202$). But it is strongly significant when the weather is bad. We explore this later in more detail with an econometric model.

Figure 3.2: Histogram of repayment decisions by communication treatment status

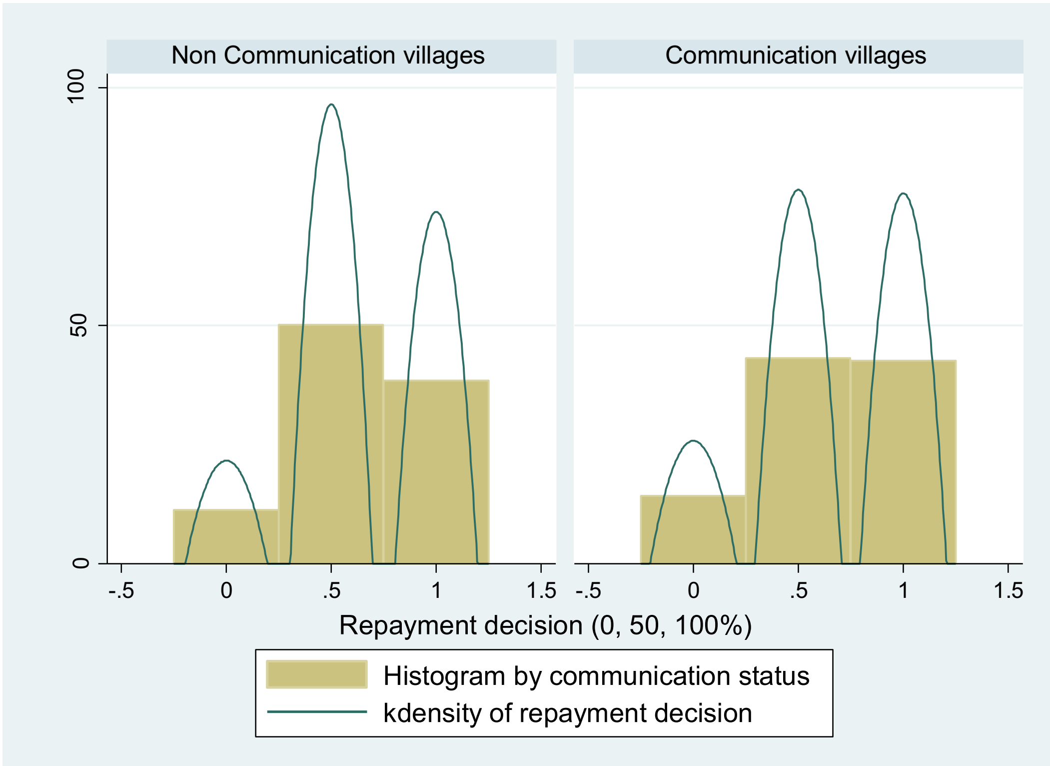
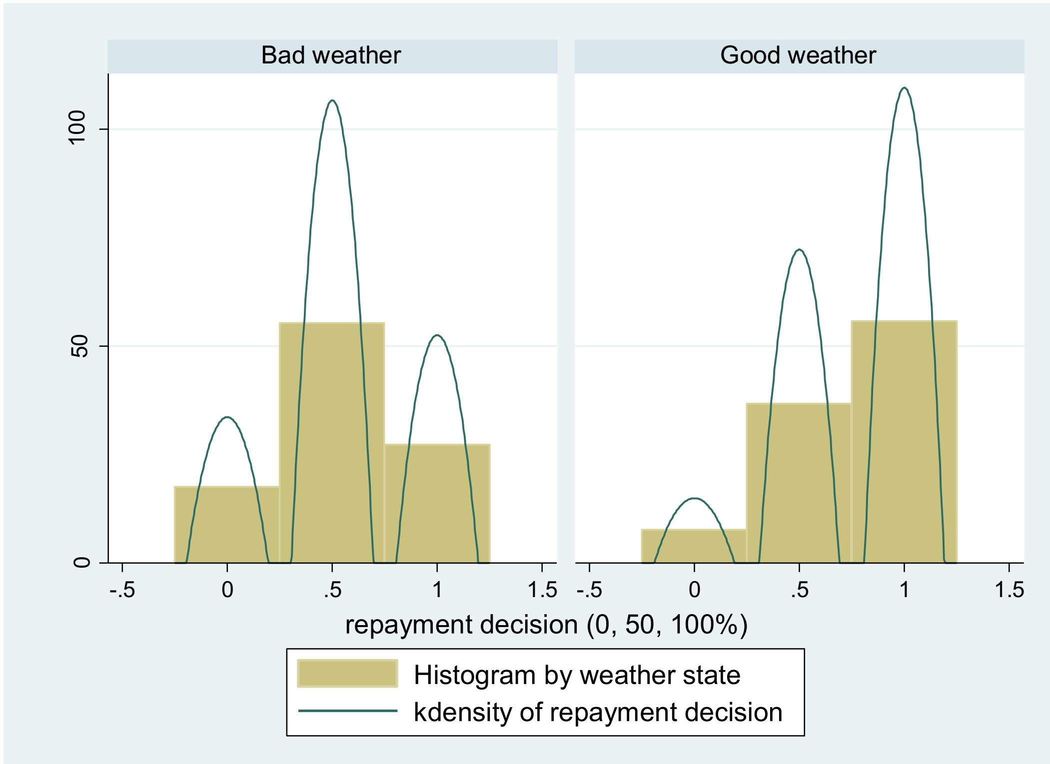


Figure 3 indicates that default rates are higher when the weather state is bad (negative shock). Again, defining 'default' as the proportion of farmers who did not repay fully (i.e. 100% of what was owed), the total default rate during bad weather rounds was 72.77% (55.32% repaid half while 17.45% did not repay at all). For good weather rounds, the default rate was lower at 44.31% (36.73% repaid half while 7.58% did not repay at all). This difference, which is statistically significant (Pr(|T| > |t|) = 0.000), suggests, as hypothesized in the theoretical model above, that profitability shocks play an important role in farmers' decisions to repay input loans. Overall, the descriptive analysis of farmer behavior is consistent with the hypotheses derived from the theoretic model.

Figure 3.3: Histogram of repayment decision by weather state



3.5.2.2. Econometric model

To test the prediction of the dynamic incentives theoretic model, the determinants of a farmer's repayment decision using the communication treatment and weather states as explanatory variables was estimated with the following specification.

$$\mathbf{Y_{it}} = \boldsymbol{\beta_0} + \boldsymbol{\beta_1} * \mathbf{T1_{it}} + \boldsymbol{\beta_2} * \mathbf{T2_{it}} + \boldsymbol{\beta_3} * (\mathbf{T1_{it}} * \mathbf{T2_{it}}) + \sum_2^{11} \boldsymbol{\delta} . \mathbf{Round}_t + \boldsymbol{\varepsilon_{it}} \qquad \textbf{(6)}$$

where:

Y_{it} represents the observed repayment decision made by farmer i in round t

T1 is the binary communication treatment variable that takes value 1 if a farmer resided in a communication village and 0 otherwise. Similarly, T2 is the binary weather state variable that takes value 1 when the weather is good and 0 otherwise. We also include an interaction term between the communication and weather state variables to see if they influence each other's effect on repayment behavior of farmers. Finally round dummies were included to control for rounds effects on farmers' behaviors.

β_1, β_2, and β_3, are the parameters to be estimated, while ε_{it} is the random error term.

We estimated the parameters of equation 7 above both as an Ordered Probit Model and a Probit model. For the Ordered Probit analysis, the dependent variable is the categorical repayment decision variable with values 0 (when no repayment was made at all), 0.5 (when 50% repayment was made), and 1 (when full repayment is made).

For the Probit analysis, the repayment decision variable is binary and takes values 1 when full repayment was made, and 0 otherwise. As such, this specification captures the probability of repaying fully, and is consistent with the definition of default used in the descriptive analysis section above.

Given that both our treatment variables were randomly assigned to farmers per round or village, our key explanatory variables are not correlated with the errors of any past, present, or future round, resulting in unbiased estimates via the strict exogeneity assumption (Wooldridge, 2010). Standard errors are clustered at the farmer level to account for the fact that farmer decisions across rounds are may be correlated.

3.5.2.3. Econometric results

Table 6 presents the results of both the Ordered Probit and Probit regressions. The results are presented overall and separately when the weather state is good and bad. Round effects are hardly significant in the overall regressions results and overall the results are consistent with the theoretic model predictions and the descriptive analysis presented above.[24] First, both the Probit and Ordered Probit regressions indicate a positive and significant effect of weather state on farmers' repayment decisions. The estimated Average Partial Effects (APE) from the Probit model indicate that farmers are about 28 percentage points more likely to repay fully in good weather state than in bad weather state. This is consistent with our research hypothesis and can be attributed to the fact the opportunity cost of repayment is higher in bad weather since yields are low. With regards to the communication treatment, the results of the Probit and Ordered Probit models estimation are also consistent with each other, and show a positive and significant coefficient for the Communication treatment variable. In particular, the APEs reported for the Probit model, though significant only at 10%, indicate that farmers' likelihood of repaying fully is on average 8.6 percentage points higher when input suppliers are able to communicate and exchange information about repayment history. This result is not only consistent with our research hypothesis but also with the findings in Greif (1993), Ghosh and Ray (1999), as well as Luoto et al. (2007), and de Janvry et al. (2010).

The estimation results separately in the good weather and bad weather states are useful for understanding the weak significance of the communication treatment in the overall estimation results. In fact, though it is a positive determinant of repayment behavior in both weather states, the communication treatment is not statistically significant in the good weather state, while it is

[24] The insignificance of the round dummies implies that the communication and weather effects were not driven by farmers behaving in a particular way during specific rounds. In particular, it indicates that the random stopping point method used during the experiment was effective in mitigating farmers' natural incentive to default in the last rounds of the game when they do not expect any future income form the relationship. We also run the regression without including the last round and the conclusion remain the same. It also might indicate that there is no significant learning effects (i.e., the farmers do not appear to be changing their behavior across rounds due to learning how the game works).

strongly significant in the bad weather state. When the weather is bad and farmers are expected to default due to the high opportunity cost of repayment, the communication treatment proves very useful by increasing farmers' likelihood of full repayment by more than 17 percentage points. But in the good weather state when farmers are less likely to default, the communication treatment does not have a strong effect.

Table 3-6: Estimation results for the determinants of farmers' repayment behavior

VARIABLES	APE estimates from Probit model Repayment decision = {0 ,100%}			Coefficients estimates from Ordered Probit model Repayment decision ={0, 50, 100%}		
	Overall	Bad weather	Good weather	Overall	Bad weather	Good weather
Weather state (1=good/0=bad)	0.279***			0.895***		
	(0.000)			(0.000)		
Communication village (0/1)	0.086*	0.172***	0.024	0.308**	0.352**	-0.052
	(0.056)	(0.001)	(0.713)	(0.026)	(0.013)	(0.735)
Interaction weather state # communication village				-0.424**		
				(0.017)		
Round ID = 2	-0.126*	-0.386***	0.130	-0.228	-0.770***	0.268
	(0.070)	(0.000)	(0.163)	(0.139)	(0.001)	(0.188)
Round ID = 3	-0.021	-0.263***	0.256**	-0.219	-0.854***	0.592**
	(0.778)	(0.007)	(0.018)	(0.214)	(0.000)	(0.022)
Round ID = 4	-0.011	-0.228**	0.166*	-0.189	-0.713***	0.232
	(0.871)	(0.026)	(0.098)	(0.255)	(0.005)	(0.333)
Round ID = 5	-0.073	-0.206**	-0.031	-0.277*	-0.547**	-0.303
	(0.316)	(0.022)	(0.829)	(0.096)	(0.010)	(0.397)
Round ID = 6	-0.064	-0.298***	0.122	-0.234	-0.750***	0.158
	(0.377)	(0.003)	(0.196)	(0.175)	(0.005)	(0.457)
Round ID = 7	-0.029	-0.213**	0.129	-0.135	-0.561**	0.216
	(0.695)	(0.024)	(0.211)	(0.443)	(0.015)	(0.347)
Round ID = 8	-0.078	-0.209**	0.034	-0.375*	-0.613**	-0.171
	(0.293)	(0.034)	(0.730)	(0.053)	(0.015)	(0.471)
Round ID = 9	-0.047	-0.172*	0.022	-0.128	-0.400*	0.014
	(0.542)	(0.062)	(0.857)	(0.467)	(0.070)	(0.956)
Round ID = 10	-0.013	-0.070	-0.040	-0.141	-0.216	-0.234
	(0.864)	(0.470)	(0.735)	(0.448)	(0.361)	(0.376)
Round ID = 11	0.069	-0.059	0.176	0.161	-0.133	0.377
	(0.492)	(0.722)	(0.175)	(0.527)	(0.759)	(0.245)

Table 3-6 (cont'd)

Number of observations	892	470	422	892	470	422

Note: pval in parentheses.
Asterisks indicate significance level. *** p<0.01, ** p<0.05, * p<0.1

Indeed, the coefficient on the interaction term between communication and weather state is negative and significant in both models. This implies that even though the communication matters for enforcing repayment of input credit, it seems to matters mostly in presence of negative productivity shocks. This suggests that in the presence of insurance mechanisms that ensure farmers against negative productivity shocks, the information exchange might be less necessary for enforcing repayment of input credits.

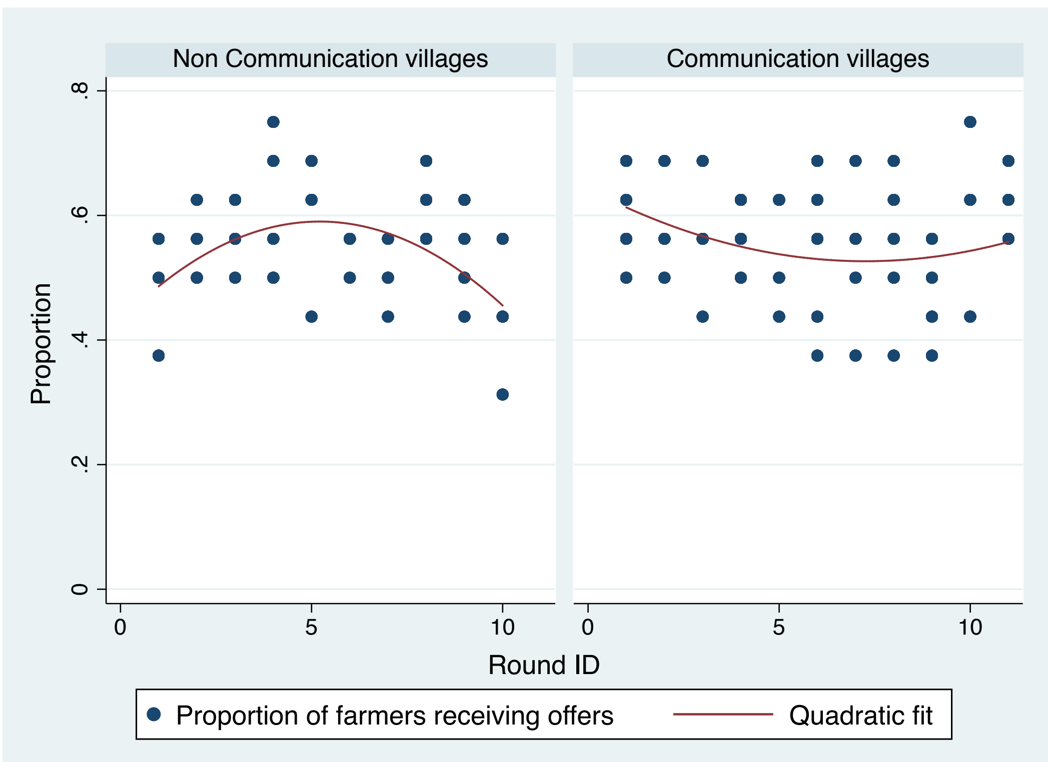
3.5.3. Brokers' behavior

3.5.3.1. Proportion of farmers receiving input on credit throughout the rounds

In this section, we explore the brokers' actions during the game and the rationale behind them. Figure 4 presents how the proportion of farmers receiving offers changes over time in both the communication and non-communication treatments. Specifically, it shows a quadratic fit by treatment group and clearly indicates that in the communication villages, the proportion of people receiving offers decreases in the early rounds of the game, then picks up in the later rounds of the experiment, while the opposite occurs in the non-communication villages. It appears that in the communication treatments, the exchange of information between brokers allows them to effectively implement the multilateral punishment strategy and ostracize defaulting farmers quickly. Once it is clearly established that defaulting is being detected and punished with high probability, the proportion of farmers receiving offers increases again and trade is sustained.

However, in the non-communication treatments, the proportion of farmers receiving offers of input loans increases in the earlier rounds. This is likely because at that early stage, brokers do not have much information about farmer's repayment history and learn about farmer credibility at a slower rate than in the communication villages. Without communication, brokers appear to have kept trying new farmers randomly each round, only avoiding those that had not repaid them in previous rounds. Farmers were then able to take advantage of this delay in information acquisition because they could default 4 times (one for each broker) before potentially being completely ostracized. This likely explains why the proportion of farmers receiving offers increases in the earlier rounds, and then decreases only in the later rounds of the experiment when sufficient information was gathered about all farmers' repayment behavior.[25]

Figure 3.4: Patterns of offers throughout the rounds of the game



[25] In every round, a maximum of 12 farmers, representing 75% of farmers in the game in each village, can receive an offer of input credit. This happens only if each of the 4 brokers make their 3 offers to all different farmers.

3.5.3.2. Brokers' Punishment strategy

As discussed earlier, the main underlying assumption of the dynamic incentives model is that brokers are collectively engaged in a multilateral punishment strategy. To test whether the brokers were actually using this punishment mechanism during the experiment, we estimated a Probit regression to test the effect of farmers' repayment history on their probability of receiving an input loan in a particular round. Specifically, we compute a credit score for each farmer that is updated in each round and takes into account all the history of offers received and repayments made. For each observation (farmer and round), we first create a repayment score for the repayment made ($SCORE_t$). It is zero if the farmer did not get any offer (or got one, but did not accept) in that round. For farmers who took offers, the repayment score takes on a value of 10, -5, or -10 for full, partial, and no repayment respectively. Then for each farmer i in round t, we create a credit score by weighting or discounting the sum of past repayment scores, were the weights are the inverse of how far back repayment was made.

$$\textit{Credit Score}_t = SCORE_{t-1}/1 + SCORE_{t-2}/2 + SCORE_{t-3}/3 + \ldots + SCORE_{t-11}/11 \qquad (7)$$

This method penalizes more recent default behavior and puts less weight on older repayment behavior.

The empirical model was specified as follow:

$$\boldsymbol{Prob(Y_t = 1) = A + B.X_t + e_t} \qquad (8)$$

where Y_t is the binary dependent variable taking values 1 when the farmer received an offer in round t, and 0 otherwise while X_t is the vector of explanatory variables in round t, and includes the farmer's updated credit score at time t, the communication treatment status of the village, the interaction between communication and credit score, and round dummies.

Table 3-7: Determinants of receiving input loan offer as function of past repayment by communication treatment

VARIABLES	Coefficients [P-values] Non Communication Villages	Communication Villages	All
Credit score	0.009	0.026***	0.009
	[0.144]	[0.001]	[0.129]
Communication			0.013
			[0.875]
Credit score * Communication			0.017[+]
			[0.101]
Round dummies			
Round ID = 2	0.165	-0.028	0.069
	[0.373]	[0.890]	[0.617]
Round ID = 3	0.217	-0.066	0.077
	[0.212]	[0.754]	[0.572]
Round ID = 4	0.305*	-0.016	0.145
	[0.097]	[0.937]	[0.291]
Round ID = 5	0.335	-0.063	0.136
	[0.131]	[0.757]	[0.364]
Round ID = 6	0.194	-0.104	0.046
	[0.261]	[0.588]	[0.719]
Round ID = 7	0.091	-0.224	-0.066
	[0.655]	[0.178]	[0.614]
Round ID = 8	0.313	-0.127	0.093
	[0.145]	[0.523]	[0.524]
Round ID = 9	0.113	-0.308	-0.097
	[0.579]	[0.130]	[0.499]
Round ID = 10	-0.095	-0.138	-0.116
	[0.641]	[0.477]	[0.404]
Round ID = 11		-0.028	0.108
		[0.909]	[0.617]
Constant	-0.063	0.221	0.072
	[0.657]	[0.120]	[0.506]
Observations	800	848	1,648

*** p<0.01, ** p<0.05, * p<0.10 [+] p<0.15

The model was estimated for the whole sample and separately for each communication treatment. The results presented in table 7 indicate that in the communication villages, past repayment behavior (captured by credit score) is a significant and positive determinant of the likelihood of getting input on credit in current periods. Farmers who have defaulted in the past are less likely to receive an offer in the current period in the communication villages. But this is not the case in the non-communication villages. This result is consistent with the idea that the punishment mechanism is more effectively implemented when input suppliers are able to communicate and exchange information about farmers. In the communication villages, such information sharing is more easily done, allowing brokers to effectively punish defaulters by not offering them input credit in subsequent periods. Brokers in the non-communication villages do not seem to have been able to implement such punishment mechanism.

Since all brokers collect information on farmer behavior over time, we would expect the extra repayment information received by brokers in the communication villages to be more important in the earlier rounds of the game. In later rounds, brokers in the non-communication villages have also collected information as they experience the behavior of farmers after giving them offers. Consequently, we expect to see a stronger effect of the credit score on the chances of getting an offer in communication villages in round 2 compared to non-communication villages. We test this by running equation model 9 for round 2 only where credit score reflect only the repayment behavior in round 1. The results presented in table 8 reflect the general results that the credit score is a significant and positive determinant of the likelihood of getting an input on credit offer. But in addition, the results from table 2 indicate a stronger and statistically significant interaction effect between communication treatment and the credit score variable.

Table 3-8: Determinants of receiving input loan offer as function of past repayment by communication treatment for round 2 only

VARIABLES	Coefficients [P-values] Non Communication Villages	Communication Villages	All
Credit score	-0.015	0.069***	-0.015
	[0.634]	[0.005]	[0.633]
Communication			0.018
			[0.934]
Interaction Credit score * Communication			0.084**
			[0.033]
Constant	0.081	0.098	0.081
	[0.577]	[0.523]	[0.576]
Observations	80	80	160

*** p<0.01, ** p<0.05, * p<0.1

From a policy point of view, this result speaks to the importance of information sharing mechanisms and institutions, for the effectiveness of collective punishment and dynamic incentive mechanisms. It also echoes the results in de Janvry et al. (2010) who find that information provided by credit bureaus allows branches of a microfinance institution to screen borrowers in and out based on repayment history thereby enforcing repayment. However it is worth noticing that their results focused on the branches of the same lender therefore not necessarily capturing collective punishment strategy amongst potentially competing lenders. Additionally, our analysis is based on the dynamic relationship between borrowers past repayment and lenders decision to consider them for loan, while de Janvry et al. (2010) compared repayment performance of borrowers in periods before and after the credit bureau was available and attributed improvement in borrowers' behavior to the ability to screen them thanks to the information made provided by the credit bureaus.

3.6. Conclusion and Implications

This paper theoretically and empirically examined the importance of communication and information exchange (about repayment history) on the effectiveness of dynamic incentives in input credit arrangements. The theoretic model predictions were tested using experimental data collected from farmers in rural Nigeria. Econometric results using both Probit and Ordered Probit approaches support the model's predictions. We find consistent evidence that information exchange among input suppliers [26] reduces default among farmers in input on credit arrangements. Productivity shocks also affect default rates, though importantly this tends to be less significant when there is information exchange among input suppliers.

The findings of this study are consistent with the literature on microfinance which has established a positive role of dynamic incentives and information sharing for the success of microfinance in situations where scoring mechanisms, collateral requirements, and sound legal systems are non-existent or weak (Tedeschi, 2006, Ghosh and Ray, 1999, McIntosh and Wydick, 2009). This study makes a contribution to this literature by providing additional evidence of the importance of information sharing for the effectiveness of dynamic incentives using experimental methods in the specific context of input credit for famers in a rural developing country setting.

Questions on how such input on credit arrangements can be implemented in practice are legitimate. The costs and other potential issues related to sharing information between input suppliers are also important. If the cost of information exchange is too high, this will increase the cost of the loan to the farmers. Therefore, it might be difficult to sustain this input on credit

[26] In the actual game we played, the participants were brokers acting on behalf of suppliers of input on credit to farmers. Therefore, we generalize the conclusion to input sellers.

arrangement without some external subsidies (from governments or development NGOs), unless the input is so profitable for farmers that they are willing to pay a high enough price for the input loan. According to Morduch (2000) input suppliers providing loans to people in more remote areas may have to make a decision to either curtail outreach to these clients or face the fact that full financial self-sufficiency may not be possible.

However, it may be possible to leverage the microfinance experience. Information sharing is already being incorporated as part of microfinance best practices. The establishment of Credit Bureaus by microfinance institutions in several regions of the globe serves as evidence (Campion and Valenzuela, 2001, de Janvry et al., 2010). Input suppliers themselves might also benefit from such a concept by establishing “input credit bureaus” that collect repayment history information about farmers to whom they provide input loans. Such information can then be shared within the network of input suppliers and play the same role as consumer credit scores in developed countries.

Alternatively, the input suppliers can rely on local village level retailers as brokers to distribute their product to farmers in very remote areas. Given that credit bureaus cannot be established everywhere, village level retailers with necessary social capital might be a potential solution since they have information about the farmers living in their communities. Also, they can more easily exchange information about repayment history with local retailers in neighboring villages to ensure defaulters do not get input loans from nearby village. This is possible because people in very remote rural areas usually know each other – they typically go to the same markets, health care facilities and places of worship. Also, with the promotion of the use of Information and Communication Technology (ICT) in rural areas, this communication and exchange of information between local retailers from different villages could be facilitated to ensure effectiveness of the dynamic incentive and solve strategic default issues.

Finally, our results indicate that it is important to think about ways to combine input credit arrangements with agricultural insurance schemes so that farmers who are unable to repay due to negative economic shocks do not face harsh punishment from input suppliers. Index based insurance schemes targeted at private input suppliers in developing countries are an option to be explored to encourage input suppliers to engage in credit arrangements with smallholders engaged in agricultural activities with non-trivial production risks and uncertainty due to weather.

APPENDICES

Appendix 1: Brokers' decision making sheet

Broker's ID: **Village name:**

Round N*:

Farmers ID	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	Total
Fertilizer Offer made (0kg, 100kg)																	
Offer Accepted/declined (1=yes/0=No)																	
Final sale realized (in Naira)																	
Farmer's Repayment behavior (0, ½, 1)																	
Net payoff to brokers (in Naira)																	

Appendix 2: Farmers' decision making sheet

Farmers' ID: **Village name:**

Round N*:

Brokers ID	1	2	3	4	Total
Offer received					
Accept/Decline (Please circle for YES and cross for NO)					

Amount owed			
Weather state (good/bad)			
Money received after harvest			
Repayment decision (please circle one)	0%	50%	100%
Net payoff to farmer			

Source: Generated by authors

Appendix 3: Implementation and sequence of actions in each round of the game

The experiment was run by a team of 6 enumerators. At the beginning of the experiment, the enumerators first identified the selected participants in the village. The selected participants who were not available were replaced by other people randomly drawn from the list of the village household heads. Then the previous instructions were presented and explained to all participants. Then the participants were separated into farmers and brokers group and received the appropriate sheets (broker sheet and farmer sheet) on which they are supposed to indicate their decisions throughout the game. Then a trial round called round 0 was executed to allow participants to get a better sense of what is going to happen during the experiment. Participants were aware that the round 0 is just a practice and that their answer to that round would not count for the payoff they would receive at the end of the game. After making sure everyone had completely understood the rules of the game, the real experiment starts with round 1 and goes down according to the following steps:

1. Broker makes offer to the farmers
2. Enumerators collect the brokers sheets then transfer offers made onto the farmers' sheets
3. Enumerators give farmers their sheets so they can examine the offers received from each broker, and make their accept/decline decisions
4. Enumerators collect the farmers sheets and transfer accept/decline decisions onto the brokers sheets
5. Enumerators calculate the amount owed by farmers to each brokers and translate onto the farmers sheets.
6. A farmer takes his turn and will flip the coin publicly to determine the weather state. This is also communicated to all players and translated onto the farmers sheets

7. Enumerators give farmers their sheets so they can make repayment decision
8. Enumerators collect the farmers' sheets and transfer repayment decision onto the brokers' sheets
9. Enumerators calculate payoffs for both farmers and brokers, and translate onto their respective sheets.

In the communication treatment villages, farmer's total repayment is reported on the brokers' public board for all the brokers to see before the beginning of the following round when they decide again offer to be made.

REFERENCES

REFERENCES

ARMENDÁRIZ DE AGHION, B. & MORDUCH, J. 2000. Microfinance beyond group lending. *Economics of transition,* 8**,** 401-420.

BESLEY, T. 1995. Savings, credit and insurance. *Handbook of development economics,* 3**,** 2123-2207.

CAMPION, A. & VALENZUELA, L. 2001. Credit bureaus: a necessity for microfinance? Microenterprise Best Practices. *Development Alternative, Inc., Bethesda, Maryland*.

CONNING, J. & UDRY, C. 2007. Rural financial markets in developing countries. *Handbook of agricultural economics,* 3**,** 2857-2908.

DE JANVRY, A., MCINTOSH, C. & SADOULET, E. 2010. The supply- and demand-side impacts of credit market information. *Journal of Development Economics,* 93**,** 173-188.

DJANKOV, S., MCLIESH, C. & SHLEIFER, A. 2007. Private credit in 129 countries. *Journal of financial Economics,* 84**,** 299-329.

DORWARD, A., KYDD, J. & POULTON, C. 1998. *Smallholder cash crop production under market liberalisation: a new institutional economics perspective*, CAB International.

FUDENBERG, D. & TIROLE, J. 1991. Game theory. 1991. *Cambridge, Massachusetts,* 393.

GHOSH, P. & RAY, D. 1999. *Information and enforcement in informal credit markets*, Boston University, Institute for Economic Development.

GREIF, A. 1993. Contract enforceability and economic institutions in early trade: The Maghribi traders' coalition. *The American economic review***,** 525-548.

GREIF, A., MILGROM, P. & WEINGAST, B. R. 1994. Coordination, commitment, and enforcement: The case of the merchant guild. *Journal of political economy***,** 745-776.

HARRISON, G. W. & LIST, J. A. 2004. Field experiments. *Journal of Economic Literature***,** 1009-1055.

HULME, D. & MOSLEY, P. 1996. *Finance against poverty*, Psychology Press.

JAPPELLI, T. & PAGANO, M. 2002. Information sharing, lending and defaults: Cross-country evidence. *Journal of Banking & Finance,* 26**,** 2017-2045.

JOHNSTON, B. F. & MELLOR, J. W. 1961. The role of agriculture in economic development. *The American Economic Review***,** 566-593.

KELLY, V., ADESINA, A. A. & GORDON, A. 2003. Expanding access to agricultural inputs in Africa: a review of recent market development experience. *Food Policy,* 28**,** 379-404.

KELLY, V. A., MORRIS, M., KOPICKI, R. J. & BYERLEE, D. 2007. *Fertilizer use in African agriculture: Lessons learned and good practice guidelines*, Washington, DC: World Bank.

LUOTO, J., MCINTOSH, C. & WYDICK, B. 2007. Credit information systems in less developed countries: A test with microfinance in Guatemala. *Economic Development and Cultural Change,* 55**,** 313-334.

MAS-COLELL, A., WHINSTON, M. D. & GREEN, J. R. 1995. *Microeconomic theory*, Oxford university press New York.

MCINTOSH, C. & WYDICK, B. 2009. What do credit bureaus do? understanding screening, incentive, and credit expansion effects. *Economics.*

MORDUCH, J. 2000. The microfinance schism. *World development,* 28**,** 617-629.

PADILLA, A. J. & PAGANO, M. 1997. Endogenous communication among lenders and entrepreneurial incentives. *Review of Financial Studies,* 10**,** 205-236.

PADILLA, A. J. & PAGANO, M. 2000. Sharing default information as a borrower discipline device. *European Economic Review,* 44**,** 1951-1980.

POULTON, C., DORWARD, A. & KYDD, J. 1998. The revival of smallholder cash crops in Africa: public and private roles in the provision of finance. *Journal of International Development,* 10**,** 85-103.

SADOULET, L. 2005. Learning from Visa®? Incorporating insurance provisions in microfinance contracts. *Insurance against poverty. Oxford University Press, Oxford***,** 387-421.

SHEAHAN, M. & BARRETT, C. B. 2014. Understanding the agricultural input landscape in sub-Saharan Africa: Recent plot, household, and community-level evidence. *World Bank Policy Research Working Paper*.

TEDESCHI, G. A. 2006. Here today, gone tomorrow: Can dynamic incentives make microfinance more flexible? *Journal of development Economics,* 80**,** 84-105.

VERCAMMEN, J. A. 1995. Credit Bureau Policy and Sustainable Reputation Effects in Credit Markets. *Economica,* 62**,** 461-478.

WOOLDRIDGE, J. M. 2010. *Econometric analysis of cross section and panel data*, MIT press.

4. ESSAY 3: THE HETEROGENEOUS WELFARE EFFECTS OF RURAL NON-FARM EMPLOYMENT: RECENT EVIDENCE FROM MALAWI

4.1. Introduction

The contribution of non-farm activities (such as non-farm wage employment and non-farm enterprises) to household income in Sub Saharan Africa (SSA) is substantial and has increased over time (Naude and Ringler, 2014; (Haggblade et al., 2010, Start, 2001, Lanjouw and Lanjouw, 2001, Lanjouw and Shariff, 2004, Reardon et al., 1998). Recent estimates indicate that 44 percent of rural African households (on average) participate in non-farm wage employment or self-employment, and the average income share from non-farm sources is 23%, with an overall positive correlation between diversification and GDP per capita (Davis et al., 2014). Consequently, the RNFE has become an essential part of discussions on poverty reduction in rural Africa; being a potential pathway out of poverty for many.

Despite the extent and growth in importance of the RNFE in SSA, there is limited rigorous and recent empirical analysis on the welfare effects of the subsector and how this varies across different kinds of rural households. Though evidence of positive correlations between RNFE participation and income exists, many studies are dated while an ongoing debate still questions whether RNFE improves welfare or if indeed it is the wealthy who are able to engage in RNFE. Consequently, this article uses a combination of empirical approaches to identify relation between RNFE participation and welfare while attempting to identify some mechanisms through which this link operates and the heterogeneity of such effects across different types of rural households.

The article makes three major contributions to the limited current knowledgebase on RNFE in SSA. First, this is the only paper in SSA (we have found) to address the endogeneity of

RNFE participation in an analysis of its welfare effects using panel data.[27] As pointed out by Barrett et al. (2001b), the usually positive relationship between income and non-farm employment participation should be interpreted with care. These results are limited by biases related to potential reverse causality since the wealthier households, or households more endowed with assets or unobservable abilities that determine wealth, are also more likely to have access to non-farm employment opportunities. Due to data limitation, very few studies in Sub-Saharan Africa have successfully done this, leaving open the important policy question of whether non-farm employment is really a route out of poverty for millions of poor households in rural Africa. Owusu et al. (2011) used a 2007 cross sectional survey of 300 households in northern Ghana to evaluate the impact of non-farm work on household income and food security. Using propensity score matching methods, they found positive effects on both income and food security and attributed poverty alleviation to non-farm activities. However, the relatively small sample used and the cross sectional nature of the data challenge the external and internal validity of these findings. Furthermore, a major critique of the propensity score matching approach is the selection-on-observable assumption and its inability to rule out unobserved factors that might be correlated with participation and welfare. More recently, Ackah (2013) used a larger sample (9,310 households) visited in 2008 in northern Ghana. They also found positive effects of both non-farm activities on household income. However, they also used cross sectional data and their sampling strategy does not allow generalization to the whole country. Much of the remaining literature on the subject is similarly based on cross sectional data and thus is largely limited in its interpretation with few attempting to explore any underlying mechanisms. Using a recently

[27] Kijima et al. (2006) would be an exception as they used panel data collected from 894 rural Ugandan households in 2003 and 2005; but they focused more on off-farm labor supply and its determinants, rather than welfare effects.

collected rich nationally representative panel dataset from Malawi, we are able to use panel estimation techniques to address the endogeneity of RNFE participation due to time invariant unobserved characteristics like ability and make broader generalizations. With rich information on household socioeconomic characteristics and agricultural and non-agricultural practices we are also able to adequately control for numerous potentially confounding factors.

The second contribution this paper makes is its exploration of the heterogeneous welfare effects of RNFE participation. Most previous efforts to explore the impact of participation in non-farm activities have focused on average effects (Ackah, 2013, Owusu et al., 2011, Matsumoto et al., 2006). However, although the average treatment effects are useful measures of the link between participation in non-farm employment and welfare outcomes, they provide an incomplete view of the relationship. While a traditional growth model might indicate larger effects at lower ends of the distribution due to diminishing marginal returns, it is also possible that RNFE participation requires certain investments or expenses less available to the poor. Furthermore, numerous studies on the determinants of household participation in non-farm activities (over the past two decades) have found evidence of entry barriers for vulnerable socio economic groups such as women or the poor (Abdulai and CroleRees, 2001, Woldenhanna and Oskam, 2001, Barrett et al., 2001a, Smith et al., 2001, Lanjouw et al., 2001, Reardon et al., 2000). The marginalized groups are either totally excluded or restricted to the least lucrative non-farm activities, despite their relatively greater need for diversification. These barriers that restrict marginalized groups participation in RNFE implies that the size of the welfare effects of participation in non-farm activities might actually be lower for these groups compared to the more privileged groups with access to high return opportunities. In either of these cases, the average treatment effects measures may hide such heterogeneity. Despite the widely recognized existence of entry barriers there are still no studies (the authors are aware of) that have investigated the distributional effects of RNFE participation at different points of the wealth distribution, which has very important implications in terms of informing the likely policies (and their targeting) that would be effective in maximizing any benefits of RNFE for rural households. This article uses a quantile regression (QR) approach to test for heterogeneous effects of RNFE participation on household consumption expenditure. The QR approach explores the relationship between non-farm employment participation and welfare at different points of the conditional distribution of the outcome (Cameron and Trivedi, 2010). Applying the QR approach within a panel framework we are able to still control for time invariant unobservable factors correlated with welfare and participation in RNFE and this is also the first paper to do this within the context of RNFE in a developing country.

The third contribution this paper makes is exploring a mechanism through which welfare effects of RNFE might occur for rural households. Since most rural households are at least partially involved in agriculture, in addition to the direct effects of income for consumption that RNFE provides for households, RNFE can also serve as a source of cash for investment in agriculture (Oseni and Winters, 2009), or rather take resources away from agricultural activities (Smale et al., 2016). In rural Malawi more than 90 percent of households are involved in agricultural crop production activities, which provides about 55 percent of total household income on average. Thus, we go a step further to explore if participation in RNFE has positive (or negative) effects on investments in agriculture. Previous studies have focused either solely on the effect of RNFE on agricultural investments (Oseni and Winters, 2009, Smale et al., 2016) or welfare directly (Ackah, 2013, Owusu et al., 2011, Matsumoto et al., 2006). We explore both in the same context. We extend the work of Oseni and Winters (2009) to another context using panel data methods rather than cross sectional data and also explore the direct effect of RNFE on welfare.

Consequently, this article significantly enriches the discussion on RNFE and rural development in SSA using recent nationally representative panel data from Malawi[28]. Keeping a clear distinction between non-farm wage employment and non-farm self-employment, we start with a descriptive analysis of patterns of participation and returns from different categories of non-farm employment, and explore evidence of persisting dualism along poverty lines. Then we analyze the determinants of households' participation in both non-farm activities using panel data econometric methods. This provides evidence of whether *push* or *pull* factors prevail in RNFE participation in rural Malawi. Finally, we investigate the impact of RNFE participation on various objective and subjective measures of welfare and poverty, using a variety of panel data econometric methods.

The rest of the paper is organized as follows. Section 2 presents the conceptual framework underlying our empirical analysis while section 3 describes the data used. This includes a general description of the patterns of and returns to participation in non-farm activities, as well as the types of non-farm activities, for the poor versus non-poor, in rural Malawi. Section 4 discusses the econometric framework used for our analyses while section 5 presents and discusses the study results. Section 6 concludes with suggestions for policy consideration and future research.

[28] https://www.malawi.gov.mw

4.2. Conceptual framework for participation in non-farm activities

As suggested by Barrett et al. (2001b), it is important to distinguish between various terminologies such as "off farm", "non agricultural" , "non-farm" employment, etc. often used synonymously in the literature to describe the RNFE. Using an adaptation of the sectorial classification in Barrett et al. (2011)[29], we define non-farm employment in this article to mean, all activities outside of crop and livestock production. Agricultural wage employment such as ganyu wage[30] is thus excluded from our non-farm employment definition. However we include other activities of the primary sector such as forestry, hunting, fishing, mining and quarrying, etc. The main reason for excluding agricultural wage employment from our analysis is because it is a special category of employment that attracts generally the poorest, and has been shown to have no significant role in household income (eg. Matsumoto et al. (2006) in Eastern Africa). Following the functional classification, we distinguish between non-farm wage (involving a wage or salary contract) and non-farm self-employment (entrepreneurial activity).

Observed patterns of participation in non-farm activities, such as wage employment and self-employment result from the combination of "*pull*" and "*push*" factors. *Push* factors relate to the need for ex-ante income smoothing strategies in the presence of binding financial constraints and limited risk mitigating solutions. Households may diversify to satisfy the need for cash to finance agricultural activities in the absence of rural financial services, or the need to feed a large household on a limited amount of land in case of crop failure. On the other hand, *pull* factors relate the desire by economically rational households to take advantage of opportunities generated by the transformation of agricultural and the rural economy as a whole.

---

[29] Barett et al. (2011) exclude completely the primary sector from non-farm employments.
[30] Ganyu labour is short-term labour hired on a daily or other short-term basis. Most commonly, piecework weeding or ridging on the fields of other smallholders or on agricultural estates.

Increased agricultural productivity from the use of modern production techniques, coupled with diminishing marginal return of labor in agricultural use, free labor for use in more productive non-farm alternatives. Also, increased urbanization, and income rise as part of the structural transformation underway in most developing countries, generates demand for non agricultural goods and services, thereby offering remunerative opportunities in the non agricultural sector for the surplus labor squeezed out from the agricultural sector (Barrett et al., 2001b, Haggblade et al., 2010). While in poor agrarian economies and for poor households, push factors might be expected to play a major role in triggering the need for diversification (Bardhan and Udry, 1999), resource constraints and entry barriers faced by the poor may also restrict their actual participation in non-farm activities. This explains why, contrary to the expectation that poor households should diversify more, we might observe more diversification by the wealthier households. When push factors do not prove significant determinants of livelihood diversification in a poor agrarian economy, it is a sign of entry barriers limiting the poor's access to diversification opportunities in spite of their needs. In what follows, we describe a simple model of income diversification under risk and resource constraints, and propose an econometric framework to analyze determinants of non-farm employment engagement.

Following Bardhan and Udry (1999), adapted by Abdulai and CroleRees (2001) we formalize households' participation in non-farm self-employment and wage employment by assuming that households allocate resources such as time and land across various activities including farm and non-farm activities. Assuming a static model[31], households choose

[31] While we recognize that a dynamic framework might add some richness to the model, our data set contains only two year periods which are nonconsecutive. Therefore, it does not allow us to test empirically for dynamic mechanisms. Hence, both our conceptual and empirical discussion, are based on a static model for the sake of consistency.

consumption C_t to maximize the following expected utility function Max $U_t = u(C_t)$ subject to the following constraints:

budget constraint: $C_t = \sum_{k=1}^{K} g_k(l_{kt}, \varepsilon_{kt}; \boldsymbol{X})$ (1)

Time endowment constraint: $\sum_{k=1}^{K} l_{kt} \leq L$ (2)

Non-negativity constraint: $l_{kt} \geq 0$, k=1…K (3)

Where l_{kt} is the amount of labor allocated to activity k at time t. $g_k(l_{kt}, \varepsilon_{kt}; \boldsymbol{X})$ is the technology constraint that characterizes the returns from investing l_{kt} units of labor in alternative k. X captures household's individual and location characteristics that influence the returns to labor use in each of the K options.

The first order conditions for the above maximization problem imply that households' allocate labor between K activities in order to equate marginal utility of allocating one unit of labor to each of them. Mathematically this implies:

$$E_t[U'(C_t).\, g'_k(l_{kt}, \varepsilon_{kt}; \boldsymbol{X})] = E_t[U'(C_t).\, g'_{-k}(l_{-kt}, \varepsilon_{-kt}; \boldsymbol{X})] \quad (4)$$

where –k refers to activities other than k. The household labor allocation decision to activity k takes into account the expected return from that activity, and the maximum expected returns from all the other possible activities. If for any activity, the household's endowments in human, financial, and physical capitals implies a low expected return from that activity compared to the others, then no labor would be allocated to that activity. This implies that dualism in the types of non-farm employment and returns along poverty lines, might justify why the poor participate less in non-farm activities than the non-poor. For the poorest, though expected marginal utility from investing labor in agriculture is low given their land constraint and limited access to other productive agricultural assets, it does not immediately translate into non-farm employment because the expected returns to the best non-farm employment they have access to, given their resources, may still be very low. The same type of argument applies when households are allocating labor between non-farm wage employment and non-farm self-employment. Different types of households' resources might matter for the non-farm self-employment compared to wage employment and therefore will affect households revealed preference for each of them. We investigate empirically, the main factors that determine households' participation in non-farm wage employment and non-farm self-employment in rural Malawi.

4.3. Data

4.3.1. Data source

This study uses data from the Malawi Integrated Households Panel Surveys (IHS3) implemented with a joint effort from the Government of Malawi through the National Statistical Office (NSO; www.nso.malawi.net), and the World Bank Living Standards Measurement Study – Integrated Surveys on Agriculture (LSMS-ISA) initiative. This is a multi-dimensional survey with detailed information about households' characteristics, activities and livelihood, agricultural practices and community level information. The first round of data includes 3,246 households, selected by means of a stratified sampling from 204 enumeration areas (EAs), and interviewed from March 2010 to November 2010 as part of the larger third Integrated Household Survey (IHS3). The same households were revisited between April and December 2013 for the second round of the panel survey. Due to attrition and split off of some households between the two survey periods, the second round IHPS dataset contains 4000 households that can be traced back to 3,104 baseline households[32].

The sample frame includes all three geopolitical regions of Malawi: North, Centre and South. The survey stratified the country into rural and urban strata. The urban stratum includes the four major urban areas: Lilongwe, Blantyre, Mzuzu, and the Municipality of Zomba. All other areas including Bomas are considered as rural areas. In this study we are interested in livelihood diversification strategies amongst rural population, which represents about 85 percent of the sample (2,633 households). The IHPS data are representative at the national, urban/rural and regional levels.

We also use auxiliary data such as the consumption aggregates developed by the World Bank LSMS team for poverty analysis in Malawi[33]. The use of these pre-generated variables allows us to compare and link our analyses to other reports and articles developed from the publicly available LSMS data, thereby contributing coherently to the large discussion about livelihoods in rural Malawi, and Africa more generally.

In line with our definition and classification of RNFE above, we consider that a household is participating in non-farm wage employment if at least one member of the household holds a non-agricultural job involving a wage or salary contract. As for non-farm self-employment it relates to whether at least one member of the household owns a business or works on their own-account.

[32] For more information on the Malawi LSMS-ISA initiative, please visit www.worldbank.org/lsms-isa

[33] The consumption aggregates data as well as a detailed discussion on how each component is calculated is in the dissemination documentation available for download along with the consumption aggregate data from the LSMS website

4.3.2. Patterns of non-farm activities in rural Malawi

Like many other parts of the developing world, the rural economy in Malawi is transforming as its economy grows. A sign of such transformation is the increasingly important role of non-farm activities in rural Malawi. Though, agriculture remains the main source of income, participation rates in non-farm wage and self-employment and income shares from those activities, have reached non-negligible levels. The proportion of households earning income from non-farm self-employment was about 18 percent in 2010, and 27 percent in 2013 (Table 1). Meanwhile, average households' income share coming from non-farm self-employment was estimated at about 7 percent and 11 percent in 2010 and 2013 respectively. As for non-farm wage employment, participation rates sat around 17 percent and 15 percent, in 2010 and 2013 respectively, with associated income shares around 8 and 6 percent on average. Compared to other Sub Saharan African (SSA) countries, rural Malawi's rates of participation in non-farm wage employment and self-employment are about average. The estimation by Davis et al. (2014) for a group of SSA countries, indicates on average 15 percent participation rate for non-farm wage employment and 34 percent for non-farm self-employment, with associated income shares of 8 percent and 15 percent.

Though participation rates are increasing for both poor[34] and non-poor households, the rates for the non-poor are much higher in both survey years and this difference is statistically significant (table 1). The differences between poor and non-poor are much starker in wage employment where participation rates for the non-poor are more than double those of the poor. The lower participation gap observed for non-farm self-employment likely reflects lower entry

[34] A household is poor if consumption expenditure per capita falls below the poverty line. See more details in section 4.

barriers to non-farm self-employment compared to non-farm wage employment. This is not surprising if such jobs require more education and better social networks. While the lower participation rates among the poor call for further attention to participation barriers, the fact that participation rates among the poor are still almost 20% is indicative of its potential importance and the need to understand if and how such households benefit from these non-farm activities.

Table 4-1: Household participation in non-farm employment in rural Malawi, 2010/2013

	Malawi				Non Poor				Poor				P-value t-test poor vs. non poor			
	Participation rates (%)		Share of Households Income from Source (%)		Participation rates (%)		Share of Households Income from Source (%)		Participation rates (%)		Share of Households Income from Source (%)		Participation rates (%)		Share of Households Income from Source (%)	
	2010	2013	2010	2013	2010	2013	2010	2013	2010	2013	2010	2013	2010	2013	2010	2013
Non-Agricultural wage	17.1	15.1	8	6.5	21.6	19	10.6	8.5	9.7	7.4	3.9	2.3	0.000	0.000	0.000	0.000
Self Employment	18.2	26.8	7.2	11.1	20.3	30.5	9	12.9	14.9	19.4	4.4	7.4	0.002	0.000	0.076	0.000

Source: Malawi IHPS (2010/11 and 2013)

4.3.2.1. Non-farm self-employment in rural Malawi

The polarized nature of income diversification patterns in non-farm activities has been demonstrated in various contexts. Reardon et al. (2000), and Davis et al. (2014), amongst others, talk about the existence of high and low return non-farm activities with the high return activities almost exclusively available to a handful of privileged. Entry barriers prevent the more marginalized groups (the poor and often women) from taking advantage of those opportunities. The next few sections present a sectorial classification of non-farm self-employment and non-farm wage employment in rural Malawi to confirm if these patterns persist.

According to Nagler and Naude (2014), rural non-farm enterprises in SSA have little potential for job creation as they are mostly informal, have low productivity, and short life spans. Table 2 tends to support this description for rural Malawi. The lifespan of the average non-farm enterprise in rural Malawi is about 10 years. Almost half of all non-farm businesses are reportedly operated from home, with only 13 percent having access to electricity, and 6 to 8 percent being formally registered. Business owners are relatively young (38 to 39 years old) and mostly uneducated (75% of business owners have no formal education). Higher level of education of the owner and formal registration of the business makes the most significant difference between enterprises owned by poor versus non-poor households. Similarly education of the business owner and age of enterprises tend to be significantly different between male and females. Female managed enterprises are on average younger than male managed and women managers are significantly less educated. These differences are important as they indicate that not only might limited assets like education affect participation but they might also affect productivity.

Table 4-2: Selected characteristics of household enterprises in rural Malawi

	Rural Malawi		Non Poor		Poor		p-value difference		Female		Male		p-value difference	
Characteristics of households' enterprises	2010	2013	2010	2013	2010	2013	2010	2013	2010	2013	2010	2013	2010	2013
Age of enterprises (years)	11.1	9.0	11.7	9.0	12.9	9.1	0.250	0.821	11.9	7.2	12.2	10.4	0.721	0.000
Outside partner (%)	3.4	3.2	4.5	3.0	0.7	3.7	0.008	0.727	1.1	1.2	4.8	4.8	0.025	0.009
Business operating premises (%)														
Home	47.0	43.3	48.1	44.4	44.6	39.8	0.592	0.356	60.0	47.0	38.6	40.5	0.000	0.162
Market place and commercial area shop	31.5	38.0	29.8	36.6	35.5	42.8	0.285	0.198	26.8	40.4	34.6	36.1	0.117	0.266
Roadside and other areas	21.4	18.6	22.1	19.0	19.9	17.4	0.671	0.712	13.2	12.6	26.8	23.4	0.006	0.002
Formal registration (%)	7.1	7.9	9.1	9.3	2.4	3.3	0.003	0.008	2.3	3.0	10.1	11.9	0.000	0.000
FBPE[a] (%)	14.3	14.3	10.1	13.4	23.7	17.5	0.001	0.282	7.4	5.7	18.7	21.2	0.007	0.000
Access to electricity (%)	8.0	13.6	13.8	15.8	0.0	0.0	0.141	0.069	0.0	27.0	14.4	8.7	0.166	0.324
Number of enterprises per household (1 to 4)	1.1	1.1	1.0	1.1	1.1	1.0	0.907	0.027	1.1	1.1	1.1	1.1	0.941	0.915
Owner if household head (%)	73.6	69.5	75.6	72.3	69.1	60.4	0.233	0.008	37.5	39.4	97.0	93.6	0.000	0.000
Age of owner (%)	38.3	38.5	37.4	38.4	40.4	39.0	0.056	0.651	39.3	38.0	37.6	39.0	0.199	0.313
Education of enterprise owner (%)														
None	77.2	74.1	74.7	70.3	82.8	86.6	0.116	0.000	81.6	80.5	74.3	68.9	0.174	0.005
PSLC[b]	10.2	12.7	9.6	14.2	11.5	8.0	0.673	0.009	11.9	11.9	9.1	13.4	0.453	0.644
JCE[c]	8.1	7.9	9.6	8.7	4.7	5.1	0.078	0.091	3.9	5.2	10.8	10.1	0.010	0.041
MSCE[d]	3.7	4.5	4.9	5.7	1.0	0.4	0.010	0.000	2.6	2.0	4.4	6.5	0.234	0.023
Non-university diploma	0.5	0.5	0.7	0.6	0.0	0.0	0.117	0.044	0.0	0.0	0.8	0.8	0.118	0.058
University diploma	0.4	0.4	0.5	0.5	0.0	0.0	0.311	0.135	0.0	0.4	0.6	0.4	0.314	0.989
Post-graduate degree	0.0	0.0	0.0	0.0	0.0	0.0		0.322	0.0	0.0	0.0	0.0		0.323

Source: Malawi IHPS (2010–11 and 2013), panel samples.
Note: a Forest Based Products Enterprise. b Primary School Leaving Certificate. c Junior Certificate Examination. d Malawi School Certificate of Education Examination

Participants in off-farm self-employment are involved in a variety of activities across several sectors or industries[35]. Table 3 indicates that about half of household enterprises in rural Malawi are in the Commerce and Tourism sector (wholesale, retail trade, restaurants, and hotels). These are typically small businesses involving selling or reselling a wide variety of products from groceries and food products to items such as clothes, shoes, etc., generating, on average, a profit of 6,000 to 7,000 MKW[36] monthly. The second most prominent sector is the Manufacturing sector, (food and non food combined) which accounts for approximately 40 percent of all household enterprises. The manufacturing sector is dominated by the Food, Beverage, and Tobacco Manufacturing, which accounts for more than half of the manufacturing sector and represents more than 25 percent of all household enterprises. It includes primarily street vendors of various food and drinks and making a profit in the range of 4,000 to 6,000MKW monthly.

The profits from non-farm self-employment are pretty low (around 6,000MKW) in rural Malawi and they tend to be lowest in the most popular sectors (Table 3). The commerce and tourism sector, which contains about half of the non-farm businesses, generates on average 55 to 65 percent of the profit reported in the construction sector. The manufacturing sector, second most important sector generates even lower revenues. Returns from the Food, beverage, and tobacco manufacturing sector are about 3000MKW monthly, and the non-food manufacturing

---

[35]The classification of non-farm enterprise activities into industry categories used here closely follows the 1992 United Nations International Standard Industrial Classification (ISIC) standards into 5 main groups. The groups include: (a) Primary sector, which comprises agriculture, livestock, hunting, fishing, and mining; (b) Food, Beverage, and Tobacco Manufacturing; (c) Non-food Manufacturing, (d) Commerce and Tourism (wholesale and retail, and restaurants and hotel businesses); and (e) Other sectors, which include construction, electricity and utilities, transportation, and other services.
[36] Kwachas (MWK), local currency used in Malawi. 1 MKW= 0.007 USD in 2010 according to Oanda.com.

sector generates about 5000 to 6000 KW monthly. These low revenues generated by non-farm activities match the description of Nagler and Naude (2014) who deemed them low productivity enterprises with little potential for job creation.

While evidence of dualism is not obvious from the observed distribution of engagement in non-farm enterprises in sectors by the poor and non-poor, a Kolmogorov-Smirnov test of equality of distribution does indicate a significant difference (at 1%) between poor and non poor in terms of distribution over the sectors of non-farm enterprises. There is also a consistently significant superiority of the profits earned by the non-poor compared to the poor across all categories of self-employment. The t-test of mean difference in table 3 show p-values less than 1% in most categories of self-employment.

Table 4-3: Distribution of non-farm household enterprises and returns by Sector, by poverty status, in Rural Malawi, 2010/2013

Sector	Rural Malawi		Non Poor		Poor		p-value difference	
	2010	2013	2010	2013	2010	2013	2010	2013
Participation	Proportion of enterprises by sector of industry (%)							
Primary sector	1.6	2.7	0.8	1.9	3.4	5.4	0.161	0.183
Food, beverage, tobacco manufacturing	26.1	25.6	22.3	25.3	34.8	26.6	0.036	0.810
Non food manufacturing	16.2	13.6	15.9	12.7	16.9	16.7	0.786	0.248
Wholesale and retail trade and restaurant & hotels	46.4	49.2	50.4	50.9	37.3	43.4	0.016	0.121
Construction, services, and other sectors	9.7	8.9	10.7	9.1	7.6	8.0	0.329	0.674
Returns	Profit of enterprises by industry (1000 KW)							
Primary sector	4.6	3.2	6.1	3.4	3.8	3.1	-	0.651
Food, beverage, tobacco manufacturing	3.8	3.7	5.1	4.3	1.7	2.1	0.038	0.006
Non food manufacturing	4.8	5.8	5.9	6.7	2.4	3.1	0.052	0.086
Wholesale and retail trade and restaurant & hotels	6.6	7	7.6	8	3.6	3.3	0.000	0.000
Construction, services, and other sectors	8.6	12.2	10	13.8	2.8	6.5	0.017	0.109
Overall	5.7	6.4	7	7.3	2.6	3.2	0.000	0.000

Source: Malawi IHPS (2010/11 and 2013), panel samples.
Note: The information in this table is at the enterprise level. Profit is the net profit generated by the enterprise over the month of operation prior to the interview, as reported by the enterprise owner. Kwachas (MWK), local currency used in Malawi. 1 MKW= 0.007 USD in 2010 according to Oanda.com.

4.3.2.2. Non-farm wage employment in rural Malawi

Non-farm wage employment is also distributed across several sectors[37] presented in table 4. The *service sector* appears to be the largest with about a third of all non-farm wage employment. It is closely followed by *transport equipment operators' sector, and laborers not elsewhere classified* and the *Professional, technical, & related worker groups* which both account for about a quarter of non-farm wage employment. Together, the remaining sectors each account for less than 7% of non-farm jobs in both 2010 and 2013. Interestingly, approximately 75% of all rural non-farm jobs are in sectors that do not require highly skilled labor and wages in those sectors are significantly lower than wages in sectors that require more skilled labor. As a result, though the average monthly wage across all non-farm jobs in rural Malawi is pretty low, around 8000 Kwachas, it masks the heterogeneity in wage employment.

Dualism in the categories of non-farm wage employment is more pronounced than for non-farm self-employment. The skilled labor jobs, which generate the highest returns, are almost exclusively accessible to the non-poor while the non-skilled labor jobs are available for all. Table 4 indicates that about 80 percent of the jobs taken by the poor are in the non-skilled labor sectors, compared to about 60% for the non-poor. In addition, and non-surprisingly, the poor earn significantly less than the non-poor in most sectors of non-farm wage employment.

---

[37] We follow the occupation codes used in the Malawi LSMS survey instruments, which includes: Relatively ***skilled*** labour jobs such as: (1) Professional, technical, & related workers; (2) Administration and managerial workers; (3) Clerical and related worker; (4) Sales workers. And relatively ***unskilled*** labour jobs such as: (1) Service workers; (2) Agricultural, animal husbandry and forestry workers, fishermen and hunters; (3) Other sectors including production and related workers, transport equipment operators and labourers not elsewhere classified.

Table 4-4: Distribution of non-farm wage employment and returns by Sector, by poverty status, in Rural Malawi, 2010/2013

Sector	Rural Malawi		Non Poor		Poor		p-value difference	
	2010	2013	2010	2013	2010	2013	2010	2013
Participation	Proportion of participants by non-farm wage employment sector (%)							
Professional, technical, & related workers	24.3	24.8	26.7	28	15.3	7.1	0.037	0.000
Administration and managerial workers	1.3	1.7	1.5	2	0.8	0	0.441	0.022
Clerical and related worker	5.7	4.6	7.2	5.4	0	0	0.000	0.000
Sales workers	4.8	5	4.9	4.6	4.3	7.4	0.848	0.496
Service workers	34.4	30.2	30.3	27.6	50.1	44.4	0.002	0.046
Forestry workers, fishermen and hunters	1.8	3.8	2.2	2.7	0.2	9.6	0.074	0.165
Transport equipment operators and laborers not elsewhere classified	27.6	30	27.2	29.6	29.3	31.5	0.745	0.785
Return	Average monthly wages of participants by Sector (1000 KW)							
Professional, technical, & related workers	16.3	16.8	16.8	17.3	12.8	6.2	0.281	0.000
Administration and managerial workers	10.6	2.9	11.5	2.9	4	-	0.012	-
Clerical and related worker	10.4	10.9	10.4	10.9	-	-	-	-
Sales workers	5.6	6.8	6.5	7.4	2	4.8	0.000	0.065
Service workers	4.7	4.9	4.6	5.2	4.9	4.1	0.659	0.131
Forestry workers, fishermen and hunters	14.6	3.3	14.8	5	4.4	0.7	0.072	-
Transport equipment operators and laborers not elsewhere classified	7.3	5.7	7.6	6.2	5.9	3	0.239	0.001
Overall	8.1	7.6	8.8	8.4	5.5	3.7	0.000	0.000

Source: Malawi IHPS (2010/11 and 2013), panel samples.
Note: Wage is the estimated monthly salary using the last payment reported by the participant member and the period of time covered by that payment. Kwachas (MWK), local currency used in Malawi. 1 MKW= 0.007 USD in 2010 according to Oanda.com.

Monthly wage earned by participants in non-farm wage employment, and t-test of difference in wages between the poor and non-poor by category of employment, reveal significant differences across poverty groups. The only sector in which we fail to reject the absence of a significant difference between wages earned by poor and non-poor is the service sector, which is the most popular sector.

The descriptive statistics on RNFE in Malawi provides some preliminary evidence of unequal access to high return non-farm opportunities as rural economies in Africa are transforming. The poor have limited access to non-farm wage employment and self-employment; and when they do participate, they tend to earn significantly less than the non-poor. The lower participation rates of the poor in non-farm activities could also be an illustration of the low returns they get from participating. In our econometric section later, we investigate the extent to which this dualism translates into differences in effects of participation in non-farm activities for the poor and the non-poor. While non-farm employment participation is believed to have a positive effect on the welfare of participants, it would come as no surprise that such effects would be heterogeneous in the presence of such dualism.

4.4. Empirical strategy

4.4.1. Determinants of participation in non-farm employment

The empirical analysis follows directly from the conceptual framework presented earlier. Since we cannot fully observe the marginal returns to labor in each activity, we postulate the following latent variables W_{itk}^{*} which characterizes the differential benefit from participation in an activity k. We drop the k subscripts for succinctness. As mentioned above, households' individual and location characteristics play an important role in determining this differential benefit. We allow for both observable and unobservable household characteristics to affect this decision. Therefore, the latent variable model is specified as follows (Green William, 2000, Wooldridge, 2010, Owusu et al., 2011):

$$\boldsymbol{W_{it}^{*} = 1[X_{it}'\beta + c_i + u_{it} > 0]} \qquad \textbf{(5)}$$

Where W_{it} is the observed participation decision of a household at time t. X_{it} is the vector of explanatory variables included in the model. Assuming a standard normal distribution for the error term u_{it} we get a Probit model for the participation decision. c_i captures unobservable characteristics of the households such as ability, networks and preferences, that can affect their employment choice and may also be correlated with some explanatory variables such as education. β is the vector of parameters of interest. The estimation of this model with pooled Ordinary Least Square (OLS) or Probit or Logit approaches without taking into account the unobserved parameter c_i will likely lead to inconsistent estimates due to omitted variables. If this was a linear model, we could use a fixed effects (FE) approach to attenuate potential biases by using variation in RNFE participation in a household over time to identify the effects of such participation on household welfare. In our case where we have some non-linear models, the use of the fixed effects method is problematic because of the incidental parameters problem (Wooldridge, 2010). Instead we use the Mundlak (1978) special case of (Chamberlain, 1982) correlated random effect (CRE) to model the relation between the unobserved time invariant heterogeneity parameter c_i and the explanatory variables X_{it} , and we get the following unobserved effects Probit response function (Green William, 2000, Wooldridge, 2010):

$$\boldsymbol{Prob(W_{it} = 1|X_{it}) = \Phi(X_{it}'\beta + c_i)} \qquad \textbf{(6)}$$

$$\boldsymbol{c_i = \psi + \overline{X_i}\xi + a_i, \quad a_i|X_i \sim Normal(0, \sigma_a^2)} \qquad \textbf{(7)}$$

The full model becomes:

$$\boldsymbol{Prob(Y_{it} = 1|X_{it}, \overline{X_i}) = \Phi(X_{it}'\beta_a + \psi_a + \overline{X_i}\xi_a),} \qquad \textbf{(8)}$$

with $\theta_a = \theta/(1+\sigma_a^2)^{1/2}$, $\theta = (\beta, \psi, \xi\ )$

Partial effects of variable x_j, from the model above, are defined as:

$$\boldsymbol{PE_{x_j}} = \frac{\partial \boldsymbol{Prob}(\boldsymbol{W_{it}=1|X_{it},\overline{X_\iota}})}{\partial \boldsymbol{x_{jt}}} = \boldsymbol{\beta_{aj}}.\,\boldsymbol{\phi}(\boldsymbol{X'_{it}\beta_a + \psi_a + \overline{X_\iota}\xi_a}) \quad \textbf{(9)}$$

The parameters of equation (8) and (9) can be consistently and efficiently estimated using random effects conditional maximum likelihood (or full MLE). However, consistency of the full MLE relies on the conditional independence assumption, thereby ruling out serial correlation in the error terms. In our case, there are reasons to believe that participation in non-farm activities in a particular year is not totally independent from participation in previous year even after controlling for observable characteristics, creating a violation of the conditional independence assumption. An alternative is to use pooled binary estimation method or partial (pooled) maximum Likelihood estimation. These estimates are consistent even under violation of the conditional independence assumption, though they are usually inefficient. More efficient estimation is possible using the Liang and Zeger (1986) Generalized Estimating Equation technique (GEE). These estimators are also robust to violation of the conditional independence assumption, though panel-robust standard errors should used for inference. Given its advantage over full MLE and pooled probit, we use mainly the GEE approach to estimate average partial effects (APEs). The GEE estimator is a more efficient version of the moment based GMM estimator, and is asymptotically equivalent to the Weighted Non Linear Square (WNLS) estimation approach (See Cameron and Trivedi (2005) P790 for a brief discussion).

An alternative way of dealing with the unobserved parameter in the binary choice model is the conditional fixed effect Logit as used by Abdulai and CroleRees (2001). Though this method effectively addresses the unobserved heterogeneity issue, partial effects are not identified and it is not robust to serial correlation in the error terms (Wooldridge, 2010), hence our preference for the CRE Probit. To ensure that the Mundlak-Chamberlain device used is not too restrictive, we also use the cluster-robust linear fixed effects model, estimated using Ordinary Least Square (OLS) on the demeaned equation, as benchmark for comparison (Wooldridge, 2010).

Notice that our main approach described above uses single equations and models separately the determinants of participation in non-farm self-employment and non-farm wage employment. This assumes that both decisions are made in complete independence from each other, implying that the residuals terms in both equations have zero correlation:

$$\boldsymbol{\rho = cov\left(u_{it}^{wage\ employment}, u_{it}^{Self-employment}\right) = 0} \qquad \textbf{(10)}$$

This does not have to be the case. It is very plausible that farmers make decisions about labor allocation to various activities simultaneously. In which case, there is some efficiency gain in estimating simultaneously both non-farm employment equations. The jointness of the decisions can be justified by the fact that households have a limited amount of labor available, and labor markets are not fully functional. Alternatively there are probably common unobserved factors that affect participation in non-farm activities. In the same spirit as seemingly unrelated regressions, system of binary response equations can be estimated that take into account the potential correlation between the residuals from both equations. Such Multivariate Probit models rely on the joint normal distribution of the errors. Maximum likelihood is applied based on the joint density function (See Green William (2000), chapter 17.5 for a discussion of multivariate Probit models):

$$\boldsymbol{\left(\begin{matrix} u_{it}^{wage\ employment} \\ u_{it}^{self-employment} \end{matrix} \middle| X_{it}, \overline{X_\iota}\right) \sim Normal\left[\binom{0}{0}, \begin{pmatrix} 1 & \rho \\ \rho & 1 \end{pmatrix}\right]} \qquad \textbf{(11)}$$

While this method might generate efficiency gains, the joint normality of the residuals from both non-farm employment equations is a very restrictive assumption since joint normality is not always guaranteed even when each marginal distribution is normally distributed (Wooldridge, 2010). Under joint normality, bivariate Maximum Likelihood estimator algorithms are available in most statistical packages such as STATA. When we have several equations, simulation-based methods such as the Geweke–Hajivassiliou–Keane (GHK) smooth recursive conditioning simulator are recommended to go around the difficulty of managing a non trivial likelihood function (Cappellari and Jenkins, 2003). Because of the potential loss of consistency in case of violation of the joint distributional assumption, we maintain single equations as our main approach, but we estimate the bivariate Probit model using traditional bivariate probit Maximum Likelihood as well as the GHK Simulated Maximum Likelihood for robustness purposes.

4.4.2. Impact of participation in non-farm activities

Next we explore if and how participating in non-farm wage and self-employment improves the welfare of participating households. Does RNFE participation relax rural households' liquidity constraints thereby allowing them to invest more in agricultural production, and achieve higher productivity? Does dualism in access and returns to non-farm employment types creates heterogeneity in the effects of non-farm employment participation across classes of wealth distribution? These questions are still under-investigated in the literature, and data and methodological limitations plague the few studies that have investigated it in other African countries (Owusu et al., 2011, Ackah, 2013, Oseni and Winters, 2009).

Appendix I presents the general framework for identification of treatment effects as applied to rural non-farm employment and informs our selection of the empirical models used. In general, the core equation to be estimated takes the form of an unobserved effects model:

$$\boldsymbol{Y_{it} = G(\beta_0 + \beta_1 SelfEmployment_{it} + \beta_2 WageEmployment_{it} + Z'_{it}\beta_3 + c_i + \varepsilon_{it})} \quad \textbf{(12)}$$

where ε_{it} is the vector of idiosyncratic errors. We allow for an unobservable household heterogeneity effect c_i to be correlated with the explanatory variables in the model. G(.) is a positive function that links the explanatory variables to the dependent variable Y_{it}. It could be a linear or non-linear function (for example: standard normal density, or tobit function) depending on the dependent variable.

To address the endogeneity of the RNFE participation decision within a panel framework, the fixed effects (FE) approach is commonly used. The advantage of the FE method is that it addresses the household's heterogeneity parameter without making any assumption about its distribution. While the CRE (described earlier) is well suited to non-linear models (and cases when explanatory variables that do not change over time are of particular interest) and can be used as an alternative approach, it restricts the distribution of the unobserved parameter (c_i) using the Mundlak-Chamberlain device. Since some of our outcome variables are better represented using non-linear models (e.g. poverty incidence (1/0), poverty gap (0-1), input purchase (1/0) value of input purchases (many zeros due to corner solution), we estimate both the CRE model and the linear FE model and compare results (see appendix 1 for full discussion).

A weakness of both the FE and the CRE approaches is that they only account for time invariant unobservable heterogeneity. Any remaining time-varying unobservable heterogeneity might still lead to inconsistent estimates unless they are effectively taken into account using appropriate instrumental variables. However, instrumental variable methods suffer the difficulty of finding a good instrument that satisfies both identification and rank conditions (See Wooldridge 2002 for a discussion): (i) they should have no partial effect on the outcome variables and should not be correlated with other factors that affect the outcomes variable; and (ii) they must be related, either positively or negatively, to the treatment indicator. Keeping in mind that "*the cure may be worse than the disease*" (Baser, 2009) if a weak instrument is used, we use the FE and CRE approaches without instrumental variable, but we include a large set of time varying explanatory variables to hopefully capture and proxy for as many sources of heterogeneity as possible. We also use a variety of robustness checks to ensure that our results are consistent across methods with different strengths and weaknesses.

4.4.2.1. Outcome variables

The first main outcome variable considered is household per capita consumption expenditure (HHPCE) as a direct measure of household welfare. It includes annual aggregate expenditures on food, non-food, durable goods and housing, deflated by spatial and temporal prices indices to adjust for cost of living differences, and adjusted for household size. The preference for this consumption-based measure as opposed to its main competitor, which is income-based, follows the smoothness argument according to which consumption is less prone to seasonal variations in livings standard, especially in rural areas of developing countries. Therefore consumption aggregates, over a relatively short period of time such as annual, offer a practical advantage over income aggregates by informing more about longer-term living standards (See Deaton and Zaidi (2002) P14 for a discussion on this). The model using HHPCE, as dependent variable, is a linear model that we estimate using FE. We also estimate the treatment effects on Log HHPCE using FE, which gives a sense of percentage effects of our treatment variable, and allow interpretation in relative terms.

For the poverty analysis, the class of poverty metrics proposed by Foster, Greer and Thorbecke (FGT, 1984) are used. Poverty incidence is based on the HHPCE, relative to the 85,852 MKW (real 2013 prices) local poverty line in rural Malawi. As opposed to the 1 or $2/day universal line, the 85,852 MKW local poverty line considered in this paper, which indicates the cost of maintaining a reference welfare level (here defined as satisfying necessary energy and nutritional requirements to have a healthy and active life) to a given person, at a given time in the specific context of Malawi (Ravallion, 1998). It comprises two principal components: food (cost of a food bundle that provides the necessary energy requirements per person) and non-food (allowance for basic non-food needs, estimated as the average non-food consumption of the population whose food consumption is close to the food poverty line). The sum of food and non-food poverty lines gives the total poverty line, which is then converted in real prices for each year using appropriate price index and regional deflators (See LSMS documentation for more details).

Household poverty incidence takes value 1 if household consumption falls below the poverty line, and 0 otherwise. The treatment effects on this variable use mainly a CRE Probit approach. Poverty gap (defined as the consumption shortfall relative to the poverty line, as a fraction of the poverty line) takes a value 0 for all non-poor households, creating a continuous set of values between 0 and 1. The squared poverty gap (sensitive to extreme poverty) takes continuous values between 0 and 1. We use the CRE fractional Probit approach for both these variables. As demonstrated by (Gallani et al., 2015), Fraction Response Model (FRM) is preferable since it overcomes the limitations of other approaches for the statistical analysis of dependent variables that are bounded in nature and present a significant number of observations at one of the boundary points. For all these three poverty measures, we also estimate the linear FE model to serve as benchmark for comparison as it requires less distributional assumptions and generally gives reasonable approximation of APEs (Wooldridge, 2010).

Subjective food security and well-being are captured by four main variables; food insecurity is a binary variable which takes value one if the household head responded yes to the question "*Did you worry that your household would not have enough food in the past 7 days?*". Similarly, food shortage takes value 1 if the household head responded yes to the question "*In the last 12 months, have you been faced with a situation when you did not have enough food to eat*?" Both are binary responses and thus estimated using the CRE Probit approach. The third variable used is food consumption adequacy describing household food consumption over the past month as 1(less than adequate for household needs); 2 (just adequate for household needs); and 3 (It was more than adequate for household needs). Similarly, the fourth variable which is income adequacy describes household perception of their current income and takes discrete values ranging from 1 (Allows you to build your savings); 2(Allows you to save just a little); 3(Only just meets your expenses); 4 (Is not sufficient so need to use savings to meet expenses); 5(Is really not sufficient so you need to borrow to meet expenses). These last 2 variables are discrete choice responses where order matters; therefore, we use ordered Probit approach with correlated random effects to control for time invariant unobserved heterogeneity. For all these four variables, we also estimate the linear FE for comparison purpose.

Finally, to capture the effects of non-farm employment on agricultural investments, we consider household purchase of seeds, fertilizer, and all inputs, as well as area of land cultivated. The effects of non-farm employment on the binary purchase decisions are first considered with a

CRE Probit model. Then, the effects on the quantity of purchases per acre of land cultivated are modeled with a CRE Tobit approach and estimated using partial Maximum likelihood.

4.4.2.2. Distributional effects of non-farm employment participation: Quantile regression approach

Though the average treatment effect is a useful way to summarize the link between participation in non-farm employment and welfare outcomes, it provides an incomplete view of the relationship. Indeed it may hide a lot of heterogeneity in terms of how different groups benefit from non-farm employment participation. Quantile regression methods address this limitation by informing about the relationship between non-farm employment participation and outcome variables at different points of the conditional distribution of the outcome (Cameron and Trivedi, 2010). As such, QR methods can help understand the distributional effects of engaging in non-farm activities. This is usually overlooked in treatment effect estimation discussions and seems quasi nonexistent in the literature on RNFE and welfare effects. Since it has very important implications in terms of what type of policies would be effective and the extent to which targeting might be necessary to improve welfare effects, we explore it. Another advantage of quantile regressions over conditional mean regressions is that the former is robust to outliers because quantiles are not sensitive to extreme values while the mean is.

Appendix 2 provides a formal presentation of the QR approach for estimating the heterogenous welfare effect of RNFE. Assuming linearity, the following conditional quantile equation is estimated for each quantile and each outcome of interest (Koenker, 2005):

$$\boldsymbol{Q_q(Y_{it}(.)|w,Z,c) = \tau_{q0} + \tau_{q1}SelfEmployment_{it} + \tau_{q2}WageEmployment_{it} + Z'_{it}\tau_{q3} + Q_q(v_{it}|w,Z,c)} \qquad \textbf{(13)}$$

where $v_{it} = c_i + u_{it}$ is the composite error. Again we model the unobserved household time invariant characteristics using the CRE device by including time averages of all the explanatory variables as additional regressors in the estimating equation. We then apply pooled quantile regression to the resulting equation. While this way of introducing time averages as regressors is a practical way of introducing CRE into quantile regressions, it is not totally consistent with the unobserved effects quantile model in equation 17. In fact, because the quantile of a sum is different from the sum of quantiles, the independence assumption after including time averages of explanatory variable is still restrictive. We recognize this as a limitation of our estimates. We explore a model without the time averages to check consistency of the results. We estimate and graphically present the quantile treatment effects for several values of q, using a generalized version of the least absolute deviation (LAD) estimation approach (Wooldridge, 2010, Green William, 2000).

4.4.3. Explanatory variables

4.4.3.1. Determinants of participation in non-farm employment

Equation 4 in our conceptual model implies that households allocate labor across different activities in order to equate marginal utility of labor across alternative employment options, given price vector, budget constraint, and time endowment. The imperfection of markets for factors such as labor, land, credit, etc. implies that household decisions are non-separable and based on unobservable shadow prices rather than market prices (De Janvry et al., 1991). Therefore the main explanatory variables included are chosen to capture the human, physical, financial, and other factors that influence the relative shadow prices or marginal values of investing labor in various activities. All the variables used in our empirical analysis and their definitions are summarized in Appendix 4.

First, we control for human capital through education and household composition. Education has consistently been proven to be an important determinant of household participation in non-farm employment as it increases household access to lucrative job opportunities. We control not only for highest education level of the head of the household, but also for maximum level of education attained by any member of the household. As for household composition, we control for the number of household member in each age group (from infant to elderly) as this captures the labor constraint and consumption needs faced by the household.

Household physical capital is captured by household ownership of various physical assets such as normalized wealth index, normalized total livestock unit[38], and land holdings. While these variables characterize usually the wealth of the household, which is often correlated with social status and access to non-farm opportunities, they may also capture the degree of household involvement in agricultural activities implying an opposite effect.

In terms of financial capital, we control for household access to credit, which also determine household's investment capabilities in various activities, be it farm or non-farm related.

Furthermore, we control for various shocks faced by household. The traditional view has it that household are pushed into non-farm employment in response to shocks to their income or consumption (Bardhan and Udry, 1999, Haggblade et al., 2010). We first introduce a binary variable to capture whether the household responded *yes* to having been affected negatively by any shock to their income in the 12 months prior to interview. Then we control for coefficient of

[38] Households' wealth and productive capital indices such as wealth index and agricultural assets index are generated using factor analysis as in FILMER, D. & PRITCHETT, L. H. 2001. Estimating wealth effects without expenditure data—or tears: An application to educational enrollments in states of india*. *Demography,* 38**,** 115-132. The index was then normalized, that is, norm_index = ((index - min(index)) / (max(index)-min(index))).

variation of rainfall in the enumeration area, as well as average 12-month rainfall, and annual mean temperature. These weather-related variables are likely to influence labor allocation to agricultural activities, thereby affecting labor allocation to non-farm employment through the time endowment constraint.

Other controls include household distance to road and population center as they affect transaction costs of searching and engaging in non-farm employment, at the same time as they affect agricultural inputs and output market participation. Value of ganyu wage salary in the community (at the enumeration area level) is included because it influences the opportunity cost of non-farm employment participation. Fertilizer being the main modern agricultural input purchased by farm households, and maize being the primary staple crop produced and consumed by households, we control for price of fertilizer and price of maize grains in the EA, both of which are expected to affect marginal returns of participation in agricultural activities, and thus non-farm activities. The economic environment is expected to affect the availability and attractiveness of non-farm employment opportunities, thereby pulling households into on farm employment participation (Haggblade et al., 2010, Reardon et al., 2007). We control for this variable for each household i by using the rate of non-farm wage employment and non-farm self-employment participation amongst the remaining households in the same enumeration area (excluding the household i).

Consistently with the correlated random effect approach, the time averages of all the explanatory variables are introduced in the estimating equation to control for time invariant heterogeneity. Time dummy taking value 1 for year 2010, and 26 dummy variables to capture the 27 district specific effects are also included whenever possible.

4.4.3.2. Impact of participation in non-farm employment

The empirical models used to analyze the impact of non-farm wage employment and non-farm self-employment on various outcomes include primarily the dummy variables capturing each type of non farm employment, in addition to which we include a set of controls Z_{it} that are expected to be correlated with both non farm employment participation and outcomes. Therefore, the empirical model used for welfare and poverty analysis includes practically all the explanatory variables used in the participation models as control, as they are all likely to affect household consumption expenditure, poverty status, and subjective welfare.

The analysis of household input purchase focuses on farm households and uses, in addition to all the variables above, the total value of crop sales by the households in the main agricultural season to capture the degree of agricultural commercialization by the household, as well as crop mix variables (share of total land cultivated in each crop type: grains, legumes, tubers, oil crops, horticulture crops, cotton crops, tobacco crops, other crops, etc.); because, the need for inputs such as fertilizer varies with the types of crop produced (De Geus and de Geus, 1967).

Again, the time averages of all the explanatory variables are introduced in the estimating equation to control for time invariant heterogeneity. Time dummy taking value 1 for year 2010, and 26 dummy variables to capture the 27 district specific effects are also included when possible.

4.5. Results and discussion

4.5.1. Determinants of participation in non-farm activities

Table 5 summarizes the determinants of participation in non-farm wage and self-employment in rural Malawi based on the unobserved effects Probit model[39]. We find little support for the push factors and early view that the poor and least endowed households diversify more due to risk aversion and capital constraints (Bardhan and Udry, 1999). Our results rather indicate the presence of potential barriers to both non-farm wage and self-employment participation. These barriers include human, financial, and physical constraints such as education, assets and credit. our results are more in line with the relatively more recent literature that reveal the presence of entry barriers preventing the poorest from participating in spite of their higher need for alternative sources of income(Abdulai and CroleRees, 2001, Reardon et al., 2000). Significant differences emerge between non-farm wage and self-employment indicating that different sets of resources matter for participation in each type of non-farm activity. While education tends to be important for engaging in non-farm wage employment, access to credit is very important for participation in non-farm self-employment. Factors like wealth, market access, and proximity to better infrastructure and opportunities are important for both non-farm wage employment participation and non-farm-self-employment.

The average partial effect of household head education is increasing with the level of education till MSCE, beyond which educational attainment level does not appear significant

---

[39] The linear FE model results are presented alongside the CRE results in the table to show consistency across methods. Also, summary statistics of the variables used in the model can be seen in appendices 8 through 11.

anymore, probably due to the low proportion (1.3%) of people in that education level category in the rural Malawi. This is consistent with several previous findings such as in Oseni and Winters (2009), Winters et al. (2007) as well as Lanjouw and Shariff (2004), and De Janvry and Sadoulet (2001), though they have often lumped non-farm wage employment and self-employment together in their analyses, thereby missing important nuances between the two types of activities. Indeed, education may exert two opposing forces influencing households' participation in non-farm-self-employment. On one hand, negative effects could exist since returns to non-farm wage employment seems higher in general than non-farm self-employment (See table 3 and 4), the more educated are pulled into non-farm wage employment, leaving the non-farm self-employment sector for the less educated. On the other hand, those who are more educated have higher expected returns from participation in non-farm self-employment due to skills acquired from their formal training. This would imply a positive effect of education on non-farm self-employment participation. This is reflected in our results, which shows that education of the household head is positive at all levels but not significant determinants of non-farm self-employment participation, while maximum education in the household shows negative effects, but still not significant. This result also explains why the poor in rural Malawi have a higher participate rate in non-farm self-employment than in non-farm wage employment.

We find that access to credit is an important determinant of households' participation in non-farm self-employment. More precisely, access to credit for relevant households would increase their participation in non-farm self-employment by about 8 to 10 percentage points on average, *ceteris paribus (see table 5).* These results are contrary to the common view that rural household enterprises require little to no capital investment and thus can provide a source of cash for the poorest who are lacking access to financial services due to failures in the rural credit market (Poulton et al., 2006, Bardhan and Udry, 1999). Viable household enterprises in the context of rural Malawi seem to require some investments that exclude or limit the participation of poor households facing credit constraints.

Wealthier household (as captured by a normalized asset index) significantly increases participation in both non-farm wage employment and self-employment. This is consistent with the findings in Oseni and Winters (2009), and add to reasons why the poorest have a more restricted access than the non-poor to non-farm employment opportunities. Assets such as tables, sewing machine, TV, bicycle, etc., can be used to operate household enterprises, commute to a non-farm wage or business place, etc., thereby increasing household likelihood (and may be expected returns) of participating in non-farm wage and self-employment. In addition, asset ownership is correlated with social status in most rural areas in Africa. So households with more assets may have stronger networks through which they can receive information about non-farm wage employment opportunities. Depending on the type of business, households with large networks may also have higher expected returns from non-farm enterprises, which increases participation.

More generally, our results reveal some evidence of a pull scenario explaining the participation in non-farm activities in rural Malawi. The participation rate in non-farm wage employment, amongst neighboring households from the same geographical area, affects, positively and significantly, each household's participation in both non-farm wage employment and non-farm self-employment. Similarly, the proportion of other households owning a business in the same geographical area significantly and positively affects each household's likelihood of participation in non-farm self-employment and non-farm wage employment. This is consistent with the pull scenario described in Haggblade et al. (2010), which stresses that the economic environment greatly determines access and participation in non-farm activities. Having a lot of households participating in non-farm wage employment or non-farm self-employment in the same locality may reflect a more vibrant economic environment with more jobs or business opportunities, which therefore increase participation in those employment categories. From a social network effects point of view, this also corresponds with what was dubbed correlated neighborhood effects by (Manski, 1993). People behave like one another when they face similar shocks or environment. However, the observed correlation might also suggest the existence of social effects (endogenous and contextual peer effects), which imply that households' decision to participate in non-farm wage activities depend on their peers behaviors and characteristics. Our analysis does not allow us to disentangle these effects and this is not the primary goal of this article. However the findings that social network effects might influence participation in non-farm self-employment (similarly to agricultural technology adoption) are intrinsically interesting and deserve further investigation. In the social network literature and technology adoption literature, Bramoulle et al. (2009) followed by Krishnan and Patnam (2014) addressed the Manski reflection problem inherent to the identification of peer effects by using average characteristics of neighbors as instruments for participation rates amongst neighbors.

Table 4-5: Average partial effects estimates of determinants of non-farm employment participation

VARIABLES	Non-farm wage employment CRE Probit model	Non-farm wage employment Linear FE model	Non-farm self-employment CRE Probit model	Non-farm self-employment Linear FE model
Age of the household head	-0.000	-0.000	0.000	0.001
	[0.863]	[0.751]	[0.988]	[0.631]
Male-headed household (0/1)	0.072***	0.072***	0.017	0.008
	[0.001]	[0.002]	[0.552]	[0.783]
Highest level of formal education acquired by household head				
PSLC	0.035	0.039	0.003	0.010
	[0.194]	[0.202]	[0.937]	[0.803]
JCE	0.130***	0.120***	0.046	0.055
	[0.007]	[0.008]	[0.365]	[0.290]
MSCE	0.236***	0.257***	0.033	0.029
	[0.004]	[0.000]	[0.651]	[0.697]
Non-University Diploma and above	0.198	0.222	0.286	0.249+
	[0.294]	[0.228]	[0.184]	[0.102]
Maximum level of formal education acquired in the household				
PSLC	0.017	0.016	0.017	0.014
	[0.478]	[0.495]	[0.578]	[0.680]
JCE	0.016	0.015	-0.041	-0.058
	[0.579]	[0.628]	[0.224]	[0.165]
MSCE	0.079+	0.102**	-0.052	-0.056
	[0.105]	[0.030]	[0.271]	[0.294]
Non-University Diploma and above	0.262	0.263+	-0.184***	-0.294*
	[0.167]	[0.127]	[0.000]	[0.095]
Number of infant (<5yo) in HH	0.001	0.006	0.022**	0.019+
	[0.903]	[0.535]	[0.036]	[0.128]
Number of children (5-14yo) in the household	0.015***	0.017**	0.001	-0.000
	[0.006]	[0.011]	[0.930]	[0.974]
Number of prime adults (15-60yo) in HH	-0.003	-0.002	0.022**	0.021*
	[0.669]	[0.793]	[0.022]	[0.065]
Number of elderly (60yo+) in HH	-0.019	-0.023+	-0.006	-0.004
	[0.383]	[0.149]	[0.844]	[0.872]
Household access to loan (0/1)	-0.006	-0.001	0.083***	0.105***
	[0.695]	[0.931]	[0.000]	[0.000]
Normalized wealth index	0.301***	0.519***	0.450***	0.567***
	[0.003]	[0.001]	[0.000]	[0.000]
Normalized TLU index	-0.432*	-0.268	0.503+	0.176
	[0.086]	[0.506]	[0.117]	[0.717]
Total land area owned by HH in Acres	0.006*	0.000	-0.000	-0.000
	[0.100]	[0.735]	[0.825]	[0.860]

Table 4-5 (cont'd)

Household was affected negatively by some income Shock During the last 12 months (0/1)	0.017	0.021	-0.018	-0.037+
	[0.297]	[0.195]	[0.420]	[0.118]
Household has a migrant network (0/1)	0.010	0.008	0.005	0.007
	[0.532]	[0.669]	[0.799]	[0.774]
Rain - EA level CoV of Dec-Jan rainfall from 1983/84 - 2012/13	-0.886***		0.229	
	[0.006]		[0.526]	
Average 12-month total rainfall (mm)	-0.000	0.000	0.000	0.000
	[0.500]	[0.721]	[0.227]	[0.741]
Annual Mean Temperature (∞C * 10)	0.000	0.000	-0.000	-0.002
	[0.838]	[0.909]	[0.839]	[0.419]
HH Distance in (KMs) to Nearest Road	-0.002	-0.004	-0.002	0.000
	[0.326]	[0.263]	[0.589]	[0.948]
HH Distance in (KMs) to Nearest Population Center with +20,000	-0.000	-0.001	-0.000	0.001
	[0.953]	[0.542]	[0.851]	[0.574]
Value of daily *ganyu* wage salary in the EA	-0.000	-0.000	0.000	0.000**
	[0.465]	[0.282]	[0.301]	[0.023]
Price of fertilizer in the EA	-0.000+	-0.000	-0.000	-0.000*
	[0.106]	[0.233]	[0.638]	[0.090]
Price of maize grains in the EA	-0.000	-0.000	0.000	0.000*
	[0.449]	[0.457]	[0.299]	[0.070]
EA neighbor's wage employment participation	0.002***	0.003***	0.001*	0.002**
	[0.001]	[0.003]	[0.088]	[0.026]
EA neighbor's self employment participation	0.001***	0.001**	0.004***	0.004***
	[0.010]	[0.016]	[0.000]	[0.000]
Time dummy (year 2010=1)	-0.002	-0.004	-0.008	-0.022
	[0.929]	[0.888]	[0.787]	[0.363]
District dummies (27 -1 dummies)	Included	Included	Included	Included
Time average of explanatory variables	Included		Included	
Constant		-0.856		0.805
		[0.276]		[0.369]
Number of observations	5,286	5,286	5,286	5,286
Number of households		2,751		2,751

Source: Generated by authors using LSMS data

Note: ***, **, *, and + indicate that the corresponding regression APE is statistically significant at the 1%, 5%, 10%, and 15% levels, respectively. The CRE model is estimated using GEE approach with cluster robust standard errors, which produces notably smaller standard errors than the pooled Probit estimates. The linear FE results are presented for comparison. Pvalues between brackets.

We find large consistency between our linear model results and the CRE Probit estimation results. The magnitudes of the APE estimates are similar across models and the signs and significance levels are consistent. More importantly, the results from the more efficient multivariate system estimates do not add any particular insight (see appendix 3). While the Simulated Maximum Likelihood (SML) estimates and the Bivariate Probit Maximum Likelihood (ML) estimates both do indicate a negative correlation ($\rho = -0.15$) and significant (at 1%) between the residuals of the non farm wage employment and non farm self employment models, the coefficient estimates are remarkably consistent in signs and significance with the main Probit results described above. This robust consistency of our estimates across methods increases our confidence in these results, which indicate that policies geared to increase rural household participation in non-farm employment might want to focus on factors such as education and access to credit for the poor, in addition to improving infrastructure and growth motors.

4.5.2. Impacts of participation in non-farm activities

Next we explore if participation in non-farm activities actually increases households' welfare. Are the poverty reduction motives of policies promoting RNFE in developing countries justified? Our results indicate a positive response to both questions. The ATT estimates indicate generally positive effects of both non-farm wage employment and non-farm self-employment on objective and subjective measures of household welfare, and this is largely consistent across different estimation techniques (see table 6). Participation in the RNFE increases household per capita expenditure and food consumption adequacy while reducing food insecurity, food shortage and income insufficiency. The magnitude of effects on self-employment is on average larger and more consistent than that from wage employment.

Table 4-6: Effect of participation in the non-farm activities on various outcomes in rural Malawi

	Effects of non-farm wage employment participation			**Effects of non-farm self-employment participation**		
VARIABLES	CRE probit/fractional probit / ordered probit	CRE Tobit	Linear FE	CRE probit/fractional probit / ordered probit	CRE Tobit	Linear FE
<u>Objective welfare outcomes</u>						
HHPCE (1000 MKW)			4.476			7.003***
			[0.158]			[0.000]
Log HHPCE			0.102**			0.129***
			[0.022]			[0.000]
Poverty incidence	-0.074**	.	-0.072**	-0.085***		-0.083***
	[0.012]		[0.022]	[0.000]		[0.000]
Poverty gap	-0.034***	-0.033**	-0.034**	-0.035***	-0.035***	-0.030***
	[0.007]	[0.011]	[0.018]	[0.000]	[0.000]	[0.002]
Squared poverty gap	-0.018**	-0.017**	-0.019**	-0.018***	-0.017***	-0.014**
	[0.021]	[0.017]	[0.045]	[0.000]	[0.000]	[0.011]
<u>Subjective welfare outcomes</u>						
Food insecurity (0/1)	-0.047+		-0.065**	-0.034+		-0.027
	[0.125]		[0.041]	[0.140]		[0.281]
Food shortage (0/1)	-0.001		-0.007	-0.006		-0.003
	[0.964]		[0.821]	[0.775]		[0.908]
Food consumption adequacy (1-3)	0.177**		0.104**	0.057		0.025
	[0.037]		[0.021]	[0.379]		[0.482]
Income inadequacy (1-5)	-0.052		-0.077	-0.094+		-0.100*
	[0.495]		[0.471]	[0.112]		[0.094]
<u>Agricultural outcomes</u>						
Fertilizer purchase decision (0/1)	0.039		0.037	0.045**		0.055**
	[0.190]		[0.303]	[0.043]		[0.036]
Inputs purchase decision (0/1)	0.046		0.037	0.053**		0.060**
	[0.160]		[0.260]	[0.031]		[0.039]
Value of fertilizer purchase (1000 MKW)		0.156	-0.638		0.388	0.136
		[0.745]	[0.419]		[0.225]	[0.786]
Value of inputs purchase (1000 MKW)		-0.124	-0.874		0.234	0.085
		[0.802]	[0.305]		[0.495]	[0.870]
Land cultivated (1000 acres)		-0.292	0.274		0.710	-0.576
		[0.699]	[0.800]		[0.561]	[0.893]

Source: Generated by authors using LSMS data

Note: ***, **, *, and + indicate that the corresponding regression APE is statistically significant at the 1%, 5%, 10%, and 15% levels, respectively. The table presents only APEs on our main treatment variables. The APEs from the CRE probit models are estimated using GEE approach. The CRE fractional Probit APEs are from GLM estimation. The unconditional APEs from CRE Tobit are estimated using pooled Tobit regression methods. The linear FE results are presented for comparison purpose. The full results tables are in appendices 2, 4, 5, and 6. P-values between brackets are based on cluster robust standard errors

Non-farm wage employment participation increases households' per capita consumption expenditure by a margin of 4,500MKW, though only significant at 16%. The log linear model indicates a statistically significant effect of 10 percent. As for non-farm self-employment the ATT estimates on the levels of HHPCE is 7,000MKW corresponding approximately to 13 percent increase in HHPCE. These numbers are not directly comparable with previous studies such as Owusu et al. (2011) and Ackah (2013) as they have lumped both types of non-farm activities together and have also focused on different outcome variables (notably household income instead of consumption expenditure). However, the general conclusion of a positive effect of non-farm employments on direct measures of household welfare seems to be an empirical regularity as described by (Barrett et al., 2001b). The t-test of equality of coefficients of non-farm self employment and non farm wage employment indicates a significant (at 1%) difference confirming that the effect of non farm self-employment on participants is higher on average than the effects of non farm wage employment participation.

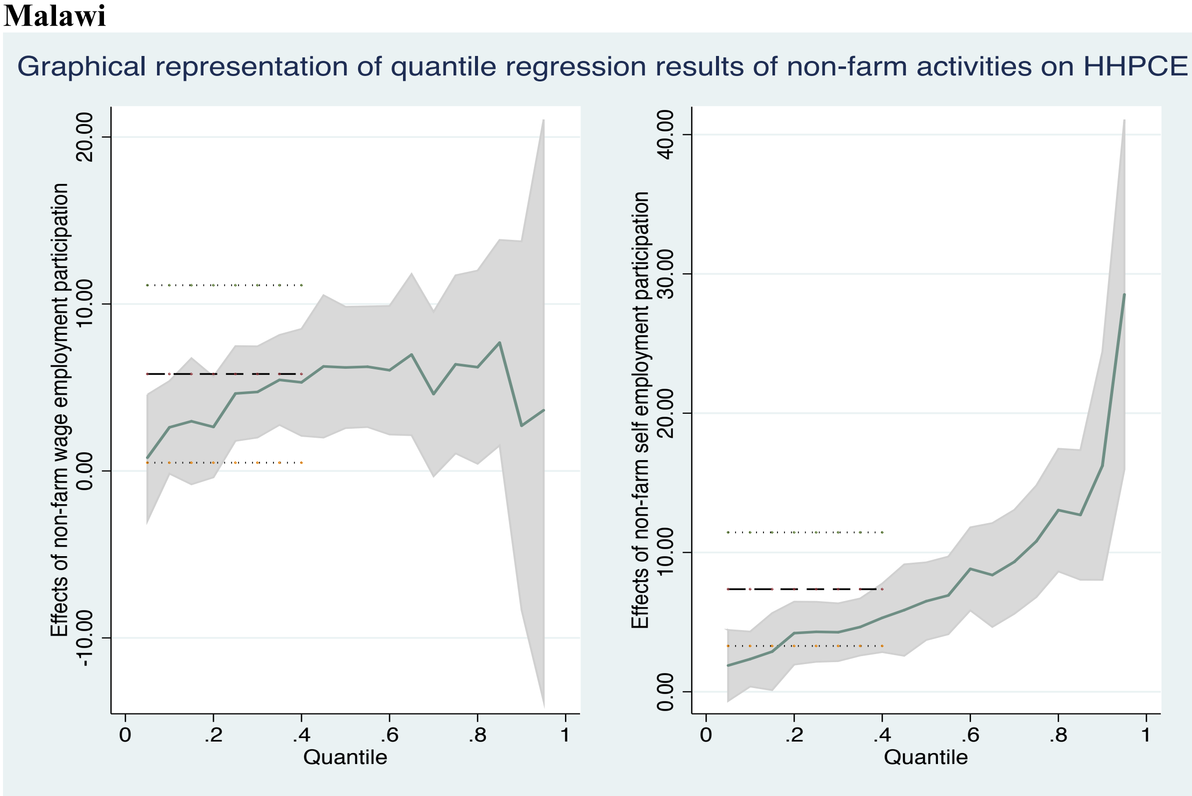
However, we also find significant heterogeneity in the welfare effect of RNFE participation. Everything else held constant, households at the bottom of the wealth distribution benefit less from both non-farm wage employment and non-farm self-employment than households at the top of the distribution (table 7 and figure 1). Though the effects of engaging in non-farm wage employment or non-farm self-employment remain positive for all classes of the welfare distribution, there is a generally increasing trend in the size of the effect as we go from lower percentiles to the top of the distribution of HHPCE.

Table 4-7: Distributional effects of participation in the non-farm activities on HHPCE in Malawi (Quantile regression results)

VARIABLES	p10	p20	p25	p50	p75	p80	p90
Effects of non-farm wage employment participation	2.611+	2.634+	4.639***	6.198***	6.386**	6.214*	2.709
	[0.114]	[0.138]	[0.006]	[0.005]	[0.016]	[0.086]	[0.714]
Effects of non-farm self-employment participation	2.347**	4.209***	4.303***	6.503***	10.804***	13.041***	16.227**
	[0.018]	[0.001]	[0.000]	[0.000]	[0.000]	[0.000]	[0.016]
Other controls included							
Time averages included							

Source: Generated by authors using LSMS data
Note: ***, **, *, and + indicate that the corresponding regression coefficient is statistically significant at the 1%, 5%, 10%, and 15% levels, respectively. The table presents quantile partial effects of our main treatment variables from 10th to the 90th percentile. The full results tables are in appendix 3. P-values between.

Figure 4.1: Quantile effects of non-farm wage and self-employment on HHPCE in rural Malawi



Source: Generated by authors using LSMS data
Note: The graph presents quantile partial effects of our main treatment variable

The quantile effects of non-farm self-employment go from a low of about 2,300MKW for the 10th percentile of HHPCE, and increases monotonically to a high of over 16,000MKW for the 90 percentile (almost 10 times the effect on the lowest quantile). As for non-farm wage employment, the effect does not even seem significant for the 10th and 20th percentiles. It starts significant from a low of 4,600MKW for the 25th percentile, to a high of a little less than 6,500 MKW for the 75th and 80th percentiles, before dropping again to a non-significant size of 2,000MKW at the 90th percentile[40]. This indicates that, for non-farm wage employment, the middle class benefits the most; while, for non-farm self-employment, the upper class benefits the most from participating. However, in both cases the poorest benefit the least from participating, which illustrates the fact that resource constraints faced by the poorest consign them to low return activities with less potential to increase their income significantly. Meanwhile the middle class and the upper class earn significantly higher returns from participating in non-farm activities because they have access to the most lucrative employment opportunities. These results are echoed by the descriptive analysis presented in section 3 about the sectors of non farm employment in which poor and non poor participate, as well as the evidence of dualism in the returns from participation. Efforts to address this could significantly increase the magnitude of RNFE effect on household welfare in rural areas.

We extend beyond HHPCE to analyze several other welfare outcomes. Our results indicate that both non-farm wage employment and non-farm self-employment affect negatively and significantly, all three FGT poverty metrics. Both the fractional Probit model and the linear

---

[40] The results from the linear quantile model without additional time averages (available from the authors upon request) also lead to the same conclusion of unequal welfare benefits from both Non-Farm Wage Employment and Non-Farm Self Employment participation. This indicate that our result is not driven by the choice of this practical way of including CRE in order to account for unobserved heterogeneity.

FE model show consistently that, on average, a household's engagement in non-farm wage employment reduces by 7 percentage points, the likelihood of its consumption expenditure falling below the poverty line. Similarly, self-employment reduces poverty incidence by a statistically significant margin of about 8.5 percentage points. This is an important margin relative to the overall poverty rate of about 37 percent in rural Malawi (See summary statistics in appendix 4). As for the depth and severity of poverty measured by poverty gap and squared poverty gap respectively, our CRE fractional Probit, Tobit, and linear FE models produce consistent results indicating significant and negative effects of both non-farm wage employment and nonfarm self-employment. In magnitude, we observe an average of about 3.3 percentage points reduction in poverty gap, and 2 percentage points reduction in squared poverty gap. This implies that though the poorest seem to benefit less from engaging in non-farm activities, the size of the effect is still enough to reduce significantly the depth and severity of poverty confirming that the rural non-farm sector can indeed be seen as a pathway out of poverty for poor rural households in Malawi.

The positive effect of non-farm activities persists in the analysis of various household subjective perceptions of food security and well-being[41]. In particular, households participating in non-farm activities feel more food and income secure than non-participating households; though this effect is not significant for non-farm self-employment participants. Both non-farm wage employment and non-farm self-employment reduce the likelihood that a household experience food shortage but this effect is not statistically significant. Similarly, the ordered Probit results show that households engaged in non-farm activities feel that both their food

[41] Recognizing the potential for strategic responses in self-reported welfare measures, we do not put much emphasis on the subjective welfare effects.

consumption and income level are more adequate for the needs of the household than non-participating households.

Several studies in the RNFE literature have explored agricultural investment as a pathway through which non-farm employment increase welfare, and concluded a positive effect. For example, Oseni and Winters (2009) concluded a positive effect of both non-farm wage employment and non farm self-employment participation on Nigerian farmers' agricultural investments (especially labor and fertilizer). Other studies such as Smale et al. (2016) found negative effects on nitrogen application on maize by Kenyan farmers, and concluded that engagement in the RNFE could be a distraction with trade-offs in labor allocation and farm investments. Our results (see bottom panel of table 6) indicate that participation in non-farm employment increases inputs use (especially inorganic fertilizer) and purchases among rural households in Malawi. These findings tend to align more with the strand of existing literature finding evidence that RNFE reduces the cash constraint for rural households at least partially engaged in agriculture such as Owusu et al. (2011) and Ackah (2013) as well as discussion in Barrett et al. (2001b) on the subject. However, while the estimates indicate that non-farm self-employment increases likelihood of purchasing external agricultural inputs, especially fertilizer, by about 5 percentage points, they show no significant effect on (unconditional) amounts of inputs purchased. As for non-farm wage employment it affects positively input purchase decision and amount but not significantly. The lack of significance of the effects of non-farm wage employment could imply that the negative effect on agriculture due to labor displacement abates the positive effect on financial constraints for agricultural investments. The area of land cultivated in the season is not significantly influenced by households' participation in either non-farm employments, which could be an illustration of land constraints faced by households due to imperfections in rural land markets (Bardhan and Udry, 1999).

It is possible that decisions to participate in non-farm employment and decision to use inputs are made simultaneously, and therefore a system estimation of non-farm employment participation and inputs use equations would improve the efficiency of the estimates; though identification relies on joint normality of the residuals from the equations in the system. For additional robustness check, we estimate a series of recursive trivariate probit models as an alternative way to explore the effects of non-farm employment participation on inputs purchase decisions (Appendix 3). In each system, we have one equation for each treatment variable (wage employment and self employment), and one equation for the outcome variable (fertilizer purchase, or inputs purchase), making a total of three equations. The equation for the outcome variable includes both treatment variables as explanatory variables, while the equations for the treatment variables do not include the outcome variable as explanatory variable. This recursive formulation (see equation 17-49 in Green William (2000)) implies a one-way causality between the treatments and outcome variables. A similar approach was used by (Smale et al., 2016) to investigate the same question in Kenya. We estimate this system, for input purchase and for fertilizer purchase decisions using GHK Simulated Maximum Likelihood (see Cappellari and Jenkins (2003) for a discussion). The estimated coefficients are consistent with the single equation results presented above suggesting a positive effect of non-farm employment participation on inputs purchase decision in general, and fertilizer purchase in particular. More interestingly while the effect of non-farm wage employment was not significant in the single equation models, the more efficient system estimations indicates a rather positive and significant effect of non farm wage employment on inputs purchase decisions. The likelihood ratio tests of joint significance of the residuals from the equations are always significant at 1%, justifying the relevance of a system estimation approach in our case.

4.6. Conclusion

This article makes several contributions to the ongoing debate around the poverty reduction potential of the RNFE. Using panel data estimation techniques applied on nationally representative data for Malawi, we find consistent evidence that the RNFE can serve as a potential mechanism to increase welfare and reduce poverty in rural Malawi. Our results are consistent across a broad suite of objective and subjective welfare and poverty measures. We conclude that, unless there are time varying unobservable heterogeneity non captured in our empirical framework, the context of rural Malawi offers consistent evidence of a positive welfare impact of participation in non-farm activities, and thus the RNFE can be rightly considered as a pathway out of poverty for rural households.

However, our results indicate the need for more attention to be paid to improving not just the access, but also the quality of the non-farm opportunities available for the poorest. Though on average, participation in non-farm wage employment and non-farm self-employment consistently has a positive effect on welfare, we find strong evidence that even when the poor participate in the RNFE, they benefit significantly less than the non-poor. Our results indicate that low education, access to credit, and assets limit their participation in non-farm employment in rural areas.

We find some important differences between non-farm wage employment and non-farm self-employment (in terms of mains determinants), which are overlooked in studies that lump both together. While education matters more for participation in non-farm wage employment than non-farm self-employment, access to credit seems more important for non-farm self-employment. This result stresses the importance of making a clear difference between different types of non-farm employment in order to get a better understanding of their determinants and roles. Given the notable heterogeneity within each type of non-farm employment, future researches should delve even deeper into those different types and explore which types of employment the poorest rural households have access to and why.

Finally, our results indicate that participation in non-farm self-employment improves the use of modern inputs (especially inorganic fertilizer). If increased agricultural productivity through modern inputs use is desired, then supporting a most inclusive development of rural enterprises seems be a potential avenue that can effectively complement interventions in other areas such as credit and insurance related programs. Other factors such as improved access to markets and better infrastructures are also likely to amplify the effects of RNFE, with real implications for broad-based economic growth and poverty reduction in rural Africa.

APPENDICES

Appendix 1: A general framework for the identification of rural non-farm employment treatment effects

Considering any outcome of interest Y, the participation effect or Average Treatment Effect on households participating in non-farm activities (ATT) is defined as:

$$\boldsymbol{ATT = E(Y_{1i} - Y_{0i}|W_i = 1) = E(Y_{1i}|W_i = 1) - E(Y_{0i}|W_i = 1)} \quad \textbf{(14)}$$

where E(.) is the expectation operator, Y_{1i} is the potential outcome of household i under participation in non-farm employment, Y_{0i} is potential outcome when household i does not participate in non-farm employment. W_i is the treatment indicator equal to 1 if household actually participates in non-farm employment, and 0 otherwise. The fundamental challenge in measuring the above ATT is related to the fact that we cannot observe Y_{0i} for a household that is participating in non-farm activities. For any individual household at a given time, he is either engaged in non-farm employment and we observe Y1i |Wi=1 or he is not engaged in non-farm employment and we observe Y0i|Wi=0. The consequent challenge for any impact evaluation exercise is the process of finding a counterfactual or a proxy for the unobserved potential outcome.

A naïve and usually wrong approach (unless participation is truly random) is to compare the observed outcomes of participants and non-participants. As explained by Angrist and Pischke (2008), this introduces a selection bias unless independence assumption in satisfied. Using the Rubin Causal Model, we can decompose the ATT in two components that make obvious the bias introduced by using the naïve approach (Rubin and John, 2011).

$$\boldsymbol{E(Y_{1i}|W_i = 1) - E(Y_{0i}|W_i = 0) = ATT + \ E(Y_{0i}|W_i = 1) - E(Y_{0i}|W_i = 0)} \quad \textbf{(15)}$$

Under the independence assumption, the potential outcome is independent of treatment status: $(Y_0, Y_1) \perp w$. This implies that the observed average outcome of non participants $E(Y_{0i}|W_i = 0)$ is an accurate estimate of the potential outcome of the participants households had they not engage in non-farm activities $E(Y_{0i}|W_i = 1)$; in which case, $Bias = E(Y_{0i}|W_i = 1) - E(Y_{0i}|W_i = 0) = 0$. If participation in non-farm activities were randomly assigned, as in an experimental setting, the independence assumption would be satisfied and we could use the naïve approach with no concern for bias. But, in our case, independence assumption is most likely violated, because rural households self-select into non-farm activities given their expected marginal utility from participation, which is in turn a function of household and community level observed and unobserved characteristics. A comparison of socio economic characteristics of participants and non-participants in non-farm wage employment and non-farm self-employment (see appendix 4) reveals that RNFE participants and non-participants in rural Malawi are different and confirms the commonly held view that that participants in non-farm employment are those likely to benefit more from it, because they have capital endowment that increase their returns to non-farm participation (Reardon et al., 1998). Not taking this into account is likely to introduce an upward bias in our estimates.

A less restrictive assumption is the conditional independence assumption (CIA), also called selection-on-observables, according to which potential outcomes are independent from treatment status once observable characteristics are controlled for, that is $(Y_0, Y_1) \perp w \mid \boldsymbol{Z}$. This implies that conditional on observed characteristics Z, the experimental conditions are restored, and selection bias disappears (Angrist and Pischke, 2008). If we could observe and control for all possible systematic differences between participant and non-participants, we could simply use the following multivariate regression models to capture ATT without being concerned about selection bias.

$$\boldsymbol{E(Y_{it}|X) = \beta_0 + \beta_1 SelfEmployment_{it} + \beta_2 WageEmployment_{it} + Z'_{it}\beta_3} \qquad \textbf{(16)}$$

where $SelfEmployment_{it}$ and $WageEmployment_{it}$ are binary variables capturing participation in non-farm self-employment and non-farm wage employment respectively by household i in period t. Z_{it} is a vector of control variables. While this is more likely to be useful when very rich information is available for households, it is limited if there are several unobserved characteristics such as households' preferences for risks, income variability, entrepreneurship that cannot be accurately observed and captured even though, they are potentially correlated with both non-farm participation decision and the outcome. Ignoring those variables can lead to omitted variables bias in our treatment effects estimate.

Fortunately, panel data such as ours offer possibilities to accommodate unobserved heterogeneity that is time invariant by means of the fixed effects and correlated random effects. In general, the core equation to be estimated takes the form of an unobserved effect model:

$$\boldsymbol{Y_{it} = G(\beta_0 + \beta_1 SelfEmployment_{it} + \beta_2 WageEmployment_{it} + Z'_{it}\beta_3 + c_i + \varepsilon_{it})} \quad \textbf{(17)}$$

where ε_{it} is the vector of idiosyncratic errors. We allow for an unobservable household heterogeneity effect c_i to be correlated with the explanatory variables in the model. G(.) is a positive function that links the explanatory variables to the dependent variable Y_{it}. It could be a linear or non-linear function (for example: standard normal density, or tobit function) depending on the dependent variable.

The advantage of the fixed effects method is that it deals effectively with the household's heterogeneity parameter ci, without making any assumption about its distribution. As described earlier, CRE plays a similar role as the FE method, except that it restricts the distribution of c_i using the Mundlak-Chamberlain device. The benefit of the CRE is that it performs well in non-linear models such as Probit and fractional response models, where FE methods bear incidental parameter issue (Wooldridge, 2010) and in cases when explanatory variables that do not change over time are of particular interest. Some of our outcome variables are better represented using non-linear models. For example, the poverty incidence is a binary variable which we capture using a Probit model; the poverty gap and squared poverty gap are captured by fractional Probit models; the inputs purchase decisions are captured by Probit models; and value of input purchases uses are captured using Tobit models to take into account the corner responses. For each of these outcomes, we estimate both the CRE model and the linear FE model and compare results.

Appendix 2: Distributional effects of non-farm employment participation: A quantile effect model

For values of q between 0 and 1, we define the conditional quantile $Q_q(Y(.)|w,Z) = F^{-1}_{Y(.)|w,Z}(q)$, as the value of outcome Y (HHPCE in our case) that splits the data into the proportions q below, and 1-q above, where $F^{-1}(.)$ represents the cumulative distribution function CDF of potential outcome Y(.). As described by Imbens and Wooldridge (2008) as well as Cameron and Trivedi (2010) we specify the q-th quantile treatment effect as the average difference between quantiles of the two marginal potential outcome distributions:

$$\boldsymbol{\tau_q(Z) = F^{-1}_{Y(1)|Z}(q) - F^{-1}_{Y(0)|Z}(q)} \tag{18}$$

Assuming linearity, the following conditional quantile equation is estimated for each quantile and each outcome of interest (Koenker, 2005):

$$\boldsymbol{Q_q(Y_{it}(.)|w,Z,c) = \tau_{q0} + \tau_{q1}SelfEmployment_{it} + \tau_{q2}WageEmployment_{it} + Z'_{it}\tau_{q3} + Q_q(v_{it}|w,Z,c)} \tag{19}$$

where $v_{it} = c_i + u_{it}$ is the composite error. Notice we allow for household unobserved heterogeneity ci to affect the conditional quantile function. However, even though we postulate a linear model, fixed effects approach allowing us to eliminate ci without restricting its conditional distribution is subject to incidental parameter problem for small T (T=2 in our case). Again we model ci using the CRE device by including time averages of all the explanatory variables as additional regressors in the estimating equation. We then apply pooled quantile regression to the resulting equation. We must recognize that with quantile regression, the independence assumption after including time averages of explanatory variable is still restrictive; because, the quantile of a sum is different from the sum of quantiles. We are therefore cautious not to make strong causal statements in terms of our quantile effects estimates. We estimate and graph the quantile treatment effects for several values of q, using a generalized version of the least absolute deviation (LAD) estimation approach (Wooldridge, 2010).

Appendix 3: Full regression tables

Table 4-8: Seemingly unrelated system equation estimates of non-farm wage employment and non-farm self-employment participation model

VARIABLES	GHK Simulated bivariate ML estimates		Seemingly unrelated bivariate ML Probit estimates	
	Wage employment	Self employment	Wage employment	Self employment
Age of the household head	-0.000	-0.000	-0.000	-0.000
	[0.982]	[0.977]	[0.984]	[0.983]
Male-headed household (0/1)	0.423***	0.061	0.423***	0.061
	[0.003]	[0.599]	[0.003]	[0.596]
Highest level of formal education acquired by household head				
PSLC	0.192	0.003	0.192	0.001
	[0.270]	[0.984]	[0.270]	[0.993]
JCE	0.621***	0.158	0.622***	0.156
	[0.004]	[0.425]	[0.004]	[0.430]
MSCE	0.991***	0.117	0.991***	0.118
	[0.001]	[0.685]	[0.001]	[0.683]
Non-University Diploma and above	0.894	0.942+	0.896	0.934+
	[0.169]	[0.127]	[0.169]	[0.130]
Maximum level of formal education acquired in the household				
PSLC	0.097	0.071	0.098	0.072
	[0.525]	[0.568]	[0.524]	[0.565]
JCE	0.102	-0.157	0.102	-0.154
	[0.579]	[0.332]	[0.580]	[0.341]
MSCE	0.408*	-0.211	0.408*	-0.211
	[0.094]	[0.360]	[0.094]	[0.360]
Non-University Diploma and above	1.099*	-1.195**	1.098*	-1.192**
	[0.066]	[0.037]	[0.067]	[0.038]
Number of infant (<5yo) in HH	0.007	0.086*	0.007	0.087*
	[0.906]	[0.064]	[0.904]	[0.062]
Number of children (5-14yo) in the household	0.087**	0.003	0.087**	0.003
	[0.024]	[0.925]	[0.024]	[0.923]
Number of prime adults (15-60yo) in HH	-0.022	0.087**	-0.022	0.087**
	[0.604]	[0.015]	[0.602]	[0.015]
Number of elderly (60yo+) in HH	-0.103	-0.016	-0.102	-0.015
	[0.499]	[0.897]	[0.502]	[0.905]
Household access to loan (0/1)	-0.042	0.329***	-0.042	0.329***
	[0.674]	[0.000]	[0.671]	[0.000]
Normalized wealth index	1.813***	1.776***	1.810***	1.771***
	[0.005]	[0.002]	[0.005]	[0.002]
Normalized TLU index	-2.788+	1.882	-2.791+	1.873
	[0.117]	[0.192]	[0.117]	[0.194]

Table 4-8 (cont'd)

Total land area owned by HH in Acres	0.042**	-0.000	0.042**	-0.000
	[0.011]	[0.794]	[0.011]	[0.788]
Household was affected negatively by some income Shock During the last 12 months (0/1)	0.104	-0.064	0.104	-0.063
	[0.364]	[0.501]	[0.364]	[0.510]
Household has a migrant network (0/1)	0.042	0.021	0.042	0.021
	[0.711]	[0.819]	[0.713]	[0.821]
Rain - EA level CoV of Dec-Jan rainfall from 1983/84 - 2012/13			-5.370***	1.022
			[0.002]	[0.454]
Average 12-month total rainfall (mm)	-0.001	0.002	-0.001	0.002
	[0.571]	[0.265]	[0.570]	[0.266]
Annual Mean Temperature (∞C * 10)	0.002	-0.002	0.002	-0.002
	[0.828]	[0.814]	[0.825]	[0.816]
HH Distance in (KMs) to Nearest Road	-0.014	-0.008	-0.014	-0.008
	[0.318]	[0.544]	[0.316]	[0.548]
HH Distance in (KMs) to Nearest Population Center with +20,000	0.001	-0.001	0.001	-0.001
	[0.936]	[0.860]	[0.938]	[0.871]
Value of daily *ganyu* wage salary in the EA	-0.000	0.000	-0.000	0.000
	[0.544]	[0.329]	[0.544]	[0.322]
Price of fertilizer in the EA	-0.001+	-0.000	-0.001+	-0.000
	[0.120]	[0.764]	[0.120]	[0.754]
Price of maize grains in the EA	-0.001	0.001	-0.001	0.001
	[0.445]	[0.404]	[0.446]	[0.398]
EA neighbor's wage employment participation	0.012***	0.006*	0.012***	0.006*
	[0.003]	[0.091]	[0.003]	[0.087]
EA neighbor's self employment participation	0.007**	0.016***	0.007**	0.016***
	[0.028]	[0.000]	[0.028]	[0.000]
Time dummy (year 2010=1)	-0.029	-0.020	-0.029	-0.019
	[0.852]	[0.878]	[0.852]	[0.883]
District dummies (27 -1 dummies)	Included	Included	Included	Included
Time average of explanatory variables	Included	Included	Included	Included
rho		-0.145***		-0.150***
		[0.000]		[0.000]
Constant	-0.885	-2.354**	-0.890	-2.358**
	[0.437]	[0.013]	[0.434]	[0.012]
Observations	5,286	5,286	5,286	5,286

Source: Generated by authors using LSMS data

Note: The symbols ***, **, *, and + indicate that the corresponding regression coefficient is statistically significant at the 1%, 5%, 10%, and 15% levels, respectively. P-values between brackets. The GHK estimates are generated using the mvprobit command with 100 draws in STATA. The bivariate ML estimates are generated using the biprobit command in STATA. The parameters in the table are coefficients and not APEs.

Table 4-9: Effect of participation in the non-farm activities on HHPCE in Malawi, FE estimates

VARIABLES	(2) HHPCE FE	(4) Log HHPCE FE
Wage employment participation	4.476	0.102**
	[0.158]	[0.022]
Self-employment participation	7.003***	0.129***
	[0.000]	[0.000]
Age of the household head	-0.053	-0.002
	[0.633]	[0.232]
Male-headed household (0/1)	2.895	0.063+
	[0.269]	[0.101]
Highest level of formal education acquired by household head		
PSLC	6.198**	0.036
	[0.048]	[0.401]
JCE	15.719***	0.187***
	[0.005]	[0.005]
MSCE	20.869*	0.197**
	[0.054]	[0.049]
Maximum level of formal education acquired in the household		
PSLC	-2.228	-0.011
	[0.350]	[0.765]
JCE	-0.755	0.020
	[0.814]	[0.667]
MSCE	-1.458	0.004
	[0.774]	[0.950]
Number of infant (<5yo) in HH	-9.266***	-0.139***
	[0.000]	[0.000]
Number of children (5-14yo) in the household	-8.872***	-0.136***
	[0.000]	[0.000]
Number of prime adults (15-60yo) in HH	-6.768***	-0.094***
	[0.000]	[0.000]
Number of elderly (60yo+) in HH	-7.836***	-0.118***
	[0.002]	[0.001]
Household access to loan (0/1)	1.635	0.034
	[0.436]	[0.221]
Normalized TLU index	140.889**	1.326**
	[0.045]	[0.032]
Normalized agricultural asset index + Normalized land holdings	41.423***	0.689***
	[0.000]	[0.000]
Household was affected negatively by some income Shock During the last 12 months (0/1)	1.644	0.027
	[0.473]	[0.446]
Household has a migrant network (0/1)	-1.124	-0.004
	[0.585]	[0.906]
Rain - EA level CoV of Dec-Jan rainfall from 1983/84 - 2012/13		

Table 4-9 (cont'd)

Average 12-month total rainfall (mm)	0.064	0.001+
	[0.228]	[0.134]
Annual Mean Temperature (∞C * 10)	0.413	0.000
	[0.261]	[0.931]
HH Distance in (KMs) to Nearest Road	0.006	-0.001
	[0.987]	[0.784]
HH Distance in (KMs) to Nearest Population Center with +20,000	-0.085	-0.003
	[0.689]	[0.160]
Price of fertilizer in the EA	-0.013	-0.000
	[0.403]	[0.604]
Price of maize grains in the EA	0.083**	0.001**
	[0.024]	[0.029]
Time dummy (year 2010=1)	-2.343	-0.037
	[0.567]	[0.623]
District dummies	Included	Included
Time average of explanatory variables		
Constant	-100.016	9.223***
	[0.354]	[0.000]
Observations	5,321	5,321
R-squared	0.209	0.249
Number of households	2,764	2,764

Source: Generated by authors using LSMS data

Note: The dependent variable in column 1 and 2 are levels of HHPCE. In column 3 and 4, the dependent variable is Log(HHPCE). The symbols ***, **, *, and + indicate that the corresponding regression coefficients are statistically significant at the 1%, 5%, 10%, and 15% levels, respectively. P-values based on clustered robust standard errors between brackets.

Table 4-10: Effects of participation in the RNFE on quintiles of HHPCE

VARIABLES	p10	p20	p25	p50	p75	p80	p90
Wage employment participation	2.611+	2.634+	4.639***	6.198***	6.386**	6.214*	2.709
	[0.114]	[0.138]	[0.006]	[0.005]	[0.016]	[0.086]	[0.714]
Self-employment participation	2.347**	4.209***	4.303***	6.503***	10.804***	13.041***	16.227**
	[0.018]	[0.001]	[0.000]	[0.000]	[0.000]	[0.000]	[0.016]
Age of the household head	-0.042	0.005	-0.013	-0.079	-0.161	-0.192+	-0.392*
	[0.524]	[0.939]	[0.829]	[0.341]	[0.198]	[0.144]	[0.057]
Male-headed household (0/1)	1.870	2.858*	1.598	2.211	6.659*	6.856*	2.388
	[0.268]	[0.063]	[0.335]	[0.399]	[0.060]	[0.091]	[0.740]
Highest level of formal education acquired by household head							
PSLC	-0.349	-1.262	-0.916	0.929	9.709*	10.191*	6.786
	[0.828]	[0.423]	[0.658]	[0.717]	[0.055]	[0.070]	[0.466]
JCE	5.945**	8.658***	7.995**	9.316**	21.622***	18.744**	20.544
	[0.011]	[0.004]	[0.011]	[0.033]	[0.006]	[0.034]	[0.210]
MSCE	2.274	4.883	5.615	16.777**	36.827***	43.716**	28.356
	[0.479]	[0.444]	[0.348]	[0.014]	[0.009]	[0.016]	[0.424]
Maximum level of formal education acquired in the household							
PSLC	0.626	0.779	0.259	-0.171	-4.342	-6.652	-1.599
	[0.734]	[0.693]	[0.915]	[0.943]	[0.316]	[0.223]	[0.829]
JCE	0.325	1.269	0.372	1.340	-5.894	-4.363	-1.536
	[0.887]	[0.539]	[0.885]	[0.703]	[0.265]	[0.521]	[0.881]
MSCE	2.153	1.823	-0.258	-0.767	-15.890+	-10.826	-0.373
	[0.397]	[0.678]	[0.957]	[0.896]	[0.135]	[0.405]	[0.983]
Number of infant (<5yo) in HH	-2.766***	-4.092***	-4.592***	-6.785***	-9.154***	-10.031***	-9.751***
	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]
Number of children (5-14yo) in the household	-3.433***	-4.063***	-4.556***	-5.737***	-6.848***	-7.026***	-8.640***
	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]
Number of prime adults (15-60yo) in HH	-2.939***	-3.856***	-4.221***	-4.003***	-4.337***	-5.659***	-8.076***
	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]
Number of elderly (60yo+) in HH	-3.744**	-7.800***	-6.678***	-4.867**	-7.138*	-6.759	-1.831
	[0.022]	[0.000]	[0.001]	[0.024]	[0.086]	[0.199]	[0.780]
Household access to loan (0/1)	2.727**	1.593	1.089	0.723	0.871	0.757	2.754
	[0.048]	[0.318]	[0.452]	[0.728]	[0.632]	[0.784]	[0.549]

Table 4-10 (cont'd)

Normalized TLU index	32.917	12.544	48.236	89.698*	99.853+	19.958	13.738
	[0.242]	[0.790]	[0.281]	[0.097]	[0.138]	[0.759]	[0.912]
Normalized agricultural asset index + Normalized land holdings	19.665***	22.493***	23.349***	27.503***	32.530***	32.966***	46.026***
	[0.000]	[0.000]	[0.000]	[0.000]	[0.001]	[0.002]	[0.006]
Household was affected negatively by some income Shock During the last 12 months (0/1)	0.662	0.474	1.783	0.960	4.360+	6.643+	5.915+
	[0.593]	[0.644]	[0.167]	[0.647]	[0.131]	[0.115]	[0.110]
Household has a migrant network (0/1)	-0.970	-1.881+	-1.215	0.127	0.220	2.956	2.504
	[0.340]	[0.121]	[0.396]	[0.948]	[0.938]	[0.361]	[0.583]
Rain - EA level CoV of Dec-Jan rainfall from 1983/84 - 2012/13	-13.093	-22.098	-28.994	-39.669*	-60.347+	-71.742	34.300
	[0.515]	[0.348]	[0.247]	[0.087]	[0.141]	[0.187]	[0.677]
Average 12-month total rainfall (mm)	0.019	0.013	0.011	0.019	0.026	0.025	0.014
	[0.443]	[0.634]	[0.733]	[0.561]	[0.508]	[0.566]	[0.857]
Annual Mean Temperature (∞C * 10)	-0.096	-0.076	-0.022	-0.003	0.201	0.528	1.146*
	[0.406]	[0.537]	[0.886]	[0.988]	[0.683]	[0.244]	[0.077]
HH Distance in (KMs) to Nearest Road	-0.100	-0.151	-0.148	-0.034	-0.189	-0.365	-0.212
	[0.640]	[0.382]	[0.437]	[0.859]	[0.694]	[0.483]	[0.772]
HH Distance in (KMs) to Nearest Population Center with +20,000	0.052	0.062	0.026	-0.109	-0.045	-0.043	-0.555+
	[0.635]	[0.642]	[0.817]	[0.431]	[0.758]	[0.870]	[0.116]
Price of fertilizer in the EA	-0.006	-0.012**	-0.011*	-0.014*	-0.018	-0.017	-0.010
	[0.390]	[0.038]	[0.064]	[0.086]	[0.173]	[0.295]	[0.731]
Price of maize grains in the EA	0.056**	0.033	0.026	0.073***	0.082***	0.097**	0.195**
	[0.024]	[0.154]	[0.323]	[0.009]	[0.003]	[0.011]	[0.010]
Time dummy (year 2010=1)	0.005	-3.124+	-3.459*	-2.524	-3.625	-2.797	4.653
	[0.998]	[0.138]	[0.075]	[0.341]	[0.297]	[0.388]	[0.571]
							[.]
Time average of explanatory variables							
District dummies							
Constant	36.094**	33.481**	30.597**	44.849***	62.689*	65.833+	34.729
	[0.011]	[0.016]	[0.028]	[0.002]	[0.079]	[0.100]	[0.489]
	5,321	5,321	5,321	5,321	5,321	5,321	5,321

Source: Generated by authors using LSMS data

Note: ***, **, *, and + indicate that the corresponding regression coefficient is statistically significant at the 1%, 5%, 10%, and 15% levels, respectively. The table presents quantile partial effects of explanatory variables from 10^{th} to the 90^{th} percentile. P-values between brackets. Standard errors are not considerably different from the cluster robust quantile results.

Table 4-11: Effects of participation in the non-farm activities on poverty incidence, gap and severity in rural Malawi

VARIABLES	Poverty incidence		Poverty gap				Poverty severity			
	CRE Probit	Linear FE	CRE Probit	Fractional CRE probit	CRE Tobit	Linear FE	CRE Probit	Fractional CRE probit	CRE Tobit	Linear FE
Wage employment participation	-0.074**	-0.072**	-0.034***	-0.034**	-0.033**	-0.034**	-0.018**	-0.018*	-0.017**	-0.019**
	[0.012]	[0.022]	[0.007]	[0.019]	[0.011]	[0.018]	[0.021]	[0.056]	[0.017]	[0.045]
Self-employment participation	-0.085***	-0.083***	-0.035***	-0.035***	-0.035***	-0.030***	-0.018***	-0.018***	-0.017***	-0.014**
	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]	[0.002]	[0.000]	[0.001]	[0.000]	[0.011]
Age of the household head	0.003**	0.002*	0.001+	0.001	0.001*	0.001	0.000+	0.000	0.000*	0.000
	[0.027]	[0.069]	[0.110]	[0.154]	[0.064]	[0.279]	[0.142]	[0.195]	[0.058]	[0.301]
Male-headed household (0/1)	-0.044+	-0.045	-0.025**	-0.026**	-0.023**	-0.022+	-0.010*	-0.011*	-0.010*	-0.008
	[0.120]	[0.206]	[0.020]	[0.020]	[0.035]	[0.103]	[0.091]	[0.077]	[0.076]	[0.315]
Highest level of formal education acquired by household head										
PSLC	0.043	0.060+	0.005	0.003	0.008	0.016	-0.000	-0.001	0.003	0.007
	[0.288]	[0.142]	[0.775]	[0.849]	[0.611]	[0.272]	[0.990]	[0.939]	[0.676]	[0.392]
JCE	-0.096*	-0.061	-0.044***	-0.046**	-0.042**	-0.015	-0.019**	-0.020**	-0.018**	-0.001
	[0.051]	[0.282]	[0.004]	[0.012]	[0.017]	[0.431]	[0.013]	[0.030]	[0.031]	[0.954]
MSCE	-0.075	-0.010	-0.037	-0.040	-0.035	0.004	-0.014	-0.015	-0.013	0.008
	[0.347]	[0.888]	[0.185]	[0.176]	[0.217]	[0.869]	[0.376]	[0.342]	[0.318]	[0.576]
Maximum level of formal education acquired in the household										
PSLC	-0.002	-0.016	0.008	0.008	0.004	-0.004	0.004	0.004	0.002	-0.003
	[0.949]	[0.644]	[0.554]	[0.565]	[0.729]	[0.750]	[0.577]	[0.598]	[0.783]	[0.694]
JCE	-0.070*	-0.069+	-0.014	-0.015	-0.020	-0.015	-0.009	-0.009	-0.011+	-0.009
	[0.077]	[0.119]	[0.328]	[0.345]	[0.162]	[0.347]	[0.280]	[0.312]	[0.113]	[0.331]
MSCE	-0.090+	-0.077	-0.026	-0.026	-0.030	-0.026	-0.012	-0.012	-0.015	-0.012
	[0.122]	[0.201]	[0.227]	[0.299]	[0.166]	[0.304]	[0.377]	[0.422]	[0.174]	[0.394]
Number of infant (<5yo) in HH	0.074***	0.078***	0.028***	0.028***	0.030***	0.030***	0.013***	0.013***	0.015***	0.014***
	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]
Number of children (5-14yo) in the household	0.074***	0.081***	0.029***	0.029***	0.030***	0.030***	0.014***	0.014***	0.015***	0.014***
	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]
Number of prime adults (15-60yo) in HH	0.052***	0.047***	0.024***	0.024***	0.024***	0.021***	0.012***	0.012***	0.012***	0.010***
	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]
Number of elderly (60yo+) in HH	0.104***	0.103***	0.040***	0.040***	0.042***	0.034***	0.019***	0.019***	0.020***	0.015**
	[0.001]	[0.003]	[0.000]	[0.000]	[0.000]	[0.010]	[0.001]	[0.002]	[0.000]	[0.048]
Household access to loan (0/1)	-0.000	-0.003	-0.015*	-0.015*	-0.010	-0.019**	-0.013***	-0.013***	-0.008*	-0.016***
	[0.984]	[0.921]	[0.091]	[0.091]	[0.277]	[0.030]	[0.008]	[0.006]	[0.066]	[0.001]
Normalized TLU index	-1.535**	-0.921*	-0.426	-0.434	-0.537*	-0.102	-0.224	-0.222	-0.262*	-0.018
	[0.045]	[0.086]	[0.181]	[0.201]	[0.056]	[0.554]	[0.217]	[0.256]	[0.057]	[0.835]
Normalized agricultural asset index + Normalized land holdings	-0.478***	-0.465***	-0.195***	-0.197***	-0.198***	-0.194***	-0.102***	-0.103***	-0.101***	-0.100***
	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]

Table 4-11 (cont'd)

Household was affected negatively by some income Shock During the last 12 months (0/1)	-0.015	-0.023	-0.006	-0.005	-0.005	-0.003	0.001	0.001	-0.000	0.004
	[0.541]	[0.477]	[0.523]	[0.608]	[0.613]	[0.818]	[0.844]	[0.856]	[1.000]	[0.593]
Household has a migrant network (0/1)	0.007	0.016	-0.006	-0.006	-0.001	-0.007	-0.004	-0.005	-0.001	-0.006
	[0.769]	[0.551]	[0.517]	[0.528]	[0.875]	[0.519]	[0.367]	[0.375]	[0.787]	[0.368]
Rain - EA level CoV of Dec-Jan rainfall from 1983/84 - 2012/13	0.900**		0.171	0.164	0.239		0.057	0.056	0.098	
	[0.018]		[0.223]	[0.431]	[0.249]		[0.448]	[0.603]	[0.317]	
Average 12-month total rainfall (mm)	-0.000	-0.001	-0.000**	-0.000*	-0.000	-0.000*	-0.000**	-0.000*	-0.000	-0.000+
	[0.745]	[0.209]	[0.020]	[0.061]	[0.201]	[0.100]	[0.013]	[0.055]	[0.190]	[0.130]
Annual Mean Temperature (∞C * 10)	0.001	-0.000	0.001	0.001	0.001	0.001	0.000	-0.000	0.000	-0.000
	[0.659]	[0.919]	[0.515]	[0.554]	[0.494]	[0.632]	[0.975]	[0.988]	[0.667]	[0.968]
HH Distance in (KMs) to Nearest Road	-0.001	-0.000	-0.001	-0.001	-0.000	-0.002	-0.001	-0.001	-0.000	-0.001
	[0.835]	[0.949]	[0.466]	[0.502]	[0.787]	[0.249]	[0.355]	[0.360]	[0.660]	[0.246]
HH Distance in (KMs) to Nearest Population Center with +20,000	-0.001	0.000	0.000	0.000	-0.000	0.001+	0.000	0.000	0.000	0.000
	[0.510]	[0.897]	[0.890]	[0.885]	[0.863]	[0.128]	[0.539]	[0.527]	[0.959]	[0.193]
Price of fertilizer in the EA	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
	[0.632]	[0.892]	[0.341]	[0.413]	[0.486]	[0.792]	[0.450]	[0.520]	[0.484]	[0.767]
Price of maize grains in the EA	-0.001**	-0.001+	-0.000***	-0.000*	-0.000*	-0.000*	-0.000***	-0.000*	-0.000*	-0.000*
	[0.022]	[0.112]	[0.003]	[0.055]	[0.054]	[0.059]	[0.008]	[0.083]	[0.063]	[0.084]
Time dummy (year 2010=1)	0.010	0.011	0.015	0.015	0.013	0.012	0.010	0.011	0.008	0.010
	[0.786]	[0.857]	[0.273]	[0.436]	[0.513]	[0.586]	[0.177]	[0.332]	[0.409]	[0.462]
Time average included	Included		Included	Included	Included		Included	Included	Included	
District dummies included	Included		Included	Included	Included		Included	Included	Included	
Constant		1.578+				0.629+				0.388*
		[0.137]				[0.131]				[0.099]
Observations	5,317	5,321	5,317	5,321	5,321	5,321	5,317	5,321	5,321	5,321
R-squared		0.130				0.141				0.107
Number of households		2,764				2,764				2,764

Source: Generated by authors using LSMS data

Note: ***, **, *, and + indicate that the corresponding regression APE is statistically significant at the 1%, 5%, 10%, and 15% levels, respectively. The table presents APEs on all explanatory variables. The APEs from the CRE probit models are estimated using GEE approach with cluster robust standard errors, which produces notably smaller standard errors than the pooled Probit estimates. The CRE fractional Probit APEs are from GLM estimation. The unconditional APEs from CRE Tobit are estimated using pooled Tobit estimation methods. The linear FE results are presented for comparison purpose. P-values between brackets are based on cluster robust standard errors

Table 4-12: Effects of participation in the non-farm activities on food security and subjective well being

VARIABLES	Food insecure		Food shortage		Adequacy of food consumption		Inadequacy of household income	
	CRE probit	Linear FE	CRE probit	Linear FE	CRE ordered probit	Linear FE	CRE ordered probit	Linear FE
Wage employment participation	-0.047+	-0.065**	-0.001	-0.007	0.177**	0.104**	-0.052	-0.077
	[0.125]	[0.041]	[0.964]	[0.821]	[0.037]	[0.021]	[0.495]	[0.471]
Self-employment participation	-0.034+	-0.027	-0.006	-0.003	0.057	0.025	-0.094+	-0.100*
	[0.140]	[0.281]	[0.775]	[0.908]	[0.379]	[0.482]	[0.112]	[0.094]
Age of the household head	0.000	-0.000	-0.002*	-0.003**	-0.002	-0.000	-0.008**	-0.009**
	[0.743]	[0.952]	[0.097]	[0.037]	[0.591]	[0.793]	[0.021]	[0.036]
Male-headed household (0/1)	0.016	0.009	-0.066*	-0.070**	0.103	0.045	-0.258***	-0.289***
	[0.596]	[0.792]	[0.050]	[0.045]	[0.268]	[0.372]	[0.002]	[0.006]
Highest level of formal education acquired by household head								
PSLC	0.071+	0.074+	0.056	0.068	-0.068	-0.027	0.051	0.040
	[0.121]	[0.146]	[0.172]	[0.185]	[0.576]	[0.691]	[0.644]	[0.749]
JCE	-0.053	-0.067	-0.042	-0.051	0.055	0.038	-0.135	-0.224
	[0.355]	[0.252]	[0.468]	[0.427]	[0.730]	[0.607]	[0.348]	[0.174]
MSCE	-0.141**	-0.119	-0.236***	-0.214**	0.344+	0.180+	-0.092	-0.162
	[0.036]	[0.151]	[0.004]	[0.018]	[0.116]	[0.108]	[0.646]	[0.480]
Maximum level of formal education acquired in the household								
PSLC	-0.029	-0.029	-0.021	-0.016	0.107	0.042	0.012	0.037
	[0.385]	[0.430]	[0.554]	[0.673]	[0.290]	[0.442]	[0.900]	[0.738]
JCE	0.087*	0.095*	0.038	0.038	0.177	0.083	-0.089	-0.067
	[0.072]	[0.055]	[0.391]	[0.410]	[0.169]	[0.236]	[0.449]	[0.596]
MSCE	0.092	0.113+	0.073	0.089	-0.069	-0.066	-0.269*	-0.257
	[0.188]	[0.137]	[0.203]	[0.155]	[0.698]	[0.504]	[0.099]	[0.151]
Number of infant (<5yo) in HH	0.023*	0.026*	0.039***	0.037***	-0.063*	-0.032+	0.066*	0.068*
	[0.073]	[0.056]	[0.003]	[0.008]	[0.093]	[0.102]	[0.051]	[0.077]
Number of children (5-14yo) in the household	0.019**	0.024***	0.027***	0.028***	-0.057**	-0.031**	0.049**	0.055**
	[0.029]	[0.007]	[0.003]	[0.008]	[0.025]	[0.032]	[0.033]	[0.037]
Number of prime adults (15-60yo) in HH	0.008	0.010	0.014	0.014	-0.073**	-0.034*	0.043*	0.052+
	[0.461]	[0.506]	[0.183]	[0.307]	[0.011]	[0.089]	[0.096]	[0.141]
Number of elderly (60yo+) in HH	0.035	0.021	0.065*	0.074**	-0.157*	-0.073+	0.144*	0.148
	[0.301]	[0.613]	[0.051]	[0.035]	[0.095]	[0.117]	[0.088]	[0.167]
Household access to loan (0/1)	0.022	0.024	0.061**	0.069**	-0.135*	-0.067*	0.248***	0.254***
	[0.377]	[0.403]	[0.014]	[0.010]	[0.052]	[0.087]	[0.000]	[0.002]

Table 4-12 (cont'd)

Normalized TLU index	-1.001+	-0.904*	-2.516***	-2.338***	1.198	1.257	-2.233**	-2.304+
	[0.103]	[0.068]	[0.000]	[0.000]	[0.295]	[0.163]	[0.032]	[0.101]
Normalized agricultural asset index + Normalized land holdings	-0.179***	-0.173**	-0.150**	-0.160**	0.876***	0.414***	-0.046	0.032
	[0.006]	[0.022]	[0.014]	[0.018]	[0.000]	[0.000]	[0.759]	[0.880]
Household was affected negatively by some income Shock During the last 12 months (0/1)	0.112***	0.107***	0.161***	0.190***	-0.274***	-0.147***	-0.041	-0.049
	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]	[0.002]	[0.537]	[0.652]
Household has a migrant network (0/1)	0.019	0.030	0.069***	0.071**	0.023	-0.001	0.223***	0.245***
	[0.465]	[0.296]	[0.006]	[0.023]	[0.754]	[0.977]	[0.001]	[0.009]
Rain - EA level CoV of Dec-Jan rainfall from 1983/84 - 2012/13	0.425		-0.545		-0.531		0.154	
	[0.270]		[0.187]		[0.616]		[0.876]	
Average 12-month total rainfall (mm)	0.001	0.001	-0.000	-0.000	-0.001	-0.001	-0.001	-0.001
	[0.291]	[0.184]	[0.564]	[0.668]	[0.613]	[0.399]	[0.670]	[0.633]
Annual Mean Temperature (∞C * 10)	-0.000	0.001	0.000	-0.001	0.006	0.003	0.001	0.003
	[0.832]	[0.784]	[0.891]	[0.618]	[0.330]	[0.450]	[0.811]	[0.688]
HH Distance in (KMs) to Nearest Road	0.006+	0.004	0.009**	0.010**	0.001	0.005	0.014+	0.022**
	[0.123]	[0.293]	[0.014]	[0.014]	[0.919]	[0.331]	[0.116]	[0.032]
HH Distance in (KMs) to Nearest Population Center with +20,000	0.002	0.000	-0.001	0.004+	-0.004	-0.003	-0.005	-0.004
	[0.337]	[0.909]	[0.744]	[0.144]	[0.404]	[0.444]	[0.294]	[0.564]
Price of fertilizer in the EA	0.000	0.000	-0.000**	-0.000	0.000	0.000	-0.000	-0.000
	[0.222]	[0.203]	[0.040]	[0.230]	[0.740]	[0.895]	[0.811]	[0.717]
Price of maize grains in the EA	0.002***	0.002***	-0.000	-0.000	-0.002**	-0.001*	0.002**	0.003+
	[0.000]	[0.002]	[0.693]	[0.864]	[0.026]	[0.078]	[0.022]	[0.115]
year_2011	0.101***	0.122**	-0.166***	-0.133***	-0.114	-0.066	0.369***	0.406***
	[0.009]	[0.036]	[0.000]	[0.005]	[0.263]	[0.335]	[0.000]	[0.002]
Constant		-2.209*		0.529		2.298+		4.227
		[0.061]		[0.663]		[0.115]		[0.169]
Observations	5,317	5,321	5,317	5,321	5,321	5,321	5,321	5,321
R-squared		0.058		0.110		0.044		0.059
Number of Households		2,764		2,764	2,764	2,764	2,764	2,764

Source: Generated by authors using LSMS data

Note: ***, **, *, and + indicate that the corresponding regression APE is statistically significant at the 1%, 5%, 10%, and 15% levels, respectively. The table presents APEs except in the ordered Probit cases. The APEs from the CRE probit models are estimated using GEE approach. The coefficients form the CRE ordered Probit are from pooled ordered Probit. The linear FE results are presented for comparison purpose. P-values between brackets are based on cluster robust standard errors.

Table 4-13: Effects of participation in non-farm activities on inputs purchases for farm households only

	Seed purchase decision (0/1)		Fertilizer purchase decision (0/1		Inputs purchase decision (0/1)		Value of seed purchase per acre of land cultivated		Value of fertilizer purchase per acre of land cultivated		Value of inputs purchase per acre of land cultivated		Land cultivated in acre	
VARIABLES	CRE probit	Linear FE	CRE probit	Linear FE	CRE probit	Linear FE	CRE Tobit	Linear FE	CRE Tobit	Linear FE	CRE Tobit	Linear FE	CRE Tobit	Linear FE
Wage employment participation	-0.007	-0.009	0.039	0.037	0.046	0.037	-0.092	-0.174	0.156	-0.638	-0.124	-0.874	-0.292	0.274
	[0.833]	[0.825]	[0.190]	[0.303]	[0.160]	[0.260]	[0.319]	[0.273]	[0.745]	[0.419]	[0.802]	[0.305]	[0.699]	[0.800]
Self-employment participation	-0.002	0.001	0.045**	0.055**	0.053**	0.060**	-0.050	-0.093	0.388	0.136	0.234	0.085	0.710	-0.576
	[0.948]	[0.984]	[0.043]	[0.036]	[0.031]	[0.039]	[0.426]	[0.292]	[0.225]	[0.786]	[0.495]	[0.870]	[0.561]	[0.893]
Age of the household head	-0.001	-0.000	-0.004***	-0.003**	-0.001	-0.000	-0.003	-0.005	-0.034+	-0.001	-0.007	-0.011	0.031	0.303
	[0.566]	[0.872]	[0.003]	[0.015]	[0.654]	[0.777]	[0.352]	[0.242]	[0.111]	[0.961]	[0.714]	[0.639]	[0.241]	[0.312]
Male-headed household (0/1)	-0.040	-0.055	-0.013	-0.003	-0.035	-0.027	-0.067	-0.078	-0.707	-1.346**	-1.105**	-1.417**	-1.272	3.526
	[0.270]	[0.188]	[0.689]	[0.942]	[0.311]	[0.535]	[0.428]	[0.348]	[0.211]	[0.026]	[0.030]	[0.023]	[0.179]	[0.623]
Highest level of formal education acquired by household head														
PSLC	0.060	0.059	-0.006	-0.017	0.022	0.017	0.143	0.140	-0.395	-0.887	-0.226	-0.764	0.895+	13.556
	[0.198]	[0.234]	[0.882]	[0.693]	[0.620]	[0.708]	[0.230]	[0.357]	[0.458]	[0.291]	[0.702]	[0.392]	[0.138]	[0.270]
JCE	-0.001	-0.028	0.075	0.077	-0.026	-0.061	0.084	0.248	0.976	1.177	0.496	1.481	1.209	6.053
	[0.989]	[0.684]	[0.182]	[0.165]	[0.684]	[0.287]	[0.603]	[0.295]	[0.260]	[0.423]	[0.602]	[0.339]	[0.270]	[0.264]
MSCE	-0.084	-0.103	0.062	0.065	-0.033	-0.028	-0.054	0.448	0.755	2.555	0.489	3.071	1.690	8.861
	[0.254]	[0.247]	[0.442]	[0.385]	[0.712]	[0.719]	[0.817]	[0.242]	[0.590]	[0.346]	[0.760]	[0.279]	[0.273]	[0.158]
Maximum level of formal education acquired in the household														
PSLC	-0.001	-0.012	0.015	0.017	-0.003	-0.010	0.007	0.005	0.243	0.151	0.058	0.199	-1.708*	-17.457
	[0.974]	[0.797]	[0.644]	[0.644]	[0.931]	[0.806]	[0.939]	[0.963]	[0.647]	[0.803]	[0.909]	[0.756]	[0.065]	[0.267]
JCE	0.123**	0.109**	-0.032	-0.002	0.059	0.067	0.223*	0.017	-1.003*	-1.702**	-0.428	-1.653*	-1.362	-9.212
	[0.015]	[0.044]	[0.457]	[0.973]	[0.221]	[0.171]	[0.087]	[0.924]	[0.079]	[0.046]	[0.506]	[0.074]	[0.159]	[0.299]
MSCE	0.167***	0.182***	-0.013	0.038	0.092+	0.122*	0.617**	0.561*	0.288	1.060	1.623	1.620	-2.040*	-11.930
	[0.010]	[0.010]	[0.822]	[0.550]	[0.148]	[0.078]	[0.014]	[0.063]	[0.819]	[0.617]	[0.310]	[0.472]	[0.096]	[0.217]
Number of infant (<5yo) in HH	0.006	0.011	-0.024*	-0.022*	-0.001	0.003	0.007	0.024	-0.214	-0.023	-0.047	-0.039	2.165	9.402
	[0.648]	[0.446]	[0.054]	[0.090]	[0.953]	[0.849]	[0.834]	[0.650]	[0.319]	[0.938]	[0.818]	[0.904]	[0.296]	[0.304]

Table 4-13 (cont'd)

Number of children (5-14yo) in the household	-0.004	-0.000	0.006	0.003	0.004	0.006	-0.026	-0.038	-0.095	-0.314+	-0.149	-0.301	0.525	2.975
	[0.689]	[0.998]	[0.529]	[0.765]	[0.664]	[0.549]	[0.403]	[0.425]	[0.505]	[0.125]	[0.345]	[0.194]	[0.460]	[0.477]
Number of prime adults (15-60yo) in HH	0.001	0.003	0.027***	0.021**	0.014	0.011	-0.005	0.006	0.189	-0.109	0.040	-0.106	0.243	-0.202
	[0.959]	[0.811]	[0.008]	[0.049]	[0.221]	[0.326]	[0.871]	[0.910]	[0.232]	[0.642]	[0.809]	[0.665]	[0.248]	[0.776]
Number of elderly (60yo+) in HH	0.027	0.021	-0.003	-0.011	-0.007	-0.014	0.004	-0.046	-0.161	-0.376	-0.312	-0.365	-0.151	-6.750
	[0.475]	[0.595]	[0.927]	[0.734]	[0.837]	[0.687]	[0.965]	[0.725]	[0.743]	[0.549]	[0.555]	[0.604]	[0.777]	[0.340]
Household access to loan (0/1)	0.006	0.022	0.013	0.017	0.030	0.046+	0.040	0.101	0.833**	1.291*	0.918**	1.170*	-0.166	1.353
	[0.809]	[0.483]	[0.566]	[0.504]	[0.263]	[0.111]	[0.587]	[0.453]	[0.035]	[0.065]	[0.025]	[0.080]	[0.896]	[0.702]
Normalized TLU index	-0.155	-0.396	0.519	0.941*	0.552	0.436	0.076	-0.580	16.426**	47.688***	14.816*	45.936***	1.380	13.952
	[0.736]	[0.604]	[0.278]	[0.067]	[0.473]	[0.433]	[0.925]	[0.812]	[0.013]	[0.005]	[0.075]	[0.008]	[0.678]	[0.549]
Normalized agricultural asset index + Normalized land holdings	0.084	0.107	0.260***	0.252***	0.210***	0.219***	0.208	0.234	2.757***	1.438	2.326**	1.643	2.436	-7.698
	[0.194]	[0.168]	[0.000]	[0.000]	[0.001]	[0.000]	[0.286]	[0.431]	[0.001]	[0.286]	[0.013]	[0.263]	[0.198]	[0.321]
Household was affected negatively by some income Shock During the last 12 months (0/1)	0.002	0.001	-0.006	0.000	-0.008	-0.005	0.004	-0.037	0.030	0.712	0.194	0.588	1.490+	4.532+
	[0.949]	[0.972]	[0.815]	[0.989]	[0.765]	[0.873]	[0.963]	[0.727]	[0.940]	[0.211]	[0.653]	[0.294]	[0.131]	[0.123]
Household has a migrant network (0/1)	-0.078***	-0.078**	0.027	0.013	-0.035	-0.046+	-0.140**	-0.023	-0.180	-0.794+	-0.622+	-0.692	0.555	4.839
	[0.006]	[0.011]	[0.281]	[0.639]	[0.204]	[0.129]	[0.034]	[0.795]	[0.662]	[0.104]	[0.135]	[0.175]	[0.363]	[0.174]
Rain - EA level CoV of Dec-Jan rainfall from 1983/84 - 2012/13	0.979**		-1.141***		0.253		2.698**		-13.870+		2.814		-25.396	
	[0.030]		[0.005]		[0.559]		[0.036]		[0.147]		[0.749]		[0.483]	
Average 12-month total rainfall (mm)	0.000	0.001	-0.001	0.000	-0.001	-0.000	-0.002	-0.004	-0.020**	-0.035	-0.026**	-0.042+	-0.009	-0.081
	[0.958]	[0.484]	[0.150]	[0.899]	[0.203]	[0.938]	[0.396]	[0.478]	[0.036]	[0.154]	[0.045]	[0.145]	[0.474]	[0.175]
Annual Mean Temperature (∞C * 10)	-0.003	-0.003	-0.007***	-0.002	-0.006**	-0.004+	-0.012**	-0.021*	-0.095***	-0.110+	-0.113***	-0.138**	0.026	0.278
	[0.310]	[0.343]	[0.004]	[0.479]	[0.025]	[0.102]	[0.021]	[0.086]	[0.001]	[0.107]	[0.000]	[0.035]	[0.609]	[0.212]
HH Distance in (KMs) to Nearest Road	-0.004	-0.009*	-0.001	-0.003	-0.002	-0.005	-0.020**	-0.036**	-0.115*	-0.266**	-0.156**	-0.302**	0.165	0.034
	[0.394]	[0.075]	[0.764]	[0.644]	[0.691]	[0.308]	[0.030]	[0.016]	[0.057]	[0.024]	[0.011]	[0.014]	[0.279]	[0.860]
HH Distance in (KMs) to Nearest Population Center with +20,000	-0.001	0.005	-0.003	-0.003	-0.003	-0.001	0.006	0.033	-0.000	0.067	0.023	0.095	-0.037	0.114
	[0.610]	[0.210]	[0.171]	[0.244]	[0.217]	[0.768]	[0.467]	[0.283]	[0.994]	[0.609]	[0.669]	[0.550]	[0.267]	[0.532]
Price of fertilizer in the EA	0.000	0.000	-0.000+	-0.000	-0.000	-0.000	0.000	0.000	-0.004*	-0.002	-0.001	-0.002	0.003	0.072
	[0.891]	[0.939]	[0.109]	[0.162]	[0.437]	[0.442]	[0.535]	[0.543]	[0.058]	[0.416]	[0.634]	[0.428]	[0.327]	[0.231]

Table 4-13 (cont'd)

Price of maize grains in the EA	0.000	-0.000	-0.000	0.000	0.000	0.000	0.001	0.001	0.006	0.011	0.010	0.013	-0.007	0.022
	[0.522]	[0.838]	[0.982]	[0.950]	[0.541]	[0.969]	[0.402]	[0.427]	[0.371]	[0.181]	[0.211]	[0.151]	[0.669]	[0.768]
Share of total land cultivated in crops														
grains crops	-0.004**	0.000	0.001	-0.006**	-0.004**	-0.004	-0.004+	0.009+	0.027	0.081*	-0.001	0.088*	0.012	-0.124
	[0.018]	[0.912]	[0.348]	[0.045]	[0.016]	[0.189]	[0.140]	[0.104]	[0.281]	[0.083]	[0.923]	[0.065]	[0.600]	[0.371]
legumes crops	-0.002	0.002	0.001	-0.006**	-0.003*	-0.004	-0.002	0.008	-0.003	0.032	-0.026*	0.037	0.018	-0.121
	[0.175]	[0.320]	[0.650]	[0.029]	[0.053]	[0.294]	[0.468]	[0.170]	[0.916]	[0.514]	[0.096]	[0.452]	[0.460]	[0.431]
tubers crops	-0.003+	0.001	-0.001	-0.008***	-0.005**	-0.005+	0.001	0.016*	-0.014	0.040	-0.030	0.054	-0.027+	-0.243
	[0.102]	[0.676]	[0.647]	[0.009]	[0.013]	[0.148]	[0.883]	[0.099]	[0.654]	[0.415]	[0.207]	[0.297]	[0.108]	[0.162]
oil crops	-0.004+		0.007***		0.001		-0.010*		0.044		-0.010		0.026	
	[0.117]		[0.007]		[0.811]		[0.080]		[0.372]		[0.828]		[0.497]	
Horticulture crops	0.001	0.005**	0.001	-0.006**	-0.002	-0.002	-0.000	0.004	-0.004	0.033	-0.019	0.037	0.010	-0.210
	[0.784]	[0.016]	[0.472]	[0.038]	[0.336]	[0.503]	[0.956]	[0.544]	[0.875]	[0.493]	[0.320]	[0.450]	[0.434]	[0.371]
cotton crops	-0.002	0.003	-0.002	-0.009***	-0.003	-0.004	-0.005	0.003	-0.060*	-0.020	-0.049***	-0.013	0.003	-0.149
	[0.389]	[0.415]	[0.346]	[0.006]	[0.227]	[0.313]	[0.228]	[0.638]	[0.092]	[0.690]	[0.010]	[0.789]	[0.927]	[0.290]
tobacco crops	-0.003*	0.001	0.007***	0.000	0.002	0.001	-0.004	0.006	0.095***	0.158***	0.065***	0.159***	0.110	0.089
	[0.089]	[0.602]	[0.000]	[0.999]	[0.366]	[0.866]	[0.170]	[0.374]	[0.001]	[0.002]	[0.000]	[0.002]	[0.294]	[0.295]
Other crops		0.005*		-0.007**		-0.001		0.006		0.044		0.046		-0.475
		[0.057]		[0.032]		[0.820]		[0.281]		[0.384]		[0.372]		[0.376]
Total value of crop sales (MKW)	0.000*	0.000**	0.000	0.000	-0.000	-0.000	0.000**	0.000	0.000***	0.000***	0.000***	0.000***	-0.000	-0.000
	[0.081]	[0.036]	[0.793]	[0.891]	[0.787]	[0.409]	[0.046]	[0.229]	[0.001]	[0.000]	[0.000]	[0.000]	[0.532]	[0.621]
Time dummy (year 2010=1)	0.025	-0.010	-0.045	-0.028	0.001	-0.015	-0.086	-0.262**	-1.800***	-2.110***	-1.294**	-2.397***	0.088	15.212
	[0.545]	[0.837]	[0.206]	[0.490]	[0.989]	[0.741]	[0.434]	[0.047]	[0.002]	[0.002]	[0.046]	[0.002]	[0.964]	[0.339]
District dummies included														
Constant		0.311		0.873		1.835+		6.362		43.646+		55.424*		-46.400
		[0.810]		[0.455]		[0.105]		[0.350]		[0.117]		[0.076]		[0.430]
Observations	5,038	5,038	5,038	5,038	5,038	5,038	5,038	5,038	5,038	5,038	5,038	5,038	5,038	5,038
R-squared		0.049		0.073		0.049		0.058		0.119		0.131		0.033
Number of households		2,675		2,675		2,675		2,675		2,675		2,675		2,675

Source: Generated by authors using LSMS data

Note: ***, **, *, and + indicate that the corresponding regression APE is statistically significant at the 1%, 5%, 10%, and 15% levels, respectively. The table presents APEs on all explanatory variables. The APEs from the CRE probit models are estimated using GEE approach. The unconditional APEs from CRE Tobit are estimated using pooled Tobit regression methods. The linear FE results are presented for comparison purpose. P-values between brackets are based on cluster robust standard errors.

Table 4-14: Multivariate recursive Probit estimation of effects of non-farm employment participation on activities on inputs purchases for farm households only.

	Fertilizer purchase			Inputs purchase		
VARIABLES	(1) Wage employment	(2) Self employment	(3) Fertilizer purchased	Wage employment	Self employment	Inputs purchased
Wage employment participation	.		0.437***	.		0.271*
			[0.002]			[0.053]
Self-employment participation		.	0.339***		.	0.237*
			[0.007]			[0.051]
Age of the household head	0.002	0.002	-0.013***	0.002	0.001	-0.002
	[0.722]	[0.743]	[0.008]	[0.761]	[0.775]	[0.679]
Male-headed household (0/1)	0.445***	0.031	-0.062	0.447***	0.026	-0.107
	[0.003]	[0.795]	[0.585]	[0.003]	[0.826]	[0.315]
Highest level of formal education acquired by household head						
PSLC	0.199	0.075	-0.039	0.208	0.077	0.053
	[0.274]	[0.629]	[0.793]	[0.255]	[0.623]	[0.717]
JCE	0.700***	0.208	0.189	0.698***	0.211	-0.085
	[0.002]	[0.305]	[0.333]	[0.002]	[0.297]	[0.666]
MSCE	1.172***	0.272	0.077	1.169***	0.277	-0.140
	[0.000]	[0.334]	[0.782]	[0.000]	[0.324]	[0.624]
Maximum level of formal education acquired in the household						
PSLC	0.127	0.044	0.038	0.122	0.041	-0.016
	[0.427]	[0.733]	[0.760]	[0.446]	[0.749]	[0.897]
JCE	0.064	-0.219	-0.133	0.068	-0.219	0.168
	[0.736]	[0.188]	[0.393]	[0.724]	[0.187]	[0.275]
MSCE	0.335	-0.365+	-0.074	0.343	-0.361+	0.275
	[0.175]	[0.115]	[0.737]	[0.166]	[0.119]	[0.213]
Number of infant (<5yo) in HH	0.030	0.087*	-0.083*	0.031	0.089*	-0.004
	[0.607]	[0.067]	[0.070]	[0.590]	[0.063]	[0.932]
Number of children (5-14yo) in the household	0.094**	0.001	0.012	0.094**	0.002	0.011
	[0.018]	[0.975]	[0.699]	[0.018]	[0.953]	[0.719]

Table 4-14 (cont'd)

Number of prime adults (15-60yo) in HH	-0.023	0.085**	0.084**	-0.019	0.085**	0.041
	[0.605]	[0.021]	[0.019]	[0.671]	[0.021]	[0.239]
Number of elderly (60yo+) in HH	-0.079	-0.023	-0.004	-0.082	-0.019	-0.027
	[0.612]	[0.855]	[0.969]	[0.600]	[0.880]	[0.804]
Household access to loan (0/1)	-0.030	0.346***	0.016	-0.025	0.346***	0.076
	[0.769]	[0.000]	[0.852]	[0.812]	[0.000]	[0.369]
Normalized TLU index	-2.637+	1.632	2.046	-2.536+	1.661	1.912
	[0.130]	[0.252]	[0.201]	[0.149]	[0.241]	[0.354]
Normalized agricultural asset index + Normalized land holdings			0.846***			0.614***
			[0.000]			[0.003]
Household was affected negatively by some income Shock During the last 12 months (0/1)	0.158	-0.052	-0.036	0.163	-0.052	-0.027
	[0.191]	[0.594]	[0.683]	[0.178]	[0.596]	[0.760]
Household has a migrant network (0/1)	0.066	-0.017	0.083	0.076	-0.016	-0.103
	[0.573]	[0.855]	[0.358]	[0.517]	[0.868]	[0.230]
Average 12-month total rainfall (mm)	-0.002	0.001	-0.002	-0.001	0.001	-0.002
	[0.416]	[0.444]	[0.199]	[0.469]	[0.426]	[0.207]
Annual Mean Temperature (∞C * 10)	0.007	-0.002	-0.022***	0.006	-0.002	-0.017**
	[0.478]	[0.843]	[0.007]	[0.498]	[0.815]	[0.036]
HH Distance in (KMs) to Nearest Road	-0.014	-0.006	-0.003	-0.015	-0.006	-0.006
	[0.334]	[0.643]	[0.831]	[0.303]	[0.635]	[0.626]
HH Distance in (KMs) to Nearest Population Center with +20,000	0.004	-0.000	-0.011*	0.004	-0.000	-0.008
	[0.561]	[0.965]	[0.070]	[0.547]	[0.957]	[0.208]
Price of fertilizer in the EA	-0.001*	-0.000	-0.001	-0.001*	-0.000	-0.000
	[0.085]	[0.473]	[0.251]	[0.084]	[0.460]	[0.513]
Price of maize grains in the EA	-0.002	0.002	0.000	-0.002	0.002	0.001
	[0.229]	[0.317]	[0.893]	[0.246]	[0.290]	[0.614]
Total land area owned by HH in Acres	-0.000	0.000		-0.000	0.000	
	[0.746]	[0.755]		[0.733]	[0.804]	
value of ganyu salary in the area	0.000	0.000		0.000	0.000	
	[0.538]	[0.373]		[0.501]	[0.370]	
EA neighbor's wage employment participation	0.010***	0.005***		0.010***	0.005***	
	[0.000]	[0.005]		[0.000]	[0.005]	
EA neighbor's self-employment participation	0.006***	0.013***		0.006***	0.013***	
	[0.004]	[0.000]		[0.006]	[0.000]	

Table 4-14 (cont'd)

Share of total land cultivated in crops						
grains crops			0.005			-0.011**
			[0.364]			[0.033]
legumes crops			0.003			-0.009*
			[0.638]			[0.088]
tubers crops			-0.003			-0.014**
			[0.714]			[0.039]
oil crops			0.024**			0.002
			[0.023]			[0.829]
Horticulture crops			0.004			-0.005
			[0.498]			[0.386]
cotton crops			-0.007			-0.010
			[0.343]			[0.163]
tobacco crops			0.022***			0.005
			[0.000]			[0.377]
Total value of crop sales (MKW)			0.000			-0.000
			[0.439]			[0.665]
Time dummy (year 2010=1)	-0.048	-0.050	-0.128	-0.048	-0.048	0.000
	[0.763]	[0.700]	[0.339]	[0.765]	[0.714]	[1.000]
Normalized wealth index	2.175***	2.298***		2.082***	2.207***	
	[0.000]	[0.000]		[0.000]	[0.000]	
District dummies	included	included	included	included	included	included
Time average of explanatory variables	included	included	included	included	included	included
atrho21			-0.107***			-0.111***
			[0.000]			[0.000]
atrho31			-0.157***			-0.068
			[0.004]			[0.191]
atrho32			-0.100*			-0.037
			[0.089]			[0.513]
Likelihood ratio test of joint significance of rhos: Chi2(3)			28.09***			16.13***
			[0.000]			[0.001]

Table 4-14 (cont'd)

Constant	-0.796	-2.522**	2.460**	-0.816	-2.553**	0.558
	[0.564]	[0.025]	[0.022]	[0.555]	[0.023]	[0.583]
Observations	5,018	5,018	5,018	5,018	5,018	5,018

Source: Generated by authors using LSMS data

Note: The symbols ***, **, *, and + indicate that the corresponding regression coefficient is statistically significant at the 1%, 5%, 10%, and 15% levels, respectively. P-values between brackets. The GHK estimates are generated using the mvprobit command with 100 draws in STATA. The parameters in the table are coefficients and not APEs.

Appendix 4: Description and summary statistics of the variables used in this article, by treatment status, household level, 2010-2013, rural Malawi

Table 4-15: Description and summary statistics of the variables used in this article, household level, 2010-2013, rural Malawi

Variables	Definition	Overall Mean	SD
Treatment variables			
Non-farm wage employment (0/1)	A member of household is engaged in non farm wage employment	15.11	35.82
Non farm self employment (0/1)	A member of household is engaged in non farm self employment	21.58	41.14
Outcome variables			
HHPCE	Household total real consumption expenditure per capita	57,759.59	46,702.55
Log of HHPCE		10.75	0.64
Poverty incidence	1[HHPCE<=Poverty Line]	37.20	48.34
Poverty gap	Consumption shortfall relative to poverty line as a fraction of poverty line		
Poverty severity	Squared poverty gap		
Food insecurity (0/1)	Did you worry that your household would not have enough food in the past 7 days	32.57	46.87
Food shortage (0/1)	In the last 12 months, have you been faced with a situation when you did not have enough food to eat	58.28	49.31
Food consumption adequacy	Which of the following describes best your food consumption over the past month		
1	1== less than adequate for household needs	41.45	49.27
2	2== just adequate for household needs	51.93	49.97
3	3== It was more than adequate for household needs	6.62	24.86
Income adequacy	Which of the following is true about your current income		
1	1==Allows you to build your savings	10.59	30.78
2	2==Allows you to save just a little	13.00	33.63
3	3==Only just meets your expenses	35.72	47.92
4	4==Is not sufficient so need to use savings to meet expenses	0.15	0.36
5	5==Is really not sufficient so you need to borrow to meet expenses	0.26	0.44
Seeds purchase decision (0/1)	The household purchased seeds for agricultural production in the rainy season	41.02	49.19
Fertilizer purchase decision (0/1)	The household purchased fertilizer for agricultural production in the rainy season	38.85	48.75
Inputs purchase decision (0/1)	The household purchased seeds, fertilizer or other chemicals for agricultural production in the rainy season	60.13	48.97
Seed purchase per acre (1000MKW)	Value of seed purchased per acre of land cultivated	0.53	1.69
Fertilizer purchase per acre (1000MKW)	Value of fertilizer purchased per acre of land cultivated	3.59	9.79
Inputs purchase per acre (1000MKW)	Total Value of seed, fertilizer, and chemicals purchased per acre of land cultivated	4.10	10.01
Land cultivated (acres)	Number of acres of cultivated land by household	4.43	52.31

Table 4-15 (cont'd)

Covariates			
Age of the household head	Age of the household head	44.04	16.52
Male headed-household	Male headed-household	76.16	42.62
Highest level of formal education acquired by household head			
None	The household head never attended formal school	77.17	41.98
PSLC	PSLC is the highest formal level of education of household head	10.03	30.05
JCE	JCE is the highest formal level of education of household head	6.87	25.30
MSCE	MSCE is the highest formal level of education of household head	4.68	21.13
Non-Univ Diploma and above	Non-Univ Diploma and above	1.25	11.10
Highest level of formal education acquired in the household			
None	No member of the household ever attended formal school	63.41	48.17
PSLC	PSLC is the highest formal level of education of the most educated household member	15.20	35.91
JCE	JCE is the highest formal level of education of the most educated household member	12.22	32.75
MSCE	MSCE is the highest formal level of education of the most educated household member	7.66	26.61
Non-Univ Diploma and above	Non-Univ Diploma and above	1.50	12.16
Size of the household	Number of people living in the household at the time of the interview	5.22	2.42
Number of infant (<5yo) in HH	Number of infant living in the household at the time of the interview	0.84	0.85
Number of children (5-14yo) in the household	Number of children living in the household at the time of the interview	1.62	1.41
Number of prime adults (15-60yo) in HH	Number of prime adults living in the household at the time of the interview	2.46	1.39
Number of elderly (60yo+) in HH	Number of elderly living in the household at the time of the interview	0.29	0.57
Household access to loan (0/1)	A member of the household received a loan in the year prior to the interview	16.63	37.24
Normalized wealth index	Principal component analysis estimate of asset index	0.05	0.08
Normalized TLU index	Total livestock unit index tlu=cattle*0.5+pigs*0.2+sheep*0.1+goats*0.1	0.01	0.03
Total land area owned by HH in Acres	Total land area owned by HH in Acres	3.35	49.13
Household was affected negatively by some income Shock During the last 12 months (0/1)	Household was affected negatively by some income Shock During the last 12 months (0/1)	86.37	34.31
Household has a migrant network (0/1)	A member of the household is living outside of the EA	35.65	47.90
Rain variability	Rain - EA level Coefficient of Variation of Dec-Jan rainfall from 1983/84 - 2012/13	0.25	0.04
Rainfall (mm)	Average 12-month total rainfall (mm)	850.32	89.24
Temperature (∞C * 10)	Annual Mean Temperature (∞C * 10)	216.31	19.12
HH Distance to Nearest Road	Household Distance in (KMs) to Nearest Road	9.84	9.85
HH Distance to Population Center	Household Distance in (KMs) to Nearest Population Center with +20,000	37.67	17.95
ganyu wage salary in the EA (MKW)	Value of daily *ganyu* wage salary in the EA (MKW)	566.22	599.87
Price of maize grains in the EA (MKW)	Price of maize grains in the EA (MKW)	198.59	112.66

Table 4-15 (cont'd)

Price of fertilizer in the EA (MKW)	Price of fertilizer in the EA (MKW)	62.27	39.34
EA neighbor's wage employment participation	Participation rate in non farm wage employment by other household in the same EA	15.11	14.90
EA neighbor's self employment participation	Participation rate in non farm self-employment by other household in the same EA	21.58	15.27
Share of total land cultivated in crops (crop mix)			
Grains	Share of total land cultivated in grains	64.53	26.22
Legumes	Share of total land cultivated in legumes	23.00	22.86
Tubers	Share of total land cultivated in tubers	1.55	7.75
Oils crops	Share of total land cultivated in oil crops	0.43	3.86
Horticulture crops	Share of total land cultivated in horticulture crops	4.03	11.10
Cotton	Share of total land cultivated in cotton	1.29	7.29
Tobacco	Share of total land cultivated in tobacco	4.19	11.79

Source: Generated by authors using LSMS data

Table 4-16: Test of balancing of covariates between non-farm wage employment participants and non-participants households, 2010-2013, rural Malawi

	Wage employment Non participants		Wage employment participants		t-test difference
Variables	Mean	SD	Mean	SD	pvalue
Treatment variables					
Non-farm wage employment (0/1)	0.00	0.00	100.00	0.00	
Non farm self employment (0/1)	21.53	41.11	21.89	41.37	0.668
Outcome variables					
HHPCE	53,903.41	40,631.58	79,420.64	67,908.37	0.000
Log of HHPCE	10.70	0.62	11.04	0.67	0.000
Poverty incidence	40.03	49.00	21.29	40.96	0.000
Poverty gap					
Poverty severity					
Food insecurity (0/1)	34.01	47.38	24.52	43.05	0.001
Food shortage (0/1)	60.24	48.94	47.25	49.95	0.000
Food consumption adequacy					
1	43.51	49.58	29.90	45.81	0.000
2	50.45	50.00	60.29	48.96	0.002
3	6.05	23.84	9.81	29.76	0.003
Income adequacy					
1	10.01	30.01	13.88	34.59	0.046
2	12.03	32.54	18.42	38.79	0.000
3	36.03	48.01	33.97	47.39	0.373
4	0.16	0.36	0.13	0.33	0.090
5	0.26	0.44	0.21	0.41	0.015
Seeds purchase decision (0/1)	39.52	48.89	50.07	50.03	0.000
Fertilizer purchase decision (0/1)	36.71	48.21	51.79	50.00	0.000
Inputs purchase decision (0/1)	58.35	49.30	70.92	45.45	0.000
Seed purchase per acre (1000MKW)	0.46	1.56	0.91	2.26	0.000
Fertilizer purchase per acre (1000MKW)	3.10	8.51	6.57	15.06	0.001
Inputs purchase per acre (1000MKW)	3.57	8.92	7.31	14.60	0.000
Land cultivated (acres)	4.63	55.98	3.20	18.04	0.194
Covariates					
Age of the household head	44.53	17.00	41.30	13.22	0.000
Male headed-household	74.25	43.73	86.84	33.82	0.000
Highest level of formal education acquired by household head					
None	82.54	37.97	47.01	49.94	0.000
PSLC	9.71	29.61	11.84	32.33	0.012
JCE	5.64	23.08	13.76	34.46	0.000
MSCE	1.96	13.86	19.98	40.01	0.000
Non-Univ Diploma and above	0.15	3.86	7.42	26.22	0.000

Table 4-16 (cont'd)

Highest level of formal education acquired in the household					
None	68.70	46.38	33.73	47.31	0.000
PSLC	15.63	36.32	12.80	33.43	0.926
JCE	11.07	31.38	18.66	38.98	0.000
MSCE	4.34	20.39	26.32	44.06	0.000
Non-Univ Diploma and above	0.26	5.05	8.49	27.89	0.000
Size of the household	5.17	2.40	5.47	2.55	0.701
Number of infant (<5yo) in HH	0.85	0.85	0.82	0.82	0.027
Number of children (5-14yo) in the household	1.62	1.41	1.66	1.41	0.826
Number of prime adults (15-60yo) in HH	2.39	1.36	2.84	1.50	0.000
Number of elderly (60yo+) in HH	0.32	0.59	0.16	0.45	0.000
Household access to loan (0/1)	15.48	36.18	23.09	42.16	0.010
Normalized wealth index	0.04	0.06	0.11	0.13	0.000
Normalized TLU index	0.01	0.03	0.01	0.02	0.901
Total land area owned by HH in Acres	3.55	52.84	2.26	16.94	0.202
Household was affected negatively by some income Shock During the last 12 months (0/1)	86.71	33.95	84.45	36.26	0.104
Household has a migrant network (0/1)	36.52	48.15	30.74	46.17	0.028
Rain - EA level CoV of Dec-Jan rainfall from 1983/84 - 2012/13	0.25	0.04	0.26	0.04	0.017
Average 12-month total rainfall (mm)	850.57	87.09	848.91	100.55	0.017
Annual Mean Temperature (∞C * 10)	215.97	18.97	218.22	19.84	0.587
HH Distance in (KMs) to Nearest Road	10.43	10.05	6.53	7.89	0.000
HH Distance in (KMs) to Nearest Population Center with +20,000	38.11	17.87	35.17	18.26	0.012
Value of daily *ganyu* wage salary in the EA (MKW)	561.67	599.41	592.09	602.18	0.820
Price of maize grains in the EA (MKW)	201.75	112.90	180.81	109.70	0.001
Price of fertilizer in the EA (MKW)	62.12	38.98	63.14	41.31	0.731
EA neighbor's wage employment participation	13.26	13.52	25.50	17.79	0.000
EA neighbor's self employment participation	20.74	14.92	26.32	16.28	0.000

Table 4-16 (cont'd)

Share of total land cultivated in crops (crop mix)					
Grains	64.31	26.17	65.90	26.52	0.370
Legumes	22.81	22.73	24.16	23.63	0.173
Tubers	1.55	7.65	1.55	8.33	0.035
Oils crops	0.48	4.09	0.13	1.90	0.107
Horticulture crops	4.14	11.28	3.38	9.93	0.271
Cotton	1.29	7.29	1.34	7.31	0.900
Tobacco	4.48	12.20	2.41	8.64	0.000

Source: Generated by authors using LSMS data

Table 4-17: Test of balancing of covariates between non-farm self-employment participants and non-participants, 2010-2013, rural Malawi

	Self-employment Non participants		Self-employment participants		t-test difference
Variables	Mean(SE0)	SD	Mean(SE1)	SD	pvalue_NFSE
Treatment variables					
Non-farm wage employment (0/1)	15.05	35.76	15.33	36.04	0.668
Non farm self employment (0/1)	0.00	0.00	100.00	0.00	
Outcome variables					
HHPCE	54,741.65	43,462.77	68,724.30	55,600.92	0.000
Log of HHPCE	10.70	0.63	10.91	0.66	0.000
Poverty incidence	39.81	48.96	27.72	44.78	0.000
Poverty gap					
Poverty severity					
Food insecurity (0/1)	33.24	47.11	30.15	45.91	0.213
Food shortage (0/1)	58.85	49.22	56.20	49.64	0.419
Food consumption adequacy					
1	42.81	49.49	36.52	48.17	0.063
2	50.90	50.00	55.70	49.70	0.369
3	6.29	24.29	7.79	26.81	0.094
Income adequacy					
1	9.98	29.98	12.81	33.44	0.108
2	11.99	32.48	16.67	37.28	0.002
3	36.88	48.25	31.49	46.47	0.003
4	0.14	0.35	0.18	0.38	0.022
5	0.27	0.44	0.21	0.41	0.013
Seeds purchase decision (0/1)	39.27	48.84	47.51	49.96	0.000
Fertilizer purchase decision (0/1)	36.73	48.21	46.71	49.91	0.000
Inputs purchase decision (0/1)	57.48	49.44	69.98	45.85	0.000
Seed purchase per acre (1000MKW)	0.49	1.66	0.65	1.78	0.011
Fertilizer purchase per acre (1000MKW)	3.48	9.62	4.00	10.38	0.215
Inputs purchase per acre (1000MKW)	3.96	9.96	4.60	10.16	0.078
Land cultivated (acres)	4.09	51.38	5.67	55.60	0.903
Covariates					
Age of the household head	44.72	17.01	41.58	14.37	0.000
Male headed-household	74.92	43.35	80.65	39.52	0.014
Highest level of formal education acquired by household head					
None	78.33	41.20	72.95	44.44	0.010
PSLC	9.70	29.61	11.22	31.58	0.194
JCE	6.27	24.25	9.05	28.69	0.106
MSCE	4.52	20.77	5.28	22.37	0.354
Non-Univ Diploma and above	1.18	10.78	1.51	12.19	0.126

Table 4-17 (cont'd)

Highest level of formal education acquired in the household					
None	65.42	47.57	56.11	49.65	0.000
PSLC	14.29	35.00	18.51	38.85	0.000
JCE	11.60	32.02	14.49	35.21	0.246
MSCE	7.26	25.95	9.13	28.81	0.248
Non-Univ Diploma and above	1.43	11.87	1.76	13.15	0.211
Size of the household	5.12	2.39	5.56	2.50	0.000
Number of infant (<5yo) in HH	0.82	0.83	0.94	0.89	0.004
Number of children (5-14yo) in the household	1.59	1.41	1.75	1.41	0.001
Number of prime adults (15-60yo) in HH	2.40	1.39	2.67	1.38	0.000
Number of elderly (60yo+) in HH	0.32	0.59	0.19	0.48	0.000
Household access to loan (0/1)	13.53	34.21	27.89	44.86	0.000
Normalized wealth index	0.05	0.07	0.08	0.11	0.000
Normalized TLU index	0.01	0.02	0.01	0.06	0.097
Total land area owned by HH in Acres	3.19	49.77	3.97	46.75	0.598
Household was affected negatively by some income Shock During the last 12 months (0/1)	85.75	34.96	88.61	31.78	0.027
Household has a migrant network (0/1)	36.75	48.22	31.66	46.53	0.070
Rain - EA level CoV of Dec-Jan rainfall from 1983/84 - 2012/13	0.25	0.04	0.26	0.04	0.002
Average 12-month total rainfall (mm)	853.06	89.51	840.36	87.59	0.012
Annual Mean Temperature (∞C * 10)	215.84	19.15	218.01	18.92	0.026
HH Distance in (KMs) to Nearest Road	10.11	9.92	8.87	9.54	0.013
HH Distance in (KMs) to Nearest Population Center with +20,000	38.05	18.05	36.28	17.55	0.343
Value of daily *ganyu* wage salary in the EA (MKW)	548.59	564.61	630.35	710.21	0.005
Price of maize grains in the EA (MKW)	196.70	111.45	205.43	116.73	0.014
Price of fertilizer in the EA (MKW)	60.87	39.22	67.36	39.37	0.000
EA neighbor's wage employment participation	14.20	14.52	18.43	15.76	0.000
EA neighbor's self employment participation	19.74	14.32	28.28	16.68	0.000

Table 4-17 (cont'd)

Share of total land cultivated in crops (crop mix)					
Grains	65.45	26.17	61.15	26.15	0.001
Legumes	22.18	22.64	26.00	23.42	0.007
Tubers	1.51	7.71	1.68	7.88	0.758
Oils crops	0.47	4.01	0.28	3.26	0.589
Horticulture crops	3.92	10.98	4.42	11.53	0.183
Cotton	1.22	7.23	1.57	7.52	0.149
Tobacco	4.36	12.04	3.57	10.79	0.078

Source: Generated by authors using LSMS data

REFERENCES

REFERENCES

ABDULAI, A. & CROLEREES, A. 2001. Determinants of income diversification amongst rural households in Southern Mali. *Food policy,* 26**,** 437-452.

ACKAH, C. 2013. Nonfarm employment and incomes in rural Ghana. *Journal of International Development,* 25**,** 325-339.

ANGRIST, J. D. & PISCHKE, J.-S. 2008. *Mostly harmless econometrics: An empiricist's companion*, Princeton university press.

BARDHAN, P. & UDRY, C. 1999. *Development microeconomics*, Oxford University Press.

BARRETT, C. B., BEZUNEH, M. & ABOUD, A. 2001a. Income diversification, poverty traps and policy shocks in Côte d'Ivoire and Kenya. *Food Policy,* 26**,** 367-384.

BARRETT, C. B., REARDON, T. & WEBB, P. 2001b. Nonfarm income diversification and household livelihood strategies in rural Africa: concepts, dynamics, and policy implications. *Food policy,* 26**,** 315-331.

BASER, O. 2009. Too much ado about instrumental variable approach: is the cure worse than the disease? *Value in Health,* 12**,** 1201-1209.

BRAMOULLE, Y., DJEBBARI, H. & FORTIN, B. 2009. Identification of Peer Effects through Social Networks. *Journal of Econometrics,* 150**,** 41-55.

CAMERON, A. C. & TRIVEDI, P. K. 2005. *Microeconometrics: methods and applications*, Cambridge university press.

CAMERON, A. C. & TRIVEDI, P. K. 2010. *Microeconometrics using stata*, Stata Press College Station, TX.

CAPPELLARI, L. & JENKINS, S. P. 2003. Multivariate probit regression using simulated maximum likelihood. *The Stata Journal,* 3**,** 278-294.

CHAMBERLAIN, G. 1982. Panel data. National Bureau of Economic Research Cambridge, Mass., USA.

DAVIS, B., DI GIUSEPPE, S. & ZEZZA, A. 2014. Income diversification patterns in rural sub-Saharan Africa: reassessing the evidence. *World Bank Policy Research Working Paper*.

DE GEUS, J. G. & DE GEUS, J. G. 1967. Fertilizer guide for tropical and subtropical farming. Cambridge Univ Press.

DE JANVRY, A., FAFCHAMPS, M. & SADOULET, E. 1991. Peasant household behaviour with missing markets: some paradoxes explained. *The Economic Journal,* 101**,** 1400-1417.

DE JANVRY, A. & SADOULET, E. 2001. Income strategies among rural households in Mexico: The role of off-farm activities. *World development,* 29**,** 467-480.

DEATON, A. & ZAIDI, S. 2002. *Guidelines for constructing consumption aggregates for welfare analysis*, World Bank Publications.

FILMER, D. & PRITCHETT, L. H. 2001. Estimating wealth effects without expenditure data—or tears: An application to educational enrollments in states of india*. *Demography,* 38**,** 115-132.

GALLANI, S., KRISHNAN, R. & WOOLDRIDGE, J. 2015. Applications of Fractional Response Model to the Study of Bounded Dependent Variables in Accounting Research. *Harvard Business School Accounting & Management Unit Working Paper*.

GREEN WILLIAM, H. 2000. Econometric analysis. *Forth Edition, Prentice Hall International, Inc, New York University*.

HAGGBLADE, S., HAZELL, P. & REARDON, T. 2010. The rural non-farm economy: Prospects for growth and poverty reduction. *World Development,* 38**,** 1429-1441.

IMBENS, G. M. & WOOLDRIDGE, J. M. 2008. Recent developments in the econometrics of program evaluation. National Bureau of Economic Research.

KIJIMA, Y., MATSUMOTO, T. & YAMANO, T. 2006. Nonfarm employment, agricultural shocks, and poverty dynamics: evidence from rural Uganda. *Agricultural Economics,* 35**,** 459-467.

KOENKER, R. 2005. *Quantile regression*, Cambridge university press.

KRISHNAN, P. & PATNAM, M. 2014. Neighbors and Extension Agents in Ethiopia: Who Matters More for Technology Adoption? *American Journal of Agricultural Economics,* 96**,** 308-327.

LANJOUW, J. O. & LANJOUW, P. 2001. The rural non‐farm sector: issues and evidence from developing countries. *Agricultural economics,* 26**,** 1-23.

LANJOUW, P., QUIZON, J. & SPARROW, R. 2001. Non-agricultural earnings in peri-urban areas of Tanzania: evidence from household survey data. *Food policy,* 26**,** 385-403.

LANJOUW, P. & SHARIFF, A. 2004. Rural non-farm employment in India: Access, incomes and poverty impact. *Economic and Political Weekly***,** 4429-4446.

MANSKI, C. F. 1993. Identification of endogenous social effects: The reflection problem. *The review of economic studies,* 60**,** 531-542.

MATSUMOTO, T., KIJIMA, Y. & YAMANO, T. 2006. The role of local nonfarm activities and migration in reducing poverty: evidence from Ethiopia, Kenya, and Uganda. *Agricultural Economics,* 35**,** 449-458.

NAGLER, P. & NAUDE, W. 2014. Non-farm enterprises in rural Africa: new empirical evidence. *World Bank Policy Research Working Paper*.

OSENI, G. & WINTERS, P. 2009. Rural nonfarm activities and agricultural crop production in Nigeria. *Agricultural Economics,* 40**,** 189-201.

OWUSU, V., ABDULAI, A. & ABDUL-RAHMAN, S. 2011. Non-farm work and food security among farm households in Northern Ghana. *Food policy,* 36**,** 108-118.

POULTON, C., KYDD, J. & DORWARD, A. 2006. Overcoming Market Constraints on Pro‐Poor Agricultural Growth in Sub‐Saharan Africa. *Development Policy Review,* 24**,** 243-277.

REARDON, T., STAMOULIS, K., BALISACAN, A., CRUZ, M., BERDEGUÉ, J. & BANKS, B. 1998. Rural non-farm income in developing countries. *The state of food and agriculture,* 1998**,** 283-356.

REARDON, T., STAMOULIS, K. & PINGALI, P. 2007. Rural nonfarm employment in developing countries in an era of globalization. *Agricultural Economics,* 37**,** 173-183.

REARDON, T., TAYLOR, J. E., STAMOULIS, K., LANJOUW, P. & BALISACAN, A. 2000. Effects of non‐farm employment on rural income inequality in developing countries: an investment perspective. *Journal of agricultural economics,* 51**,** 266-288.

RUBIN, D. B. & JOHN, L. 2011. Rubin Causal Model. *In:* LOVRIC, M. (ed.) *International Encyclopedia of Statistical Science.* Berlin, Heidelberg: Springer Berlin Heidelberg.

SMALE, M., KUSUNOSE, Y., MATHENGE, M. K. & ALIA, D. 2016. Destination or Distraction? Querying the Linkage Between Off-Farm Work and Food Crop Investments in Kenya. *Journal of African Economies*.

SMITH, D. R., GORDON, A., MEADOWS, K. & ZWICK, K. 2001. Livelihood diversification in Uganda: patterns and determinants of change across two rural districts. *Food Policy,* 26**,** 421-435.

START, D. 2001. The rise and fall of the rural non‐farm Economy: Poverty Impacts and Policy options. *Development policy review,* 19**,** 491-505.

WINTERS, P., DAVIS, B., CARLETTO, G., COVARRUBIAS, K., STAMOULIS, K., QUINONES, E., ZEZZA, A. & DI CARACALLA, V. D. T. 2007. Assets, activities and rural poverty alleviation: Evidence from a multicountry analysis. *FAO, Rome*.

WOLDENHANNA, T. & OSKAM, A. 2001. Income diversification and entry barriers: evidence from the Tigray region of northern Ethiopia. *Food Policy,* 26**,** 351-365.

WOOLDRIDGE, J. M. 2010. *Econometric analysis of cross section and panel data*, MIT press.