Doctoral Thesis

博士論文

Financial Behavior and Constraints of Individuals and Households in Developing Countries: Evidence from Ethiopia and Nigeria

（発展途上国における個人と家計の金融行動と制約：

エチオピアとナイジェリアの事例から）

劉 恩珍

Doctoral Thesis

博士論文

Financial Behavior and Constraints of Individuals and Households in Developing Countries: Evidence from Ethiopia and Nigeria

（発展途上国における個人と家計の金融行動と制約：
エチオピアとナイジェリアの事例から）

劉 恩珍

Graduate School of Frontier Sciences

The University of Tokyo

Abstract

Financial sector development has been considered as one of the promising ways for both poverty reduction and sustainable economic development in developing countries. By using appropriate financial services, such as savings, credit and insurance, individuals can plan for their long-term goals, start businesses, and manage unexpected risk. Expanding access to formal financial institutions to all ("financial inclusion") has gained increasing importance as a policy objective among policymakers for decades. Despite the progress made towards financial inclusion, many people still lack access to formal financial services, and even people with access to formal financial services in developing countries, their usage and take-up rate of financial services are very low. To include them into formal financial system, the demand side of financial services also needs to be carefully considered. Thus, this study aims to explore 1) the determinants of the choice of savings methods to understand which savings devices and financial service qualities the poor demand, and 2) the determinants and predictors of different types of credit constraints based on loan applications.

This study first explores the determinants of the saving method choice among bank accounts, Rotating Savings and Credit Associations (ROSCAs) and cash savings at home using primary data collected from production workers in on cut flower farms in Ethiopia. Unique contextual features of this sample are that workers have full access to formal banks as well as informal savings groups and they are familiar with using bank accounts, which allows us to focus on the demand side of workers' saving methods. By using the multivariate probit model and seemingly unrelated regression model, this study finds that the usage and the amount of bank savings increases in asset ownership. This study further finds that the amount of ROSCAs savings among savers increases in worker with risk-aversion, suggesting that the workers in my sample value the social insurance aspect of ROSCAs. This study also finds

that the usage and amount of ROSCAs savings increases for more impatient workers. Overall, informal saving groups work as insurance and commitment tools, complementing formal financial institutions.

Using the same case of cut flower production workers in Ethiopia, this study also explores domestic migrants' saving behavior since more than half of workers are rural-to-urban migrants in this sample. Domestic migration within developing countries has received less attention compared to international migration. Since empirical evidence of the internal migrants' employment situation in the labor market in sub-Saharan Africa is scarce, this study also aims to analyze differences in wages and productivity between migrants and non-migrants. An extended Mincer-type human capital wage regression is used to estimate the wage equation and Blinder-Oaxaca decomposition approach is used to estimate the rate of return differences across migrants and non-migrants. This study finds that even though migrant workers show higher work performance than non-migrants, they are likely to be paid less. With field observation and descriptive analysis on financial behavior of migrants, this study discusses their motivations of internal migration even though they may face differential treatment in wages.

Furthermore, to better understand the demand side of microcredit program, this study investigates the determinants and predictors of different types of credit constraints using nationally representative household data from Nigeria. Direct elicitation approach is used to measure detailed non-pricing credit constraints (quantity, risk, and transaction-cost rationing) households may face. This study finds that households in the south region of Nigeria are more likely to face quantity rationing. This study also finds that the probability of being risk rationed decreases if household heads engage in waged labor, while the probability of being transaction-cost rationed decreases when households owns their non-farm enterprises. For lower income households, the probability of being risk rationed increases if households engage in agriculture. By using supervised machine learning approach, this study also finds that the probability of being credit constrained can be predicted by not only households' credit market participation and but also community level data such as climatology, terrain, and crop season parameters; precipitation, elevation and average timing of onset of greenness decrease in day of year. These selected predictors could be used for other datasets which does not contain information on households' credit constraints status for better targeting.

Overall, this study investigates the determinants of savings and credit constraints for workers and households in developing countries, focusing on the demand side of financial services. In addressing these issues, the uniqueness of the study relative to existing studies lies on the use of rich data on the financial behavior of workers in a particular industry, which allows me to focus on the demand-side saving behavior, as well as the national-representative data which allows me to provide comprehensive analyses on credit constraints. The findings from this study are expected to help policymakers and practitioners consider financial services that are tailored to the financial needs and preferences of the poor to include the poor in formal financial systems.

Acknowledgement

First and foremost, I would like to express my special appreciation and thanks to my supervisor Prof. Aya Suzuki for her patience, encouragement, and all the support she gave me. She has been a wonderful role model and mentor for me, and her words gave me the confidence and motivation to become a better researcher.

I would also like to thank my committee members, Prof. Kazushi Takahashi, Prof. Hiroyuki Nakata, Prof. Masahide Horita and Prof. Maiko Sakamoto for their invaluable comments and suggestions.

I am also very grateful to Prof. Yukichi Mano and Dr. Girum Abebe for sharing their insights during the fieldwork in Ethiopia. Their passion in research and energetic characteristics was such a great inspiration to me.

A very special thanks to Prof. Geunwoo Lee for his advice and feedback on my research and for helping me in numerous ways during various stages of my PhD. I am also grateful to Prof. Yuri Kim for her support and encouragement.

I would also like to thank my dear friends at the department, Dr. Sijia Zhao, Kikuchi san, Dr. Emmanuel Apiors, Nobuyuki, and Lingfeng for their immeasurable support from my first year at The University of Tokyo and their friendship. I have learned a lot from them. I wish to give special thanks to Aly, Nisa and Suhyoon for many unforgettable memories and their friendship. I am also grateful to Suzuki lab members, especially Ikuya for his support during writing my thesis and Mayuko, Kotoko for their support and friendship.

I am indebted to all my friends and family in Japan who opened their homes to me during my time at The University of Tokyo and who were always so helpful in numerous ways. Special thanks to Ino sensei, Yayoi san, Seehuang, Elena, Nagisa, Koto san, Hosana, Frederick, Emi san, Rieko san, Mai, Yuki, IBF church, and Mikachu family members.

I gratefully acknowledge the funding received towards my PhD from The University of Tokyo, Japan Student Services Organization, The Korean Scholarship Foundation, Otuska Toshimi Scholarship Foundation.

I wish to express my deepest gratitude to Soyeon Kim, Mina Park, Immanuel Ghana team members and Pastor Youngwhan Yoon, who always remember me in their prayers and have supported me spiritually and mentally.

I would also like to say a heartfelt thanks to my dear parents, Hyeongnam Ryu and Youngjoo Lee for their love, supports, and always believing in me to follow my dreams. My appreciation also goes to my lovely siblings Sujin Ryu and Giwoong Ryu.

Finally, I thank my God, my good Father, for giving me unfailing love, countless blessings, and strength. Your faithfulness and love are the foundation of my life. I have experienced Your guidance day by day during my PhD life. Thank you, Lord.

Table of Contents

List of Tables

List of Figures

1. Introduction

1.1. Background

Financial inclusion has gained increasing importance as one of the promising ways for both poverty reduction and opportunities for economic growth. The World Bank launched the Universal Financial Access 2020 initiative, which aims to ensure that all people will have access to a transaction account by 2020, as a first step toward broader financial inclusion wherein everyone has access to appropriate and safe financial services (CGAP, 2009). By using appropriate financial services, such as savings, credit and insurance, individuals can plan for their long-term goals, start businesses, and manage unexpected risk. Financial inclusion does not only aim to expand access to formal finance, but also to encourage use of financial services.

There have been many efforts to improve access to appropriate and safe financial services, many people in developing countries still lack access to formal financial services, and even people with access to formal financial services, their usage of formal financial institutions is low. For example, according to the World Bank Global Financial Inclusion database, 35 percent of adults owned a transaction account in formal banks and only 26 percent of adults saved at formal banks in Ethiopia in 2017. Dupas et al. (2018) argue that simply expending access to basic bank accounts will appear unattractive to the poor. For including the poor into the formal financial system, financial services should be more tailored to the poor's specific needs to attract them. So, the demand side of financial services needs to be considered.

Also, the take-up rate of microcredit program is normally 20-25 percent, which is very low (Banerjee, 2013). One of the reasons for unsatisfactory take-up rate of microcredit program could be mis-targeting. Microcredit program might be more effective when credit-constrained households are well-targeted for microcredit programs and demand of borrowers should be reflected in designing microcredit program. Thus, to solve low take-up problem of microcredit program, strengthening empirical methods for identifying credit-constrained households would be needed for better targeting those who credit-constrained. Identifying and distinguishing different types of credit constraints is important since each of credit constraints requires different policy implications (Boucher et al., 2009).

1.2. Objectives and Research Questions

To include the poor into formal financial system, the demand side of financial services needs to be carefully considered. First, this study aims to explore the determinants of the choice of savings methods to understand which savings devices and financial service qualities the poor demand. I consider the case of cut flower production workers in Ethiopia since this sample has a unique contextual features in that workers have full access to banks as well as informal saving groups and they are familiar with using bank accounts, which allow me to focus on the demand side of workers' saving methods.

The data collected in January 2018 from 709 unskilled and low-income production workers in cut flower farms in Ethiopia. Ethiopian commercial banks have recently been opening branches in many cities and holding seminars to help workers on cut flower farms open bank accounts because cut flower farm workers earn regular incomes. Cut flower workers live in peri-urban areas near cut flower farms where they can access to banks, and more than half of the workers in my sample are paid via bank account. Using this data, this study investigates the determinants of the savings choice made among formal bank accounts, informal saving groups and home savings which are three most commonly used saving methods among cut flower workers. Using the same case of cut flower production workers in Ethiopia, this study also explores internal migrants' saving behavior since more than half of workers are rural-to-urban migrants in this sample.

Furthermore, this study also aims to investigate the determinants of detailed non-pricing credit constraints based on direct elicitation approach of identifying credit constraints. This study also aims to explore the possible strong predictors of the households' credit constraints status for targeting by supervised machine learning approach because a general household survey does not often contain this directly elicited credit constraint variable.

1.3 Outline of the Study

The reminder of this study proceeds as follows. In Chapter 2, I report on the determinants of the choice between formal bank accounts and informal saving methods. Saving behaviors and working environment of internal migrants are described in Chapter 3. Chapter 4 investigates the determinants and predictors of detailed non-pricing credit constraints. Finally, Chapter 5 concludes the study.

2. The Determinants of the Saving Method Choice: The Case of Unskilled Workers in Ethiopia's Cut Flower Industry

2.1. Introduction

Expanding access to formal financial institutions to all (the process of "financial inclusion") has gained increasing importance as a policy objective among policymakers. Providing poor households access to basic bank accounts has attracted much attention from government agencies and multilateral institutions working in international development as a first step toward financial inclusion.[1] Recent studies have shown that access to formal savings accounts encourages the poor to save more, invest more in their businesses, and even consume more, leading to welfare improvement (Ashraf et al., 2006; Dupas and Robinson, 2013; Prina, 2015). Despite the progress made toward financial inclusion, many people still lack access to formal banks. For example, according to the World Bank Global Financial Inclusion database, 35 percent of adults (41% for male and 29% for female, respectively) owned a transaction account in formal banks and only 26 percent of adults saved at formal banks in Ethiopia in 2017. Many people in developing countries still rely on informal methods of savings (e.g., joining informal savings groups, investing in assets such as livestock and jewelry, or storing money at home). These informal savings instruments co-exist with formal financial institutions in developing countries, even in urban areas where people can easily access to formal financial institutions.

Dupas et al. (2018) argue that simply expanding access to basic bank accounts using policies such as offering subsidies to those who open accounts is ineffective in helping the poor in developing countries and is unlikely to improve welfare significantly based on

[1] The World Bank launched the Universal Financial Access 2020 initiative, which aims to ensure that all people will have access to a transaction account by 2020, as a first step toward broader financial inclusion wherein everyone has access to appropriate and safe financial services (CGAP, 2009).

empirical evidence drawn from Uganda, Malawi, and Chile.[2] Expanding access to savings services that are tailored to the financial needs and preferences of the poor has proved a more effective way of including the poor in formal financial systems (Dupas and Robinson, 2013; Dupas et al., 2018). To understand which savings devices and financial service qualities the poor demand, more studies need to focus on the saving behaviors of the poor in the context of the coexistence of formal and informal financial institutions.

The question of what factors determine the choice of savings instruments is worth exploring because the answers will help policymakers and practitioners consider what action is needed to expand access to formal finance among the poor and to encourage use of financial services. I explore the determinants of the choice between formal bank accounts and informal saving methods (informal savings groups and cash savings at home) using primary data collected in January 2018 from 709 unskilled and low-income production workers laboring on cut flower farms in Ethiopia. Ethiopian commercial banks have recently been opening branches in many cities and holding seminars to help workers on cut flower farms open bank accounts because cut flower farm workers earn regular incomes (albeit relatively low incomes). Cut flower workers live in peri-urban areas near cut flower farms, and more than half of the workers (around 54 percent) in my sample are paid via bank accounts; thus, they have access to banks.[3] However, despite greater availability of formal finance, the workers tend to save more by using informal saving groups known as “ROSCAs” (Rotating Savings and Credit Associations, known as *Equb* in Ethiopia).

The ROSCA is the most common informal financial institution in developing countries. Individuals who agree to meet regularly and save money together contribute a fixed amount

[2] Measuring the impact of savings is difficult since outcomes tend to be estimated noisily and are likely to be diffuse because purposes of savings are heterogeneous and informal savings channels are many and varied. (Dupas et al., 2018; Karlan and Morduch, 2010).

[3] Regarding the supply of formal financial services, distance to the bank is often regarded as a major barrier to financial inclusion in the rural areas of developing countries.

to a pot, and each individual receives the pot in turn. People use ROSCAs for several reasons, including to purchase indivisible durable goods (Besley, et al., 1993; Anderson et al., 2009; the “early pot motive”), to cope with risk (Calomiris and Rajaraman, 1998; Klonner, 2003), to use it as a commitment device (Ashraf et al., 2006; Gugerty, 2007), or to protect savings by social pressure (Anderson and Baland, 2002; Ambec and Treich, 2007). ROSCAs can bring benefits for their participants, yet they also have several limitations such as insecurity (potential breaches of trust), inflexibility in terms of liquidation, and limited growth potential. Keeping cash at home is also one of the commonly used informal saving methods. Some might be motivated to do so due to potential bank closings and malfunctioning ATM machines. However, cash stored at home is vulnerable to theft and disaster like flooding or fire and it is a lack of investment returns.

While many studies examine ROSCAs and the effect of offering bank accounts individually, few studies have investigated the savings choice between formal financial institutions and informal savings groups. Moreover, there is limited literature on how bank savings and ROSCA participation interact with each other. Kedir et al. (2011) examine the relationship between households’ savings decisions and wealth using household panel data drawn from urban Ethiopia, finding that savings in formal and informal financial institutions coexist. They find that households tend to use only ROSCAs when their wealth level is low and that they use bank accounts and ROSCAs simultaneously once their wealth crosses a certain threshold. Carpenter and Jensen (2002) find significant differences between bank account and ROSCA use depending on income, education, and literacy levels using Pakistan household data.

This chapter examines saving decisions among workers in the same industry who have full access to banks as well as informal savings groups with more focus on the amount of actual savings in different saving methods, while previous studies simply use binary outcome data of individual's financial institutions participation and samples of heterogeneous occupations in the community context (where income levels also vary). It also uses unique data that include detailed information on financial activities (the savings amount and participation data on formal and informal saving methods), risk preferences, time preferences, financial literacy, social networks, and workers' socio-economic characteristics, which allow us to focus on the demand side of workers' saving methods.

Using a multivariate probit model and seemingly unrelated regression model, I find that ROSCAs work as insurance and commitment tools based on social networks on the farms, which complement formal financial institutions. The use of bank accounts increases when the workers own land assets which can be regarded as buffer stocks in times of need. On the other hand, among savers, more risk-averse workers tend to use ROSCAs, indicating that ROSCA savings might be a tool for risk coping strategy for workers. In addition, impatient workers are more likely to save in ROSCAs, which indicates that the poor use ROSCAs as commitment tools for savings.

The rest of this chapter proceeds as follows. Section 2.2 presents my hypotheses, which are based on the literature and field observations. Section 2.3 describes the data and presents contextual features and model specification. Section 2.4 outlines the estimation results. Finally, Section 2.5 concludes this chapter by describing its policy implications.

2.2. Hypotheses

I explore the determinants of the savings choice made among formal bank accounts, ROSCAs and home savings which are three most commonly used saving methods among cut flower workers by considering six hypotheses established based on the literature and my field observations. The first hypothesis concerns the relationship between households' asset variables and workers' saving behaviors. When households that rely very little on insurance or social protection face unexpected income shocks, selling assets is one way of providing self-insurance (Deaton, 1992; Udry, 1994; Diagne, 1999; Heltberg, 2013). I can thus expect that workers with higher asset holdings may tend to save more by using bank accounts than by using ROSCA because they have less of a need to rely on the social insurance feature of ROSCAs. They may prefer safer and easier saving methods (i.e., bank savings) in order to be able to liquidize their savings at any time. We thus proposed the following:

Hypothesis 1: Self-insured workers with higher asset ownership tend to save more in bank accounts.

Second, I consider how workers' risk preferences relate to their saving behaviors. While the relation between risk preference and saving propensity has been examined, the differential effects of risk preferences on saving decisions have not been examined rigorously. Theoretically, risk preference may affect saving method choices in two directions: Risk-averse workers may prefer to use bank savings if they consider ROSCAs or home savings risky, or they may prefer to use ROSCAs if they value the social insurance feature of ROSCAs.

Risk-averse workers may consider ROSCAs risky because of the potential risk of loss and risky investment feature of ROSCAs. The first type of risk is that of losing savings. Each ROSCA member contributes the same amount at each meeting and receives a lump sum amount when it is their turn. This continues until all members have received a lump sum once in one cycle. Because of this rotational structure of ROSCAs, members who have received the pot earlier may have an incentive to default.[4] To avoid this risk, workers may prefer to save via bank accounts which carry less of a loss risk. However, it is also true that ROSCAs can mitigate this risk through social connectedness and reputational effects. The ROSCA participants who default are sanctioned socially and are excluded from further ROSCA participation (Anderson et al., 2009; Besley et al., 1993). Thus, the risk of loss may not be an important factor in saving choice, except cash savings at home.

The second type of risk is investment risk. Kedir et al. (2011) suggest that individuals join ROSCAs from a "risky investment motive" because, under binding borrowing constraints, participants who take the pot at the beginning of a cycle can use the lump sum to finance high yield capital goods. However, participants who receive late payouts lose the interest income they would have gained if they had saved in bank accounts because ROSCAs provide zero interest. Kedir et al. (2011) argue that ROSCAs represent risky investments because their payment timing is random, and their theoretical model assumes that a household has to allocate its savings between high-risk high-return ROSCAs and low-risk low-return bank accounts.

On the other hand, workers may participate in ROSCAs from an insurance motive against idiosyncratic risks, and ROSCAs may work as informal risk-sharing mechanisms for them in the absence of a formal insurance system. If this holds true, risk-averse workers may prefer to participate in ROSCAs. Klonner (2003) offers a theoretical model of how bidding ROSCAs serves as a risk-sharing device, and Calomiris and Rajaraman (1998) show that bidding ROSCAs in India have insurance components. Even in random ROSCAs case, participants

[4] Wright and Mutesasira (2001) survey 1,500 individuals in Uganda and find that 27 percent of those who had used ROSCAs had lost their money, although the loss represented only 6 percent of the total savings in ROSCAs in the period.

may ask other members to switch the pot receipt order or may borrow money from them in an emergency. Even though workers may not be able to cope with covariate risks effectively in economically and socially homogenous groups, like cut flower farm workers (Calomiris and Rajaraman, 1998), participating in ROSCAs may enable access to funds in the face of unexpected idiosyncratic risks such as illness. Thus, the way risk preferences relate to saving method choice is likely to differ according to the ROSCA's devices and functions that workers consider important. If risk-averse workers are less likely to use ROSCAs, it may indicate that workers associate the risk of losing their savings or of losing interest income with ROSCAs. On the other hands, if risk-averse workers tend to save in ROSCAs, it may indicate that workers value its insurance aspects. We thus propose the following:

Hypothesis 2: Risk preferences affect differentially the choice of saving method.

Third, workers who are impatient and have more present-biased time preferences may save more via ROSCAs since they may want commitment mechanisms in order to overcome self-control problems. Workers who put a high value on the present (compared to the future) and exhibit time-varying discount rates are more likely to consume their money impatiently, and thus may have a preference for commitment-saving devices in order to save more effectively (Ashraf et al., 2006; Gugerty, 2007). Workers who join ROSCAs commit to contributing a fixed amount of savings and to a specific deposit date in advance; moreover, they cannot withdraw their savings freely. This inflexibility works as a commitment mechanism, and the negative economic or social punishments associated with failure to meet their savings goals is likely to encourage workers to save regularly (Ambec and Treich, 2007; Anderson et al., 2009). However, workers who prefer flexibility of being able to withdraw and deposit their savings freely may choose to use bank accounts or cash savings at home. We thus propose the following:

Hypothesis 3: ROSCAs savings increase with impatience and present-biased time preferences because workers need for the commitment properties offered by ROSCAs.

Fourth, workers with high incomes may tend to save more in bank accounts, and savings in ROSCAs may first increase as households' incomes increase and then decrease at a specific threshold (in an inverted U-shaped relationship). Some workers with higher wealth levels may tend to reduce savings in ROSCAs or withdraw from ROSCA participation since they can purchase indivisible durable goods directly. Individuals with high incomes may choose to buy durable goods directly rather than join ROSCAs. Kedir et al. (2011) and Kedir and Ibrahim (2011) show this non-linear quadratic relationship between household expenditure and ROSCA participation. We thus propose the following:

Hypothesis 4: Bank savings increase with income level, while there is an inverted U-shaped relationship between income level and ROSCA savings.

Finally, we add a variable reflecting the individual's social network to my estimation. As ROSCAs are formed via pre-existing social networks, workers with a high degree of social connectedness on cut flower farms may be more likely to participate in ROSCAs, and their savings in ROSCAs may be higher. Regarding other individual-level variables, we predict that workers with high levels of education and financial literacy tend to use bank accounts more, following the literature (Carpenter and Jensen, 2002; Kedir et al., 2011). We thus propose the following:

Hypothesis 5: ROSCAs savings increase with social connectedness in the workplaces.

Hypothesis 6: Bank savings increase with high levels of education and financial literacy.

2.3. Data and Estimation Strategy

2.3.1. Data

The cut flower industry in Ethiopia is a labor-intensive export-oriented industry. It is one of the priorities of the government of Ethiopia, which is the second-largest cut flower exporter in Africa after Kenya. The cut flower sector creates a large number of low-skilled, low-income jobs. Most of the production workers on cut flower farms are women who have low education levels and lack other job opportunities.

To investigate the cut flower industry and its workers, I conducted a survey on workers along with a census of flower farms in January 2018 in collaboration with Hitotubashi University and the Ethiopian Development Research Institute. This survey was conducted on randomly chosen four farms located in different cut flower clusters: Holeta, Sebeta, and Bishoftu. All the production workers on the four farms were interviewed in the survey, totaling 710 workers. The data include detailed information on the workers' financial activities, risk preferences, time preferences, financial literacy, social networks, and socio-economic characteristics. We dropped one observation of missing information, and this leaves us with a total sample of 709 workers.

To measure risk preferences, I conducted a risk preference game (see Table 2.1) with real payoffs based on the expected utility theory for the sake of the simplicity, following Suzuki (2015).[5] I asked each worker to choose either Project A, in which the worker was guaranteed a payoff, or Project B, in which the worker had a 50 percent chance of winning or of losing and receiving nothing. I offered eight games sequentially with increasing payoffs for Project A. Thus, workers who switched their choice from Project B to Project A earlier can be considered risk averse.

[5] Prospect theory make use of three parameters; nonlinear weighting of probabilities, aversion to lass compared to gain, and risk aversion while expected utility theory simply uses risk aversion as one main parameter (Tanaka et al. 2010). Expected theory approach could be regard as a special case of prospect theory (Suzuki, 2015).

Table 2.1. Risk Preference Game

	Project A	**Project B**	
	You obtain for sure:	50% chance of obtaining:	50% chance of obtaining
RG1	5 Birr	40 Birr	0 Birr
RG2	10 Birr	40 Birr	0 Birr
RG3	15 Birr	40 Birr	0 Birr
RG4	20 Birr	40 Birr	0 Birr
RG5	25 Birr	40 Birr	0 Birr
RG6	30 Birr	40 Birr	0 Birr
RG7	35 Birr	40 Birr	0 Birr
RG8	40 Birr	40 Birr	0 Birr

I also conducted a time preference game (see Table 2.2) to measure impatience and present-biased time-inconsistent preferences (Tversky and Kahneman, 1986; Benzion et al., 1989; Ashraf et al. 2006). In this game, the worker was asked to choose Option A, in which the worker could receive a payment today, and Option B, in which the worker could receive a payment three months later. I offered 11 games sequentially in which the payment for Option B increased. The impatience level can be measured where workers switched their choice from Option A to Option B. The worker could choose Option A (receive a payment today) in all games, which could be considered the highest impatience level. I offered a similar game to measure present-biased time-inconsistent preferences in which the payment timing for Options A and B were changed to three months and six months later, respectively. Workers who exhibited less patience for current trade-offs than future trade-offs could be considered to be present-biased time-inconsistent.

Table 2.2. Time Preference Game

	Option A **Today**	**Option B** **3 months later**
TP1	20 Birr	21 Birr
TP2	20 Birr	22 Birr
TP3	20 Birr	24 Birr
TP4	20 Birr	26 Birr
TP5	20 Birr	28 Birr
TP6	20 Birr	30 Birr
TP7	20 Birr	32 Birr
TP8	20 Birr	34 Birr
TP9	20 Birr	36 Birr
TP10	20 Birr	38 Birr
TP11	20 Birr	40 Birr
	Option A **3 months later**	**Option B** **6 months later**
TP21	20 Birr	21 Birr
TP22	20 Birr	22 Birr
TP23	20 Birr	24 Birr
TP24	20 Birr	26 Birr
TP25	20 Birr	28 Birr
TP26	20 Birr	30 Birr
TP27	20 Birr	32 Birr
TP28	20 Birr	34 Birr
TP29	20 Birr	36 Birr
TP30	20 Birr	38 Birr
TP31	20 Birr	40 Birr

Workers' social connectedness is measured using the social network variable through the random matching within sample technique (Conley and Udry, 2010; Maertens and Barrett, 2013; Murendo et al., 2018). This method is more time-efficient than a census and is better at capturing both strong and weak network links. However, this technique may omit key network nodes, resulting in omitted variable bias. Each worker was matched with five other workers randomly drawn from the same cut flower farm and was asked whether he/she knew each of the matched workers. This variable is used as a proxy variable for workers' social connectedness on their farms (Murendo et al., 2018).

Table 2.3. Descriptive Statistics (N=709)

Variable	Mean	SD	Min	Max
= 1 if female	0.85	0.36	0	1
Age	26.8	8.62	16	60
= 1 if married	0.38	0.48	0	1
= 1 if Oromo ethnic group	0.76	0.43	0	1
Years of schooling	4.46	4.13	0	16
= 1 if migrant	0.58	0.49	0	1
Household size	2.34	2.01	0	9
Risk averse index (8: the most risk averse)	4.01	2.7	0	8
Impatience index (12: the most impatient)	10.12	3.48	1	12
= 1 if present-biased time-inconsistent preference	0.14	0.35	0	1
Financial literacy (5: the highest number of correct answers)	1.82	1.11	0	5
Social network size (5: the highest degree of social connectedness)	1.51	1.59	0	5
The total hectares of farmland a worker owns (ha)	0.18	0.71	0	10
= 1 if a worker owns livestock	0.23	0.42	0	1
Food expenditure (birr, month)	769.87	564.83	0	8,000

Table 2.3 reports the descriptive statistics. As is true of the cut flower sector, the percentage of female workers is very high. The average age is about 27 years old, and the average years of schooling is 4.5. More than half the workers are migrants in the sample. Cut flower farms are located near the capital city of Addis Ababa, in close proximity to the international airport, and cut flower workers live in peri-urban areas near the farms.

The average monthly income of the workers in the sample is around 1,130 birr, which is approximately 42 USD (1 USD is approximately 27 birr as of January 2018). Around 19 percent of workers had sent money to family members or relatives in the past 12 months, and the average annual remittance is around 1,250 birr (46 USD as of January 2018).

Table 2.4. Descriptive Statistics of Savings by Saving Methods (birr, year)

Saving methods	No. of workers (%)	Mean	SD	Min	Max
The total savings in any method	604 (85%)	2,941.75	4,100.37	10	56,000
Savings in bank accounts	430 (61%)	1,449.85	3,111.15	10	50,000
Savings in ROSCAs	317 (45%)	2,571.55	2,543.24	160	24,000
Savings in cash at home	205 (29%)	1,649.77	2,346.27	5	18,200

Table 2.4 presents the descriptive statistics on savings by different saving methods. The average total amount of annual savings via bank accounts, ROSCAs or cash savings at home

is approximately 2,942 birr (109 USD as of January 2018). More than half of the workers (around 54 percent) in the sample are being paid via banks, and around 72 percent own a savings account. Around 61 percent of workers save via bank accounts, and around 45 percent of the workers in the sample participate in ROSCAs. Approximately 29 percent of workers keep their savings at home. The average ROSCAs savings is the highest, while the average bank savings is the lowest. Figure 2.1 shows distributions of the amount of savings by saving methods. The ROSCAs savings are relatively tightly distributed, while bank savings are widely distributed. Workers often use three different saving methods at the same time or use two different methods simultaneously. Table 2.5 presents descriptive statistics of seven combinations of three different saving methods which is mutually exclusive.

Table 2.5. Descriptive Statistics of Mutually Exclusive Savings Choices (birr, year)

Saving methods (mutually exclusive)	No. (%)	Total savings	Bank	ROSCAs	Cash
Bank only	172 (24%)	1,158	1,158	-	-
ROSCAs only	79 (11%)	2,320	-	2,320	-
Cash at home only	62 (9%)	1,684	-	-	1,684
Bank and ROSCAs simultaneously	148 (21%)	4,487	1,691	2,795	-
Bank and cash simultaneously	53 (7%)	3,057	1,797	-	1,260
ROSCAs and cash simultaneously	33 (5%)	4,348	-	2,331	2,017
Bank, ROSCAs and cash simultaneously	205 (8%)	5,622	1,382	2,478	1,763

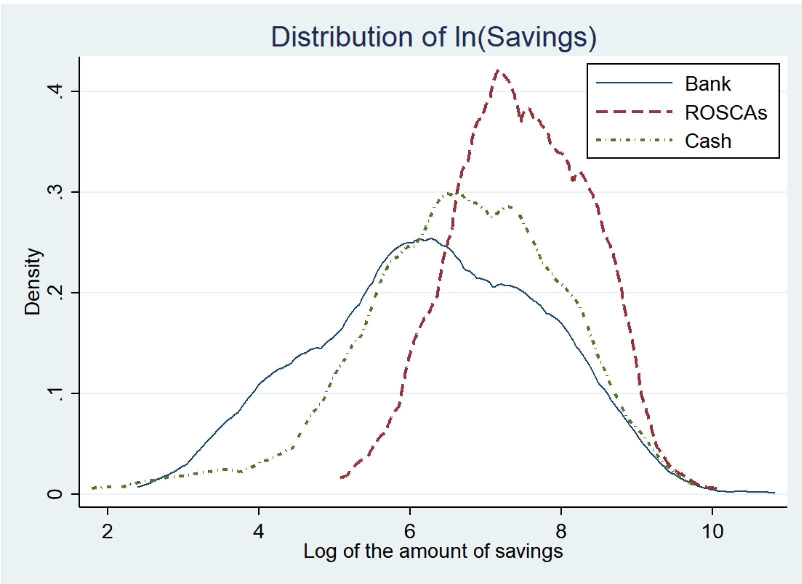

Figure 2.1. Distribution of Log of the Amount of Savings by Saving Methods

Regarding ROSCAs savings, among the workers who answered that they had participated in ROSCAs in the past 12 months, almost all (94%) participated in only one savings group, and approximately 83 percent of workers answered that they had formed their group with other workers at the same farm. Further, 83 percent of the workers answered that they had never participated in ROSCAs before they joined their farm. The average number of ROSCA members on the cut flower farms is around six, and the average one-time payment is around 497 birr (around 18 USD as of January 2018).

As for the major purpose of savings, more than 72 percent of workers answered that they saved money to cope with emergencies. Other responses are as follows: “To buy durables” (6%); “To start or grow a business” (6%); “Migrate to work overseas” (4%); “Asset building” (3%); “For children’s future (3%); “Medical” (2%); “Education” (2%); and “Others” (1%).

2.3.2. Unique Contextual Features

I explore the determinants of the savings choice among formal bank accounts, ROSCAs and home savings which are three most commonly used saving methods using unique contextual features of production workers laboring on cut flower farms in Ethiopia. First, workers laboring cut flower farms have full access to formal banks as well as informal savings groups. Since cut flower workers live in peri-urban areas near cut flower farms, where commercial banks and ATMs are accessible, distance to bank is not a barrier for them as opposed to the rural area of developing countries. Further, unlike rural areas wherein geographical distance between households and irregular income might make it difficult to organize and sustain informal savings groups, regular income and working environment with concentration of individuals who are paid at the same time may stimulate to organize and maintain informal saving groups. Based on the survey data, more than half of the production workers at cut flower farms are migrants from rural areas, and around 83 percent answered that they had never participated in ROSCAs before they joined the cut flower farm. Many unskilled migrant workers migrated to find jobs in labor-intensive industries such as manufacturing and horticulture, and these migrant workers tend to lack social networks in their new settlements. Working on the cut flower farms allow these workers to easily form informal savings groups at their workplace.

Second, cut flower workers are familiar with using bank accounts and there is less concern about trust issues in formal banking institutions. More than half of the workers (around 54 percent) in the sample are being paid via banks, and around 72 percent own bank accounts. Approximately 72 percent of workers answer that bank savings is the most preferred method of saving. ROSCAs savings are most preferred by 21 percent of workers, and only 6 percent of workers choose cash savings at home as the most preferred saving method.

These contextual features allow us to examine the demand side of the saving methods used among the poor when they have full access to formal and informal financial institutions, and the analysis of saving behaviors of cut flower production worker could provide an in-depth description of financial behaviors and need of the poor.

2.3.3. Estimation Strategy

To investigate the determinants of the saving choice among three different methods, the following model is estimated using multivariate probit model and seemingly unrelated regression (SUR) model:

$$Y_{ij} = \beta_0 + \beta_1 X_{ij} + \beta_2 H_{ij} + \beta_3 R_{ij} + \beta_4 T_{ij} + \mu_j + \varepsilon_{ij} \tag{2.1}$$

where i denotes a worker; j denotes the worker's farm; Y_{ij} is a binary variable indicating if a worker saves via bank accounts/ROSCAs/cash savings at home (multivariate probit model) or the log of total savings via bank accounts/ROSCAs/cash savings at home (SUR); X_{ij} is individual characteristics such as gender, age, education, ethnicity, marital status, and migrant status; H_{ij} captures the wealth variables of the worker's household such as land ownership, livestock ownership, and total amount of food expenditure; R_{ij} is the worker's risk preference index; T_{ij} is the worker's time preference and impatience index; μ_j is a farm fixed effect; and ε_{ij} is an error term.

If the decisions to save in banks, ROSCAs, and cash savings at home are correlated, conducting OLS on three equations separately may yield inefficient results. Thus, we estimate three equations using Zellner's seemingly unrelated regression (SUR) to enhance efficiency.

2.4. Estimation Results

I examine the determinants of the saving method choice among bank accounts, ROSCAs and cash savings at home using multivariate probit estimation and seemingly unrelated regression (SUR) estimation. Table 2.6 presents the multivariate probit estimation results of the determinants of using three different saving methods.[6] The likelihood ratio test for the null hypothesis that all correlation coefficients are zero is rejected at the 1 percent significance level (χ^2 = 12.41). Among three covariances of the error terms between two saving methods, only the covariance of the error terms between bank accounts and ROSCAs is positive and significant, suggesting that unobserved factors which increase the probability of using bank accounts increase the probability of using ROSCAs. This indicates bank accounts and ROSCAs are complements. Other covariances of the error terms are positive but insignificant. Table 2.7 presents the seemingly unrelated regression estimation results. The p-

[6] The probit estimation results of the determinants of mutually exclusive savings choices (the total of 7 combinations of 3 different saving methods) are reported in Table A.1. in the Appendix.

value of the Breusch–Pagan test for error independence is 0.017, which indicates that the estimation employing SUR is better than that using OLS.[7]

Table 2.6. Determinants of Savings Choice (Multivariate Probit Estimates)

Dependent variable: = 1 if save using	**(1) Bank**	**(2) ROSCAs**	**(3) Cash at home**
Individual-level characteristics			
= 1 if female	-0.315 *	0.104	0.089
	(0.162)	(0.142)	(0.151)
Years of schooling	0.043 **	0.032 **	0.019
	(0.017)	(0.014)	(0.015)
Asset and income proxy variables			
The total hectares of farmland a worker owns	0.127	0.112	0.118
	(0.091)	(0.070)	(0.072)
= 1 if a worker keeps livestock	0.418 ***	0.119	0.503 ***
	(0.157)	(0.145)	(0.157)
The log of total food expenditures	0.152	0.197 **	-0.225 **
	(0.097)	(0.093)	(0.104)
Risk and time preferences			
Risk averse (8: the most risk averse)	0.015	0.031	-0.034
	(0.022)	(0.019)	(0.021)
Impatience (12: the most impatient)	-0.022	0.037 **	-0.007
	(0.016)	(0.015)	(0.016)
= 1 if present-biased time-inconsistent preference	0.127	0.120	0.343 **
	(0.164)	(0.142)	(0.148)
Cognitive skills			
Financial literacy (5: the highest number of correct answers)	0.069	0.038	-0.052
	(0.058)	(0.051)	(0.055)
Correlation coefficients:			
rho21 (Bank and ROSCAs)		0.236 *** (0.073)	
rho31 (Bank and Cash)		0.064 (0.079)	
rho32 (ROSCAs and Cash)		0.108 (0.068)	
Log pseudo-likelihood		-1121.495	
Wald chi2 (54)		483.85 ***	
Observations		709	

Note: 1. Robust standard errors in parentheses *** p<0.01, ** p<0.05, * p<0.1.
2. Each regression includes farm fixed effects. Controls include age, age squared, marital status, migrant, Oromo ethnic group, and household size.

[7] There may be possible selection bias if saver group is systematically different from non-saver group. Heckman two-step estimation results cannot reject the hypothesis of independence between the errors of the selection equation and the outcome equation, suggesting that we do not necessarily consider endogenous sample selection.

Table 2.7. Determinants of Savings Choice (Seemingly Unrelated Regression Estimates)

Dependent variable: The log of the amount of savings using	**(1)**	**(2)**	**(3)**	**(4)**	**(5)**	**(6)**
	Bank		**ROSCAs**		**Cash at home**	
Individual-level characteristics						
= 1 if female	-0.845 ***	-0.813 ***	0.106	-0.008	0.134	0.126
	(0.287)	(0.289)	(0.378)	(0.378)	(0.294)	(0.296)
Years of schooling	0.120 ***	0.122 ***	0.089 **	0.084 **	0.016	0.016
	(0.030)	(0.030)	(0.039)	(0.039)	(0.031)	(0.031)
Asset and income proxy variables						
The total hectares of farmland a worker owns	0.341 **	0.345 **	0.238	0.222	0.317 **	0.315 **
	(0.154)	(0.154)	(0.202)	(0.201)	(0.157)	(0.158)
= 1 if a worker keeps livestock	0.699 **	0.704 **	0.524	0.507	0.978 ***	0.977 ***
	(0.301)	(0.301)	(0.396)	(0.394)	(0.308)	(0.308)
The log of total food expenditures	0.197	0.205	0.599 **	0.569 **	-0.504 ***	-0.506 ***
	(0.188)	(0.188)	(0.247)	(0.246)	(0.192)	(0.192)
Risk and time preferences						
Risk averse (8: the most risk averse)	0.044	0.045	0.083	0.082	-0.085 **	-0.085 **
	(0.041)	(0.041)	(0.053)	(0.053)	(0.041)	(0.041)
Impatience (12: the most impatient)	-0.042	-0.044	0.094 **	0.101 **	-0.014	-0.013
	(0.030)	(0.030)	(0.040)	(0.040)	(0.031)	(0.031)
= 1 if present-biased time-inconsistent preference	0.077	0.063	0.449	0.501	0.802 ***	0.805 ***
	(0.299)	(0.299)	(0.393)	(0.392)	(0.306)	(0.306)
Cognitive skills						
Financial literacy (5: the highest number of correct answers)	0.112	0.121	0.114	0.083	-0.022	-0.024
	(0.106)	(0.107)	(0.140)	(0.140)	(0.109)	(0.109)
Social networks						
Social network size (5: the highest degree of social connectedness)		-0.080		0.284 ***		0.019
		(0.084)		(0.110)		(0.086)
R^2	0.322	0.322	0.109	0.118	0.211	0.211
Observations	709					

Note: 1. Standard errors in parentheses *** p<0.01, ** p<0.05, * p<0.1.
2. Each regression includes farm fixed effects. Controls include age, age squared, marital status, migrant, Oromo ethnic group, and household size.

2.4.1. Risk Coping Mechanisms

Column (1) of Table 2.6 and columns (1) and (2) of Table 2.7 show that savings using bank accounts significantly increase for workers who own land and livestock, whereas there is no significant relationship between land and livestock ownership and the amount of savings using ROSCAs. Land and livestock assets could be regarded as buffer stocks in times of need since liquidizing assets is used as a risk coping strategy in developing countries (Deaton, 1992). This means that workers who have buffer stocks which would help them cope with unexpected shocks are more likely to use formal bank accounts to save rather than ROSCAs. This result supports Hypothesis 1. To put it another way, those who are more insured with livestock or land assets are not likely to engage in ROSCA. This may be suggesting that workers who use ROSCAs are valuing the social insurance aspects of ROSCAs. Interestingly, cash savings at home (column (3) of Table 2.6 and columns (5) and (6) of Table 2.7) also have a positive relationship with asset ownership, suggesting that self-insured workers with illiquid asset ownership are more likely to use the saving methods with more flexibility in terms of liquidation.

Table 2.8. Determinants of Savings Choice Using Saver Sub-Sample (Seemingly Unrelated Regression Estimates)

Dependent variable:	**(1)**	**(2)**	**(3)**	**(4)**	**(5)**	**(6)**
The log of the amount of savings using	**Bank**		**ROSCAs**		**Cash at home**	
Individual-level characteristics						
= 1 if female	-1.121 ***	-1.069 ***	-0.309	-0.472	0.003	0.009
	(0.294)	(0.296)	(0.408)	(0.409)	(0.322)	(0.325)
Years of schooling	0.093 ***	0.094 ***	0.040	0.036	-0.017	-0.016
	(0.030)	(0.030)	(0.042)	(0.042)	(0.033)	(0.033)
Asset and income proxy variables						
The total hectares of farmland a worker owns	0.384 ***	0.390 ***	0.226	0.205	0.297 *	0.298 *
	(0.148)	(0.148)	(0.205)	(0.204)	(0.162)	(0.162)
= 1 if a worker keeps livestock	0.460	0.471	0.182	0.147	0.989 ***	0.990 ***
	(0.293)	(0.293)	(0.407)	(0.404)	(0.321)	(0.321)
The log of total food expenditures	0.244	0.258	0.760 ***	0.716 ***	-0.358 *	-0.356 *
	(0.188)	(0.188)	(0.261)	(0.259)	(0.206)	(0.206)
Risk and time preferences						
Risk averse (8: the most risk averse)	0.057	0.057	0.112 **	0.112 **	-0.079 *	-0.079 *
	(0.041)	(0.041)	(0.057)	(0.057)	(0.045)	(0.045)
Impatience (12: the most impatient)	-0.068 **	-0.070 **	0.074 *	0.081 *	-0.023	-0.023
	(0.031)	(0.031)	(0.043)	(0.043)	(0.034)	(0.034)
= 1 if present-biased time-inconsistent preference	-0.062	-0.080	0.327	0.382	0.686 **	0.684 **
	(0.296)	(0.296)	(0.412)	(0.410)	(0.325)	(0.325)
Cognitive skills						
Financial literacy (5: the highest number of correct answers)	0.105	0.115	0.082	0.051	-0.106	-0.105
	(0.106)	(0.106)	(0.148)	(0.147)	(0.117)	(0.117)
Social networks						
Social network size (5: the highest degree of social connectedness)		-0.107		0.335 ***		-0.012
		(0.085)		(0.118)		(0.093)
R^2	0.345	0.347	0.149	0.160	0.283	0.283
Observations	604					

Note: 1. Standard errors in parentheses *** p<0.01, ** p<0.05, * p<0.1.
2. The p-value of the Breusch–Pagan test for error independence is 0.001
3. Each regression includes farm fixed effects. Controls include age, age squared, marital status, migrant, Oromo ethnic group, and household size.

I conducted the same analyses only using saver subsample to compare three different saving method choices among savers. Table 2.8 presents the estimation results. Columns (1), (2), (5) and (6) of Table 2.8 show that within saver group, workers with higher land assets tends to save more in bank accounts while workers with higher livestock assets are more likely to keep higher cash savings at home.

Table 2.8 shows that among savers, workers who are more risk-averse have a significantly increased ROSCAs savings, while this is not the case for bank savings and cash savings. In addition, risk-taker workers tend to increase their cash savings at home, which weakly suggests that one of motivations for cash savers to keep their savings at home may be for immediate investment in the optimum time to invests. This supports Hypothesis 2.

As mentioned in Section 2.2, if risk-averse workers consider the risk of losing their savings or of losing interest income, they may have a tendency to increase their bank savings. However, if risk-averse workers consider insurance component of ROSCAs, they may prefer to use ROSCAs more. ROSCAs have an insurance feature whereby group members who have not received pot money can obtain money in times of unexpected need if other group members are willing to switch the receipt order. A significant positive relationship between risk aversion and ROSCAs savings indicates that ROSCAs are used more for insurance purposes. Though ROSCAs cannot effectively insure against covariate risks in socially homogenous groups like cut flower farm workers, ROSCAs can provide insurance against idiosyncratic risks such as illness (Calomiris and Rajaraman, 1998). Participating in ROSCAs might be a risk-coping strategy for workers given that formal insurance systems are underdeveloped in Ethiopia.

Regarding loss risk, although savings via bank accounts carry less of a loss risk, ROSCAs in cut flower farms also seem to be a relatively low-risk saving method because ROSCAs on cut flower farms are relatively small, and trust among cut flower workers is relatively high according to my field observations. Kedir et al. (2011) suggest that individuals join ROSCAs from a risky investment motive. However, the ROSCA sample Kedir et al. (2011) studied comprised urban ROSCAs, which typically last for around two years, suggesting that the average number of participants is around 24 people. On the other hand, my sample is drawn from peri-urban areas, and ROSCAs at cut flower farms are relatively small, consisting of an average of around six members, and the average monthly contribution per person is approximately 497 birr (roughly 18 USD as of January 2018). According to the ROSCA classification Bisrat et al. (2012) used,[8] ROSCAs at cut flower farms can be classified as small *Equbs* with regards to monthly contributions and number of participants. Since *Equbs* at cut flower farms are relatively small, cut flower farm workers are unlikely to participate in ROSCAs from a risky investment motive. In addition, 60 to 80 percent of the sample used by Kedir et al. (2011) reported that their objective for saving was to purchase capital assets or to start a business, while around 73 percent of savers in the sample reported that their main purpose was to cope with emergencies, implying that the risky investment motive may not be relevant to my sample.

Table 2.9. The Relationship Between the Perceived Credit Limits and the Propensity of Using Bank Accounts/ROSCAs/Cash (OLS Estimates)

Dependent variable: The log of the perceived credit limits	**(1) Bank**	**(2) ROSCAs**	**(3) Cash at home**
Propensity of using bank accounts/ROSCAs/cash	-1.677 (2.476)	16.539 ** (7.641)	7.145 ** (2.957)
Other control variables		YES	
Farm fixed effect		YES	
Observations		709	

Note: 1. Robust standard errors in parentheses *** p<0.01, ** p<0.05, * p<0.1.
2. Each regression includes farm fixed effects. Controls include gender, education, land asset, livestock asset, income proxy variable, risk preference, impatience index, time-inconsistent preference, financial literacy, age, age squared, marital status, migrant, Oromo ethnic group, and household size.

[8] Bisrat et al. (2012) classified ROSCAs in Ethiopia into three categories based on their size: large, medium, and small. The pot money in large *Equbs* is collected weekly, and the minimum contribution per person is 1,000 birr. The monthly contribution of medium-sized *Equbs* is between 500 to 1,000 birr, and the contribution per person in small *Equbs* is less than 500 birr.

To further investigate whether workers used ROSCA savings from an insurance motive, I examine the relationship between the propensity of using each saving method and the perceived credit limit which is based on workers' responses to the following question: "What is the maximum amount that you could borrow within a week if you are faced with an emergency, such as illness?" The results of Table 2.9 show a positive relationship between the perceived credit limit and the propensity to save in ROSCA.[9] On the other hand, the propensity to save in banks is negatively correlated to the perceived credit limit (not significant). This indicates that the amount of money workers believe they could borrow in an emergency is likely to increase if the propensity to save in ROSCAs increases. This result suggests that ROSCAs can serve as an insurance substitute in the context where formal insurance systems are weak, and that cut flower workers are motivated to use ROSCAs as an insurance tool against unexpected risks.[10] The propensity to save in cash at home is also positively related to the perceived credit limits. The possible reason may be related to the fact that the average of cash savings is relatively higher than that of bank savings (Table 2.4 and Figure 2.1) and cash savings are easy to liquidize compared to other methods, which may increase the possibility of mutual lending, but it needs further investigation.

2.4.2. Commitment Tools

Columns (3) and (4) of Table 2.7 show that workers who are impatient tend to save more using ROSCAs, although we find no significant relationship between present-biased time-inconsistent preference and ROSCAs savings. This finding partially supports Hypothesis 3. Impatient workers are more likely to join ROSCAs because they potentially have a

[9] The propensity of using bank accounts/ROSCAs/cash savings at home is calculated based on the results of the multivariate probit estimation of Table 2.6.

[10] Social interactions at the workplace can play a significant role in risk-coping strategies. Around 50 percent of workers answered they had someone at their workplace whom they could rely on for financial support in times of need, and around 68 percent answered that they had someone they could rely on for moral support at their workplace.

preference for commitment feature of ROSCAs in order to save regularly (Ashraf et al., 2006; Gugerty, 2007). Workers participating in ROSCAs are likely to set savings goals together with other members and strive to meet these goals. The economic and social punishments associated with default (e.g., loss of reputation) provide incentives to do so.

Columns (1) to (4) of Table 2.8 also show that among savers, impatient workers are more likely to use ROSCAs, while patient workers tend to save using bank accounts. I further find a significant positive relationship between present-biased time-inconsistent preference and cash savings at home. One possible explanation is that workers are not sophisticated enough to be aware of their time-inconsistent preferences. If individuals with present bias are sophisticated enough to be aware of their present-biased preferences, they may prefer to take advantage of commitment devices such as ROSCAs to overcome their self-control problems. However, if individuals are (partially) naïve, they underestimate their demand for commitment devices (Ashraf et al., 2006; Kremer et al., 2019). This may partially explain why present-biased preference is not significantly associated with ROSCAs savings, but it correlates with cash savings at home, which is inconsistent with literature.

Workers who need for the commitment properties may participate in ROSCAs not only to cope with self-control problem, but also sometimes to respond to bargaining power within a household (Anderson and Baland, 2002; Ambec and Treich, 2007; Gugerty, 2007). Anderson and Baland (2002) find an inverted U-shaped relationship between female income share of household and ROSCA participation, suggesting that motivation for women to join ROSCAs is weak if women have sufficient bargaining power in the household.

To check whether women workers' use of ROSCAs is associated with their bargaining power within a household, I estimated the same model with adding women's income share of household dummies using married women and unmarried women subsamples. Women's income share of household is a proxy for their bargaining power in household. Column (2) of

Table 2.10 shows that there is an inverted U-shaped relationship between married women workers' income share of household and ROSCAs savings. At low income share, women workers will not have enough bargaining power, however, as their income share increases, ROSCAs savings increase significantly to protect their savings. At the very high levels of income share, their ROSCAs savings decrease since women may have enough bargaining power (not significant). Even for unmarried women, there could be pressure on their savings from their household members' consumption. Although we find no similar results with married women's case, I find a significant negative relationship when unmarried women have high income share of household, suggesting if unmarried women have sufficient bargaining power, they have weak motivations to commit their savings against intrahousehold conflict by using the commitment properties of ROSCAs (Column (5) of Table 8). Interestingly, if unmarried women have a high bargaining power within the household, their cash savings at home significantly increase while there is no relationship between their income share and bank savings, suggesting that unmarried women prefer to keep their savings at home once they have enough bargaining power.

Table 2.10. Determinants of Savings Choice Using Married and Unmarried Women Sub-Sample (OLS estimates)

Dependent variable: The log of the amount of savings using	**(1)**	**(2)**	**(3)**	**(4)**	**(5)**	**(6)**
		Married Women			**Unmarried Women**	
	Bank	**ROSCAs**	**Cash**	**Bank**	**ROSCAs**	**Cash**
Individual-level characteristics						
Years of schooling	0.165 ***	0.243 ***	0.075	0.085	-0.008	0.013
	(0.060)	(0.074)	(0.053)	(0.055)	(0.058)	(0.049)
Asset & income proxy variables						
The total hectares of farmland a worker owns	0.219	0.291	0.265 *	0.963 **	0.456	-0.165
	(0.151)	(0.177)	(0.159)	(0.375)	(0.509)	(0.631)
= 1 if a worker keeps livestock	0.835 *	-0.227	0.501	0.197	0.424	1.223 *
	(0.456)	(0.660)	(0.460)	(0.467)	(0.824)	(0.686)
The log of total food expenditures	0.136	-0.347	-1.610 ***	-0.091	0.419	-0.669 ***
	(0.443)	(0.482)	(0.425)	(0.252)	(0.299)	(0.255)
Risk and time preferences						
Risk averse (8: the most risk averse)	0.040	-0.012	0.087	0.064	0.106	-0.212 ***
	(0.073)	(0.093)	(0.067)	(0.078)	(0.086)	(0.073)
Impatience (12: the most impatient)	-0.039	0.139 **	-0.009	-0.125 **	0.113	-0.068
	(0.055)	(0.061)	(0.050)	(0.063)	(0.072)	(0.066)
= 1 if present-biased time-inconsistent preference	0.379	0.383	0.972 *	-0.033	0.415	0.850
	(0.492)	(0.711)	(0.548)	(0.546)	(0.674)	(0.623)
Cognitive skills						
Financial literacy (5: the highest number of correct answers)	-0.099	-0.249	-0.181	0.389 *	0.553 **	-0.198
	(0.218)	(0.266)	(0.191)	(0.206)	(0.215)	(0.187)
Income share						
Income share > 0 & < = 25%	0.573	2.124	-0.451	-0.700	-0.343	0.944
	(1.098)	(1.348)	(0.796)	(0.706)	(0.959)	(0.911)
Income share > 25 & < = 50%	0.286	2.109 ***	0.340	-0.681	-0.355	1.406 **
	(0.541)	(0.779)	(0.561)	(0.472)	(0.656)	(0.543)
Income share > 50 & < = 75%	-0.429	1.993 *	-0.194	-2.087	-0.949	-0.651
	(0.708)	(1.097)	(0.609)	(1.616)	(1.318)	(1.273)
Income share > 75 & < = 100%	0.790	-0.958	-0.251	0.087	-1.315 **	1.238 **
	(0.972)	(1.274)	(1.047)	(0.559)	(0.598)	(0.518)
R^2	0.386	0.259	0.313	0.323	0.154	0.236
Observations		222			245	

Note: 1. Robust standard errors in parentheses *** p<0.01, ** p<0.05, * p<0.1.
2. Each regression includes farm fixed effects. Controls include age, age squared, migrant, Oromo ethnic group, and household size

2.4.3. Income Levels

I also examine whether income relates to saving choice differentially. I find that if household expenditure increases (i.e., the income proxy variable),[11] ROSCA savings increase significantly, while there is no significant relationship between household expenditure and

[11] We use total food expenditure as a proxy for household income. The results of using workers' monthly income are similar to the results shown in Table 2.6 and Table 2.7.

the amount of savings in bank accounts (Columns (1) to (4) of Table 2.7). However, the squared term of the income proxy variable is not significant, suggesting that there is no inverted U-shaped relationship between income and ROSCA savings. Thus, Hypothesis 4 is not supported. We also find that cash savings at home have a significant negative relationship with income level.

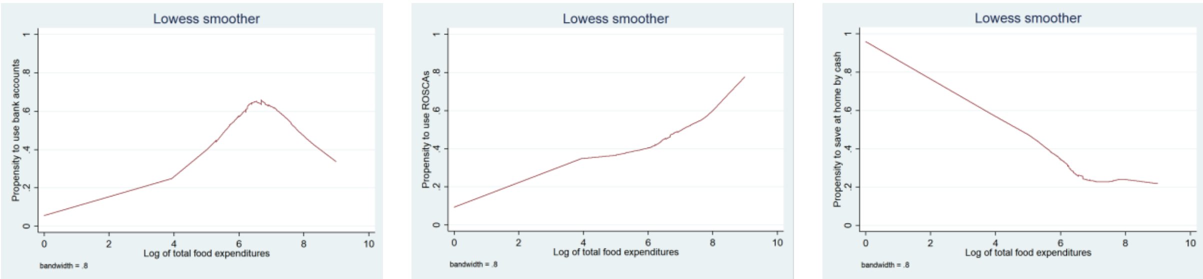

Figure 2.2. The Relationship Between Household Expenditure Variable and the Propensity to Save in Bank Accounts/ROSCAs/Cash (Lowess Regression Function)

To investigate the relationship between the household expenditure variable and workers' savings decisions, I use a locally weighted regression which tends to follow the data by providing desirable smoother in a nonparametric way (Cameron and Trivedi, 2010). Figure 2.2 shows that the propensity to save in both ROSCAs and bank accounts increase as household expenditure increases but the propensity to save in bank accounts and household expenditure variables show a negative relationship after a threshold is reached. This suggests that cut flower production workers tend to save in ROSCAs if their income increases but this is not the case for bank savings. This sample comprises low-income and unskilled workers in a homogeneous occupation whose income variation is relatively narrow, which means that their average income could be lower than the threshold at which ROSCA savings are reduced, as observed by Kedir et al. (2011) and Kedir and Ibrahim (2011). This positive relationship between income and ROSCA participation among average-income individuals is supported theoretically (Ambec and Treich, 2007) and is consistent with the empirical literature (Anderson and Baland, 2002; Levenson and Besley, 1996). Regarding bank savings, around

54 percent of the workers in this sample are paid via bank accounts as some cut flower farms pay their workers that way. Some low-income workers who cannot join ROSCAs for lack of enough money for savings may hold their wages in their bank accounts by default. After the threshold, bank savings decrease because workers may start joining ROSCAs. This may explain the inverted U-shaped relationship between the propensity to save in bank accounts and household expenditure, which contrasts with other studies' finding that income level has a significant positive effect on the usage of bank accounts (Carpenter and Jensen, 2002; Kedir et al., 2011).[12] Regarding cash savings at home, the propensity to keep savings at home decreases as household expenditure increases. It is also worth noting that workers with higher asset ownership tend to save more in bank accounts or home savings, while workers with higher income level are more likely to use ROSCAs savings.

2.4.4 Social Connectedness at Workplaces and Other Variables

Columns (2), (4) and (6) of Table 2.7 include a social network size variable measured via a random matching within sample technique following Conley and Udry (2010), Maertens and Barrett (2013), and Murendo et al. (2018). Each worker was matched with five other workers randomly drawn from the same cut flower farm and was asked whether the worker knew each of the matched workers. This network size variable should not be interpreted quantitatively but can be interpreted as a proxy variable for workers' social connectedness on their farms (Murendo et al., 2018). This variable could be endogenous, but we can still use it to investigate the relationship between social networks and saving behaviors, which has rarely been examined in the literature. The results show that workers with high degrees of

[12] The literature on savings has focused on the community context where heterogeneities in households' income and occupations are significant. Kedir et al. (2011) and Kedir and Ibrahim (2011) find that the propensity of households using ROSCAs first increases and then decreases after a threshold is reached, which means that some households with high wealth levels withdraw from ROSCA participation. On the other hand, the propensity of households using bank accounts increases as household wealth increases.

social connectedness at their workplaces have a significantly positive relationship with savings in ROSCAs, which confirms Hypothesis 5. Working on the cut flower farms allow workers to easily form social networks at their workplace (Getahun and Villanger, 2018). Suzuki at al. (2018) discuss how social interactions, facilitating ROSCAs, can lead to higher saving rates on cut flower farms in Ethiopia.

Moreover, male workers and workers with a high level of education are more likely to use bank accounts, while financial literacy has no significant relationship with workers' saving behaviors (see Tables 2.6 and 2.7), partially supporting Hypothesis 6. I also find that a high level of education has a positive relationship with ROSCAs savings. This result contrasts with the literature's finding that financial literacy influences the usage of formal financial institutions (Carpenter and Jensen, 2002; Kedir et al., 2011). This discrepancy might occur because the financial literacy of workers in this sample is relatively low (with an average financial literacy test score of 1.82 out of 5), and they may not be required to have advanced financial literacy at their level of savings. However, Ethiopia's commercial banks regularly hold seminars to promote the use of formal bank accounts, and more than 70 percent of workers have bank accounts and know how to use them. Therefore, a low level of financial literacy might not be preventing unskilled low-income workers from saving in formal bank accounts. Rather, they may choose either formal or informal saving methods strategically and seek to link the two methods.

2.5. Chapter Summary

I have found that ROSCAs work as insurance and commitment tools, complementing formal financial institutions. This result suggests that vulnerable populations such as unskilled production workers use both formal and informal methods of savings strategically and allocate their savings by linking between them, since formal financial institutions and ROSCAs provide different functions. Though we cannot rigorously rule out investment motives for their savings, the results suggest that cut flower farm workers choose ROSCA savings even with availability of formal saving methods because ROSCAs serve as an insurance tool and commitment device.

This chapter contributes to the literature on saving behaviors by providing empirical evidence on the risk-sharing mechanism and commitment motives of informal saving groups for people with full access to formal banks. This sample is drawn from low-income employees in a labor-intensive production industry with access to formal banks. Thus, my results may not be generalizable to other occupational groups since saving behaviors can vary across occupations (Dupas and Robinson, 2013). Future research should seek to understand saving method preferences and saving habits in developing countries and disentangle their motives. Exploring the mechanism and motivations regarding informal financial institutions is important because the result may assist in designing financial inclusion polices that are effective for the poor.

Results of this chapter indicate that policy interventions aimed at improving access to formal financial institutions might be more effective if they were followed by formal risk-sharing instruments, as vulnerable people who struggle to cope with risks tend to choose informal savings group from insurance motive. Moreover, bank accounts with commitment features may attract savers participating in informal saving groups from commitment motives to overcome their self-control problems.

3. Internal Migrants' Saving Behavior and Working Environment

3.1. Introduction

Many unskilled migrant workers migrated from rural area to find jobs in labor-intensive industries such as manufacturing and horticulture sector in developing countries. However, internal migration within developing countries has received less attention compared to international migration and empirical evidence of the internal migrants' working environment in the labor market in sub-Saharan Africa is scarce. Internal migrant plays a central role in the urbanization process and is often viewed as the labor market adjustment to the inter sectoral shift in importance from agriculture to manufacturing and services (ILO, 2016). Considering the importance of the non-agricultural labor market for poverty reduction, it is important to understand financial aspects and working condition of internal migrants.

Using the same case of cut flower production workers in Ethiopia as the previous chapter, this chapter first explores saving behaviors and financial aspects of internal migrants since more than half of workers are rural-to-urban migrants in this sample. To my knowledge, there is very limited literature regarding saving behaviors of internal migrants since most of the literature investigate remittance of international migrants and its effect on their rural households.

This chapter also aims to analyze differences in wages and productivity between migrants and non-migrants to understand internal migrants' working environment. Labor market discrimination defined as a situation in which persons who are equally productive receive different wages (List and Rasul, 2011). There are two main economic models of discrimination. The first model is taste-based discrimination, which means employers maximize their utilities based on their prejudice towards minorities (Becker, 1957). The second one is statistical discrimination, which indicates that employers maximize firm profits by discrimination based on observable characteristics because of imperfect information on skills or productivity of minorities. Previous studies focus on wage differentials between migrants and non-migrants based on occupational differences (Meng and Zhang, 2001; Demurger et al., 2009; Lee, 2012). However, this chapter focuses on wage differentials against migrants within homogenous industry context.

The Mincer-type wage regression is estimated, and the Blinder-Oaxaca decomposition approach is used to test for possible wage differentials. This chapter finds that earnings of workers are significantly different between migrant and non-migrant and migrants are likely to be paid less. However, migrants are more likely to be productive and less likely to be absent. The Blinder–Oaxaca decomposition results reveal that about half of wage differentials is unexplained even when non-cognitive skills are controlled, supporting the idea of wage differentials and the possible differential treatment in compensation against migrants in cut flower industry in Ethiopia. With field observation and descriptive statistics on financial behavior of migrants, even though they may face discrimination or suffer from disadvantages in the labor market, higher levels of subjective wellbeing and descriptive analysis results on financial behaviors suggest that rural to urban migration would be investment for better job opportunities.

This chapter contributes to the literature on labor discrimination by providing rich empirical evidence using various earnings and work performance variables: detailed information on earnings, productivity, attendance, and subjective well-being. This chapter focuses on wage differentials between migrants and non-migrants within homogenous industry while previous studies focus on wage differentials based on occupational differences.

The rest of this chapter proceeds as follows. Section 3.2 presents the snapshot of saving behaviors and financial aspects of migrants and summary statistics. Section 3.3 presents model specification and methodologies. Section 3.4 outlines the estimation results. Finally, Section 3.5 concludes this chapter.

3.2. Snapshot of Saving Behaviors of Migrants and Descriptive Statistics

3.2.1. Snapshot of Saving Behaviors of Migrants

In the outskirts of the capital Addis Ababa, where many rural-to-urban migrants settle in, one starting point into the city's formal labor market is horticulture. The cut flower industry is one of the labor-intensive export-oriented industries in Ethiopia, which is the second-largest cut flower exporter in Africa after Kenya. There are three different cut flower clusters: Holeta, Sebeta, and Bishoftu. Since the cut flower sector have created a large number of low-skilled low-income jobs, many of its production workers are migrant workers who migrated from rural areas to find jobs.

With the same dataset as the previous chapter, Table 3.1 presents descriptive statistics of saving behavior by migrants. I define migrant workers who are not born in the village that he or she currently reside in, and most of migrants in this sample migrated for take a job offer or looking for work. The ratio of ROSCAs savings to total amount of savings of migrants is higher than non-migrants. In the survey, even though 74 percent of migrant workers answered that the most preferred savings method is bank accounts, they tend to use ROSCAs savings more. On the other hand, non-migrants tend to use bank accounts more.

Table 3.1. Savings Behavior of Migrants (year, birr)

Variables	(1) Total (N = 709)		(2) Non-migrant (N=300)		(3) Migrant (N= 409)		(4) Diff. (2)-(3)
	Mean	SD	Mean	SD	Mean	SD	
Usage of saving methods							
Bank accounts	0.61	0.02	0.67	0.03	0.56	0.02	0.10***
ROSCAs savings	0.45	0.02	0.43	0.03	0.47	0.03	-0.04
Cash at home	0.29	0.02	0.34	0.03	0.27	0.02	0.08**
Amount of savings							
Log of bank savings	3.81	0.12	4.19	0.18	3.53	0.16	0.67***
Log of ROSCAs savings	3.34	0.14	3.23	0.22	3.42	0.19	-0.19
Log of Cash savings at home	1.93	0.12	2.27	0.19	1.68	0.15	0.59**
Ratio of amount of savings to total savings							
Bank savings/total savings (%)	0.42	0.02	0.43	0.03	0.42	0.02	0.01
ROSCAs savings/total savings (%)	0.38	0.02	0.34	0.02	0.42	0.02	-0.07**
Cash savings/total savings (%)	0.19	0.01	0.23	0.02	0.16	0.02	0.06**

Note: 1. *** p<0.01, ** p<0.05, * p<0.1.

Table 3.2. Association between Migrant and Savings Behavior (Probit Estimates)

Dependent variable: =1 if workers have bank accounts but did not use it	
=1 if migrant	0.349 *** (0.150)

Note: 1. N=709. Robust standard errors in parentheses *** p<0.01, ** p<0.05, * p<0.1.
2.Regression includes farm fixed effects. Controls include gender, education, land asset, livestock asset, income proxy variable, risk preference, impatience index, time-inconsistent preference, financial literacy, age, age squared, marital status, Oromo ethnic group, and household size.

Table 3.2 shows that migrant workers tend not to use bank accounts even though they have it. In the previous chapter, I explored the determinants of the choice of savings methods, and each regression in the previous chapter includes migrant dummy variable. However, I could not find any statistically significant migrant dummy variable.

Regarding financial aspects of migrants, the number of migrants who send money to their family members relatively low. Among the total of 410 migrant workers in this sample, 95 migrants (23%) send money to their family members, and 38 migrants (10%) even received money from others. Migrant workers spent 2.8 times more housing rent than non-migrants and they have less livestock than non-migrants. Migrants' social connectedness within their farm is relatively low and their perceived credit limit (the amount of money they could borrow if they are faced with an emergency) is much lower than non-migrants. However, migrants' proportion of gaining or sharing useful information in their social network within farm is relatively higher than locals. This may indicate that migrants participate in ROSCAs more because of the insurance feature of ROSCAs and information sharing. Interestingly, more than 30 percent of migrant workers have concrete plans to migrate outside Ethiopia (Mostly, for working as a domestic worker in UAE), and this applies the same way for non-migrants as well.

3.2.2. Descriptive Statistics

Table 3.3 reports the descriptive statistics of variables used for wage regression estimation for migrants and non-migrants and Figure 3.1 shows histogram of log of monthly earnings by migrant group. Monthly wages of migrants are much lower than those of non-migrants, and the gap is even bigger for the upper quantile. For worker's performance variables, monthly average productivity data is collected as cut flower farms have their own neat productivity scoring system and some farms keep data on actual production from individual workers and target levels. The descriptive statistics show no mean difference in monthly productivity between migrants and non-migrants. As for attendance, migrants are more likely to do perfect attendance (Cut flower farms have a 6-day workweek). In addition, migrants' subjective well-being is higher than non-migrants. Education level of migrants is higher than non-migrants, while working experience of non-migrants is higher than migrants. Note that, 83 percent of workers answered that they never worked in other cut flower farms prior to joining their current farm, and 70 percent of workers answered that they never had any work experience at all. Non-migrants in this sample is more likely to be Oromo ethnic group since cut flower farms are located near the capital city of Addis Ababa where Oromo ethnic group is dominant. In this sample, Oromo, Amhara, and other ethnic groups account for 76 percent, 14 percent, and 10 percent respectively. As for big-five personality dimensions, which is popular in measuring non-cognitive skills in the context of workplace, migrants show higher big-five dimensions except for extraversion.

Table 3.3. Descriptive statistics

Variables	(1) Total (N = 710) Mean	SD	(2) Non-migrant (N=300) Mean	SD	(3) Migrant (N= 410) Mean	SD	(4) Diff. (2)-(3)
Workers' Earnings							
Monthly wages (birr, month)	1130.14	302.58	1200.11	327.98	1078.94	271.76	121.17***
Q10	900		950		860		
Q50	1053		1100		1050		
Q90	1500		1700		1291.5		
Bonus (birr, year)*	366.66	484.90	372.66	497.58	360.57	472.51	12.10
Total income (birr, year)	13853.95	3654.83	14755.35	3944.64	13194.39	3278.47	1560.96***
Work performance and subjective well-being							
Monthly average productivity*	0.73	0.16	0.73	0.14	0.73	0.18	0.00
=1 if take no days off from work (not missed a day)	0.36	0.48	0.28	0.03	0.42	0.02	-0.14***
Subjective well-being (1-10, 10: the best possible life)	4.85	2.21	4.45	0.10	5.14	0.12	-0.69***
Individual characteristics							
Age	26.80	8.61	26.95	8.73	26.70	8.54	0.26
= 1 if female	0.84	0.37	0.79	0.41	0.88	0.33	-0.09***
Years of schooling	4.45	4.13	3.90	3.91	4.85	4.25	-0.95***
Years of working experience	4.73	3.59	5.16	3.63	4.41	3.54	0.75***
= 1 if married	0.37	0.48	0.35	0.48	0.39	0.49	-0.04
= 1 if Oromo ethnic group	0.76	0.43	0.90	0.30	0.65	0.48	0.25***
= 1 if heard this job openings through friends and relatives	0.58	0.49	0.55	0.50	0.60	0.49	-0.05
Big-Five personality dimensions							
Extraversion	4.16	0.96	4.24	0.90	4.10	1.00	0.14*
Agreeableness	5.60	1.14	5.46	1.13	5.71	1.14	-0.25***
Conscientiousness	5.45	1.15	5.27	1.15	5.58	1.13	-0.31***
Emotional Stability	4.60	1.15	4.48	1.05	4.68	1.22	-0.19**
Openness to Experience	4.59	1.20	4.39	1.06	4.75	1.28	-0.36***

Note: 1. *** p<0.01, ** p<0.05, * p<0.1. 2. The total number of observations for bonus variable is 566 (285 non-migrants and 281 migrants) since one farm have no bonus scheme. 3. The total number of observations for productivity variable is 677 (283 non-migrants and 394 migrants) due to data unavailability.

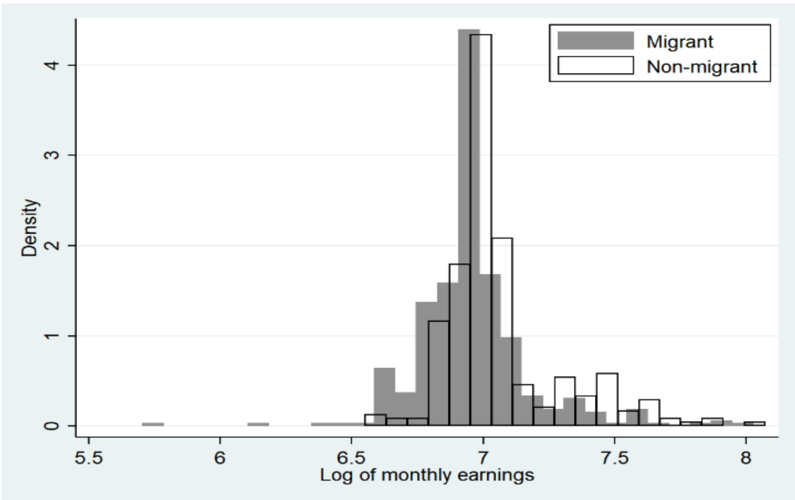

Figure 3.1. Histogram of Log of Monthly Earnings by Migrant

3.3. Empirical Model and Estimation Strategy

3.3.1. Mincer Wage Regression Estimation

To investigate whether there are earnings differentials and work performance difference between migrants and non-migrants, I start with the standard Mincer wage regression with migrant dummy variable.

$$Y_{ij} = \beta_0 + \boldsymbol{\beta_1 migrant_{ij}} + \beta_2 age_{ij} + \beta_3 age^2_{ij} + \beta_4 education_{ij} + \beta_5 experience_{ij} + \beta_6 experience^2_{ij} + \beta_7 X_{ij} + \mu_j + \varepsilon_{ij} \quad (3.1)$$

where i denotes a worker; j denotes the worker's farm; Y_{ij} is workers' earnings variable (the log of monthly wages, bonus, total income) and work performance variable (productivity, attendance, subjective well-being); X_{ij} is individual characteristics such as gender, marital status, ethnicity, new worker dummy variable); μ_j is a farm fixed effect; and ε_{ij} is an error term. To estimate this simplest form of Mincer regression, ordinary least squares (OLS) is employed. If migrant groups face wage differentials, a coefficient of migrant variable would have a minus sign.

$$Y_{ij} = \beta_0 + \boldsymbol{\beta_1 migrant_{ij}} + \boldsymbol{\beta_2 migrant_{ij} \times social\ link_{ij}} + \boldsymbol{\beta_3 social\ link_{ij}} + \beta_4 age_{ij} + \beta_5 age^2_{ij} + \beta_6 education_{ij} + \beta_7 experience_{ij} + \beta_8 experience^2_{ij} + \beta_9 X_{ij} + \mu_j + \varepsilon_{ij} \quad (3.2)$$

The second model is an extended Mincer wage regression with the interaction term of migrant variable and social link dummy variable. This social link variable indicates whether workers heard job openings of this cut flower farm they are working through their friends or relatives. The interaction term of migrant variable and social link dummy variable is added to investigate whether the effect of migrant group is mitigated or aggravate if migrant worker heard job opening through their personal networks. The existing literature suggests that migrant workers in developing countries tend to depend on information on employment opportunities through their personal networks (Banerjee, 1984; Munshi, 2003). Mano et al. (2011) shows that the initial wages of the workers who were recruited with personal networks are significantly lower than that of the formally-recruited (but the negative effect of social links on wages disappears over time) in the context of cut flower industry in Ethiopia, supporting the information-cost hypothesis which predicts low initial wages for the referred applicants since they tend to be unable to find jobs elsewhere so their reservation wages are low (Antoninis, 2006). If this holds true, a coefficient of social link and a coefficient of the interaction terms of migrant and social link variables would have a negative sign.

3.3.2. The Blinder-Oaxaca Decomposition

I also employ the Blinder-Oaxaca decomposition approach which is widely used in literatures on discrimination to analyze the wage differentials (Blinder, 1973; Oaxaca, 1973; Meng and Zhang, 2001; Lee, 2012). This approach is to decompose wage differentials between groups into what can be explained by observables and what cannot be explained by observables. This unexplained differential is regarded as discrimination. More specifically, this approach is to explain the distribution of the outcome variable by a set of factors that vary systematically associate with socioeconomic status. Variations in earnings of workers can be explained by variations in observable qualities (e.g. education, work experience). Even if this inequalities in observable dimensions managed to be eliminated by some policy interventions, inequalities between migrants and non-migrants may remain. This unexplained portion can be interpreted as discrimination. I use twofold decomposition and the general equation for the decomposition is used:

$$\ln(Earnings_{ij}{}^{\mathrm{NM}}) - \ln(Earnings_{ij}{}^{\mathrm{M}}) = (\bar{X}^{\mathrm{NM}} - \bar{X}^{\mathrm{M}})\hat{\beta}^{\mathrm{NM}} + \bar{X}^{\mathrm{M}}(\hat{\beta}^{\mathrm{NM}} - \hat{\beta}^{\mathrm{M}}) \qquad (3.3)$$

Superscript N stands for migrants and NM for non-migrants. The second term, $\bar{X}^{M}(\hat{\beta}^{NM} - \hat{\beta}^{M})$, is unexplained wage difference by the regression model and this term represents the rate of return differences between migrants and non-migrants.

3.4. Estimation Results

Table 3.4. Mincer Wage Regression (OLS Estimates)

VARIABLES	(1) Log of monthly wages	(2)	(3) Log of bonus	(4)	(5) Log of total income	(6)
= 1 if migrant	-0.006	-0.040**	-0.380*	-0.961***	-0.011	-0.051***
	(0.013)	(0.020)	(0.230)	(0.295)	(0.013)	(0.020)
Age	0.015***	0.015***	-0.069	-0.092	0.014***	0.014***
	(0.005)	(0.005)	(0.072)	(0.071)	(0.005)	(0.005)
Age squared	-0.000***	-0.000***	0.001	0.002	-0.000***	-0.000***
	(0.000)	(0.000)	(0.001)	(0.001)	(0.000)	(0.000)
= 1 if female	-0.050***	-0.050***	-0.142	-0.174	-0.054***	-0.054***
	(0.016)	(0.016)	(0.293)	(0.287)	(0.016)	(0.016)
Years of schooling	0.004***	0.005***	0.065**	0.070**	0.005***	0.005***
	(0.002)	(0.002)	(0.031)	(0.031)	(0.002)	(0.002)
Years of working experience	-0.001	-0.001	0.505***	0.527***	0.002	0.001
	(0.009)	(0.008)	(0.159)	(0.161)	(0.008)	(0.008)
Years of working experience	0.003***	0.003***	-0.033***	-0.033***	0.003***	0.003***
squared	(0.001)	(0.001)	(0.011)	(0.011)	(0.001)	(0.001)
= 1 if married	0.018	0.016	0.052	-0.006	0.020	0.018
	(0.014)	(0.014)	(0.225)	(0.222)	(0.013)	(0.013)
= 1 if Oromo ethnic group	0.027	0.028	-0.340	-0.276	0.023	0.024
	(0.021)	(0.021)	(0.363)	(0.369)	(0.021)	(0.021)
= 1 if new worker less than	0.006	0.005	-2.240***	-2.272***	-0.019	-0.021
one year	(0.014)	(0.014)	(0.333)	(0.335)	(0.014)	(0.014)
= 1 if have social links		-0.045**		-1.395***		-0.057***
		(0.022)		(0.289)		(0.021)
Interaction term of migrant		0.064**		1.191***		0.077***
and social links		(0.026)		(0.410)		(0.025)
Constant	6.597***	6.627***	2.955**	4.270***	9.094***	9.133***
	(0.082)	(0.085)	(1.325)	(1.303)	(0.083)	(0.085)
Observations	710	710	566	566	710	710
R-squared	0.494	0.499	0.301	0.325	0.509	0.516

Note: 1. Robust standard errors in parentheses *** p<0.01, ** p<0.05, * p<0.1. 2. Each regression includes farm fixed effects. 3. Social link variable means whether a worker heard job openings through friends and relatives.

Table 3.4 presents the estimation results of the simple Mincer wage regression and the extended equation. The log of monthly wages and yearly bonus and total income (The sum of wages and bonus) are used as dependent variables. Migrant variable negatively relates with workers' wages. Earnings are significantly different between migrants and non-migrants and migrants are likely to be paid less.

Columns (2), (4), and (6) of Table 3.4 show a negative relationship with social link and workers' earnings, which is consistent with the existing empirical literature (Mano et al., 2011). Interestingly, the coefficient of the interaction term of migrant and social link dummy variable is significantly positive, suggesting that the negative effect of migrant variable is mitigated if migrant worker heard job opening through their personal networks. Note that, correlation coefficient between migrant and social link is less than 0.05, which is very low.

Table 3.5. The Blinder–Oaxaca Decomposition (twofold decomposition)

Dependent Variable: Log of monthly wages	**(1) Mincer wage equation**	**(2) Mincer wage equation with Big-Five personality dimensions**
Predicted mean of wages for non-migrants	1166.89 ***	1166.89 ***
Predicted mean of wages for migrants	1053.57 ***	1053.57 ***
Difference (%)	10.76 ***	10.76 ***
Explained (%)	4.92 ***	5.60 ***
Unexplained (%)	5.57 ***	4.88 ***

Note: 1. *** p<0.01, ** p<0.05, * p<0.1.

The decomposition result from column (1) of Table 3.5 shows that the predicted average monthly wage is 1167 birr for non-migrants and 1054 for migrants, yielding around 11 percent of wage gap. Adjusting migrants' endowments levels to the levels of non-migrants would increase migrants' earnings by 4.92 percent. A gap of 5.57 percent, which is about half the wage gap remains unexplained, supporting the idea of wage discrimination against migrants. Column (2) of Table 3.5 is the decomposition result using Mincer wage equation with Big-five personality dimensions, resulting the decrease in the unexplained portion, but still a substantial discrimination exists.[13]

[13] Reweighted Oaxaca-Blinder decomposition using a Recentered Influence Functions (RIFs) as outcome

To examine differences in work performance between migrants and non-migrants, I use two different work performance variables as dependent variables (monthly average productivity, attendance) as well as worker's subjective well-being. Table 3.6 shows that migrants are more likely to be productive than non-migrants and less likely to be absent. In addition, migrants tend to have a higher level of satisfaction of their life. The negative relationship between migrant and wage (Table 3.4) and the positive relationship between migrant and work performance (Table 3.6) suggest possible differential treatment in cut flower farm in Ethiopia.

Table 3.6. Workers' Performance and Subjective Well-being (OLS and Probit Estimates)

VARIABLES	(1) Monthly average productivity	(2) =1 if take no days off from work (not missed a day)	(3) Subjective well-being
	OLS	Probit	OLS
= 1 if migrant	0.024*	0.309**	0.368**
	(0.013)	(0.129)	(0.172)
Age	0.008*	-0.007	-0.068
	(0.004)	(0.038)	(0.066)
Age squared	-0.000*	-0.000	0.001
	(0.000)	(0.001)	(0.001)
= 1 if female	-0.103***	-0.312**	0.045
	(0.016)	(0.141)	(0.228)
Years of schooling	-0.001	-0.035**	-0.012
	(0.002)	(0.015)	(0.023)
Years of working experience	0.018**	-0.236***	-0.169*
	(0.008)	(0.066)	(0.091)
Years of working experience squared	-0.001	0.014***	0.009
	(0.001)	(0.004)	(0.006)
= 1 if married	-0.001	-0.080	0.072
	(0.013)	(0.120)	(0.172)
= 1 if Oromo ethnic group	0.010	-0.054	0.009
	(0.020)	(0.160)	(0.245)
= 1 if new worker less than one year	0.021	0.532***	-0.052
	(0.018)	(0.155)	(0.237)
Constant	0.587***	0.583	7.236***
	(0.075)	(0.637)	(1.141)
Observations	677	710	710
R-squared	0.176		0.083

Note: 1. Robust standard errors in parentheses *** p<0.01, ** p<0.05, * p<0.1. 2. Each regression includes farm fixed effects.

variables are reported Table A.2. in the Appendix.

3.5. Chapter Summary

This chapter found that migrant workers are less likely to use bank savings even though they have bank accounts and tend to use ROSCAs savings more while non-migrants tend to use bank accounts. Overall, migrant workers face more vulnerable financial conditions than locals.

This chapter also found that earnings of workers are significantly different between migrant and non-migrant and migrants are likely to be paid less. However, migrant workers are more likely to be productive and less likely to be absent. The Blinder–Oaxaca decomposition results reveal that around half of wage differentials is unexplained even when non-cognitive skills are controlled, supporting the idea of differential treatment against migrants in cut flower industry in Ethiopia. Even though migrant workers may face differential treatment in compensation, higher levels of subjective wellbeing and descriptive analysis results on financial behaviors of migrants suggest that rural to urban migration would be investment for better job opportunities.

This chapter contributes to the literature on internal migrants by providing rich empirical evidence using various earnings and productivity outcome variables. This could also shed light on policies for the vulnerable young women migrants in developing countries.

4. Determinants and Predictors of Credit Constraints: Empirical Evidence from Nigeria

4.1. Introduction

For the last two decades, there have been heated debates on the impact of microfinance. Many researchers have explored the impact of microfinance programs through Randomized Control Trials (RCTs). However, there has not been so much effect on the sustained consumption gain as a result of access to microcredit and the take-up rate of microcredit program is normally 20-25 percent, which is very low (Banerjee, 2013). Existing literature also finds improved access to credit does not reveal significant effects on education and health as well. (Kaboski and Townsend, 2012; Banerjee, et al., 2015a; Banerjee, et al., 2015b). One of the reasons for those unsatisfactory results could be mistargeting. Microcredit program might be effective when credit-constrained households are well-targeted for microcredit programs. Thus, to solve low take-up problem of microcredit program, strengthening empirical methods for identifying credit-constrained households would be needed for better targeting.

Identifying and distinguishing different forms of credit constraints is important since different types of credit constraints require different policies (Boucher et al., 2009). For example, simply expanding supply of credit may not be appropriate for those who are facing binding demand-side constraints since they may not apply for credit even though they can access to credit market. Demand side constraints may occur when potential borrowers overestimate risk or transaction cost they may face. This misperception leads deserving potential borrowers to refrain from borrowing, which may adversely affect the efficiency of resource allocation. Thus, understanding what factors are associated with different forms of credit constraints which discourage potential borrowers in developing countries would be crucial. However, the existing literature is limited in providing robust evidence on the determinants of different kind of credit constraints in different settings.

Thus, this chapter has two separate purposes. Firstly, it aims to investigate the determinants of detailed non-pricing credit constraints and its effect on household welfare by using direct elicitation approach of identifying credit constraints. This is important because most of existing studies on credit constraints rely on relatively small dataset limited to agriculture context in rural area. Secondly, it aims to explore the possible strong predictors of households' credit constraints for targeting. Shortlisting the best predictors of credit constraints is different from obtaining variables that are associated with credit constraints by regression methods from the first purpose of this chapter. The selected predictors can be applicable to other datasets which lack information on credit constraints status to predict the households' credit constraints status. Since the directly elicited credit constraints variable is not often available in general household survey, the selected predictors may provide useful information for better and more appropriate targeting of microcredit program beneficiaries.

Using a nationally representative large dataset, this chapter finds that the probability of being risk rationed decreases if a household head engages in waged labor, while the probability of being transaction-cost rationed decreases when a household owns their non-farm enterprises. Loss averse is found to be positively related with each credit rationing except transaction-cost rationing. For lower income households, the probability of being risk rationed increases if a household engages in agriculture. This chapter further finds that households in the south of Nigeria are more likely to face quantity rationing.

For identifying the best predictors of the household's credit constraints status, one of the supervised Machine Learning (ML) approaches, the Least Absolute Shrinkage and Selection Operator (LASSO) with 10-fold cross-validation is employed since it outperforms Ordinary Least Squares (OLS) regression in terms of out-of-sample prediction. This chapter finds that credit constraints status of household can be predicted by not only households' credit market participation and but also community level data such as climatology, terrain, and crop season parameters. For improving the accuracy of targeted microcredit programs, these selected predictors can be applicable to other datasets which lack information on household's credit constraints status.

The rest of this chapter proceeds as follows. Section 4.2 provides existing literature on identification of credit constraints. Section 4.3 describes the data and estimation strategy. Section 4.4 discusses the estimation results and Section 4.5 concludes this chapter with remarks on policy implications.

4.2. Credit Constraints

Previous studies have used mainly two ways of identifying credit constraints empirically. The first one is the indirect method, which identifies the presence of credit constraints from violations of the assumptions of life-cycle permanent income hypothesis (LC/PIH) (Diagne et al, 2000). The idea is that in the absence of credit constraints, transitory income fluctuations should not affect consumption because households can smooth their consumption by using credit. For example, Zeldes (1989) exogenously splits the sample of households from the United States on the basis of an income to wealth ratio and investigated the violation of LC/PIH occurs in credit-constrained households. Garcia et al. (1997) employs an endogenous split approach with unobserved regime to control the endogenous problem resulting from an exogenous split. However, households with credit constraints may be able to smooth their consumption with precautionary saving, thus, it is hard to say that the violation of LC/PIH can be the evidence for credit constraints (Diagne et al, 2000).

The second method of identifying credit constraints is the direct elicitation approach, which is based on survey responses on households' perceptions of their credit-constrained status (Diagne et al, 2000; Simtowe et al., 2008). This approach was first applied by Jappelli

(1990). To investigate whether a household is credit-constrained, households are directly asked about their credit history such as loan applications and rejections, and nonborrowers' perceptions towards credit market. Jappelli (1990) classifies households in the United States as credit-constrained if they had a loan application rejected or did not attempt to apply for credit because they believed to be rejected. Theoretical justification of the direct elicitation approach lies in the extended version of LC/PIH which explicitly allows for the probability of being credit constrained, and it identifies whether households' credit demand exceeds the credit supply available (Diagne et al, 2000; Simtowe et al., 2008).

Boucher et al. (2009) extended this direct elicitation approach and categorize non-pricing credit constraints into quantity rationing, risk rationing and transaction-cost rationing. Quantity rationing is supply side constraints, while risk and transaction-cost rationing can be regarded as demand side constraints. According to their classification, quantity rationing occurs when a borrower has effective demand for credit, but they face binding credit limit that is set by a lender. On the other hand, for demand-side constrained borrowers, their effective demand is reduced by risk or transaction costs they perceive, which leads them to voluntarily withdraw their loan application. For example, potential borrowers refrain from applying for a loan if they overestimate risk or transactions costs they may face. Boucher et al. (2009) empirically find that ignoring demand-side credit constraints would result in underestimation of credit constraints using Peruvian agriculture data.

Although there is rich literature on credit constraints, most of empirical studies conducted in developing countries make use of relatively small sample and are limited to agriculture context in rural area. Thus, this chapter attempts to provide more comprehensive analyses on the issue of credit constraints, using a nationally representative large dataset, which is not limited to only agriculture in rural area. Non-agricultural activities such as non-farm enterprise and waged employment are considered, and observations in urban area are also included to investigate which policy should be implemented to relax credit constraints in different settings in developing countries. This kind of analysis was not available in the existing literature as they rely solely on data from agricultural in rural area. Also, with using supervised machine learning approach, this chapter explores which variables are strong predictors of households' credit constraints status since the directly elicited credit constraints variable is not often available in general household survey. To my knowledge, this is first attempt in literature on credit constraints.

4.3. Data and Estimation Strategy

4.3.1. Data

To investigate the determinants and predictors of different forms of credit constraints, this chapter uses the General Household Survey (GHS) 2015/2016, the third wave of the nationally-representative household survey in Nigeria conducted by the Nigeria National Bureau of Statistics with the support of the World Bank. This dataset includes information on consumption, socio-economic characteristics, and credit application history of around 4,600 households in Nigeria. This dataset also includes various community level information and non-cognitive traits of household head, which are not often available in large scale nationally representative household survey. This allows me to conduct comprehensive analyses on the issue of credit constraints as well as machine learning approach to select the strong predictors of credit constraints status. Since community level data is important in my analysis, I dropped 368 household observations that community level data is missing.

Direct elicitation approach is used to identify different forms of credit constraints. Following Boucher et al. (2009), non-pricing rationing is categorized into risk rationing, transaction-cost rationing, and quantity rationing based on credit application history and the nonborrowers' perception of credit market. Distinguishing different types of credit constraints is important for designing and targeting effective policy since different forms of credit constraints may require different policy interventions. Overall credit constraints include all three types of credit rationing. Figure 4.1 shows how non-pricing rationing is categorized and Table 4.1 presents the number of households which face each credit rationing. Based on direct elicitation approach, around 20 percent of households are facing any forms of credit rationing. To be specific, risk-constrained, transaction-cost constrained, and quantity constrained households account for around 7 percent, 1 percent, and 13 percent respectively.

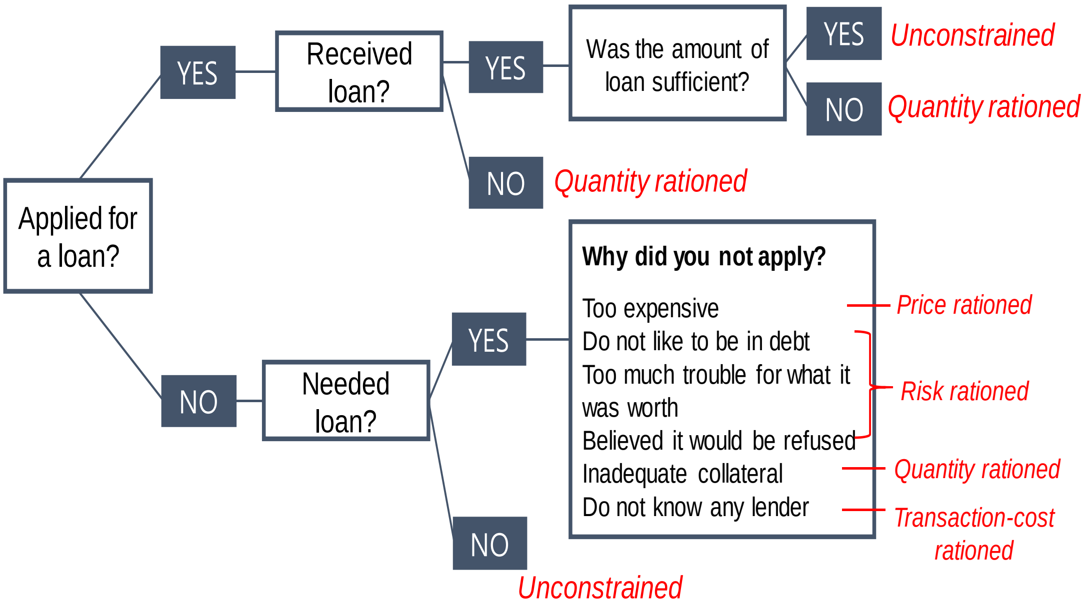

Figure 4.1. Direct Elicitation Approach
Source: Nigeria National Bureau of Statistics 2015

Table 4.1. Non-Pricing Credit Rationing Categories (among a total of 4,232 HHs)

Non-pricing rationing	Number of HHs	%
Risk rationed	279	6.6
Transaction-cost rationed	38	0.9
Quantity rationed	539	12.7
Overall credit constraints	856	20.2

Total consumption expenditure is calculated as a sum of all food, non-food expenditure, education, and total imputed rent. Regarding non-cognitive traits of household head, trust variable takes a value of one if household head answered that most people can be trusted. Loss aversion takes a value of one if household head chose the answer that he or she prefers to choose safe investments that keep their money secure rather than a small chance of losing money for the sake of potentially high profit.

Regarding risk preference, a household head is asked to choose between two options.

Q1. Option1: You receive 200 Naria for sure. (go to **Q2**)

Option2: I flip a 1 Naira Coin. If it shows Sir Herbert Macaulay, you get 600 Naira. If it's the coat of arms, you get 50 Naira. (go to **Q3**)

Q2. Option1: You receive 200 Naria for sure.

Option2: I flip a 1 Naira Coin. If it shows Sir Herbert Macaulay, you get 800 Naira. If it's the coat of arms, you get 50 Naira.

Q3. Option1: You receive 200 Naria for sure.

Option2: I flip a 1 Naira Coin. If it shows Sir Herbert Macaulay, you get 400 Naira. If it's the coat of arms, you get 50 Naira.

The respondent who choose Option 1 for Q1 and Q2 is categorized as the most risk averse. By combining answers from Q1 to Q3, the respondent who choose Option 1 for Q1 and Option 2 for Q2 is regarded as highly risk averse and the respondent who choose Option 2 for Q1 and Option1 for Q3 is classified as moderate risk averse. The respondent who choose Option 2 for Q1 and Option2 for Q3 is regarded as risk-taker.

To classify respondents with present-biased time-inconsistent preferences, a household head is asked to choose between two options for the following questions.

Q1. Option1: You receive 1000 Naira today. (go to **Q2**)

Option2: You receive 2000 Naira in 1 month. (go to **Q3**)

Q2. Option1: You receive 1000 Naira today. (go to **Q4**)

Option2: You receive 2500 Naira in 1 month. (go to **Q4**)

Q3. Option1: You receive 1000 Naira today.

Option2: You receive 1500 Naira in 1 month.

Q4. Option1: You receive 1000 Naira in 1 year. (go to **Q5**)

Option2: You receive 2000 Naira in 1 year and 1 month. (go to **Q6**)

Q5. Option1: You receive 1000 Naira in 1 year.

Option2: You receive 2500 Naira in 1 year and 1 month.

Q6. Option1: You receive 1000 Naira in 1 year.

Option2: You receive 1500 Naira in 1 year and 1 month.

By combining answers from Q1 to Q3, the respondent who choose Option 1 for Q1 and Option 1 for Q2 is regarded as the most impatient. The respondent who choose Option 1 for Q1 and Option 2 for Q2 is considered as highly impatient and the respondent who choose Option 2 for Q1 and Option 1 for Q3 is categorized as moderately impatient. The respondent who choose Option 2 for Q1 and Option 2 for Q3 is regarded as patient. From Q4 to Q6, the similar questions are asked by changing the payment timing to one year later. The respondent who is less patient for current trade-offs than future trade-offs is considered as present-biased time inconsistent.

Table 4.2 reports the descriptive statistics. Around 77 percent of households in Nigeria have their own house. The average age of household heads is around 53 years old, and about 20 percent of household heads are female. The average years of schooling of household heads is about 8 years. The average number of adult males and females, which can be regarded as human assets, is 1.7 and 1.8 respectively (The average number of people in household is around 5.8). Regarding non-cognitive traits of household heads, around 32 percent of household heads show present-biased time inconsistent preference and about 18 percent of household heads trust other people in general. As for risk preference, the average of risk preference index is around 3.5. To be specific, about 79 percent of household heads show the most risk averse. Highly risk averse, moderate risk averse, and risk-taker household heads account for 4 percent, 5 percent, and 11 percent respectively. Approximately 70 percent of household heads show risk aversion, which suggests that majority of them are reluctant to take a small chance of losing money for the sake of potentially high profit. Around 70 percent of households reside in rural area. Regarding occupations household heads engaging in, around 14 percent of household heads are engaging in waged employment. About 47 percent of household own their non-farm enterprises and about 65 percent of household also engage in agricultural activities.

Table 4.2. Descriptive Statistics (N=4,232)

Variable	Mean	SD	Min	Max
Total consumption expenditure per capita (Naria)	154726	165175	12459	3802534
Total assets (Naria)	261716	2919740	0	171000000
Total livestock value (Naria)	156930	845096	0	27900000
House ownership	0.77	0.42	0	1
Age of household head	53.29	14.62	15	103
Female household head	0.19	0.4	0	1
Years of schooling of household head	7.97	6.79	0	17
Number of adult males (over 15 years old)	1.67	1.26	0	12
Number of adult females (over 15 years old)	1.8	1.26	0	11
Present bias preference	0.32	0.47	0	1
Risk preference (4: the most risk averse)	3.52	1.01	1	4
Trust	0.18	0.38	0	1
Loss aversion	0.69	0.46	0	1
Rural	0.71	0.46	0	1
Enterprises	0.47	0.5	0	1
Waged employment	0.14	0.35	0	1
Agriculture	0.65	0.48	0	1

Table 4.3 presents the descriptive statistics by different forms of credit constrained households. Households are categorized as risk rationed, transaction-cost rationed, quantity rationed household, and non-constrained household. T-test is performed using all the other

groups as a reference group. Compared to other groups, risk-constrained households tend to have higher loss aversion. For transaction-cost-constrained households, however, they possess lower loss aversion compared to other groups. They tend to live in rural area and engage in agricultural activities rather than non-farm enterprise or waged labor. Quantity-constrained households have less livestock while they tend to have higher education level compared to other groups. Quantity-constrained households also possess higher loss aversion and tend to reside in urban area. They are more likely engage in non-farm enterprises or waged employment. Households that do not face any forms of credit constraints show lower total consumption expenditure per capita and higher total livestock value. The non-constrained household head tend to be female and older, and they have lower education level compared to other constrained groups. They also tend to have present-biased time inconsistency, while they are less likely to show loss aversion.

Table 4.3. Descriptive Statistics by Different Credit Constraints Groups

Variable	Risk rationed	Transaction-cost rationed	Quantity rationed	Non-constraints
Total consumption expenditure per capita (Naria)	169,754	178,494	158,663	152,587 *
Total assets (Naria)	199,304	100,643	299,576	262,642
Total livestock value (Naria)	112,228	326,331	70,632 **	172,495 **
House ownership	0.77	0.84	0.7 ***	0.78 ***
Age of Household head	52	50	52 **	53 ***
Female Household head	0.17	0.15	0.15 ***	0.2 ***
Years of schooling of HH head	8.04	6.71	8.83 ***	7.84 **
Number of adult males (over 15 years old)	1.73	1.71	1.72	1.65
Number of adult females (over 15 years old)	1.76	1.87	1.77	1.81
Present bias preference	0.29	0.29	0.3	0.33 *
Risk preference (4: the most risk averse)	3.55	3.55	3.52	3.52
Trust	0.15	0.11	0.18	0.18
Loss aversion	0.75 **	0.55 *	0.73 *	0.68 **
Rural	0.7	0.89 **	0.66 **	0.71
Enterprises	0.49	0.32 **	0.51 *	0.47
Waged employment	0.11	0.03 **	0.17 **	0.14
Agriculture	0.68	0.79 *	0.62	0.65
Observations	279	38	539	3,376

Note: 1. *** p<0.01, ** p<0.05, * p<0.1. 2. T-test is performed using all the other groups as a reference group.

4.3.2. Estimation Strategy

Firstly, to explore the determinants of each credit rationing, probit model is used and the average marginal effect is calculated. Secondly, the effect of credit constraints on household welfare is estimated to investigate whether directly elicited credit constraints variables are correlated with household welfare. Since one of the objectives of this chapter is to investigate the possible strong predictors of households' credit constraints, checking whether there is association between directly elicited credit constraints and household's welfare is worth exploring. The following model is estimated using instrument variables regression to address potential endogeneity of credit constraints variable.

$$Y_{ij} = \beta_0 + \beta_1 \mathrm{C}_{ij} + \beta_2 X_{ij} + \mu_{\mathrm{k}} + \varepsilon_{ij} \qquad (4.1)$$

The dependent variable, Y_{ij} is household welfare proxied by total consumption expenditure level for a household i in local government area j. The binary variable C_{ij} takes value one if a household is credit constrained. X_{ij} is household characteristics such as physical and human assets, household heads' characteristics and non-cognitive traits. μ_{k} is a zone fixed effect, and ε_{ij} is an error term.

To control possible endogeneity of credit constraints variable, two instrumental variables for the household's credit constraints status are used. Whether the community where a household is residing in has a formal bank and average enhanced vegetation index which measures health vegetation in the past five years may relate with credit constraints status but may not necessarily directly associate with consumption expenditure level of household.

Thirdly, to identify strong predictors of the household's credit constraints status, Supervised Machine Learning approach is employed. While many econometric methods are used for parameter estimation (βˆ) with focus on the relationship between outcome variables and independent variables, Machine Learning approach is used for the purpose of prediction ($\hat{y}$). Supervised Machine Learning algorithms finds functions that predict well out-of-sample by fitting complex and flexible functional forms to the data while managing overfitting by regularization and empirical tuning (Mullainathan and Spiess, 2017). Because of its advantages, Supervised Machine Learning algorithms has recently gained popularity in predicting poverty and shortlisting variables for targeting when outcome variables (such as income levels) are lacking. For examples, Blumenstock et al. (2015) predict poverty in Rwanda using mobile phone data and Jean et al. (2016) predict local poverty level through satellite imagery. Kshirsagar et al. (2017) and Knippenberg et al. (2019) employ Machine Learning algorithms to predict poverty status and food insecurity respectively when information on outcome data is missing.

One of the Supervised Machine Learning approaches, the Least Absolute Shrinkage and Selection Operator (LASSO) with 10-fold cross-validation is used for producing predictions of the household's credit constraints status, fitting a function to a training set, which can make predictions for new dataset which does not contain information on credit constraints status.

$$\text{Minimize:} \quad \frac{1}{n}\sum_{i=1}^{n}(y_i - \acute{x}_i\beta)^2 + \lambda\sum_{j=1}^{p}|\beta_j| \tag{4.2}$$

LASSO is a linear regression which penalizes additional parameters by including the penalization term λ through cross-validation. In cross-validation, the data is repeatedly separated into training and validation data. The data is repeatedly set into 10-fold of approximately equal size and treat one fold as the validation data and others as the training data. Penalization term λ is chosen to optimize predictive performance.

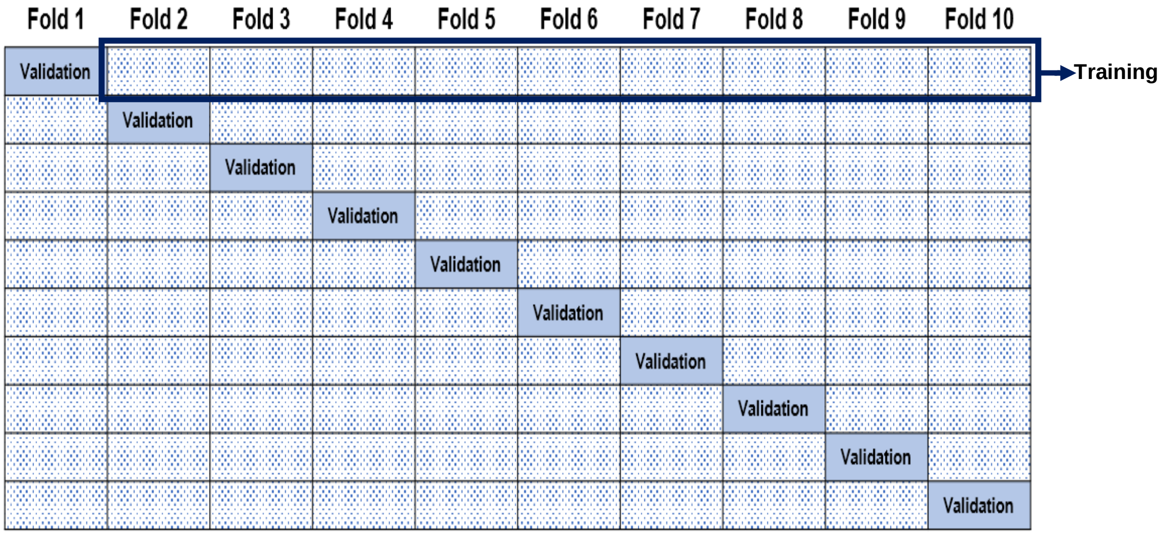

Figure 4.2. 10-Fold Cross-Validation

4.4. Estimation Results

4.4.1. Determinants of Credit Constraints

Table 4.4 reports the average marginal effect of probit estimates. The probability of being risk rationed increases with households with more livestock assets and loss averse. (column (1) of Table 4.4). Risk-aversion increases the probability of being risk rationed but is not statistically significant. Livestock assets can be distinguished from other assets in that they are investment assets since households can produce by-product from their livestock, and households also face risks of losing their livestock for various causes such as an infectious disease. One explanation may be that this feature of livestock assets may increase the perceived overall risk, and this may lead households to reluctant to borrow money from credit market, but it needs more investigation. The probability of being risk rationed increases with loss aversion, suggesting if households are reluctant to bear the minimum risk of losing their collaterals, they are less likely to participate in credit market. If households have fewer adult females and reside in urban area, they are more likely to be risk rationed.

Column (2) of Table 4.4 shows that the probability of being transaction-cost rationed increases with households with higher livestock value and younger household head. If household head are less loss averse, the probability of being transaction-cost rationed decreases. In other words, if household with a willingness to take a risk of losing their money for the sake of higher potential profits tends to be transaction-cost rationed. One possible explanation is that a household with less loss aversion may seek high-risk and high-return opportunities, which make them hard to find potential lenders. The probability of being quantity rationed and overall credit-constrained (including all three types of rationing; risk, transaction-costs, and quantity raining) increase with household heads who are male, younger, and loss averse (Columns (3) and (4) of Table 4.4).

Regarding the relationship between each rationing category and occupations households engage in, if a household head engages in waged employment, the probability of being risk rationed decreases, while the probability of being transaction-cost rationed decreases with households owning non-farm enterprises. Waged employment may ensure financial stability compared to non-farm enterprises or agricultural sector, which may increase households' capacity of taking risks.

Table 4.4. Determinants of Credit Constraints (Marginal Effects of Probit Estimates)

Variables	(1) Risk rationed	(2) Transaction-cost rationed	(3) Quantity rationed	(4) Overall credit constraints
Log of total assets	0.005	-0.001	-0.003	0.000
	(0.003)	(0.001)	(0.004)	(0.005)
Log of total livestock value	0.002 **	0.0004 *	-0.001	0.001
	(0.001)	(0.000)	(0.001)	(0.001)
House ownership	-0.008	0.001	-0.018	-0.024
	(0.010)	(0.004)	(0.015)	(0.018)
Age of HH head	-0.000	-0.0002 *	-0.001 **	-0.001 ***
	(0.000)	(0.000)	(0.000)	(0.000)
Female HH head	-0.004	-0.003	-0.049 ***	-0.056 ***
	(0.010)	(0.005)	(0.017)	(0.020)
Years of schooling of HH head	-0.000	-0.000	0.000	-0.000
	(0.001)	(0.000)	(0.001)	(0.001)
Number of adult males	0.001	0.000	0.007	0.008
	(0.003)	(0.001)	(0.004)	(0.006)
Number of adult females	-0.006 *	0.001	0.003	-0.002
	(0.003)	(0.001)	(0.005)	(0.006)
Present bias preference	-0.010	-0.002	-0.012	-0.023
	(0.010)	(0.003)	(0.012)	(0.016)
Risk preference	0.002	0.001	-0.002	0.000
	(0.004)	(0.001)	(0.005)	(0.006)
Trust	-0.010	-0.006	0.011	-0.007
	(0.011)	(0.004)	(0.013)	(0.017)
Loss aversion	0.025 ***	-0.006 **	0.028 **	0.045 ***
	(0.009)	(0.003)	(0.013)	(0.015)
Rural	-0.026 **	0.004	-0.008	-0.031
	(0.011)	(0.005)	(0.016)	(0.019)
Enterprises	0.003	-0.006 **	0.015	0.011
	(0.008)	(0.003)	(0.011)	(0.014)
Waged employment	-0.027 **	-0.013	0.012	-0.020
	(0.013)	(0.008)	(0.016)	(0.020)
Agriculture	0.006	-0.000	0.016	0.021
	(0.010)	(0.004)	(0.015)	(0.019)
Observations	4,232	4,232	4,232	4,232

Note: 1. Standard errors adjusted for 404 clusters at local government area in parenthesis *** p<0.01, ** p<0.05, * p<0.1.
2. Each regression includes zone fixed effects.

Table 4.5. Determinants of Credit Constraints Using Lower Income Sub-Sample (Marginal Effects of Probit Estimates)

Variables	(1) Risk rationed	(2) Transaction-cost rationed	(3) Quantity rationed	(4) Overall credit constraints
Log of total assets	0.005	-0.002	0.005	0.007
	(0.004)	(0.001)	(0.006)	(0.007)
Log of total livestock value	0.001	0.000	-0.002 *	-0.001
	(0.001)	(0.000)	(0.001)	(0.002)
House ownership	-0.023 *	0.005	-0.024	-0.043 **
	(0.012)	(0.006)	(0.018)	(0.022)
Age of HH head	-0.000	-0.000	-0.001	-0.001 *
	(0.000)	(0.000)	(0.001)	(0.001)
Female HH head	0.012	0.001	-0.032	-0.019
	(0.014)	(0.007)	(0.022)	(0.025)
Years of schooling of HH head	-0.000	-0.000	-0.000	-0.000
	(0.001)	(0.000)	(0.001)	(0.001)
Number of adult males	0.001	0.000	0.006	0.007
	(0.004)	(0.001)	(0.005)	(0.007)
Number of adult females	-0.004	0.001	0.002	-0.000
	(0.004)	(0.002)	(0.006)	(0.007)
Present bias preference	-0.006	-0.008 *	-0.009	-0.021
	(0.011)	(0.004)	(0.014)	(0.019)
Risk preference	0.001	0.000	-0.007	-0.006
	(0.005)	(0.002)	(0.006)	(0.007)
Trust	-0.023 *	-0.006	0.025	-0.006
	(0.012)	(0.005)	(0.015)	(0.020)
Loss aversion	0.011	-0.011 ***	0.040 ***	0.037 **
	(0.010)	(0.004)	(0.015)	(0.018)
Rural	-0.021	0.000	0.015	-0.003
	(0.014)	(0.006)	(0.021)	(0.032)
Enterprises	-0.001	-0.009 **	0.016	0.006
	(0.009)	(0.004)	(0.014)	(0.016)
Waged employment	-0.020	-0.013	-0.003	-0.028
	(0.017)	(0.010)	(0.020)	(0.025)
Agriculture	0.027 *	-0.005	0.002	0.020
	(0.014)	(0.006)	(0.019)	(0.023)
Observations	2,908	2,908	2,908	2,908

Note: 1. Standard errors adjusted for 393 clusters at local government area in parenthesis *** p<0.01, ** p<0.05, * p<0.1.
2. Each regression includes zone fixed effects.

Since this dataset is national-wide data, income variation is large. For focusing more on vulnerable income groups, the same analyses for the lower income group subsample is conducted to investigate the determinants of each credit rationing among lower income group. The threshold of the average income of household proxied by total consumption expenditure per capita is used to classify lower income group. Table 4.5 shows that among the lower income group, the probability of being risk rationed decreases when households own their house and household heads tend to trust others. The probability of being transaction cost rationed decreases with household heads with present-biased preference and loss aversion. Loss aversion is positively associated with the probability of being quantity rationed and overall credit constrained. When it comes to the lower income group, the probability of being risk rationed increases if a household engages in agriculture, while households with non-farm enterprises decreases the probability of transaction cost rationing.

As there is huge economic and social imbalance between the north and the south of Nigeria and the south is much richer than the north, I plot the probability of being credit constrained for different reasons by geographical regions in Figure 4.3. It shows that the propensity of quantity rationing in rural area are lower than those in urban area, while the propensity of transaction-cost rationing in rural area is higher than in urban area. The gap of the propensity of quantity rationing is much bigger between south region and north region of Nigeria, which suggests that borrowers in the south region of Nigeria demand more credit than credit limit which lender's willingness to offer. Households residing in the richer south of Nigeria may have more investment opportunities, resulting in high chance of facing quantity rationing. I also plot the probability of different types of credit rationing by regions by the quartile of assets in Figure 4.4. It also shows that the propensity of quantity rationing higher as asset quartiles increase in the south of Nigeria.

Table 4.6 and 4.7 presents the average marginal effect of probit estimate using North/South and Urban/Rural subsamples. Risk rationing and transaction-cost rationing are combined to demand-side constraints in this sub-sample analysis since the event of transaction-cost rationing in sub-sample is rare and this leads to failure of estimation. Columns (2) and (4) of Table 4.7 shows that households in the south of Nigeria are more likely to face quantity rationing. Interestingly, trust decreases the probability of facing demand-side constraints, while trust increases the probability of facing quantity rationing (i.e. supply-side constraints) both in the north of Nigeria and rural area.

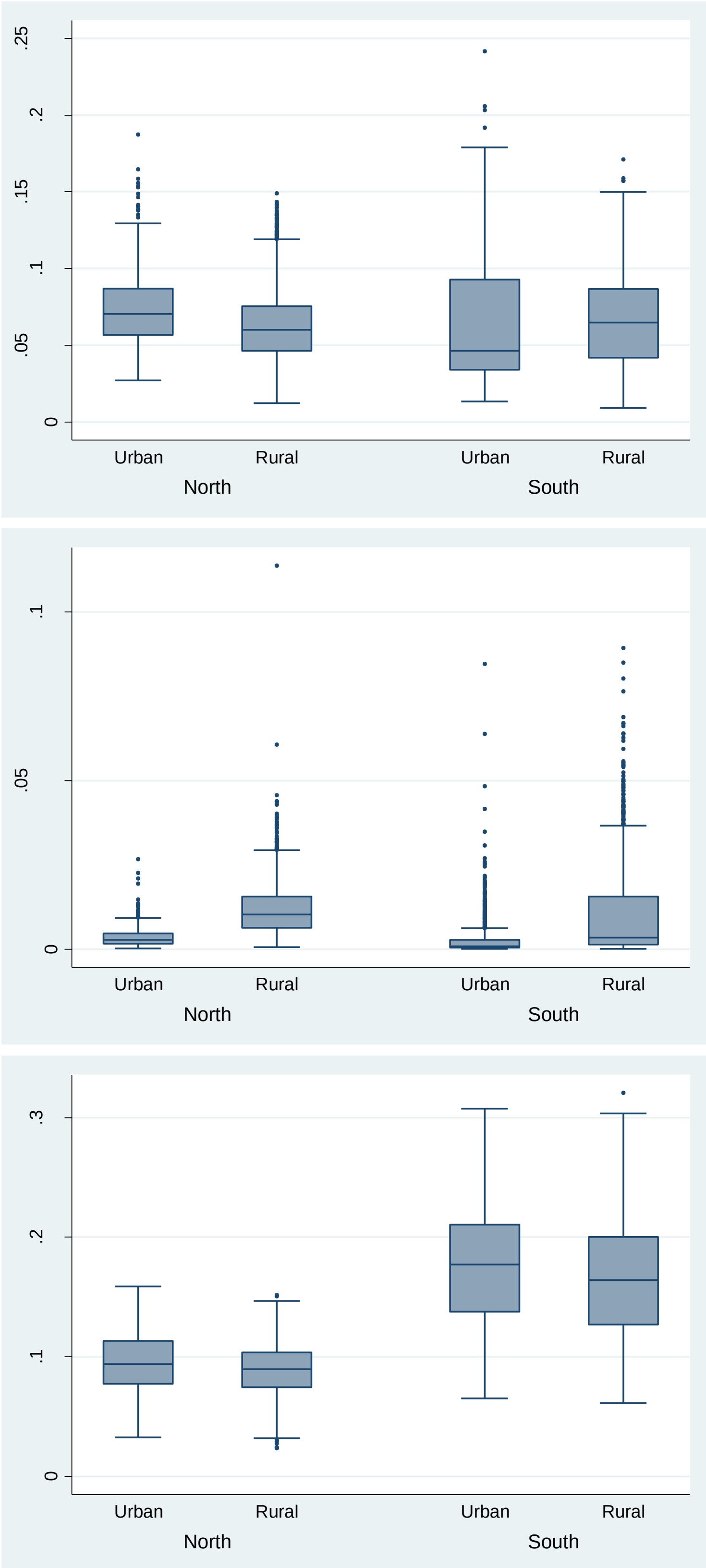

Figure 4.3. Propensity of Each Rationing Category by Regions

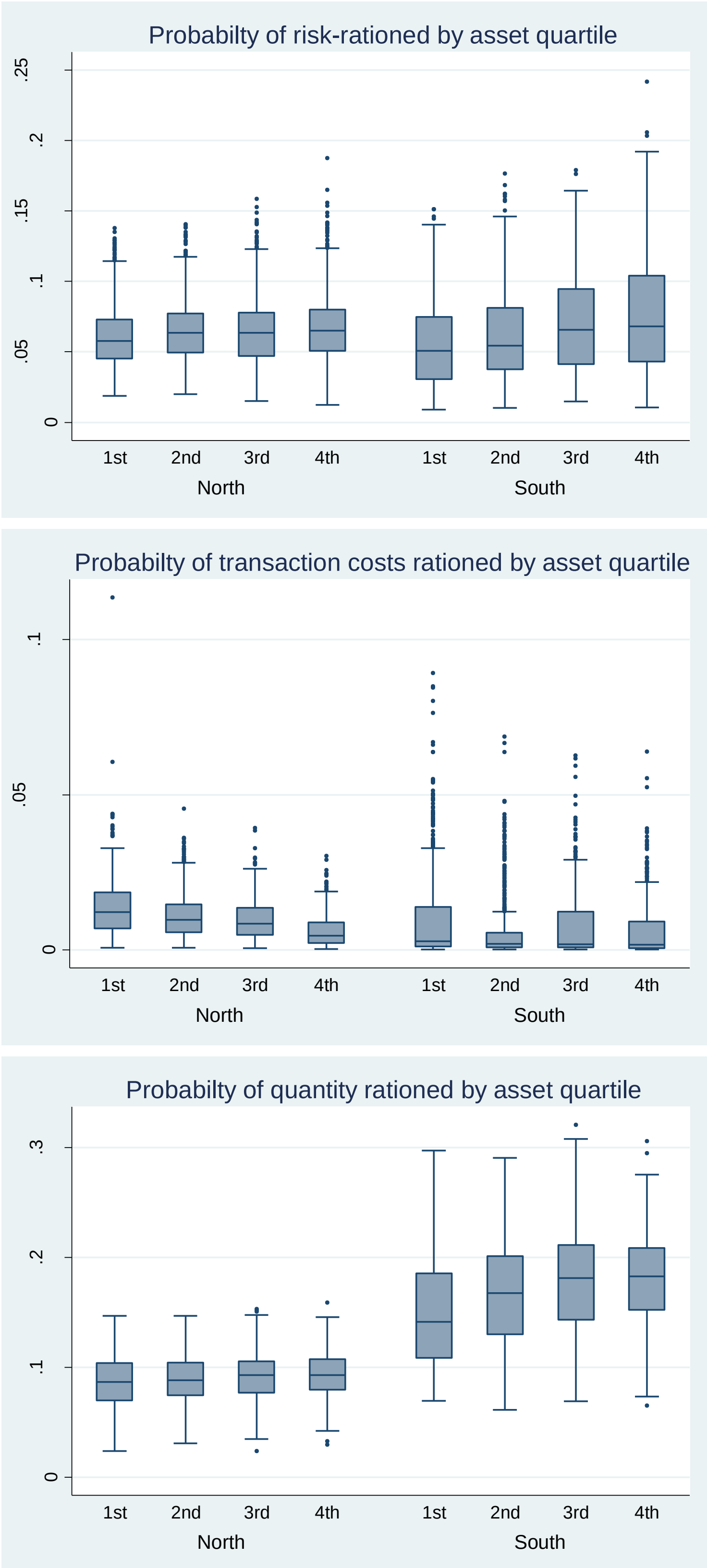

Figure 4.4. Propensity of Each Rationing Category by Asset Quartiles

Table 4.6. Determinants of Credit Constraints Using North/South Sub-Samples (Marginal Effects of Probit Estimates)

Variables	(1)	(2)	(3)	(4)
	North		South	
	Demand-side constraints	Quantity rationed	Demand-side constraints	Quantity rationed
Log of total assets	-0.008	-0.005	0.013 ***	-0.004
	(0.005)	(0.006)	(0.005)	(0.006)
Log of total livestock value	0.002 **	0.000	0.001	-0.003
	(0.001)	(0.001)	(0.001)	(0.002)
House ownership	-0.002	-0.038 *	-0.006	-0.009
	(0.019)	(0.020)	(0.013)	(0.022)
Age of HH head	-0.001 **	-0.001	-0.000	-0.001 *
	(0.000)	(0.001)	(0.000)	(0.001)
Female HH head	0.033 *	-0.034	-0.018	-0.049 **
	(0.019)	(0.027)	(0.014)	(0.023)
Years of schooling of HH head	0.000	-0.001	-0.001	0.004 **
	(0.001)	(0.001)	(0.001)	(0.002)
Number of adult males	0.002	-0.001	0.006	0.017 **
	(0.005)	(0.005)	(0.005)	(0.008)
Number of adult females	0.001	0.003	-0.006	0.004
	(0.005)	(0.005)	(0.005)	(0.008)
Present bias preference	-0.009	-0.037 **	-0.012	0.019
	(0.014)	(0.015)	(0.014)	(0.019)
Risk preference	0.006	-0.002	-0.003	-0.003
	(0.006)	(0.006)	(0.005)	(0.008)
Trust	-0.041 **	0.027 *	0.012	-0.013
	(0.017)	(0.014)	(0.017)	(0.022)
Loss aversion	0.008	0.045 ***	0.013	0.006
	(0.013)	(0.016)	(0.014)	(0.020)
Enterprises	-0.013	0.008	0.011	0.026
	(0.012)	(0.014)	(0.012)	(0.018)
Waged employment	-0.005	0.027	-0.054 **	0.000
	(0.018)	(0.018)	(0.022)	(0.027)
Agriculture	0.012	0.001	0.007	0.036
	(0.018)	(0.020)	(0.015)	(0.023)
Rural	-0.027	-0.002	0.010	-0.007
	(0.019)	(0.025)	(0.016)	(0.022)
Observations	2,221	2,221	2,011	2,011

Note: 1. Standard errors adjusted for clusters at local government area in parenthesis *** p<0.01, ** p<0.05, * p<0.1.

Table 4.7. Determinants of Credit Constraints Using Urban/Rural Sub-Samples (Marginal Effects of Probit Estimates)

Variables	(1)	(2)	(3)	(4)
	Urban		Rural	
	Demand-side constraints	Quantity rationed	Demand-side constraints	Quantity rationed
Log of total assets	0.000	-0.009	0.006	0.000
	(0.006)	(0.008)	(0.005)	(0.005)
Log of total livestock value	0.002	-0.001	0.002 *	-0.001
	(0.002)	(0.003)	(0.001)	(0.001)
House ownership	-0.000	-0.033	-0.001	-0.006
	(0.015)	(0.025)	(0.016)	(0.019)
Age of HH head	-0.000	-0.002 **	-0.001 *	-0.001 *
	(0.001)	(0.001)	(0.000)	(0.000)
Female HH head	0.014	-0.046	-0.012	-0.048 **
	(0.021)	(0.031)	(0.016)	(0.020)
Years of schooling of HH head	0.001	0.001	-0.000	-0.000
	(0.002)	(0.002)	(0.001)	(0.001)
Number of adult males	0.000	0.020 **	0.004	0.002
	(0.006)	(0.009)	(0.004)	(0.005)
Number of adult females	0.001	0.004	-0.006	0.001
	(0.005)	(0.009)	(0.005)	(0.005)
Present bias preference	0.010	0.013	-0.022 *	-0.019
	(0.017)	(0.021)	(0.012)	(0.014)
Risk preference	-0.003	0.010	0.002	-0.007
	(0.007)	(0.011)	(0.005)	(0.005)
Trust	0.019	-0.033	-0.032 **	0.026 *
	(0.019)	(0.027)	(0.015)	(0.014)
Loss aversion	0.037 **	0.028	-0.000	0.024 *
	(0.017)	(0.025)	(0.011)	(0.014)
Enterprises	0.022	0.014	-0.015	0.012
	(0.017)	(0.021)	(0.010)	(0.013)
Waged employment	-0.014	-0.013	-0.045 **	0.029
	(0.021)	(0.029)	(0.019)	(0.019)
Agriculture	0.007	0.029	0.006	0.013
	(0.020)	(0.024)	(0.014)	(0.019)
North	0.007	-0.097 ***	-0.013	-0.087 ***
	(0.018)	(0.029)	(0.013)	(0.018)
Observations	1,243	1,243	2,989	2,989

Note: 1. Standard errors adjusted for clusters at local government area in parenthesis *** p<0.01, ** p<0.05, * p<0.1.

4.4.2. The Effects of Credit Constraints on Household Welfare

This section investigates the effect of credit constraints identified using directly elicitation approach on household welfare. Two instrumental variables are used for this analysis to control possible endogeneity of credit constraints variable; whether the community where households are residing in has a formal bank and average enhanced vegetation index which measures health vegetation in the past five years. These instruments may relate with credit constraints status but may not necessarily directly relate with household welfare proxied by consumption expenditure level of household.

Columns (1) and (2) of Table 4.8 present that both quantity ratioing and overall credit constraints have a significant negative relationship with consumption expenditure level of household. The instrumental variables are jointly significant in the first stage (F-statistics is 11.3) and pass the Sargan test for overidentification, which means instrumental variables used here are valid.

Household welfare is positively associated with total physical assets, while it is negatively related with livestock assets. Household welfare also has a positive relationship with age of household head, female household head, education level of household head. Number of adult males and adult female in household and whether households reside in rural area have a negative relationship with household welfare.

Table 4.8. The Effects of Credit Constraints on Household Welfare (IV Estimates)

Dependent variable: Total consumption expenditure per capita	**(1) Quantity rationed**		**(2) Overall credit constraints**	
Credit constraints	-402,672	***	-345,066	***
	(125,506.758)		(109,428.733)	
Log of total assets	29,577	***	30,397	***
	(2,393.523)		(2,486.523)	
Log of total livestock value	-1,870	***	-1,019	
	(581.511)		(643.994)	
House ownership	-16,667	*	-16,112	
	(9,865.762)		(10,043.181)	
Age of HH head	1,006	***	958	***
	(310.333)		(322.549)	
Female HH head	26,729	**	27,832	**
	(11,398.366)		(11,615.202)	
Years of schooling of HH head	2,905	***	2,629	***
	(453.965)		(460.230)	
Number of adult males	-12,452	***	-12,627	***
	(2,603.805)		(2,699.513)	
Number of adult females	-29,407	***	-31,205	***
	(2,665.519)		(2,711.386)	
Present bias preference	2,678		-987	
	(6,124.306)		(6,370.698)	
Risk preference	-3,362		-2,348	
	(3,027.774)		(2,999.588)	
Trust	2,024		-4,388	
	(7,243.587)		(7,191.663)	
Loss aversion	5,385		9,779	
	(7,088.761)		(8,113.100)	
Rural	-40,405	***	-46,564	***
	(8,009.286)		(8,196.855)	
Constant	-90,637	**	-80,074	**
	(37,058.267)		(39,672.778)	
Observations	4,232		4,232	

Note: 1. Robust standard errors in parenthesis *** p<0.01, ** p<0.05, * p<0.1.
2. Each regression includes zone fixed effects.

4.4.3. Possible Strong Predictors of Households' Credit Constraints

This section explores which variables are strong predictors of households' credit constraints status in Nigeria among 87 of potential predictors available in the dataset using supervised machine learning approach. Since the directly elicited credit constraints variable not often available in general household survey, if credit constraints status can be predicted by a subset of observable characteristics, it could be efficiently used for improving targeting for microcredit program when lacking information on households' credit constraints status. One of the advantages of using machine learning approach is that all the variables available including even unexpected variables within dataset can be explored as a potential predictor since machine learning approach manages to fit complex and flexible functional forms, offering a set of methods that outperform OLS in terms of out-of-sample prediction.

To identify the best predictors by selecting a subset of variables within dataset, least absolute shrinkage and selection operator (LASSO) with 10-fold cross-validation is used. Package glmnet (Friedman et al., 2010) in R program is employed. Using a random 2:1 training set and test set split, I first estimate the algorithm using training set, and then evaluate the predictions using the hold-out sample set following Kshirsagar et al. (2017).

Figure 4.5 illustrates 10-fold cross-validation to select λ for prediction of overall credit constraints. The red dotted line indicates the cross-validation curve and it also shows upper and lower standard deviation curves along the λ sequence. I choose the value of λ, which gives the most regularized model such that error is within one standard error of the minimum since the aim of this section is to identify the subset of best predictors (Hastie et al., 2009; Friedman et al., 2010).

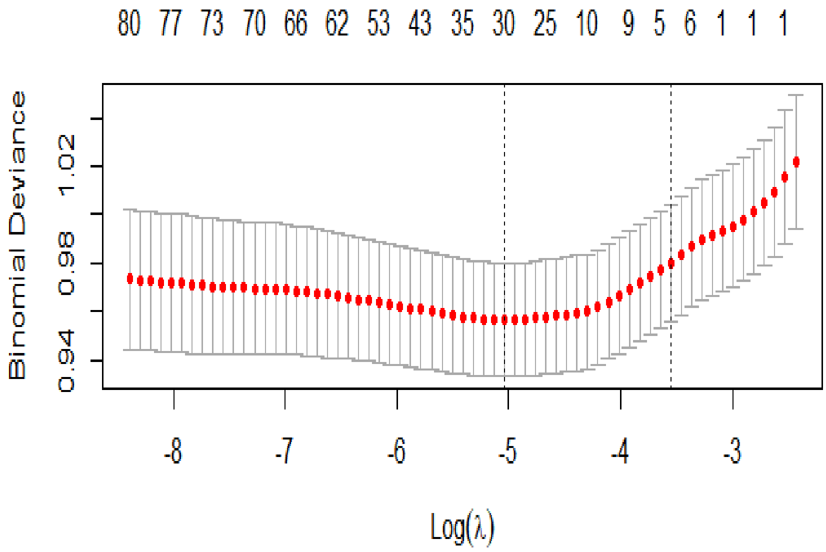

Figure 4.5. 10-Fold Cross-Validation to Select Lambda (λ)

Table 4.9 reports the selected predictors of each rationing category by LASSO algorithm. For overall credit constraints, whether households participate in semiformal credit market and whether households participate in informal credit market are selected among household level variables. Among community level variables, distance to the capital of the state of residence, precipitation and the average timing of onset of greenness decrease in day of year (in the reference period of from 2001 to 2015) are selected, which may relate with information sharing and agricultural outputs. The average timing of onset of vegetation greenness decrease is one of the crop season parameters of community level, and it is the one of the key transition dates in the annual cycles of vegetation growth, reflecting spatial patterns in climate.

For quantity rationing, participation of each credit market is unsurprisingly a strong predictor, but the average timing of onset of greenness decrease in day of year is also selected. For risk rationing, interestingly, four of household level variables including grass roof and the right to use land as collaterals as well as a community level variable related with terrain (elevation) are selected. For transaction cost ratioing, only grass roof is selected as a predictor.

Predictors of each credit constraints status selected by LASSO algorithm are different from determinants used in the previous section 4.4.1. The selected predictors by LASSO could hardly be interpreted in the similar way with the estimated parameters by econometric methods because Machine Learning approach does not produce consistent estimates of the underlying parameters due to lack of standard errors on the coefficients. In addition, the selected predictors can be unstable when the variables are correlated with each other, substituting with each other in prediction, which could produce similar predictions using very different variables while not necessarily affecting the performance of prediction (Mullainathan and Spiess, 2017).

For assessing accuracy of prediction, the out-of-sample mean squared error (MSE) is computed using the hold-out sample set and the square of the Pearson correlation coefficient (i.e. the correlation between the actual and predicted outcome in the hold-out sample) are reported. Although, the out-of-sample performance is quite low, these results suggest that the probability of being credit constrained can be predicted by not only households' credit market participation and but also community level data such as climatology, terrain, and crop season parameters. These selected predictors can be applicable to new datasets which lack information on household's credit constraints status for improving the accuracy of targeted microcredit programs.

Table 4.9. Predictors of Each Rationing Category Selected by LASSO Algorithm

Overall credit constraints	Coefficients	Quantity rationing	Coefficients
Household level variables		*Household level variables*	
Participation of semiformal credit market	0.20411	Participation of formal credit market	0.04379
Participation of informal credit market	0.03392	Participation of semiformal credit market	0.30152
		Participation of informal credit market	0.12897
Community level variables		*Community level variables*	
Household distance to the capital of the State of residence	-0.00009	Average timing of onset of greenness decrease in day of year 1-356, within growing season, averaged by state (2001-2015)	0.00013
Precipitation of wettest month, from monthly climatology (mm)	0.00008		
Average timing of onset of greenness decrease in day of year 1-356, within growing season, averaged by state (2001-2015)	0.00021		
N	4,232	N	4,232
Selected lambda (λ)	0.02868	Selected lambda (λ)	0.02752
Mean squared error (MSE)	0.14964	Mean squared error (MSE)	0.09084
Squared correlation between the fitted value & actual value (R^2)	0.20507	Squared correlation between the fitted value & actual value (R^2)	0.36807

Risk rationing	Coefficients	Transaction cost rationing	Coefficients
Household level variables		*Household level variables*	
The roof of the main dwelling is predominantly made of grass	0.00085	The roof of the main dwelling is predominantly made of grass	$2.789864e^{-16}$
Participation of semiformal credit market	-0.02912		
Participation of informal credit market	-0.02392		
Have the right to use land as collateral	0.00099		
Community level variables			
Elevation (m)	-0.00001		
N	4,232	N	4,232
Selected lambda (λ)	0.013	Selected lambda (λ)	0.00651
Mean squared error (MSE)	0.0612	Mean squared error (MSE)	0.00895
Squared correlation between the fitted value & actual value (R^2)	0.11376	Squared correlation between the fitted value & actual value (R^2)	0.04266

4.5. Chapter Summary

This chapter found that the probability of quantity rationing increases with household in the south area in Nigeria where relatively larger investment opportunities are available compared to the north area. Loss averse is found to be positively related with each credit rationing except for transaction-cost rationing. This chapter also found that if household heads engage in waged labor, the probability of being risk rationed decreases, while the probability of being transaction-cost rationed decreases with households with non-farm enterprises. For the lower income households, the probability of being risk rationed increases if households engage in agriculture. LASSO algorithm results suggest that the probability of being credit constrained can be predicted by not only households' credit market participation and but also community level data such as climatology, terrain, and crop season parameters; precipitation, elevation and average timing of onset of greenness decrease in day of year. These selected predictors could be applicable to other datasets which does not contain information on households' credit constraints status for improving targeting for microcredit program beneficiaries.

Findings from this chapter suggests that policies which aim to offer microcredit program for lower income households which engage in agriculture may be more effective if policies focus on relaxing risk-rationing constraints. For example, microcredit programs linked with insurance scheme would be more attractive for those who withdraw their credit application because they consider an accompanying risk of loan application is higher than they can benefit from loan. Conventional microcredit program may be more appropriate for household living in the south area of Nigeria with effort to relax quantity rationing. For example, credit scoring using the usage history of mobile phone is being considered recently in order to provide information on qualified borrowers to lenders so that lenders would more easily lend their money in the context of developing countries wherein borrowers tend to lack of collaterals and lenders would face asymmetric information.

5. Conclusion

In order to include the poor into the formal financial system, financial services should be more tailored to the poor's specific needs to attract them. Thus, the demand side of financial services needs to be carefully considered.

In Chapter 2, to find the determinants of the choice between formal bank accounts and informal saving methods, this study uses primary data collected from unskilled and low-income production workers in cut flower farms in Ethiopia. By using the multivariate probit model and seemingly unrelated regression model, I find that having greater assets (i.e., those who are more self-insured) leads to more bank savings but does not affect the choice of ROSCA savings. I further find that risk-aversion positively affects ROSCA savings among savers but not bank savings, suggesting that the workers value the social insurance aspect of ROSCAs. I also find that more impatient workers and those with greater social connectedness tend to save more with ROSCAs. Overall, informal saving groups work as insurance and commitment tools based on social networks within farms, complementing formal financial institutions, and that workers use these methods of saving strategically.

In Chapter 3, using the same dataset as the previous chapter, I explore saving behaviors of internal migrants and their working environment. I find that earnings of workers are significantly different between migrant and non-migrant and migrants are likely to be paid less. However, migrants are more likely to be productive and less likely to be absent. The Blinder–Oaxaca decomposition results reveal that about half of wage differentials is unexplained even when non-cognitive skills are controlled, supporting the idea of wage differentials and the possible differential treatment in compensation against migrants in cut flower industry in Ethiopia. With field observation and descriptive statistics on financial behavior of migrants, even though they may face discrimination or suffer from disadvantages in the labor market, higher levels of subjective wellbeing and descriptive analysis results on financial behaviors suggest that rural to urban migration would be investment for better job opportunities.

In Chapter 4, I investigate the determinants and predictors of detailed non-pricing credit constraints using a nationally representative large dataset in Nigeria. I find that households in the south region of Nigeria are more likely to face quantity rationing. I also find that the probability of being risk rationed decreases if household heads engage in waged labor, while the probability of being transaction-cost rationed decreases when households owns their non-farm enterprises. Loss averse is found to be positively related with each credit rationing except transaction-cost rationing. For lower income households, the probability of being risk rationed increases if households engage in agriculture. By using supervised machine learning approach, I further find that the probability of being credit constrained can be predicted by not only households' credit market participation and but also community level data such as climatology, terrain, and crop season parameters; precipitation, elevation and average timing of onset of greenness decrease in day of year. These selected predictors could be used for other datasets which does not contain information on households' credit constraints status for better targeting.

To sum up, this study investigates the determinants of savings choices and credit constraints in developing countries, focusing on the demand side of financial services. This study contributes to the literature on financial inclusion by emphasizing on demand-side saving behavior by using rich data on financial behavior of workers and unique contextual features, as well as providing comprehensive analyses on different types of non-pricing credit constraints using the national-representative large dataset. The findings from this study are expected to help policymakers and practitioners consider financial services that are tailored to the financial needs and preferences of the poor to include the poor in formal financial systems.

Finally, there are several limitations to this study. First, the results on savings may not be generalizable to other occupational groups since saving behaviors can vary across occupations. Future research should seek to understand saving method preferences and saving habits in developing countries and disentangle their motives. Second, the direct identification approach of credit constraints needs be improved more. Especially, distinguishing different types of discouraged borrowers is challenging issue. Thus, a further study should be conducted to better understand the underlying motivations of nonborrowers.

References

Ali D. A., K. Deininger, and M. Duponchel. (2014). "Credit Constraints and Agricultural Productivity: Evidence from rural Rwanda." *Journal of Development Studies* 50(5): 649-665.

Ambec, S. and N. Treich. (2007). "ROSCAs as financial agreements to cope with self-control problems." *Journal of Development Economics* 82(1): 120-137.

Anderson, S. and J.-M Baland. (2002). "The Economics of Roscas and Intrahousehold Resource Allocation." *Quarterly Journal of Economics* 117(3): 963-995.

Anderson, S., J.-M Baland, and K. O. Moene. (2009). "Enforcement in Informal Saving Groups." *Journal of Development Economics* 90(1): 14-23.

Antoninis, M. (2006). "The wage effects from the use of personal contacts as hiring channels." *Journal of Economic Behavior and Organization* 59: 133-146.

Ashraf, N., D. Karlan, and W. Yin. (2006). "Tying Odysseus to the Mast: Evidence from a Commitment Savings Product in the Philippines." *Quarterly Journal of Economics* 121, 673-697.

Banerjee, A. V. (2013). "Microcredit Under the Microscope: What Have We Learned in the Past Two Decades, and What Do We Need to Know?" *Annual Review of Economics* 5: 487-519.

Banerjee, A., D. Karlan and J. Zinman. (2015a). "Six Randomized Evaluations of Microcredit: Introduction and Further Steps." *American Economic Journal: Applied Economics* 7(1): 1-21.

Banerjee, A., E. Duflo, R. Glennerster and C. Kinnan. (2015b) "The Miracle of Microfinance? Evidence from a Randomized Evaluation." *American Economic Journal: Applied Economics* 7(1) 22-53.

Benzion, Uri, A. Rapoport, and J. Yagil. (1989). "Discount Rates Inferred from Decisions: An Experimental Study," *Management Science*, XXV, 270-284.

Becker, G. (1958). "The Economics of Discrimination." *Social Forces* 37(2):180–181.

Besley, T., S. Coate, and G. Loury. (1993). "The Economics of Rotating Savings and Credit Associations." *American Economic Review* 83(4): 792-810.

Bisrat, A., K. Kostasa and L. Feng. (2012). "Are there Financial Benefits to Join RoSCAs? Empirical Evidence from Equb in Ethiopia." *Procedia Economics and Finance* 1, 229-238.

Blinder, A. (1973). "Wage Discrimination: Reduced Form and Structural Estimates," *Journal of Human Resources* 7: 436-55.

Blumenstock, J. E., G. Cadamuro, and R. On. (2015). "Predicting Poverty and Wealth from Mobile Phone Metadata." *Science* 350(6264): 1073-76.

Boucher, S. R., C. Guirkinger, and C. Trivelli. (2009). "Direct Elicitation of Credit Constraints: Conceptual and Practical Issues with an Application to Peruvian Agriculture." *Economic Development and Cultural Change* 57(4): 609-640.

Calomiris, C. W. and I. Rajaraman. (1998). "The Role of ROSCAs: Lumpy Durables or Event Insurance?" *Journal of Development Economics* 56, 207-216.

Cameron, A. C. and P. K. Trivedi. (2010). *Microeconometrics Using Stata*. Texas: Stata

Press Publication.

Carpenter, S. B. and R. T. Jensen. (2002). “Household Participation in Formal and Informal Savings Mechanisms: Evidence from Pakistan.” *Review of Development Economics,* 6(3): 314-328.

CGAP. (2009). *Financial Access 2009: Measuring Access to Financial Services Around the World.* Washington, DC: World Bank.

Chiburis, R. C., J. Das, and M. Lokshin. (2012). “A Practical Comparison of the Bivariate Probit and linear IV Estimators.” *Economics Letters* 117(3): 762-766.

Conley, T. G. and C. R. Udry. (2010). “Learning about a New Technology: Pineapple in Ghana.” *American Economic Review* 100(1): 35-69.

Deaton, A. (1992). “Saving and Income Smoothing in Cote D’ivoire.” *Journal of African Economies* 1(1): 1-24.

Démurger, S., M. Gurgand, L. Shi, and Y. Ximing. (2009). “Migrants as Second-Class Workers in Urban China? A Decomposition Analysis.” *Journal of Comparative Economics* 37:610-628.

Diagne, A. (1999). “Determinant of Household Access to and Participation in Formal and Informal Credit Markets in Malawi.” FCND Discussion Paper No. 67, Food Consumption and Nutrition Division, International Food Policy Research Institute.

Diagne, A., M. Zeller and M. Sharma. (2000). “Empirical Measurements of Household's Access to Credit and Credit Constraints in Developing Countries: Methodological Issues and Evidence.” FCND Discussion Paper No. 90, International Food Policy Research Institute, Washington. DC.

Dupas, P. and J. Robinson. (2013). “Savings Constraints and Microenterprise Development: Evidence from a Field Experiment in Kenya.” *American Economic Journal: Applied Economics* 5(1): 163-192.

Dupas, P., D. Karlan, J. Robinson, and D. Ubfal. (2018). “Banking the Unbanked? Evidence from Three Countries.” *American Economic Journal: Applied Economics* 10(2): 257-297.

Friedman J, Hastie T, Tibshirani R (2010). “Regularization Paths for Generalized Linear Models via Coordinate Descent.” *Journal of Statistical Software* 33(1): 1–22

Garcia, R., A. Lusardi and S. Ng, (1997). “Excess Sensitivity and Asymmetries in Consumption: An Empirical Investigation,” *Journal of Money, Credit, and Banking* 29(2): 154-176.

Getahun, T. D. and E. Villanger. (2018). “Labour-Intensive Jobs for Women and Development: Intra-Household Welfare Effects and its Transmission Channels.” *Journal of Development Studies* 54(7): 1232-1252.

Gugerty, M. K. (2007). “You Can’t Save Alone: Commitment in Rotating Savings and Credit Associations in Kenya.” *Economic Development and Cultural Change* 55, 251-282.

Hastie T, Tibshirani R, Friedman J (2009). *The Elements of Statistical Learning: Prediction, Inference and Data Mining*. 2nd edition. Springer-Verlag, New York.

Heltberg, R. (2013). “Shocks and Coping in Sub-Saharan Africa,” Living Standards Measurement Study Brief Series, World Bank.

Hjort, J. (2014). “Ethnic divisions and production in firms.” *Quarterly Journal of Economics*

129(4): 1899-1946.

International Labour Office. (2016). “The Role of Internal Migration in Access to First Job: A Case Study of Uganda.” Technical Brief No.5.

Jappelli, T. (1990). “Who Is Credit-Constrained in the U.S. Economy?” *Quarterly Journal of Economics*, Vol.105, No.1, pp.219-234.

Jean, N., M. Burke, M. Xie, W. M. Davis, D. B. Lobell, and S. Ermon. (2016). “Combining Satellite Imagery and Machine Learning to Predict Poverty.” *Science* 353(6301): 790-94.

Kaboski, J.P. and R.M.Townsend. (2012). “The Impact of Credit on Village Economies.” *American Economic Journal: Applied Economics* 4(2): 98-133

Karlan, D. and J. Morduch. (2010). Access to Finance. *Handbook of Development Economics* 5(C), 4703-4784.

Karlan, D., A. L. Ratan and J. Zinman. (2014). “Savings by and for the Poor: A Research Review and Agenda.” *Review of Income and Wealth* 60(1): 36-78.

Kedir, A. M., R. Disney, and I. Dasgupta. (2011). “Why Use ROSCAs When You Can Use Banks? Theory, and Evidence from Ethiopia.” Discussion Papers in Economics 11/32, Department of Economics, University of Leicester.

Kedir, A. M. and G. Ibrahim. (2011). “ROSCAs in Urban Ethiopia: Are the Characteristics of the Institutions More Important than those of Members?” *Journal of Development Studies* 47(7): 998-1016.

Klonner, S. (2003). “Rotating Savings and Credit Organizations When Participants Are Risk Averse.” *International Economic Review* 44(3): 979-1006.

Knippenberg, E., N. Jensen, and M. Constas. (2019). “Quantifying Household Resilience with High Frequency Data: Temporal Dynamics and Methodological Options.” *World Development* 121: 1-15

Kremer, M., G. Rao and F. Schilbach. (2019). “Behavioral Development Economics.” Handbook of Behavioral Economics, D. Bernheim, S. DellaVigna, and D. Laibson (eds.).

Kshirsagar, V., J. Wieczorek, S. Ramanathan, and R. Wells (2017). “Household Poverty Classification in Data-Scarce Environments: A Machine Learning Approach.” arXiv preprint arXiv:1711.06813.

Lee, L. (2012). “Decomposing Wage Differentials between Migrant Workers and Urban Workers in Urban China’s Labor Markets.” *China Economic Review* 23: 461-70.

Levenson, A.R. and T. Besley. (1996). “The Anatomy of an Informal Financial Market: Rosca Participation in Taiwan.” *Journal of Development Economics* 51(1): 45-68.

List, J.A. and I. Rasul, (2011). “Field Experiments in Labor Economics.” Handbook of Labor Economics, Volume 4a, 103-228.

Luke, N. and K. Munshi. (2006). “New roles for marriage in Urban Africa: Kinship networks and the labor market in Kenya.” *Review of Economics and Statistics* 88(2): 264-282.

Maertens, A. and C. B. Barrett. (2013). “Measuring Social Networks' Effects on Agricultural Technology Adoption.” *American Journal of Agricultural Economics* 95(2): 353-359.

Mano, Y., T. Yamano, A. Suzuki, and T. Matsumoto. (2011). “Local and Personal Networks in Employment and the Development of Labor Markets: Evidence from the Cut Flower Industry in Ethiopia.” *World Development* 39(10): 1760-1770.

Meng, X. and J. Zhang. (2001). "The Two-Tier Labor Market in Urban China: Occupational Segregation and Wage Differentials between Urban Residents and Rural Migrants in Shanghai." *Journal of Comparative Economics* 29(3): 485-504.

Murendo, C., M. Wollni, A. D. Brauw, and N. Mugabi. (2017). "Social Network Effects on Mobile Money Adoption in Uganda." *The Journal of Development Studies* 54(2): 327-342.

Mullainathan, S. and J. Spiess. (2017). "Machine Learning: An Applied Econometric Approach." *Journal of Economic Perspectives* 31(2):87-106.

Munshi, K. (2003). "Networks in the modern economy: Mexican migrants in the US labor market." *Quarterly Journal of Economics* 118(2): 549-597.

Murphy, A. (2007). "Score Tests of Normality in Bivariate Probit Models." *Economics Letters* 95(3): 374-379.

Nigeria National Bureau of Statistics. General Household Survey, Panel (GHS-Panel) 2015-2016. Ref. NGA_2015_GHSP-W3_v02_M. Dataset downloaded from www.microdata.worldbank.org on 2019.11.11

Oaxaca, R. (1973). "Male–Female Wage Differentials in the Urban Labour Market." *International Economic Review* 14: 693-709.

Prina, S. (2015). "Banking the Poor via Savings Accounts: Evidence from a Field Experiment." *Journal of Development Economics* 115, 16-31.

Schwab, S. (1986). "Is Statistical Discrimination Efficient?" *The American Economic Review* 76(1): 228-234.

Simtowe, F., A. Diagne and M. Zeller, (2008). "Who is Credit Constrained? Evidence from Rural Malawi," *Agricultural Finance Review* 68(2): 255-272.

Suzuki, A. (2015). "Risk on Dynamic Behavior of Farmers in the Export Market: A Case from the Pineapple Industry in Ghana." 2014 AAEA Annual Meeting. Minneapolis: Agricultural & Applied Economics Association.

Suzuki, A., Y. Mano, and G. Abebe. (2018). "Earnings, Savings, and Job Satisfaction in a Labor-Intensive Export Sector: Evidence from the Cut Flower Industry in Ethiopia." *World Development* 110, 176-191.

Tversky, Amos, and D. Kahneman. (1986). "Rational Choice and the Framing of Decisions." *Journal of Business*, LIX, S251-278.

Tanaka, T., C. Camerer, and Q Nguyen. (2010). "Risk and time preferences: Linking experimental household survey data from Vietnam." *American Economic Review* 100, 557-571.

Udry, C. (1994). "Risk and Insurance in a Rural Credit Market: An Empirical Investigation in Northern Nigeria." *The Review of Economic Studies* 61(3): 495-526.

Wright, G. and L. Mutesasira. (2001). *The relative risks of savings to poor people*. MicroSave Briefing Note 6.

Zeldes, Stephen P., (1989). "Consumption and Liquidity Constraints: An Empirical Investigation," *Journal of Political Economy* 97(2): 305-346.

Appendix

Table A.1. Determinants of Mutually Exclusive Savings Choices (Probit Estimates)

Dependent variable: =1 if save using	**(1) Bank only**	**(2) ROSCAs only**	**(3) Cash at home only**	**(4) Bank & ROSCAs**	**(5) Bank & cash**	**(6) ROSCAs & cash**	**(7) Bank, ROSCAs & cash**
Individual-level characteristics							
= 1 if female	-0.083	0.299	0.475 *	0.002	-0.130	0.380	-0.157
	(0.161)	(0.234)	(0.255)	(0.162)	(0.197)	(0.306)	(0.183)
Years of schooling	-0.009	-0.018	0.025	0.047 ***	0.007	0.006	0.016
	(0.016)	(0.021)	(0.023)	(0.017)	(0.020)	(0.024)	(0.020)
Asset and income proxy variables							
The total hectares of farmland a worker owns	-0.041	-1.417 **	-0.374	0.018	-0.146	0.113	0.232 ***
	(0.083)	(0.689)	(0.240)	(0.071)	(0.156)	(0.103)	(0.080)
= 1 if a worker keeps livestock	-0.273	-0.664 **	0.452	0.219	0.532 ***	-0.201	0.282
	(0.166)	(0.307)	(0.283)	(0.155)	(0.200)	(0.424)	(0.178)
The log of total food expenditures	-0.104	0.190	-0.602***	0.133	0.076	-0.174	0.238 *
	(0.101)	(0.140)	(0.160)	(0.123)	(0.134)	(0.143)	(0.133)
Risk and time preferences							
Risk averse (8: the most risk averse)	-0.013	0.019	-0.067 **	0.032	-0.017	-0.025	0.012
	(0.022)	(0.030)	(0.032)	(0.023)	(0.028)	(0.034)	(0.030)
Impatience (12: the most impatient)	-0.024	0.071 ***	0.002	0.000	-0.037 *	-0.020	0.016
	(0.016)	(0.024)	(0.023)	(0.017)	(0.022)	(0.022)	(0.023)
= 1 if present-biased time-inconsistent preference	-0.253	0.034	0.421 **	0.045	-0.008	-0.247	0.440 **
	(0.168)	(0.222)	(0.206)	(0.165)	(0.213)	(0.299)	(0.187)
Cognitive skills							
Financial literacy (5: the highest number of correct answers)	0.027	0.045	-0.131	0.044	-0.011	-0.086	0.026
	(0.058)	(0.078)	(0.089)	(0.060)	(0.078)	(0.094)	(0.062)
Observations				709			

Note: 1. Robust standard errors in parentheses *** p<0.01, ** p<0.05, * p<0.1.
2. Each regression includes farm fixed effects. Controls include age, age squared, marital status, migrant, Oromo ethnic group, and household size.

Table A.2. The Blinder–Oaxaca Decomposition (Recentered Influence Functions (RIFs) decomposition)

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
		Q10			Q50			Q90	
	overall	explained	unexplained	overall	explained	unexplained	overall	explained	unexplained
Non-migrant group	6.868***			7.011***			7.446***		
	(0.016)			(0.015)			(0.038)		
Migrant group	6.765***			6.986***			7.164***		
	(0.012)			(0.006)			(0.036)		
Difference	0.103***			0.024			0.281***		
	(0.017)			(0.016)			(0.047)		
Explained	0.016			0.024**			0.107***		
	(0.011)			(0.010)			(0.039)		
Unexplained	0.087***			0.000			0.174***		
	(0.023)			(0.015)			(0.054)		
Age		-0.000	-0.157		0.000	0.022		0.012	0.491
		(0.006)	(0.288)		(0.004)	(0.246)		(0.039)	(0.876)
Age squared		0.000	0.029		-0.001	-0.053		-0.016	-0.263
		(0.006)	(0.109)		(0.004)	(0.104)		(0.051)	(0.398)
= 1 if female		0.002	0.068*		0.002	-0.018		0.012	-0.025
		(0.003)	(0.038)		(0.002)	(0.023)		(0.008)	(0.116)
Years of schooling		-0.002	-0.001		0.003	-0.016		0.007	-0.042
		(0.004)	(0.020)		(0.003)	(0.019)		(0.008)	(0.057)
Years of working experience		0.018	-0.031		0.015	0.004		0.013	0.094
		(0.011)	(0.090)		(0.010)	(0.057)		(0.027)	(0.225)
Years of working experience squared		-0.007	0.010		-0.002	0.002		0.046	-0.088
		(0.006)	(0.037)		(0.004)	(0.023)		(0.038)	(0.169)
= 1 if new worker less than one year		-0.005	-0.005		0.007*	-0.035***		-0.005	0.025
		(0.005)	(0.020)		(0.004)	(0.011)		(0.006)	(0.020)
= 1 if married		0.001	0.006		0.001	-0.003		-0.004	-0.010
		(0.002)	(0.015)		(0.001)	(0.008)		(0.005)	(0.047)
		0.008	-0.079*		-0.000	-0.040**		0.042*	0.090
		(0.011)	(0.041)		(0.006)	(0.019)		(0.024)	(0.077)
Observations	710	710	710	710	710	710	710	710	710