An Analysis of Households' Credit Markets in Ethiopia and Malawi

Eleonora Fichera, B.A., M.Sc.

Thesis submitted to the University of Nottingham
for the degree of Doctor of Philosophy

June 2010

Do not ask us the word which in every way
our shapeless soul perhaps measures, and in letters of fire
may declaim it and shine like a crocus
lost in the centre of a dusty field.

Ah! the man who goes away sure,
to others and to himself a friend,
and cares not about his shadow which the dog days
reflect across a plasterless wall!

Ask us not for the formula to open worlds for you,
only some syllable distorted and dry like a twig.
This alone is what we can tell you today,
that which we are not, that which we do not want.

(Eugenio Montale, Ossi di seppia, 1925)

Abstract

The aim of this thesis is to analyse formal and informal credit in Ethiopia and Malawi. As credit markets in developing economies are dominated by informal institutions, the analysis of the interaction between formal and informal institutions is crucial to understanding how welfare improvements can be achieved.

The thesis begins with an explanation of the motives for demanding credit. It then focuses on analysing the existence, diffusion and persistence of informal finance in developing economies. Much research on this topic remains hamstrung by the quality and availability of data and by the lack of empirical models, constraining the meaningful identification of the characteristics of the localities where informal institutions operate.

The central idea of the first essay is to develop an empirical model that explains the determinants of participation in informal credit arrangements. We adopt an endogenous switching regression model of access to informal credit where the availability of a particular type of informal arrangement varies across clusters in rural Ethiopia. This strategy allows for taking into account substitutability between sources as well as household and cluster socioeconomic characteristics.

The second essay exploits the idea that banks can crowd out informal borrowing in Malawi by creating microfinance institutions that acquire information in innovative ways. We adopt propensity score matching and find that the creation of a specific microfinance programme reduces informal borrowing.

The third essay uses the credit limit variable to test liquidity constraints and the spillover hypotheses in Malawi. A ten percent increase in the informal credit line increases households' demand for informal credit by more than nine percent. We also find that a 10 percent increase in the credit limit of a microfinance programme reduces the informal demand by four percent, partly explaining the coexistence of formal and informal credit institutions.

Acknowledgements

The doctoral thesis is part of a long journey. I wish to thank the people that pushed, that pulled and that accompanied me along my life's journey up to now.

To my partner, who has listened to my struggles and has just been there for me even in times of self-hatred. The word "love" cannot adequately express what I feel for you...

To my parents, for having cultivated the seek of knowledge above all.

A special thought goes to my high school teacher of philosophy, Silvana Salandra. Thank you so much for encouraging my appetite for research and for pushing me to seek questions rather than "right" answers.

Thanks to my school friends for making me feel home when we are together. In particular, I wish to thank Antonella, Christian, Nama, Sergio, Valentina and Vittoria.

Thank you to my spiritual teacher, Orsi. Your soul has been here many more times than mine and I feel lucky to have known you.

To my undergraduate friends and classmates, especially Chiara and Elena. We have struggled and joyed together for our successes and failures.

I wish to thank my first (and best) teacher of Development Economics, E. La Ferrara. I would not be at this point had I not attended your lectures. The memory of them helps me to go through my most apathetic periods.

I am grateful to my undergraduate supervisor Giovanni Bruno. Thank you for being so generous with your time: I have learnt so much from you. Your suggestions have always been on my best interest and not yours.

I also thank my supervisors, Richard Disney and Trudy Owens for their support. Their comments have made me a rigorous researcher.

I cannot avoid mentioning the many friends I have met in the UK. My japanese friend, Asako, with whom I shared the toughest challenges during the master. I send all my affection to my african brothers, Henry and Vincent, may our doors be always open to each other. Special thanks go to Michael who has been my "ambulance" and technical advisor. To my sweet bengali friend, Bidisha: I will never forget you. Thanks to my housemate, Johannes, for leaving sweets and chocolate in front of my door during many sleepless nights.

Finally, but not less importantly, I would like to thank the italian group: Alessia, Amanda, Arcangelo, Daniele and Mirona, Elena. Thanks for the lovely dinners and nights out without which the PhD would have been much more stressful.

"In the end, though, maybe we must all give up trying to pay back the people in this world who sustain our lives. In the end, maybe it's wiser to surrender before the miraculous scope of human generosity and to just keep saying thank you, forever and sincerely, for as long as we have voices" [E. Gilbert, 2006].

Contents

List of Figures

List of Tables

Abbreviations

ATE	**A**verage **T**reatment **E**ffect
ATT	**A**verage **T**reatment effect on **T**reated
CBM	**C**ommercial **B**ank of **M**alawi
CIA	**C**onditional **I**ndependence **A**ssumption
CM	**C**onditional **M**oment
ERHS	**E**thiopian **R**ural **H**ousehold **S**urvey
EUHS	**E**thiopian **U**rban **H**ousehold **S**urvey
FIML	**F**ull **I**nformation **M**aximum **L**ikelihood
FMHFS	**F**inancial **M**arkets and **H**ousehold **F**ood **S**ecurity
GTZ	**G**erman agency for **T**echnical **C**ooperation
HWS	**H**uber **W**hite **S**andwich
IFAD	**I**nternational **F**und for **A**gricultural **D**evelopment
IFPRI	**I**nternational **F**ood **P**olicy **R**esearch **I**nstitute
IIA	**I**ndependence of **I**rrelevant **A**lternatives
IV	**I**nstrumental **V**ariable
LAD	**L**east **A**bsolute **D**eviation
LC	**L**ife **C**ycle
MAR	**M**issing **A**t **R**andom
MCAR	**M**issing **C**ompletely **A**t **R**andom
MK	**M**alawian **K**wachas
MMF	**M**alawi **M**udzi **F**und
MNP	**M**ulti**N**omial **P**robit
MPC	**M**arginal **P**ropensity to **C**onsume
MRFC	**M**alawi **R**ural **F**inance **C**ompany
MSCE	**M**alawi **S**chool **C**ertificate of **E**ducation
MUSCCO	**M**alawi **U**nion of **S**avings and **C**redit **C**ooperatives
NGOs	**N**on **G**overnmental **O**rganisation**s**
OLS	**O**rdinary **L**east **S**quare
PAs	**P**easant **A**ssociation**s**
PC	**P**rincipal **C**omponent
PIH	**P**ermanent **I**ncome **H**ypothesis
PMERW	**P**romotion of **M**icro **E**nterprises for **R**ural **W**omen
PSID	**P**anel **S**tudy of **I**ncome **D**ynamics
RoSCAs	**Ro**tating **S**avings and **C**redit **A**ssociation**s**
SB	**S**tandardised **B**ias
RUM	**R**andom **U**tility **M**odel
SACCOs	**S**avings and **C**redit **C**ooperative**s**

To J. M. K.,

"God dwells within you, as you."

Chapter 1

Introduction

> *"A variety of institutions contribute to the process of development precisely through their effects on enhancing and sustaining individual freedoms as well as substantive opportunities".*
>
> Amartya Sen (1999)

1.1 Motivation

Rural households in developing economies have volatile and low incomes. These households suffer from income shocks due to fluctuations in weather and consumption prices and from health shocks due to infectious diseases. As they try to smooth income by adopting traditional production and employment choices and by diversifying economic activities, they obtain low returns for low risk strategies. In the presence of income shocks these households also try to smooth consumption by borrowing and saving from formal and informal credit arrangements.

In some way we can argue that borrowing is used by households as a saving strategy. An example could be taken considering households who borrow to acquire a tractor. The objective of the household is to create a self-commitment device to save for their old days. A tractor is a good basis for a self-commitment device because people increase their production and if they don't repay they lose the tractor again. This mechanism indirectly improves the welfare of households in two ways.

First, access to credit creates funds that alleviate households vulnerability to income shocks by facilitating risk-coping strategies. Credit will be available to cushion consumption against income shocks. Availability of credit can also avoid the adoption of low risk and low return strategies by providing incentives to undertake riskier technologies.

The second channel through which access to credit affects household welfare is by enhancing investments in human and physical capital [Binswanger and Khandker, 1995; Heidhues 1995; Nwanna, 1995]. Access to credit can raise productivity and reduce labour intensive technologies by decreasing the opportunity costs of capital intensive assets compared to family labour.

For these reasons, financial institutions have been regarded as a contributing factor to economic growth and development. Most of government interventions in rural credit markets are based on this premise and they have been further justified on the basis of improving the distribution of rural incomes. However, several interventions up to the 1990s have not really succeeded in fulfilling these objectives. Commercial, agricultural banks and other formal institutions fail to cater for the credit needs of smallholders due to a number of reasons: they lack appropriate informational sharing mechanisms and methods for dealing with asymmetries in credit markets; environments are very risky and markets are interlinked; there are few scale economies and weak legal systems [Bardhan and Udry, 1999; Besley, 1994; Gosh et al., 1999; Ray, 1997].

It is generally the terms of the contracts set by standard formal financial institutions that have created the myth that the poor are not bankable, and since they cannot afford the required collateral, they are considered uncreditworthy [Adera, 1995]. Despite efforts to overcome the widespread lack of financial services, especially among smallholders in developing countries, and the expansion of credit in the rural areas of developing countries, the majority still have only limited access to bank services to support their consumption and production decisions [Braverman and Guasch, 1986].

Thus, it is increasingly being recognised that formal institutions alone cannot achieve welfare improvements especially in the poorest rural areas of developing countries. As Rodrik et al. (2004) pointed out on the relation between formal institutions and development, "desirable institutional arrangements have a large element of context specificity, arising from differences in historical trajectories, geography and political economy or their initial conditions..." Hence, whether or not formal institutions improve welfare and encourage development is as much a question of the incentives and enforcement mechanisms of the institutions themselves as the environment they operate in (often dominated by the presence of informal credit arrangements) [e.g. Durlauf and Fafchamps, 2005; Fafchamps, 2006].

Since the effectiveness of formal credit institutions depends on informal arrangements, social norms, existing levels of social capital and markets linkages, analysing the factors that affect the formation and the access to informal institutions is crucial to understanding how the interaction between formal and informal credit institutions can be harnessed to effect desirable policy objectives.

In recent years this has indeed been the premise of the so-called "microfinance revolution" [Armendariz and Morduch, 2005]. By mimicking and exploiting the features of informal lending, banks can design contracts that harness local information and give borrowers incentives to use their own information on their peers to the advantage of the bank.

1.2 Objectives of the thesis

Broadly speaking, the objective of this thesis is to analyse formal and informal credit markets in Ethiopia and Malawi. More specifically, the thesis addresses the following research questions: Why do households participate in informal credit institutions? Do governments displace the informal loan market by introducing formal credit institutions? Why do formal and informal credit markets coexist?

Each of these questions is the focus of three self-contained essays: one focusing on Ethiopia and the other two on Malawi. As we recognise that Ethiopia and Malawi are two different countries, we make no attempt to compare them.

The setting

Ethiopia has one of the largest concentrations of poor people on the planet. It ranks 170 out of 177 countries in the 2006 United Nations Human Development Report. 31 million people live on less than half a dollar a day and between 6 and 13 million people are at risk of starvation each year. Poverty in Ethiopia affects the majority of the population: 81 percent of the 71.3 million people live below a poverty line of two U.S. dollars a day.

Livelihoods are predominantly based on agriculture, which accounts for 85 percent of employment, 45 percent of national income and over 90 percent of export earnings. Life expectancy is 48 years (UNICEF, 2004), under five mortality is 123 per 1,000 live births, and an estimated 1.4 percent of the adult population are living with HIV/AIDS (Demographic and Health Survey 2005). Food security is a major challenge. 15 million people are at risk from food insecurity, and over 8 million people are classed as chronically food insecure.

Malawi is one of the ten poorest countries in the world. It ranks 165 out of 177 countries according to the UN's Human Development Index. Around 60 percent of the population live below the poverty line. The population of around 13 million people (UN Population Division, 2005) is fast growing and young: less than three percent is over 65 years.

Malawi's economy is critically dependant on agriculture which accounts for 40 percent of GDP and over 90 percent of exports. Tobacco is the principal export (accounting for around 60 percent of export earnings), making Malawi vulnerable to tobacco price shocks. Life expectancy at birth has fallen from around 45 years in 1990 to around 37 years today. Malawi suffers from one of the worst HIV/AIDS epidemics in the world with around one million people infected. Food security does not exist, even during good harvests. Agricultural development has been hampered by recurring droughts and environmental degradation (deforestation, land degradation and water pollution).

The data

In spite of the diversities between these two countries, the widespread use of informal credit in Ethiopia and the government interventions in credit markets in Malawi represent the ideal environment for answering the above mentioned research questions. The two household surveys used in this thesis, the Ethiopian Rural Household Survey and the Malawi Rural Financial Markets and Household Food Security, are very rich data sets containing information about social and economic characteristics of the households as well as localities, and borrowing behaviour from formal and informal lenders. As a consequence, they constitute an invaluable source of information to analyse the characteristics and interaction of the formal and informal credit sectors.

Objectives

The specific objectives of each essay can be summarised as follows. The central idea of the first essay is to develop an empirical model that can be of use in analysing the determinants of participation in informal credit arrangements. We adopt an endogenous switching regression model of access to informal credit where the availability of a particular type of informal arrangement varies across clusters in rural Ethiopia. This strategy allows for taking into account substitutability between sources as well as household-based and cluster-based socioeconomic characteristics.

The second essay exploits the idea that banks can acquire the local information they lack (and that is readily available to informal lenders) in innovative ways. By creating microfinance institutions, banks can crowd out informal borrowing. We adopt a policy evaluation technique to test the effectiveness of this policy in Malawi.

Finally, the third essay uses information on the credit limit to explain the coexistence of formal and informal credit sources in Malawi.

Although the essays are self-contained and focus on two different countries, a unified story can be drawn from the thesis. If participation in informal arrangements depends on the socioeconomic characteristics of households as well as clusters, one way for banks to enter this market and exploit local information is to give borrowers incentives to use their existing social linkages to the advantage of the banks. But information problems are only part of the story, other market failures such as weak legal enforcement and the low level of social capital may force the banks to ration credit and cause the persistence of informal credit institutions. In addition, if the "social" motive[1] for participation in informal arrangements prevails over the "economic" motive, segmentation occurs despite banks' attempt to enter the market and complete crowding out will not be achieved.

The next section explains in detail the analysis and the contribution of each essay.

[1]See the next section for a summary of the sociological or cultural motive. A more detailed explanation of this approach is also contained in the second chapter.

1.3 Plan of the thesis

The three self-contained essays of this thesis focus on participation in informal credit in rural Ethiopia, effect of microfinance institutions on informal borrowing and the coexistence of formal and informal credit in Malawi. More specifically, the plan of the thesis can be summarised as follows.

The second chapter reviews the theoretical and empirical literature on credit markets comparing developed and developing countries. It provides a link between the three essays of this thesis.

While credit markets in developed countries are dominated by the formal sector, in developing economies - in particular sub-Saharan African countries - most of the loans originate from informal sources. After highlighting risk and acquisition of durable goods as motives for seeking credit (whether it be formal or informal), the literature review focuses on two theories for the existence and diffusion of informal credit in developing countries: the economic approach; and the cultural or sociological approach.

The economic approach maintains that informal finance arises as a response to credit market failures. It is argued that market imperfections are more important in developing economies at present for a variety of reasons. In developing economies such as in sub-Saharan Africa informational sharing mechanisms tend to be small scale and localised, markets are tightly interlinked, low levels of wealth limit the provision of collateral and there are few scale economies [Bardhan and Udry, 1999; Besley, 1994; Gosh et al., 1999; Ray, 1997]. In these circumstances, informal credit arrangements have an advantage as they exploit low transaction costs [Kochar, 1997; Udry, 1990], screening is performed through established relationships with borrowers [Aleem, 1990], and credit contracts are flexible and customised with a chance to renegotiate repayments [Baydas et al., 1995].

The cultural or sociological approach, by contrast, sees informal institutions as far less purposive than rational individuals engaged in maximising behaviour within some constraints [Aryeetey and Udry, 1995; Azam et al., 2001; Fafchamps, 2002; Fafchamps and Lund, 2003; Platteau, 2004; Udry, 1990]. According to this theory, norms of reciprocity, intergenerational altruism and obligation involve households without having been consciously devised [Granovetter, 1995].

Despite the numerous financial reforms aimed at facilitating the diffusion of formal credit institutions in developing countires, we still observe the coexistence of formal and informal credit arrangements. The literature typically focuses on two research areas, the "spillover" or "residuality" theory and the markets segmentation theory.

This thesis specifically tests the "spillover" theory maintaining that the informal sector exists to satisfy the unmet demand for credit resulting from credit rationing in the formal sector [for example, Banerjee and Duflo, 2001; Bell et al., 1997; Besley, 1994; Bose and Cothrem, 1997; Eswaran and Kotwal, 1989].

On the other hand, according to the market segmentation theory the informal sector may be the preferred source of credit for its unique characteristics, for the social preferences of the borrowers and for the specific purpose it is used [Barslund and Tarp, 2006; Mohieldin and Wright, 2000].

The relative advantage of the informal sector over formal institutions may be an object of concern as it can cause market inefficiency. This motivation together with distributional issues, vulnerability and poverty reduction call for government interventions in credit markets. We look at two policies that could address these issues. The first endeavours to create links between local moneylenders and banks. The second intervention creates government-sponsored microfinance institutions. This thesis specifically tests the effectiveness of the latter policy.

Chapter three is the first empirical essay and addresses the following research question: "Why do households participate in informal credit institutions?" The chapter uses as its primary source panel household data from the Ethiopian Rural Household Survey (ERHS, 1994-1997). The contribution is to build a unified empirical model capable of overcoming several limitations of the literature on this topic. We argue that the endogenous switching regression model with principal components is able to identify the following groups of factors that affect participation in informal credit.

The first group - household-based determinants such as wealth and demographic characteristics - has been largely discussed in the literature [for example, Bose, 1998; Kochar, 1997; Pal, 2002; Ravi, 2003; Ray, 1997]. However, a limitation of these studies is that a high degree of collinearity between household-specific variables limits the significance of individual regressors. We overcome this problem by constructing principal components of wealth variables.

The second group - idiosyncratic and aggregate shocks - has been analysed in the literature as a motive for participation in credit markets [Bardhan and Udry, 1999; Binswanger and Rosenzweig, 1993; Platteau and Abraham, 1987; Ruthenberg, 1971; Townsend, 1994]. However, data availability limits the identification of different types of shocks which may affect access to credit. The rich data in the ERHS allows for the distinction between aggregate and idiosyncratic shocks, the former operating at the cluster level and the latter at the household level.

The third group - cluster-based determinants such as demographic, infrastructural and geographical characteristics - is often ignored in the literature due to limited data and lack of appropriate models able to identify such characteristics. Knowledge of these cluster-level determinants is as important as knowing why households utilise such institutions in clusters where they are available. With access to the village studies provided by the ERHS, we have been able to identify dimensions of heterogeneity of access - most notably social, geographic and economic characteristics - which may operate at a cluster level, but are not identified at a household level [e.g. Fafchamps and Gubert, 2007]. The endogenous switching regression specification allows us to model the demand for a particular type of informal credit as endogenously determined by household-based and cluster-based determinants. Then, the access to informal credit is allowed to vary across endogenously different clusters.

The fourth chapter is a policy-oriented empirical essay answering the following question: "Do governments displace the informal loan market by introducing formal credit institutions?" A policy that arises in response to market failures (one of the causes for the diffusion of informal credit) aims at creating microfinance institutions that will acquire information in innovative ways. By mimicking and exploting some of the features of informal lending, banks can design credit contracts that harness local information and give borrowers incentives to use their own information on their peers to the advantage of the bank [Armendariz and Morduch, 2005; Ray, 1997].

This essay evaluates the effectiveness of this policy by testing whether microfinance institutions actually crowd out access to informal loans in Malawi. We adopt propensity score matching to identify a causal relationship between access to formal credit programmes and a reduction of informal borrowing. Propensity score matching is implemented to match participants in microfinance programmes with households that have similar observed characteristics (the so-called control group) and have been past participants, but are not current members. We use the Malawi Rural Financial Markets and Household Food Security (FMHFS, 1995), a rich survey containing information about households' borrowing behaviour.

The chapter introduces several innovations to the literature on crowding out. First, few empirical studies have tested the crowding out hypothesis in the context of group-lending institutions [for example, Mckernan et al., 2005].

Second, following the evaluation literature on training programmes [for example, Brodaty et al., 2001; Frölich et al., 2004], we develop a model with multiple treatments where households are classified as members of one, or more than one, group-lending programme. This approach allows for a comparison between the effectiveness of different credit programmes as well as between different groups of households. Does crowding out differ with the economic status of the household? In particular, are relatively constrained (unconstrained) households more (less) likely to reduce borrowing from informal lenders [Cox et al., 1998; Cox and Jimenez, 2005; Navajas et al. 2003]?

Third, nearly all the literature has focused on crowding out in the context of realised transfers. Yet households' demand for informal loans is also affected by the membership in a microfinance programme not just by the actual borrowing [Cox and Fafchamps, 2008]. We evaluate the effects of both being a borrower and a member of microfinance programmes.

Fourth, most of the literature is only concerned with crowding out of the supply of informal loans. This chapter disentangles demand and supply by employing outcome variables such as demand and credit limit of informal loans[2]. Such detailed data is uncommon in many developed and developing countries.

Finally, we develop a rigorous sensitivity analysis by adopting a number of matching algorithms and by testing for hidden biases arising from unobservable factors that affect simultaneously the assignment into one of the programmes and the outcome variable.

Chapter five is the third empirical essay and addresses the question: "Why do formal

[2]The credit limit variable is extensively explained in chapter five. As it refers to the maximum amount the borrower thinks the lender is able to lend, it can be thought of as being the "supply" of informal loans.

and informal credit markets coexist?" In spite of recent financial liberalisation aimed at broadening formal credit markets and in spite of interest rate differentials, in sub-Saharan Africa formal and informal credit institutions persist in the same market. The aim of this chapter is to motivate the partial crowding out effect found in the previous essay. By using information on the credit limit provided in the Malawian survey we test the spillover hypothesis, that is, the informal sector arises from a spillover demand from the rationed formal sector.

This chapter also tests the liquidity constraints hypothesis, that is, an increase in the credit limit should also affect the demand of liquidity constrained households. As the spillover effect results from the existence of liquidity constraints, the spillover and the liquidity constraints hypotheses are linked together.

The chapter makes several contributions to the literature. First, it extends Diagne (1999) and Diagne et al. (2000) approach by differentiating credit limits supplied by one or more credit programmes.

Second, unlike previous studies that adopt a reduced form specification in which demand and supply are collapsed into a single variable, we disentangle demand and supply equations in two ways. The data set allows for the identification of the demand equation and the supply equation (which is the credit limit equation) for both applicants and non-applicants to formal and informal lenders. In addition, following Diagne (1999) and Grant (2007) we apply a number of exclusion restrictions to identify demand and supply equations such as seasonal dummies and village characteristics.

Finally, we perform several robustness checks by addressing specification issues that may seriously affect the results (for example, heteroskedasticity, non-normality and selectivity).

Chapter six concludes the thesis with a brief summary of the findings. Limitations of the approaches adopted in the thesis are also discussed. Finally, the chapter provides some concluding remarks.

Chapter 2

Literature review

2.1 Introduction

In light of the research objectives outlined in the previous chapter, the literature review is focused on the following issues. First, it gives an overview of credit market institutions in Africa. Informal and formal credit arrangements are discussed in detail with specific reference to Ethiopia and Malawi, the two countries on which this thesis is focused.

Second, this chapter describes two motives for credit highlighted in the literature: risk-coping and acquisition of durable goods.

Third, the literature review proceeds with an analysis of the motives for demanding informal credit. It specifically focuses on the economic or market failure approach and the sociological approach.

Finally, this chapter provides some motivations for government interventions in credit markets arguing in favour of the creation of microfinance institutions.

2.2 Credit market institutions in Africa

To illustrate the relative importance of formal and informal credit, figure 2.1 shows the percentage share of formal and informal loans in selected countries in Africa and Asia by way of comparisons between poor and non-poor households (except for Ethiopia where rural and urban households are reported). The poor belong to the lowest quartile of income (or consumption expenditure) in their respective countries. The non-poor are the three other quartiles. Most of the loans originate from informal sources (with the exception of Malawi and Ghana[1]).

The dominance of informal credit sources is especially evident amongst poor households. In comparison to non-poor households, the poor obtain a smaller share of their loans from the formal sector in five countries (Egypt, Madagascar, Malawi, Nepal and Pakistan). The poor obtain a similar share of formal loans in Cameroon and obtain a marginally larger share of formal loans in Bangladesh and Ghana.

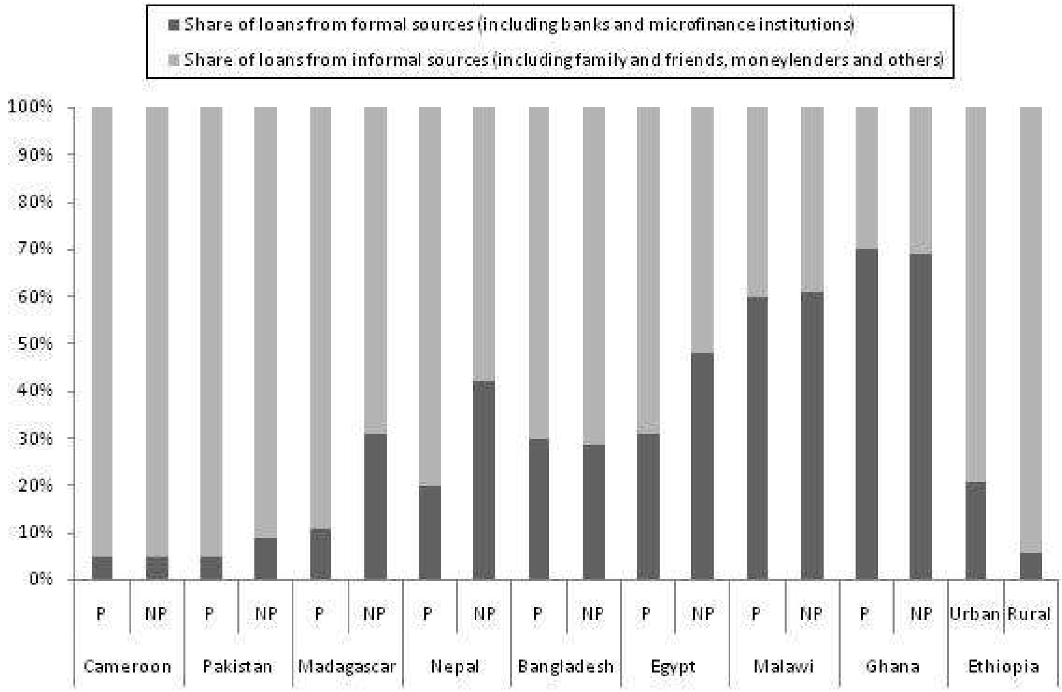
The importance of informal credit arrangements remains high even in Egypt where formal financial institutions are relatively widespread. In Ethiopia the informal sector is paramount especially in rural areas and in Bangladesh, group-based credit programmes play a significant role in providing credit to the poor. The villages selected for the surveys in Ghana and Malawi benefited from government-sponsored credit programmes which may explain the relatively higher proportion of formal loans.

Typically the literature defines informal institutions as those made up of a set of behaviours based on socially-shared rules, usually unwritten, that are enforced outside officially-sanctioned channels[2] [Helmke and Levitsky, 2003; Pejovich, 1999]. For instance,

[1] The data has been taken from studies on household surveys conducted by the International Food Policy Research Institute [Zeller and Sharma, 1998] and by the University of Addis Ababa and Oxford University for Ethiopia.

[2] Some question the need to distinguish between formal and informal institutions. Ostrom (2005) suggests the use of the term "shared strategies" since people respond to the known rules in the same way whether they are enforced by formal institutions or not.

FIGURE 2.1: Share of formal and informal loans in selected countries in Africa and Asia



Note: "P" and "NP" refer to poor and non-poor households respectively. Source: Zeller and Sharma (1998), Ibrahim, Kedir and Torres (2007) and our own calculations from the ERHS.

informal finance consists of often unrecorded lending activities that take place outside formal financial institutions.

Informal financial institutions vary in their operational features: some are community or group-based whilst others are individual. They vary in their scope - some are involved in either savings or lending, whilst others are involved in both. They can involve cash, payment in-kind transactions or both. Moneylenders, friends and relatives, rotating savings and credit associations (RoSCAs) and self-help groups are examples of informal credit institutions.

Numerous studies have shown that informal financial arrangements across Africa exhibit diversity [Aryeetey and Hyuha, 1991; Chipeta and Mkandawire, 1991; Soyibo, 1996]. Each informal arrangement typically covers a limited geographical area and often takes place among people linked by contracts (i.e. landlord-tenant), among kinsmen or people in the same locality. On the other hand, formal finance comprises of institutions regulated by the government and the Central Bank operating within the regulatory framework of the financial system and generally provides services on a more geographically dispersed basis. Examples include commercial banks, agricultural banks and government-sponsored microfinance institutions.

The large spectrum of credit institutions can be stratified according to two criteria. The first stratifies according to an increasing level of formality as referred to above. The second criterion refers to the degree of social cohesion, ordered in a roughly decreasing level of lender-borrower closeness and exogeneity of the lending methodology [Robinson and Schmid, 1988; van Bastelaer, 2000; Woolcock and Narayan, 2000].

FIGURE 2.2: Segments of financial systems

Tier	Definition	Institutions	Principal Clients
Informal institutions	Unregistered, with unwritten socially-shared rules	Friends & relatives	family members or friends with "obligation" to reciprocate
		Mutual help associations (i.e. RoSCA-type groups)	usually village members for burial activities or to finance durables
		Moneylenders	tenants or clients within the geographical area covered by the moneylender
Formal institutions	Licensed by the Central Bank and regulated by the financial system	Group-lending	poor groups in a village
		Agricultural Banks	businesses for agricultural investments
		Commercial Banks	businesses, corporations and government

Source: Own classification.

Figure 2.2 illustrates informal and formal credit institutions ordered in an increasing degree of formality and decreasing level of social cohesion. In this context, informal institutions appear at the top of the scale due to their low level of formality and high degree of social cohesion. As formal institutions have a high level of formality and low degree of social cohesion, they appear at the lower end of the scale of credit institutions.

2.3 Informal credit institutions

2.3.1 Friends and relatives

Friends and relatives provide sources of financial (and other) help through the highest degree of social cohesion. The fact that both lender and borrower know each other dispenses with the key features of formal credit transactions such as ensuring credit worthiness or demanding collateral and guarantees. Usually loans are supplied without interest repayments or regular repayment schedules and transaction records are not made. Such lenders provide loans on a need-based manner [Fafchamps, 2008]. Sanctions often include denial of future loans or other social costs (such as “bad” reputation within the community) in case of default.

A disadvantage of this source of lending is the limited and often irregular supply of loans. That is, only when the individual has surplus funds is the loan made.

Credit given by friends and relatives often implies an obligation for future reciprocation. Thomas and Worrall (2000) note that if the costs of giving are covered by the perceived benefits of future reciprocity, then often these forms of informal credit are likely to be more effective than other informal or formal institutions.

The importance of these types of informal networks is now recognised especially in small communities. For example, Sahlins (1972) reported a mechanism of “generalised reciprocity” in which those with high income help those with low income. Similar mutual help contracts have been described by Platteau and Abraham (1987) who found evidence of reciprocal credit among a community of fishermen in an Indian village. Udry (1990) found evidence of credit with repayments contingent on the realization of production in Northern Nigeria. Reciprocity may be stronger among ethnically homogenous groups, family, clan or religious affiliations because these groups can threaten to impose larger punishments on individuals breaking the mutual insurance arrangements.

Gächter and Herrmann (2009) demonstrate with laboratory experiments in Russia and Switzerland that many people are "strong reciprocators" who cooperate and punish others even if there are no gains from future reciprocity or other reputational gains. They show that patterns of strong reciprocity can be explained by cultural differences across the two countries.

2.3.2 Mutual help associations

Further down the scale of credit institutions, there are a series of mutual, often local, helping associations. For example, in Ethiopia there are a number of mutual assistance associations called *iddir* and the RoSCA-type called *equb* [Aredo, 1993; Mauri, 1987].

i) RoSCAs

The nature of RoSCAs was originally analysed within the anthropological literature. Geertz (1962) described RoSCAs as a "middle-rung" institution. He defined it as "a shift from a traditionalistic agrarian society to an increasingly fluid commercial one", and as an "educational mechanism in terms of which peasants learn to be traders, not merely in the narrow occupational sense, but in the broad cultural sense". Regarding their functionality, Geertz (1962) pointed out that RoSCAs are "... a lump sum fund composed of fixed contributions from each member of the association which is distributed, at fixed intervals and as a whole, to each member of the association in turn". Ardener (1964), however, argued that this definition is too restrictive and defined RoSCAs as "an association formed upon a core of participants who agree to make regular contributions to a fund which is given, in whole or in part, to each contributor in rotation". The fund or "pot" is allocated to one member by drawing or bidding. At the end of each round past winners are excluded from receiving the pot.

Although little is known about their origin, RoSCAs are not restricted to any country. Ardener (1964) had already reported a well-developed RoSCA in Asia at the end of the nineteenth century. In Ghana, RoSCAs are found in larger towns but not in rural areas. In Egypt, RoSCAs are known as *gameya* and have existed for more than fifty years. Rural areas see membership confined to women, but in urban areas men, women and children belong to them [Ardener, 1964].

The principal function of RoSCAs is to assist capital-formation, or more simply to create savings. As some individuals feel that they would struggle to save if they were not committed to such a group, contributions can be seen as a form of forced saving. In addition, women may see it as a way to prevent their husbands using family savings for personal consumption (for instance on alcohol or cigarettes).

In relatively recent years, RoSCAs have been subject to a more formal economic analysis. For example, Besley et al. (1993) analyse the economic rationale of RoSCAs. This paper compares the random and bidding RoSCAs. In the former "people commit to putting fixed sum of money into a "pot" for each period of life of the RoSCA. The pot is randomly allocated to one of the members. In the next period, the process repeats itself, except that the previous winner is excluded from the draw of the pot. The process continues until each member of the RoSCA has received the pot once".

The bidding RoSCA is similar to the random RoSCA, except that the pot may be obtained earlier if one member bids more than the others. "The bidding process merely establishes the priority".

Besley et al. (1993) show that both random and bidding RoSCAs improve members welfare compared to the autarky level. However, in the bidding RoSCA each member has a different rate of nondurable consumption during the accumulation period (i.e. those who get the pot earlier, make higher contributions and consume less of the nondurable; the last member getting the pot makes no contribution and must have greater nondurable consumption during accumulation than under autarky). Moreover, Besley et al. (1993) demonstrate that random RoSCAs are better than bidding RoSCAs because the random allocation dominates ex ante the bidding one (ex post this may not be true for the last "bidder").

Besley et al. (1996) show that RoSCAs in Taiwan allow members to reduce the time to acquire a durable good. This thesis pins down the factors affecting the formation of RoSCAs in rural Ethiopia.

ii) Iddirs

RoSCAs are only one of a range of indigenous voluntary organizations and associations existent in developing countries. For instance, in rural Ethiopia the most widespread self-help association is the *iddir*.

Iddirs are indigenous voluntary associations primarily established to provide mutual aid in burial matters, but also to address other community concerns such as financial needs in case of poor health conditions [Pankhurst and Mariam, 2000]. Household members pay monthly fixed contributions. Whenever a member of the *iddir* dies, the association uses the money for the ceremonial expenses. Since their introduction at the beginning of the twentieth century, *iddirs* have become more formalized. They involve regular meetings, they have a chairman or "judge" and there are well defined rules to regulate how funds will be collected and disbursed. With regard to membership structure, *iddirs* are open to anyone regardless of socio-economic status, religion, gender and ethnic affiliation. Many *iddirs* help members who face economic problems by giving them benefits without requiring any contribution. These associations have appropriate incentives and enforcement techniques because they are well integrated within the local communities. For example, a person who does not belong to an *iddir* is considered a disgrace to his or her family. In comparison to some formal sources, RoSCAs and *iddirs* are less impersonal.

2.3.3 Moneylenders

Moneylending is characterised by a more exogenous lending methodology and often by a lower level of social cohesion. Stiglitz (1990) noted that "the local moneylenders have one important advantage over the formal [lending] institutions: they have more detailed knowledge of the borrowers. They therefore can separate out high-risk and low-risk borrowers and charge them appropriate interest rates". Moneylenders provide flexible contract terms and dispense with the need for collateral due to the information they possess on borrowers. They usually charge high rates of interest in comparison to formal lending institutions.

Moneylenders may borrow from banks during high demand for credit by using their own funds as security, thus creating a channel where formal funds are injected into the informal sector.

Mansuri (2007) reported that often moneylenders' primary activity is not lending: loans are means of obtaining a return on other transactions in which both lender and borrower are involved. The interweaving of activities between borrowers and lenders allows the lender to gather information about the borrowers' ability to repay. The relationship between moneylender and borrower is reminiscent of a patron-client vertical interaction. It is intrinsically unequal as the moneylender has access to several methods (such as lowering the wage if the moneylender is also the employer) to ensure repayment [van Bastelaer, 2000]. Badhuri (1973) observed that perpetual indebtedness of the borrower as a consequence of high interest rates is characteristic of a semi-feudal environment. The loan is used as a way to secure asset transfers or long-term relationships with the borrowers.

2.4 Formal credit institutions

As stated above, formal or institutional lenders can be placed further down in the scale of credit institutions when classified in this way due to their low degree of social cohesion.

2.4.1 Group-lending

The subsector of formal institutions closest to informal credit arrangements is microfinance[3]. It is usually based on the group-lending approach that assists those poor designated as "safe" borrowers (i.e. able to repay small uncollateralized loans). Microfinance uses a lending methodology that relies on traditional and personal interactions among borrowers. Hence, group-lending relies on a similar level of social cohesion that forms the basis of RoSCAs[4]. Ghatak (1999) suggested that group-lending institutions that use joint liability schemes can deal with the major problems faced by institutional lenders using local information and social sanctions[5]. For example, positive assortative matching (borrowers match with their same "type" and they form homogeneous groups) allows group members to reduce the risk of default by one (or more) of them; dynamic incentives facilitate enforcement of payments when a defaulting member is excluded from future loans.

[3]"Microfinance" is a more recent concept than "microcredit". It was developed in the 1990s to indicate both the microsaving and the microcredit components of a financial service.

[4]Some would classify group-lending either as an informal or as a semi-formal credit institution. In light of the above definition, in this thesis we consider group-lending a formal institution as it is regulated by the government with a set of codified rules.

[5]Joint liability refers to the fact that if one group member defaults, the other members are liable to repay the loan.

However, there are also some disadvantages to group-lending, for example group size decreases once social sanctions are applied [Impavido, 1998]. Second, the degree to which group members know each other and interact on a regular basis also affects the performance of the group. Third, group repayments are negatively affected by aggregate shocks (i.e. shocks that affect all members of a community).

Although the group-based approach has developed in the 1970s, the concept is a century old. Ghatak and Guinnane (1998) and Woolcock and Narayan (2000) pointed out the existence of a German credit cooperative in the mid-nineteenth century. Today the most studied example of group-based lending is the Grameen Bank in Bangladesh. The Grameen Bank was founded in 1976 by Mohammad Yunus, a professor at the University of Chittagong, as a research project. By 1994, the Grameen Bank had served half of all villages in Bangladesh, with a total membership of more than two million, of which 94 percent were women. It uses group-lending and joint liability schemes where small uncollateralized loans are repaid in weekly instalments. If any member of the group defaults, the whole group is denied future credit. Using this approach, the Grameen Bank has consistently reported repayment rates in excess of 95 percent.

Since its foundation, the Grameen Bank model has been exported to many countries throughout Africa, Latin America and Asia. It was replicated in Malawi (one of the countries on which this thesis is focused) in 1987 when the World Bank and the International Fund for Agricultural Development (IFAD) funded the Mudzi Fund. Another replication of the model was founded in 1986 when Bolivian business leaders established a non-profit microlending entity called PRODEM. In 1992, PRODEM became Bancosol after a privatisation process. By 1997 Bancosol was the first microfinance institution to issue dividends to shareholders.

2.4.2 Agricultural banks

Agricultural and commercial banks are the credit institutions with the lowest degree of social cohesion and higher level of formality.

Agricultural banks were created in low-income countries after World War II in an attempt to develop the agricultural sector. Large state agricultural banks were subsidised to induce farmers to irrigate, apply fertilizers, and adopt new crop varieties and technologies. The goal was to increase land productivity and labour demand, thus pushing up agricultural wages.

Critics of the agricultural state banks argue that subsidized credit failed to improve the well-being of poor households for the following reasons [Armendariz and Morduch, 2005]. First, the interest rate acted as a rationing criterion: only those with the most worthy projects were willing to pay for credit. The rationing mechanism broke down when the subsidized interest rate fell below the market rates of interest. In this context, credit was allocated to unproductive recipients. Note that this is the opposite of the Stiglitz-Weiss'[6] (1981) "story" where rationing excludes the most risky projects.

Second, because of the subsidized funds flowing from the government, bankers had no incentives to collect savings deposits. Poor households were thus left with relatively inefficient saving mechanisms.

Third, state banks were inevitably linked with the political process and thus tended to forgive repayments before the elections. This allowed the powerful access to cheap funds which were meant for the poor and removed incentives previously created to build efficient institutions.

Finally, critics argue that credit is a fungible financial tool and should not be delivered as a specific input into a particular production process (for instance to buy fertilizers

[6]The Stiglitz-Weiss' model is discussed in more detail later on.

for farm production).

On the other hand, recent empirical work by Burgess and Pande (2005) showed net positive average impacts of India's Integrated Rural Development Programme (IRDP) on the poor. According to Burgess and Pande (2005), the expansion of access to informal finance enabled people to increase non-agricultural production activities. As the economic returns from these activities were higher than those from agricultural activities, the IRDP was able to reduce rural poverty. Nevertheless, the programme was ended in 1990 because the expansion of rural bank branches was too expensive. High default rates and subsidised interest rates are testimony of the fact that rural branches were a policy vehicle for costly redistribution of resources to rural areas.

Binswanger and Khandker (1995) found that between 1972-1973 and 1980-1981 state agricultural banks in India had increased rural wages and employment. However, as they found only modest impacts on agricultural output, they concluded that the costs of such government programmes were much higher than the economic benefits.

In 1970 the Agricultural and Development Bank was established in Ethiopia. It is government-owned and provides short term loans to the agricultural sector, medium and long-term loans to individuals, cooperatives and agricultural projects as well as special credit lines for microenterprises. Additionally, the Agricultural and Development Bank offers banking services like current and saving accounts.

2.4.3 Commercial banks

Private, domestic commercial banks are a relatively recent phenomenon in many developing countries, especially in Africa. From the 1950s to the 1970s, banks were predominantly owned by the government or by other foreign commercial banks. The existing local banks were typically relatively small and often served a closed set of business groups.

In most developing countries, up to the 1980s, it was the highly regulated formal financial markets that were responsible for the inadequate development of privately-owned commercial banks due to their interest rate ceilings, high reserve requirements and directed credit lines. Banks could not charge sufficiently high interest rates to cover the costs and risks of lending to a large clientele.

In the 1980s the financial liberalization process allowed private domestic commercial banking to expand rapidly. New private banks were used to obtain funds for businesses and corporations.

At that time the government of Malawi, for example, implemented measures to liberalize its financial sector. Reforms included the elimination of agricultural subsidies and interest rate controls along with the removal of exchange control regulations and restrictions on capital movements. This liberalization process has increased the number of players in Malawi's financial markets. It has eight commercial banks providing savings, lending and other investment products. Two major banks, the National Bank of Malawi and Stanbic Bank, dominate the financial sector with 58 percent of the sector's assets and 59 percent of its deposits[7]. The National Bank of Malawi is predominantly owned by companies with significant government shareholdings. The Standard Bank of South Africa now holds a 60 percent shareholding in Stanbic Bank (formerly known as the Commercial Bank of Malawi when it was controlled by the government).

In Ethiopia, the financial liberalization process initiated in 1992 was much less radical than elsewhere in Africa. The commitment to continued government ownership of existing financial institutions was still strong and the government was reluctant to allow foreign banks in Ethiopia.

[7]This data has been reported by the United Nations Capital Development Fund, 2006.

Since the financial reforms began, new financial institutions have been allowed to operate in Ethiopia. Six private banks and eight insurance companies now operate alongside public ones. Although the government-owned Commercial Bank of Ethiopia (CBE) remains the country's largest commercial bank, its dominance is declining as private banks and competition from international banks grow.

From 1994 the CBE obtained greater autonomy in its lending decisions and acquired its own Board of Directors. A few years ago, the government restructured the CBE and signed a contract with the Royal Bank of Scotland for management consultancy services. In January 2009, the Commercial Bank of Ethiopia received regulatory approval to open a branch in Southern Sudan.

2.5 Why do households demand credit?

There are two primary motives for households seeking credit. First, households use credit to cope with shocks that may appear in their lifetime ("the risk motive"). Secondly, credit may be required to purchase "lumpy" assets - typically, durable goods. Whilst the empirical analysis carried out in this thesis explicitly focuses on the risk motive for credit, the following sub-sections will analyse each of these motives in turn.

2.5.1 Risk

Risk pervades all of life's activities. It causes fluctuations in income and health. These fluctuations can be predictable or unpredictable. Risk can affect us individually (so-called *idiosyncratic* shocks such as illness) or can affect the entire community (so-called *aggregate* shocks such as natural disasters or fluctuations in prices that affect the entire economy).

Although risk is paramount in all societies, the types of risk to which poor rural economies are exposed, are quite different from those that can be observed in developed countries. For instance, in developing economies there is a higher incidence of infectious diseases and natural disasters. On the other hand, the impact of business cycle fluctuations, technological obsolescence and stock market fluctuations are less severe in poor rural economies [Fafchamps, 1999].

In order to better understand the link between risk and credit demand, it is necessary to consider risk-coping strategies. There are two mutually non-exclusive ways of dealing with income and health fluctuations: a) managing risk before income shocks occur (*ex ante* risk management) through income smoothing mechanisms; and b) coping with risk *ex post* through intertemporal consumption smoothing and risk sharing strategies.

In the absence of perfect insurance markets, households may adopt *ex ante* strategies to reduce the variability of income. The choice of occupation according to expected earnings and strategic migration of family members can be considered income smoothing strategies. In an agricultural economy risk management strategies might include crop and field diversification.

Ex post risk coping strategies involve intertemporal consumption smoothing (by saving and borrowing) and risk-sharing mechanisms (self or mutual insurance). The primary distinction between these two strategies is that intertemporal smoothing enables the household to attenuate the effects of income shocks on consumption over time. Risk-sharing, by contrast, spreads the effects of income shocks across households. Thus, risk sharing can be viewed as the cross-sectional counterpart of intertemporal consumption smoothing [Cochrane, 1991].

Intertemporal consumption smoothing may be achieved by accumulating and selling assets and also by storing goods for future consumption [Alderman and Paxson, 1992].

Risk-sharing may be accomplished through formal institutions and informal arrangements. Examples of the former include insurance and futures markets whilst examples of the latter include state-contingent transfers and remittances between friends and relatives.

2.5.1.1 Intertemporal consumption smoothing

Intertemporal smoothing allows consumption to be insulated from the effects of income fluctuations by using saving and credit transactions. In Friedman's permanent income hypothesis (PIH) model, consumers try to smooth out spending based on their estimates of permanent income [Friedman, 1957]. Only if there has been a change in permanent income will there be a change in consumption. Indeed, the PIH states that transitory changes in income do not affect consumer spending behaviour in the long run.

The permanent income hypothesis (PIH) is derived from a partial equilibrium model which involves a representative household, taking prices as given. A representative household maximises expected utility subject to a constraint where the household receives a random income y and decides how to allocate its resources between consumption and net saving for the next period. The solution to the problem is given by the following equation:

$$E\left\{\frac{u^{'}(c_{t+1})}{u^{'}(c_t)}\right\} = \frac{1}{\beta(1+r_t)} \tag{2.1}$$

where r is the interest rate and β is the discount factor (bounded between zero and one).

Equation 2.1 shows that the intertemporal ratio of marginal utilities depends on the discount factor and the interest rate. Whenever $\beta(1 + r) = 1$, the marginal utility of consumption is a martingale process. Households save over time in order to create a buffer stock for precautionary reasons. In addition or alternatively households borrow from credit markets.

If households experienced an adverse shock, the average propensity to consume would increase. In the light of the uncertainty surrounding the adverse shock, households would be prepared to take loans, even with high interest rates, to get through the bad period.

As referred to above, households could also cope with risk by creating buffer stock savings. Following Deaton (1992), suppose that the marginal utility of consumption is convex. Note that the convexity of the marginal utility (third derivative of the utility) tells us how prudent households are. This concept is different from the degree of risk aversion (second derivative of the utility). Only in the special case of iso-elastic utility are the two concepts equivalents. In addition, assume that the variability of consumption increases, thus creating more uncertainty. The increase in (mean-preserving) spread will increase the expected value and the marginal utility of consumption. As a consequence, consumption decreases and savings increase. When households are more prudent, an increase in uncertainty enhances precautionary savings [Banks et al., 2001].

The PIH has been criticised for the assumption of perfect capital markets [Maki, 1993]. When capital markets are imperfect, the optimal consumption path is different to the one specified in the PIH because households cannot create a "cushion" of marketable assets and do not have the capacity to borrow up to the value of prospective lifetime wealth against future earnings [Ishikawa, 1974; Pissarides, 1978].

The deviation from the life cycle (LC) and permanent income hypothesis (PIH) has been used to indirectly infer the presence of credit constraints. One of the testable implications of the LC/PIH is that in the absence of liquidity and borrowing constraints, transitory income shocks do not affect consumption [Deaton, 1992; Hall, 1978].

Empirical tests for the presence of credit constraints based on the LC/PIH use household consumption and income data to look for a significant dependence (or "excess sensitivity") of consumption on transitory income. Evidences of liquidity constraints as a result of imperfect capital markets have been provided in both developed and developing countries as shown below.

However, the LC/PIH approach to detect credit constraints may be inconclusive. First, deviations from the LC/PIH can result from prudent or cautionary behaviour even if the borrower is not credit constrained [Carroll, 1991; Kimball, 1990; Zeldes, 1989b]. Secondly, if conditions of uncertainty are negatively correlated with wealth, then current income will be negatively correlated with consumption growth even without borrowing constraints [Carroll, 1991]. Finally, Deaton (1990) pointed out that the effect of income shocks on consumption also depends on the initial asset position of the borrower. Hence, deviation from the LC/PIH is neither a sufficient nor a necessary condition for being credit constrained.

A relatively voluminous literature on developed countries has linked the failure of the PIH to the presence of liquidity constraints [Bernanke, 1984; Hall and Mishkin, 1982; Hayashi, 1987; Jappelli and Pagano, 1989; King, 1986; Zeldes, 1989]. For instance, Jappelli and Pagano (1989) found that countries characterised by high excess sensitivity of consumption to current income are also those where consumers borrow less from capital markets. Italy, Spain and Greece are examples of countries with high excess sensitivity and Sweden and United States have a low excess sensitivity. They concluded that the low levels of consumer debt observed in countries where the excess sensitivity of consumption is high can be interpreted as evidence that liquidity constraints are at the root of the empirical failures of the LC/PIH in time-series tests.

Several studies in developing countries have rejected the PIH [Morduch, 1992; Paxson 1992; Rosenzweig and Binswanger, 1993]. Paxson (1992) has shown that deviation in average rainfall is reflective of transitory income shocks affecting Thai rice farmers. She used the deviation from average rainfall to calculate the marginal propensity to save transitory income. Households saved around three-quarters to four-fifths of transitory income which is less than the marginal propensity to save predicted by the PIH (which would be equal to one). Morduch (1992) has found in the International Crop Research Institute for the Semi-Arid Tropics (ICRISAT) data that consumption smoothing is real and significant for the comparatively better off households, while landless and small farmers do not show the same pattern. Rosenzweig and Binswanger (1993) showed that poor households are more constrained in their ability to insulate their consumption from income risk. The literature provides several explanations for this.

First, the lack of collateral and the high transaction costs limit poor households' access to credit markets. Credit market imperfections result in collateralised lending which creates difficulties for asset-poor households [Eswaran and Kotwal, 1989]. In addition, the presence of fixed transaction costs per loan makes borrowing harder for poor households [Morduch, 1995].

Secondly, the scarcity and indivisibility of assets, together with the fixed costs of storage limit poor households' ability to save. Access to relatively safe and profitable assets is often limited. The lumpiness of assets causes intertemporal consumption smoothing to be harder. For example, during the 1984-1985 famine in Ethiopia, prices collapsed because many households were selling assets. Whenever a common negative shock occurs, incomes are low and returns on assets are also low. The covariance between asset values and income due to common shocks makes consumption smoothing more problematic for low income households [Dercon, 2002].

2.5.1.2 Risk-sharing

From the previous sub-section it has emerged that intertemporal consumption smoothing is more problematic in many developing economies where collateralised lending limits the access to credit markets, credit rationing is pervasive and where income shocks are correlated with asset prices (such as livestock). The existence of liquidity constraints affects the ability of households to transfer resources across time periods, as well as across uncertain states of nature, relative to income. As a result, consumption (and thus saving) tends to be highly correlated with current income, rather than permanent income.

Intertemporal consumption smoothing, however, is not the only strategy that households can adopt to cope with risk *ex post*. Households can risk share with unknown economic agents through private or government insurance schemes, or through participation in financial markets. Alternatively they can protect consumption against income fluctuations by sharing risk with friends and kin.

The formal insurance schemes analysed within the literature in developed countries typically take the form of bankruptcy laws [Fay et al., 2002], insurance within a firm [Guiso et al., 2005], government public policy programmes such as unemployment insurance [Engen and Gruber, 2001], Medicaid [Gruber and Yelowitz, 1999] and food stamps [Blundell and Pistaferri, 2003]. However, there is now strong evidence against complete consumption insurance provided by formal schemes [Attanasio and Davis, 1996; Attanasio and Weber, 1992; Cochrane, 1991].

In developing economies such as in sub-Saharan Africa where informational sharing mechanisms tend to be small scale and localised and the legal systems are weak, enforcement problems and information asymmetries severely limit the use of formal insurance schemes.

However, consideration can be given to the informal insurance mechanisms between kin groups, friends, relatives and members of a community and their ability to cope with risk.

Most analyses of risk sharing in a developing country context stem from Townsend's (1994) model of insurance in India. Consider a model with N households that live in the same village. There are T periods in which shocks may occur with a probability of π_s. Suppose that in each state of the nature, s, each household i receives an exogenous income, y_{is}, and consumes an amount c_{ist}. The utility function takes the following functional form:

$$U_i = \sum_{t=1}^{T} \beta^t \sum_{s=1}^{S} \pi_s u_i(c_{ist}) \tag{2.2}$$

and displays the usual properties: twice continuously differentiable and intertemporally separable. The Pareto efficient allocation can be thought of as a maximization problem of a social planner that gives a weight λ_i to each household i with $0 < \lambda_i < 1$ and $\sum \lambda_i = 1$:

$$\max_{c_{iht}} \sum_{i=1}^{N} \lambda_i U_i$$

$$\text{s.t.}$$

$$\sum_{i=1}^{N} c_{ist} = \sum_{i=1}^{N} y_{ist} \qquad \forall \quad i, s, t$$

$$c_{ist} \geq 0 \qquad \forall \quad s, t$$

The solution to the model gives:

$$\frac{u_i^{'}(c_{ist})}{u_i^{'}(c_{jst})} = \frac{\lambda_j}{\lambda_i} \qquad \forall \quad j, i, s, t$$

This implies that in the village there exists a co-movement of households' marginal utilities and consumption levels. In a Pareto-efficient allocation of risk within a community, households can achieve full (idiosyncratic) risk sharing and the only risk they face is aggregate risk.

In developing countries the full insurance hypothesis has been largely rejected [Deaton, 1992; Grimard, 1997; Morduch, 1995; Townsend, 1994; Udry, 1994]. For example, Deaton (1992) and Grimard (1997) analysed the patterns of consumption within villages in Cote d'Ivoire and found no evidence of full risk-sharing.

Grimard (1997) points out that the rejection of the full insurance hypothesis is due to its strong theoretical implication - namely, the fact that the household's entire consumption is determined by the group's aggregate resource constraint. According to the full insurance hypothesis, a household which unexpectedly enjoys a rise in its individual permanent income must share the entire rise with the community. But the full insurance hypothesis ignores the fact that moral hazard and enforcement costs may affect the outcome of the insurance scheme.

In the presence of non-competitive markets with information and enforcement obstacles, a Pareto efficient allocation cannot be achieved. However, households within a community, relatives or other social groups may share risk through informal arrangements that approximate the Pareto-efficient allocation of risk. In these circumstances, mutual insurance can be undertaken as the information amongst people is good, income is difficult to hide and behaviour can be monitored. As mentioned earlier, there is empirical evidence on the existence of these institutions in Thailand [Townsend, 1994], among fishing communities in Southern India [Platteau and Abraham, 1987] and northern Nigeria [Udry, 1990].

2.5.2 Durable goods

The second motive for credit is the purchase of durable goods [for example a car as in Attanasio et al., 2008]. Not only does the ownership of these goods yield a flow of consumption services over several periods, but also it improves households' wealth. The utility maximization problem in the presence of durable goods can be modified as follows [Bertola et al., 2006]:

$$\max E_t \sum_{j=0}^{\infty} \beta^j u(c_{t+j}, d_{t+j}) \tag{2.3}$$

where d represents the durable goods. An additional complication to the standard model is that durables are endogenous to the household's optimization problem. The budget constraint is modified to include the purchase of durable goods (given by g):

$$A_{t+1} = (1 + r_{t+1})(A_t + y_t - c_t - g_t) \tag{2.4}$$

where A is the level of assets, y is the income and r is the interest rate determined in the credit markets. The stock of durables d can then be modelled as the amount of goods at any point in time, plus new durable purchases, minus the depreciation.

The household's optimal plan involves equating the marginal utilities of consumption between periods and also equating the marginal utilities of durable and nondurable consumption.

In agrarian societies where there is a delay between the start of production and the realisation of output, credit transactions also serve to finance durables used for farm production (for instance the purchase of fertilizers or farm equipment).

The financial institutions that provide credit for the purchase of durables are varied and often complex[8]. There are the formal financial institutions mentioned above such as agricultural and commercial banks, and government-sponsored microfinance institutions. There are also specialist informal institutions such as moneylenders and mutual help groups.

The mutual help groups RoSCAs are specifically formed for the purchase of durables and have the advantage of reducing the time it takes to acquire a particular asset. Besley et al. (1993) analysed the economic rationale of RoSCAs and showed that where a group of individuals wish to gain access to an indivisible durable consumption good, and have no access to external finance, a RoSCA provides a means of realising gains from intertemporal trade. Besley et al. (1993) provided an example: "[...] consider 10 individuals each of whom wishes to own a durable that costs $100. Left to their own efforts, they can save $10 per week over 10 weeks. However, they can do better by pooling their joint savings. One (lucky) individual can get the durable after one week instead of waiting for 10 weeks. The same is true for the second individual etc. Only the last person would get the good in 10 weeks. This is a Pareto improvement as nobody will be worse off".

2.6 Why do households demand informal credit?

From the previous sections it has emerged that in developing economies formal credit markets and insurance are not as widespread as informal arrangements. The literature provides two main theories for the existence and diffusion of informal credit in developing countries, the *economic* approach and the *cultural* or sociological approach. The next subsections will outline these approaches as they will be the object of the empirical analyses of this thesis.

[8]See sections 2.2-2.4 for a more detailed description of the various financial institutions.

2.6.1 The *economic* approach: market failure

The *economic* approach maintains that informal finance arises as a response to credit market failures. The literature typically points out that loan contracts may be affected by adverse selection, moral hazard and enforcement problems.

First, the characteristics of a credit contract can select certain types of borrowers and, hence, may influence the distribution of lender's profit ("adverse selection" problem).

Second, the terms of the loan contract may also affect the performance of the borrower and, in turn, the distribution of lender's profits. The lender faces a "moral hazard" problem. When unobservable actions or efforts are taken by borrowers after the loan has been disbursed but before project returns are realised, the lender faces an *ex ante* moral hazard; on the other hand, when unobservable actions or efforts are taken by borrowers after the loan and projects are realised, the lender faces an *ex post* moral hazard or enforcement problem [Armendariz and Morduch, 2005].

Moral hazard and adverse selection problems make it difficult for the formal credit market to clear through prices. In an attempt to avoid default on lending funds, lenders ration the supply of credit.

There are two forms of credit rationing [de Meza and Webb, 2005]. First, at a given interest rate, applicants willing to take larger loans will be denied [Jaffee and Russell, 1976] and second at a given interest rate, amongst applicants who appear to be identical, there are inconsistencies in that some will get the loan and others will not [Stiglitz and Weiss, 1981].

Stiglitz and Weiss (1981) argued that the lender's interest rate has a dual role of sorting potential borrowers and affecting the actions of borrowers. As high interest rates attract fewer borrowers of worse quality, it is advantageous for lenders to set the interest rate as low as possible to be attractive to "good" borrowers. In this context, credit rationing arises not as a market "disequilibrium" but because lenders set interest rates to obtain the right "mix" of borrowers thus limiting the risk of default from "bad" borrowers .

Bester (1985) used a hidden information model to show that in equilibrium there is no credit rationing if banks compete by simultaneously choosing the rate of interest and collateral requirements used to evaluate the risk of a potential borrower. Borrowers are then sorted according to their riskiness through contracts that stimulate self-selection. For example, borrowers with a low probability of default are more inclined to choose a contract with a lower interest rate and higher collateral than borrowers with a high probability of default.

Stiglitz and Weiss (1981) emphasised that credit rationing arises especially where collateral is limited. However, de Meza and Webb (1987) take the view that even if collateral is limited, credit rationing will not arise if borrowers differ in ability rather than intrinsic risk. Furthermore, de Meza and Webb (2005) argued that credit rationing breaks down because decisions about the loan size or about the time at which the project starts are endogenous. For example, suppose that a borrower can reduce the loan amount through self-finance and that the interest rate remains the same with the reduced loan. The lower repayment will reduce moral hazard, making the borrower more attractive to the bank. Lensink and Sterken (2001, 2002) have applied this rationale to the situation in which a borrower can decide to delay the start of the project.

It is often argued that market imperfections and, consequently, credit rationing, are less important in developed economies in recent years for a variety of reasons[9]. Developed economies have appropriate informational sharing mechanisms (for example credit scoring) and methods for dealing with informational asymmetries in credit markets if and when they arise. The provision of collateral-based contracts and the existence of

[9]See Bertola et al. (2006) for a detailed description of the topic.

other instruments like credit bureaus are examples of these methods [Cole and Mishler, 1998; Jappelli and Pagano, 2003; Padilla and Pagano, 1997, 2000; Pagano and Jappelli, 1993, 1999]. Also, formal institutions can develop in such markets because of scale economies and the relative lack of vulnerability of credit markets to adverse economic shocks [Carpenter and Jensen, 2002; Hoddinott et al., 2000].

In contrast, in developing economies such as in sub-Saharan Africa, informational sharing mechanisms tend to be small scale and localised, markets are tightly interlinked, low levels of wealth limit the provision of collateral and there are few scale economies [Bardhan and Udry, 1999; Besley, 1994; Gosh et al., 1999; Ray, 1997]. In these circumstances, informal lending arrangements such as family and friends, and the development of local arrangements such as rotating saving and credit associations (RoSCAs) have an advantage. They exploit low transaction costs [Kochar, 1997; Udry, 1990], screening is performed through established relationships with borrowers [Aleem, 1993], and credit contracts are flexible and customised with a chance to renegotiate repayments [Baydas et al., 1995].

Kochar (1997) argued that informal loans, in particular those from friends or relatives, may be cheaper than formal loans and thus preferred by borrowers. Chung (1995) and Mushinski (1999) pointed out that high transaction costs in the formal sector may discourage households from taking formal loans. Barham et al. (1996) called these households "transaction cost-rationed" in the formal sector. As a consequence of the higher effective costs of formal loans, these households may take an informal loan despite its higher interest rate.

2.6.2 The *cultural* or sociological approach: the role of social norms

The cultural or sociological approach sees informal institutions as less purposive than rational individuals engaged in maximizing behaviour within some constraints.

The cultural view is that people engage within social milieus made up of associations that often vary between geographical areas. Many established relationships and norms are simply accepted as the "natural state" of affairs. Norms of reciprocity, intergenerational altruism and obligation involve households without having been consciously devised [Granovetter, 1985].

According to this approach, markets are bound up with networks of personal relations, kinship and reciprocal norms that are more extensive than in formal contracts [Aryeetey and Udry, 1995; Azam et al., 2001; Fafchamps and Lund, 2003; Platteau, 2004; Udry, 1990]. Understanding the characteristics of localities and how norms, obligations and networks work is as important as pinning down the economic rationale for their formation.

The issue on how social factors affect the expansion of informal credit arrangements is not unique to developing countries. For instance, Guiso et al. (2004) showed that informal institutions are more likely to develop in areas where there is less social capital[10]. In the Italian regions where social capital is lower, the risk of default is higher and hence formal credit arrangements are less developed. In this context, informal credit institutions compensate for the low social capital by relying on networks of personal and kinship relations.

There is another explanation for the formation of informal groups that emphasizes not standard economic arguments based on imperfect information, but rather behavioral

[10]The definition of "social capital" is taken from Bertola et al. (2006) and refers to the set of rules of conduct that bind people to obey to legal norms. However, this concept has been previously discussed by others [for example, Coleman, 1988; Putnam, 1993; van Bastelaer, 2000].

explanations organized around self-commitment [Bertrand et al., 2004]. If people have quasi-hyperbolic preferences, they find it difficult to save [Fafchamps and Lund, 2003]. In this case, groups like RoSCAs offer an incentive for members to commit to save in order to contribute to the pot.

Several models of saving and credit under time-inconsistency highlight an agents need for commitment and tendency to overborrow. For example, Laibson (1997) and Harris and Laibson (2000) have extensively studied lifetime saving under hyperbolic discounting. Krusell and Smith (2003) solve a Ramsey-style model when agents are time-inconsistent. Ashraf et al. (2005) find, in a field experiment, that agents most interested in commitment savings devices are those who face relatively greater time inconsistency in their preferences and are aware of it. Among other papers, Thaler and Bernartzi (2004) provide empirical evidence of the value of commitment in a range of informal financial settings.

2.7 The coexistence of formal and informal credit institutions

Despite the fact that several financial reforms have facilitated the diffusion of formal credit institutions in developing countries, we still observe the persistence of both formal and informal credit sectors in the same areas. This section aims to explain how the coexistence of formal and informal credit institutions occurs. The literature typically focuses on two research areas, the “spillover” or “residuality” theory and the market segmentation theory. Whilst an overview of both theories will be given, this thesis specifically tests the former theory.

2.7.1 The "residuality approach" or "spillover theory"

The so-called "residuality approach" or "spillover theory" maintains that the informal sector exists to satisfy the unmet demand for credit resulting from credit rationing in the formal sector [for example, Banerjee and Duflo, 2001; Bell et al., 1997; Besley, 1994; Bose and Cothren, 1997; Eswaran and Kotwal, 1989].

The spillover theory is linked to the market failure view and compatible both with the Stiglitz and Weiss (1981) and de Meza and Webb (1987) hypotheses outlined in sub-section 2.6.1. Quantity-constraints on banks' loans, as a result of the lack of collateral and of information problems, induce borrowers to resort to the informal sector. This view is based on the assumption that formal institutions are the cheapest available source of credit. Therefore, there is a natural ordering of credit sources whereby a borrower who uses secondary sources (informal credit) is assumed to be unable to satisfy his/her financial needs from the primary source (formal credit). The borrower is said to experience credit rationing with regard to the primary source[11].

The spillover hypothesis implies that a test of credit rationing is necessary. A direct test of credit rationing has been developed by Diagne (1999) and by Diagne et al. (2000) by using data on the maximum credit limit that the lender is willing to offer. This credit limit is used to detect whether formal credit markets are rationed.

Bell (1990) examined the interactions between institutional and informal credit sources in India. The model shows that, if formal credit is rationed and the informal lender is able to offer a contract which is preferred by the borrower, there is a spillover of demand from the formal to the informal market. This implies that if institutional lenders do not give as much credit as a borrower desires, the borrowers will turn to informal lenders. By using data from the Punjab region in India, Bell (1990) supported the conclusion

[11]It might be possible, however, that he/she is also rationed on the use of the secondary source.

of the model. The informal interest rate in equilibrium may be higher than the formal interest rate depending on the default rate, the cost of entry for new moneylenders and, consequently, on the level of competition.

Banerjee and Duflo (2001) showed that an expansion in the availability of bank credit leads to a fall in the firm's borrowing from the market as long as the bank is the cheapest credit source.

Boucher and Guirkinger (2007) demonstrated that the informal sector is the recipient of the spillover in demand for households with an intermediate level of wealth. The informal sector has the effect of relaxing the formal sector's quantity-rationing for these households.

2.7.2 Market segmentation

Rather than assuming perfect fungibility of credit (whether it be formal or informal), the second explanation for the coexistence of formal and informal credit sources maintains that markets are segmented. This means that no single type of credit can meet the needs of potential borrowers, and no single type of credit is accessible to everyone [Hoff and Stiglitz, 1990]. The reason for market segmentation, according to this theory, is not formal quantity-constraints on credit supplies, but the unique characteristics of the formal and informal sectors that inhibit the substituting of one credit source for the other[12].

The concept of segmented markets typically refers to the variation in preferences among consumers in different economic strata, for example consumers will use a loan in differing ways [Aryeetey and Udry, 1995; Barslund and Tarp, 2006; Mohieldin and

[12] Arguably, the spillover of formal credit demand to the informal sector leads to market segmentation. In this thesis, the author takes the view that market segmentation can itself be a cause, and not a consequence of the coexistence of formal and informal credit. As market segmentation can arise independently of spillover effects, it is separately treated as a theory for the coexistence of formal and informal credit.

Wright, 2000].

Aryeetey and Udry (1995) in their study of Ghana pointed out that "the variation in the types of informal financial units derives from the fact that such units are purpose-oriented". Mohieldin and Wright (2000) found that in Egypt the formal sector services loans for investment purposes, while the informal sector provides loans to aid consumption smoothing. Barslund and Tarp (2006) evidenced that in rural Vietnam formal loans are used almost entirely for production and asset accumulation, while informal loans are used for consumption smoothing.

Market segmentation may also arise as a result of asymmetric information between lenders. Indeed, a lender may have better knowledge about a potential borrower's creditability, or have better access to this information. This kind of information asymmetry may limit competition between lenders and may lead to a monopoly in particular segments of the market.

Social factors affecting the demand and supply of credit may also provoke market segmentation. In line with the *cultural* or sociological approach previously described, the importance of networks of personal relations and the degree of social capital in a community are factors that contribute to the segmentation of formal and informal credit sources.

2.8 Why intervene in credit markets?

The literature focuses on four motives for interventions in credit markets. The first motive is market failure. As previously discussed, scarcity of collateral, weak legal institutions and covariant risk environments render market failures particularly severe in developing countries [Udry, 1994]. In this context, the rationale for government intervention is to achieve market efficiency. Udry (1994) points out that a Pareto improvement must be sought taking into account enforcement problems and imperfections of information. By this standard, governments should aim at achieving a *constrained* Pareto efficiency. Applying this criterion, he argues, "narrows the field for market failure, but it still leaves room for a fairly broad array of cases in which resources could end up being inefficiently allocated".

Due to imperfect information, lenders who have better access to information may obtain market power. More specifically, village moneylenders are often seen as monopolists for their ability to gain access to local knowledge. The informal financial transactions offered by them have often been characterized as exploitative for the exceedingly high interest rates.

Udry (1994) and Armendariz and Morduch (2005) explained that the presence of high interest rates does not automatically imply inefficiency. However, an argument for intervention can be made if the monopolist moneylender is not able or willing to discriminate in the price charged to each borrower. Whenever the moneylender lends to the point where the marginal value of credit to each borrower is the same we obtain a "discriminating monopoly" outcome. In this case, loans are efficiently supplied even if the lender is said to be exploitative [Basu, 1989].

Distribution issues may represent the second motivation for government interventions in credit markets. To explain the importance of distributional issues and their link with efficiency let us analyse a typical agency problem highlighted by Udry (1994). Suppose there are two farmers: one has a high quality investment project and the other has some financial capital. Without adequate information about the quality of the project, the relatively wealthier farmer is unwilling to lend money to the other farmer. In this case, only if the financial capital is redistributed from the wealthier to the other farmer will the project be undertaken. In this simple example redistribution policies may reduce information problems in the economy.

The third motive for intervention in credit markets is the mitigation of vulnerability. Morduch (1999) and Dercon (2002) pointed out that public credit schemes can enhance the safety net to include particularly vulnerable households. In order to avoid the displacement of pre-existent local risk-sharing mechanisms, Dercon (2002) suggested encouraging the formation of group-based credit and savings schemes. In this way, groups can cope with idiosyncratic shocks by building up resources in good years that can be used in bad years. At the same time the government credit scheme can provide funds even when several common shocks occur and members are unable to insure each other.

Poverty reduction represents another motivation for government interventions in credit markets. Public credit schemes can displace informal transfers between equally poor households. For instance, Cox and Jimenex (1997) and Morduch (1999) observed in South Africa that intergenerational altruism from young to old equally poor households often impedes poverty reduction. In this context, displacing informal transfers may help keeping funds among younger households thus encouraging investment in human capital accumulation and other productive activities.

2.8.1 Two examples of government interventions in credit markets

In this sub-section we consider two government interventions which attempt to solve some of the problems outlined above. The first endeavours to create links between local moneylenders and banks. The second intervention creates government-sponsored microfinance institutions and this thesis specifically tests the effectiveness of these institutions.

(i) Creating links between local moneylenders and banks

Agency theory highlights a trade-off between resources and information. Banks have adequate resources to finance a large number of projects, but a lack of information about lenders' riskiness. Moneylenders, by contrast, have access to local information, but lack adequate resources.

The banks could hire local moneylenders to disburse loans and collect payments. In this way banks could circumvent their information problems by taking advantage of moneylenders' knowledge of the local market. The idea seems promising but it creates another problem, one of agency. How do the banks ensure that the moneylenders honestly and reliably carry out their principal's objectives?

Alternatively banks could make more financial capital available to moneylenders with the expectation that they would then lend the funds to local borrowers. At face value, this "trickle-down" policy also seems promising, but the increase in financial capital may push the interest rates up thus harming poor borrowers.

(ii) Government sponsored microfinance institutions

The second intervention aims at creating government-sponsored institutions (for example microfinance institutions) that mimic some of the features of informal arrangements [Armendariz and Morduch, 2005; Ray, 1997]. Banks could design contracts to harness local information and give borrowers incentives to use the information they hold on their peers to the advantage of the bank. For instance, group-lending institutions adopt a joint-liability scheme that creates incentives for safe borrowers to form a group with similarly safe borrowers ("positive assortative matching[13]").

Evaluating the effectiveness of this policy partly involves establishing whether microfinance institutions can "displace" the use of informal loans. This is called the "crowding out hypothesis".

[13] See sub-section 2.4.1 for a description of this mechanism.

The empirical literature on crowding out in the last fifteen years is quite large [see table C4.1 for a summary of the available literature]. Most of this literature tests the crowding out hypothesis by means of simple probit or tobit regressions where the dependent variables are private transfers of remittances and the independent variables include, among other controls, some form of public transfers (i.e. public pensions). The main problem faced by these simple models is endogeneity: the beneficiaries of public transfers schemes are not randomly selected. It is therefore important to use instrumental variables in the evaluation of the effects that public transfers have on private transfers.

We mention here some studies that do manage to establish a *causal* effect of public transfers on private transfers by means of instrumental variables or randomisation. Rosenzweig and Wolpin (1994) found a small trade-off between government aid provided to young women and parental aid in the United States. Cox et al. (2004) showed that in Philippines crowding out varies with the income of the targeted population. Jensen (2003) found that government old age pensions in South Africa crowd out private transfers. McKernan et al. (2005) showed that microcredit programmes in Bangladesh partially crowd out informal arrangements.

The papers by Albarran and Attanasio (2002) and Attanasio and Rios-Rull (2000) overcome the endogeneity problem by evaluating the effect of the *randomised* programme PROGRESA in Mexico. For example, the results obtained in Albarran and Attanasio (2002) indicate a substantial amount of crowding out of PROGRESA on local insurance arrangements. However, a recent paper by Kaboski and Townsend (2006) used pre and post programme data on the Thailand's Million Baht Village Fund programme and found no evidence of crowding out or of substitution away from other credit sources.

Given the above discussion on credit sources, it should be said that whether or not microfinance institutions will work in "displacing" informal loans depends on several factors including the design of the programmes, the target groups and on the communal norms and characteristics of the localities where these programmes are adopted. The extent to which the macro-level norms guide micro-level behaviour will depend on the larger context of social and economic change. While appropriate reforms could improve the economic context, the endowment of social capital evolves more slowly [Marchesi, 2002]. As argued by Williamson (2000), social capital is not the objective of a policy reform but a constraint to it.

2.9 Conclusion

Credit markets in Africa are dominated by informal institutions especially amongst poor households in Malawi and rural households in Ethiopia. Informal (formal) credit institutions have been classified according to a high (low) level of social cohesion and low (high) degree of formality. Examples of informal sources are friends and relatives, RoSCAs, self-help groups and moneylenders. Microfinance institutions, commercial and agricultural banks are examples of formal credit institutions.

The literature identifies two motives for credit: risk coping and acquisition of durable goods. Intertemporal consumption smoothing and risk sharing are two risk coping strategies discussed in detail.

According to Townsend's (1984) model of insurance, in a Pareto-efficient allocation of risk within a community, households achieve full (idiosyncratic) risk sharing and the only risk they face is aggregate risk. Asymmetric information and enforcement problems make risk-sharing contracts hard to implement. However, there is evidence that in small communities where information cannot be hidden and behaviour can be monitored an approximate Pareto efficient allocation of risk can be achieved.

Two theories for the existence and diffusion of credit are identified in the literature: the *economic* approach and the *cultural* or sociological approach. Both theories are specifically analysed in the thesis. The first one maintains that informal finance arises as a response to credit market failures in the form of adverse selection, moral hazard and enforcement problems. The sociological view is that people engage with milieus of associations that vary between geographical areas.

Despite the financial liberalisation process facilitating the diffusion of formal credit institutions, in developing economies formal and informal arrangements still coexist. The literature review outlined two theories for the persistence of formal and informal credit sectors. The first one, the "spillover" or "residuality" theory, maintains that the informal sector exists to satisfy the unmet demand for credit resulting from credit rationing in the formal sector. Rather than assuming perfect fungibility of credit, the second theory claims that markets are segmented. The unique characteristics of formal and informal sectors are to be held responsible for the coexistence of the two credit sources.

Finally, the literature of this thesis is motivated to intervene in credit markets due to its implications for efficiency and distribution, mitigation of vulnerability, and poverty reduction. There is a strong argument in favour of the creation of microfinance institutions that can displace access to informal loans.

Chapter 3

Access to informal credit in rural Ethiopia

"Destiny, I feel, is also a relationship - a play between divine grace and willful self-effort. Half of it you have no control over; half of it is absolutely in your hands, and your actions will show measurable consequence".

Elizabeth Gilbert (2006)

3.1 Introduction

Why do households participate in informal credit institutions? The standard economic argument is that informal finance arises as a response to credit market failures [Bell et al., 1997; Besley, 1994; Eswaran and Kotwal, 1989; Kochar, 1997; Pal, 2002]. The sociological approach, by contrast, maintains that markets are bound up with networks of personal relations, kinship and reciprocal norms that are more extensive than in formal contracts [Aryeetey and Udry, 1995; Azam et al., 2001; Fafchamps and Lund, 2003; Platteau, 2004; Udry, 1990].

It is often argued that market imperfections are less severe in developed economies at present for a variety of reasons[1]. Developed economies have appropriate informational sharing mechanisms (i.e. credit scoring) and methods for dealing with informational asymmetries in credit markets if and when they arise, such as, provision of collateral-based contracts and other instruments like credit bureaus [Cole and Mishler, 1998; Jappelli and Pagano, 2003; Padilla and Pagano, 1997, 2000; Pagano and Jappelli, 1993, 1999]. Also, formal institutions can develop in such markets because of scale economies, the relative lack of vulnerability of credit markets to adverse economic shocks and a better endowment of legal enforcement and social capital [Carpenter and Jensen, 2002; Guiso et al., 2004; Hoddinott et al., 2005].

In contrast, in developing economies such as in Africa, informational sharing mechanisms tend to be small scale and localised, markets are tightly interlinked and risky, low levels of wealth limit the provision of collateral, there are few scale economies and inefficient legal enforcement, and there is a smaller average endowment of social capital [Bardhan and Udry, 1999; Besley, 1994; Gosh et al., 1999; Ray, 1997]. In these circumstances, informal lending arrangements such as family and friends, and the development of local arrangements such as rotating saving and credit associations (RoSCAs) have an advantage. Indeed, these informal institutions may persist in the market even when formal institutions such as rural credit banks increase their market penetration [Carpenter and Jensen, 2002; Conning, 2001; Dasgupta et al., 2007; Madestam, 2005].

There is a large recent literature on the determinants of participation in informal credit arrangements in developing countries [e.g. Bose, 1998; Diagne, 1999, 2000; Kochar, 1997; Nagarajan et al., 1995; Pal, 2002; Ravi, 2003; Ray, 1997; Udry, 1994; Zeller, 1994].

Many studies adopt a reduced form specification in which variables that affect the

[1]See Bertola et al. (2006) for a detailed description of the topic.

demand for credit by different households and the supply of credit by various institutions are collapsed into a single equation estimating the probability of participation. A common characteristic of these studies is that a high degree of collinearity between household-specific variables (such as components of wealth, income and other household characteristics) limits the significance of individual regressors. Data availability impedes the identification of idiosyncratic and aggregate shocks which affect access to and choice of credit sources. Also, the effect that the existence of specific credit sources has on the relative substitutability between loans is often ignored.

Another weakness of such studies is that the reduced form specification fails to identify other dimensions of heterogeneity of access - most notably social, economic and geographic - which may operate at a cluster level, but which are not identified at a household level (other than through crude proxies such as ethnicity and religion). For example in rural Ethiopia, which is the location examined in the present chapter, the nature of local credit markets varies widely across geographical areas. The use of RoSCAs, known as *equbs*, is far more pervasive in the south than in the north of the country - indeed in some northern localities there is no evidence of such informal institutions being used at all. Non-participation may therefore not be an outcome of household choice but of cluster characteristics. Knowledge of these cluster-level differences is as important as knowing why households utilise such institutions in clusters where they are available.

The present chapter uses as its primary source panel household data from the Ethiopian Rural Household Survey (ERHS, 1994-1997). In light of the previous discussion, it makes several innovations to the now well-established literature on the use of informal institutions. First, it presents a model where the demand for informal credit of a particular type is only observed in clusters where such credit is supplied. Several empirical versions of this model will be implemented.

Second, by using principal components analysis, primarily on household wealth-holdings and expenditure, this chapter demonstrates how it is particular associations between components of wealth and expenditure that have a highly significant impact on the use of informal arrangements, when compared with standard regression models which specify the determinants of household use of informal institutions as linear combinations of underlying assets.

Third, household-based and cluster-based determinants (such as shocks and socio-economic characteristics) of the observed use of informal institutions will be explicitly differentiated.

Finally, this chapter develops a model in which household and cluster determinants affect the demand for a particular type of informal credit thus allowing the relative substitutability of informal credit sources to endogenously vary across clusters.

The structure of the chapter is as follows. Section 3.2 outlines the data management strategy. Section 3.3 reports the descriptive statistics. Several empirical specifications are identified in section 3.4. Section 3.5 concludes.

3.2 Data description and management

3.2.1 The Ethiopian Rural Household Survey

Ethiopia is divided into nine ethnically-based administrative regions and subdivided into 68 zones and two chartered cities: Addis Ababa and Dire Dawa[2]. It is further subdivided into 550 *weredas* and six special *weredas*. *Weredas* are divided into Peasant Associations (PA) or *Kebeles* which were constituted after the 1974 revolution.

The Ethiopian Household Survey is composed of a rural and an urban part, separa-

[2]A map of Ethiopia is displayed in figure A.1 in appendix A.

tely undertaken by the Department of Economics at Addis Ababa University (AAU). The urban surveys were done in collaboration with the Department of Economics of Göteborg University and Michigan State University, while the rural surveys were done in collaboration with the Centre for the Study of African Economies (CSAE), Oxford University and the International Food Policy Research Institute (IFPRI).

The Ethiopian Rural Household Survey (ERHS)[3] is a unique longitudinal household data set taken in 15 villages. This thesis utilises four of the following five rounds: 1989, 1994 (which includes two rounds), 1995, 1997. The first round in 1989 involved six farming Peasant Associations in Central and Southern Ethiopia. In this wave, Peasant Associations (PAs) were selected among those afflicted by famine in 1984-1985 and by other droughts between 1987 and 1989. Households were then randomly selected within each Peasant Association. The second survey produced in 1994 includes 15 Peasant Associations across four regions and the total sample comprises 1,477 households. The 1994 survey includes two rounds: the first round includes data from March to July, the second round considers the period from September to January[4]. The sample constituting the six villages present in the 1989 round was re-randomised by including an exact proportion of newly formed or arrived households and by replacing the lost households with similar ones. The new nine villages were purposely selected to represent: the diversity of farming systems in the country, a representative number of landless households and an exact proportion of female headed households.

However, the sample does not include pastoral households or urban areas. Hence, this data, which can be considered representative of non-pastoral households, cannot be

[3]Funding for data collection was provided by the Economic and Social Research Council (ESRC), the Swedish International Development Agency (SIDA) and the United States Agency for International Development (USAID); the preparation of the public release version of these data was supported, in part, by the World Bank. AAU, CSAE, IFPRI, ESRC, SIDA, USAID and the World Bank are not responsible for any errors in these data or for their use or interpretation.

[4]The Ethiopian year follows the Julian calendar: the year starts on the 11th of September. Hence, the second round covers the period 1994-1995.

fully representative of the country as a whole.

While it is relatively easy to compare and merge the two rounds in 1994, 1995 and 1997, the use of the first round is quite tricky. In fact, the 1989 survey addressed a narrower range of topics because at the time when the survey was undertaken, there was no intention of producing a panel data set. So, the present chapter focuses only on the last four rounds.

3.2.2 Village Studies

The originality of the ERHS consists of the implementation of village studies for all 15 communities alongside survey data[5]. These studies are based on interviews by graduate students and qualitative fieldwork. They describe the location, seasonal activities, events, local organizations and institutions, values and beliefs of the villages.

The large variety of topics in the survey allows the data to be used by other disciplines in a variety of different ways. Complementary anthropological and sociological techniques enable a wider range of comparisons to be made.

The first issue in the data management is the choice of the unit of analysis. Different approaches focus on different entities: households, individuals or communities. While anthropologists are more interested in communities and social networks defined in the survey as Peasant Associations, economists focus on households or individuals.

Considering the household as the unit of analysis is quite restrictive (i.e. no intra-household issues), but easy to interpret. However, one should be aware of the fact that there are different definitions of "household" across different areas.

[5]A map of the surveyed sites is displayed in appendix A.

1. In northern Ethiopia the household is viewed as an economic unit of people living, eating and working together rather than a unit based on kinship. New households are formed by separation from the original household and through marriage. Usually older households are wealthier than newly formed ones that lack resources.

2. In southern Ethiopia household membership is related to kinship. Co-residence is not a necessary criterion for household membership. A married son not entirely economically independent is still considered a member of the household.

The second issue is spatial: choice and characteristics of the sampled villages. As mentioned above, the survey sites were chosen by IFPRI in 1989 from more vulnerable areas. From 1994 new sites were added to represent the diversity of the country. Table C3-1 in appendix C summarises several characteristics of the Peasant Associations. There are eight quite wealthy PAs and seven vulnerable ones; six PAs are adjacent to all-weather roads and have relatively easy access to towns. Coffee and *teff*[6] are produced in most of the richer PAs, cereals in the poorer ones.

The third issue is time intervals between rounds. It should be pointed out that the rounds were not uniformly conducted at the same periods of the year (see table C3-2). Hence, any comparison between rounds should be made with care.

Unit of analysis, spatial issue and time intervals are characterised in detail in the village studies, taking into account seasonal activities and events, farm economy, off-farm income activities, consumption, local institutions and organization, beliefs and values. Appendix A contains a brief summary of each survey site.

[6] *Teff* is an annual grass, a species of lovegrass native to northern Ethiopia. It is an important food grain and used to make injera.

3.2.3 Missing Data

The problem of missing data affects most surveys, especially those from developing countries. In panel data sets, households may drop out or may move away to another location or may not be available at the time of interview. When data is collected by questionnaires, households may be unwilling or unable to respond to some questions. Moreover, data set responses may contain outliers and implausible values that, when set equal to missing, may cause bias to the results.

Unfortunately, some information provided in the ERHS is either missing or implausible[7]. Hence, it is important the way in which missing values are dealt with.

There are several methods of dealing with missing data[8]. We chose to replace missing data by *hotdeck* imputation because it has the advantage of taking into account the uncertainty of imputed values and is computationally easier than multiple imputation techniques as decribed in appendix A. *Hotdeck* imputation replaces missing values with a single random draw from an imputation class. Within each imputation class a missing observation on X is replaced by randomly sampling a single observed value of X (with replacement) from that class [Paul et al., 2003]. When the missing mechanism is either completely at random (MCAR) or at random (MAR) and the model is correctly specified, *hotdeck* imputation gives unbiased coefficient estimates[9]. A limitation of this technique is that it is statistically inefficient because it uses a single draw. Also, standard errors are biased because the estimates are given by a re-sample of the data.

Imputation was only utilised when there were outliers or when there were other information available. For example, the value of cows was imputed whenever the respondent

[7] In most cases the percentage of missing values does not exceed six percent. In the empirical analysis we do not use variables such as 'number of rooms in the house' that has a percentage of missing values of around 70 percent.

[8] A short description of different mechanisms for coping with missing data is provided in appendix A.

[9] MCAR refers to the fact that the probability that a case is missing is independent of households' characteristics. When the probability that cases are missing is not independent of some subset of households' characteristics, then the missing process is called 'missing at random' (MAR).

provides the number of cows owned. An objection might be that the decision of non-reporting the value is non-random, but, as outlined earlier, missing randomness is not a testable assumption. We have used region, peasant association and education of the household head as classes in the *hotdeck* imputation. There are two reasons for the choice of such imputation classes. Firstly, we wanted to have proxies of the characteristics of the villages (relatively similar villages can be given the same value of assets, for example) and of the characteristics of the households (i.e. education might affect the rate and quality of response). Secondly, these two classes were the ones with lower rate of missings.

3.2.4 Price Indices

In order to adequately compare monetary values that differ across rounds and clusters, price indices need to be used to deflate households' expenditure.

While in developed countries the main source of price variation is time, in developing countries spatial price variation is more crucial [Deaton, 1997]. Poorly developed infrastructures, high transport costs, and poor distribution systems are to blame. So, for instance, in developing countries where often roads are not available and markets are not integrated, there is little chance for arbitrage between geographical areas.

The ERHS indirectly provides some data on prices. Each household states the amount of food consumed in a local unit of measure and the expenditure in local currency. The ratio of these two variables (after having converted quantities in kilograms) is a measurement of price or unit value[10]. Caution is needed in interpreting prices as unit values for a number of reasons [Deaton, 1997]. Firstly, they are affected by quality as well as by the actual price that the consumer faces in the market. Secondly, unit values are not

[10]The conversion factors for each local unit of measure is provided in the data.

available for those households who do not report their expenditure on each food item. Finally, unit values might be affected by measurement error especially when they are obtained by dividing expenditure by quantity. As quantity is not reported in standard units, it may also be affected by measurement error [Capeau and Dercon, 1998; Kedir, 2005]. Nevertheless, looking at unit values provides some information about price vari-

TABLE 3.1: Log unit values of some food items by regions and peasant associations; rural Ethiopia, 1997

Regions	**Maize**	**Barley**	**Salt**	**Potatoes**	**Beef**
Amhara	-0.04	0.42	-0.97	-0.75	0.06
	(0.53)	*(3.18)****	*(15.00)****	*(2.24)***	*(0.39)*
Oromiya	-0.21	-0.18	-1.05	-0.95	0.69
	*(2.22)***	*(1.45)*	*(15.86)****	*(2.86)****	*(4.68)****
Separ	-0.37	-0.10	0.94	-0.82	0.16
	*(5.75)****	*(1.40)*	*(14.00)****	*(2.47)***	*(1.88)**
R^2	0.14	0.18	0.22	0.02	0.03
N. obs.	307	140	1261	269	207
			F-statistics		
Tigray	0.98	18.46	71.18	n/a	0.40
Amhara	19.86	72.53	46.91	15.96	1.13
Oromiya	27.73	0.00	41.86	28.08	17.23
Separ	21.51	1.27	154.44	104.33	72.44

Note: The top panel shows coefficients and absolute t-values (in brackets) from a regression of log-unit values on regions for each food item. The bottom panel shows the F-statistics for cluster (peasant association) dummies within each of the four regions. The regressions and F-statistics use data from only those households who report expenditure on the considered food items.
***$p < 0.01$, **$p < 0.05$, *$p < 0.1$.

ation.

The top panel of table 3.1 gives regressions of logarithms of unit prices for some food items on a set of dummies indicating regions: Amhara, Oromiya, Separ and Tigray (omitted region). These very simple regressions aim at capturing prices differences *between* regions [as in Deaton, 1997]. Almost all regions are cheaper than Tigray for all food items with the exception of beef. While salt and potatoes are cheaper in Oromiya than elsewhere, maize is cheaper in Separ.

The bottom panel of table 3.1 displays the values of F-statistics for the regressions of logarithms of unit prices on a set of peasant associations (clusters) dummies. These regressions look at differences in unit prices between clusters *within* the same region. Tigray is composed of two peasant associations, Amhara and Oromiya include four peasant associations and Separ has five peasant associations (see map in appendix A or table C3-1). However, either because some households do not report expenditure for that particular item or because some of the food items are not produced in that area, not all clusters have been included in each regression. The bottom panel tests the null hypothesis of no difference in logarithms of unit prices *between* clusters. The hypothesis can be rejected for most of the items in all clusters except those in Tigray (for maize and beef).

Given the spatial price variation, it is not appropriate to use a consumer price index based on the country as a whole to deflate monetary values. Instead, the survey data can be used to calculate a Fisher index.

Disney et al. (2004) constructed an "ideal Fisher index" that uses unit values from the Ethiopia Urban Household Survey (EUHS), but accounts for selection between metric and non-metric units, quality effects, and correlations between purchases of items and unit values. They adopted this spatial price deflator to calculate food poverty measures in urban Ethiopia. We do not make such "quality adjustment" because we use deflated wealth values simply to construct principal components and not to estimate poverty measures. Deaton (1997) points out that if the response of quality to income is close to zero, then it is possible that the effect of quality on prices will also be negligible.

The Fisher index we use can be written as follows[11]:

$$P_F = \sum_{n=1}^{N} w_n \left(\frac{p_n^1}{p_n^0} \right) \qquad \forall \text{ n} = 1, 2, \ldots, 47$$

where p^0 is the price in the base year (1997) and N=*47* are the included food items.

[11] Balk (2004) shows that the Fisher price index can be expressed as a weighted arithmetic average of price relatives.

Table 3.2: Characteristics of Peasant Associations

PA	**N. of villages**	**Number of households**	**Predominant ethnic group**	**% of predominant ethnic group**	**Distance to nearest town (Km)**	**Tot. land (ha)**	**Tot. irrigated land (ha)**	**Tot. rainfed land (ha)**	**Tot. grazing land (ha)**
Harresaw	3	800	Tigrai	99	17	500	15	200	20
Geblen	3	2150	Tigrai	35	18	3100	100	1000	1500
Dinki	8	138	Argoba	67	10	798	6	132	150
Yetmen	4	n/a	Amhara	100	0	n/a	n/a	n/a	n/a
Shumsha	9	800	Amhara	100	12	2400	1	800	800
Sirbana G.	n/a	250	Oromo	60	15	2400	0	1960	240
Adele K.	25	900	Oromo	99	13	1004	30	733	23
Korodegaga	1	245	Oromo	100	25	1200	500	500	39
T.Ketchema	2	450	Oromo	61	12	800	0	696	40
Imdibir	10	790	Gurage	100	1	60	0	4.8	80
Azr Deboa	16	842	Kembata	98	4.5	1012	0	766	87
Adado	> 10	1000	Gedeo	100	11	800	0	700	0
Gara Godo	4	1900	Wolaita	93	13	800	0	600	100
Doma	4	200	Gamo	98	3.5	3600	2000	650	300
Debre Birhan	3	450	Amhara	87	10	2280	13	960	840

Source: Own calculation based on ERHS, community data.

The Fisher index has been calculated according to the 2000/2001 Tanzanian Household Budget Survey as a weighted average of median unit prices for each food item, where the weights are given by the median quantity of each food item. It includes 47 food items most commonly used by the households: white *teff*, barley, coffee, potatoes, maize, milk, bread etc. As said before, unit prices are calculated by dividing expenditure by the quantity converted in standard units (kilograms) after having deleted outliers. Then, median unit prices and quantities are obtained for each food item. The Fisher index is a weighted average of these two measures. Table C3-3 shows the values of the index for each peasant association.

3.3 Descriptive statistics

This section provides some descriptive statistics both of the community and of the household data. The subsection entailing community statistics describes the general characteristics of the survey sites. The household subsection is divided into two parts. The first part considers the composition of the household. The second part analyses the household borrowing behaviour, the reasons for borrowing and the characteristics of borrowers.

3.3.1 Community level-data

The community survey includes information about the peasant associations (PAs). As already mentioned, a PA is an administrative unit of one or a small number of villages. Table 3.2 shows some general characteristics of the peasant associations. The biggest PA is Adele Kebe (with 25 villages) situated in the Oromiya region. There are approximately 900 households and the total land size is around 1000 hectares of which

Table 3.3: Institutions by Peasant Associations

PA	N. of banks within PA	Distance (Km) to nearest bank outside PA	N. of agric. office within PA	Distance (Km) to nearest agric. office outside PA	N. of agric. coop. within PA	Distance (Km) to nearest agric. coop. outside PA	N. of NGOs within PA	Distance (Km) to nearest NGO outside PA	Tot. gov. hospitals within PA	Distance (Km) to nearest gov. hosp. within PA
Harresaw	0	44	1	n/a	1	n/a	0	17	0	n/a
Geblen	0	18	1	n/a	0	n/a	0	5	0	18
Dinki	0	70	0	5	0	75	0	10	0	70
Yetmen	0	17	1	n/a	0	3	0	17	0	75
Shumsha	0	120	1	n/a	1	n/a	0	9	0	120
Sirbana G.	0	15	1	n/a	1	n/a	0	n/a	0	15
Adele K.	0	7	1	11	0	n/a	0	25	0	25
Korodegaga	0	25	1	0	0	n/a	0	10	0	25
T.Ketchema	0	12	0	12	0	5	0	n/a	0	2
Imdibir	0	35	1	n/a	0	n/a	1	n/a	0	18
Azr Deboa	0	4.5	0	2	0	2	0	4.5	0	61
Adado	0	22	0	11	1	n/a	0	17	0	22
Gara Godo	0	27	1	n/a	1	n/a	0	13	0	27
Doma	0	60	0	5	0	n/a	0	n/a	0	80
Debre Birhan	0	10	1	n/a	0	n/a	1	n/a	0	10

Source: Own calculation based on ERHS, community data.

30 hectares is irrigated, 733 hectares is rain fed and 23 hectares is grazing land. Other big PAs include Azr Deboa (16 villages), Adado (more than 10 villages) and Imdibir (10 villages) all situated in Separ. They have no irrigated land and the total land size is 1,012, 800 and 60 hectares respectively. Korodegaga has only one village and it is located in the Oromiya region.

Table 3.3 displays the access of PAs to formal sector institutions. There are no banks within any of the PAs. Azr Deboa is the nearest PA to a bank (4.5 kilometers), but it has no institution within itself. Adele Kebe is quite close to a bank (7 kilometers) and it also has an agricultural office. Shumsha, located in Amhara (northern Ethiopia), is the most distant to a bank (120 kilometers), but it has an agricultural office and an agricultural cooperative within the PA. While most of the PAs have an agricultural office, only few of them have an NGO (Imdibir and Debre Berhan). The closest NGO is five kilometers from Geblen. There are no government hospitals within the surveyed peasant associations. Almost all hospitals are distant more than 10 kilometers from the PA. T. Ketchema and Debre Birhan are the closets PAs to the hospitals (2 and 10 kilometers respectively).

3.3.2 Household level-data

The rural household data set used here is an unbalanced panel involving four rounds (1994-1997). It excludes the 1989 round because it has a restricted number of PAs. Overall there are 1,457 households[12], 94 percent of which (or 1,372) are present in all rounds, and 15 peasant associations in four regions. For brevity purposes, we focus on the characteristics of the households in each of the four regions.

[12]It is lower than the number of interviewed households (1,477) because not all households demographic characteristics were reported.

3.3.2.1 Households' characteristics

Table 3.4 reports the number of surveyed households and the percentage of female headed households by round and region. The table shows that attrition is very low: only 5.8 percent of households leave or reenter the sample after leaving. While 44 percent of households in Tigray are female headed, only around eleven percent of female households are located in Separ.

Table 3.4: Households' characteristics by region

REGIONS	*Number of surveyed households* Rounds				*% of female headed households* Rounds			
	1	**2**	**3**	**4**	**1**	**2**	**3**	**4**
Tigray	150	149	145	144	45	44	44	43
Amhara	471	468	460	446	24	24	23	24
Oromiya	381	382	383	375	23	23	23	26
Separ	440	441	441	427	10	10	10	15

Source: Own calculation based on ERHS, community data.

Table 3.5 shows the household composition by region. On average, households have approximately five members in Tigray and Amhara. Larger households (with approximately seven members) can be found in southern Ethiopia. The average number of children per household is one across all age groups (i.e. between zero and five, six and 10, 11 and 17). While only two percent of household heads in Separ attended an adult literacy program, approximately 28 percent of household heads in Amhara participated in an adult literacy program (ALP).

Table 3.5: Households' composition by region

REGIONS	*Household size*		*N. of children 0-5*		*N. of children 6-10*		*N. of children 11-17*		*% of HH head*
	N. obs	**Mean (std. dev.)**	**N. obs**	**Mean (std. dev.)**	**N. obs**	**Mean (std. dev.)**	**N. obs**	**Mean (std. dev.)**	***with ALP***
Tigray	588	5.1 (3)	588	1.1 (1)	588	0.8 (1)	588	0.8 (1)	0.74
Amhara	1845	5.2 (3)	1845	1.0 (1)	1845	0.8 (1)	1845	0.9 (1)	27.94
Oromiya	1521	7.4 (4)	1521	1.3 (1)	1521	1.2 (1)	1521	1.4 (1)	11.39
Separ	1749	7.3 (3)	1749	1.3 (1)	1749	1.2 (1)	1749	1.2 (1)	1.63

Source: Own calculation based on ERHS.

Table 3.6 reports the mean and standard deviation of assets and expenditure for each region. Assets have been classified into three categories: equipment, house assets and other assets[13]. On average, households in Oromiya own the highest value of equipment. Tigray is the region where households own the lowest value of equipment and house assets. Non-food expenditure has been divided in three types: expenditure in clothes, furniture and ceremonials; expenditure in health or education for anyone outside the household; and expenditure in health or education for household members. While households in Oromiya spend more on non-food items, households in Amhara have the highest food expenditure. Also, Tigray is the region with the lowest non-food expenditure (in clothes, furniture and ceremonials; and in health or education).

TABLE 3.6: Households' assets and expenditure by region

REGIONS	***Tigray***		***Amhara***		***Oromiya***		***Separ***	
	N. obs	**Mean (std. dev.)**	**N. obs**	**Mean (std. dev.)**	**N. obs**	**Mean (std. dev.)**	**N. obs**	**Mean (std. dev.)**
Equipment	584	82 (57)	1822	97 (104)	1515	108 (150)	1738	85 (112)
House assets	584	58 (103)	1822	221 (298)	1515	221 (378)	1738	276 (611)
Other assets	584	156 (250)	1822	124 (165)	1515	145 (247)	1738	66 (119)
Expenditure 1[a]	586	63 (88)	1839	100 (107)	1517	183 (210)	1744	125 (140)
Expenditure 2[b]	7	9 (15)	36	15 (33)	119	20 (40)	106	13 (21)
Expenditure 3[c]	4	23 (35)	18	16 (38)	92	33 (43)	58	23 (47)
Food exp.	584	32 (37)	1768	35 (39)	1499	25 (33)	1727	22 (28)

Source: Own calculation based on ERHS. Note: all values in local currency (1 birr=0.1143$), deflated by using the Fisher Index (1997 base year). Expenditure deflated by the square root of households' size. [a]expenditure in clothes, furniture and ceremonials;[b]expenditure in health or education for anyone outside the household; [c] expenditure in health or education for household members. Food expenditure refers to the week before the interview. All other expenditures refer to a four month period before the interview. The value of assets in each round, subsequent to the first, corresponds to the value of the previous round and the current purchases after subtracting amount sold.

3.3.2.2 Households' borrowing behaviour

The Ethiopian Rural Household Survey contains information about households' borrowing behaviour. This section outlines the characteristics of both credit suppliers and

[13]Equipment includes hoe, plough, hammer, saddle, cart, weaving, mill, sickle, chopper, and spade. House assets include beds, tables, fanos, radio, leather sofa, iron, shelves, woven straw table, and leather. Other assets include jewerly, sword, hive, beehive and barrel.

borrowers. It describes the composition and distribution of credit sources, the motives for borrowing and the extent to which shocks affect rural households. Finally, it describes the characteristics of borrowing households in terms of wealth and demographic factors.

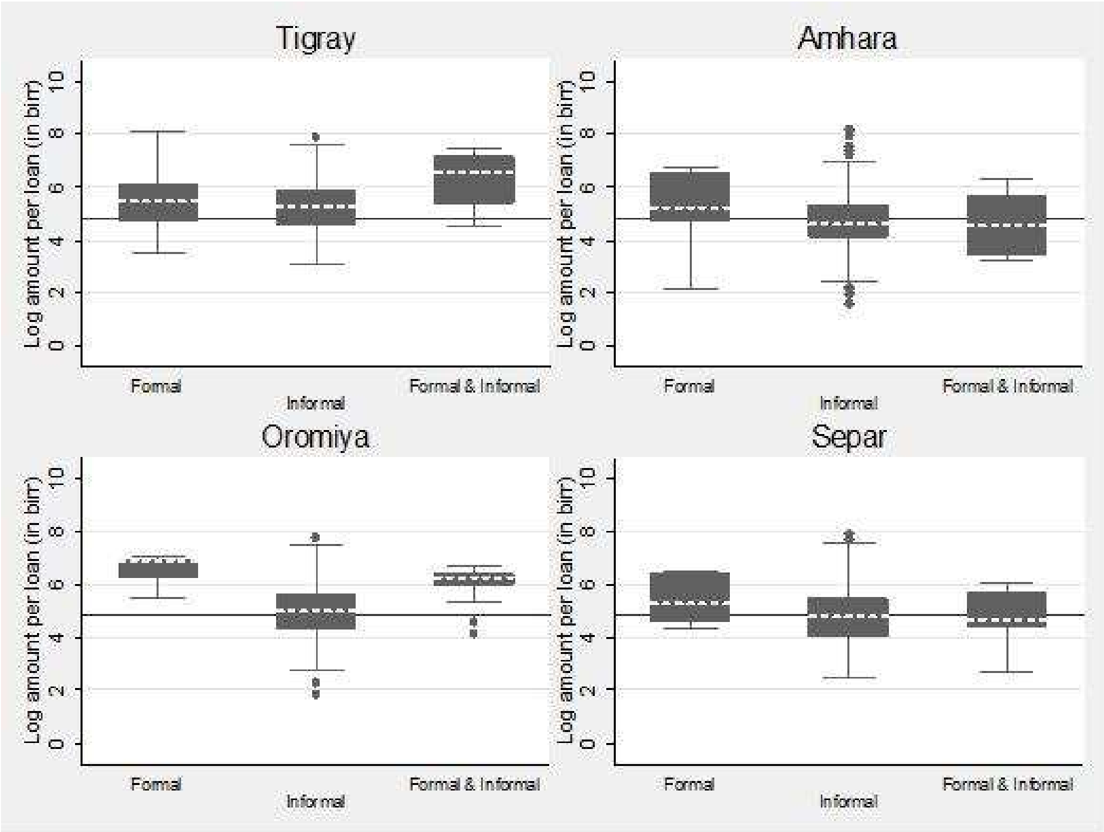
Who are the suppliers of credit?

There are two credit sources in the regions. Formal or institutional lenders include banks and NGOs. Informal lenders include: friends and relatives, moneylenders, *equbs*, *iddirs*. All these credit institutions were described in the previous chapter.

Figure 3.1 shows the distribution of formal and informal sources by region. Box plot diagrams are interpreted as follows: for each box, 50 percent of cases have values within the box and the dotted horizontal line is the median. The length of the box is the inter-quartile range and the lower boundary (upper boundary) of the box is the 25$^{\text{th}}$ (75$^{\text{th}}$) percentile. The black line is the overall median. The circles are extreme values (not outliers).

Across all rural areas the median amount of credit per loan (in logarithm) is 4.8 birr. Households can borrow only from the informal sector, only from the formal sector or they can borrow from both informal and formal lenders. In Oromiya the box plots are quite narrow, meaning that the distribution of formal, informal and formal and informal loans is less spread than in the other regions. In this region, households borrow more from formal and from both formal and informal sources than in other rural areas. However, the distribution of formal loans is negatively skewed because most of the cases inside the box fall below the median line. In Amhara and Tigray households borrow less from formal sources than in other regions and there is a large spread in the distribution of formal loans. In Separ, households borrow more from formal sources. The median amount they

Figure 3.1: Distribution of formal and informal credit by region in rural Ethiopia



Source: Own calculation based on ERHS.

borrow is approximately the same (around 4.9 birr) across the three loan categories.

In sum, box plots have shown that access to credit sources varies across regions. In the following descriptive statistics we pin down two reasons for this variation. First, the availability of different lenders varies across regions thus affecting the substitutability between sources. Second, as households' wealth varies across regions so does the access to credit markets.

So far, the informal sector has been considered as a homogeneous group. However, informal lenders are of different types: some of them (i.e. friends and relatives) may not require collateral or interest payments while some others (i.e. moneylenders) may do so. Also, their availability varies across geographical areas.

Figure 3.2 displays the distribution of credit providers by region. Not all informal sources are utilised by households in each region which explains the high variation in access to credit observed in the box plots. For example, *equbs*[14] are only used in southern Ethiopia[15] (i.e. Oromiya and Separ). In all the regions more than 50 percent of the loans are supplied by friends and relatives. The second most used credit sources are NGOs (27 percent) in Tigray, followed by moneylenders in Amhara and Separ (approximately 16 percent); and *iddirs* (around 13 percent) in Separ. In southern Ethiopia (i.e. Oromiya and Separ) borrowing from friends and relatives is more common than in northern Ethiopia (i.e. Tigray and Amhara).

Any analysis of access to informal credit cannot neglect the social characteristics that affect the availability of such sources in different clusters. For instance, according to Bevan and Pankhurst (1996), in southern Ethiopia household membership is not related to the economic status of the member but to kinship, establishing an "obligation" to reciprocate. In rural Ethiopia formal sources are not as common as informal ones. Approximately, one percent of loans are supplied by banks in Amhara, Oromiya and Separ. In Tigray, where the NGO is the second largest credit provider, households do not borrow from banks at all.

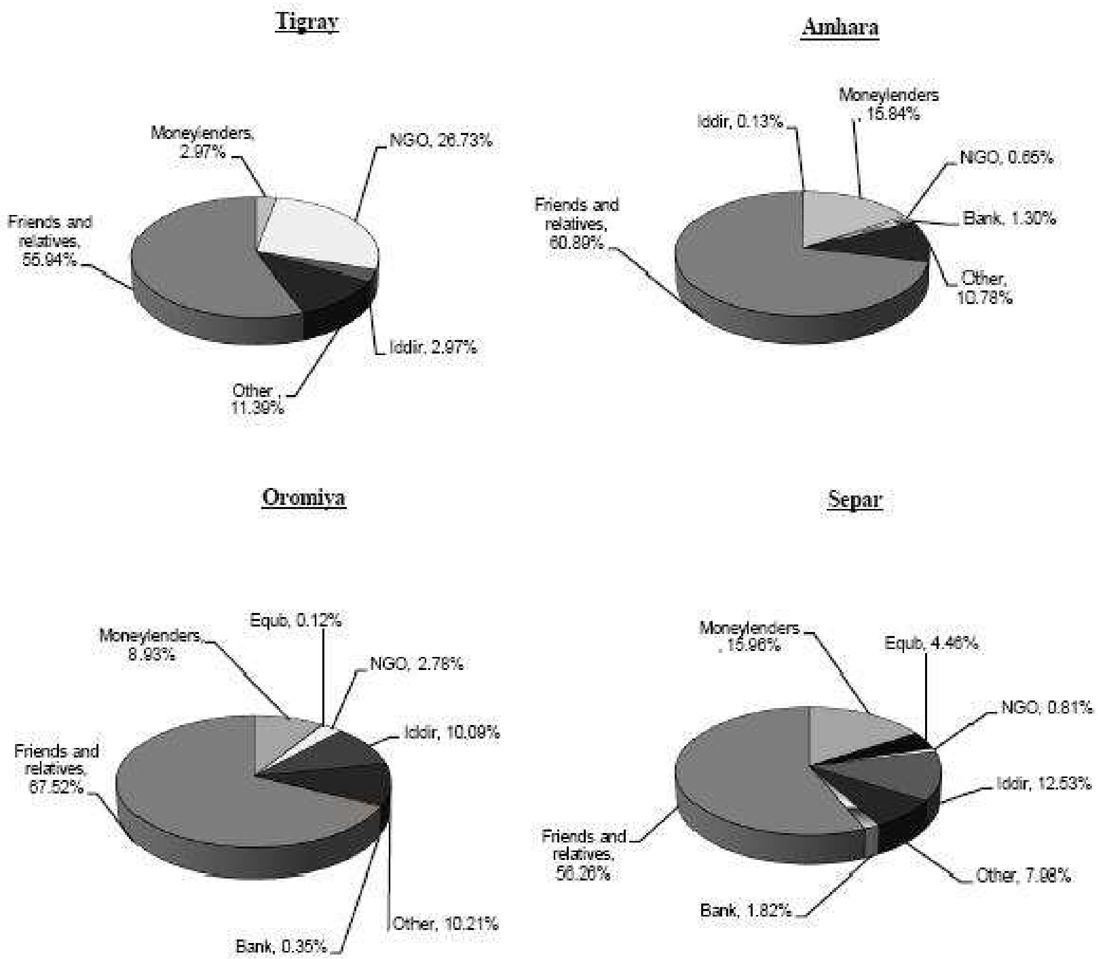
For what purposes do rural households borrow?

As mentioned in the previous chapter, agricultural households borrow for two reasons: a) to finance the acquisition of farm equipment; and b) to cope with shocks. The ERHS contains information about the reasons for which households borrow.

[14]By contrast, the other self-help group (i.e. *iddir*) is used in all regions.

[15]Indeed, the village studies report that *equbs* are not available in Haresaw and Geblen (located in Tigray) and Dinki (located in Amhara). Figure 3.2 shows that the surveyed households in Amhara do not borrow from *equbs* even though this source is available in all the PAs except Dinki.

FIGURE 3.2: Distribution of loan sources by region

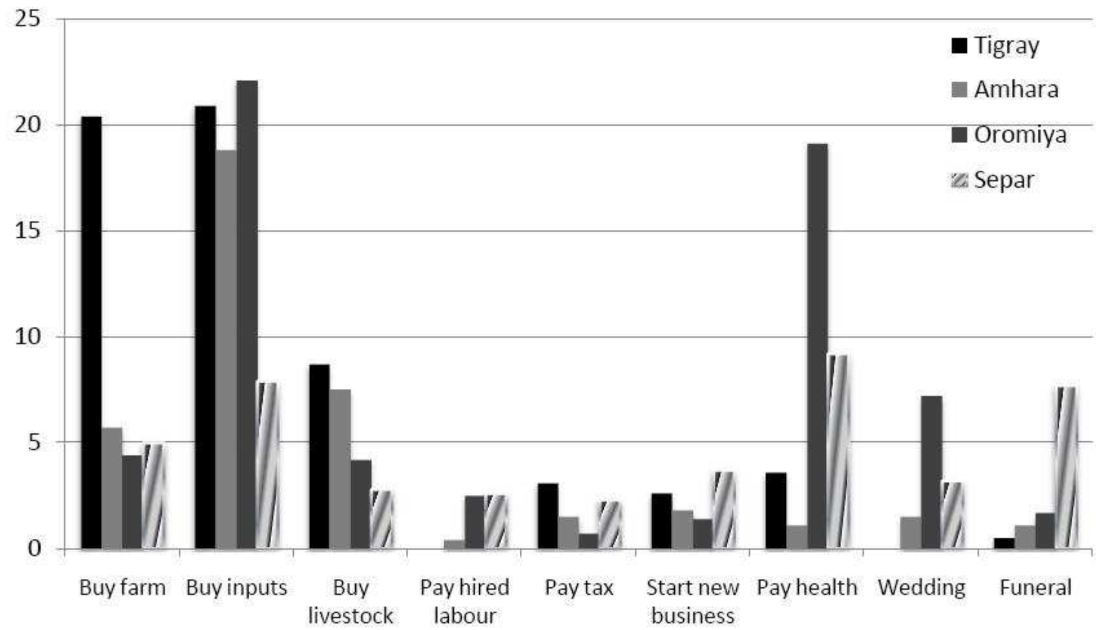


Source: Own calculation based on ERHS.

By looking at figure 3.3, the most common reason for borrowing in Tigray, Amhara and Oromiya is to buy farm inputs. In addition, rural households use credit to buy new farms. While more than 20 percent of loans in Tigray are used to buy new farms, only five percent of loans in Amhara and Separ are used for the same purpose. Approximately nine percent of loans in Tigray and Amhara are used to buy livestock.

The second reason for borrowing is to cope with shocks. The ERHS has information on health and income shocks (in terms of income needed for weddings and funerals). While in Separ almost 10 percent of loans are used to pay for health expenses, in Oromiya almost 20 percent of loans are used for the same purpose. In Oromiya and Separ a large

Figure 3.3: Reasons for borrowing by region



Source: Own calculation based on ERHS.

percentage of loans are used for ceremonies: approximately seven percent of loans are used for weddings and funerals. Indeed, in these two regions a large percentage of loans are obtained from *iddirs*, institutions specifically created for burial activities (see figure 3.2).

The ERHS contains more specific data on the type of health shocks households face (tables C3-4 and C3-5). A high percentage of borrowing households in Amhara, Oromiya and Separ have children under the age of seven who have health problems. Also, in these regions households spend more on medicines.

To what extent have shocks affected the village or the borrowing household?

In the previous chapter we mentioned an area of research that links the access to credit with risk-sharing and risk-coping strategies [e.g. Bardhan and Udry, 1999; Binswanger and Rosenzweig, 1986; Platteau and Abraham, 1987; Ruthenberg, 1971; Townsend, 1994]. According to this literature, risk-sharing institutions can cope with idiosyncratic

TABLE 3.7: Extent of income and health shocks by region

affected...	**Tigray**	**Amhara**	**Oromiya**	**Separ**
formal				
beyond *wereda*	44.4	57.7	28.6	75.0
some in the *wereda*	15.9	23.1	14.3	21.4
everyone in the PA	28.6	23.1	22.9	17.9
some in the PA	23.8	11.5	20.0	35.7
few in the PA	22.2	15.4	22.9	21.4
household only	61.9	38.5	54.3	60.7
informal				
beyond *wereda*	51.2	51.5	20.2	59.3
some in the *wereda*	18.3	12.3	11.3	13.4
everyone in the PA	22.1	25.5	32.3	24.5
some in the PA	26.7	14.2	19.3	29.2
few in the PA	21.4	12.0	14.4	20.5
household only	61.1	62.7	58.2	72.9

Source: Own calculation based on ERHS. Note: % of borrowing households displayed.

shocks, but not with aggregate (undiversifiable) shocks.

The ERHS provides data on the extent to which income and health shocks have affected the household over a period of 20 years. Each household is asked the extent to which health or income shocks have affected the community. As displayed in table 3.7, aggregate shocks affect everyone in the peasant association, idiosyncratic shocks affect only the household. The top and bottom panels of table 3.7 report the extent of the shocks for households borrowing from formal and informal credit sources respectively.

A high percentage of households borrow from either formal or informal credit sources after they have been affected by an idiosyncratic shock in rural Ethiopia. Approximately 63 percent and 73 percent of households who have been affected by an idiosyncratic shock borrow from informal lenders in Amhara and Separ respectively. Less than 33 percent of households who have been affected by aggregate shocks borrow from either formal or informal lenders in rural Ethiopia.

The ERHS specifically asks households the type of income shocks. Table 3.8 shows

Table 3.8: Shocks by region

loss of ...	**Tigray**	**Amhara**	**Oromiya**	**Separ**
harvest	8.7	30.0	24.3	37.1
oxen	14.7	27.7	26.0	31.6
livestock	13.1	32.7	20.9	33.3
land	14.7	24.7	33.7	26.8
labour	7.9	19.0	26.0	47.1
assets	5.5	28.4	33.2	33.0
income from political event	6.3	24.2	25.9	43.5
income from military event	15.7	31.4	17.3	35.7

Source: Own calculation from ERHS. Note: % of borrowing households displayed.

several income shocks that affected borrowing households. A large proportion of borrowing households have been affected by income shocks in Amhara, Oromiya and Separ. For example, 47 percent of borrowing households have been affected by a loss of labour force in Separ[16]. About 33 percent of households who faced a loss of land and assets had access to credit in Oromiya.

What are the characteristics of the borrowing households?

Dasgupta et al. (2007) showed that rich households borrow only from formal lenders in urban Ethiopia. Households with an intermediate level of wealth borrow from both formal and informal sources. Poor households borrow only from informal lenders.

In table 3.9 we display some characteristics of households who borrow only from formal lenders, only from informal sources or from both formal and informal lenders in rural Ethiopia. We find that larger and older households (in terms of age of the household's head) borrow only from formal lenders or from both formal and informal sources.

In all regions except Separ, a large percentage of female headed households borrow only from informal lenders. We also find that richer households who have more land, more valuable assets and spend more on food and non-food items borrow only from

[16]A loss of labour force results from death or illness of a household's member.

Table 3.9: Selected characteristics of borrowing households

Characteristics:	**Tigray**	**Amhara**	**Oromiya**	**Separ**
Only formal				
Household size	6.1 (3)	6.8 (2)	8.1 (2)	9.6 (4)
Female head (%)	27.1 (45)	12.5 (34)	42.9 (51)	0.0
Age household head	52.7 (15)	41.8 (13)	53.4 (18)	48.4 (8)
Land size owned (ha)	0.5 (0.3)	4.4 (3)	2.6 (2)	1.1 (1)
Value of assets†	306.1 (273)	600 (470)	504.7 (450)	926.6 (687)
Food expenditure†	23.7 (18)	45.5 (29)	16.2 (9)	29.9 (16)
Non-food expenditure†	70.8 (94)	122 (92)	141 (134)	166.9 (93)
N. of observations	*59*	*16*	*14*	*7*
Only informal				
Household size	5.7 (3)	5.3 (2)	7.8 (4)	7.6 (3)
Female head (%)	39.8 (49)	24.6 (43)	20.5 (40)	11.9 (32)
Age household head	47.1 (16)	46.2 (15)	48.1 (15)	46.1 (15)
Land size owned (ha)	0.4 (0.3)	2.3 (2)	1.8 (1)	0.8 (1)
Value of assets†	298.9 (284)	415.5 (426)	491.3 (624)	441.5 (728)
Food expenditure†	36.6 (43)	31.5 (32)	27.7 (34)	21.5 (23)
Non-food expenditure†	87.1 (116)	98 (102)	212.4 (225)	133.7 (154)
N. of observations	*133*	*690*	*733*	*882*
Formal and informal				
Household size	8.8 (1)	6.3 (2)	10.7 (5)	8.0 (2)
Female head (%)	0.0	0.0	33.3 (48)	4.8 (21.8)
Age household head	57.0 (12)	48.1 (11)	46.2 (12)	48.3 (16)
Land size owned (ha)	0.6 (1)	2.5 (2)	3.3 (2)	0.8 (1)
Value of assets†	204.1 (88)	582.4 (528)	515.1 (429)	543.8 (327)
Food expenditure†	33.1 (12)	35.5 (24)	17.9 (15)	27.9 (28)
Non-food expenditure†	43.5 (46)	89.8 (84)	181.5 (146)	122.9 (71)
N. of observations	*4*	*11*	*21*	*21*

Source: Own calculation based on ERHS. Note: † Values in local currency (1 birr=0.1143$), deflated by using the Fisher Index (1997 base year). Expenditure also deflated by the square root of households' size. Standard deviation in brackets.

formal lenders in Amhara and Separ. On the other hand, in Tigrai and Oromiya households with more valuable assets and more land borrow only from formal lenders or from both formal and informal sources. In these regions, however, households who borrow only from informal lenders have higher food and non-food expenditure.

Table 3.9 has shown that average land size varies between different credit sources, although this does not demonstrate a causation between the two variables. In order to establish whether richer households have a higher probability to access collateralised lending (i.e. formal credit), in figure 3.4 we plot the size of total landholdings at time *t-1* against the predicted probabilities of the access to different loans sources for each loan "type" at time *t*.

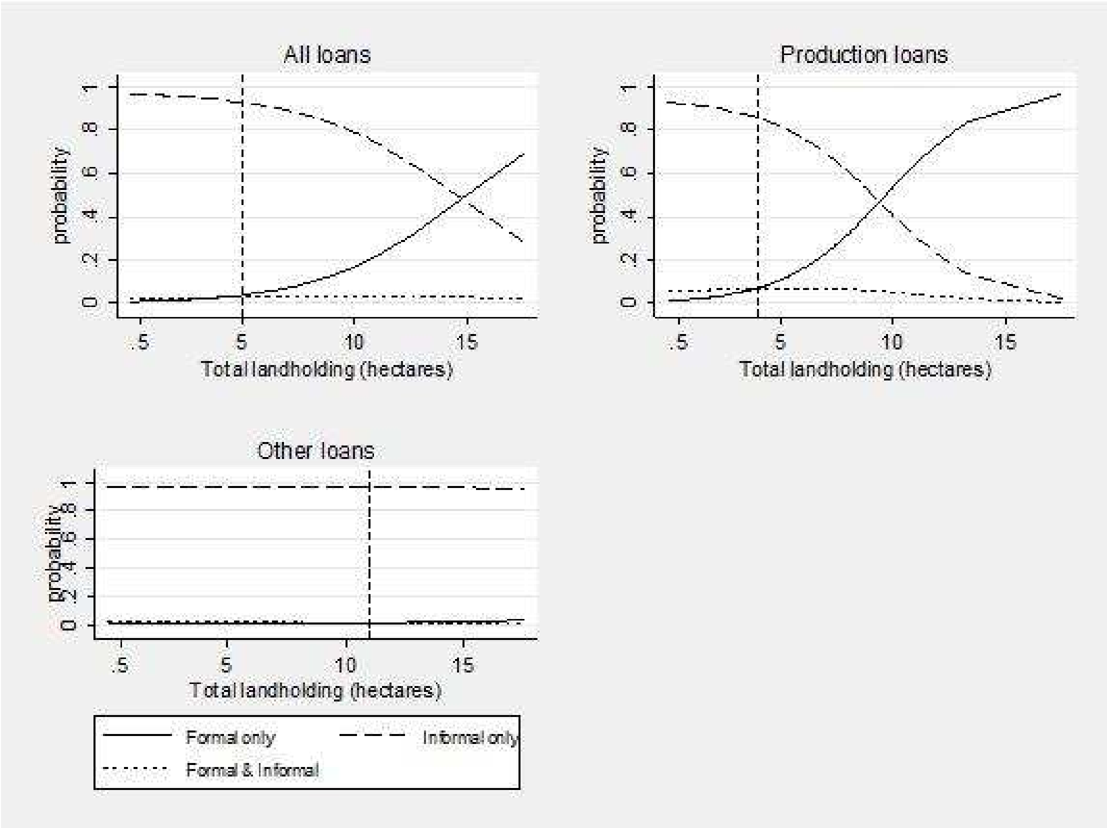
Figure 3.4 can also establish whether the probability of having access to different credit sources varies with the purpose for taking loans. As mentioned in the previous chapter, there is some literature that claims that different credit institutions serve different purposes [Aryeetey and Udry, 1995; Barslund and Tarp, 2006; Mohieldin and Wright, 2000]. For example, Mohieldin and Wright (2000) found that in Egypt the formal sector services loans for investment purposes, while the informal sector provides loans for consumption smoothing.

The predicted probabilities have been obtained from a series of multinomial regressions with different credit sources as the dependent variable and total landholding as the independent variable[17].

We have chosen this specification as opposed to a bivariate probit model with the view of looking at the determinants of borrowing from only formal or informal sources and of borrowing from both sources as a way of comparison with Dasgupta et al. (2007). The

[17] We have run three multinomial regressions according to the "types" of loans. The results are shown in table C3-6 in appendix C. The reduced form regressions report significance of lagged land size in the access to formal credit (with the exception of the category "other loans").

FIGURE 3.4: Predicted probabilities - different loan "types"



Source: Own calculation based on ERHS.

regressions have been confined to the subset of clusters where there are *equbs* in order to avoid the bias that arises from aggregating all clusters regardless of the availability of credit sources. Different loan "types" can be classified as follows: 1) production loans are used to buy farm equipment, to buy livestock or a new farm, to start a new business and to pay for hired labour; 2) "ceremonies" loans are used for weddings, funerals and health expenses; and 3) other loans are used to pay for taxes and other goods. Because loans used for weddings, funerals and health expenses are only supplied by informal lenders, we could not plot them against different sources.

For all loans, when households' landholding is greater than five hectares, the predicted probability of borrowing from only formal sources sharply increases, while the predicted probability of borrowing from only informal sources decreases. The predicted probability of borrowing from only formal sources exceeds the predicted probability of borrowing from only informal sources when land size is approximately fifteen hectares. However, only eight percent of the households actually have more than five hectares of land and less than one percent of respondents own more than fifteen hectares.

The second graph plots the predicted probabilities for production loans. Above four hectares the probability of borrowing from only informal sources declines, while the probability of borrowing from only formal sources increases. Whenever total landholding is higher than ten hectares the probability of borrowing from formal sources exceeds the probability of borrowing from informal lenders. The predicted probability of borrowing from both formal and informal sources slightly decreases when households' landholding is greater than eight hectares.

In the last graph the predicted probability of borrowing from different sources is less steep. Above ten hectares the probability of borrowing from only informal sources slightly decreases and the probability of borrowing from only formal sources slightly increases.

We find that the probability to borrow from formal lenders is higher when loans are used for production than for other purposes. The probability to borrow from the informal sector, by contrast, is higher for loans not used for farm investments.

What "story" can be told from these descriptive statistics?

In conclusion, the availability of credit varies across southern and northern regions of rural Ethiopia. According to the economic and cultural approaches explained in the previous chapter, there are indeed specific economic and social characteristics that affect households' participation in informal arrangements. Southern rural Ethiopia is the most densely populated area. Households have more children and own more valuable assets. Also, we observe a considerably higher food and non-food expenditure. The most common sources of credit are friends, relatives and moneylenders. This can be explained by the fact that in southern Ethiopia household membership is not related to the economic status of the member, but to kinship which establishes an "obligation" to reciprocate [Bevan and Pankhurst, 1996]. A self-help institution (i.e. *equb*) is only available in southern rural Ethiopia.

Households' participation in informal credit can be also related to the fact that borrowing is purpose-oriented [Aryeetey and Udry, 1995; Barslund and Tarp, 2006; Mohieldin and Wright, 2000]. In Ethiopia, rural households borrow to buy inputs, farms, or to pay for health expenses, with some differences across regions. In the regions where NGOs (i.e. in Tigray) or banks (i.e. in Separ) are available, rural households borrow from them to buy farm inputs. Mostly, households prefer to borrow from friends and relatives.

Informal loans are less likely to be used for farm investments. They are also less likely to require land (larger land size does not increase the probability of borrowing). On the other hand, the more land households own, the higher is the probability to borrow from only formal lenders, especially when considering production loans. Above a threshold of approximately ten hectares the probability of borrowing from formal lenders sharply exceeds the probability of borrowing from informal credit institutions. We find that richer households, who have more valuable assets and spend more on food and non-food items, borrow only from formal lenders or from both formal and informal sources in rural Ethiopia.

Finally, the data supports the hypothesis that either formal or informal credit is used to cope with shocks especially when they are idiosyncratic. This result supports the literature on risk-coping strategies mentioned in the previous chapter [e.g. Bardhan and Udry, 1999; Binswanger and Rosenzweig, 1986; Platteau and Abraham, 1987; Ruthenberg, 1971; Townsend, 1994].

3.4 Econometric Analysis

How can we model households' decision to participate in informal credit?

There are several issues to be considered. First, there is selection bias arising from the fact that, when modelling the amount of debt, those who demand some credit are not representative of the full sample, but systematically differ from the full sample. Second, the participation decision depends on the wealth characteristics of the household, but individual regressors tend to be highly collinear. Third, participation in informal credit depends on the relative substitutability of specific credit sources and their availability in different clusters. Fourth, there are dimensions of heterogeneity of access to credit - most notably social, economic and geographic - which may operate at the cluster level, but are not identified at the household level. Finally, the extent to which shocks affect the cluster and the household not only determines the access to credit but also the choice of a specific credit source.

The aim of this chapter is to show the drawbacks of standard modelling approaches of the participation in informal credit. We argue that an endogenous switching regression model is a superior specification in addressing each of the above mentioned issues. We lead to this empirical specification by using two models: the logit and the Heckman selection model. This approach allows us to highlight the advantages of the endogenous switching regression model compared to the reduced form logit specification and the Heckman model whenever selection bias is not severe.

In sub-section 3.4.1 we show how some of the problems of standard approaches can be overcome through two specifications. The first one addresses collinearity of wealth-holding variables and expenditure by adopting principal components analysis. The second specification shows that cluster differences are significant in explaining participation in informal credit. The underlying assumption of the logit models is that the availability of informal credit sources is exogenous to cluster level and household level characteristics.

The Heckman selectivity model deals with the possible selection bias arising from the fact that, when modelling the amount of informal credit, we only observe those who have positive debt and not those who, despite having positive propensity to borrow, did not get any credit. However, this model does not explain the factors that cause the access to a specific credit source to differ across clusters (sub-section 3.4.2).

After having shown the drawbacks of the previous approached, we conclude the empirical part with our main results shown by the endogenous switching regression model. We claim that endogeneity in the availability of a particular type of informal finance affects households' participation in the informal credit sector. The endogenous switching regression model is able to address this issue thus solving the identification problem of the demand for specific informal credit sources (sub-section 3.4.3).

3.4.1 Standard approaches: logit models

This sub-section models the probability of taking out informal credit as a function of shocks and of households demographic and collateral characteristics. The empirical specification includes the following wealth and expenditure variables: equipment, house assets (i.e. furniture) and valuables (i.e. jewels and gold), value of livestock, land size, number of plots and quantity of harvested crops, food and non food expenditure. We adopt principal components analysis to avoid collinearity between wealth variables. Principal component analysis is a statistical technique that linearly transforms a set of correlated variables into a smaller set of uncorrelated components. Appendix B describes how we constructed principal components of wealth variables in the ERHS.

By comparing a model that specifies the determinants of households' probability to access informal sources with a model that creates components for a subset (i.e. wealth and expenditure) of the same determinants, this sub-section shows that collinearity and the number of variables in the model can be drastically reduced[18]. A similar approach has been adopted in the Demographic Health Survey (2006) to create socioeconomic indicators in Ethiopia[19].

Suppose that households' participation in informal arrangements depends on a set of households' characteristics. The standard random utility model (RUM) argues that a household borrows from informal sources if the utility of borrowing is greater than the utility of not borrowing [McFadden, 1984].

In table C3-7 we report a logit model. We chose the logit model as opposed to a probit model because it allows an easier interpretation of the coefficients in terms of odds ratios. The model can be written in the following form:

$$\Pr\left(I_{i,t}=1|X_{i,t},Z_{i,t},S_{i,t}\right)= \text{ F}\left(\alpha_i+\beta X_{i,t}+\gamma Z_{i,t}+\vartheta S_i+\sum_{s=1}^{T-1}\varphi_s\tau_{i,s}+\xi P_i\right)$$
$$\forall \quad i=1,\ldots,N \text{ and } t=\text{1994a,1994b,1995,1997} \tag{3.1}$$

where subscript i indicates each household. The dependent variable, I_i, indicates whether household i at time t borrows from informal sources. Household-specific characteristics[20] $X_{i,t}$, include age of household head and its squared value, household size and its squared value, a dummy indicating whether the household is female headed, number of children between 0 and 5, 6 and 10, 11 and 17 years old, a dummy indicating whether the head of the household has attended school and whether the household head belongs to an ethnic minority. Principal components of assets and expenditure variables

[18] Indeed, wealth variables are highly correlated with each other as shown in table B3-1 of appendix B.

[19] "Constructing socio-economic status indices: how to use principal components analysis" [Vyas and Kumaranayake, 2006].

[20] Some of them (i.e. gender) do not depend on time.

are denoted by $Z_{i,t}$.

The probability of borrowing from informal sources depends on a number of demand shocks (for example, diseases that affect the harvest, land lost for disputes with relatives and illness of the husband). S_i is a vector indicating whether household i has been affected by one of the above mentioned demand shocks in the last twenty years. Time and cluster dummies are denoted by τ and P respectively.

The results are shown in table C3-7 of the appendix C because they represent an inferior model that allows us to justify the model specification in the following sections. We compare a standard regression (model I) with a principal component regression (model II)[21]. Standard errors have been adjusted by using the robust and cluster option in Stata, which is a generalization of the Huber/White/Sandwich (HWS) estimate of variance [Deaton, 1997; Rogers 1993; Williams 2000][22]. It obtains robust variance estimates that adjust for within-cluster correlation. The principal component regressions are also adjusted for the inclusion of principal components.

In the first model, most of the assets and expenditure variables (except for land size and harvested crops) are not significant. By including components' scores, model II produces a significant assets and expenditure indicator. In particular, the third component places more weight on the quantity of harvested crops[23]. Its correlation[24] with the probability of borrowing from informal sources is negative and significant (the coefficient is 0.70). The fact that the first component is not significant does not necessarily mean that principal components analysis is not useful for several reasons. First, principal components have the advantage of reducing the number of variables in a model in

[21]Note that the number of observation is low because only 3,000 households answer the question of informal credit. Then missing observation on assets causes further reduction in the number of observations. As explained in the descriptive section, non-random missingness is not a testable assumption.

[22]Note we have not corrected the standard errors for the inclusion of three components rather than the full number of components that explain the entire variance of the regressors. However, we have included the three components that explain most of the variance.

[23]Appendix B explains how to interpret each component score in the ERHS.

[24]In the following sections we explain why we do not try to describe causality.

addition to solving collinearity problems. Second, although the first component explains most of the variance of the regressors, according to the scree plot criterion illustrated in Appendix B all the three components should be included and considered altogether as reducing the variance of the individual regressors. Third, regardless of the proportion of the variance explained, each component places different weights to the variables explained. The fact that only the third component may indeed indicate that quantity of harvested crops is the most important explanatory factor. Finally, an F-test of joint significance shows that each of the weatlh variable is jointly significant. This means that we cannot simply exclude insignificant regressors but we have to take into account that they may be correlated between each other[25].

Is the previous specification the most appropriate to identify access to informal credit? It implies that the determinants of credit are common across all localities. But because not all informal sources are available across rural Ethiopia, we proceed with a test of differences between two groups of clusters: northern rural Ethiopia (where there are no *equbs*) and southern rural Ethiopia (where there are *equbs*)[26].

A standard Chow-type test can be used to test for differences in the slope parameters across the two groups[27].

The models in table C3-8 of Appendix C have been estimated by using a logit specification for consistency with the other models throughout the chapter. The probability of borrowing from informal lenders in each group of clusters solely depends upon households' characteristics, components of wealth-holdings and shocks. The Chow-type (likelihood-ratio) test reported in table C3-8 rejects the hypothesis that the estimated coefficients (i.e. household characteristics, wealth components and shocks) are equal

[25]The test statistic is significant at the one percent level with a value of the statistics equal to 120.74.
[26]Figure 3.2 shows that all informal credit sources except *equbs* are available in all clusters.
[27]It can be shown that a Chow-test on the two groups is equivalent to an F-test of joint significance of each additional coefficient in a regression that nests the restricted regression. The proof is provided in appendix B.

across clusters where *equbs* are either present or not. This shows that a model of the participation in informal credit cannot aggregate different clusters. The next two sub-sections suggest ways to overcome this problem.

3.4.2 Selectivity models

This section looks not only at the probability of taking out a loan but also at the quantity of loan taken. There are three issues that arise in modelling the amount of informal credit held by households. First, the amount of informal debt observed in practice results from the interaction of demand and supply. Second, in models dealing with the amount of informal credit, only those who actually applied for informal credit are retained in the sample. When the dependent variable measures values, the standard OLS regression is subject to possible sample selection bias. Finally, in light of the results presented in the previous section, an analysis of the access to informal sources that aggregates the data at the national level ignores clusters' heterogeneity and leads to biased results.

In this sub-section we address the above mentioned issues by modelling the amount borrowed from the informal sector with a selection equation for access to credit in clusters with and without *equbs*. We do not use a likelihood ratio test to compare this model with the previous ones because selectivity models adopt a different approach adding to the analysis of the probability to take a loan, the analysis of the quantity of credit. The aim of this section is rather more general. Given that in the previous analysis we have shown that clusters cannot be aggregated, we now use the selectivity model to analyse the determinants of informal credit (both the probability of taking out credit and its quantity) in each of the two groups of clusters with and without *equbs*.

Suppose that the amount household i borrows from informal sources can be represented by the following equation[28]:

$$\ln\, Q^{I*}_{i,t} = \alpha_{0i} + \beta X_{i,t} + \delta D^{F}_{i,t-1} + \gamma Z_{i,t} + \vartheta S_i + \sum_{s=1}^{T-1} \varphi_s \tau_{i,s} + \xi south + u_{i,t}$$

$$\forall \quad i = 1, \dots, N \text{ and } t = \text{1994a,1994b,1995,1997} \qquad (3.2a)$$

where the included regressors are the same as the ones in model 3.1. In model II of tables 3.10 and C3-9 to C3-11 in appendix C, we partition the vector of lagged formal credit dummies in $D^{F}_{i,t-1} = [\text{Bank}_{t-1}, \text{NGO}_{t-1}]$. Also, we have included dummies for the rounds (τ) and a dummy indicating whether household i lives in the South (*south*) in clusters where *equbs* exist and dummies for specific peasant associations in clusters where *equbs* do not exist. The selection equation can be defined as:

$$I^{*}_{ij(i),t} = \alpha_{1i} + \beta X_{i,t} + \vartheta_1 S_i + \chi C_{j(i)} + v_{ij(i),t}$$

$$\forall \quad i = 1, \dots, N \text{ and } j(i) = 1, \dots, 15; \quad t = \text{1994a,1994b,1995,1997} \qquad (3.3b)$$

where:

$$I_{ikj(i),t} = 1.(I^{*}_{ij(i),t} > 0)$$

$I^{*}_{ikj(i),t}$ indicates the latent demand for informal credit with $k = 1$ for southern Ethiopia (where there are *equbs*) and k=2 for northern Ethiopia (where there are no *equbs*). In other words, with this equation we determine who applied for informal credit by looking at the probability that households borrow from informal sources in the two groups of clusters with and without *equbs*.

This specification has two advantages. First, it allows for cluster-level variations in

[28] A full derivation of the general model is described in Appendix B.

borrowing strategies irrespective of the availability of a specific informal credit source. Second, it also specifically identifies the fact that substitutability between credit sources is different in clusters with and without *equbs* thus affecting the demand for informal credit itself.

The error terms u_i and v_i have a bivariate normal distribution with covariance cov $(u_i, v_i) = \sigma_{uv}$. The observability criterion for the selectivity model is:

$$Q_i^I = Q_i^{I*}.1(I_i^* > 0) \tag{3.3}$$

that is, we only observe the amount of credit of those who borrow from informal sources. We cannot observe those households who, despite having positive propensity to borrow, could not have access to credit (i.e. rationed households). In other words, the sample of households is affected by a selection problem [Heckman, 1979].

There are two ways in which this model can be estimated: a) by using a full-information maximum likelihood (FIML) selectivity model; or b) by using a two-step selection model. Table 3.10 reports the results of the two-step estimation in clusters with *equbs*[29]. The following analysis focuses on the two-step model because the hypothesis of independent equations could not be rejected[30].

Identification requires that the selection equation 3.3b includes at least one regressor that is not present in equation 3.2a. Indeed, cluster-specific characteristics, $C_{j(i)}$, and a dummy indicating idiosyncratic shocks are assumed to affect the probability of borrowing from informal lenders. The vector $C_{j(i)}$ represents characteristics that vary only across clusters j, but not across households (i.e. number of villages, distance to the

[29]The two-step estimation for clusters without *equbs* is reported in table C3-11 in appendix C. It will not be discussed here because the first-stage results are the same as the ones presented in the text, but with opposite sign. Tables C3-9 and C3-10 - namely, those reporting the FIML results - have been placed in appendix C because they are not much different from the two-step estimation.

[30]The likelihood ratio test of independent equations has been rejected only at the 10% level in Model II when clusters with *equbs* have been selected (see table C3-9).

Table 3.10: Selectivity models - 2 Step estimation (PA has Equbs)

Log(informal credit)	**Model I**		**Model II**	
	1^{st}*stage*	2^{nd}*stage*	1^{st}*stage*	2^{nd}*stage*
hh characteristics:				
age head	0.01	0.02	0.01	0.03
	(0.02)	*(0.01)*	*(0.02)*	*(0.01)**
age head squared	-0.0002	-0.0002	-0.0001	-0.0003
	(0.00)	*(0.00)**	*(0.00)*	*(0.00)**
hh size	0.33	0.03	0.34	0.03
	*(0.05)****	*(0.03)*	*(0.05)****	*(0.04)*
hh size squared	-0.004	-0.001	-0.01	-0.002
	*(0.00)***	*(0.00)*	*(0.00)***	*(0.00)*
female head	0.06	-0.03	0.09	0.04
	(0.11)	*(0.06)*	*(0.13)*	*(0.10)*
number of children	-0.19	-0.001	-0.19	0.03
	*(0.04)****	*(0.02)*	*(0.04)****	*(0.03)*
head schooling	0.97	-0.001	1.05	0.01
	*(0.11)****	*(0.06)*	*(0.13)****	*(0.09)*
head ethnic minority	0.25	-	0.32	-
	*(0.12)***		*(0.14)***	
bank (lagged)	-	-	-	-0.38
				(0.45)
NGO (lagged)	-	-	-	1.16
				*(0.63)**
PCs of hh assets:				
assets & exp. (pc1)	-	0.17	-	0.17
		*(0.01)****		*(0.02)****
assets & exp. (pc2)	-	-0.08	-	-0.06
		*(0.02)****		*(0.04)**
assets & exp. (pc3)	-	0.02	-	0.07
		(0.02)		*(0.04)**
shocks:				
household only	0.45	-	0.53	-
	*(0.09)****		*(0.11)****	
land slide	-	0.59	-	0.73
		*(0.26)***		*(0.33)***
harvest diseases	-	-0.07	-	-0.21
		(0.05)		*(0.07)****
land taken by cooperative	-	-0.07	-	0.87
		(0.52)		*(0.90)*
head imprisoned	-	0.30	-	-0.87
		(0.52)		*(0.91)*
assets resettlements	-	-0.33	-	-1.54
		(0.64)		*(0.90)**
banditry		-1.39		-1.68
		(0.90)		*(0.91)**
PA characteristics:				
n. villages in PA	0.09	-	0.10	-
	*(0.01)****		*(0.01)****	
dist. nearest bank all weather road	0.01	-	0.01	-
	*(0.00)****		*(0.00)****	
n. of agricultural offices in PA	0.25	-	0.13	-
	*(0.10)***		*(0.12)*	
irrigated land (ha)	0.001	-	0.001	-
	*(0.00)****		*(0.00)****	
rain fed land (ha)	0.002	-	0.002	-
	*(0.00)****		*(0.00)****	
south	-	0.17	-	-0.16
		*(0.07)***		*(0.12)*

round 2	-	-0.87	-	-0.20
		*(0.06)****		*(0.09)***
round 3	-	-0.65	-	0.01
		*(0.07)****		*(0.09)*
round 4	-	-0.61	-	-
		*(0.06)****		
constant	-3.21	5.00	-3.82	4.41
	*(0.46)****	*(0.25)****	*(0.56)****	*(0.40)****
Mills ratio	-0.14		-0.22	
	(0.11)		*(0.12)**	
N. Obs	1,940		1,063	

Source: own calculation from ERHS. Standard errors in parenthesis. †p-value
***$p < 0.01$,**$p < 0.05$,*$p < 0.1$

nearest bank interacted with a dummy indicating whether there is an all weather road, number of agricultural offices and size of irrigated and rain fed land in hectares). We thus claim that the chosen cluster characteristics are exogenous and do not affect unobservable factors included in the quantity of credit. For instance, we can think of the distance to the bank as a quasi-experiment where location of the household is exogenous to household choice. We also use an individual level variable such as idiosyncratic shocks as an addictional selection variable.

In addition, the probability of borrowing from informal sources depends on a set of households' characteristics (i.e. age, household size, number of children and dummies indicating whether the household is female headed, whether the household head has some school education and whether he/she belongs to an ethnic minority). We included a dummy that takes value one when the household has been affected by an idiosyncratic shock. Given that household i borrows from informal sources in cluster j with or without *equbs*, the amount of credit (in logarithm) depends on assets and expenditure components, as well as on households' characteristics and shock dummies.

For each group of clusters with and without *equbs* we have estimated two models. The first model includes the set of covariates described above. The second model adds dummies for participation in formal credit (banks and NGOs). We lagged these dummies by one period in order to avoid reversed causality with the dependent variable.

Theory leaves the sign of the relationship between informal and formal sectors indeterminate [McKernan et al., 2005]. Formal credit may substitute for informal arrangements, but may also be complementary. Assets acquired through formal credit may improve the credit-worthiness of households increasing their access to informal loans. On the other hand, some literature found evidence of crowding out, that is, increases in access to informal credit result in reductions of formal loans, or *vice versa*. In order to test the crowding out hypothesis, McKernan et al. (2005) modelled informal transfers in Bangladesh as a function of credit programs. They avoid endogeneity of formal credit by using a quasi-experimental approach on the basis of the programs eligibility criteria [following Pitt and Khandker, 1998]. Our approach is similar in that it also models informal arrangements as a function of formal credit, but we deal with endogeneity by lagging formal credit dummies.

In tables 3.10 and C3-11 in appendix C we report the two-step Heckman models for clusters with and without *equbs*, respectively. As the coefficients of the first stage regressions in the two groups of clusters have opposite signs, but the same value, we hereby report the results for the clusters with *equbs* and comment on the reasons for which the coefficients differ in sign.

Considering the latent demand for informal credit (first stage regression), we show three sets of results entailing households' characteristics, incidence of shocks and clusters' characteristics.

With regard to households characteristics, we find that the probability of borrowing from informal sources increases when the household head belongs to an ethnic minority[31] in clusters where there are *equbs* and decreases in clusters where there are no *equbs*. This result can have several explanations[32]. Firstly, for borrowing households the existence of an additional credit source such as *equbs* changes the relative substitutability between different informal sources. Secondly, the existence of *equbs* signals that different socio-economic characteristics of the two groups of clusters might affect the borrowing behaviour. As mentioned by Raturi and Swami (1999) credit markets may discriminate in terms of ethnicity. Members of ethnic minorities perceived to be dishonest or unproductive may be discouraged to take loans. For example, Munnel et al. (1996) found in U.S. that African-American applicants are less likely to receive loans *ceteris paribus*. Fafchamps (1997) and Raturi and Swami (1999) found that in Zimbabwe, black-owned firms, are substantially less likely to receive credit. According to La Ferrara (2003), ethnic minorities may be excluded from other sources of informal credit and they may rely on self-help groups.

Other households' characteristics are significant. For example, household size increases the probability of borrowing from informal sources in clusters where there are *equbs*, but at a decreasing rate (i.e. the coefficient of the squared value is negative).

The fact that the household head has some school education has a positive and highly significant (at one percent level) impact on the probability of borrowing from informal lenders. Again, the coefficient in clusters with *equbs* displays an opposite sign to the one in clusters with no *equbs*. It may be a result of unobservable cluster differences or it may be explained by the fact that some education is required to participate in *equbs* where usually one member is supposed to keep track of the other members' contributions.

With regard to the incidence of shocks, we find that when household i has been

[31]That is, when his or her ethnicity is not prevalent in that cluster. This variable results from a combination of individual level data (ethnic group of household head) and cluster level data obtained fom the village studies (prevalent ethnicities in the clusters).

[32]Because the availability of *equbs* is not random, we cannot attribute cluster differences only to the existence of *equbs*. The fact that households endogenously choose to set up a RoSCA group could rather indicate that cluster-specific socioeconomic characteristics affect this choice.

affected by an idiosyncratic shock, the probability of borrowing from informal sources increases in clusters with *equbs*. This result confirms the well-established literature arguing that aggregate shocks impede risk pooling strategies [Bardhan and Udry, 1999; Hoddinott et al., 2005; Ray, 1997]. However, the coefficient is negative when considering clusters with no *equbs*. There could be three explanations for this result: a) the existence of *equbs* facilitates risk pooling strategies when shocks are idiosyncratic; b) *equbs* exist in clusters that are more prone to idiosyncratic shocks; and c) the existence of *equbs* signals a society where mechanisms of reciprocity are more common. According to van Bastelaer (2000), RoSCAs can be seen as "a widespread way to crystallize social relations in an informal - yet often formally run - system of internal credit delivery". Van Bastelaer (2000) pointed out that RoSCAs help its members to build up trust.

Finally, we find that all peasant associations' (PA) characteristics significantly affect the probability of borrowing from informal sources in clusters with and without *equbs*. For example, the larger the distance to the bank, the higher is the probability of borrowing from informal sources in clusters where there are *equbs* (the coefficient is positive and highly significant). This variable has been interacted with a dummy indicating whether there is an all-weather road because distance itself may not reveal the accessibility of banks[33]. The same coefficient in table C3-11 is negative and significant.

Also, in clusters where there are *equbs* the size of irrigated and rain fed land has a positive (negative in clusters with no *equbs*) effect on the probability of borrowing from informal lenders. As mentioned earlier, this result might reflect the existence of a more developed farming society which, in turn, affects access to informal credit and in particular to *equbs*. The same explanation could be used for the positive coefficient on the number of agricultural offices in model I of table 3.10.

[33] A mud road relatively close to the bank may be less accessible than a far all weather road.

From the second stage of the Heckman model we single out four main results regarding households' characteristics, collateral components, extent of shocks and substitutability between formal and informal sources. For brevity purposes in the following discussion we only comment on the results for clusters with *equbs* (table 3.10).

Model II of table 3.10 shows that the amount of credit households borrow from informal lenders increases when the age of the household head increases, but at a decreasing rate (the squared value is negative).

With regard to collateral characteristics, we find that principal components are significant. The first component indicates that an overall increase in assets and expenditure is positively correlated[34] with the amount of credit obtained from informal sources in clusters with and without *equbs*. The second component indicates that the more farm assets (i.e. land) the household has, the lower the amount of credit borrowed from informal lenders. Wealthier households borrow less from informal sources and may have access to formal loans. The third component is only weakly significant in model II of table 3.10. It indicates that the quantity of harvested crops is positively related to the amount of informal debt.

Most of the shocks are significant in model II of table 3.10. Shocks which are more likely to affect the entire community such as harvest diseases, banditry and resettlement of assets have a negative impact on the amount of informal debt. The opposite is true for shocks that are more likely to be idiosyncratic (i.e. land slide).

Finally, we find no evidence of crowding out[35]. The lagged dummy indicating whether household i borrowed from banks[36] has no impact on the amount borrowed from informal lenders in clusters with and without *equbs*. This result can have two explanations.

[34]We do not talk about causation here, because there could be reversed causality between components of expenditure/wealth and credit.

[35]However, as we will show in the fourth chapter, a *causal* test of crowding-out should find appropriate counterfactuals.

[36]The NGO dummy is positive but significant only at the 10% level in model II of table 3.10.

First, formal and informal loans may be independent of each other because they are purpose-oriented (as explained in the previous chapter). Second, the result may indicate that there is no long-run effect of formal credit on access to informal loans, but there might be short-run effects that are not captured by the lagged variable[37].

3.4.3 Endogenous switching regression models

Two issues have emerged from the previous models. First, there are significant cluster differences in the participation in informal credit arrangements. These differences reflect socio-economic factors that are endogenous to the clusters themselves. Second, we find that those who demand informal credit are representative of the full sample of borrowing and non-borrowing households. In other words, the two-step Heckman models show no evidence of selection bias[38].

As there is no evidence of selection bias, we claim that the most appropriate model of households' participation in informal arrangements in rural Ethiopia is a switching regression with endogenous criterion [Lee, 1978; Maddala, 1983]. The endogenous switching regression models for mixed continuous and discrete variables consist of joint estimation of the probability that in cluster j *equbs* are available and the amount of informal credit borrowed.

Following Duong and Izumida (2002), we estimate an endogenous switching regression by using a two-step Heckman model where the selection equation determines the switching group[39]. In other words, this model allows to take clusters heterogeneity into account by substituting the selection equation of the previously estimated Heckman

[37] Because the lagged dummy refers to the previous round which can be up to two years before time t.

[38] That is, we can model the amount of credit from informal lenders without worrying about the selection of households who have access to credit. The inverse Mills ratio is not significant in any of the models except in model II of table 3.10 where its significance level is low (10 percent).

[39] In table C3-12 we also show the results for the one step Heckman model with lagged formal credit dummies in clusters with and without *equbs*. The results are not very different.

model (which we have shown that does not cause selection bias) with the availability of a particular type of informal credit which is endogneously determined by the characteristics of the clusters themselves. More formally[40], let E^* be the function of a vector of the exogenous household socioeconomic situation and clusters characteristics:

$$E^* = \alpha_{0i} + \beta X_i + \vartheta I_i + \chi C_{j(i)} + \sum_{s=1}^{T-1} \varphi_s \tau_{i,s} + v_{ij(i)} \tag{3.4a}$$

Define $E_{j(i)} = 1$ when cluster j has *equbs* iff $E^* > 0$ and $E_{j(i)} = 0$ when cluster j has no *equbs* iff $E^* \leq 0$, where *j(i)* indicates the jth cluster where household i lives. Households' characteristics are defined by X; I is a dummy indicating whether household i has been affected by an idiosyncratic shock and C is a vector of cluster characteristics.

The model can be postulated for any household i [Lee, 1978; Maddala, 1983]:

$$\ln Q^{I*}_{1i,t} = \alpha_{1i} + \beta_1 X_{1i,t} + \delta_1 D^F_{1i,t-1} + \gamma_1 Z_{1i,t} + \vartheta_1 S_{1i} + \sigma_{1v}\lambda_{1i} + \sum_{s=1}^{T-1} \varphi_s \tau_{i,s} + \xi_1 south + u_{1i,t} \text{ iff } E_{j(i)} = 1$$

$$\ln Q^{I*}_{2i,t} = \alpha_{2i} + \beta_2 X_{2i,t} + \delta_2 NGO_{i,t-1} + \gamma_2 Z_{2i,t} + \vartheta_2 S_{2i} + \sigma_{2v}\lambda_{2i} + \sum_{s=1}^{T-1} \varphi_s \tau_{i,s} + \xi_2 P + u_{2i,t} \text{ iff } E_{j(i)} = 0$$

$$\forall \quad i = 1, \ldots, N \text{ and } t = \text{1994a,1994b,1995,1997} \tag{3.5b}$$

where *X, Z, S* and τ have been defined in sub-section 3.4.1. We include the partitioned vector of lagged formal credit dummies $D^F_{i,t-1} = [\text{Bank}_{t-1}, \text{NGO}_{t-1}]$ in clusters where *equbs* exist. Because households do not borrow from banks in northern Ethiopia we only include a lagged dummy for access to NGOs in clusters where *equbs* do not exist. In order to avoid collinearity, we include dummies of peasant associations instead of the dummy "south" in clusters where *equbs* are not available. The inverse Mills ratio is denoted by λ. Q^{I*}_{1i} and Q^{I*}_{2i} are the two possible values of the dependent variables - amount borrowed from informal lenders - depending on the values of E^*.

By using Monte Carlo simulations, Kimhi[41] (1999) pointed out that standard errors

[40] Omitting time subscripts.
[41] The proof entailed endogenous switching regression models with discrete variables.

should be corrected when estimating a two-step endogenous switching regression. Table 3.12 displays models with and without the standard error correction obtained by using bootstrapping methods[42].

The first stage regression - namely, explaining whether the cluster has *equbs* - in table 3.11 also adopts a standard errors correction for intra-cluster correlation[43]. Equation 3.4a is used to estimate the first-stage regression (i.e. probability that there are *equbs* in cluster j)[44]. The switching regression is a function of households' characteristics (i.e. age, gender, schooling and ethnicity of household head, household size and a dummy indicating whether the household has been affected by an idiosyncratic shock) as well as cluster-specific characteristics (number of villages and agricultural offices in the PA, size of irrigated and rain fed land, distance to the nearest bank interacted with a dummy indicating whether there are all-weather and dry roads). Hence, table 3.11 reports the factors affecting the formation of *equbs* in southern Ethiopia.

As Carpenter and Jensen (2002) pointed out, the formation of RoSCAs is affected by two factors. Firstly, there must be a sufficiently large number of people, living in the same location (or in the vicinity), who are willing to form a group. However, the likelihood that RoSCAs exist does not monotonically increase with the number of people. There will be a turning point at which the increase of people will not allow social enforcement and screening [Ghatak and Guinnane, 1999]. Secondly, additional factors such as income sources and variability affect group formation. Indeed, variability of income is strictly linked to the extent of shocks. Villages that are affected by aggregate

[42]Kimhi (1999) used the Murphi-Topel (1985) correction to take into account the fact that the second stage regression included a predicted term from the first stage estimation. However, in sufficiently large samples bootstrapping gives asymptotically equivalent results.

[43]The table does not show the marginal effects for comparability purposes with the STATA generated two-step Heckman. Both sign and significance in the marginal effects do not differ from the standard coefficients. The number of observations is very different because model I has been estimated in conjunction with the second stage of the two-step Heckman.

[44]We do not show the results for clusters without *equbs* because they are exactly the same as the ones in table 3.11, but with opposite sign.

TABLE 3.11: Endogenous switching regression models (first stage)

Pr(PA has Equbs)	**Model I: uncorrected std. errors**	**Model II: cluster-corrected std. errors**
hh characteristics:		
age head	0.03	0.01
	(0.02)	*(0.01)**
age head squared	-0.0004	-0.0001
	*(0.00)**	*(0.00)**
hh size	0.52	0.33
	*(0.06)****	*(0.12)****
hh size squared	-0.01	-0.002
	*(0.00)***	*(0.00)*
female head	0.22	0.01
	*(0.13)**	*(0.03)*
number of children	-0.28	-0.21
	*(0.05)****	*(0.05)****
head schooling	1.65	1.52
	*(0.14)****	*(0.21)****
head ethnic minority	1.68	1.54
	*(0.22)****	*(0.41)****
household only (shock)	0.71	0.55
	*(0.11)****	*(0.09)****
PA characteristics:		
n. villages in PA	0.27	0.24
	*(0.03)****	*(0.19)*
distance to nearest bank	0.07	0.07
***all weather road**	*(0.01)****	*(0.03)***
n. agricultural offices	-0.30	-0.36
in PA	*(0.16)**	*(1.02)*
irrigated land (ha)	0.01	0.01
	*(0.00)****	*(0.00)***
rain fed land (ha)	0.004	0.004
	*(0.00)****	*(0.00)***
round 2	0.69	0.03
	*(0.12)****	*(0.01)****
round 3	0.83	0.08
	*(0.12)****	*(0.03)****
round 4	-	-0.17
		*(0.05)****
Constant	-9.23	-6.46
	*(0.72)****	*(2.03)****
N. Obs.	1,612	5,003

Source: own calculation from ERHS. Note: std. errors in parentheses adjusted for within-cluster correlation in model II. ***$p < 0.01$,**$p < 0.05$,*$p < 0.1$

shocks will not allow pooling of resources hence discouraging group formation. In other words, *equbs* have an insurance role in clusters that are less prone to aggregate shocks. Anthropologists also argue that the existence of *equbs* might be linked to immigration (i.e. more accessible villages had contact with immigrants who used RoSCAs) or to a more developed society where cash is available.

We discuss three main results from the first stage estimation of the demand for *equbs* regarding households' characteristics, incidence of shocks and clusters' characteristics.

With regard to households' characteristics, table 3.11 shows that as household size increases, the probability that the PA has *equbs* increases as well (with decreasing rate in model I). The age of the household head positively and significantly (only at the ten percent level in model II) affects the existence of *equbs*, but at a decreasing rate. This effect is however quite small. The number of children has a negative and highly significant impact on the probability that *equbs* exist in the PA. As Carpenter and Jensen (2002) pointed out, it is the number of adults that should affect the existence of *equbs*.

The probability that the PA has *equbs* is also positively and significantly affected by the fact that the household head has some school education and belongs to an ethnic minority. Credit markets may indeed discriminate in terms of ethnicity [Raturi and Swami, 1999]. Hence, members of ethnic minorities excluded by other credit sources may be more willing to form self-help groups. Unfortunately, there is no data about group members, but it is very likely that *equbs* are formed among homogenous ethnic members [Ghatak and Guinnane, 1999; La Ferrara, 2003].

The incidence of shocks affects the demand for insurance arrangements such as *equbs*. We find that the existence of idiosyncratic shocks positively affects the probability that the PA has *equbs*. As mentioned in the previous chapter, the creation of risk pooling strategies depends on whether shocks affect the entire community or not [Bardhan and Udry, 1999; Hoddinott et al., 2005; Ray, 1997].

Not only do households' characteristics, but also clusters' characteristics affect the availability of *equbs*. In this first stage regression we single out three factors: demographics, infrastructures and geographical characteristics.

First, as the number of villages in the PA increases, the probability that *equbs* exist increases. The existence of risk pooling strategies is, in fact, affected by the diversification of incomes of participants [e.g. Fafchamps and Gubert, 2007]. The larger the number of villages, the higher the probability that farm incomes are not correlated, thus improving the role of *equbs* as an insurance mechanism.

Second, the demand for *equbs* is affected by the existence of other credit institutions such as banks. We find the more distant the bank is, the higher the probability that the PA has *equbs*. Unlike RoSCAs, accessibility to banks depends on physical access (i.e. having a bank branch). This means that as the distance to the bank increases, rural households will have to bear (often substantial) transportation costs to gain access to it [Carpenter and Jensen, 2002].

Third, geographical characteristics determine whether the cluster is more prone to aggregate shocks thus affecting the demand for risk-pooling institutions such as *equbs*. We find that the larger the rain fed land (and hence the lower the probability of an aggregate (i.e. weather) shock), the higher is the demand for *equbs*. In a Townsend-type world, the lower the covariance of incomes, the higher the probability that farmers engage in risk-sharing strategies.

Another "story" could be the fact that if the PA has more irrigated or rain fed land it increases the chances of farming and harvesting, and this may affect the need of farming equipment. Besley et al. (1994) showed that RoSCAs allow individuals to have access to an indivisible durable good by reducing the time of its acquisition. Following the anthropological literature, this result could be explained by the fact that a more developed society where cash is available (i.e. captured by a more developed farming environment) increases the probability that *equbs* exist [Geertz, 1962].

Equation 3.5b is the second-stage regression and reports the amount of credit borrowed from informal lenders given the endogenous availability of *equbs* in cluster j. It

TABLE 3.12: Endogenous switching regression models (second stage)

Log(amount informal credit	Model I=with *equbs* uncorrected std. errors	Model I=with *equbs* corrected std. errors	Model II=without *equbs* uncorrected std. errors	Model II=without *equbs* corrected std. errors
hh characteristics:				
age head	0.03	0.05	0.32	0.04
	(0.02)	*(0.02)***	*(0.16)**	*(0.02)***
age head squared	-0.0003	-0.0005	-0.0003	-0.0004
	(0.00)	*(0.00)***	*(0.00)**	*(0.00)***
hh size	0.02	0.02	-0.20	0.03
	(0.04)	*(0.04)*	*(0.46)*	*(0.04)*
hh size squared	-0.002	-0.002	0.002	-0.002
	(0.00)	*(0.00)*	*(0.02)*	*(0.00)*
female head	0.05	0.01	-0.23	0.06
	(0.11)	*(0.13)*	*(0.85)*	*(0.13)*
number children	0.01	0.01	0.12	0.01
	(0.03)	*(0.03)*	*(0.34)*	*(0.03)*
head schooling	0.03	-0.02	-0.37	0.13
	(0.10)	*(0.10)*	*(1.51)*	*(0.10)*
credit sources:				
bank (lagged)	-0.30	-0.30	-	-
	(0.53)	*(0.53)*		
NGO (lagged)	1.22	-0.84	-1.89	-0.71
	*(0.74)**	*(0.60)*	*(0.71)****	*(0.67)*
PCs of hh assets:				
assets & exp. (pc1)	0.16	0.16	0.40	0.17
	*(0.02)****	*(0.03)****	*(0.50)*	*(0.03)****
assets & exp. (pc2)	-0.07	-0.08	-0.55	-0.06
	*(0.04)**	*(0.05)**	*(0.59)*	*(0.04)*
assets & exp. (pc3)	0.06	0.06	0.40	0.06
	(0.04)	*(0.05)*	*(0.44)*	*(0.05)*
shocks:				
land slide	0.71	1.07	1.90	0.97
	*(0.38)**	*(0.38)****	*(1.41)*	*(0.33)****
harvest diseases	-0.29	-0.27	0.27	-0.31
	*(0.09)****	*(0.09)****	*(1.38)*	*(0.09)****
land taken by	0.96	1.05	-	1.05
cooperative	*(1.05)*	*(0.52)***		*(0.51)***
head imprisoned	0.91	0.84	-	0.79
	(1.05)	*(0.41)***		*(0.41)**
assets resettlement	-1.62	-1.55	-	-1.41
	(1.06)	*(0.76)***		*(0.69)***
banditry	-1.64	-1.59	-	-1.45
	(1.06)	*(0.78)***		*(0.70)***
south	-0.24	-0.20	-	-
	*(0.13)**	*(0.14)*		
Haresaw	-	-	-0.96	-0.85
			(0.96)	*(0.58)*
Geblen	-	-	0.46	0.72
			(2.19)	*(0.66)*
round 2	-0.12	-0.05	1.42	-0.01
	(0.10)	*(0.13)*	*(0.78)**	*(0.13)*
round 3	0.09	0.16	2.06	0.18
	(0.11)	*(0.11)*	*(1.06)**	*(0.11)**
round 4	-	0.00	-	0.001
		(0.12)		*(0.12)*
λ *(Mills)*	-0.28	-0.62	0.01	0.01
	*(0.08)****	*(0.15)****	*(0.98)*	*(0.01)*
Constant	4.52	3.94	-2.64	3.65

	*(0.46)****	*(0.48)****	*(4.44)*	*(0.48)****
N. Obs	1,612	758	4,149	758

Source: own calculation from ERHS. Note: std. errors in parenthesis corrected by bootstrapping (1,000 replications). $^{***}p < 0.01$,$^{**}p < 0.05$,$^{*}p < 0.1$.

has been estimated for each $k = 1, 2$, that is, for each subset ($p < 15$) of clusters with and without *equbs*. Table 3.12 displays the second stage regression of the amount of credit (in log) borrowed from informal lenders for the two groups of clusters with and without *equbs* (models I and II, respectively). The uncorrected model has approximately the same results as the corrected model.

We highlight four results entailing households' demographics, substitutability with formal credit sources, collateral components and income shocks.

With regard to the variables *age* and *age squared*, we find that they are significant when the standard errors are corrected for by the inclusion of a predicted term (i.e. the Mills ratio). The coefficients can be interpreted in two ways. The "experience effect" indicates that the household's head has more capability in obtaining information or simply an enlarged social capital. The "income effect", as described by Attanasio et al. (2000), may arise from the fact that young households are more likely to be credit constrained because income in the early periods of their lives is generally low. On the other hand, the negative sign of *age squared* indicates that as the household head becomes older both the income and the probability of repayment are not likely to increase, reducing the amount of credit obtained from informal lenders.

Another "story" entails the demand side and shows that the household head may actually need less credit as he gets older. The ambiguity in the interpretation depends on the fact that the dependent variable (amount borrowed) does not allow disentangling demand issues from supply issues.

Second, we find no evidence of crowding out as shown by the insignificant lagged formal credit dummies in the corrected models. However, the NGO coefficient is very significant in model II with the uncorrected standard errors. This result should be biased because the standard errors are not corrected for by the inclusion of a predicted term from the first stage regression.

As for the collateral components, table 3.12 shows that an overall increase in assets and expenditure - represented by the first principal component - is positively associated with the amount of credit obtained from informal sources in clusters with and without *equbs*. The second component indicates that the more farm assets (i.e. land) the household has, the lower the amount of credit borrowed from informal lenders in clusters where there are *equbs* (the coefficient is only significant at the ten percent level).

Finally, we find that all shocks are significant after the standard error correction. Shocks that affect the entire community (i.e. harvest disease, assets resettlement and banditry) have a negative impact on the amount of informal debt in clusters with and without *equbs*. The contrary is true for idiosyncratic shocks (i.e. land slide, land taken by cooperative and head imprisoned).

3.5 Conclusion

This chapter has analysed the determinants of households' participation in informal arrangements by using a panel data of 15 peasant associations in rural Ethiopia (ERHS, 1994-1997).

According to the market failure view [Bardhan and Udry, 1999; Besley, 1994; Gosh et al., 1999; Ray, 1997], informal credit arrangements have an advantage in developing economies such as in sub-Saharan Africa because informational sharing mechanisms tend to be small scale and localised, markets are tightly interlinked and highly risky, low levels of wealth limit the provision of collateral, there are few scale economies, inefficient legal systems and low endowments of social capital.

In this chapter we have identified three groups of factors that affect households' participation in informal credit arrangements. The first group - household-based determinants such as wealth and demographic characteristics - has been well discussed within the large literature on this topic [for example, Bose, 1998; Kochar, 1997; Pal, 2002; Ravi, 2003; Ray, 1997]. However, a limitation of these studies is that a high degree of collinearity between household-specific variables (such as components of wealth, income and other household characteristics) limits the significance of individual regressors.

The second group - cluster-based determinants such as demographic, infrastructural and geographical characteristics - is often ignored by the literature due to limited data and lack of appropriate empirical models able to identify such characteristics. Knowledge of these cluster-level differences is as important as knowing why households utilise such institutions in clusters where they are available.

The third group - idiosyncratic and aggregate shocks - has been analysed by the literature as a motive for participation in credit markets [e.g. Bardhan and Udry, 1999; Binswanger and Rosenzweig, 1993; Platteau and Abraham, 1987; Ruthenberg, 1971; Townsend, 1994]. However, data availability limits the identification of cluster level and household level shocks which may affect access to credit.

In this chapter we have been able to address the above-mentioned limitations of the literature by "importing" the endogenous switching regression model from the labour economics literature. We have led to this empirical specification by two "inferior" models: the logit and the Heckman selection model. This approach allows us to highlight the advantages of the endogenous switching regression model compared to the reduced form logit specification and the Heckman model whenever selection bias is not severe.

We have adopted two logit specifications. In the first one we have used principal components analysis, primarily on household wealth-holdings and expenditure, to show how it is particular associations between components of wealth and expenditure that have a highly significant impact on the use of informal arrangements, when compared with standard regression models which specify the determinants of household use of informal institutions as linear combinations of underlying assets.

In the second specification, with access to the village studies provided by the ERHS, we have been able to identify dimensions of heterogeneity of access -most notably geographic, social and economic characteristics- which may operate at a cluster level, but which are not identified at a household level (other than through a crude proxy such as ethnicity). The underlying assumption of this model is that the availability of informal credit sources of a particular type (i.e. *equbs*) is exogenous to cluster level and household level characteristics. This specification points out that there are significant differences between southern (where there are *equbs*) and northern Ethiopia (where *equbs* are not available). These differences affect the access to and the substitutability between credit sources.

After showing with a Heckman model that sample selection bias does not seem to affect our analysis, we have modelled the participation in informal credit through a switching regression with endogenous criterion [Lee, 1978; Maddala, 1983]. The endogenous switching regression models for mixed continuous and discrete variables consist of joint estimation of the probability that in cluster j *equbs* are available (the switching group) and the amount of informal credit borrowed. This specification allows us to model the demand for a particular type of informal credit (i.e. *equbs*) as endogenously determined by household-based and cluster-based determinants. Then, access to informal credit is allowed to differ across endogenously different clusters.

Compared to the Heckman model, the endogenous switching regressions allow us to explain the determinants of the formation of *equbs*. There is no substantial difference between the Heckman model and the endogenous switching regression model in the factors affecting informal debt holding (with the exception of risk factors which are more significant in the endogenous switching model).

We found that access to informal credit is significantly determined by both cluster-based and household-based characteristics. Income diversification (proxied by the number of villages), availability of formal institutions (proxied by the distance to the bank) and incidence of aggregate shocks (proxied by the size of rain fed land) are all factors that positively and significantly determine the demand for informal arrangements such as *equbs*.

Conditional on the endogenously determined socio-economic characteristics, we have then modelled the amount of informal debt held by households. The results have shown that idiosyncratic shocks significantly increase informal debt holding. This confirms the literature that claims that informal credit arrangements are mostly effective in settings where incomes are not highly correlated [Binswanger and Rosenzweig, 1986; Ruthenberg, 1971; Townsend, 1994; Udry, 1999].

Note that the concepts of shock and risk are quite distinct. Shocks can affect behaviour even if they were unanticipated, that is, even if people never expected the shock to happen, and took no precaution against it. People respond to shocks minimising their adverse effects or maximising their beneficial effects. But this does not imply that their behaviour is affected by risk. In this chapter we have focused on the effect of shocks as events that happened in the past and that were beyond the control of individuals.

Wealth components have a significantly positive effect on the access to informal credit.

Finally, following McKernan et al. (2005), we have also tested the crowding out hypothesis by including a lagged vector of formal credit dummies. There is no evidence of crowding out whatsoever. We have provided two explanations. Firstly, formal and informal loans may be independent of each other because they are purpose-oriented [Aryeetey and Udry, 1995; Barslund and Tarp, 2006; Mohieldin and Wright, 2000] as evidenced by the descriptive analysis. Secondly, the result may indicate that there is no long-run effect of formal credit on access to informal loans, but there might be short-run effects that are not captured by the lagged variables.

The main limitation of this chapter is the fact that we do not use panel data methods despite having four rounds. In an attempt to generate an improvement in efficiency, we increase the sample size by pooling the data. However, this formulation does not distinguish in any way between two different households and the same household at two points in time.

Chapter 4

Does the introduction of microfinance crowd out informal loans in Malawi?

"To say that without collateral, banking cannot be done is more stupid than saying that human beings cannot fly because they have no wings. Human beings are creative [...]".
Prof. M. Yunus

4.1 Introduction

Do governments displace the informal loan market by introducing formal credit institutions? The World Development Report states that "*informal and formal strategies are not independent: public policies and the availability of formal mechanisms heavily influence how extensively informal arrangements are used and which kinds are used*" [World Bank, 2001, p.140].

As discussed in the second chapter, the sociological and economic approaches explain the existence and diffusion of informal credit in developing countries. The sociological view is that people engage in social networks of personal relations and kinship. Norms of reciprocity, intergenerational altruism and obligation involve households without having been consciously devised.

The economic view maintains that scarcity of collateral, poor legal enforcement, covariant risk environments and informational problems characterise developing economies such as sub-Saharan Africa. The local information that is required in these economies precludes efficient market coverage from large credit institutions. Banks have funds to lend, but lack adequate information and enforcement mechanisms to recover the loans.

One of the policies that arises as a response to these market failures aims at creating microfinance institutions that will acquire information in innovative ways[1]. By mimicking and exploiting some of the features of informal lending, banks can design credit contracts that harness local information and give borrowers incentives to use their own information on their peers to the advantage of the bank [Armendariz and Morduch, 2005; Ray, 1997]. For instance, in group-lending programmes borrowers who cannot offer any collateral are asked to form small groups. Group members are held jointly liable for the debts of each other. Formally speaking, what joint liability does is to make any single borrowers' terms of repayment conditional on the repayment performance of other borrowers in a pre-specified and self-selected group of borrowers.

In this chapter, we evaluate the effectiveness of this policy by testing whether microfinance institutions actually crowd out access to informal loans in Malawi[2]. We use the Malawi Rural Financial Markets and Household Food Security Survey (FMHFS, 1995) conducted by IFPRI in cooperation with the Rural Development Department of Bunda

[1]See sub-section 2.8.1 in the second chapter for an overview of two public interventions in credit markets.

[2]In the second chapter we have discussed the following four motivations for public interventions aimed at displacing informal loans: enhancing efficiency, distributional motives, mitigation of vulnerability and poverty reduction.

College of Agriculture. The survey contains information about households' borrowing behaviour from both informal lenders and group-lending institutions[3].

Prior to 1995, interventions in rural finance markets in Malawi had no structured policy basis[4]. In 1995 the Malawi Government published the Policy Framework for Poverty Alleviation programme (PAP). Some of the strategies proposed in the PAP centred on the provision of credit facilities and the promotion of micro and small enterprises. The new credit facilities were created by the government of Malawi and some of them received funds from the World Bank. Loans were delivered to small groups for farming activities, such as the acquisition of agricultural inputs (i.e. fertilizers, seeds and farm equipment), or for small-scale trading activities such as sale of products.

The relatively large literature on crowding out in the last fifteen years has found no consensus on the effect of government sponsored programmes on pre-existent private schemes [see table C4.1 for a summary of the available literature]. Most of this literature tests the crowding out hypothesis by means of simple regressions where the dependent variables are private transfers or remittances and the independent variables include, among other controls, some form of public transfers (e.g. public pensions). Typically either probit or tobit models are used, although recently there have been attempts in using non-parametric specifications [Jensen, 2003]. The problem of these studies is the endogeneity bias that arises from non-random selection of participants in the public programme. Some other studies have resolved this issue by means of instrumental variables, randomised or pre and post programme participation data [e.g. Albarran and Attanasio, 2002; Attanasio and Rios-Rull, 2000; Cox et al., 2004; Jensen, 2004; Kaboski and Townsend, 2006; McKernan et al., 2005; Rosenzweig and Wolpin, 1994].

[3]The programmes are: the Malawi Rural Finance Company (MRFC), the Malawi Muzdi Fund (MMF), the Promotion of Micro-Enterprises for Rural Women (PMERW) and the Malawi Union of Savings and Credit Cooperatives (MUSCCO). See sub-section 4.2.2 or appendix A for a full description.

[4]The Government of Malawi/UNICEF (1993) Report and many other studies have provided a basis for the development of a policy framework for poverty alleviation [Bokosi and Khalil-Edriss, 2003].

This chapter tests the hypothesis that the group-lending institutions created in Malawi in 1995 reduce the use of informal credit. Like some of the above mentioned studies [for example, Albarran and Attanasio, 2002; Attanasio and Rios-Rull, 2000; Kaboski and Townsend, 2006] we adopt policy evaluation techniques in order to identify a *causal* relationship between access to formal credit programmes and reduction in use of informal loans. Indeed, the evaluation of the impact of group-lending institutions on the access to informal loans requires the use of an untreated group similar to the group of treated households who participate in group-lending. We choose past members of group-lending institutions as the untreated group. The empirical analysis outlines the motives underlying the choice of this untreated group. Then, propensity score matching is implemented to match participants in group-lending institutions with households that have similar observed characteristics (the so-called "control group"), but are not members of any group-lending institutions.

The chapter introduces several innovations to the literature on crowding out. First, few empirical studies have tested the crowding out hypothesis in the context of group-lending institutions [for example, McKernan et al., 2005]. Although Morduch (2000) has recognized the importance of analysing the role of group-lending institutions in markets where there are a variety of other lenders, most of the economic literature on group-lending institutions has been concerned with the impact of these institutions on clients [e.g. Morduch, 1998; Pitt and Khandker, 1998; Wydick, 1999] and with the ability of joint-liability schemes to overcome information problems affecting formal lenders [Besley and Coate, 1995; Ghatak, 1999; Stiglitz, 1990].

Second, following the evaluation literature on training programmes [for example, Brodaty et al., 2001; Frölich et al., 2004], we develop a model with multiple treatments where households are classified as members of one, or more than one, group-lending programme. This approach allows a comparison between the effectiveness of different credit programmes as well as between different groups of households. Does crowding out differ with the economic status of the household? In particular, are relatively constrained (unconstrained) households more (less) likely to reduce borrowing from informal lenders [Cox et al., 1998; Cox and Jimenez, 2005; Navajas et al. 2003]? For instance, according to Navajas et al. (2003) less wealthy households switch to group-lending institutions to reduce borrowing costs.

Third, we evaluate the effects of both being a borrower and a member of group-lending programmes. This allows us to test the crowding out hypothesis even in presence of *expected* transfers. Nearly all the literature has focused on crowding out in the context of *realised* transfers. Yet households' demand for informal loans is also affected by their membership of a microfinance programme and not just by actual borrowing [Cox and Fafchamps, 2008].

Fourth, most of the literature is only concerned with crowding out of the supply of informal loans. This chapter disentangles demand and supply by employing outcome variables such as demand and credit limit of informal loans[5]. Such detailed data is uncommon in many developed and developing countries.

Finally, we develop a rigorous sensitivity analysis by adopting a number of matching algorithms and by testing for hidden biases arising from unobservable factors that affect simultaneously the assignment into one of the programmes and the outcome variable.

To sum up, this chapter aims at testing whether the introduction of microfinance institutions crowd out demand and supply of informal credit. We evaluate both the effect of membership and borrowing from microfinance instutions by using propensity score matching.

[5]The credit limit variable is extensively explained in chapter five. As it refers to the maximum amount the borrower thinks the lender is able to lend, it can be thought of being the "supply" of informal loans.

The structure of the chapter is as follows. In the next section we describe the data set. Section three contains the descriptive statistics. The evaluation strategy is explained in section four. The selectivity issue is addressed in section five. Section six concludes.

4.2 Data description and management

4.2.1 The Malawi Rural FMHFS survey

The Malawi Rural Financial Markets and Household Food Security survey (FMHFS)[6] was conducted by IFPRI in cooperation with the Rural Development Department of Bunda College of Agriculture as a part of a study on the determinants of access to and participation in existing formal and informal credit and saving programmes and their effects on agricultural productivity, income generation and food security. The Malawi FMHFS was collected in 1995 involving three rounds: the first round took place between February and April, the second one in July-August and the last round in November-December. The survey includes detailed information on land tenure and agricultural production, assets, food and non-food consumption, credit and savings, wage and self-employment income.

The sample includes 404 households in 44 villages in five districts of Malawi[7]. The data was collected using a stratified sampling procedure according to programme membership and then a random selection within each stratum. Half of the stratum-selected sample participated in four credit groups: the Malawi Rural Finance Company (MRFC), the Malawi Muzdi Fund (MMF), the Promotion of Micro-Enterprises for Rural Women

[6]Funding for this research has come from the Rockefeller Foundation, GTZ/Malawi through the Ministry of Women, Children Affairs, Community Services, Social Welfare (MOWCACDSW), UNICEF/Malawi and USAID/Malawi.

[7]The five districts are: Dowa, Nkhotakota, Rumphi, Mangochi and Dedza. A map is displayed in appendix A.

(PMERW) and the Malawi Union of Savings and Credit Cooperatives (MUSCCO). The other half of the sample either has previously participated in one of the credit programmes or has never participated in a formal credit programme.

The credit and savings module was administered to every member of the household who was over 17 years of age. Information about credit characteristics was collected for the following loans: rejected loans, not demanded and granted loans for any credit programme, formal or informal loan source.

4.2.2 The microfinance credit programmes

In 1995 the government of Malawi published the Policy Framework for Poverty Alleviation (PAP) which was centred on the provision of credit facilities and the promotion of micro and small enterprises. The credit facilities were supported by the government of Malawi on a policy basis and some of them received funds from the World Bank. This thesis focuses on four microfinance programmes[8].

The Malawi Rural Finance Company (MRFC) is funded by the World Bank. It provides in-kind seasonal agricultural loans for fertilizers, seeds and pesticides for hybrid maize and tobacco. It also offers short-term (two years) and medium-term (five years) loans for farm equipment. The targeted people are jointly liable groups of 5-10 smallholder farmers. Moreover, the MRFC offers two savings deposit services to its borrowers: ordinary and contract savings accounts. With contract savings account, clients can choose the amount and timing of deposits. For honouring commitments, they can either get a bonus interest or earn a collateral-free loan limit.

The Malawi Mudzi Fund (MMF) most closely resembles micro-credit facilities as it was designed as a replica of the Grameen Bank in Bangladesh. It is funded by the World

[8]A more extensive description of the credit programmes is provided in appendix A.

Bank and by the International Fund for Agricultural Development (IFAD). The targets of the MMF are poor households with less than one hectare of land. It provides loans for non-farm income-generating activities. First-time Mudzi borrowers are not required to provide collateral but are required to form groups, which are based on the Grameen model of groups of five and "centres" of 20-25, and undergo a six-month training period, to qualify for loans.

The Malawi Union of Savings and Credit Cooperatives (MUSCCO) is the apex organization for Savings and Credit Co-operatives (SACCOs) which offer a range of services including credit, savings and insurance. It was created in 1980 and financially supported by the United States Agency for International Development. The MUSCCO is the principal Malawian financial institution actively promoting savings mobilization. It has not experienced the default rates that have characterized other lending operations, primarily because it is member-based and funds loaned represent members own savings.

The Promotion of Microenterprises for Rural Women (PMERW) is a credit programme financially supported by the German Agency for Technical Cooperation (GTZ). It was started in 1986 by the Ministry of Women and Children's Affairs and Community Services. The most recent version of the credit programme targets women groups of 5-10 who are skilled in business activities. The structure is similar to the saving and credit clubs except that members can borrow up to MK 1,000 and they can receive loans directly from the Central Bank of Malawi. Credit members are selected among those who have excellent credit and business management skills.

4.2.3 Missing Data

As described in the second chapter, missing values have been replaced by using *hot-deck* imputation. Because the Malawi FMHFS is a short "panel", recall questions are less affected by missing values and there is a significant consistency of data over time. Imputation has only been used to replace outliers which never exceeded three percent of the data.

4.2.4 Price index and weights

Unlike the Ethiopian Rural Household survey described in the second chapter, the Malawi Financial Markets and Household Food Security survey is a long cross-section that covers three different seasons within the same year. In this context, the Fisher index is inadequate to measure seasonal price variation because it ignores the effect of seasonal variation in consumption [Rothwell, 1958]. In other words, the Fisher index cannot measure correctly price changes between months (when items and weights of the market basket are different). Conceptually, the problem is identical to that of measuring differences in price levels between two countries having different market baskets. Hence, in order to maintain a significant seasonal price variation that can potentially affect households' demand for credit, we have not deflated values with the Fisher index (or indeed with any other index).

Despite the fact that there are numerous credit programmes in Malawi, credit programme participation is still not high. Indeed, according to Diagne (1999) only 12 percent of the 4,699 households enumerated in the 44 villages covered in the village census were current members of a credit programme. This figure cannot be representative of credit membership in Malawi because it includes villages that were specifically hosting credit programmes. The fact that many villages in Malawi do not host credit programmes and the low participation in hosting villages rules out the use of random sampling. Since the purpose of the study was to evaluate these credit programmes, the only feasible alternative to include enough credit programme participants was to stratify along the programme membership status variable.

The data was collected adopting a choice-based sampling procedure where households were selected according to their participation in credit programmes. The information about the villages in which the four credit programmes were operating was obtained from the national headquarters and district offices of these credit institutions. This was then followed by a village survey before a sample frame of villages hosting present members from each credit programme was selected for the study. The selection of the area was done according to the location of the credit programmes. The next step consisted of carrying out a village census where all households were listed and information on whether a household was a present or past member of a credit programme was collected [Bokosi and Khalil-Edriss, 2003]. Thus, the survey was stratified along programme membership with random selection within each stratum [Diagne, 1999]. About half of the sample households participated in the four credit programmes and the other half were equally divided between past members and non-participants.

Manski and Lerman (1977) showed that choice-based sampling produces inconsistent estimates. In order to correct for this inconsistency, we use choice based sampling weights[9]. The weights are defined as follows:

$$\omega = \frac{H(j_i)}{Q(j_i)} \tag{4.1}$$

where $Q\left(j_i|\beta_0\right) = \dfrac{N_j}{N}$ is known and represents the decision-making population selecting the jth alternative (i.e. programme membership). $H(j_i) = \dfrac{n_j}{n}$ is the choice-based sampling ratio; N_j is the size of the population defined by programme j and n_j is the size of the sample stratum; n and N are the total sample and population sizes,

[9]Appendix B describes how these weights produce consistent estimators.

respectively. This is the probability weight used in the survey where the population frequencies have been obtained by the village census conducted prior to the survey.

4.3 Descriptive statistics

This section provides some descriptive statistics of the community and of the household characteristics. The statistics have been weighted to correct for endogenous sampling. The subsection entailing community statistics describes the general characteristics of the survey sites. The household subsection is divided in two parts. The first part considers the composition of the household. The second part describes households' borrowing behaviour.

4.3.1 Community level

The community survey was undertaken in 1995 and includes information about the demographic characteristics, infrastructures and agricultural production of the villages. For brevity purposes, the statistics displayed in this section cover the five surveyed districts[10]. As shown in table 4.1, one district is located in the North, one in the Centre, and three districts are located in the South of Malawi[11]. While Dowa and Mangochi include three and five villages respectively; Nkhotakota, Rumphi and Dedza include seven, nine and twenty villages respectively. Over 1,000 households live in Mangochi, Nkhotakota and Dedza. Taking into account both the number of villages and households, Mangochi is the most populated district. Indeed, southern Malawi is the most populous area.

The major religion is Christian: 88 percent and 89 percent of people are Christians in

[10]The districts are an aggregation of villages.

[11]See map in appendix A for more details.

Dedza and Rumphi respectively. Overall, more than 70 percent of people are Christians in the five surveyed districts. Fewer people are Muslims: 18 percent in Mangochi and 8 percent in Nkhotakota.

The districts differ in agricultural characteristics. No cultivable land is irrigated in Rumphi, while 32 percent and 10 percent of the total cultivable land is irrigated in Mangochi and Nkhotakota respectively.

Table 4.1: Characteristics of the districts

District	Location	N. of villages	N. of HHs	% of Christians	% of Muslims	% of trad. religion	N. of shallow tube wells	Prop. irrigated land in tot. cult. land	Prop. hybrid maize in tot. maize area
Dowa	South	3	700	87	0	10	0	2	50
Mangochi	Centre	5	4,717	81	18	4	0	32	94
Nkhotakota	South	7	1,815	74	8	9	14	10	75
Rumphi	North	9	895	89	0	0	5	0	38
Dedza	South	20	2,508	88	0.1	12	15	6	20

Source: Own calculation based on MRFMHFS, community data.

Mangochi and Nkhotakota have the highest proportion of hybrid maize cultivation in total maize area: 94 percent and 75 percent respectively.

Table 4.2 shows the availability of institutions in each district. Farmers' clubs are particularly widespread in Rumphi and Dedza. There are not many modern private medical practitioners: Mangochi has six and Nkhotakota has only one. The other districts have no private medical practitioners. On the other hand, traditional healers are more diffused: there are 14 in Mangochi and Dedza respectively and four in Nkhotakota. NGOs are present only in Mangochi and Nkhotakota. Shops are widespread in all the districts. Again, Mangochi and Nkhotakota have the highest number of shops: 35 and 76 respectively.

Table 4.3 describes some of the characteristics of agricultural production in each district. In all the districts, December is one of the months of the hungry season. The prices per kilo of maize vary across seasons and districts. For example, in Mangochi the

Table 4.2: Institutions by district

District	N. of farmers'	N. of churches	N. of mosques	N. of modern private medical practitiones	N. of traditional healers	N. of NGOs	N. of shops
Dowa	30	9	0	0	1	4	5
Mangochi	84	23	11	6	14	10	35
Nkhotakota	76	26	2	1	4	6	76
Rumphi	233	5	2	0	1	0	24
Dedza	215	29	1	0	14	3	33

Source: Own calculation based on MRFMHFS, community data.

price per kilo of maize in July is more than twice the price in January. Also, in Mangochi 100 percent of produced food grains are consumed, while 75 percent are consumed in Dowa[12].

Table 4.3: Agricultural production by district

District	Months of hungry season	% of major food grains consumed	price of local maize per kilo (MK)			
			July	April	January	October
Dowa	February, December	75	1.3	1.1	1.0	0.9
Mangochi	November, December	100	2.0	1.4	0.2	0.0
Nkhotakota	January, November, December	86.4	1.9	0.9	0.9	0.8
Rumphi	December	76.1	0.6	0.8	0.6	0.0
Dedza	October, November, December	79.8	1.4	1.2	1.1	0.5

Source: Own calculation based on MRFMHFS, community data.

Table 4.4 displays the availability of tarred or gravel roads in each district. In Mangochi, all roads to the government office, credit office, post office and commercial bank are tarred or gravel. Also, 80 percent of roads to the primary school and health centre are tarred. The distance is relatively small: less than 10 Km to a government office, credit office, post office, primary school and health centre. However, the commercial bank is more distant at 102 Kms. In Nkhotakota, 71 percent of roads to the government office, credit office, post office and health centre are tarred, while 86 percent and 57 percent of roads to the primary school and commercial bank, respectively, are tarred or

[12]Food grains include local and hybrid maize, beans, cassava, rice and nuts.

Table 4.4: Infrastructures by district: tarred or gravel road

District	% to gov. office	distance to gov. office (Km)	% to credit office	distance to credit office (Km)	% to post office	distance to post office (Km)	% to primary school	distance primary school (Km)	% to commercial bank	distance commercial bank (Km)	% to health centre	distance to health centre (Km)
Dowa	0	20	0	6	0	20	33	4	0	20	0	20
Mangochi	100	10	100	0.3	100	8	80	2	100	102	80	8
Nkhotakota	71	4	71	4	71	1	86	1	57	69	71	24
Rumphi	0	6	22	4	11	4	22	1	0	29	0	6
Dedza	70	26	80	9	20	23	30	3	55	44	55	34

Source: Own calculation based on MRFMHFS, community data.

Table 4.5: Credit sources by district

	MRFC		MMF		MUSCCO		PMERW1		PMERW2		PSACA		N. of money lenders in the district	N. of money lenders out of district who lend to HHs in district
	N. of groups	Average n. of years of existence	N. of groups	Average n. of years of existence	N. of groups	Average n. of years of existence	N. of groups	Average n. of years of existence	N. of groups	Average n. of years of existence	N. of groups	Average n. of years of existence		
Dowa	1	10	0	0	2	5	0	n.a.	0	n.a.	1	n.a.	2	1
Mangochi	4	4.2	4	1	0	0	3	n.a.	1	n.a.	0	n.a.	3	0
Nkhotakota	3	2.4	0	0	1	0.7	3	n.a.	4	n.a.	1	n.a.	5	0
Rumphi	4	1.6	1	n.a.	0	0	3	n.a.	3	n.a.	1	n.a.	0	0
Dedza	5	3.1	0	0	0	0	0	n.a.	0	n.a.	4	n.a.	0	1

Source: Own calculation based on MRFMHFS, community data. n.a.=not available.

gravel. In Dowa, there are no tarred roads to the government office, credit office, post office, commercial bank and health centre.

Table 4.5 shows the existence of formal credit groups[13] and informal moneylenders in each district. Not all formal credit programmes are available at the national level. MRFC groups are present in all districts and, overall, they existed for more than two years. MUSCCO groups exist only in Dowa and Nkhotakota. These credit groups are relatively younger than MRFC groups. MMF groups are much less widespread: they only exist in Mangochi and Rumphi and they are of relatively recent formation. Although Dedza has more villages than other districts, it only hosts MRFC groups. In Dowa, Mangochi and Nkhotakota there are a few moneylenders (two, three and five, respectively). In Dedza there are no local moneylenders, but one outside moneylender operates in the district.

To sum up, southern and central Malawi is more populated. It has a higher proportion of irrigated land and consumes more food grains than the northern parts of the country. Mangochi, situated in the Centre, is the district where there are more modern medical practitioners and where institutions such as schools, post offices and government offices are connected by tarred or gravel roads to the villages. The major religion is Christian. In many districts, December is the month of the hungry season.

4.3.2 Household level

Household data covers three rounds in 1995: February-April, July-August and November-December. It includes 404 households, 44 villages in five districts. It is a strongly balanced short "panel".

[13] A detailed description of credit groups is provided in appendix A.

4.3.2.1 Households' characteristics

Table 4.6 reports some of the households' characteristics. By comparing the number of actual households as reported in table 4.1 to the number of surveyed households, it is possible to point out how many households the survey has covered. In Dowa and Mangochi eight percent of actual households have been interviewed; in Nkhotakota and Dedza four percent; in Mangochi only two percent.

TABLE 4.6: Households' characteristics by district

Distict	N. of surveyed households per round	% of female headed households per round	Distance to parents village village (Km)				Education of household head	
			Head		Spouse			
			Mean	*Std. error*	*Mean*	*Std. error*	% with professional training	% with adult literacy certificate
Dowa	56	34.1	39.1	1.2	39.3	1	1	0.3
Mangochi	102	29.9	136.6	33.5	107.2	33	13.5	0.2
Nkhotakota	70	36	289.1	79.1	221.2	70.7	14	7.8
Rumphi	75	21.6	5.1	3.5	15.7	9.5	24.5	0
Dedza	101	36.8	23.9	8.4	76.4	45	3.4	3.1
N. of obs.	*404*	-	*135*		*135*		-	

Source: Own calculation based on MRFMHFS, community data.

Table 4.6 also shows that in Dowa, Nkhotakota and Dedza more than 30 percent of surveyed households are female headed. In Mangochi and Nkhotakota the average distance to the parents' village is lower for the spouse than for the head of the household. While in Dowa only one percent and 0.3 percent of households' heads have, respectively, a professional training or an adult literacy certificate, in Rumphi almost 25 percent of household heads had professional training.

Table 4.7 reports the household composition by district. On average, households have approximately five members in Dowa and Mangochi, four members in Dedza and six members in Rumphi and Nkhotakota. In all districts the average number of children between 0-5, 6-10, 11-17 is one. A large majority of households' heads in Dowa, Rumphi and Dedza are employed in agriculture.

Table 4.7: Households' composition and occupation by district

District	**Household size**		**N. children 0-5**		**N. children 0-6**		**N. children 11-17**		**Occupation of household head**
	Mean	**Std. error**	**Mean**	**Std. error**	**Mean**	**Std. error**	**Mean**	**Std. error**	**% employed in agriculture**
Dowa	4.5	0.35	0.9	0.14	0.7	0.2	0.8	0.2	93.2
Mangochi	4.7	0.55	0.6	0.12	0.9	0.1	1.2	0.4	22.5
Nkhotakota	6.2	0.75	1.5	0.35	0.9	0.1	1.6	0.3	50.4
Rumphi	5.6	0.47	0.8	0.17	0.9	0.2	1.4	0.2	92.1
Dedza	4.2	0.29	0.7	0.12	0.7	0.1	0.8	0.1	93.5
N. of obs.	*1,212*		*1,212*		*1,212*		*1,212*		

Source: Own calculation based on MRFMHFS. Weighted results.

4.3.2.2 Households' borrowing behaviour

The Malawi FMHFS contains information about households' borrowing behaviour. This section identifies the characteristics of both credit suppliers and borrowers.

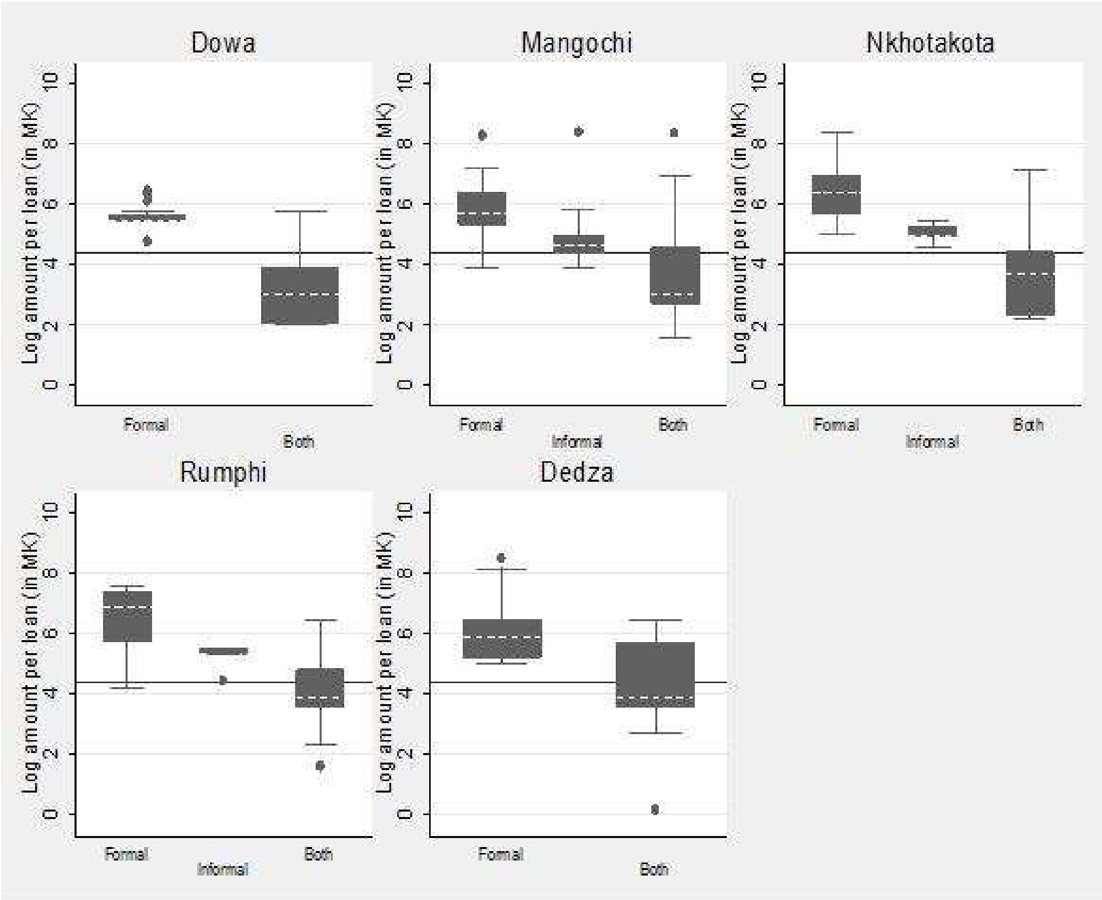
Who are the suppliers of credit?

There are two credit sources in the districts. Formal or institutional lenders include the four credit programmes described in sub-section 4.2.2, the Central Bank of Malawi (CBM) and World Vision (i.e. a NGO). A more detailed description of the credit groups is also provided in Appendix A. Informal sources include: friends and relatives and other informal lenders (i.e. moneylenders and traders).

Figure 4.1 shows the box plots of formal and informal sources by district[14]. Across all districts the median amount of credit per loan (in logarithm) is 4.41 MK. Households can borrow only from the formal sector, only from the informal sector or they can borrow from both informal and formal lenders. The median formal credit per loan in each district is higher than the overall median. In Mangochi, Nkhotakota and Rumphi the informal box plots are quite small, meaning that the distribution of informal loans is less spread than the distribution of other loans.

[14]As already explained in the third chapter, the dotted horizontal line is the median and 50 percent of cases have values within the box plots. The length of the box is the inter-quartile range and the upper boundary (lower boundary) of the box is the 75^{th} (25^{th}) percentile. The black line is the overall median. The circles are extreme values but not outliers.

FIGURE 4.1: Distribution of formal and informal credit by district in rural Malawi

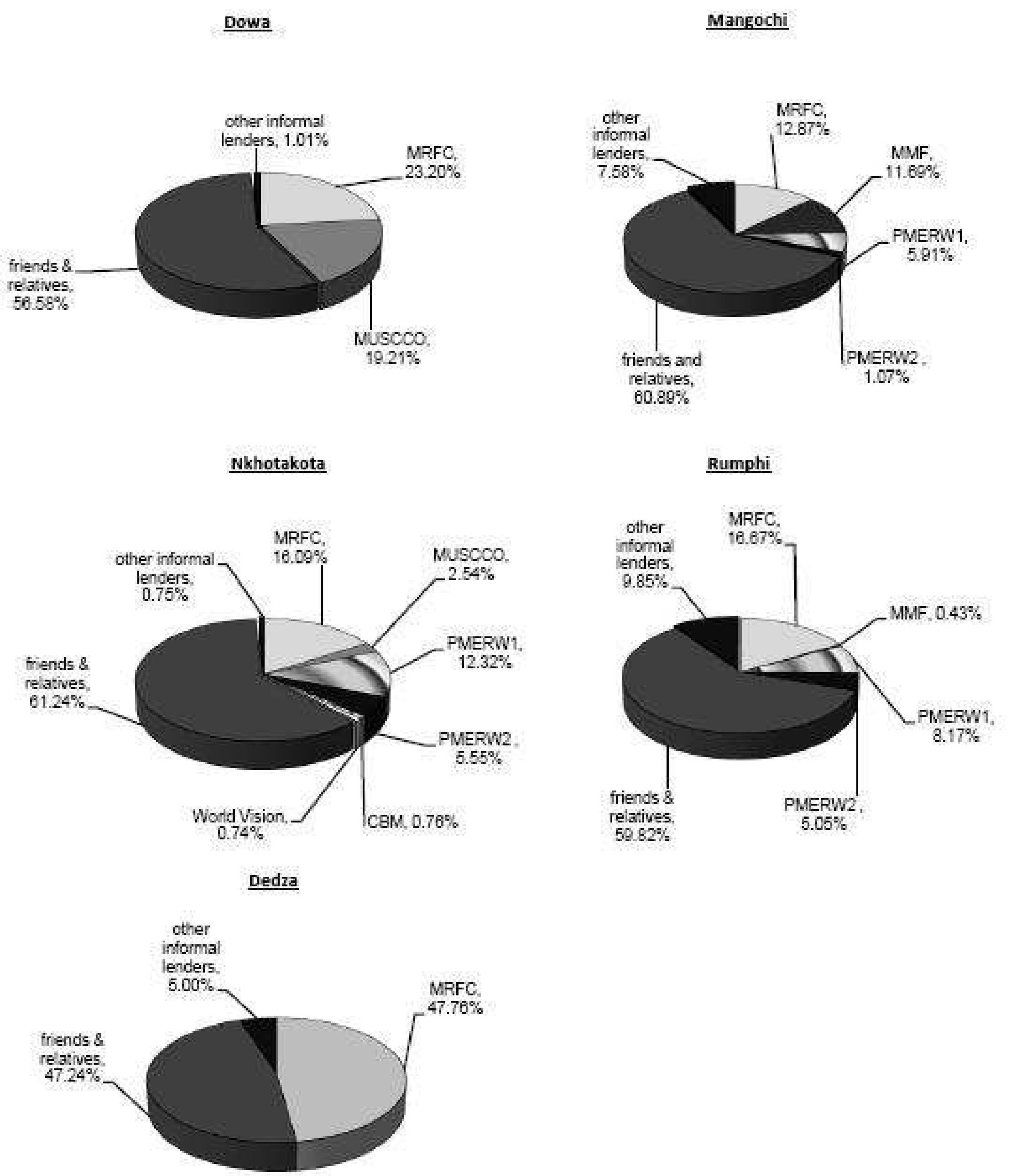


Source: Own calculation based on MRFMHFS. Note: all values in local currency, 15 Malawian Kwachas (MK)=1 US$. Malawi's per capita GNP is US$ 170 (approx. 2,550 MK. World Bank, 1997).

On the other hand, in Dowa the distribution of formal loans is very concentrated (the logarithmic value of most loans is 5.8 MK). The distribution of formal and informal loans, by contrast, is disperse and negatively skewed in Mangochi, Rumphi and Dedza. In all the districts, households borrow more from formal lenders. The median amount of formal and informal loans in all the districts is significantly lower than the median across all credit sources. For instance, in Dowa the median formal and informal loans (in logarithm) is around three MK (lower than the overall median of 4.41 MK).

Figure 4.2 displays the distribution of loans by source and district. Not all loans are utilised in all districts. While friends and relatives, MRFC and other informal lenders are used in all districts, households borrow from the other formal credit programmes in

FIGURE 4.2: Distribution of loan source by district



Source: Own calculation based on MRFMHFS.

Mangochi, Nkhotakota and Rumphi. The MRFC programme is the most diffused among formal credit sources. For instance, in Dedza approximately 48 percent of the loans are provided by the MRFC. Despite the presence of formal credit programmes, friends and relatives are the most widespread source of credit in all districts (supplying around 60 percent of the loans) except for Dedza. As already outlined in table 4.5, Dedza is the only district where only one formal credit programme (i.e. MRFC) is available.

What are the characteristics of those who ask for credit?

The dataset allows the researcher to identify different groups of households: those who borrowed from at least one of the credit programmes; those rejected from formal lenders; non-applicants (i.e. households who never participated in a formal credit programme); and past members (i.e. households who once were members of one of the credit programmes). Table 4.8 displays the characteristics of the four groups: as a result of the sample design the majority of households are borrowers[15], non-applicants and past participants are approximately 311 and 159 respectively across all rounds, but there are only 51 rejected applicants.

Among the four groups, non-applicants have the highest percentage of female headed households (38 percent) while only 11 percent of past participant households have a female head. Households who participate in credit programmes have, on average, the highest number of members (around six).

From the composition of the household it is clear that both borrowers and past members have a higher number of children between 0 and 15 years of age. Almost 87 percent of past participants have a household head employed in agriculture. However, only 49 percent of participant households have a household head who is mainly employed in agriculture. This can be explained by the fact that the household head of participant households may be employed in small trade or in other activities.

Rejected applicants have the highest land size (2.4 hectares)[16]. The share of land in total assets owned by borrowers is lower than rejected households and past participants (approximately 58 percent versus 69 and 64 percent respectively). Also, participants

[15]The number of households within each group refers to respondents. As pointed out in sub-section 4.2.4, in each round half of the surveyed households were interviewed among programme participants and the other half was equally divided between past participants and non-participants.

[16]This is a result of the ceiling on land set by the programmes eligibility criteria.

Table 4.8: Characteristics of households' groups

	Participants	**Rejected from formal lenders**	**Non-applicants**	**Past participants**
Households demographics:				
Female headed HHs (%)	22.0 (541)	20.5 (51)	38.1 (311)	10.6 (159)
Average HH size	5.8 (541)	5.3 (49)	4.5 (311)	5.7 (159)
Average number of children 0-15	3.0 (541)	2.6 (49)	2.2 (311)	3.2 (159)
Household's head main occupation: agriculture (%)	48.8 (540)	71.9 (51)	69.9 (307)	86.7 (159)
Households assets and shock:				
Households affected by negative income/health shocks (%)	59.4 (535)	42.9 (48)	53.9(309)	51.6 (154)
Average land size (ha)	2.2 (517)	2.4 (47)	1.4 (288)	2.1 (153)
Share of land owned by spouse	16.2 (517)	3.2 (47)	16.2 (288)	24.5 (153)
Average value of house (MK)	1055 (517)	700 (47)	463 (288)	846 (153)
Share of assets held as land (%)	56.4 (517)	62.9 (47)	56.8 (288)	63.7 (153)
Average food expenditure (MK)	13.7 (517)	10.4 (47)	12.3 (288)	9.5 (153)

Source: Own calculation based on MRFMHFS. Note: household types are defined according to participation in the credit programmes. Weighted results. Number of observations in parentheses. Expenditure deflated by the square root of households' size.

have a higher value of house compared to the other three groups of households.

These results may indicate that programme participants are relatively better off than the other groups. They are not mainly employed in agriculture and their assets are not primarily composed of land. However, it is only possible to determine a correlation (not causation) between households' level of wealth and their membership. Participants in credit programmes have a considerably higher food expenditure compared to the other groups. Also, they seem to have been affected by negative income or health shocks to a greater extent than the other three groups[17].

[17]Negative shocks include: natural disasters affecting crops, illness or death of a household's member, death of livestock, environmental degradation (i.e. erosion or deforestation), unavailability of inputs and fewer members of working age.

To summarise the households descriptive statistics, we find that households in southern Malawi are larger. They are more educated and less likely to be employed in agriculture than in northern Malawi.

We also find that the distribution of formal and informal loans differs across districts. In particular, some households borrow only from informal sources in Mangochi, Nkhotakota and Rumphi. In these districts, the distribution of formal loans is generally less concentrated than the distribution of other loans.

Despite the existence of formal credit programmes, households mostly borrow from friends and relatives in Malawi. Those who borrow from the credit programmes are relatively better off in terms of value of the house and food expenditure compared to past participants, rejected applicants and non participants.

4.4 The evaluation problem

The standard model in the evaluation literature [Roy, 1951; Rubin, 1974] involves the estimation of the effect of participation in a programme on a hypothetical outcome. Lechner (1999a) and Imbens (2000) allow the model to include multiple treatments. Several applications of this extended model have been used to evaluate the effects of different training programmes [e.g. Brodaty et al., 2001; Dorsett, 2001; Frölich et al., 2004; Larsson, 2000; Lechner, 2000 and 2001].

Suppose we want to evaluate the effect of different credit programmes on the amount borrowed from informal lenders. Consider the outcomes of T mutually exclusive groups to be denoted by Q^m, Q^l. We define two groups: $T = m, l$.

The first group, $T = m$, denotes two treatments: either membership only in the MRFC[18] or membership in the MRFC and in other formal credit programmes (i.e. par-

[18] A full description of credit programmes is contained in Appendix A.

ticipation in at least two credit programmes). The main advantage of our choice of treatments is that it allows us to keep many observations while evaluating the effect of participation in one or more formal credit programmes on the access to informal sources. On the other hand, by pooling together different credit programmes in the second treatment we neglect heterogeneity of credit institutions.

The second group, $T = l$, denotes the case of *no treatment*. Our untreated group is composed of past members of credit programmes. There are two reasons why we did not use households who never participated in any credit programme as the untreated group. First, we argue that past members are the appropriate untreated group because they have the same (time-invariant) unobservable characteristics (i.e. entrepreneurship ability) as the group of participants. Note that because all the credit programmes deliver loans for farming activities, unobservable factors like entrepreneurship may affect selection into the programmes. Second, we are not able to identify the outcome for the group of households who never participated in credit programmes because they do not borrow from informal lenders[19].

Following Lechner's (1999a) approach we denote participation in a particular treatment with the variable $T \in m, l$, with $T = l$ for the no treatment option. The average treatment effects for any pair-wise comparison between treatments m and l are given by:

$$\gamma_0^{m,l} = E\left(Q^m - Q^l | X\right) = EQ^m - EQ^l \tag{4.2a}$$

$$\vartheta_0^{m,l} = E\left(Q^m - Q^l | T = m, X\right) = E\left(Q^m | T = m, X\right) - E\left(Q^l | T = m, X\right) \tag{4.2b}$$

where equation 4.2a denotes the average effect (ATE) of treatment m relative to treatment l for the population; and equation 4.2b is the average treatment effect on house-

[19]The reason for not borrowing from informal lenders may be either voluntary or not.

holds treated by programme m (ATT).

An evaluator's "classic problem" is to identify $E\left(Q^l|T=m\right)$ since the difference between $E\left(Q^m|T=m\right)$ and $E\left(Q^l|T=m\right)$ cannot be observed for the same household. Rubin (1974) solves the identification problem by defining the conditional independence assumption (CIA). The CIA states that, given a set of observable covariates X in a particular attributes space χ, potential treatment outcomes are independent of the participation status. Lechner (1999a) formalized the CIA for the multiple treatment case:

$$Q^m, Q^l \qquad \coprod T|X=x, T \in m, l; \qquad \forall\ x \in \chi \tag{4.3}$$

In order to overcome the "curse of dimensionality"[20], Rosenbaum and Rubin (1983) suggest using balancing scores $b(X)$. In the multiple treatment case, the propensity score[21], $P^{m|ml}(X)$, is a type of balancing score defined as the probability of participation in programme m for household i conditional on the participation in m and l given a set of observed covariates X such that the conditional distribution of X given the propensity score is independent of the assignment into the treatment[22].

The main drawback of this procedure is that the CIA only holds after controlling for observable characteristics. Indeed, Heckman et al. (1997) show that even after conditioning on a set of observables, outcomes of participants and non-participants may still be significantly different for a variety of reasons. For example, selection in the programmes may be conditioned on a series of unobserved characteristics and differences in outcomes may arise when participants and non-participants live in different regions. In order to check the robustness of our untreated group of past members, we deal with the (possible) selection on unobservables at the end of this chapter.

[20]Conditioning on all relevant covariates is problematic when there is a high dimensional vector X.
[21]The conditional probability $P^{m|ml}(X) = \frac{P^m(X)}{P^m(X)+P^l(X)}$.
[22]see appendix B for a more detailed description of the topic

The approach taken in this evaluation consists of four stages. First, we estimate the propensity scores of participation. These can be obtained either by using multinomial models or by a series of binary choice models. In the second stage we perform the Mahalanobis metric matching algorithm with propensity score. The third stage is to estimate the average effects of participating in one or more credit programmes relative to past membership. The outcome of interest is the amount households borrow from informal sources. Hence, for those in option m, the mean effect of option m rather than option l is estimated as the mean difference in the amount borrowed from informal lenders between households in option m and the matched households in option l. The final stage ensures that the results do not depend on the methodological assumptions of our evaluation procedure. Indeed, the sensitivity analysis adopts several specifications: a) it changes the regressors of the model and the matching algorithm; b) it changes the definition of treatment and outcome; and c) it changes the model used to estimate the propensity scores.

4.4.1 First stage: estimation of the propensity scores

In this stage we estimate the propensity scores of participation. Consider T mutually exclusive treatments denoted by $T = m = 1, 2$ where $T = 1$ is membership only in the MRFC[23]; and $T = 2$ is membership in the MRFC and in other formal credit programmes (i.e. participation in at least two credit programmes). $T = l$ denotes the case of *no treatment*. Our untreated group is composed of past members of credit programmes[24]. The propensity scores are the predicted values $\hat{T}^k_{ij(i)} = \hat{P}\left(T = k | T = m, l\right)$ where i indicates the ith household, *j(i)* indicates the village where household i lives and $k = m, l$. In this context, m is either participation in the MRFC only or participation in more than one

[23] A full description of credit programmes is contained in Appendix A.

[24] See above for the reasons why we chose past members as untreated group.

programme and l is past membership. More formally, the propensity scores are given by:

$$\hat{P}^{m|ml}(x) = \frac{\hat{P}^{m}(x)}{\hat{P}^{m}(x) + \hat{P}^{l}(x)}$$

In the case of multiple treatments, the first decision to make is whether the conditional participation probabilities should be estimated for each combination of treatments as binary choices or whether they should be modelled with a multinomial model including all relevant choices. Both approaches have advantages and disadvantages and although we decided to model the decision process with a series of logit models, we also checked the robustness of our results in a multinomial context.

The use of the multinomial logit model can be ruled out because of the violation of the *Independence of Irrelevant Alternatives* (IIA). According to the IIA, the inclusion or exclusion of some programmes does not alter the relative probability of a choice programme to another. However, as outlined by Larsson (2000), the IIA may not hold in a multiple programmes context because the relative probabilities of one choice to another may change whenever programmes are at least partly substitutes to each other[25]. Hence, if we were to use a multiple choice model, the multinomial probit model (MNP) would be the best one because it would overcome the IIA assumption. However, since the MNP could not be fitted by our data we follow Lechner (2001) in using a series of logit models.

Bryson et al. (2002) highlighted two shortcomings in using a series of binary choices. Firstly, as the number of treatments increases the number of models increases as well [i.e. for T choices we need $\frac{1}{2}(T(T-1))$ models[26]]. Secondly, the choice in each model is conditional on being in one of the selected treatment groups. On the other hand,

[25] See later for an example of change in relative probabilities.

[26] T is the number of different options including the "*no treatment*" option.

Lechner (2001) found little difference in the performance of the MNP model and the series of binomial models. In particular, the matching quality (measured by the standardised bias[27]) achieved with the MNP is not much different from the series of binary choices. The latter approach is more flexible because it allows modelling each of the binary choices with a different set of covariates. Also, the binary models are more robust to errors since a mis-specification in the model of any pair of treatments will not compromise the other choices [Dorsett, 2001].

After a decision about the model specification has been made, we need to choose the variables to be included in the model. From a theoretical point of view, only those variables that affect both the participation decision and the outcome should be included. We suspect that, in anticipation of participation, poor households decrease their effort to increase income (i.e. job search or effort to increase production). Ashenfelter (1978) discovered a similar result evaluating the treatment effects on earnings (the so-called Ahsenfelter's Dip[28]). Later research found a rise in unemployment shortly before participation in a labour market programme as a result of anticipation effects [see for example Fitzenberger and Prey, 2000; Heckman et al., 1999; Heckman and Smith, 1999].

In order to avoid a reversed causality between the covariate X and the participation decision, variables should be either fixed over time (i.e. gender) or should be measured before participation. Because the data does not contain information about the starting date of membership, the following relatively static variables were included: household and community characteristics, and semi-fixed factors that affect eligibility such as land size. As described in sub-section 4.2.2, most of the credit programmes set some eligibility criteria: credit is delivered to small farm holders and poor households. While we can

[27] See below for a definition of this statistic.

[28] This effect can be ignored if the introduction of a new programme is unanticipated. Ideally, this hypothesis could be tested by looking at households' income before and after the creation of microfinance institutions. However, because the data entails only one year (1995) and some of the programmes were created before 1995, we cannot test this hypothesis.

include land size as a covariate in the estimation of the propensity score because it does not change much over time, we cannot use agricultural income because it displays variability across seasons and it both affects and is affected by participation and outcome.

The final issue arising in this evaluation problem is that samples are choice-based [Smith and Todd, 2005]. As mentioned above, choice-based sampling leads to an over-sampling of participants relative to the eligible households in the population. In appendix B we describe how the likelihood function should be changed. Also, in the previous section we pointed out that sampling weights are required to consistently estimate the probability of participation in the credit programmes. So, the propensity scores will be defined as:

$$\omega \hat{P}^{m|ml}(x) = \omega \frac{\hat{P}^m(x)}{\hat{P}^m(x) + \hat{P}^l(x)}$$

where:

$$\omega = \frac{H(j_i)}{Q\left(j_i|\beta_0\right)}$$

as described in sub-section 4.2.4. However, Heckman and Todd (2005) showed that matching methods[29] can be still applied even with the propensity scores without weights since the ranking of the observations is just shifted by a scalar (where the scalar is defined by the weight, ω) and the same observations will be matched. We check the robustness of our results by dropping the weights in the sensitivity analysis. Frölich et al. (2004) found that dropping the sampling weights does not change the results of their evaluation of a Swedish rehabilitation policy. Diaz and Handa (2006) do not adjust for choice-based sampling in the estimation of the propensity scores of participation in the PROGRESA programme in Mexico.

We model the choice of participation with a series of logit models where the treatments are: membership in one programme only (i.e. MRFC) and membership in more than

[29]They show it for nearest neighbour matching.

TABLE 4.9: Selected characteristics by treatment groups prior to matching

	(I) MRFC vs. Past members		(II) 2nd programme vs. Past members		Group comparisons (I)		Group comparisons (II)	
	treated	*untreated*	*treated*	*untreated*	**t-stat.**	% $\|bias\|$	**t-stat.**	% $\|bias\|$
households characteristics:								
household size	5.73	5.80	6.44	5.69	-0.30	3.40	2.64***	31.90
age head	49.16	44.34	46.33	43.33	3.16***	35.80	2.25**	23.70
female head†	-	-	0.31	0.19	-	-	2.30**	27.70
n. of children 6-10	0.92	1.13	1.16	1.07	-2.09**	23.70	0.76	8.70
n. of days sick (HH head)[1]	-	-	1.72	2.55	-	-	-1.65	17.10
education & occupation of HH's head:								
msce certificate†	0.01	0.03	0.03	0.05	-1.10	12.50	-1.02	10.80
professional training†	0.15	0.14	-	-	0.28	3.10	-	-
occupation in agriculture†	0.86	0.80	-	-	1.73*	19.70	-	-
contract labourer†	-	-	0.01	0.03	-	-	-1.55	14.90
households assets:								
land size (ha)	2.27	2.39	2.12	2.83	-0.44	5.00	-2.64***	24.00
share of land owned by spouse (%)	32.73	21.92	14.09	10.65	2.32**	26.30	0.96	11.60
n. of gifts	0.10	0.09	-	-	0.41	4.60	-	-
characteristics of the community:								
total n. of households	201.72	246.19	468.04	299.99	-1.40	15.90	3.49***	43.50
electricity†	0.17	0.23	-	-	-1.28	14.50	-	-
distance to the government (Km)	22.50	15.29	10.20	8.93	4.52***	51.20	1.47	16.80
distance to the credit office (Km)	10.95	6.40	3.01	3.53	2.87***	32.5	-0.94	11.10
N. of observations	153	158	383	94				

Source: own calculation from MRFMHFS.†dummy variables.[1] month prior to interview.***$p < 0.01$,**$p < 0.05$,*$p < 0.1$.Total number of observations: Model I=311, Model II=477.

one programme. The untreated group is past membership. As mentioned above, the choice of the second treatment ($T = 2$) allows keeping many observations, but it neglects programmes heterogeneity. In order to (partially) overcome the latter problem, we include covariates that affect the eligibility to all programmes. Also, because we have observed in table 4.5 that only one credit programme (MRFC) is available in the district of Dedza, we estimate the propensity scores conditional on the existence of the programmes in the districts.

Table 4.9 contains the descriptive statistics of the covariates included in the logit estimation of the propensity scores, separately for each treatment group. The first two columns display the mean for each covariate in the two treatment groups. Although most of the literature performs a pair-wise comparison of each of the T treatments, we have only compared participation in the MRFC and in more than one programme with past participation. Indeed, participation in more than one programme is not an appropriate *untreated* group for the membership in the MRFC because eligibility varies across credit programmes. Instead, past members are the proper untreated[30] group of *potential* members of a credit programme.

The last column of table 4.9 shows two statistics to compare treated and untreated groups over the set of covariates used to estimate the propensity scores. These statistics are the two-sample t-test and the percentage of absolute bias between treated and untreated groups. This is a common approach in policy evaluation [e.g. Caliendo et al. 2005; Lechner, 1999b; Sianesi, 2004]. We use the two-sample t-test to compare the means of the covariates for treated and untreated households. Also, Rosenbaum and Rubin (1985) used the standardised bias indicator to assess the distance in marginal distributions of the X covariates. The standardised bias is given by:

[30]See above a full justification of the choice of untreated group.

$$SB = 100 * \frac{\left(\bar{X}^m - \bar{X}^l\right)}{\sqrt{\frac{1}{2}\left[V_m(X) - V_l(X)\right]}}$$

where X_m and V_m are the mean and variance of the group in treatment m before matching and X_l and V_l are the analogue for the untreated group. Hence, the above equation can be defined as the difference of sample means in the treated and untreated groups as a percentage of the square root of the average of sample variances in both groups [Caliendo and Kopeinig, 2005].

Based on these statistics, we see that there is a moderate to large difference between the treatment and control groups. In particular, households participating in either the MRFC only or in any other credit programme have a significantly older household's head compared to past members. The bias attached to the MRFC treatment and to the second programme is approximately 36 percent and 24 percent, respectively, meaning that there is a quite large difference in the means of treated and untreated groups. Comparing households participating in the MRFC programme with the untreated group of past members, we found that they have significantly less children - between 6 and 10; they are more likely to have a household head employed in agriculture; the spouse owns a higher share of land; and they live in villages that are more distant to government and credit offices. On the other hand, by comparing participants in more than one credit programme with past members we found that they have significantly larger families; they are more likely to have a female head; they have less land; and they live in more populated villages.

Hence, the aim of propensity score matching is to use an appropriate control group that would attenuate the differences between treated and untreated households by reducing the standardised bias and by eliminating the significance of the two sample t-test.

As mentioned above, the propensity scores of participation are obtained from a series of logit models defined as follows:

$$T^{*k}_{ij(i)} = x'_{ij(i)}\beta_0 + C_{j(i)}\beta_1 + u_i \tag{4.6}$$

where the subscript $i = 1, 2, \ldots, N$ indicates the ith household; and *j(i)* indicates the village where household i lives. Also, $T^{*k}_{ij(i)}$ is the unobserved propensity to participate where $T^{k}_{ij(i)} = 1.(T^{*k}_{ij(i)} > 0)$; $k = m, l$ indicates each treatment group (i.e. m=MRFC only or 2^{nd} programme; and l=past members). The model includes a vector $x_{ij(i)}$ of households' characteristics, education and occupation of household head and a vector $C_{j(i)}$ of community characteristics that vary only across villages but not across households. In addition, we include district dummies and round dummies (i.e. we pool across different seasons).

The propensity scores are the predicted values, $\hat{P}^{m|ml}(x)$, estimated from equation 4.6 where in our case m is either participation in the MRFC only (model I) or participation in more than one programme (model II) and l is past membership. Notice that model II in table 4.11 has been estimated conditional on the existence of at least one additional programme in the district (i.e. we have excluded the district of Dedza where only the MRFC programme exists). Also, as mentioned above, the choice-based corrected probabilities, $\omega\hat{P}^{m|ml}(x)$, are obtained by using the Manski-Lerman weights (1977).

Before presenting the logit regression, we run a linear probability model (i.e. an OLS on a binary choice variable). The results shown in table 4.10 are the same in significance and sign as the ones using a logit model (displayed in table 4.11). Different regressors have been used for the two models for two reasons. First, the two programmes had different eligibility as well as utilisation. So, for example whilst gender was not a selection criterion in the MRFC, it was in one of the other credit programmes. Also, whilst

Table 4.10: Linear probability models of participation

Pr(participation in ...)	**MRFC vs. Past member (I)**	**2nd programme vs. Past member (II)**
households characteristics		
hh size	0.03	0.05
	*(0.02)****	*(0.01)****
age head	0.00	0.002
	(0.002)	*(0.00)*
female head†	-	0.05
		(0.07)
n. of children 6-10	-0.07	-0.04
	*(0.03)****	*(0.04)*
n. of days sick (HH head)[1]	-	-0.02
		(0.01)***
education & occupation of HH head		
msce certificate†	-0.080	0.21
	*(0.31)***	*(0.24)*
professional training†	0.23	-
	*(0.10)****	
occupation in agriculture†	0.06	-
	(0.06)	
contract labourer†	-	0.21
		(0.16)**
households assets		
land size (ha)	0.02	0.01
	(0.01)	*(0.01)*
share of land owned by spouse (%)	0.00	0.00
	(0.00)	*(0.00)*
n. of gifts	0.05	-
	(0.07)	
community characteristics		
total n. of households	0.00	0.00
	*(0.00)****	*(0.00)*
electricity†	0.09	-
	(0.07)	
distance to government office (Km)	0.01	-0.01
	*(0.00)****	*(0.00)*
distance to credit office (Km)	0.00	0.00
	(0.00)	*(0.01)*
Dowa†	0.38	0.39
	*(0.16)***	*(0.12)****
Nkhotakota†	0.02	-0.01
	(0.07)	*(0.12)*
Rumphi†	0.18	-0.01
	*(0.08)***	*(0.11)*
round 2†	0.04	0.03
	(0.05)	*(0.07)*
round 3†	0.05	0.11
	(0.05)	*(0.07)**
Constant	-0.30	0.09
	*(0.12)***	*(0.15)*
N. Obs	312	477
R^2	0.44	0.41

Source: own calculation from MRFMHFS. Note: Robust std. errors in ().Weighted regression. †dummy variables.[1] month prior to interview.***$p < 0.01$,**$p < 0.05$,*$p < 0.1$.

Table 4.11: Series of logit models of participation

Pr(participation in ...)	**MRFC vs. Past member (I)**	**2nd programme vs. Past member (II)**
households characteristics		
hh size	1.24	1.35
	*(0.11)***	*(0.13)****
age head	1.00	1.01
	(0.01)	*(0.01)*
female head†	-	1.30
		(0.58)
n. of children 6-10	0.54	0.81
	*(0.13)****	*(0.19)*
n. of days sick (HH head)[1]	-	0.92
		(0.03)**
education & occupation of HH head		
msce certificate†	0.002	0.21
	*(0.00)****	*(0.24)*
professional training†	4.23	-
	*(2.18)****	
occupation in agriculture†	2.06	-
	(1.15)	
contract labourer†	-	0.21
		(0.16)**
households assets		
land size (ha)	1.14	1.04
	(0.11)	*(0.07)*
share of land owned by spouse (%)	1.00	1.00
	(0.00)	*(0.01)*
n. of gifts	1.3	-
	(0.68)	
community characteristics		
total n. of households	1.00	1.00
	*(0.00)***	*(0.00)*
electricity†	1.79	-
	(0.86)	
distance to government office (Km)	1.06	0.97
	*(0.02)****	*(0.03)*
distance to credit office (Km)	1.01	1.03
	(0.02)	*(0.04)*
Dowa†	12.89	12.55
	*(14.08)***	*(12.85)***
Nkhotakota†	1.85	0.80
	(1.34)	*(0.53)*
Rumphi†	4.91	0.37
	*(2.92)****	*(0.24)*
round 2†	1.37	1.25
	(0.49)	*(0.50)*
round 3†	1.63	2.09
	(0.61)	*(0.92)**
N. Obs	311	477
Pseudo R^2	0.39	0.29

Source: own calculation from MRFMHFS. Note: odds ratios displayed and robust std. errors in ().
Weighted regression.†dummy variables.[1] month prior to interview.***$p < 0.01$,**$p < 0.05$,*$p < 0.1$.

the MRFC were delivered exclusively for agricultural reasons, the same is not true for the other credit programmes. This is why, for example, occupation in agriculture was excluded. Second, we tested for the significance of the regressors and we excluded only insignificant regressors (the inclusion of which did not change the results of the other regressors).

Table 4.11 displays the odds ratio, e^{β}, for each of the two logit models. As explained in the third chapter, the coefficients should be interpreted as follows: for a unit change in the regressor, the odds are expected to change by a factor e^{β}, holding all other variables constant. The correspondent coefficient, β, can be found by taking the logarithm of the odds ratio. The sign of the coefficient is positive when $e^{\beta} > 1$ and negative otherwise[31].

We briefly comment on the sign of the coefficients because the regression is only used to predict the propensity scores. The probability of participating only in the MRFC programme increases for larger families, but with fewer children - between 6 and 10. The household head with a MSCE certificate[32] is less likely to be a member of the MRFC, but he is more likely to participate in the programme if he has professional training[33]. Members of the MRFC programme are more likely to live in villages that are larger and more distant to government offices. An increase in household size significantly increases the probability of participating in more than one programme. Moreover, the probability of being a member of more than one programme decreases when the household head is a contract labourer or has been frequently sick.

[31]For binary variables, going from 0 to 1, one can interpret the odds ratios directly without any transformation. A transformation is needed to compare continuous variables (measured on different scales) with each other. One way is to compare a one standard deviation change in two differently measured variables.

[32]MSCE=Malawi School Certificate of Education corresponds to high school certificate at age 16-17.

[33]As said earlier, we did not include income in order to avoid reversed causality. The inclusion of education and professional training which are higly correlated with income partially controls for this omitted variable.

4.4.2 Second stage: matching algorithm

After having obtained the scores from the first stage, we perform matching. Matching involves selecting a *control* group from a pool of *untreated* households in which the distribution of observed variables is as similar as possible to the distribution in the treated group.

In the multiple treatment case, as noted by Vinha (2006), because matching is performed on more than one conditional probability, a nearest neighbour algorithm is usually adopted. In general, the treatment impact in the common support region is given by:

$$\hat{M}(T) = \frac{1}{N_m} \sum_{i \in I_m \in C} \left[Q_i^m - \sum_{j \in I_l} w_{ij} Q_j^l \right] \tag{4.7}$$

where N_m is the subset of households used for matching in treatment m, I_m and I_l are the full set of households available in options m and l respectively, Q_i^m and Q_j^l indicate the amount borrowed from informal sources by household i treated in programme m and by control household j in option l, respectively. In equation 4.7 the weight function w_{ij} determines a nonparametric regression that gives a higher weight to nonparticipant households j the stronger the similarity to participants i in terms of observable characteristics X. So, w_{ij} is the weight given to the control household j compared to household i in the treatment group, such that $\sum_{j \in I_m} w_{ij} = 1$. That is, the weights of the control households sum to one for each treatment observation. Only treatment households within the common support region are used.

Different matching algorithms employ different forms of the weighting function w_{ij}. In this sub-section we use Mahalanobis metric matching with propensity scores. This algorithm is implemented by randomly ordering households and then calculating the distance between the first treated households and all the controls, where the Mahalanobis distance between a treated household i and a control household j is defined by:

$$d(i,j) = \left(P_i^m - P_j^l\right)' V^{-1} \left(P_i^m - P_j^l\right) \tag{4.8}$$

where P_i^m and P_j^l are the propensity scores in options m and l for treated household i and control household j. V is the sample covariance matrix from the full set of households[34].

According to this matching algorithm, the control household j with the minimum Mahalanobis distance is used as a match for treated household i and both households are removed from the pool. The process goes on until matches are found for all treated households.

How do we assess the quality of matching?

There are several ways to assess the quality of matching. Figure 4.3 shows the distribution of the predicted propensity scores between treated and untreated groups for the participation in only the MRFC (upper panel) and in more than one programme (lower panel). Good matching is achieved whenever the distributions of treated and untreated groups are similar. While there is good overlap between the distributions of the propensity score in the two treatment groups, it can be seen that for values of the propensity score higher than 0.5 the number of households who participate in the MRFC only is larger than past members. An opposite trend can be seen in the lower panel, that is, for values of the propensity score lower than 0.4 the number of households who

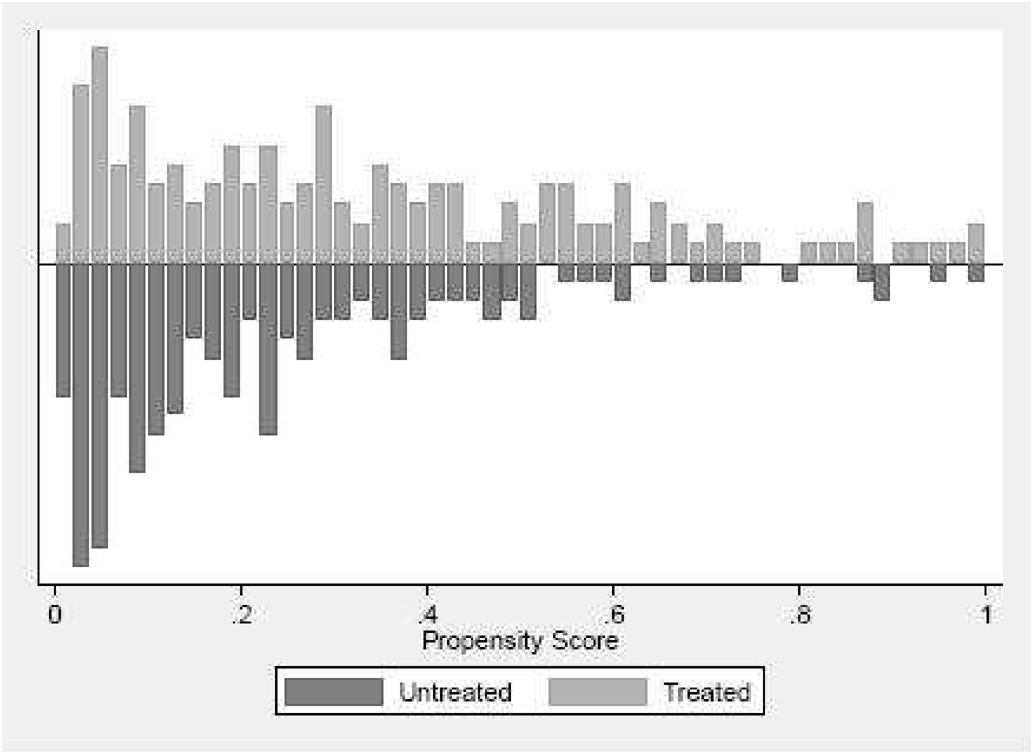
[34]Rubin (1980) showed the covariance matrix to be:

$$V = \frac{\{(N_m - 1)V_m + (N_l - 1)V_l\}}{(N_m + N_l - 2)}$$

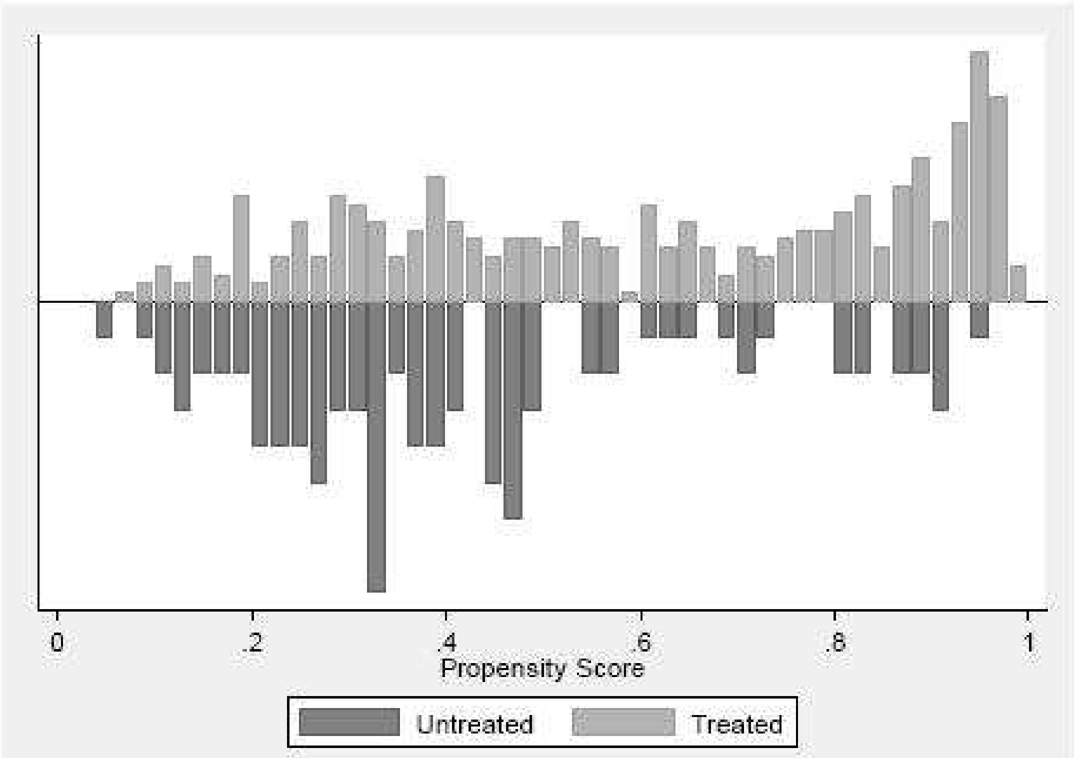
where N_k is the number of observations in treatment k and V_k is the sample covariance of the relevant propensity scores, P, in the treatment group $k = m, l$.

Figure 4.3: Bar charts of propensity scores

MRFC vs. Past members



2nd Program vs. Past members



Source: Own calculation based on MRFMHFS.

participate in more than one programme is lower than past members.

Overall, the samples are well matched on the propensity scores with quite similar distributions in the treated and untreated groups. In addition to the analysis of the overlapping regions, we check that matching has reduced the bias between treatments and controls for the set of covariates[35] used to estimate the propensity scores.

Table 4.12 displays the reduction of bias and the two-sample *t*-test for the selected characteristics that had a bias higher than 10 percent prior to matching. In the policy evaluation literature [e.g. Austin and Mamdami, 2006; D'Agostino, 1998; Manca and Austin, 2008] a standardised bias higher than 10 percent (and sometimes 20 percent) is taken to denote high imbalance in a covariate between treatment and controls. If matching has worked, the covariates should be balanced and no significant differences should be found after matching.

The first two columns of table 4.12 show the mean of treated and control groups for the two models: MRFC versus past members (model I) and 2^{nd} programme versus past members (model II). The last two columns display the group comparisons for each of the models based on the *t*-test and on the absolute percentage reduction of bias obtained by comparing the standardised bias (SB) of treated and control groups before and after matching. The higher the reduction of bias the better balance has been achieved on that covariate. In some cases[36] there is an enormous reduction of bias showing that the matching procedure is able to balance the observed characteristics of treatment and control groups. However, even after matching some regressors are still significant (i.e. age of the head in model I; and female head, share of land owned by spouse and total number of households in model II). In the fourth stage, we perform a sensitivity analysis by dropping these significant regressors to see whether our results change.

[35] The matching algorithm is conditional on the districts where control and treated households live.
[36] Where we achieved a 100 percent reduction of bias in some regressors.

Table 4.12: Selected characteristics by treatment groups after matching (when bias prior to matching > 10%)

	(I) MRFC vs. Past members		(II) 2nd programme vs. Past members		Group comparisons (I)		Group comparisons (II)	
	treated	*controls*	*treated*	*controls*	**t-stat.**	**% reduc. $\|bias\|$**	**t-stat.**	**% reduc. $\|bias\|$**
households characteristics:								
household size	-	-	6.44	6.31	-	-	0.78	82.20
age head	49.16	46.29	46.33	45.51	1.90*	40.4	0.95	72.6
female head†	-	-	0.31	0.25	-	-	2.02**	45.20
n. of children 6-10	0.92	0.87	-	-	0.56	75.3	-	-
n. of days sick (HH head)[1]	-	-	1.72	1.72	-	-	-0.00	100.00
education of households head:								
msce certificate†	0.01	0.01	0.03	0.03	0.00	100.00	0.21	100.00
occupation in agriculture†	0.86	0.88	-	-	-0.35	81.80	-	-
contract labourer†	-	-	0.01	0.01	-	-	0.00	100.00
households assets:								
land size (ha)	-	-	2.12	2.00	-	-	0.98	83.60
share of land owned by spouse (%)	32.73	35.05	14.09	9.79	-0.45	78.60	2.05**	24.80
characteristics of the community:								
total n. of households	201.72	212.19	468.04	409.08	-0.34	76.50	1.93**	64.90
electricity†	0.17	0.16	-	-	0.31	77.4	-	-
distance to the government (Km)	22.50	21.10	10.20	9.91	0.86	80.50	0.55	77.10
distance to the credit office (Km)	10.95	10.91	3.01	3.26	0.02	99.20	-0.73	51.20

Source: own calculation from MRFMHFS.†dummy variables.[1] month prior to interview.***$p < 0.01$,**$p < 0.05$,*$p < 0.1$.

4.4.3 Third stage: estimation of the average effects

In the third stage we estimate the average treatment effects on the amount households borrow from informal lenders. It is important to highlight that our outcome variable refers to a variety of lenders: friends, relatives and other informal lenders such as moneylenders and traders. In sub-section 4.3.2.2, we showed that the majority of households borrow from friends and relatives.

Table 4.13 reports the results for the average effects: the upper panel displays the estimated pair-wise average treatment effects on treated (ATT) and the lower panel shows the pair-wise average effects for the population. Each estimated average treatment effect on treated is reported in absolute and relative terms. As Larsson (2000) pointed out, the absolute size of the effects allows for a comparison of the effects between treatment and control households. In addition, the relative effects in percentage points indicate how considerable is the size of the effect.

TABLE 4.13: Average effects from Mahalanobis matching

	(a) Average Treatment Effect on Treated (ATT): $\vartheta_0^{m,l} = E[Q^m \mid T=m, P^{m\mid ml}(X)] - E[Q^l \mid T=m, P^{m\mid ml}(X)]$			
Outcome:	***(I) m=MRFC; l=Past Members***		***(II) m=2^{nd} programme; l=Past Members***	
	Difference	*t-stat.*	*Difference*	*t-stat.*
Credit from informal lenders†	-25.44	-2.06**	1.29	0.04
	(-73.98%)		**(1.95%)**	
	(b) Average Treatment Effect (ATE): $\gamma_0^{m,l} = E\left(Q^m - Q^l\right) = EQ^m - EQ^l$			
Outcome:	***(I) m=MRFC; l=Past Members***		***(II) m=2^{nd} programme; l=Past Members***	
	Difference		*Difference*	
Credit from informal lenders†	-23.15		4.14	

Source: own calculation from MRFMHFS.†Value in MK. 15 Malawian Kwachas (MK)=1 US$ Malawi's per capita GNP is US$ 170 (approx. 2,550 MK. World Bank, 1997).**$p < 0.05$

First, let us describe the effect of participating only in the MRFC programme on the amount households borrow from informal sources. Compared to past members, those who participate only in the MRFC borrow significantly less from informal lenders. In other words, we found strong evidence of crowding out of informal loans as in some of the above mentioned empirical studies [i.e. Attanasio and Rios-Rull, 2000; McKernan et al., 2005]. In particular, membership in one microfinance programme reduces the borrowing from informal sources by almost 2 U.S. dollars (approximately 25.5 MK). In relative terms it reduces the amount members borrow from informal lenders by 284 percent. The average treatment effect on the population confirms the above results, but with a slightly smaller impact of approximately 1.5 U.S. dollars (23.2 MK). On the other hand, there is no evidence of crowding out when households participate in more than one credit programme.

An explanation for the above result is the different nature of the two groups and is in line with what was found by Navajas et al. (2003) and Cox and Jimenez (2005). Table 4.14 compares households participating only in the MRFC with households participating in more than one credit programme. The latter group turns out to be relatively better off in terms of having a more valuable house, larger plots of land and higher food expenditure. So, we could interpret the participation in more than one credit programme as an indicator of being a relatively less constrained household. As in Navajas et al. (2003), we found that less capitalized borrowers switch from an informal credit contract to a loan contract provided by a microfinance institution (i.e. the MRFC). We will further discuss this issue in the conclusions. The insignificant effect for the group of households participating in more than one microfinance programme may also be affected by the fact that we pool different types of programmes in the treated group and past members of several programmes in the control group. Unfortunately, we do not have enough observations to disentangle the effect of each microfinance programme and hence we cannot further investigate this issue.

Table 4.14: Characteristics of groups of borrowers

	MRFC only	2nd Program	Past participants
Households demographics:			
Female headed HHs (%)	23.1 (153)	21.4 (388)	10.6 (159)
Average HH size	5.2 (153)	6.2 (388)	5.7 (159)
Average number of children 0-15	2.7 (153)	3.2 (388)	3.2 (159)
Household heads main occupation: agriculture (%)	53.2 (153)	46.1 (387)	86.7 (159)
Households assets and shock:			
Households affected by negative shocks (%)	58.9 (152)	59.7 (383)	51.6 (154)
Average land size (ha)	1.9 (150)	2.4 (367)	2.1 (153)
Share of land owned by spouse	23.3 (150)	11.7 (367)	24.5 (153)
Average value of house (MK)	1046 (150)	1061 (367)	846 (153)
Share of assets held as land (%)	56.1 (150)	57.9 (367)	63.7 (153)
Average food expenditure (MK)	10.7 (150)	15.5 (367)	9.5 (153)

Source: Own calculation based on MRFMHFS. Note: household types are defined according to participation in the credit programs. Weighted results. Number of observations in parentheses. Expenditure deflated by the square root of household's size.

To conclude, it is worth highlighting again that the validity of our results rests on the assumption that selection into the programme is based on observable characteristics. We can offer two arguments in support of this. First, we have included a range of factors such as age, household size, occupation, education, land size that is likely to affect participation. Second, we have argued that the choice of past membership as comparison group should take into account unobservable factors such entrepreneurial ability which is assumed to be fixed over time.

4.4.4 Fourth stage: sensitivity analysis

The last stage of this evaluation method involves checking the robustness of our results. We perform three types of sensitivity analysis: a) change in the model specification and matching algorithm; b) change in treatment and outcome definition; and c) change of the model used to estimate the propensity scores. Let us analyse each of these robustness checks.

a) Change in the model specification and matching algorithm

Table 4.15 shows the sensitivity checks for the two groups of treatments: MRFC participants in the upper panel and participants in more than one programme in the lower panel. The last two rows in each panel show the absolute and relative values of the average treatment effects together with the t-statistic in parenthesis.

In model I, we drop the sampling weights. Heckman and Todd (2005) showed that with nearest neighbour algorithms[37] it does not matter whether matching has been performed on the odds ratio without weights since the ranking of the observations is identical and the same neighbour will be selected. As in Frölich et al. (2004), we found that the most significant results remain largely unchanged. Indeed, the average treatment effect on treated (ATT) households in the MRFC programme is still negative and significant. Both the relative and absolute effects remain almost unchanged. Although the ATT on households participating in more than one programme changed sign, it is still insignificant.

In model II, we drop the variables that remained significant after matching in stage two[38] (see table 4.12). Again, the results remain unchanged: a negative and significant ATT in panel (a) and a positive and insignificant effect in panel (b).

The last two columns of table 4.15 report the average treatment effects on treated households obtained after performing two different matching algorithms. Nearest neighbour matching involves finding for each treated household, the control household with the most similar propensity score. We have implemented this procedure with replace-

[37]Mahalanobis metric matching is a type of nearest neighbour matching.

[38]We dropped household head in the MRFC treatment and female head, share of land owned by spouse and total number of households in the 2^{nd} programme treatment.

TABLE 4.15: Sensitivity analysis of ATT to changes in model and matching algorithm

	(a) m=MRFC; l=Past Members			
Outcome:	**Different model specifications:**		**Different matching algorithms:**	
Credit from	*Model (I):*	*Model (II):*	*Model (III):*	*Model (IV):*
informal lenders†	*no weights*	*drop sign. vars.*	*Nearest Neighbour*[1]	*Kernel matching*
ATT	-25.29	-27.86	-21.46	-20.14
	*(-2.24)***	*(-2.22)***	*(-1.98)***	*(-2.38)***
% points	**(-73.86%)**	**(-75.69%)**	**(-71.58%)**	**(-70.27%)**

	(b) m=2ndprogramme; l=Past Members			
Outcome:	**Different model specifications:**		**Different matching algorithms:**	
Credit from	*Model (I):*	*Model (II):*	*Model (III):*	*Model (IV):*
informal lenders†	*no weights*	*drop sign. vars.*	*Nearest Neighbour*[1]	*Kernel matching*
ATT	-4.19	9.02	31.23	31.00
	(-0.13)	*(0.29)*	*(0.83)*	*(0.84)*
% points	**(-5.85%)**	**(15.44%)**	**(71.25%)**	**(70.32%)**

Source: own calculation from MRFMHFS.[1]Nearest Neighbour has been performed with caliper and replacement †Value in MK. 15 Malawian Kwachas (MK)=1 US$ Malawi's per capita GNP is US$ 170 (approx. 2,550 MK. World Bank, 1997).**$p < 0.05$. t-stat. in parenthesis.

ment, that is, while each treated household has only one match, the control household may be matched to more than one treated household. Dehejia and Wahba (2002) found that nearest neighbour with replacement produces a better matching.

In order to improve the quality of the match, we have also selected control households within a preset amount (or caliper) of the treated household's estimated propensity score. In other words, the nearest neighbour matching with replacement and caliper imposes an *a priori* common support region. More formally, keeping the same notation as before, for a pre-specified $\delta > 0$, treated household i is matched to untreated household j such that:

$$\delta > \left| P_i^m - P_j^l \right| = \min_{k \in C} \left\{ \left| P_i^m - P_j^l \right| \right\}$$

where P^k, with $k = m, l$ are the propensity scores for the two options and C is the set of neighbours of treatment households in the untreated group. Smith and Todd (2005) pointed out that a drawback of this algorithm is that it is difficult to determine *a priori* the size of caliper. We set our caliper $\delta = 0.02$ as a result of a maximization in the bias reduction and a minimization of loss of observations[39]. Again, model III in table 4.15

[39]We lose 4 observations in models III and IV of panel (a) and 45 observations in models III and IV of panel b.

confirms the results obtained by Mahalanobis matching for both treatment groups.

To further check the robustness of our results we perform a non-parametric estimator, the Kernel matching. As Smith and Todd (2005) pointed out, Kernel matching is like a weighted regression where the counterfactual outcome is constructed with a weighted average of all households in the control group. Unlike the nearest neighbour with replacement, the main advantage of this approach is that the variance is smaller as a result of the use of more information. Heckman et al. (1997) derived the asymptotic distribution of this estimator[40].

In other words, we have associated to the outcome Q_i^m of treated household i in treatment option m a matched outcome given by a kernel weighted average of the outcome of all untreated households, where the weight given to the untreated group j is proportional to the closeness between i and j. The application of the Kernel algorithm involves the choice of the Kernel function K and of the bandwidth h.

DiNardo and Tobias (2001) showed that the choice of Kernel does not affect the results. We have used a standard Epanechnikov Kernel[41]. As shown by Silverman (1986) and Pagan and Ullah (1999), the choice of bandwidth involves a trade-off between bias and variability. A large bandwidth decreases the variance by providing a better fit with a smoother density function. On the other hand, as the bandwidth increases the bias increases as well. We set the bandwidth to be equal to the caliper size in model III.

[40] The Kernel matching can be written as follows:

$$\hat{Q}_i^m = \frac{\sum_{j \in \{C\}} K\left(\frac{P_i^m - P_j^l}{h}\right) Q_j^l}{\sum_{j \in \{C\}} K\left(\frac{P_i^m - P_j^l}{h}\right)}$$

where the outcome of control household j in treatment option l is weighted by:

$$w_{ij} = \frac{K\left(\frac{P_i^m - P_j^l}{h}\right)}{\sum_{j \in \{C\}} K\left(\frac{P_i^m - P_j^l}{h}\right)}$$

[41] The Epanechnikov Kernel is given by: $K(u) \propto (1 - u^2)$ if $|u| < 1$, zero otherwise.

Once again, the average treatment effect on treated households in the MRFC programme is negative and significant. The absolute effect is slightly smaller with a value of 1.3 U.S. dollars (approximately 20 MK). The ATT in panel (b) is still positive but not significant.

To sum up, our results remain unchanged even after modifying the specification of the model or of the matching algorithm. Membership in the MRFC credit programme has a crowding out effect on informal sources, the absolute size of the effect ranges between 1.3 and 2 U.S. dollars according to different specifications. In relative terms, it reduces the amount members borrow from informal lenders by more than 70 percent in all the above mentioned specifications.

b) Change in treatment and outcome definition

In this section we change the definition of treatment. Previously, we estimated the effect of merely being a member of a microfinance programme on the amount households borrow from informal lenders. Now, what happens if we apply a stricter definition of treatments, that is, what happens if we define a treated household to be both a member and borrower from the microfinance programmes?

In order to answer this question, we repeat the above mentioned three stages of the evaluation procedure with the new definition of treatments. The logit models include different regressors and the results are reported from tables C4-2 to C4-4 of appendix C together with the indicators used to assess the quality of matching. Table 4.16 reports the average effects for the two groups of newly defined treatments. Although the magnitude of the ATT for the MRFC treatment is about the same as the ones described above (i.e. the crowding out effect is approximately 1.6 U.S. dollars), the significance has decreased.

The above result can be explained by the fact that not all the covariates included in the model achieved a good matching performance. As shown in table C4-4, for some variables such as "number of gifts" and "age head squared" the reduction of bias between treatment and control groups is around 30 percent. However, the relative crowding out effect is still quite large and above 70 percent.

In addition, we change the outcome variable. We now know that participation in the MRFC programme reduces the amount borrowed from informal lenders. But does this happen because households demand less or because informal lenders give them less credit (or both)? This ambiguity arises from the fact that demand and supply issues cannot be disentangled by simply looking at the amount borrowed from informal lenders.

The second row of table 4.16 looks at whether crowding out applies also to the demand for credit to informal lenders. The logit models we used are the same as the ones shown in table 4.11. We find a very large and significant reduction in the demand for informal finance for households who participate in the MRFC (-75.22 percent). There is no evidence of crowding out for households who participate in more than one credit programme (panel (II)).

The third row of table 4.16 disentangles the supply from the demand of informal loans by looking at the credit limit. As the credit limit variable is the maximum amount that the borrower thinks the lender is willing (or able) to lend, it can be thought to be the "supply" of informal loans[42]. This approach allows testing whether *transfers* from informal lenders are crowded out by the introduction of microfinance programmes.

There is an abundant theoretical literature that stems from Becker's (1974) and Barro's (1974) model of altruistic transfers. This literature suggests that public transfers crowd out private transfers motivated by altruism. Whilst some studies supported the altruism hypothesis [Lee et al., 1994; McGarry and Schoeni, 1995a, b; Secondi, 1997], some others found evidence against pure altruism [Altonji et al., 1992; Bernhem et al.,

[42]However, caution should be used in interpreting it as a supply function because the lender is not price taker in this market. Chapter five describes the credit limit variable in more detail.

TABLE 4.16: Sensitivity analysis of ATT to changes in treatment and outcome definition

	(a) Average Treatment Effect on Treated (ATT): $\vartheta_0^{m,l} = E[Q^m \mid T=m, P(X)] - E[Q^l \mid T=m, P(X)]$			
Outcome:	*(I) m=MRFC; l=Past Members*		*(II) m=2^{nd} programme; l=Past Members*	
	Difference	*t-stat.*	*Difference*	*t-stat.*
Credit from informal lenders†	-24.34	-1.75**	39.07	0.68
	(-70.63%)		**(54.91%)**	
Demand from informal lenders†	-29.75	-2.30**	4.45	0.14
	(-75.22%)		**(6.15%)**	
Credit limit of informal lenders†	-67.62	-1.55	28.02	0.31
	(-69.14%)		**(19.50%)**	
	(b) Average Treatment Effect (ATE): $\gamma 0^{m,l} = E\left(Q^m - Q^l\right) = EQ^m - EQ^l$			
Outcome:	*(I) m=MRFC; l=Past Members*		*(II) m=2^{nd} programme; l=Past Members*	
	Difference		*Difference*	
Credit from informal lenders†	-13.54		60.87	
Demand from informal lenders†	-27.92		6.35	
Credit limit of informal lenders†	-76.33		12.05	

Source: own calculation from MRFMHFS. †Value in MK. 15 Malawian Kwachas (MK)=1 US$ Malawi's per capita GNP is US$ 170 (approx. 2,550 MK. World Bank, 1997).**$p < 0.05$, *$p < 0.1$.

1985; Cox, 1987; Cox and Rank, 1992; Hayashi, 1995].

Although most of the loans in Malawi are supplied by friends and relatives, we cannot specifically support the altruism hypothesis because we had to aggregate informal loans (including those given by moneylenders) due to lack of observations. Taking this into consideration, the results in table 4.16 show no significant evidence of crowding out of the supply of informal loans.

c) Change of the model used to estimate the propensity scores

The propensity score could also be estimated by using a multinomial logit model with three alternatives:

$$Pr\left(T_i = m\right) = \frac{\exp(x_i \beta_m)}{\sum_{l=0}^{L} \exp(x_i \beta_l)}$$

where m denotes the treatment choice and i the household; x is the vector of covariates including household characteristics and assets, education and occupation of the household head. The model in table 4.17 also includes some characteristics of the villages that do not vary across households. Here we denote the choice alternatives as follows: past members ($T = 0$), participation only in the MRFC programme ($T = 1$) and participation in more than one programme ($T = 2$). Thus, in the above equation $L = 2$.

The coefficients are expressed as odds ratios. In order to interpret them, we choose past membership as the comparison base group (or outcome). We have not applied the sampling weights because they would impede[43] the estimation of the IIA test.

The results are more or less the same as the ones shown in the two logit models of table 4.11. In particular, as the number of children increases, the probability of being a member of the MRFC programme or member of more than one credit programme descreases. The characteristics of the community are also significant. For instance, the likelihood of participating in the MRFC programme increases with the distance to the government office. This is not an unrealistic result since households are more willing to form credit groups than incurring in often substantial transportation costs.

Unlike in the series of logit models, the conditional predicted scores cannot be obtained directly, but need to be calculated from the unconditional predicted probabilities as follows:

$$\hat{P}^{m|ml}(x) = \frac{\hat{P}^m(x)}{\hat{P}^m(x) + \hat{P}^l(x)}$$

for each of the treatments where m is either participation in MRFC only or participation in more than one programme and l is past membership[44].

[43]The Hausman command generates a negative χ^2-statistic, rendering the test infeasible. This happens because the variance of the difference of the coefficient vectors is not positive definite in finite samples. A solution could be to use *suest*, but we have shown that excluding the weights does not affect the results. So, we choose to slightly change the specification by dropping the weights.

[44]It is just the same as the previous equation, but with different notation.

Table 4.17: Multinomial logit model of participation - base outcome: past members

Pr(choice=...)	**MRFC**	**2nd programme**
households characteristics		
household size	1.09	1.17
	(0.08)	*(0.07)****
age head	1.02	1.01
	*(0.01)***	*(0.01)*
n. of children 6-10	0.73	0.70
	*(0.12)**	*(0.10)***
education & occupation of HH head		
msce certificate†	0.41	0.56
	(0.39)	*(0.37)*
households assets		
land size (ha)	1.04	0.95
	(0.06)	*(0.05)*
share of land owned by spouse (%)	1.00	1.00
	(0.00)	*(0.00)*
community characteristics		
total n. of households	1.00	1.00
	*(0.00)**	*(0.00)*
distance to government office (Km)	1.07	1.00
	*(0.02)****	*(0.02)*
distance to credit office (Km)	0.99	0.98
	(0.01)	*(0.01)*
Mangochi†	8.12	18.66
	*(5.67)****	*(11.30)****
Nkhotakota†	2.08	3.74
	(1.05)	*(1.45)****
Rumphi†	2.93	4.34
	*(1.34)***	*(1.62)****
round 2†	1.26	1.11
	(0.35)	*(0.30)*
round 3†	1.47	1.75
	(0.45)	*(0.46)***
N. of obs.	700	
Pseudo-R^2	0.18	

Source: own calculation from MRFMHFS. Note: odds ratios displayed and std. errors in (). ***$p < 0.01$,**$p < 0.05$.

Table 4.18: Test for Independence of Irrelevant Alternatives - IIA

Outcome	***Small-Hsiao test of IIA assumption*** H_0***: Odds(outcome-j vs. outcome-k) are independent of other alternatives***					
(N=700)	*lnL(full)*	*lnL(omit):*	χ^2*:*	*Degrees of freedom*	*Prob.*$> \chi^2$	*Evidence*
MRFC only	-170.82	-146.69	48.26	15	0.00***	against H_0
2^{nd} **programme**	-110.35	-89.01	42.67	15	0.00***	against H_0

Source: own calculation from MRFMHFS.**$p < 0.05$.

The multinomial logit model requires the *Independence of Irrelevant Alternatives* (IIA)[45] assumption to hold. For example [Larsson, 2000], suppose that a household has three choices: no programme, programme 1 and programme 2, with respective probabilities 3/10, 6/10 and 1/10. The IIA property states that if we drop programme 2, the relative probability of programme 1 to no programme, $6/3 = 2$, does not change so that the new probabilities become 1/3 and 2/3 for no programme and programme 1, respectively. However, if the programmes are at least partly substitutes to each other, the new probabilities may be expected to be nearer to 3/10 and 7/10.

We could not apply the Hausman test because the variance-covariance matrix is not positive and hence we used the Small-Hsiao test[46]. The test strongly rejects the *Independence of Irrelevant Alternatives'* hypothesis (table 4.18). This is why our preferred specification is the series of logit models. Nevertheless, we want to check the robustness of the previous results and hence we repeat the three stages of our evaluation method. The indicators of matching quality are reported in Appendix C (see tables C4-5 and C4-6). The average effects are reported in table 4.19.

Membership in the MRFC credit programme reduces the borrowing from informal sources by 2 U.S. dollars (approximately 30 MK). In relative terms, it reduces the amount members borrow from informal lenders by more than 75 percent. This effect is larger than the one found previously. The average treatment effect on the population confirms the above results, with an impact of around 1.6 U.S. dollars (approximately 24 MK). The results for the participation in more than one programme (second column of table 4.19) are very similar to those obtained in the specification without sampling weights. In particular, there is no significant effect of participating in more than one microfinance programme on the amount households borrow from informal lenders.

[45]This test is explained in Appendix B.
[46]See Appendix B for a description of this test.

TABLE 4.19: Average effects from Mahalanobis matching

	(a) Average Treatment Effect on Treated (ATT): $\vartheta_0^{m,l} = E[Q^m \mid T = m, P^{m\mid ml}(X)] - E[Q^l \mid T = m, P^{m\mid ml}(X)]$			
Outcome:	***(I) m=MRFC; l=Past Members***		***(II) m=2^{nd} programme; l=Past Members***	
	Difference	*t-stat.*	*Difference*	*t-stat.*
Credit from informal lenders†	-29.92	-2.48**	-23.34	-0.75
	(-76.97%)		**(-25.92%)**	

	(b) Average Treatment Effect (ATE): $\gamma_0^{m,l} = E\left(Q^m - Q^l\right) = EQ^m - EQ^l$	
Outcome:	***(I) m=MRFC; l=Past Members***	***(II) m=2^{nd} programme; l=Past Members***
	Difference	*Difference*
Credit from informal lenders†	-23.83	9.41

Source: own calculation from MRFMHFS.†Value in MK. 15 Malawian Kwachas (MK)=1 US$ Malawi's per capita GNP is US$ 170 (approx. 2,550 MK. World Bank, 1997).**$p < 0.05$

4.5 Polychotomous selection model

This section discusses whether the results obtained by matching on the propensity scores are appropriate for this evaluation problem. Propensity score matching is no *panacea* to all problems. There are a number of issues to be considered. Firstly, not only is the conditional independence assumption impossible to test, but it also leaves some uncertainty about the inclusion of all variables affecting selection in the various credit programmes. If there are unobservable factors that affect simultaneously the assignment in one of the programmes and the outcome variable, a hidden bias may arise. Secondly, even assuming we have controlled for all variables affecting participation, there are several issues concerning the choice of matching algorithm and of the discrete choice model. Although the second issue has been extensively addressed by changing the specifications of our evaluation method, the selection on unobservables issue has not yet been discussed.

Following Lee (1983) we extend the Heckman selection model to a polychotomous case with continuous dependent variable. The model we estimate consists of two steps. The first step is the same as the one applied for estimating the propensity scores. In particular, we estimate a binary choice model for each of the treatments:

$$T^{*k}_{ij(i)} = x'_{ij(i)}\beta_0 + C_{j(i)}\beta_1 + u_i \tag{4.10}$$

where $k = m = 1, 2$ denotes the following treatments: participation in the MRFC programme only and participation in more than one credit programme; *j(i)* indicates the cluster where household i lives. This is exactly the same model in equation 4.6.

The second step involves the estimation of the outcome equation adjusted for selection[47]. To be precise, the amount households borrow from informal sources can be estimated by using the following OLS model:

$$Y_{ij} = x'_i\beta_0 + T^m_i\alpha_1 + T^l_j\alpha_2 + \lambda^1_i\beta_1 + \lambda^2_j\beta_2 + u_i \tag{4.11}$$

where in this case i and j indicate the two different treatment households. If the *lambda* coefficients are significant then selection is based on unobservables. While Larsson (2000) estimated the first step with a multinomial logit model, we have used a series of binary models because of the rejection of the IIA. The first step estimation coincides with the estimation of the propensity scores. The results of the second step are in table C4-7 of appendix C. We have included all the variables affecting both the participation in the MRFC and in more than one programme. Identification requires including one variable in the first step that is not contained in the second step. Indeed, in equation 4.11 we have not included the community characteristics.

The inverse of Mills ratio is not significant for both treatments. Hence, we can conclude that selection is not based on unobservables. A possible explanation of this result

[47] By using the Mills' ratio obtained from the first stage. This can be written as:

$$\lambda^k_{ij} = \frac{\varphi(\gamma' W_{ij})}{\Phi(\gamma' W_{ij})}$$

where $W = (x, C)$ and k is defined above. We then obtain as many Mills ratios as treatments.

lies in our choice of untreated group. As mentioned above, because past members have the same (fixed) entrepreneurship ability as participants we have been able to compare two groups with the same observed as well as unobserved characteristics. However, this model has some drawbacks: a) the number of observations is quite low; and b) we could not include the programme dummies because of collinearity.

4.6 Conclusion

The role of microfinance institutions in markets where there are other informal lenders is relevant at the policy level. A government that wants to reach small borrowers could create new lending institutions that mimic the features of informal lending arrangements. Indeed, it is recognized that informal lenders overcome moral hazard and adverse selection problems by using localised informational arrangements and interlinkages. Microfinance institutions could also reach small borrowers by adopting joint liability schemes that enable borrowers to select safe fellow group members so to avoid the risk of default. So, would these institutions have any effect on households' access to informal sources? Or would they just serve a different segment of households leaving the competition in the credit market unchanged?

This chapter has addressed the following question: "Is there evidence of crowding out of group lending on informal credit?" Several empirical papers have been developed in the last fifteen years. However, only some of them have been able to establish a *causal* relation between introduction of microfinance programmes and reduction of informal loans. In none of these studies an increase in the supply of formal credit completely substitutes informal loans. For example, Attanasio and Rios-Rull (2000) showed that the introduction of Mexico's PROGRESA programme partially crowded out local insurance.

Nearly all the surveys have focused on the *realised* transfers rather than *potential* transfers [Cox and Fafchamps, 2008]. Yet, households' access to informal credit is affected both by access and membership in microfinance programmes.

We focused on the 1995 credit policy in Malawi and its effects on the access to informal loans. We used a rich financial survey: the Malawi Rural Financial Markets and Household Food Security Survey (FMHFS, 1995) conducted by IFPRI in cooperation with the Rural Development Department of Bunda College of Agriculture. The survey contains information about households' borrowing behaviour from both informal lenders and from microfinance institutions. Like some of the above mentioned studies [e.g. Attanasio and Rios-Rull, 2000; Kaboski and Townsend, 2006] we adopted policy evaluation techniques in order to identify a *causal* relationship between access to government sponsored credit programmes and informal loans. We used propensity score matching to determine the existence and the extent of the impact of group lending institutions on the access to informal loans.

The evaluation approach consisted of four stages. First, we obtained the propensity scores from a series of logit models. In the second stage we performed matching with the Mahalanobis metric algorithm. The third stage estimated the average treated effect (ATE) and the average treatment effects on treated households (ATT) who participate in one, or more than one, credit programme relative to past-membership. The outcome of interest is the amount households borrow from informal sources. The final stage ensured that the results were not dependent on the methodological assumptions of the evaluation procedure.

The chapter has developed a rigorous sensitivity analysis by performing the following robustness checks. It has changed the model specification and matching algorithm; the definition of treatment and outcome; and the model used to estimate the propensity scores.

We have found strong evidence of crowding out of group lending on informal sources. The results show that participation in one microfinance programme (i.e. the MRFC) has a negative and significant effect on the borrowing from informal sources. The absolute size of the effect ranges between 1.3 and 2 U.S. dollars according to different specifications. In relative terms, it reduces the amount members borrow from informal lenders by more than 70 percent in all the specifications.

Most of the literature focuses only on the crowding out effect of the supply of informal loans. The rich data set allows the researcher to disentangle the demand and supply of informal loans. The results show that the MRFC credit programme reduces the demand for informal credit. This is evidence of the fact that the MRFC programme and informal loans are, at least partly, substitutable.

This chapter has also innovatively applied the multiple treatments model of the labour economics literature [for example, Brodaty et al., 2001; Frölich et al., 2004] to test the crowding out hypothesis. This allows a comparison between the effectiveness of different credit programmes as well as between households that differ in their economic status.

The results show no significant crowding out effect of membership in more than one credit programme on the access to informal loans.

There are several explanations for the above results. As participants in more than one credit programme turn out to be relatively better off, we interpret this result as evidence that crowding out is affected by the credit constraints that arise from households' wealth heterogeneity. This is in line with findings of Navajas et al. (2003) who showed that less capitalized borrowers switch from an informal credit contract to a loan contract provided by microfinance institutions. Relatively wealthier households, by contrast, may not substitute one source for the other but simply increase the overall demand for credit once the supply of formal loans increases.

Secondly, the other credit programmes may not be substitute for informal loans as they serve different purposes[48]. Because all the programmes deliver loans for investments purposes, households keep on borrowing from informal sources to smooth consumption and use microfinance to buy inputs for the farm. However, while this explains multiple-borrowing, it does not explain why only poorer households reduce their borrowing from informal sources.

Finally, the insignificant crowding out effect of households participating in more than one credit programme may be affected by the fact that we pool different types of programmes. Unfortunately, we do not have enough observations to disentangle the effect of each credit programme and hence we cannot further investigate this issue.

The first two explanations seem to be most plausible. In particular, households could use multiple-borrowing because of market segmentation and because of credit rationing from formal credit programmes. It is actually possible to test the rationing hypothesis by using the information on credit limit provided by the survey. If a change in the credit limit of participants in the MRFC programme has a significant effect on the demand to informal sources, then we can interpret this result as evidence of the existence of credit constraints [Diagne, 1999; Diagne et al., 2000; Gross and Souleles, 2002]. We will deal with this issue in the next chapter.

In conclusion, formal credit, informal institutions and group lending programmes all entail a mix of social, political and economic incentives that are contingent on the local

[48]We have highlighted the importance of market segmentation in the second chapter. But it could also be the case that the other credit programmes are more expensive and this impedes crowding out of informal finance. We will look at this issue in the next chapter.

context. Only when microfinance programmes are designed in such a way as to meet local needs will microfinance hold greater potential for displacing informal finance.

Chapter 5

Credit constraints in Malawi

"The more constraints one imposes, the more one frees one's self. And the arbitrariness of the constraint serves only to obtain precision of execution".

I. Stravinsky (1882-1971)

5.1 Introduction

Why do formal and informal credit markets coexist? In spite of recent financial liberalisation aimed at broadening formal credit markets and interest rate differentials, in Africa, formal and informal credit sectors persist in the same market[1]. Two main explanations are offered by the literature. First, the informal sector may be the recipient of "spillover" demand from the rationed formal sector [Banerjee and Duflo, 2001; Bell et al., 1997; Besley, 1994; Eswaran and Kotwal, 1989]. The theoretical assumption of the spillover view is that informal credit is more expensive than formal loans. Therefore, according to this view, there is a natural ordering of credit sources whereby a borrower who uses secondary sources (i.e. informal credit) is assumed to be unable to satisfy his financial needs from the primary sources (i.e. formal credit). The borrower is said to be

[1]See chapter two for a more detailed discussion of this issue.

credit rationed with regard to the primary source[2]. Indeed in developing economies, such as in Africa, formal credit rationing is extensive because of information asymmetries, lack of collateral and legal enforcement [e.g. Ibrahim et al., 2007; Ghosh et al., 1999; Zeller, 1994].

Several empirical studies have found evidence of “spillover” effects. Bell et al. (1997) develop a model where a private and an institutional lender coexisted in a credit market because of spillovers from the latter source to the former. This model is discussed in more detail later. They used a cross-sectional switching regression model for demand and supply functions of credit in rural Punjab and showed that the formal credit market was responsible for most rationing. Banerjee and Duflo (2001) described a model of credit rationing in the context of firms and showed that an expansion in the availability of bank credit leads to a fall in the firm's borrowing from the market as long as the bank is the cheapest credit source.

An alternative explanation for the coexistence of formal and informal sectors is the occurrence of market segmentation. According to this view, the unique characteristics of the informal and formal credit sectors inhibit the substitution of one source for the other. As a result, the informal sector need not be the sector of last resort, but instead the preferred sector.

The chapter tests two hypotheses: 1) the spillover hypothesis; and 2) the liquidity constraint hypothesis[3]. This technical chapter aims at explaining the result of the previous chapter which showes that an increase in the supply of credit within a village cause a (partial) shift from informal sources to government-sponsored institutions. In this context, the spillover hypothesis implies that there is a certain degree of substitutability

[2] It might be possible, however, that she is also rationed on the use of the secondary source.

[3] See chapter two for a discussion of the literature on liquidity constraints in both developed and developing countries.

between the MRFC programme and informal credit and that a reduction of demand for the latter can be achieved by increasing the ceiling on the MRFC programme[4]. We also want to provide evidence for the existence of liquidity constraints in the credit provided both by government-sponsored programmes and by informal lenders[5]. As the spillover effect results from the existence of liquidity constraints, the two hypotheses are linked together.

The chapter uses the Malawi Rural Financial Markets and Household Food Security Survey (FMHFS, 1995), an original data set that contains information on credit limits and on the credit demand for both rejected applicants and borrowers, for formal and informal credit sources in Malawi[6].

We use information on the credit limit to test the above specified hypotheses. This is a direct approach to test for liquidity constraints rather than reduced form models using qualitative indicators[7] as developed by, amongst others, Jappelli (1990)[8]. Researchers at the International Food Policy Research Institute (IFPRI) use the credit limit concept to measure rationing [Diagne, 1999; Diagne et al., 2000; Zeller and Sharma, 1998]. The idea is very similar to the qualitative approach insofar as the household is asked to report the maximum amount that a lender is willing to lend, which is the credit limit of the respondent with regard to that particular lender. Thus, the authors define a household being credit constrained if "the optimal amount borrowed when borrowing under a credit constraint is strictly less than the optimal amount that would be borrowed if the credit constraint did not exist" [Diagne et al., 2000]. In other words, the household

[4]Appendix A contains a description of the credit programmes.

[5]According to Hayashi (1987), quantity constraints on the amount of borrowing (credit rationing) are a type of liquidity constraints.

[6]It is the same data set used in chapter four. Thus, the reader is invited to consult the previous chapter for a description of the data set.

[7]This indicator looks at whether households would have liked to borrow more.

[8]As explained in the second chapter, other approaches look at: a) significant dependence of consumption on transitory income [see for example, Jappelli and Pagano, 1989; Hayashi, 1987; Zeldes, 1989]; and b) qualitative questions on whether the borrower would have liked to borrow more [Feder et al., 1990; Jappelli, 1990].

is credit constrained if the optimal loan size is strictly less than the credit limit.

Diagne (1999) used the credit limit in Malawi as the dependent variable in a reduced-form recursive system of simultaneous equations aimed at capturing substitutability effects between formal and informal sources. A recent paper by Gross and Souleles (2002) used U.S. credit card panel data to see whether changes in liquidity have real effects (i.e. a direct test of liquidity constraints[9]). They included a very rich set of control variables and instrumental variables to overcome the endogeneity problem caused by the fact that banks might increase credit supply when credit demand is expected to rise. They showed that increases in credit limits generate a significant rise in debt.

We provide evidence of spillover effects and liquidity constraints in the formal and informal credit markets in Malawi by making the following contributions to the literature. First, whilst previous studies in Malawi [e.g. Nankumba, 1980; Reeser et al., 1989] and in other developing economies [e.g. Bose, 1998; Pal, 2002; Ravi, 2003] adopted a reduced form specification in which variables that affect the demand for credit by different households and the supply of credit by various institutions are collapsed into a single equation, we have been able to identify both demand and supply equations. The very rich data set we use allows for the identification of the demand equation for both applicants and non-applicants; and the supply equation (the credit limit equation) of formal and informal lenders. The demand and supply equations are identified by restrictions whereby the demand equation is set to depend on costs of loans, households' demographics and assets; whereby the supply equation is a function of village characteristics affecting all lenders, controls for competition and seasonality of loans.

Both the approach and the methodology adopted in this chapter differ from those of

[9]They answered the following question: "*When someone's credit limit (credit line) increases, what fraction of that extra liquidity does she uses to borrow and spend?*" According to the Permanent Income Hypothesis the answer should be zero because there are no liquidity constraints.

Diagne (1999) and Diagne et al. (2000) insofar as the credit limits supplied by one or more credit programmes are explicitly differentiated[10]. The fact that households may tailor their demand to the expected credit limit raises issues on the exogeneity of the credit limit itself - an identification problem that we explicitly take into account. In addition, several robustness checks are performed by addressing specification issues that may seriously affect the results. Heteroskedasticity, non-normality and selection problems can affect credit demand making conventionally used censored regression models inconsistent.

The outline of the chapter is as follows. Section two describes a model of spillover in presence of a rationed institutional credit market. Section three defines the credit limit variable in the Malawi survey. This variable is then used to identify constrained households in the descriptive statistics. In section four we explain the empirical strategy. Section five concludes.

5.2 A model of spillover

The aim of this section is to provide a theoretical model that guides the specification of the econometric estimation. This model is an intuitive and graphical application of Bell et al.'s (1997) model of spillovers in the credit market.

Suppose that a household produces farm output by using its endowments of land, $\bar{A}$, and labour, $\bar{L}$, which it supplies inelastically. The prices for labour (L), inputs (N) and formal credit are given by (w, p, r_f), respectively. The household's production function is given by $\epsilon F(L, N; \bar{A})$. It displays the usual properties: increasing, strictly concave and differentiable function. The output is affected by an i.i.d. (random) shock, ϵ, that

[10] The rationale of this approach lies in the justification of the results in the fourth chapter, where participation in the MRFC programme crowded out informal loans. Instead, Diagne (1999) aggregates all formal sources and separately provides control dummies for credit programmes.

represents the state of nature and is revealed only after the allocation of inputs. In addition to farm income, the household has a riskless income from other sources, $\bar{Y}$. So, the household's endowment vector is $Z \equiv \left(\bar{A}, \bar{L}, \bar{Y}\right)$.

In the presence of a loan, Q, the liquidity constraint on the purchase of variable inputs is given by:

$$wL + pN = w\bar{L} + Q \equiv K \tag{5.1}$$

and the household consumes:

$$Y_D \equiv \epsilon F(L, N; \bar{A}) + \bar{Y} \tag{5.2}$$

Suppose that household's preferences over lotteries are represented by a von Neumann-Morgenstern utility function $U(Y)$, and suppose that the present discounted values of the household's expected utility at the optimum is given by V_D. Because the household takes the prices of inputs and labour as given, we can replace F in equation 5.2 with $G(K; \bar{A}) \equiv \max_{L,N} F$ subject to equation 5.1. Given K, the optimal amounts of L and N are functions of K, w, p and $\bar{A}$.

For simplicity, Bell et al. (1997) assumed that the only collateral demanded by formal or informal lenders is the crop itself. While the informal lender can seize the entire crop whenever the loan is not repaid, the formal lender cannot do the same. Only informal lenders are assumed to exercise debt seniority. However, all lenders can deny future loans should the household default. In other words, loans supplied by formal lenders are riskier because these lenders cannot appropriate the crop.

If the household decides to repay the loan at the end of the season it remains eligible for a new loan in the next period. Bell et al. (1997) pointed out that in a stationary equilibrium, this requires choosing a minimum level of income under which default will occur. In other words, the household's choice variable determines the state of nature, $\epsilon_f \geq 0$. Thus, consumption can be written as follows:

$$Y = Y_D \qquad \text{if } \epsilon \leq \epsilon_f; \qquad Y = Y_D - (1 + r_f)Q_f \qquad \text{otherwise} \tag{5.3}$$

The probability of default is given by:

$$\Delta_f \equiv \Pr(\epsilon \leq \epsilon_f) \tag{5.4}$$

The value of the life-time expected utility can be written as follows:

$$V_f = EU(Y) + \Delta_f \delta V_D + (1 - \Delta_f)\,\delta V_f \tag{5.5}$$

or

$$V_f = \frac{[EU(Y) + \Delta_f \delta V_D]}{[1 - (1 - \Delta_f)\,\delta]}$$

where δ is the household's discount factor.

The household's problem involves the choice of a set (Q_f, ϵ_f) to maximize V_f subject to equations 5.1 and 5.4. Bell et al. (1997) showed that V_f is concave and monotonically increasing in Q_f.

In order to avoid infinite borrowing and to mitigate informational problems, formal lenders impose a ceiling on the supply of credit. Hence, the household will compare the life-time expected utility from borrowing up to the limit and choosing the associated value ϵ_f with the utility obtained from choosing an interior solution $\left(Q_f^0, \epsilon_f^0\right)$. In this case, the associated demand for formal credit is given by:

$$D_f^0 = D_f^0\left(w, p, r_f; \bar{Z}\right) \equiv Q_f^0$$

As mentioned above, the informal lender, in addition to refusing the supply of any future loan, can take the entire crop should the household be unable to repay with the farm income. The household cannot default and it will certainly choose to repay if the value of the crop is higher than the amount due to the lender. In this case, household's consumption can be written as:

$$Y = \bar{Y} \text{ if } \epsilon G(.) \leq (1 + r_i)Q_i; \qquad \text{and } Y = Y_D - (1 + r_i)Q_i \text{ otherwise} \tag{5.3'}$$

The probability of default can be defined as follows:

$$\Delta_i \equiv \text{ Pr}\,(\epsilon \leq \epsilon_i) \tag{5.4'}$$

where $\epsilon_i \equiv \frac{(1+r_i)Q_i}{G(w\bar{L}+Q_i;\bar{A})}$. As noted by Bell et al. (1997), because of the different lenders' policies regarding collateral it is evident from a comparison of equations 5.3 and 5.4 with equations 5.3' and 5.4' that even if $(Q_f, r_f) = (Q_i, r_i)$, the household will obtain different levels of life-time expected utility. Also, Bell et al. (1997) showed that because the maps of iso-V_f and iso-V_i contours are not the same, the household's optimal choice of credit at the same interest rate will be different (i.e. $D^0_f(r;.) \neq D^0_i(r;.)$).

When the household borrows from both sources, there are three possible cases: both loans are repaid; the formal loan is defaulted; and both loans are defaulted[11]. In this case:

$$Y_{fi} \equiv \epsilon G\left(w\bar{L} + Q_f + Q_i; \bar{A}\right) - (1 + r_i)\,(Q_f + Q_i) + (r_i - r_f)\,Q_f + \bar{Y} \tag{5.6}$$

And household's consumption is given by:

$$Y = \bar{Y} \text{ if } \epsilon_i > \epsilon; \qquad Y = Y_{fi} + (1+r_f)Q_f \text{ if } \epsilon_f > \epsilon \geq \epsilon_i; \qquad Y = Y_{fi} \text{ otherwise} \tag{5.3''}$$

[11]Households cannot default only on the informal sector because informal lenders are assumed to be able to seize the entire crop.

Turning to the supply side, the formal lender can set a ceiling on the credit offered to the household. As mentioned above, the credit limit reflects the availability of funds, informational problems and the necessity to avoid unlimited borrowing from the household. Let the credit limit set by the formal lenders be R_f, then the household's opportunity set is given by:

$$S_f \equiv \{(Q_f, r) : 0 \leq Q_f \leq R_f, r = r_f\} \tag{5.7}$$

Figure 5.1 shows two cases: in the first panel, the household realises its notional demand for credit at the formal interest rate; in the second and third panel, the household is rationed in the formal demand for credit.

It is assumed that informal lenders form a system of monopolies with exclusive contracts with each household. However, the large number of informal lenders ensures a relatively high degree of competition. Also, informal lenders are assumed to be risk-neutral and their opportunity cost of funds is constant at r_0 with $r_0 > r_f$. The informal lender has debt seniority over the formal lender and its expected profit from a loan Q_i at the interest rate r_i is:

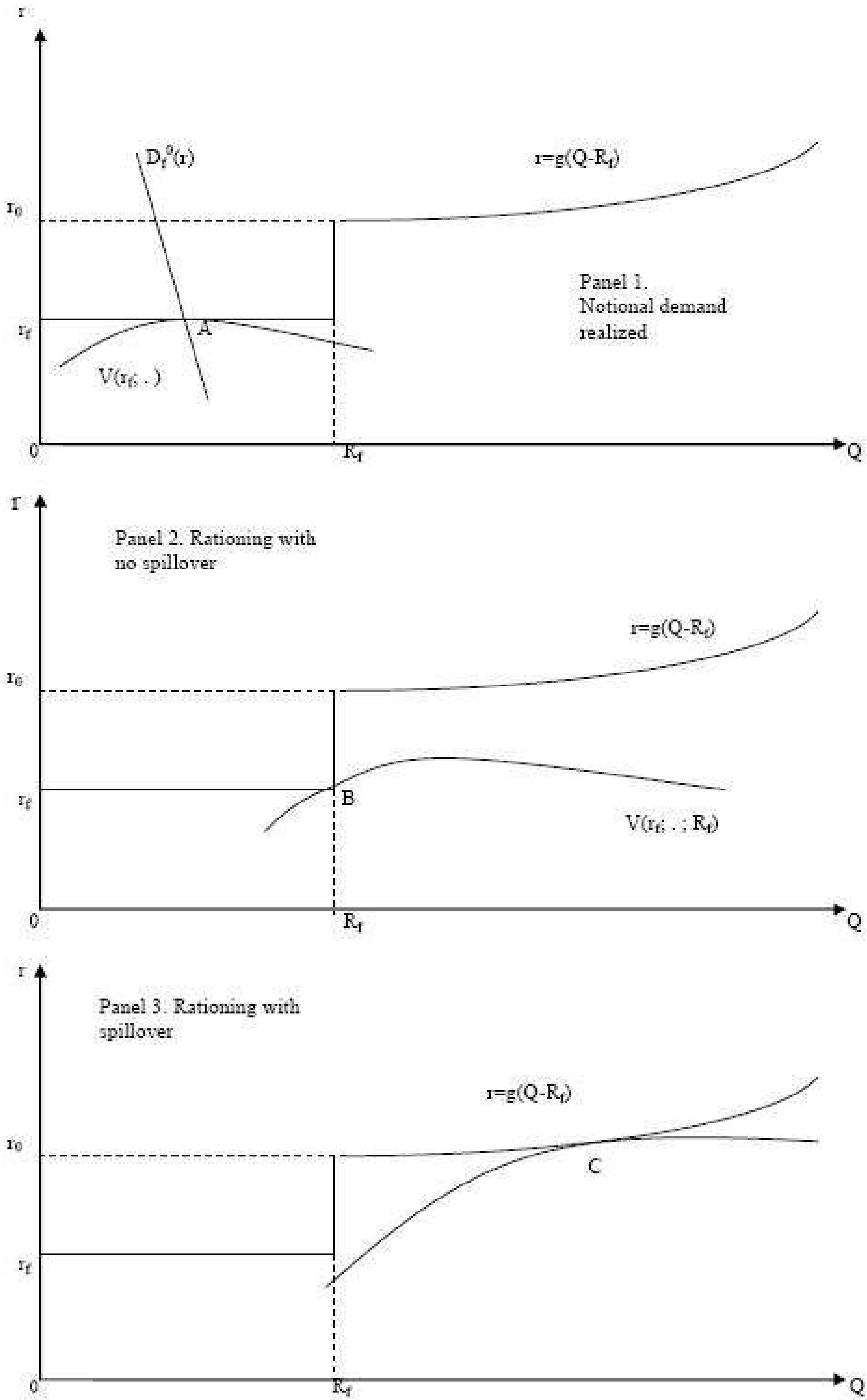
$$E\pi = \int_0^{\hat{\epsilon}_i} \left[\epsilon G\left(K; \bar{A}\right) - (1 + r_i)Q_i\right] h(\epsilon)d\epsilon + (r_i - r_0)Q_i \tag{5.8}$$

Free market entry implies that the expected profit is equal to zero. Let the zero expected profit contour in (Q_i, r_i)-space be given by:

$$r_i = g(Q_i; x) \tag{5.9}$$

where x denotes the household's characteristics that the lender can observe and use to design a credit contract.

FIGURE 5.1: The borrower's credit demand



Source: Bell, Srinivasan, Udry, 1997.

As pointed out by Bell et al. (1997), since the integrand in equation 5.8 is negative in the interval $[0, \epsilon_i]$, then $r_i > r_0$ for all $Q_i > 0$. As Q_i becomes very small, ϵ_i decreases as well, so that the informal lender breaks even on a very small loan at a rate just above its opportunity costs of funds, r_0.

Equations 5.7 and 5.9 give the boundary of the household's opportunity set in the space of $Q \equiv (Q_f + Q_i)$, the total amount borrowed from the formal and informal sectors, respectively, and the interest rate, r. Bell et al. (1997) defined this schedule as a reverse L-shaped offer curve from the formal sector and a zero-expected profit contour of the informal lender. In figure 5.1, $g(Q_i; x)$ has origin in $(R_f, 0)$ to denote the fact that the household seeks credit from the formal lender first.

Turning to the demand side, Bell et al. (1997) pointed out two cases. Firstly, if the formal credit sector imposes a credit limit (i.e. $R_f > 0$), then in the region bounded to the left by the vertical line through $(R_f, 0)$ and below by $g(Q_i; x)$, the life-time expected utility contours will differ from the pure cases in which the household approaches one of the two sectors. Secondly, if the optimal amount of credit from the informal sector is $Q_i^0 > 0$, then the upwardly sloped $g(Q_i; x)$ curve implies that Q_i^0 is less than the household's notional demand for informal credit at the rate $r_i = g(Q_i; x)$. This is the spillover case shown in panel 3 of figure 5.1.

The implications of Bell et al. (1997) model are as follows. First, the demand equations for formal and informal credit will have different parameter values for the same regressors. Second, the demand for formal (informal) credit should include the credit limit for informal (formal) credit which can be interpreted as a supply function of credit[12]. Third, if the spillover hypothesis is true, then any change in the informal (formal) credit limit should have a significant impact on the demand for formal (informal)

[12]However, it is not a proper supply function because the lenders are not price-takers in their respective segments of the market.

credit. Thus, we have testable implications for the econometric equations regarding the demand for credit and the credit limits.

5.3 Descriptive statistics

5.3.1 The credit limit variable in the Malawi rural FMHFS

The Malawi Rural Financial Markets and Household Food Security survey[13] (MRFMH FS) contains information about the expected credit limit a borrower faces. The survey asked: "How much do you think you could possibly borrow from this lender at a time?"

The maximum amount a credit applicant can borrow is a function of both borrower's and lender's characteristics. For example, an informal lender with restricted availability of funds could be able to lend less than a formal lender. However, the credit limit depends also on factors outside the control of both borrower and lender. The occurrence of aggregate negative shocks may reduce the availability of loans while increasing the demand for credit. Hence, the credit limit is a random variable, the realised value of which cannot be precisely known by either the borrower or the lender.

The applicant's demand for credit depends on his expectation about the maximum amount he or she can borrow. At the time of borrowing, only the lender knows the actual value of the credit limit. For instance, if the demand exceeds the supply, the borrower does not get the chance to know the actual credit limit set by the lender. However, as argued by Diagne (1999), it does not matter whether the expected credit limit coincides with its actual value because ultimately it is the expectation about the maximum amount the applicant can borrow that affects his behaviour.

How then does the borrower form his expectations? In a repeated interactions envi-

[13]The data set is described in section 4.2.1 of chapter four.

ronment where lenders and borrowers in the same village are likely to set up different contracts, the borrower can form his expectation by learning about his realised credit limit from previous loans obtained by that credit source at a particular time. Even if the borrower does not ask for a loan, the expectation about the credit limit can be formed by looking at other borrowers[14]. Also, some government and NGO supported credit programmes set a fixed credit limit which is known to all applicants. Hence, in order to justify an econometric analysis based on the expected credit limit we have to assume that the borrower has accurate information to predict the maximum amount he is able to borrow. Diagne (1999) showed that marginal effects can be obtained even when the realized value of the credit limit is not observed[15].

5.3.2 The characteristics of borrowers

In addition to the credit limit, households provide information about their demand for credit. The question asked in the survey for each credit source (whether it be formal or informal) is[16]: "How much (credit) did you ask for?" This function shows the amount the borrower is *willing* to borrow, but not the amount she is *able* to borrow. Hence, the extent of the credit constraint is determined by the difference between what the borrower is willing and able to borrow.

The amount the borrower effectively borrows is a function of the credit limit, the interest rate and the amount asked for.

Credit limit, credit demand and amount borrowed can be used to identify four types of households as in table 5.1: households with no access; rejected; rationed; and borrowers. We outline two credit sources for each type of household: credit programmes[17] and

[14]It is very likely that in the same village information about credit is shared among borrowers.
[15]Proof in Appendix B.
[16]The questionnaire does not specify whether the demand is at any interest rate or not.
[17]See appendix A for a description of the credit programmes in Malawi.

informal lenders.

Households with no access to credit have a zero demand. This group may include "discouraged borrowers" as defined by Feder et al. (1990), Jappelli (1990) and Zeller (1994)[18]. The second type of households consists of rejected applicants. Note, while rejected households have applied for a loan but have been turned down, households with no access have not applied for a loan[19]. The third group of households has been rationed by the lenders. Their demand for credit exceeds the credit limit. Finally, successful applicants are those who actually received a loan.

TABLE 5.1: Constrained households by source of credit

Type of households:	**Credit programmes**	**Informal**
Households with no access	39.7 (481)	21.4 (259)
Rejected households	4.2 (51)	6.2 (75)
Rationed households	12.3 (149)	4.6 (56)
Households with successful applications (debt | debt>0)	28.8 (349)	15.8 (191)

Source: own calculation from MRFMHFS. Note: Percentage displayed and number of households in parenthesis. Percentage of the sampled households over the three rounds (N=1,212).

Approximately 40 percent of households have not applied to at least one of the credit programmes - the highest group of constrained households in table 5.1. Also, 21 percent of households have not applied to informal lenders. A higher proportion of households have been rationed by credit programmes compared to informal lenders (12.3 percent against 4.6 percent). On the other hand, more applicants are rejected from informal lenders. This can be explained by the fact that group-lending programmes target poorer

[18] Jappelli (1990) states "If there is a cost of apply, consumers with high probability of loan denials may not apply because they perceive that, if they do, they will be refused loans. We refer to these consumers as discouraged borrowers." In our case households have zero demand because their expected limit was zero or because they had other reasons for not applying.

[19] Figure 5.2 displays the reasons for not applying to formal loans (including credit programmes and other formal sources).

TABLE 5.2: Selected characteristics of constrained households

Type of households:	*Credit programmes*			*Informal credit*		
	No access households	**Rejected**	**Rationed households**	**No access**	**Rejected households**	**Rationed households**
Household size	5.7 (2.4)	6.0 (2.6)	6.0 (2.5)	5.8 (2.3)	6.0 (2.5)	6.1 (2.5)
Female head (%)	27.4 (44.7)	15.7 (36.7)	20.1 (40.2)	25.1 (43.4)	28.0 (45.2)	26.8 (44.7)
Age household head	45.0 (13.0)	45.9 (13.0)	44.5 (11.3)	46.1 (12.8)	45.8 (13.5)	44.4 (13.9)
Number of children 0-15	3.0 (1.8)	2.9 (1.8)	3.1 (1.8)	3.1 (1.7)	3.2 (1.8)	3.2 (1.6)
Head with primary school (%)	78.0 (41.5)	80.4 (40.0)	75.2 (43.3)	76.1 (42.8)	68.0 (47.0)	75.0 (43.7)
N. of observations:	*481*	*51*	*149*	*259*	*75*	*56*
Land size (ha)	2.1 (2.1)	3.07 (4.0)	2.0 (2.1)	2.1 (1.3)	2.0 (1.6)	1.9 (1.1)
Share of land owned by spouse (%)	17.3 (35.2)	6.6 (19.4)	15.9 (33.0)	19.1 (36.5)	12.8 (29.8)	10.4 (27.0)
Share of land in total assets (%)	61.8 (24.1)	57.0 (24.0)	59.0 (23.2)	58.8 (22.1)	61.5 (22.7)	60.2 (22.9)
N. of observations:	*478*	*50*	*149*	*258*	*75*	*56*
Food expenditure†	13.5 (13.6)	15.0 (14.1)	13.4 (14.7)	13.6 (14.4)	11.4 (10.0)	12.6 (10.9)
N. of obs.	*440*	*48*	*143*	*253*	*72*	*52*
Non-food expenditure†	105.7 (151.8)	144.7 (187.3)	126.9 (188.4)	97.7 (128.8)	116.5 (171.6)	90.3 (143.6)
N. of obs.	*451*	*51*	*149*	*259*	*75*	*56*

Source: Own calculation based on MRFMHFS. Note:†in local currency, 15 Malawian Kwachas (MK)=1 US$. Malawi's per capita GNP is US$ 170 (approx. 2,550. World Bank, 1997). Standard deviation in brackets. Expenditure deflated by the square root of households' size.

households who may not be considered safe borrowers by informal lenders. Almost 29 percent of sampled households borrowed from at least one of the credit programmes and approximately 16 percent borrowed from informal lenders.

Table 5.2 displays the mean and standard deviation of some selected characteristics of constrained households. The three groups of households with zero demand, rejected and rationed households do not differ from each other in terms of demographic characteristics. Households who have been rejected or have not applied to credit programmes are more likely to have a household head holding primary school education. On the other hand, the three groups differ in terms of wealth and especially with regard to asset holdings. As expected, households rejected by credit programmes own larger plots. This is due to the fact that the government-sponsored programmes target small farmers. Informal lenders reject or discourage households with a larger share of land owned by the spouse. Households who have been rejected by credit programmes spend more on food and non-food items.

TABLE 5.3: Credit limit by quintiles of land

Land quintiles:	**Credit programmes**	**Informal**	*N. of obs.*[1]	
poorest 5^{th}	493 (770)	328 (1560)	82	
2^{nd} **poorest** 5^{th}	491 (984)	101 (254)	100	
middle 5^{th}	586 (1008)	87 (201)	97	100
2^{nd} **richest** 5^{th}	682 (1065)	82 (193)	78	82
richest 5^{th}	1005 (1785)	81 (177)	119	120

Source: Own calculation based on MRFMHFS. Note: all values in local currency, 15 Malawian Kwachas (MK)=1 US$. Malawi's per capita GNP is US$ 170 (approx. 2,550. World Bank, 1997).[1]The two columns of observations correspond to the following groups: credit programmes and informal sources. Std. deviation in parentheses.

Table 5.3 reports the average credit limit by quintiles of land. It is evident that the formal credit limit is, on average, an increasing function of land size. That is, the credit limit of credit programmes is larger for the households highest quintiles of land[20]. On the other hand, table 5.3 shows that the credit limit supplied by informal lenders does not increase with the size of land. Probably other factors such as kinship or interlinkages affect the size of the credit limit in the informal credit market.

5.3.3 The behaviour of borrowers

Table 5.4 displays the mean and standard deviation of the amount borrowed, of the demand and the credit limit for each credit source in the five districts[21]. In the fourth chapter we have described different sources of credit: formal lenders (i.e. group-lending programmes, the Commercial Bank of Malawi and World Vision); and informal lenders (i.e. friends and relatives and other informal lenders such as moneylenders and traders).

The average credit limit shown in table 5.4 can be exceeded by (or can exceed) the average demand. This can happen either because the distributions are skewed differently or because households are constrained. Indeed, table 5.1 has shown that some households have been rationed by lenders.

Table 5.4 shows that, on average, households in Nkhotakota borrow more from formal lenders (with the exception of the credit programmes MRFC and MUSCCO[22]). By contrast, households in Dedza on average borrow more from the MRFC programme and from friends and relatives. The demand for credit[23] to the MRFC is higher in Mangochi, but on average households in Nkhotakota demand more credit to the other formal credit sources (excluding the demand to MUSCCO which is higher in Dowa). Moreover, the credit limit is higher in Nkhotakota than in other districts with the exception of the

[20]Recall that the credit limit refers to the applicant's expectation of the amount the lender will be able/willing to lend. It refers to borrowers and non successful applicants.

[21]Appendix A contains a map of the districts in Malawi.

[22]Credit programmes are described in appendix A.

[23]As explained is sub-section 5.3.2 the demand for credit determines the amount a household is willing to borrow given its expectation on the credit limit.

Table 5.4: Households' indebtedness by district and source of credit

Districts:	Dowa			Mangochi			Nkhotakota			Rumphi			Dedza		
	Borrow	**Demand**	**Limit**	**Borrow**	**Demand**	**Limit**	**Borrow**	**Demand**	**Limit**	**Borrow**	**Demand**	**Limit**	**Borrow**	**Demand**	**Limit**
MRFC	161	229	161	146	670	224	471	521	549	31	29	35	701	137	1035
	(37)	(86)	(37)	(53)	(425)	(102)	(94)	(143)	(122)	(26)	(30)	(29)	(171)	(95)	(384)
MMF	0	0	0	417	629	419	0	0	0	0	0	0	0	0	0
				(72)	(110)	(73)									
MUSCCO	218	230	270	0	0	0	7	21	7	0	0	0	0	0	0
	(105)	(113)	(143)				(7)	(22)	(7)						
PMERW 1	0	0	0	31	140	31	161	157	202	14	18	14	0	0	0
				(19)	(109)	(19)	(69)	(76)	(89)	(12)	(22)	(12)			
PMERW 2	0	0	0	3	7	3	138	151	138	0	0	0	0	0	0
				(3)	(7)	(3)	(76)	(155)	(76)						
CBM	0	0	0	0	0	0	769	576	961	0	0	0	0	0	0
							(554)	(584)	(705)						
World Vision	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Friends & relatives	15	4	24	57	5	149	24	15	40	48	93	48	94	122	94
	(10)	(4)	(14)	(38)	(5)	(112)	(13)	(11)	(18)	(7)	(7)	(7)	(26)	(35)	(26)
Other informal lenders	0	0	0	13	4	16	6	0	19	0	0	0	0	0	0
				(9)	(4)	(10)	(5)		(16)						

Source: Own calculation based on MRFMHFS. Note: all values in local currency, 15 Malawian Kwachas (MK)=1 US$. Malawi's per capita GNP is US$ 170 (approx. 2,550. World Bank, 1997). Loan values include cash and in-kind amount. In-kind values have been provided by the respondent. Standard deviation in brackets. CBM=Central Bank of Malawi.

TABLE 5.5: Rejected applicants' characteristics by district and source of credit

Districts:	**Dowa**		**Mangochi**		**Nkhotakota**		**Rumphi**		**Dedza**	
	Demand	**Limit**	**Demand**	**Limit**	**Demand**	**Limit**	**Demand**	**Limit**	**Demand**	**Limit**
MRFC	40 (25)	0	-	-	-	-	270 (75)	0	69 (38)	20 (17)
MMF	-	-	387 (231)	0	-	-	-	-	-	-
MUSCCO	204 (61)	7 (7)	-	-	-	-	0.3 (0.3)	0	-	-
PMERW 1	-	-	-	-	25 (25)	0	7 (7)	0	-	-
PMERW 2	-	-	-	-	12 (13)	0	-	-	-	-
NABW	-	-	57 (55)	19 (18)	178 (182)	0	-	-	-	-
World Vision	-	-	65 (64)	16 (16)	971 (778)	49 (53)	-	-	-	-
Other NGO	-	-	-	-	191 (203)	0	1242 (1303)	0	7 (8)	0
Other government prog.	-	-	23 (25)	17 (19)	78 (84)	0	4969 (5213)	0	83 (75)	0
Friends & relatives	104 (40)	32 (22)	60 (43)	21 (16)	118 (41)	25 (11)	28 (15)	1 (1)	13 (7)	1 (1)
Other informal lenders	-	-	5 (5)	0.1 (0.2)	33 (25)	4 (4)	43 (37)	0	42 (35)	0.5 (0.5)

Source: Own calculation based on MRFMHFS. Note: all values in local currency, 15 Malawian Kwachas (MK)=1 US$. Malawi's per capita GNP is US$ 170 (approx. 2,550. World Bank, 1997). Loan values include cash and in-kind amount. In-kind values have been provided by the respondent.

limit set by the MRFC, MUSCCO and by informal lenders.

Table 5.5 shows the demand and the credit limit of rejected applicants for each credit source and district. Households who have been turned down by lenders provide the amount of credit they asked for and the maximum credit (i.e. the credit limit) they expected to face were they involved in a credit transaction. The questions asked in the survey are: "What is the maximum amount you could possibly borrow from a lender if you really wanted?" and "How much did you ask as loan amount?"

It is evident that there are two types of rejected applicants. There are households who could not borrow because their demand had been completely rejected (i.e. lenders would not offer them any amount of credit). For instance, the MRFC in Rumphi does not provide credit despite the fact that on average rejected households asked for 270 MK. In addition, there are cases in which the average demand is higher than the average credit limit. As mentioned previously, this can be explained either by the fact that the distributions may be differently skewed or by the fact that some households demand more than the limit.

To sum up, demand, amount borrowed and credit limit vary across districts and credit sources. In particular, Nkhotakota is the district where households on average borrow more from most of the credit programmes and where the demand and credit limit are higher on average than in other districts. There is also evidence that rejected applicants are those who have zero credit limit (i.e. lenders would not supply any loan).

5.3.4 Borrowing costs

Table 5.6 reports two main types of borrowing costs: the monthly interest rate and other costs such as travel costs and fees, low or no wage. As informal credit contracts are usually interlinked with other contracts (i.e. landlord-tenant), the landlord/lender may not charge an interest rate but may decrease or cancel the wage of the borrower/tenant for the delivery of the loan. Hence, the interest rate alone cannot truly reflect the cost of the credit contract. However, because the data set only provides the above mentioned aggregate variable "other costs" we cannot separately identify each non-interest charge.

TABLE 5.6: Costs of borrowing

Type of costs:	**Mean**	**Min.**	**Max.**	**St. Dev.**
Interest rates from:				
MRFC programme	3.77 (77)	0.04	6.21	1.40
More than one programme	3.02 (196)	0.42	38.04	3.64
Informal lenders	19.61 (33)	1.79	55.02	17.68
Loan costs for:				
MRFC programme	0.21 (82)	0	10	1.3
More than one programme	0.10 (223)	0	20	1.4
Informal	0.33 (450)	0	64	1.3
Percentage of households with interest-free formal loans	12.9 (50)			
Percentage of households with interest-free informal loans	98.1 (387)			
Percentage of households with zero costs on formal loans	97.6 (439)			
Percentage of households with zero costs on informal loans	98.7 (444)			

Source: own calculation from MRFMHFS. Loan costs include: travel, fees, no or low wage etc.% monthly interest rate. Number of respondents in brackets. Note: all values in local currency, 15 Malawian Kwachas (MK)=1 US$. Malawi's per capita GNP is US$ 170 (approx.2,550 MK). World Bank, 1997.

As households are unwilling or unable to provide the interest rate for each credit source, the Malawi FMHFS does not report the value of the interest rate. However, the survey asks for the amount borrowed and the amount to be repaid together with the date by which the loan has to be given back and the date of receipt. By using this information we have calculated the monthly interest rate[24].

[24]The interest rate $i = \frac{(\text{amount to be repaid-amount borrowed})}{(\text{amount borrowed})}$. The timing is calculated as the difference between due date and date of loan receipt converted into months (about 30.4375 days). Then,

The informal interest rate is determined by friends, relatives and other informal lenders. We also distinguish between the interest rate charged to participants in the MRFC only and to participants in more than one credit programme. In table 5.6 we report the interest charges for those who actually had to pay interest on the loan. We find that informal lenders charge higher interest rates than either the MRFC programme or the other formal credit programmes[25].

It is also evident that 98.1 percent of households face interest-free informal loans while only 13 percent of households have a zero interest rate on their formal loan.

As expected, friends and relatives, who dominate the credit sector, lend without requiring any interest. Despite the large percentage of interest-free informal loans, interest rates do vary from 0.04 to 38 percent in the credit programmes and from almost 2 to 55 percent in the informal sector (presumably due to the fact that not all informal loans are provided by friends and relatives.). High variability in the interest rates is very common in the credit markets of developing countries [Banerjee and Duflo, 2001; Fafchamps, 2000].

Loan costs show extreme variability as well. The majority of loans have virtually zero costs. On average, informal loans carry higher additional costs compared to formal loans (provided either by the MRFC only or by more than one programme). Participants in the MRFC only face slightly higher costs than participants in more than one programme.

In addition to economic costs, informal loans carry social costs that are not shown

the interest rate for the formal (informal) sector is an average of the interest rates offered by each source: banks, credit programmes and NGOs for the formal sources; friends, relatives and other informal lenders for the informal sources.

[25] However, whenever we do include the borrowers who obtained interest-free loans from informal lenders as in table C5-1 in appendix C, we find that the interest rate charged by the MRFC is lower than the informal one, but the interest rate faced by participants in more than one credit programme is higher than the informal interest rate. The interest rate differentials across sources is affected by the characteristics of the borrowers and by the design of the credit contracts. Note also that in the following empirical analysis we aggregate interest rates and costs across different formal sources to keep a higher number of observations.

in the survey. The fact that informal networks create an "obligation" to reciprocate is recognised both by economists [for instance, Platteau and Abraham, 1987; Sahlins, 1972; Udry, 1990] and by anthropologists [for example, Levi-Strauss, 1949; Mauss, 1925]. We can conclude that informal loans are the most expensive available credit source, but they also display higher variability.

Table 5.7 correlates the interest rates to land quintiles. There is no evidence of any particular relationship between the two variables especially for participants in more than one programme and for borrowers from informal lenders. The interest rate charged to participants in the MRFC only, by contrast, increases with land size and is larger for the three highest households quintiles of land.

Table 5.7: Interest rate by quintiles of land

Land quintiles:	**MRFC only**	**More prog.**	**Informal**	*N. of obs.*		
poorest 5th	0.2 (0.7)	2.5 (1.6)	1.9 (8.3)	75	43	75
2nd poorest 5th	0.6 (1.6)	3.2 (5.6)	1.5 (6.4)	96	45	96
middle 5th	0.8 (1.6)	2.6 (2.9)	1.7 (7.5)	91	35	91
2nd richest 5th	0.8 (1.7)	3.0 (4.3)	1.5 (7.8)	74	33	74
richest 5th	0.8 (1.7)	2.7 (2.5)	0.8 (5.4)	113	53	113

Source: Own calculation based on MRFMHFS. Note: the formal interest rate is disaggregated according to participation in the MRFC or more than one credit programme. The first and second (third) column of N. Obs. refers to the MRFC and more than one credit (informal) source. Std. deviation in parentheses.

In figure C5-1 in appendix C we plot credit limit (in logarithm) and interest rate for the formal and informal credit sectors[26]. We find that the informal sector is mainly characterised by interest free loans. Few cases show that a higher credit limit is correlated with a lower interest rate. This relationship is more apparent in the formal sector where a higher credit limit is associated with a lower interest rate. Indeed, both credit limit

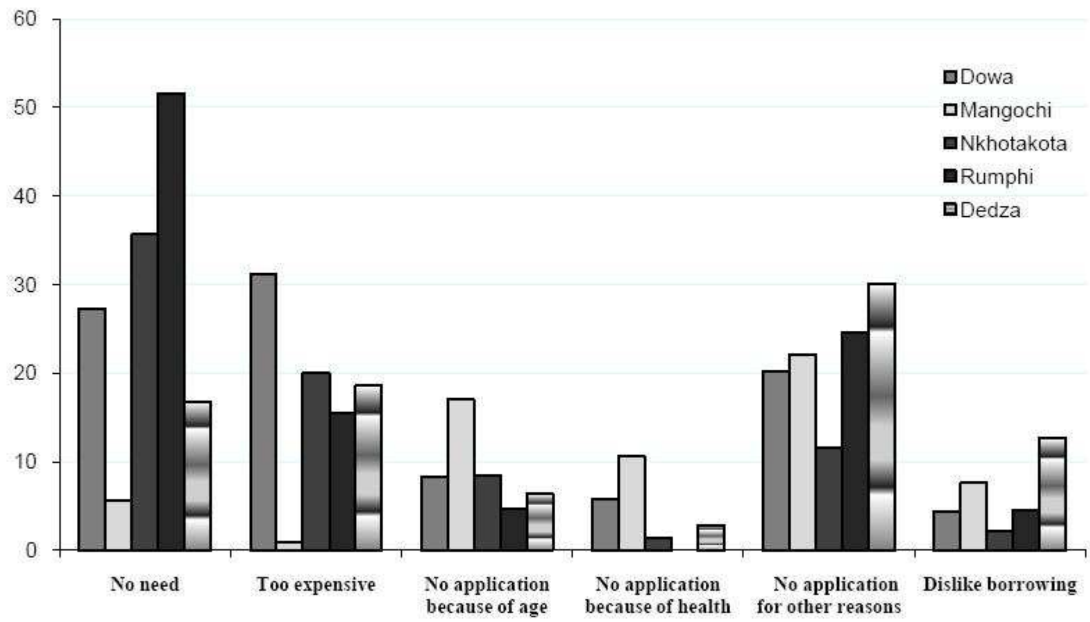
[26]We have aggregated formal sources such as credit programmes, banks and NGOs to keep a higher number of observations.

and interest rate are linked with the riskiness of the borrower[27].

5.3.5 Why don't households borrow?

Despite the availability of formal credit, households borrow mainly from informal sources and in particular from friends and relatives[28]. Figure 5.2 displays the reasons why households do not borrow from formal lenders.

FIGURE 5.2: Reasons for not borrowing from formal sources by district



Source: Own calculation based on MRFMHFS.

In Nkhotakota and Rumphi households report that the major reason for not borrowing is that they have no need to do so. On the other hand, in Dowa the most important reason for not borrowing from formal sources is that the loans are too expensive.

There are several reasons for which households are discouraged from applying for formal credit. Almost ten percent of borrowers in all districts do not apply because of their age. Some households (from five to ten percent across all districts) report to

[27] Banks tend to apply a higher credit limit (i.e. less rationing) and a lower interest rate to allegedly safer borrowers.

[28] This has been discussed in the fourth chapter.

dislike borrowing. Less than ten percent of households across all districts do not apply for health problems.

5.4 Econometric analysis

The aim of this section is to test whether borrowers are credit constrained ("the liquidity constraints hypothesis") and whether they respond to credit constraints by substituting one form of credit for another ("the spillover hypothesis"). The Malawi FMHFS data set gives an opportunity to discuss this issue because it introduced new credit programmes which offer potential scope for substitution with informal sources.

As discussed in the descriptive analysis, we have several measures of "credit constraints". First, households report the maximum amount they think the lender would give them (i.e. the credit limit) and the credit demand for each source of credit (whether it be formal or informal). Second, the Malawi FMHFS provides information on the credit limit and credit demand of both borrowers and rejected applicants. Therefore, we can possibly identify two types of people: a) those who were rejected and answered as to what they asked for; and b) those who received what they asked for.

If households are liquidity constrained then an increase of the borrowing limit should affect credit demand. In addition, if the spillover hypothesis explains multiple borrowing, then any change in the credit limit of one sector should have an impact on the demand for credit from the other sector. As a result, the informal sector arises as a spillover from the rationed formal sector demand.

The central issue of the following econometric analysis is that households may have tailored their credit demand to what they expected they could receive, making hard the direct inference of who is "unconstrained". We address this technical problem in sub-section 5.4.1. In sub-section 5.4.2 several specification tests are performed. Finally, in subsections 5.4.3 and 5.4.4, we check the robustness of the results by adopting alternative specifications that overcome the limitations of the censored regression model adopted in sub-section 5.4.1.

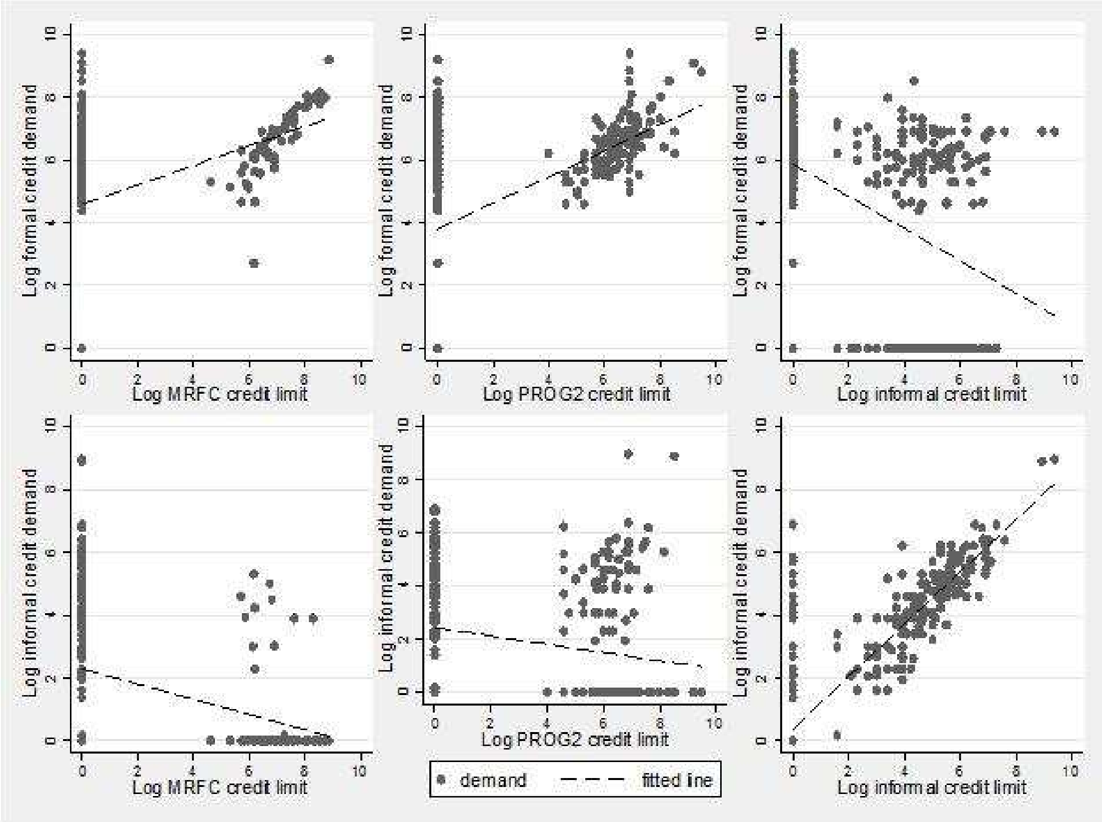
5.4.1 Estimation strategy

In order to investigate the relation between credit limit and demand, in figure 5.3 we plot one against the other for each credit source (i.e. informal and formal). As expected, we observe a certain degree of correlation between the two variables. In particular, in the first and second graph of the top panel there is a positive correlation between formal credit demand and formal credit limit (from the MRFC programme and the other credit programmes); in the last graph of the bottom panel there is also a positive correlation between informal credit demand and credit limit. The other graphs show no evidence of correlation between credit demand and credit limit. More important for our analysis is the fact that there are several data points in which a higher demand for credit is associated with a low credit limit, indicating that the latter might be exogenous. This confirms the descriptives in tables 5.4 and 5.5 where, on average, the demand for credit exceeded the credit limit.

Two issues should be pointed out. First, the scatter plots cannot give any information on the direction of causality. In other words, saying that credit limit and credit demand are correlated does not necessarily imply that the former exogenously causes the latter. Second, measurement error may also affect the credit limit biasing any analysis of its relation with the demand for credit.

In order to address these issues and, at the same time, test for the liquidity constraints

Figure 5.3: Scatter plot of credit demand versus credit limit by source



Source: Own calculation based on MRFMHFS.

and spillover hypotheses, we use a two-step tobit model[29] where the credit limit is endogenous to the demand. Let the demand for credit[30] of household i be denoted by $D^k_{i,t}$ with k=*formal*, *informal* and the household's credit limit (or credit line) be given by $L_{i,t}$. The formal sector consists of all the government sponsored credit programmes, the Commercial Bank of Malawi and World Vision; the informal sector includes friends and relatives and other informal sources such as moneylenders and traders.

Unlike Gross and Souleles (2002), we do not have a large panel and so the model is estimated by pooling across the rounds. More formally, let:

[29]It would have been more efficient to use a maximum likelihood estimation of the tobit, but the model turned out to be computationally heavy.

[30]The "demand" variable has been constructed as follows. For those who borrow it is equivalent to the amount of debt, whilst for those who have been rejected it is determined by their demand for credit. This approach aims at reducing sample selection bias. Later on, we check whether our main results change whenever debt is used as dependent variable.

$$D^{*}_{i,t} = \alpha_0 + X^{'}_{i,t}\beta_0 + L_{i,t}\beta_1 + \gamma^{'}_0 time_t + u_{1i,t} \qquad \text{with } u_i \sim N(0,\sigma^2) \tag{5.10}$$

The above equation could be estimated by pooled instrumental variables (with the endogenous credit limit) if all potential borrowers demanded credit.However, the demand is censored[31] to the left at zero[32].

Despite the fact that we have included the demand from rejected applicants in place of the missing values in $D^{k}_{i,t}$, the demand from discouraged borrowers is not observable and hence the variable is censored. Including the expected demand of rejected applicants in place of missing values has the advantage of increasing the number of observations, but could cause biases. Tables C5-2 and C5-3 in appendix C report the second and first stage regressions of the instrumental variable tobit model where the dependent variable is the amount of debt. The results show that there is still significant evidence of spillover effects and liquidity constraints.

Let the endogenous vector L_i be modelled as follows:

$$L_{i,t} = \alpha_1 + Z^{'}_{i,t}\delta + u_{2i,t}$$

where endogeneity arises because of the correlation between u_{1i} and u_{2i}. Decompose the error term as follows [Wooldridge, 2002]:

$$u_{1i,t} = \rho u_{2i,t} + \epsilon_{i,t} \tag{5.11}$$

Under the assumption that u_{1i} and u_{2i} are jointly normally distributed, u_{2i} and ϵ_i

[31]However, note that we cannot observe $D^{k*}_{i,t}$ but only:

$$D^{k}_{i} = \max\{0, D^{k*}_{i}\}$$

[32]The demand could also be censored to the right at the value of the limit. However, we have observed in figure 5.3 that the demand for credit exceeds the limit.

are uncorrelated by definition and ϵ_i is also normally distributed[33]. Replacing it in the model:

$$D^*_{i,t} = \alpha_0 + X^{'}_{i,t}\beta_0 + L_{i,t}\beta_1 + \gamma^{'}_0 time_t + \rho u_{2i,t} + \epsilon_{i,t}$$

$$L_{i,t} = \alpha_1 + Z^{'}_{i,t}\delta + u_{2i,t} \tag{5.12}$$

Smith and Blundell (1986) proposed a two-step procedure because u_{2i} is unobservable. This is the method we use to produce the results reported in tables 5.8 and 5.9. In the first step, we estimate δ by OLS and predict u_{2i}:

$$\widehat{u}_{2i,t} = L_{i,t} - Z^{'}_{i,t}\widehat{\delta} \tag{5.13}$$

In the second step, we use $\widehat{u}_{2i,t}$ in the model for $D^{k*}_{i,t}$ above and estimate by tobit. The condition for the identification of the above model is that factors affecting the demand equation do not enter the credit limit equation or *vice versa*.

Which exclusion restrictions should be made? Ideally, we would want factors affecting the supply of loans but not unobserved factors related to the demand. As we have no information on lenders (except for the number of moneylenders), we use the characteristics of the villages. However, because the characteristics of the villages affect the demand for loans too, we select variables that can be thought of as “natural experiments” such as the number of tube wells in the village, or months for the hungry season that are exogenous to unobserved factors affecting households’ demand for credit.

As in Diagne (1999), in the first regressions (table 5.9) we include the characteristics of the villages that affect all lenders such as the number of deep tube wells, the number of members in farm clubs, the distance of the village to the nearest commercial bank, the average price of maize in one of the hungry seasons (i.e. February) and the number

[33]We will test the normality assumption in the context of a simple tobit model later.

of households who own land greater than five acres.

Following Grant (2007), we have also included a seasonal dummy for the hungry season of February, the number of households in the village, and the number of moneylenders as measure of local competition.

Table 5.8 - namely, the second stage regressions - shows the results of two models where the dependent variable is the logarithm of the credit demand for informal and formal credit, respectively. As in Castronova and Hagstrom (2004) and Ibrahim et al. (2007), we look at the response of demand to the following groups of variables: a) households characteristics that include education of the household head and the occupation of the spouse in a small trade; b) proxies for current resources such as total value of assets, food and non-food expenditure, size of land and share of land in total assets, proportion of land owned by the spouse; and a proxy for vulnerability (i.e. number of negative shocks); c) prices indicated by the interest rate and other costs of formal and informal loans; d) a regional dummy indicating whether the household is located in the South and the proportion of Christians in the village where household i lives. All these groups are included in vector X.

In addition, we use a partition of the vector $L_i = \left[L_i^I, L_i^{MRFC}, L_i^{PROG2}\right]$ which contains the credit limit faced by the households who borrow from the informal sector, from the MRFC only and from more than one credit programme[34]. Unlike Diagne (1999) we disaggregate the credit limit of different credit programmes in order to test the hypothesis outlined in chapter four, that is, households would decrease informal credit were they not rationed by the MRFC programme.

The coefficient β_1 could be interpreted as the fraction of a credit line that is borrowed or the marginal propensity to consume (MPC) out of liquidity for each credit source[35]

[34] Appendix A contains a description of the credit programmes.

[35] Of course this is only valid if we assume that the (self-reported) credit limit is truly exogenous. We will question this assumption later in this chapter. The use of "demand" in the way defined earlier may

[Campbell and Mankiw, 1990; Gross and Souleles, 2002]. As outlined in the introduction, a direct test of liquidity constraints looks at whether changes in liquidity (measured by the credit limit) have any effect on the demand of household i (dDemand/dLimit). According to the permanent income hypothesis, where there are no liquidity constraints, the predicted coefficient on dDemand/dLimit is zero.

5.4.1.1 Results

Table 5.8 reports the results of the instrumental variable regressions.

The most important findings entail the liquidity constraints and spillover hypotheses. Liquidity constraints occur whenever there is a significant impact of the credit limit on the credit demand. Spillovers exist because there are negative terms on the limits across loan sources.

There is strong evidence of the existence of liquidity constraints in the demand for formal and informal credit. Table 5.8 shows that in model I the effect of the informal credit limit is quite large and significant at one percent level. In particular, a ten percent increase in the informal credit line would increase the demand for informal loans by more than nine percent. This is evidence of a high degree of liquidity constraints.

Similarly, we find that even households who participate in credit programmes are constrained in their demand for formal credit. More specifically, from table 5.8 it is evident that an increase of ten percent in the MRFC credit line would increase the demand for formal credit by almost four percent. Also, households who participate in more than one credit programme are constrained. Indeed, a ten percent increase in the credit limit of more than one credit programme increases the demand for formal

---

cast doubt on the interpretation of the coefficients β_1 as evidence for liquidity constraints. However, we do think that replacing unobserved debt of non-borrowers with their demand is appropriate in resolving sample selection.

Table 5.8: Instrumental variables tobit - 2^{nd} stage regressions

	MODEL I: Log(informal credit)	**MODEL II: Log(formal credit)**
hh characteristics:		
hh size	-0.07	-0.03
	(0.14)	*(0.11)*
age head	0.01	0.01
	(0.01)	*(0.01)*
female head	0.10	-0.74
	(0.84)	*(0.74)*
n. children 0-15	0.20	0.18
	(0.17)	*(0.14)*
head primary education†	-0.38	-0.42
	(0.35)	*(0.27)*
spouse employed in small trade	-0.14	-0.02
	(0.45)	*(0.34)*
Assets, expenditure and shocks:		
land size (ha)	-0.03	-0.05
	(0.07)	*(0.07)*
land share owned by spouse (%)	0.004	-0.001
	(0.00)	*(0.00)*
land share in total assets (%)	0.02	-0.01
	*(0.01)**	*(0.01)*
value of assets (MK)	0.0001	0.00004
	(0.00)	*(0.00)*
food expenditure (MK)	0.01	0.0003
	*(0.01)**	*(0.00)*
non food expenditure (MK)	-0.0002	0.001
	(0.00)	*(0.00)****
number of negative shocks	-0.22	0.06
	(0.14)	*(0.10)*
Costs of loans:		
formal interest rate (%)	0.07	0.06
	(0.07)	*(0.05)*
informal interest rate (%)	0.02	0.05
	(0.02)	(0.02)
formal loan costs	0.01	-0.04
	(0.06)	*(0.06)*
informal loan costs	0.01	0.02
	(0.03)	*(0.03)*
% Christians in the same village	0.02	0.0002
	*(0.01)***	*(0.01)*
South†	0.39	0.35
	(0.36)	*(0.31)*
round 2†	-0.86	-0.12
	(0.59)	*(0.41)*
round 3†	-4.97	-0.48
	(0.35)	*(0.44)*
log informal credit limit (MK)	0.94	-0.49
	*(0.22)****	*(0.17)****

log MRFC credit limit (MK)	-0.51	0.36
	*(0.29)**	*(0.17)***
log 2nd program credit limit (MK)	-0.14	0.32
	(0.24)	*(0.13)***
N. Obs.	256	256

Source: own calculation from MRFMHFS. †dummy variables.***$p < 0.01$,**$p < 0.05$,*$p < 0.1$.Marginal effects and standard errors displayed.

credit by 3.2 percent. The liquidity constraint effect in the formal credit programmes is lower than the informal one, because we have partitioned the credit limit of different credit programmes, whilst the dependent variable aggregates all formal sources to keep a higher number of observations.

More importantly, even extra liquidity from the credit programmes has a significant impact on the demand for informal and formal credit, signalling a certain degree of substitutability between sources. We find that an increase of the credit limit given by the MRFC reduces the demand for informal credit (although there is weak significance only at ten percent). In other words, ten percent extra liquidity from the MRFC programme reduces the informal demand by approximately four percent. That is, households' demand for informal loans corresponds to almost half of any expected increase in the credit line set by the MRFC programme. This is evidence of spillover effects and substitutability between the MRFC programme and informal loans[36].

However, there is no effect of the credit limit set by all the other programmes on the demand for informal credit. This is in support of the results in the fourth chapter where we found that introducing the MRFC programme would reduce the amount borrowed from informal lenders. The impact of participation in the MRFC on the amount households borrow from informal sources depends on its credit limit. However, the effect is not very significant. Perhaps there are factors other than the credit limit and specific

[36]Tables C5-2 and C5-3 in appendix C report the second and first stage regressions of the instrumental variable tobit model where the dependent variable is the amount of debt. The results show that there is still significant evidence of spillover effects and liquidity constraints.

Table 5.9: Instrumental variables tobit - 1^{st} stage regressions

	L^I	L^{MRFC}	$L^{2nd}prog.$
hh characteristics:			
hh size	-0.003	-0.15	0.22
	(0.13)	*(0.15)*	*(0.17)*
age head	-0.01	0.02	0.02
	(0.01)	*(0.01)*	*(0.02)*
female head†	1.55	-0.68	-1.01
	(0.86)	*(0.99)*	*(1.10)*
n. children 0-15	0.001	-0.01	0.16
	(0.17)	*(0.19)*	*(0.21)*
head primary education†	0.04	0.57	-0.26
	(0.34)	*(0.39)*	*(0.43)*
spouse employed in small trade†	-0.66	-1.06	2.37
	*(0.39)**	*(0.45)***	*(0.50)****
Assets, expenditure and shocks:			
land size (ha)	0.18	-0.05	-0.06
	*(0.08)***	*(0.09)*	*(0.10)*
land share owned by spouse (%)	-0.003	0.001	-0.003
	(0.00)	*(0.01)*	*(0.00)*
land share in total assets (%)	-0.01	0.002	0.01
	(0.01)	*(0.001)*	*(0.01)*
value of assets (MK)	-0.0001	0.00004	0.0002
	(0.00)	*(0.00)*	*(0.00)***
food expenditure (MK)	-0.01	-0.001	0.005
	*(0.00)**	*(0.01)*	*(0.01)*
non food expenditure (MK)	0.001	-0.0004	-0.004
	*(0.00)***	*(0.00)*	*(0.00)*
number of negative shocks	0.12	-0.26	-0.12
	(0.12)	*(0.14)*	*(0.16)*
Costs of loans:			
formal interest rate (%)	-0.05	0.09	0.14
	(0.04)	*(0.05)**	*(0.06)***
informal interest rate (%)	0.09	-0.01	-0.01
	*(0.02)****	*(0.02)*	*(0.03)*
formal loan costs	-0.06	0.05	-0.08
	(0.07)	*(0.08)*	*(0.09)*
informal loan costs	0.07	-0.01	-0.01
	*(0.03)***	*(0.04)*	*(0.04)*
% Christians in the same village	0.003	0.01	-0.03
	(0.01)	*(0.01)*	*(0.02)*
South†	-0.99	2.73	-1.72
	*(0.60)**	*(0.69)****	*(0.76)***
round 2†	-1.33	-0.36	0.88
	*(0.43)****	*(0.50)*	*(0.55)*
round 3†	-1.23	1.68	-0.25
	*(0.44)****	*(0.55)*	*(0.55)*
community characteristics:			
number of deep tube wells	0.43	-0.46	-0.09

	*(0.18)***	*(0.23)****	*(0.23)*
avg. price of maize in October	0.05	-1.46	0.46
	(0.70)	*(0.81)**	*(0.90)*
distance to commercial bank	0.04	-0.01	-0.00
	*(0.01)****	*(0.01)*	*(0.01)*
n. of members in farms clubs	-0.001	0.004	0.004
	(0.01)	*(0.02)***	*(0.02)*
n. of households	-0.003	0.0004	0.001
	*(0.01)****	*(0.00)*	*(0.00)*
n. of HHs with land > 5 acres	-0.02	0.01	0.03
	(0.01)	*(0.01)*	*(0.01)*
n. of moneylenders	0.08	-0.40	-0.23
	(0.29)	*(0.33)*	*(0.37)*
hungry season (February)†	1.92	-3.23	3.21
	*(0.68)****	*(0.78)****	*(0.87)****
Constant	1.91	0.24	-1.04
	(1.45)	*(1.67)*	*(1.86)*
N. Obs.		**256**	

Source: own calculation from MRFMHFS. †dummy variables.***$p < 0.01$,**$p < 0.05$,*$p < 0.1$.
Note: first stage regressions are the same for formal and informal credit.The dependent variables are in log form.

to the MRFC programme that affect the household's decision to reduce informal credit. For example, the MRFC can offer medium-term loans (5 years) for acquiring farm equipment.

The evidence of spillover effects can be explained in light of the analysis in the previous chapter and of the descriptive statistics outlined in sub-section 5.3.4. Relatively poorer households reduce the amount they borrow from informal lenders whenever the MRFC programme is made available[37].

Table 5.8 also shows that increasing the informal credit limit would reduce the demand for formal credit. This could be interpreted as evidence of spillover effects from the informal to the formal sector. We will discuss this result more in detail in the next sub-section where we perform the exogeneity tests.

A comparison between the effects of the MRFC credit limit in model II with the informal credit limit in model I, shows that households are more constrained in the informal

[37]Recall from the fourth chapter that participants in the MRFC programme are relatively worse off than participants in more than one programme.

credit sector. There are two interpretations of this result. As friends and relatives are the largest informal credit source, it is possible that larger constraints arise because of their limited availability of funds. Secondly, in model II the dependent variable includes all formal credit sources whilst the credit limit refers only to the MRFC programme. As a consequence, the small effect of the MRFC programme reflects the fact that we only capture part of the demand to the formal sector.

Other additional results can be summarised as follows. We find that an increase in non-food expenditure has a positive and significant relation with the demand for formal credit. By contrast, households' food expenditure is not correlated with the demand for formal loans. This can be explained by the fact that credit programmes are usually delivered for production investments. We also find that an increase in the percentage of Christians in the village increases the demand for informal credit.

The first stage regressions in table 5.9 display each partition of the credit limit vector $L_i = \left[L_i^I, L_i^{MRFC}, L_i^{PROG2}\right]$. We only briefly discuss this table because our main results have been reported in the second stage regressions. It is important to point out that almost all instruments are significant. For instance, the greater the distance to the commercial bank, the higher is the informal credit limit. Informal lenders reduce their credit supply as the number of households in the locality increases. This can be interpreted as evidence of limited availability of funds from informal lenders.

There are two additional points that we should consider before questioning the specification of our model. Firstly, because we are essentially estimating a cross-section regression, we could encounter an inverse causality problem between assets, food, non-food expenditure and the demand for credit. For instance, does the demand for credit influence food expenditure or *vice versa*? Secondly, assets, food and non-food expenditure could be affected by measurement error. While the second problem can be solved by instrumental variables, the first problem cannot really be overcome without any time variation[38]. So, in tables C5-6 and C5-7 of appendix C we check whether our results depend on the inclusion of these (possibly) endogenous variables. It is evident that even after dropping assets, food, and non-food expenditure the results do not change.

5.4.2 Specification tests

We now turn to questioning the models by performing three specification tests as shown in table 5.10. First, we test the validity of our instruments by using an overidentification test. Second, we test whether the vector L_i is exogenous by performing the Wald test and the Smith Blundell test. Finally, we challenge the normality assumption on which the censored regression models are based[39].

We use the Amemiya-Lee-Newey as an overidentification test. The *chi*-squared statistics shown in table 5.10 are: 2.33 with a p-value of 0.94 and 5.28 with a p-value of 0.51 which means that the validity of the instruments is not rejected for models I and II respectively.

The Smith-Blundell test of exogeneity is a simple t-test of the null hypothesis that:

$$H_0 : \rho = 0$$

where ρ is the correlation between u_{1i} and u_{2i} as shown in equation 5.11. In addition, we consider the Wald test value following the ivtobit estimation. Both the Smith-Blundell test and the Wald test of exogeneity reject the hypothesis of endogeneity in the demand for informal credit as shown in table 5.10. This is not surprising because, by definition,

[38] The survey has been collected in three different seasons. However, we could not solve the simultaneity problem because our sample size was severely reduced by the inclusion of lagged assets and expenditure variables.

[39] The rejection of normality makes the censored regression model inconsistent. The tobit model is derived in Appendix B. The marginal effects of the underlying regressions are displayed in tables C5-4 and C5-5 of appendix C.

TABLE 5.10: Specification tests

	Tests after ivtobit and tobit			
	Degrees of freedom	χ^2	*Prob.*$> \chi^2$	*Evidence*
Model I: informal credit				
Amemiya-Lee-Newey overidentification test	7	2.33	0.94	accept H_0
Wald test of exogeneity	3	0.61	0.89	accept H_0
Smith-Blundell test of exogeneity	-	1.84	0.14	accept H_0
Test of normality	-	-	0.00***	reject H_0
Model II: formal credit				
Amemiya-Lee-Newey overidentification test	6	5.28	0.51	accept H_0
Wald test of exogeneity	3	1.01	0.80	accept H_0
Smith-Blundell test of exogeneity	-	3.02	0.03**	reject H_0
Test of normality	-	-	0.00***	reject H_0

Source: own calculation from MRFMHFS.***$p < 0.01$.

the credit limit variable in the survey indicates the maximum value that the borrower *expects* to receive from the lender. The timing of the events is clear: the household's demand is determined by its *expectation* of the credit limit offered by the lender[40]. In addition, most of the instruments are relevant as shown by their significance in the first stage regressions in table 5.9.

However, the Smith-Blundell test shows evidence of endogeneity of the informal and formal credit limits in the formal credit demand equation. There are several explanations for the contrast between the Wald test and the Smith Blundell test. First, it could be that this weak evidence results from only one of the regressors being endogenous (for example, the informal credit limit). Second, it could be that aggregating different formal sources in the dependent variable does not give a clear direction in the causality between credit limit and demand. The test also suggests a weak evidence of exogenous causality of informal credit on the formal credit demand in model II. This could be explained by the fact that, because informal credit is not the cheapest source, it is not the preferred credit sector and, hence, reverse causality may occur.

[40]Assuming that expectations are correct is not too restrictive because information asymmetries are not severe in small villages. Also, credit programmes set an official credit limit.

We also look for the relevance of instruments by using Shea partial R-squared [Shea, 1996] that is used whenever a regression includes more than one endogenous regressor[41]. The Shea test rejects the weak instrument hypothesis for the informal and the more than one programme credit limit (at the 10 percent level) and for informal credit limit (at the five percent level).

The weak evidence of endogeneity of the credit limit could suggest the use of a simpler tobit model. If we were to use the tobit model, however, the concern regarding the normality assumption of the error term would become crucial. The standard censored regression model is not consistent if the error term is not normally distributed. This inconsistency arises because the density of D_i given the covariates (which we call x for simplicity) depends on the fact that D_i^* given x is distributed as a normal [Wooldridge, 2002]. There are several ways to test for normality after a censored regression model [Greene, 2003; Pagan and Vella, 1989]. In this context, we perform a conditional moment (CM) test of normality. The approach taken in the literature is to see whether the third and fourth moments are 0 and $3\sigma^4$ as it should be in a normal distribution[42].

It is evident from table 5.10 that we strongly reject the null hypothesis of normality in both models. Hence, the censored regression model is not appropriate with our data.

As pointed out by Deaton (1997), there are two approaches to take whenever the censored regression model fails to be consistent. The first is to use an estimation strategy that is not affected by the distribution of the error term. The second approach completely

[41]Note that as we could not perform the test in the case of the tobit model, we have performed the test in the case of a linear two stage instrumental variable model. The results of the credit limit coefficients are not very different from the tobit (see Appendix C table C5-8).

[42]More formally, under the normality assumption we should have:

$$E[(D_i - x_i^{'}\beta)^3] = 0 \qquad \text{and} \qquad E[(D_i - x_i^{'}\beta)^4 - 3\sigma^4] = 0$$

The test statistic (See Greene (2003) for a derivation of the moment conditions) is given by:

$$C = i^{'}M[M^{'}M - M^{'}G(G^{'}G)G^{'}M]^{-1}M^{'}i$$

where, as described by Greene (2003), the rows of G are the terms of the gradient of the log-likelihood function and the rows of M contain the sample moment conditions.

abandons censored regression models. In the next two sub-sections we will look at both approaches to check the robustness of the results.

5.4.3 Quantile censored regression

Even when there is no violation of the normality assumption, the tobit model yields biased estimates in the presence of heteroskedasticity. For instance, in the case of demand for informal credit, there are non-applicants among the poor, but there are also many relatively wealthier households who choose not to ask for credit. So, not demanding credit is partly a matter of income and partly a matter of taste [Deaton, 1997].

A simple way to test for heteroskedasticity involves the Hausman test[43]. In this case, the Hausman test can be used to test between two estimators where in the null hypothesis of no heteroskedasticity, H_0, the first estimator, $\widehat{\beta}^{TOBIT}$, is both consistent and efficient, but under H_1 it becomes inconsistent. The second estimator, denoted by $\widehat{\beta}^{CQ}$ and estimated through censored quantile regression[44], is consistent under both H_0 and H_1, but inefficient under H_0. Hence, the test statistic can be written as follows:

$$\left[\sqrt{N}\left(\beta^{CQ}-\beta^{TOBIT}\right)\right]'\left(V^{CQ}-V^{TOBIT}\right)^{-1}\left[\sqrt{N}\left(\beta^{CQ}-\beta^{TOBIT}\right)\right]\sim\chi^2_k$$

where k is the number of coefficients identified in both estimations.

Table 5.11 displays the results of the Hausman test for both models[45] of formal and

[43]This test is discussed in appendix B. If heteroskedasticity fails for the tobit model, the censored quantile regression model is more appropriate. We are aware of the fact that the endogenous censored quantile regression should have been performed (although we found only weak evidence of endogeneity), but it turned out to be computationally heavy with our data. Moreover, we are only interested in looking at the variation of the distribution of credit demand and not at the size of the coefficients.

[44]We discuss this estimator later on.

[45]The regressors used in the estimation are the same as the ones used in tables C5-4 and C5-5 of appendix C.

Table 5.11: Hausman test of heteroskedasticity

	Degrees of freedom	χ^2	**Prob.**$> \chi^2$	**Evidence**
Model I: informal credit H_0: **no heteroskedasticity**	20	87.11	0.00***	reject H_0
Model II: formal credit H_0: **no heteroskedasticity**	20	121.89	0.00***	reject H_0

Source: own calculation from MRFMHFS.***$p < 0.01$.Without bootstrapped std. errors.

informal credit. It is evident that we can strongly reject the assumption of no heteroskedasticity. Once again, the censored regression model is not appropriate for these demand equations.

Since both normality and heteroskedasticity are rejected, censored quantile regressions, although generating "reduced form" parameters represent a more suitable alternative[46]. There are several advantages in using the censored quantile regression, also known as censored least absolute deviation estimator (LAD) implemented by Powell (1984). Firstly, it is a non-parametric estimator where non-parametric refers to the distribution of the error term. In other words, it does not require normality of the residuals. Secondly, it does not assume homoskedasticity[47] or symmetry of the errors. Finally, in the case of a particular quantile, the median regression, the estimates are less sensitive to the presence of outliers in the dependent variable, a common occurrence in developing country data.

Consider a simplified[48] version of the latent variable model in equation 5.10, where the median of u is zero:

$$D^{k*} = x_i\beta + u_i \qquad \mathrm{Med}(u_i \mid x_i) = 0 \tag{5.14}$$

[46]Note, that given the weak evidence on the exogeneity of the credit limit we, hereby, abstract from using endogenous censored quantile models.

[47]As pointed out by Deaton (1997), when the errors are heteroskedastic the quantile regressions for percentiles other than the median will not be parallel to the regression line, but will diverge for bigger value of the regressors.

[48]We have omitted time dummies for simplicity.

Table 5.12 shows the results of two models where the dependent variable is the logarithm of the credit demand to informal and formal lenders just as in table 5.8. The models include the following variables[49]: a) households characteristics that include education of the household head, the occupation of the spouse in a small trade, household size and number of children between 0 and 15; b) proxies for current resources such as total value of equipment (i.e. tractors, threshers etc.), food and non-food expenditure, size of land and share of land in total assets, proportion of land and livestock owned by the spouse, and a proxy for vulnerability (i.e. number of negative shocks); c) a regional dummy indicating whether the household is located in the South and the proportion of Christians in the village where household i lives.

As shown in table 5.8, we use a partition of the vector $L_i = \left[L_i^I, L_i^{MRFC}, L_i^{PROG2}\right]$ which contains the credit limit faced by the households who borrow from the informal sector, from the MRFC only and from more than one credit programme. All these groups of variables are included in vector $x_i = [X_i, L_i]$ of equation 5.14.

The results of the ivtobit are also supported by this specification. We find again evidence of liquidity constraints. An increase in the credit limit of informal lenders and of credit programmes (either the MRFC only or the other programmes) would increase the demand for informal and formal credit, respectively. Although some of the coefficients are slightly smaller than those shown in table 5.8 (i.e. the MRFC and other credit programmes limits in model II), the statistical significance of the MRFC credit limit is higher. Also, with regard to the substitutability of sources, we confirm the results in table 5.8, that is, households in the median distribution of informal credit demand would ask of informal lenders less credit if the MRFC programme increased its credit limit.

Additional results include a positive and significant relationship between food and

[49] We could not include exactly the same variables as in table 5.8 for a problem of convergence of the models.

Table 5.12: Quantile (median) censored regression

	MODEL I: Log(informal credit)	**MODEL II: Log(formal credit)**
hh characteristics:		
hh size	-0.08	0.05
	(0.07)	*(0.12)*
age head	0.003	0.39
	(0.01)	*(0.67)*
female head	-0.17	0.01
	(0.37)	*(0.01)*
n. children 0-15	0.14	0.02
	*(0.08)**	*(0.15)*
head primary education†	0.13	-0.28
	(0.16)	*(0.30)*
spouse employed in small trade†	0.08	0.08
	(0.21)	*(0.38)*
Assets, expenditure and shocks:		
land size (ha)	0.01	0.06
	(0.04)	*(0.08)*
land share owned by spouse (%)	-0.002	0.002
	(0.00)	*(0.00)*
livestock share owned by spouse (%)	0.002	-0.001
	(0.00)	*(0.00)*
land share in total assets (%)	-0.001	-0.005
	(0.00)	*(0.01)*
value of equipment (MK)	-0.00003	-0.0002
	(0.00)	*(0.00)*
food expenditure (MK)	0.01	0.003
	*(0.00)**	*(0.00)*
non food expenditure (MK)	-0.0001	0.001
	(0.00)	*(0.00)***
number of negative shocks	0.06	-0.03
	(0.05)	*(0.11)*
% Christians in the same village	0.004	0.01
	(0.00)	*(0.01)*
South†	0.14	0.37
	(0.13)	*(0.33)*
round 2†	0.01	-2.31
	(0.12)	*(0.98)***
round 3†	0.57	-0.11
	(0.30)	*(0.31)*
log informal credit limit (MK)	1.01	-0.20
	*(0.05)****	*(0.10)****
log MRFC credit limit (MK)	-0.08	0.37
	*(0.04)***	*(0.14)****
log 2nd programme credit limit (MK)	-0.01	0.40
	(0.03)	*(0.15)****
Constant	-1.24	3.04
	*(0.51)***	*(1.31)****
N. Obs	158	341
pseudo-R^2	0.67	0.26

Source: own calculation from MRFMHFS. Coefficients displayed and bootstrapped std. errors (100 replications) in parenthesis.***$p < 0.01$,**$p < 0.05$,*$p < 0.1$.†dummy variables.

non-food expenditure with the demand for informal and formal credit respectively, but the coefficient of non-food expenditure is less significant than in table 5.8.

In sum, this sub-section shows that quantile censored regression models can be used for testing the liquidity constraints and spillover hypotheses whenever the rejection of the heteroskedasticy and normality assumptions renders standard censored regression models inconsistent. The results confirm the existence of liquidity constraints and spillover effects.

5.4.4 Selectivity models

The second way to model the demand for credit abandons the censored model specifications. Apart from the inconsistency caused by heteroskedasticity and non-normality of the error terms, probably the main drawback of the censored regression models is that they characterise censored observations as a corner solution. In other words, the censored regression models do not explain the behaviour of those who do not apply for credit[50]. The Heckman model overcomes this problem by accounting for selectivity in the credit demand through a mixed continuous and discrete choice model. This approach is exactly the same as the one we adopted in the third chapter where here we can model a proper demand for credit.

Suppose that the demand for credit from household i can be modelled by the following equation[51]:

$$D_{i,t}^{k*} = \alpha_{0i} + X_{i,t}^{'}\beta_0 + \text{Ł}_{i,t}\beta_1 + \gamma_0^{'} time_t + u_{1i,t} \tag{5.15a}$$

[50]They may not apply either because they are constrained or because they simply do not want to apply. Recall that the "demand" variable has reduced selection bias by replacing missing debt with the amount rejected households asked for. Despite this, we may still have selection problems because we do not observe households who did not apply at all.

[51]A full derivation of the general model is described in Appendix B.

just like the ivtobit model in equation (5.10). We can now model the selection equation as follows:

$$
\begin{aligned}
I^{k*}_{ij(i),t} =& \alpha_1 + \delta X_{i,t} + \chi C_{j(i)} + \gamma_0^{'} time_t + v_{ij(i),t} \\
& \forall \quad i = 1, \dots, N \text{ and } j(i) = 1, \dots, 44
\end{aligned} \tag{5.15b}
$$

The error terms u_i and v_i have a bivariate normal distribution with covariance $\text{cov}(u_i, v_i)$. The observability criterion for the selectivity model is:

$$D^k_i = D^{k*}_i.1\left(I^{k*}_i > 0\right) \tag{5.16}$$

that is, we only observe the demand for credit of the applicants (whether rejected or not). We cannot observe households who, despite having positive propensity to demand, could not have access to credit. In other words, the sample of households is affected by a selection problem [Heckman, 1979].

As described in the third chapter, there are two ways in which this model can be estimated: a) by using a full-information maximum likelihood (FIML) selectivity model; or b) by using a two-step selection model. Table 5.13 displays the results of selectivity models for the informal and formal credit demand. We use a two-step estimation for the informal credit demand because we cannot reject the hypothesis that equations 5.15a and 5.15b are independent of each other. On the other hand, the rejection[52] of the hypothesis of independent equations supports the estimation of the formal demand through a full information maximum-likelihood model. We report the results of the FIML for the informal demand and of the two-step estimation for the formal demand in table C5-9 of appendix C.

[52]We can only reject it at the ten percent level. However, table C5-9 in appendix C shows that the results do not change once the model is estimated with a two-step selectivity model.

Table 5.13: Selectivity models

	MODEL I: 2 step estimation		MODEL II: FIML	
	Pr(Informal)	**Log(informal credit)**	**Pr(formal)**	**Log(formal credit)**
hh characteristics:				
hh size	0.07	0.32	0.18	0.11
	(0.06)	*(0.20)*	*(0.05)****	*(0.11)*
hh size squared	-	-0.03	-	-0.01
		*(0.01)***		*(0.01)*
age head	-0.01	-0.07	0.004	0.02
	*(0.01)**	*(0.04)*	*(0.00)*	*(0.02)*
age head squared	-	0.001	-	-0.0001
		*(0.00)**		*(0.00)*
female head†	-1.39	0.18	-2.02	-1.26
	*(0.26)****	*(0.68)*	*(0.23)****	*(0.38)****
n. children 0-15	-0.06	0.21	-0.08	0.08
	(0.07)	*(0.12)**	*(0.05)*	*(0.07)*
head can read and write†	-0.05	0.41	-0.04	0.02
	(0.15)	*(0.21)**	*(0.11)*	*(0.12)*
spouse does household work†	-	0.42	-	0.17
		*(0.22)**		*(0.11)*
head employed in agriculture†	-	0.54	-	0.14
		*(0.21)***		*(0.11)*
Assets, expenditure and shocks:				
land size (ha)	-	0.06	-	0.01
		(0.04)		*(0.04)*
land share in total assets (%)	-	0.001	-	-0.002
		(0.01)		*(0.00)*
value of assets (MK)	-	0.0001	-	0.0001
		*(0.00)**		*(0.00)****
food expenditure (MK)	-	0.01	-	0.001
		*(0.00)***		*(0.00)*
non-food expenditure (MK)	-	0.0001	-	0.0004
		(0.00)		*(0.00)***
number of negative shocks	0.19	-	-0.08	-
	*(0.05)****		*(0.04)**	
% people in trad. religion in village	-	-0.02	-	-0.01
		*(0.01)***		*(0.01)**
South†	0.13	-0.22	-0.05	0.40
	(0.22)	*(0.21)*	*(0.17)*	*(0.12)****
round 2†	-0.82	-0.18	-1.99	-0.35
	*(0.13)****	*(0.30)*	*(0.13)****	*(0.33)*
round 3†	-1.68	0.92	-1.90	-0.19
	*(0.19)****	*(0.58)*	*(0.12)****	*(0.39)*
log informal credit limit (MK)	-	0.39	-	-0.06
		*(0.05)****		*(0.02)****
log MRFC credit limit (MK)	-	-0.12	-	0.11
		*(0.06)***		*(0.02)****
log 2nd programme credit limit (MK)	-	0.06	-	0.09
		(0.03)		*(0.01)****
Village characteristics:				
number of deep tube wells	0.26	-	0.11	-
	*(0.09)****		*(0.09)*	
electricity†	0.54	-	0.02	-
	*(0.25)***		*(0.22)*	
farm clubs†	0.05	-	0.57	-
	(0.25)		*(0.23)***	
traditional healers†	-0.38	-	0.36	-
	*(0.17)***		*(0.12)****	

price of maize (july)	-0.81	-	0.27	-
	*(0.35)***		*(0.29)*	
distance to credit office	0.01	-	-0.01	-
	(0.01)		*(0.01)*	
distance to comm. bank (Km)	0.01	-	-0.01	-
	(0.01)		*(0.01)***	
n. of clubs memb.	-0.01	-	-0.01	-
	(0.01)		*(0.01)*	
n. of households	0.001	-	-0.001	-
	(0.00)		*(0.00)**	
n. of HHs with land btw 3-4.99 acres	-0.01	-	0.005	-
	(0.00)		*(0.00)**	
hungry season (February)†	1.08	-	-1.07	-
	*(0.34)****		*(0.29)****	
n. of moneylenders in the village	0.22	-	0.31	-
	*(0.09)***		*(0.07)****	
constant	0.26	1.34	-0.31	4.22
	(0.41)	*(1.18)*	*(0.34)*	*(0.70)****
N. Obs	961		946	
Mills ratio	0.30		-	
	(0.35)			
LR test of ind. equs.	-		**(0.10‡)***	

Source: own calculation from MRFMHFS. Coefficients displayed and standard errors in parenthesis.
$^{***}p < 0.01$,$^{**}p < 0.05$,$^{*}p < 0.1$.†dummy variables.$^{\ddagger}p$-value

For comparability purposes, we include most of the regressors that we used in table 5.8: a) households characteristics that include education of the household head, occupation of the spouse within the household and occupation of the household head in agriculture; b) proxies for current resources such as total value of assets, food and non-food expenditure, size of land and share of land in total assets, proportion of land owned by the spouse; and a proxy for vulnerability (i.e. number of negative shocks); c) prices indicated by the interest rate and other loan terms; d) a regional dummy indicating whether the household is located in the South and the proportion of people who follow traditional religion in the village.

Identification requires that the selection equation 5.15b includes at least one regressor that is not present in equation 5.15a. Indeed, village-specific characteristics, $C_{j(i)}$, and the number of negative shocks have been considered in equation 5.15b. Also, equation 5.15a includes households' assets not included in equation 5.15b. The vector $C_{j(i)}$ represents characteristics that vary only across villages, but not across households (i.e.

number of deep tube wells, dummies indicating whether the village has electricity, farm clubs and traditional healers, price of maize in July, distance to the credit office and to the commercial bank, number of households, of clubs' members and number of households with land between three and 4.99 acres, a dummy for the hungry season of February). The probability of demanding credit depends on a set of households' characteristics (i.e. age, household size, number of children and dummies indicating whether the household is female headed, whether the household head can read and write). We also include the number of negative shocks in the last seven years. Given that household i has a positive demand, the amount of credit (in logarithm) asked of formal and informal lenders depends on the value of assets and expenditure as well as on households' characteristics.

The most important result is that the selectivity models in table 5.13 strongly support the liquidity constraints and spillover hypotheses. Because the credit limit coefficients are not included in the selection equation, we can interpret them as in a standard regression model. A ten percent increase in the informal credit limit would increase households demand for informal credit by almost four percent. Similarly, increasing the credit limits set by the MRFC and other credit programmes would have a positive and significant impact on the demand for formal credit. We can also confirm spillover effects, that is, a ten percent increase of the ceiling set by the MRFC programme would reduce households' demand for informal credit by 1.2 percent.

Other important results entail households' characteristics, shocks and village characteristics. According to the selection equations in models I and II, a female headed household has a negative probability of demanding credit from informal and formal lenders, respectively. The more shocks the household has faced in the last seven years, the higher is the probability of demanding credit of informal lenders (the coefficient is highly significant at the 1 percent level). However, the same variable has a negative impact on the probability of asking for credit from informal sources, but it is only significant at ten percent level. This result states that households rely on informal sources to cope with shocks[53] [see for example, Bardhan and Udry, 1999; Hoddinott et al., 2005; Ray, 1997].

We also find that in the hungry season of February households are more likely to ask for informal credit and less likely to ask for formal credit. An explanation of this result could be that informal credit is used primarily for consumption that needs to be supported particularly in the hungry season.

Most of the village characteristics have a significant impact on both the demand for informal and formal credit. For example, the existence of farms' clubs in the village has a positive and significant effect at the five percent level of the demand for formal credit. Farms' clubs allow people to form group-lending institutions.

Networks created by the clubs help farmers to choose members of a jointly liable group. Ghatak (1999) showed that in jointly liable groups, members match with their same "type" and form homogeneous groups (positive assortative matching).

We also show that as the distance to the nearest commercial bank increases, the probability to ask for formal credit decreases (the coefficient is significant at the 5 percent level). In other words, not only is the positive demand for credit affected by households' characteristics, but also by the characteristics of the villages where households live as we have already seen in the third chapter.

The second and fourth columns of table 5.13 report the results of the demand for credit taking selectivity into account. As we have already found in chapter three, the higher the value of assets, the higher the demand for formal credit. Relatively wealthier households rely more on formal lenders. The selectivity models confirm the results of the ivtobit model, that is, non-food expenditure has a positive relation with the demand

[53]Note, unlike the ERHS we do not know whether these shocks are idiosyncratic or not.

for formal credit. These results may indicate that formal credit is used for farm investments.

It is evident that the demand for formal credit is heterogeneous across the country: households who live in the South have a significantly higher demand for formal credit.

In sum, evidence on the liquidity constraints and spillover hypotheses in Malawi has been confirmed even after taking selectivity of credit demand into consideration. The similarity in the statistical significance between the results of different specifications may be related to the fact that selectivity is not severe (i.e. the Mills' ratio is not significant) and this may be due to the way we have constructed our dependent variable (i.e. "demand" for credit). However, the size of the liquidity constraints and the spillover effects is smaller than in the previous specifications. A ten percent increase in the informal credit limit would increase households demand for informal credit by almost four percent. Similarly, increasing the credit limits set by the MRFC and other credit programmes would have a positive and significant impact on the demand for formal credit. We can also confirm spillover effects, that is, a ten percent increase of the ceiling set by the MRFC programme would reduce households' demand for informal credit by 1.2 percent.

5.5 Conclusion

Using the Malawi Rural Financial Markets and Household Food Security Survey (FMHFS, 1995), an original data set that contains information on credit limit, credit demand of both rejected applicants and borrowers for formal and informal credit sources, we test the following two theories. First, the "liquidity constraints" theory, that is, an increase in the credit limit should affect the demand of liquidity constrained households.

Second, the “spillover” theory, that is, any change in the credit limit of one sector has an impact on the demand for credit of the other sector. The theoretical assumption of this view is that formal credit is the cheapest available source, but it is rationed. Hence, the informal sector arises as a spillover from the rationed formal credit market [Banerjee and Duflo, 2001; Bell et al., 1997; Besley, 1994; Eswaran and Kotwal, 1990].

As the spillover effect results from the existence of liquidity constraints, the two hypotheses are linked together. We use the credit limit to detect for liquidity constraints. This approach has been developed by researchers at the International Food Policy Research Institute (IFPRI) [Diagne, 1999; Diagne et al., 2000; Zeller and Sharma, 1998] in an attempt to overcome the disadvantages of qualitative studies provided by, amongst others, Jappelli (1990). Both applicants and non-applicants were asked the maximum amount they expected a lender would be willing to lend, which is the credit limit of the respondent with regard to that particular lender. So, whenever the demand exceeds the credit limit the household is said to be credit rationed.

We make several contributions to the literature. First, our approach differs from those of Diagne (1999) and Diagne et al. (2000) because we explicitly differentiate credit limits supplied by one or more credit programmes. The rationale of this approach lies in the explanation of the results of the fourth chapter where we found that the introduction of the MRFC programme partially crowds out access to informal loans. In this chapter we have shown that the partial crowding out can be explained to a certain extent by the existence of spillover effects. A ten percent increase in liquidity from the MRFC programme reduces the informal demand by approximately four percent. The spillover effect can be explained by the fact that the MRFC programme is cheaper than informal loans. This result supports the spillover theory, that is, households resort to the informal sector only after having been rationed by the cheaper formal sector.

Second, we provide evidence of liquidity constraints in both the formal and informal sector. Following Gross and Souleles (2002), we look at the significance of the marginal propensity to consume out of liquidity interpreted as dDemand/dLimit. In particular, a ten percent increase in the informal credit line would increase the amount households ask of informal lenders by more than nine percent. This is evidence of a high degree of liquidity constraints. Households are also constrained in their demand for formal credit. A ten percent increase in the MRFC credit limit increases the demand for formal credit by almost four percent. In addition, an increase of ten percent in the credit line set by more than one credit programme would increase the demand for formal credit by 3.2 percent.

Third, unlike previous studies that adopted a reduced form specification in which demand and supply are collapsed into a single variable, we have been able to disentangle demand and supply equations in two ways. The very rich data set allows the identification of the demand equation and the supply equation (which is the credit limit equation) for both applicants and non-applicants to formal and informal lenders. In addition, following Diagne (1999) and Grant (2007) we apply a number of exclusion restrictions to identify demand and supply equations such as seasonal dummies and village characteristics.

Finally, we perform several robustness checks by addressing specification issues that may seriously affect the results (for example, heteroskedasticity, non-normality and selectivity). Both the liquidity constraints and the spillover hypotheses hold even after changing the estimation methods.

The identification and specification issues, as well as the complexity of the problem we have attempted to analyse, requires a cautious interpretation of the results. We can pin down several drawbacks of our approach. First, we neglect the choice based sampling[54] of the survey. There are two reasons for this decision: a) Diagne (1999) showed that the sampling correction is only necessary when the credit programme dummies are included[55]; and b) it is not possible to apply the weights in a censored regression model. Hence, our results should not be interpreted as representative of the country.

Second, although we use a very rich data set, the cross sectional analysis carries many problems - the most severe one being simultaneity. In particular, we show that credit limit and credit demand are correlated, but we find weak evidence of reversed causality between the two variables. Nevertheless, a more correct approach to test for liquidity constraints should have looked, as in Gross and Souleles (2002), at the effect of the credit line variation over time on the credit demand.

Third, the lack of appropriate information on lenders' characteristics casts doubt on the relevance of the instruments. We only find weak evidence of their relevance and thus caution should be used to interpret the results of causal effects.

Fourth, the replacement of unobserved debt for non-borrowers with their demand for credit may introduce ambiguity in the interpretation of the coefficient of the credit limit variables as evidence for liquidity constraints. We support the results in two ways. The so-called "demand" variable used in the estimation allows for a sample selection correction. In addition, we check whether the results are driven by the way we deal with non-borrowers and we find that whenever debt holding is used as dependent variable there is still evidence of liquidity constraints and spillover effects.

Finally, there are measurement errors that could affect assets and expenditure variables. We adopt a simple solution that drops these variables and we show that our main

[54] We have explained this concept in the fourth chapter.

[55] He showed that the correction for choice-based sampling consists only of replacing the programme dummies by the corresponding estimated choice-based corrected conditional probabilities.

results remain unchanged: households in Malawi are constrained and would ask less of informal lenders were they not rationed by the MRFC credit programme.

Chapter 6

Conclusions

6.1 Introduction

Access to credit can be an important instrument for improving the welfare of poor households through at least two channels. The first channel is direct and allows coping with risk by smoothing consumption against shocks. The second channel through which access to credit improves welfare is by enhancing investments in physical and human capital.

These objectives have been the premise for government interventions in credit markets. However, the creation of formal credit institutions such as commercial and agricultural banks have failed to cater for the credit needs of smallholders. The success of these institutions depends on their incentives and enforcement mechanisms as well as the environment they operate in. As credit markets in developing economies are dominated by informal credit institutions, the analysis of the interaction between formal and informal credit institutions is crucial to understanding how policy objectives (such as welfare improvement) can be achieved.

The objective of this thesis is to contribute toward this debate by analysing the determinants of access to informal credit institutions in rural Ethiopia, the effectiveness of the Malawian government credit policy aimed at displacing informal borrowing by creating microfinance institutions, and the reasons for the persistence of informal institutions in Malawi where formal institutions are available.

The remainder of this chapter is organised as follows. Section two presents the main findings, the contributions and the limitations of each empirical chapter. Section three summarises some concluding remarks.

6.2 Main findings

This section summarises the main findings, the contributions and limitations of each chapter. The following sub-sections correpond to the three empirical chapters each entitled with the research question addressed by the relevant chapter.

6.2.1 Why do households participate in informal credit institutions?

The first empirical chapter analyses the determinants of households' participation in informal arrangements by using a panel data set of 15 peasant associations in rural Ethiopia (ERHS, 1994-1997).

The literature highlights two motives for the existence and diffusion of informal credit in developing economies. The economic approach maintains that informal finance has an advantage over formal institutions as it can overcome informational and enforcement problems arising from credit market failures [Bardhan and Udry, 1999; Besley, 1999; Gosh et al., 1999; Ray, 1997]. The cultural or sociological approach, by contrast, sees markets as bound up with networks of personal relations, kinship and reciprocal norms that are more extensive than in formal contracts [Aryeetey and Udry, 1995; Azam et al., 2001; Fafchamps and Lund, 2003; Platteau, 2003; Udry, 1990].

This chapter identifies three groups of factors pertaining to the above mentioned motives for the diffusion of informal arrangements. The first group - household-based determinants such as wealth and demographic characteristics - has been well discussed within the large literature on this topic [for example, Bose, 1998; Kochar, 1997; Pal, 2002; Ravi, 2003; Ray, 1997]. However, a limitation of these studies is that a high degree of collinearity between household-specific variables (such as components of wealth, income and other household characteristics) limits the significance of individual regressors.

The second group - cluster-based determinants such as demographic, infrastructural and geographical characteristics - is often ignored by the literature due to limited data and lack of appropriate empirical models able to identify such characteristics. Knowledge of these cluster-level differences is as important as knowing why households utilise such institutions in clusters where they are available.

The third group - idiosyncratic and aggregate shocks - has been analysed by the literature as a motive for participation in credit markets [e.g. Bardhan and Udry, 1999; Binswanger and Rosenzweig, 1993; Platteau and Abraham, 1987; Ruthenberg, 1971; Townsend, 1994]. However, data availability limits the identification of cluster level and household level shocks which may affect access to credit.

The contribution of this chapter is to address the above-mentioned limitations of the literature. We address collinearity of wealth by using principal component analysis in a logit regression. Sample selection is dealt with a Heckman selection model. Then, we identify cluster-based and household-based determinants of participation in informal credit by adopting an endogenous switching regression model with principal components.

The logit specification shows that principal components can deal with collinearity between wealth variables. It also points out that significant differences between southern and northern Ethiopia influence the existence of a particular informal credit arrangement (i.e. the RoSCA-type institution called *equb*). These differences affect the access to and the substitutability between credit sources.

After finding no evidence of sample selection bias in a Heckman model of informal debt holding, we model households' participation in informal credit adopting a switching regression with endogenous criterion [Lee, 1978; Maddala, 1983]. The endogenous switching regression models for mixed continuous and discrete variables consist of joint estimation of the probability that in cluster j *equbs* are available (the switching group) and the amount of informal credit borrowed. This specification allows modelling the demand for a particular type of informal credit (i.e. *equbs*) as endogenously determined by household-based and cluster-based determinants. Then, the access to informal credit is allowed to differ across endogenously different clusters.

We find that access to informal credit is significantly determined by both cluster-based and household-based characteristics. Income diversification (proxied by the number of villages), availability of formal institutions (proxied by the distance to the bank) and incidence of aggregate shocks (proxied by the size of rain fed land) are all factors that positively and significantly determine the demand for informal arrangements such as *equbs*.

Conditional on the endogenously determined socio-economic characteristics, we then model the amount of informal debt held by households. The results show that idiosyncratic shocks significantly increase the access to informal finance. This confirms the literature stating that informal credit arrangements are mostly effective in settings where incomes are not highly correlated [Binswanger and Rosenzweig, 1986; Ruthenberg, 1971;

Townsend, 1994; Udry, 1999]. Wealth components are also positively correlated with the access to informal credit.

To sum up, in rural Ethiopia the participation in informal credit arrangements is not only determined by factors identifiable at the household level (such as wealth, demographic characteristics and idiosyncratic risk), but also by cluster-based characteristics (such as income diversification, aggregate shocks and geographical factors). Knowledge of these cluster-level differences is as important as knowing why households utilise such institutions in clusters where they are available.

6.2.1.1 Limitations

There are several limitations in this chapter that are due to data availability. We use a household panel data set that includes four rounds. In an attempt to generate an improvement in efficiency, we increase the sample size by pooling the data. However, this formulation does not distinguish in any way between two different households and the same household at two points in time.

The same motivation (i.e. the attempt to maintain a large sample size) has influenced the decision not to lag expenditure and wealth variables. Indeed, we make no attempt to establish a causal relation between participation in informal credit and principal components of wealth.

Another limitation of this chapter is the assumption that clusters' characteristics are fixed over time. Because the village studies were taken at one point in time, some of the clusters' characteristics such as distance to the bank and number of households in the cluster are considered to be fixed. This is not a strong assumption for at least two reasons. First, these particular characteristics should not significantly vary across time. Second, the household survey covers rounds relatively close to each other (for example, the first and second round were undertaken in the same year).

There is scope for extending the analysis carried out in this chapter by including two more recent rounds when they become publicly available. This will enable us to use panel data techniques and resolve possible simultaneity problems.

6.2.2 Do governments displace the informal loan market by introducing formal credit institutions?

If the market failure view mentioned in the previous sub-section holds, then it is the information on individual borrowers and localities required in developing economies that precludes efficient market coverage from large formal credit institutions. Banks have funds to lend, but lack adequate information and enforcement mechanisms to recover the loans.

One of the policies that arises as a response to these market failures aims at creating microfinance institutions that will acquire information in innovative ways. By mimicking and exploiting some of the features of informal lending, banks can design credit contracts that harness local information and give borrowers incentives to use their own information on their peers to the advantage of the bank [Armendariz and Morduch, 2005; Ray, 1997]. For instance, in group-lending programmes borrowers who cannot offer any collateral are asked to form small groups. Group members are held jointly liable for the debts of each other. Formally speaking, what joint liability does is to make any single borrowers terms of repayment conditional on the repayment performance of other borrowers in a pre-specified and self-selected group of borrowers.

The second empirical chapter evaluates the effectiveness of this policy by testing whether the microfinance institutions created by the government of Malawi in 1995 under the Policy Framework for Poverty Alleviation (PAP) crowd out access to informal credit. We use a rich financial survey: the Malawi Rural Financial Markets and Household Food Security Survey (FMHFS, 1995) conducted by IFPRI in cooperation with the Rural Development Department of Bunda College of Agriculture. The survey contains information about households' borrowing behaviour from both informal lenders and microfinance programmes[1].

Like other studies on crowding out [e.g. Attanasio and Rios-Rull, 2000; Kaboski and Townsend, 2006] we adopt policy evaluation techniques in order to identify a *causal* relationship between access to government sponsored credit programmes and informal loans. We use propensity score matching to determine the existence and the extent of the impact of group lending institutions on the access to informal loans.

The evaluation approach consists of four stages. First, we obtain the propensity scores from a series of logit models. In the second stage we perform matching with the Mahalanobis metric algorithm. The third stage estimates the average treated effect (ATE) and the average treatment effects on treated households (ATT) who participate in one, or more than one, credit programme relative to past-membership. The outcome of interest is the amount households borrow from informal sources. The final stage ensures that the results do not depend on the methodological assumptions of the evaluation procedure.

One contribution of the chapter is to adopt a rigorous sensitivity analysis by performing the following robustness checks. It changes the model specification and matching algorithm; the definition of treatment and outcome; and the model used to estimate the propensity scores.

We find strong evidence of crowding out of group lending on informal sources. The results show that participation in the MRFC microfinance programme significantly reduces borrowing from informal sources (by approximately two U.S. dollars). In relative

[1] A description of the credit programmes is contained in sub-section 4.2.2 of chapter four.

terms, it reduces the amount members borrow from informal lenders by more than 70 percent in all the specifications.

Some of the existing literature adopts reduced form equations whereby the supply function cannot be disentangled from demand shifts. Therefore, the coefficients cannot be interpreted as pure substitutability between credit suppliers. In this empirical essay we make an attempt to separate out demand and supply. The rich data set provides information on the amount households asked of informal lenders and the maximum amount they think they will be able to borrow (the credit limit variable[2]). The results show that the MRFC credit programme reduces the demand for informal credit. This is evidence of the fact that the MRFC programme and informal loans are, at least partly, substitutable. On the supply side, we find no significant evidence on crowding out of informal lenders.

Another contribution of this chapter is the test of crowding out in presence of *expected* transfers. Nearly all the literature has focused on crowding out in the context of *realised* transfers. Yet households' demand for informal loans is also affected by the membership in a microfinance programme not just by the actual borrowing [Cox and Fafchamps, 2008]. We find evidence of crowding out for both borrowers and members of the MRFC credit programme.

Finally, this chapter innovatively applies the multiple treatments model found in the labour economics literature [for example, Brodaty et al., 2001; Frölich et al., 2004] to test the crowding out hypothesis. This allows a comparison between the effectiveness of different credit programmes as well as between households that differ in their economic status.

Whilst we find significant evidence of crowding out of one credit programme (i.e. the

[2]The credit limit can be interpreted as a supply function. However, caution is needed in interpreting this variable as the lender is not price taker in this market.

MRFC), there is no significant effect of membership in more than one credit programme on the access to informal loans. We provide several explanations for this result.

As participants in more than one credit programme turn out to be relatively better off, we interpret this result as evidence that crowding out is affected by the credit constraints that arise from households' wealth heterogeneity. This is in line with findings of Navajas et al. (2003) who showed that less capitalized borrowers switch from an informal credit contract to a loan contract provided by microfinance institutions. Relatively wealthier households, by contrast, may not substitute one source for the other but simply increase the overall demand for credit once the supply of formal loans increases.

Secondly, the other credit programmes may not be substitute for informal loans as they serve different purposes[3] [Mohieldin and Wright, 2000] or are more expensive.

Finally, the insignificant crowding out effect of households participating in more than one credit programme may be affected by the fact that we pool different types of programmes. Unfortunately, we do not have enough observations to disentangle the effect of each credit programme and hence we cannot investigate this issue further.

The first two explanations seem to be most plausible. In particular, households could use multiple-borrowing because of market segmentation and because of credit rationing from formal credit programmes. It is actually possible to test the rationing hypothesis by using the information on credit limit provided by the survey. If a change in the credit limit of participants in the MRFC programme has a significant effect on the demand to informal sources, then we can interpret this result as evidence of the existence of credit constraints [Diagne, 1999; Diagne et al., 2000; Gross and Souleles, 2002]. We deal with this issue in the last chapter.

In summary, the creation of the MRFC programme in Malawi displaces access to

[3]We have highlighted the importance of market segmentation in the second chapter.

informal loans. This intervention in the credit markets is important to achieve market efficiency (through the displacement of exploitative informal lenders); distributional objectives (through the provision of credit to poor but entrepreneurial farmers); mitigation of vulnerability (through the provision of funds that are independent of common shocks at the cluster level); and poverty reduction.

6.2.2.1 Limitations

Data availability also affects this chapter. First, the second treatment group (i.e. participants in more than one credit programme) in the propensity score matching results from an aggregation of several credit programmes with different eligibility criteria. Unfortunately, the lack of a sufficient number of observations for each programme impedes the creation of a more "robust" treatment group.

Second, the control group of past members also aggregates different credit programmes. The fact that the majority (but not all) of past members participated in a previous version of the MRFC programme makes the control group to be most appropriate for the first treatment (i.e. participants in the MRFC). However, we do not have enough observations to separate out past members of different credit programmes.

Finally, in an attempt to avoid reversed causality between participation in credit programmes and wealth we have not included any income variable. However, it could be that the eligibility criteria between past members and current participants changed over time. Unfortunately, because we have one cross section we cannot test this hypothesis. Nevertheless, one could say that since we have included education, which is higly correlated with income, we can partially control for the households' economic status. Another way to overcome this problem could be to construct a pseudo-panel by merging this cross section with other household surveys available in Malawi. Admittedly, it would be difficult to merge similar households given that this financial survey focuses on specific villages.

6.2.3 Why do formal and informal credit markets coexist?

In spite of recent financial liberalisation aimed at broadening formal credit markets and in spite of interest rate differentials, in Africa the formal and informal credit sectors persist in the same market[4]. Two main explanations are offered by the literature. First, the informal sector may be the recipient of "spillover" demand from the rationed formal sector [Banerjee and Duflo, 2001; Bell et al., 1997; Besley, 1994; Eswaran and Kotwal, 1989]. The theoretical assumption of the spillover view is that informal credit sources are more expensive than formal loans. Therefore, according to this view, there is a natural ordering of credit sources where a borrower who uses secondary sources (i.e. informal credit) is assumed to be unable to satisfy his financial needs from the primary sources (i.e. formal credit). The borrower is said to be credit rationed with regard to the primary source[5]. Indeed in developing economies, such as in Africa, formal credit rationing is extensive because of information asymmetries, lack of collateral and legal enforcement.

An alternative explanation for the coexistence of formal and informal sectors is the occurrence of market segmentation. According to this view, the unique characteristics of the informal and formal credit sectors inhibit the substitution of one source for the other. As a result, the informal sector need not be the sector of last resort, but instead the preferred sector.

This last chapter tests two hypotheses: 1) the spillover hypothesis; and 2) the liquidity constraints hypothesis. The goal of the chapter resides in an attempt to motivate the

[4]See chapter two for a more detailed discussion of this issue.

[5]It might be possible, however, that she is also rationed on the use of the secondary source.

result of the previous chapter where we found that an increase in the supply of credit within a village causes a (partial) shift from informal sources to a government-sponsored institution (i.e. the MRFC programme). In this context, the spillover hypothesis implies that there is a certain degree of substitutability between the MRFC programme and informal credit and that a reduction of demand for the latter can be achieved by increasing the ceiling on the MRFC programme. We also want to provide evidence for the existence of liquidity constraints in the credit provided by government-sponsored programmes and by informal lenders. As the spillover effect results from the existence of liquidity constraints, the two hypotheses are linked together.

The chapter adopts the Malawi Rural Financial Markets and Household Food Security Survey (FMHFS, 1995) which contains information on credit limit, credit demand of both rejected applicants and borrowers for formal and informal credit sources in Malawi. We use the information on the credit limit as a direct test for liquidity constraints.

The contributions to the literature can be summarised as follows. First, our methodology differs from those of Diagne (1999) and Diagne et al. (2000) as we explicitly differentiate between the credit limits supplied by one or more credit programmes. The rationale of this approach lies in the explanation of the results of the previous chapter.

The results show that the partial crowding out can be explained to a certain extent by the existence of spillover effects. A ten percent increase in liquidity from the MRFC programme reduces the informal demand by approximately four percent. The relatively small effect can be explained by the fact that the coexistence of formal and informal credit sources can also be due, to a certain extent, to market segmentation.

Second, we provide evidence of liquidity constraints in both the formal and informal sectors by adopting the credit limit method. A ten percent increase in the informal credit line would increase the amount households ask of informal lenders by more than nine percent. Households are also constrained in their demand for formal credit. A ten percent increase in the MRFC credit limit increases the demand for formal credit by almost four percent. In addition, an increase of ten percent in the credit line set by more than one credit programme would increase the demand for formal credit by 3.2 percent.

Third, unlike previous studies that adopted a reduced form specification in which demand and supply are collapsed into a single variable, we are able to disentangle demand and supply equations in two ways. The data set allows for the identification of the demand equation and the supply equation (which is the credit limit equation) for both applicants and non-applicants to formal and informal lenders. In addition, following Diagne (1999) and Grant (2007) we apply a number of exclusion restrictions to identify demand and supply equations such as seasonal dummies and village characteristics.

Finally, we perform several robustness checks by addressing specification issues that may seriously affect the results (for example, heteroskedasticity, non-normality and selectivity). Both the liquidity constraints and the spillover hypotheses hold even after changing the estimation methods.

To conclude, this chapter explains the coexistence of formal and informal credit markets with the spillover hypothesis. It also provides evidence for the existence of liquidity constraints. As the spillover effect results from the existence of liquidity constraints, the two hypotheses are linked together. We find that a ten percent increase in liquidity from the MRFC programme reduces the informal demand by approximately four percent. We also find that households are constrained in their demand for formal credit. A ten percent increase in the MRFC credit limit increases the demand for formal credit by almost four percent. In addition, an increase of ten percent in the credit line set by more than one credit programme would increase the demand for formal credit by 3.2

percent.

6.2.3.1 Limitations

Several drawbacks of the approach taken in this last chapter can be highlighted. First, we neglect the choice based sampling of the survey. There are two reasons for this decision: a) Diagne (1999) showed that the sampling correction is only necessary when the credit programme dummies are included[6]; and b) it is not possible to apply the weights in a censored regression model. Hence, our results should not be interpreted as representative of the country.

Second, although we use a very rich data set, the cross sectional analysis carries many problems - the most severe one being simultaneity. In particular, we show that credit limit and credit demand are correlated, but we find no evidence of reversed causality between the two variables. Nevertheless, a more correct approach to test for liquidity constraints should have looked, as in Gross and Souleles (2002), at the effect of the credit line variation over time on the credit demand.

Third, the replacement of unobserved debt for non-borrowers with their demand for credit may introduce ambiguity in the interpretation of the coefficient of the credit limit variables as evidence for liquidity constraints. We support the results in two ways. The so-called "demand" variable used in the estimation allows for a sample selection correction. In addition, we check whether the results are driven by the way we deal with non-borrowers and we find that whenever debt holding is used as dependent variable there is still evidence of liquidity constraints and spillover effects.

Finally, there are measurement errors that could affect assets and expenditure variables. We adopt a simple solution that drops these variables and we show that our main

[6]He showed that the correction for choice-based sampling consists only of replacing the programme dummies by the corresponding estimated choice-based corrected conditional probabilities.

results remain unchanged: households in Malawi are constrained and would ask less of informal lenders were they not rationed by the MRFC credit programme.

6.3 Concluding remarks

Despite the limitations outlined in the previous section and the fact that Ethiopia and Malawi are different countries, a unified story can be drawn from this thesis.

As participation in informal arrangements depends on the socioeconomic characteristics of households as well as clusters (third chapter), one way for banks to enter this market and exploit local information is to give borrowers incentives to use their existing social linkages to the advantage of the banks (fourth chapter). But information problems are only part of the story, other market failures such as weak legal enforcement and the low level of social capital may force the banks to ration credit and cause the persistence of informal credit institutions (fifth chapter). In addition, if the "social" motive for participation in informal arrangements prevails over the "economic" motive, segmentation occurs despite banks' attempt to enter the market and complete crowding out will not be achieved (third chapter).

In conclusion, whether microfinance programmes or, indeed, any formal credit institution will be able to enter a rural credit market and "displace" informal loans depends on several factors including the design of the programmes, the target groups and on the communal norms of the localities where these programmes are adopted. The extent to which the macro-level norms guide micro-level behaviour will depend on the larger context of social and economic change [Fafchamps, 2006; Durlauf and Fafchamps, 2005]. While appropriate reforms could improve the economic context, the endowment of social capital evolves more slowly [Marchesi, 2002]. As argued by Williamson (2000), social capital is not the objective of a policy reform but a constraint to it.

Appendix A

Data description and management

ETHIOPIA

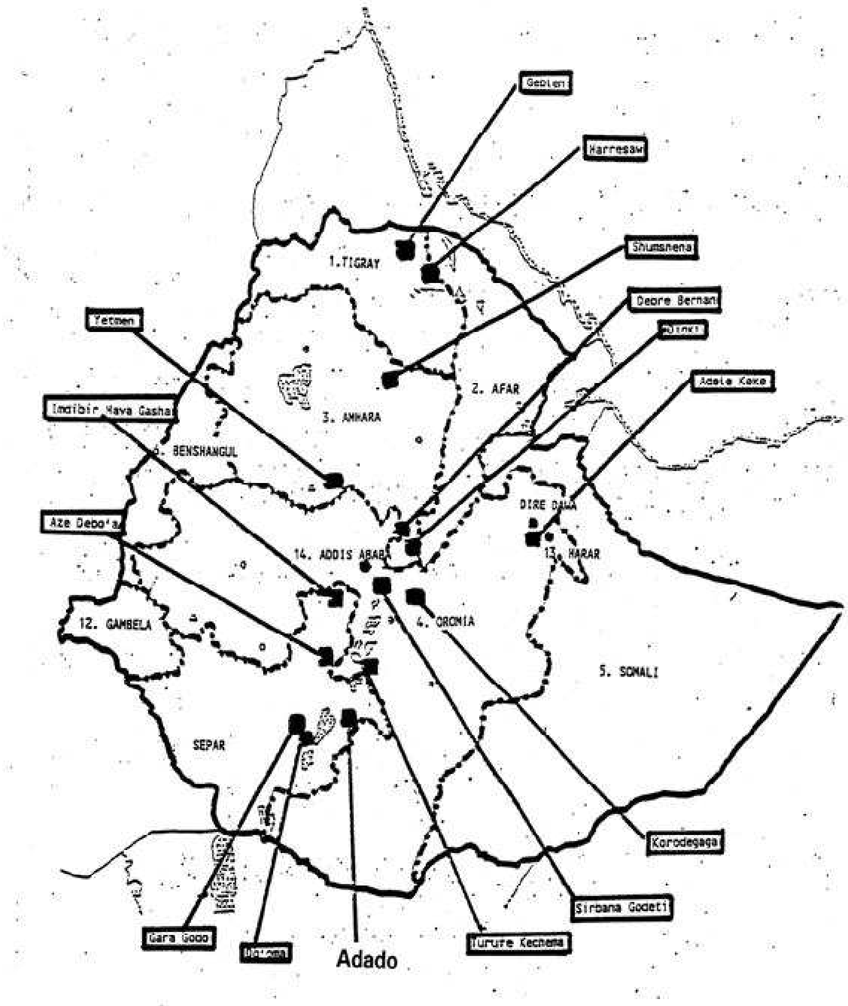
Structure of the ERHS

Structure of the data set

The data set contains the following sections:

- *Household Data:* each survey round contains data entailing a large variety of topics;
- *Questionnaires:* describe the different topics corresponding to each survey round in the household data and consider community level questions;
- *Community Data:* contain information on the 15 communities from Round 4;
- *Village Studies:* is a qualitative description of the 15 villages considered in the survey;
- *Aggregates:* contain aggregate information on anthropometrics, land, consumption and health;
- *Conversions:* contain information on conversion of local units of measure to standard units.

FIGURE A.1: Surveyed sites in Ethiopia



Source: Village Studies in ERHS data.

Village studies: summary

<u>*Consumption, assets, wealth and poverty*</u>

Adado

From May to September people rely on *enset* and cabbage. *Enset* is the emergency staple food in the area: this plant is usually drought resistant and available when there is a short supply of other crops. Between October and April people eat cow beans, horse beans, barley, maize and meat. The hungry season is September. Husband and wife eat from the same plate, while children eat separately. In 1983, a major famine affected the area and made even *enset* scarce.

Livestock consist of cattle and sheep and they constitute an important source of income. Sheep can be easily sold in the local market for cash needs. The source of wealth in the community is trade in agricultural products, farming and land. The rich buy coffee and other crops to be sold when the price increases. Respondents have ranked households according to their wealth in four categories. The wealthy are described as having many plots of land; having sufficient coffee and *enset*; saving in cash; having galvanized iron roofs to their houses; having sufficient clothes; being able to borrow from outside the household; possessing the surrounding crops; being able to produce a wide range of products for consumption and sale and being able to employ daily labourers in their farming land. The middle category have fewer plots than the rich; they cannot overcome the problem caused by coffee disease; they employ less labourers than the rich; have enough land to some extent; have a roof made from thatch and occasionally galvanized iron; do not borrow from outside the household. The poor have very small plots; or they have many children with not enough land; some are employed in the surrounding service cooperative coffee mill; some work for individuals; have houses whose walls are made from bamboo and the roof from *enset* leaves. The very poor are those who have no land or crops for consumption; have no house or one made of twigs and *enset* leaves; are employed in coffee milling or in local households and in the town; are unable to borrow since they have no collateral.

Adele Kebe

Emergency food in the area is provided by sweet potatoes and wheat flour. The fruit and berries of wild food such as cactus are also eaten during famine. Between September and November people eat rice, spaghetti and macaroni since it is the wedding season. Milk is drunk from September to December and eggs are eaten from October to December and from February to June. People usually eat the same kind of food. The husband eats before the others, then children eat and the wife eats only after she has fed the family. They eat from different plates. Livestock are used as assets for saving and investments. Epidemics have killed many animals, but now the Ministry of Agriculture has a vaccination programme. Many people sold their livestock because of drought in 1994/1995. The wealthiest people in the community are those who have *chat* plantations, who have a good number of livestock or those who trade beside farming. The source of their wealth is the land which is planted with *chat* and crops like sorghum and maize. In their home one could find: a wooden or iron bed, trays, cups, tape recorder, radio set, bags, mattress, kerosene lamp, thermos flask etc. The poorest people are those who do not have fertile land or only a very small amount of it. They are often widows, or they have many children and/or they are old and sick. In their home one could find: skin mats, wooden plates, cooking pans, kettle.

Aze Deboa

People eat a variety of crops like potato, tef, barley, maize, products of *enset*, wheat and others. But in emergency situations such as drought and famine *kocho*, *enset* and *kolo* (roasted cereal) predominate in the diet. The hungry season is between January and May, when crops become expensive and there is only *enset*. Two serious famines occurred in 1985 and 1994. Aze Deboa is average in terms of wealth when compared to other villages in the area. The wealthiest people in the community are those who have many and quality cattle, sheep and goats, coffee, other trees and enset. The major sources of their wealth is the hard wok. Equipment in a wealthy house consists of: table, chairs, clay pots, cup, wooden or metal bed, tape recorder, knife, bottles etc. The poorest households are those with small and infertile land, demobilised soldiers, formerly resettles residents, those with few animals and those who do not work hard. Equipment is a poor house consists of: clay pot, glass, wooden bed, mattress and of other things.

Debre Berhan

The main staple is barley. In good years people mix barley with wheat, sorghum and horse beans between December and June. The hungry season is between July and November. Sometimes during holidays and some religious festivals people eat meat, chicken and eggs. Married couples eat from the same plate, while children eat separately. Investing in livestock (especially small ones) is the most common form of saving. According to respondents, the wealthiest people are those who own from 3 to 4 pairs of oxen, more than 5 cows, at least 2 horses and 2 mules, 10 sheep and those who are ready to do any kind of work. In their home one could find: an iron bed, a hand gun, a radio and kitchen utensils. Poorer households do not have oxen or land. They may be widows, war victims, old, sick etc. In their home one could find a woken bed and some kitchen utensils.

Dinki

When there is no drought people with sauce prepared from beans and chickpeas, sometimes bananas, pepper or sugar mixed with pepper as a substitute for sauce. In times of war or drought, wild food is eaten. They eat *tef* between November and May, sorghum between Decembee and April and maize between December and July. The hungry season is between July and September. In Muslims households, husbands eat alone while their wives eat together with children. The village is poorer than surrounding villages. There are no wealthy people in Dinki. This is due to the landscape, shortage of rainfall, and successive failure of the *belg* rainfall and harvest. Poor households have too many children, they cannot help others, they give land for fixed rent or they have no land at all, they do not have oxen.

Doma

Maize is the main cereal, but also sweet potato, *enset* and milk are eaten frequently. They eat bananas, milk, butter, eggs, chicken and meat during all months except March and April. The hungry season is from March to July. The difference in diet between the richer and the poorer households lies in the variety. There are three meals per day in most households, but the poorest ones. The parents and older children eat from the same plate, while younger children eat separately. The level of cash saving is low, the better off households accumulate savings in form of livestock. After 1990 when the price of cattle went up due to epidemics, people are reluctant to save in livestock. Some have begun to save in form of consumer goods such as radios or tape recorders. Only few households save through *Equbs* and the amount is very low. The majority of people are poor. The area suffered from successive droughts. The wealthiest people are those who own more livestock and grain and have more access to irrigated land. They save and they sell at the appropriate time when prices are higher. In their homes wealthy households have table, bench, chests, chair, plates, pots, radio and wooden beds etc. The poorest people in the community are those who have no oxen and no land, or no irrigated land. In their houses one would find only some of the above mentioned items.

Gara Godo

People consume a range of crops such as cassava, yam, sweet potato, soya beans, horse beans, chick peas, cow peas, bananas, meat etc. Root crops are the emergency crops. The hungry season is between February and May. Everyone in the house eats the same food at the same time. The wealthiest people in the community are those who have 1 or 2 pairs of oxen, more than 2 hectares of land, 10 or more heads of cattle, and some cash. The source of their wealth is usury, speculation (i.e. buying coffee at a low price and selling it when the price rises) and cultivation of others' land under sharecropping arrangements. The middle wealth-category includes those with one ox or a pair of oxen, 1 hectare of land, and a few head of cattle, sheep and goats. The poor households are those with no farm stock, no cattle and very small amount of land. They are not self-sufficient and they tend to have many children.

Geblen

People eat maize in September and October and barley between September and January. For the rest of the year they eat wild food if there is no money to buy food. The hungry season is between February and May. The area has been affected by all the famines that affected Ethiopia. There are very few livestock: the common ones being sheep and goats. People sell their livestock when in need of cash. There are no wealthy households in the village. The village is the poorest in the woreda because: a) of the forced evacuation to the South-West of Ethiopia during the 1984 drought; b) of unpredictable weather conditions due to vicinity to the Red Sea; c) of the soil quality (i.e. mainly stone). A self-sufficient household is one that has a pair of oxen, a medium is one with one oxen and a poor household is one with no livestock.

Imdibir

There are no emergency crops as the area never experienced crop failure due to reliance on *enset* which is eaten with butter, vegetables or lentils. They eat maize between July and September. In September they also eat butter and milk. All household's members eat the same kind of food. The hungry season is from April to August. Sheep are raised for income. The wealthy households are those with more land and cattle and those who grow cash crops. They work, but also they hire outside labourers. They migrate for trade and receive remittances from their children. The poor people in the community have a small amount of cattle and land.

Harresaw

The area has suffered from successive famines. The emergency crop is a type of barley. People also eat wild foods and roots of local plants. They eat barley and wheat (which is received in the form of aid) in all months but September and April. All members of the family eat from the same plate at the same time. People usually eat once per day. People own sheep and some goats for cash purposes. Harresaw is poorer than the surrounding villages. The very wealthy households own more than two oxen and have relatives working in Saudi Arabia. Equipment in their houses include: radio, iron beds, drinking vessels and glasses, carpets, high quality blankets, big pots and chairs. Poor households do not own land or livestock. The assets of a poor family include a traditional skin mat, low quality blankets and pottery cups.

Korodegaga

The most important crops are maize, barley and wheat. The most common food is porridge with milk. Butter and milk are eaten except between March and May. Chicken and eggs are eaten a little in April and between June and August. The size of the family determines whether members eat all together or not. There are no emergency crops and no wild food is eaten. Savings take the form of investing in livestock production. Cattle, goats and sheep are raised for cash purposes. The main source of cash for food is firewood, but people also sell their cattle to buy food. Wealth is determined by the number of livestock owned. A household with 5 or 6 hectares of land and many goats, sheep, cows, oxen etc. is very rich. Few households own a tape recorder or a radio. Poor households are characterised by lack of food or of cattle, goats and sheep. The main sources of wealth are via marriage, inheritance, hard work and economising.

Shumsheha

In this village food shortage is very common. Most people eat home-grown food for six months. For the remaining 6 months they eat imported wheat. Traditionally household's members eat from the same plate. People are very poor and eat only once or twice per day. The reasons of poverty are: not fertile land and not favourable weather conditions. Saving in the form of cash is very uncommon: most people have nothing to save and the relatively better off invest their money in livestock. The relatively wealthier households own two pair of oxen, cows, goats, sheep and donkeys. Rich peasants usually diversify crops, have fertile land and are young and strong. They have tin-roofed houses and few of them own a radio. All kitchen utensils are locally produced. Poor households do not own livestock or land at all. They are usually disabled or old.

Sirba and Godeti

Cereals and pulses are the most common crops. There are no emergency crops as the area never suffered from famine. Since there are differences in economic standing, villagers do not eat the same kind of food. Differences can be found in livestock products: meat, eggs and milk are mainly consumed by the rich. Different households have different habit: members may or may not eat from the same plate. Due to scarcity of grazing land, cattle is declining. Oxen are fattened for sale, while sheep, goats and chickens are sold only when there is need of cash. Wealth is based upon the possession of land and oxen, the ability and the desire to work hard. Stratification is discernable in the village. By village standards there are rich people and poor ones. The wealthy households own a lot of land which may be leased, have enough oxen, work hard, rent oxen, inherit land, practice trading and purchase land. In a wealthy house one would find: a radio and tape recorders, chairs, stools, tables, iron or wooden bed and mattress, sheets and blankets, vessels, bottles, glasses etc. Poor households have no land or oxen. They have a very large family, they are old or sick. In a poor home one would find: a traditional bed, tin cans for drinking, vessels.

Turufe Kecheme

People eat and produce a large range of foods: *teff*, millet, barley, maize, wheat, horse beans, sorghum, potatoes, *enset*, etc. Almost everyone in the village eats the same kind of food. Parents and older children eat from the same plate, while younger children eat separately. People keep cattle, sheep and chickens and they do not sell them even if it is profitable. Accumulation of wealth is mainly based on agriculture, but few people accumulate wealth by trading. Assets in a wealthy home may include wooden beds, table, chairs, bench, mattress, glasses, plates, a radio etc. Poor lack oxen and agricultural implements. They cannot work on others' land because they are old or sick.

Yetmen

There are no emergency crops. People eat meat, milk products and eggs in September and October. They also eat eggs in January and meat in January, February and from May to July. The hungry season is in August and September. Almost all the people eat together. People rear cattle, mules, donkeys, horse, chickens, sheep and goats. Livestock may be sold for cash needs. The wealthiest people in the community are the owners, traders, moneylenders and the skilled ones. They own large amounts of livestock. Equipments in their houses are: tables, chairs, dishes, glasses, trays and various baskets. Poor people have small amounts of livestock and they may be landless. They are usually hired to work for others for a daily wage. Poor include sick and old people, those who are unable to work and widows. In a poor house one would find: dish made of clay and other household goods made from reeds.

Savings, credit and investment

Adado

Strategies to survive crisis include saving cash, migration, marketing of *enset* and fruit and shifting cultivation between highlands and low lands. Social security is provided by the help of neighbours or/and by the following local organizations: *iddir*, *equb*, *mehber* and Peasant Association. The number of members in the *iddir* varies between 40 and 200. members in the different *iddir* pay between 0.5 to 1 *birr* a week, 1-2 *birr* for a fortnight and 2 *birr* a month. On death the following sums are paid: son or daughter - 5-10 *birr*; husband or wife - 100 *birr*; a relative - 40 *birr*. The *iddir* also provides *enset* for the funeral. There are *equb* among retail traders. They contribute between 2 and 5 *birr* a week. There is also a group called *edigret* ("development") which is like *equb*, but it is once per year.

Adele Kebe

As most people lead a subsistence life, savings and substantial investments are not affordable. People invest in livestock. Another form of savings is represented by *equb*, which are not very frequent. The main source of credit is richer neighbours who lend with interest (i.e. for a loan of 50 *birr*, the repayment is given by a quintal of sorghum whose value is 150 *birr*). While women borrow from shopkeepers and amongst themselves, men do not lend to friends and relatives.

Aze Deboa

Savings and investments are closely tied in this village. A person can get help from an *iddir* if he is a member. There are religious, village and clan *iddir*. They are useful for funerals and house-building. Members usually vary from 30 to 40 and contributions are only made in time of distress. The typical traditional women's practice is a "kembatigna wijo" (a butter *equb*). Between 4 and 10 women join together, collect a certain amount of butter (usually 1 Kg) weekly and give it to one member each week. This continues until everybody has received a share. They sell this butter and buy goats, sheep or even cattle so that the butter saved in the form of *equb* ends up in being invested. There are also *equb* which are money counterparts of women's wijo. In an *equb* any number of people join a group, collect some amount of money weekly or monthly and give it to one member at a time. This continues until everybody receives his own share. *Equb* are often differentiated according to wealth. A businessman's group may contribute more than 100 *birr* a month, while small scale traders contribute 2-5 *birr*. Even government employees such as teachers practice this form of credit. Men, women or children sometimes practice local banking. They put money in a box and break it only when they want to use the saved money. There are moneylenders in the village and they charge an interest rate of about 10% a month. People borrow money from friends and relatives.

Debre Berhan

The most common way to save is by investing in livestock particularly in small stock. Savings in the form of cash, jewellery and other form is rarely practised. In periods of good harvest, farmers try to sell part of their production to buy livestock. People have formed *iddir*, *equb* and *mehber* to help each other during crises. They serve as life and property insurance. When there is a major accident, the victim will also be assisted by people who do not belong to the same institution. The *iddir* and *equb* have their own rules as to how much to give for what types of crisis. *Iddir* is a territorially based voluntary association of peasants formed for mutual help and cooperation. The primary function is to help household members in case of death of a member, loss of property, accidents etc. For instance, the members of an *iddir* collect a fee of 25 cents monthly and 20 Kg of beans annually from each household head. Widows and divorced household heads pay half of this amount. The money and the grain is stored in the house of the treasurer for later distribution to members who faced problems according to the regulations of the *iddir*. Credit needs are seasonal: in May/June to buy seeds; from August to October for consumption and some borrow between June and August.

Moneylenders charge 10% interest rate a month until they repay. if they take cash from grain traders they repay a fixed amount of grain when they have their harvest.

Dinki

All the people in the village have a social obligation to help each other. When an oxen dies the price will be levied on the people and everyone contributes. If a house burns down people help to re-build it. There are no *iddir* or *equb* in Dinki.

Doma

The level of cash savings is low since most settlers are drought affected people. Some of the wealthier households used to save in the form of livestock, but after a disease killed most of them savings in this form have been discouraged. Some people have begun to save in consumer durable such as tape recorders and radios. Only few households belong to an *equb*. People borrow cash and grain from friends, relatives and moneylenders. Interest rate varies from 0 to 100%. Also, there are various *iddirs* to support families of the dead.

Gara Godo

There is a large variety of mutual support networks in Gara Godo. Virtually all households belong to an *iddir*. On the surface it is a burial organisation which provides support to households in times of death and funerals. However, at a closer look, the *iddir* is a multipurpose organisation. Its principal function is the mutual exchange of labour (i.e. house building). In addition, members of an *iddir* assist those who cannot cultivate the land because ill or old. Members of the *iddir* will help for a share of the harvest. *Iddir* also provides credit services: needy households have access to small loans without interest rate. *Equb* is a traditional rotative credit scheme which involves cash or production such as butter. the butter *equb* involves only women. This is because women are responsible for the management of livestock products. An *equb* acts as a local bank without any interest rate. People borrow from moneylenders for production purposes: interest rate is 100% and people need collateral. The Ministry of Agriculture provided credit of 78 birr for 75% of households in 1993 for production.

Geblen

The Ministry of Agriculture has a credit programme to encourage investment in land and trade. More than 10% of households have taken loans. Loans are provided for seed and fertiliser for those who have land, oxen, donkeys and money for potential trader chickens, sheep and goats for those who are old and ploughing tools. If a house burns down relatives help to build a new roof. Beside relatives, also *iddir* help in times of crisis. However, *iddir* are not common in the village. There are no local moneylenders.

Imdibir

In times of personal crisis people get help from relatives, neighbours, friends and *iddir*. Most people belong to an *iddir*, the number of members may range from 100 to 300 and contributions may range between 2 and 3 *birr* a month. Depending on their financial situation, *iddir* give about 1000 *birr* to people whose house burnt down. Most people belong to more than one *iddir* on the basis of clanship or neighbourhood. The contribution to an *equb* is usually between 3 and 5 *birr* a week and the number of members may range between 70 and 80. Payments are made weekly and even people from outside the PA may participate in an *equb*. The decision about which *equb* to join depends on the amount of money collected in each respective *equb* as there are rich *equb* and poor *equb*. People borrow much from moneylenders (up to 300 *birr*), friends ad relatives and *iddir*. The *iddir* often charges an interest rate of 10% per annum.

Harresaw

There is no *equb* and there are no moneylenders in the PA. People only borrow from close friends and relatives. In times of crisis, people receive help form neighbour and from *iddirs*.

Korodegaga

Most households in Korodegaga hold savings in the form of livestock. *Equbs* are usually formed by women and do not play an important role in the community. There are many *iddirs* to which people participate regardless of their socio-economic status. There are no moneylenders and people mainly borrow from friends and relatives without any interest payment. If a person loses the house all of the residents have an obligation to contribute money and the members of the *iddir* are responsible for building it again.

Shumsheha

Livestock is used for saving. Recently, *equbs* have been introduced and all households are member of a least one *iddir*. In times of crisis, people get help from friends and relatives. In case of a serious crise, people get help from the government and NGO.

Sirba Godeti

There are some *equbs*. *Equbs* are operational in the periods after the harvest. There are also *iddirs* for both males and females. *Iddirs* are based on assistance during times of disaster or mourning. They provide assistance in burial matters, material help in case of lost house. Members are usually from the same ethnic group.

Turufe Kecheme

There are people in the PA who store their savings in government banks. During the rainy season, some lend money for profit. Few peasants also invest their money in the trade of herd. There are a number of *equbs* and *iddirs*. Women's *iddir* collect butter for weddings.

Yetmen

Peasant in Yetmen invest money in buying agricultural inputs and other commodities. Seasonal needs for credit are: September/October for food; April/May for weddings; June for food, fertilisers and seeds; August for food. There are several *equbs* in the village.

Managing missing data

The mechanisms for coping with missing data can be classified according to the probability of response. Consider a simple bivariate model where Y denotes the response variable some values of which can be missing, X denotes household income and Z denotes any other set of characteristics other than Y and X. Define R_h as the probability that the hth household either responds ($R_h = 1$) or does not respond ($R_h = 0$). There are three mechanisms of missing:

1. The probability that $R_h = 0$ is independent of Y, Z and X. This is called missing completely at random (MCAR) mechanism;

2. The probability that $R_h = 0$ is independent of X, but not of (some subset of) Y and Z. This is called missing at random (MAR) mechanism;

3. The probability that $R_h = 0$ depends on X and (some subset of) Y and Z. This is called missing not at random (MNAR) mechanism.

Often, the assumptions from the missing data mechanism are not statistically testable. Most empirical works assume a MAR mechanism, whereas the MCAR is quite unrealistic.

Imputation techniques are a set of rules based on observed values for replacing missing values. These techniques can be evaluated for the extent to which they attenuate coefficient [b] and standard error [SE(b)] bias, and for the extent to which they generate accurate variances [Var(b)]. Note that the underlying assumption of these techniques is the perfect specification of the relevant model.

Below, we will briefly summarise different techniques to deal with missing data. In the ERHS we used *hotdeck* imputation as explained in the second chapter.

1. *Casewise deletion*

   This is the simplest (and most common) method. It requires any case that contains missing on one or more of its variables to be deleted. The assumptions of this technique are: i) the missing mechanism is completely at random (MCAR) or ii) the model is perfectly specified and the missing values in X are not correlated with Y. Under either of the two assumptions this method leads to unbiased coefficient estimates. Also, the coefficient standard errors will be valid for a reduced size sample.

   We have not used this method because it relies on restrictive assumptions. Firstly, it is very unlikely that missing values are completely random. Secondly, deleting cases can result in a very small sample of data remaining. This would be the case with the ERHS as it is severely affected by missing values. Thirdly, the general objection to imputation techniques (see later) is the assumption of perfect specification of the model.

2. *Mean imputation*

   This technique replaces each missing value for a given variable with the observed mean for that variable. We have not used mean imputation because it produces: i) biased coefficient estimates in linear regression models even when there is MCAR; ii) small standard errors.

3. *Mean imputation with a dummy*

   Mean imputation with a dummy is an extension of mean imputation. This technique replaces missing values with the observed mean, but it also includes a dummy variable which takes on value 1 whenever the observation is missing and 0 otherwise. On the one hand, this approach has the advantage of testing for the missing mechanism: if the dummy variable is significant then missing is not completely at random. On the other hand, it might be difficult to interpret the dummy in a regression model. Moreover, the inclusion of dummies for each variable containing missing creates problems in terms of reduced degrees of freedom. We have excluded this approach because the ERHS contains many variables with some missing values.

4. *Conditional mean imputation*

   This method replaces missing values for some variable X with means of X conditional on other variables in the data set. These means are the predicted values from a regression of X on other covariates. The resulting coefficient estimates from a linear regression model are biased but consistent. In addition, standard errors will be too small because they do not take into account the uncertainty of imputed values. Hence, we have not used this technique.

5. *Multiple imputation*

   The above described techniques do not take into account the uncertainty of imputed values. Techniques like *hotdeck* and multiple imputation introduce a random component to imputation.

   In 1987, Rubin proposed a new technique that takes into account the variability of the imputation process. This technique produces M data sets from M imputations of missing observations. The researcher then estimates the relevant model M times using each of the imputed datasets. The estimate of the kth coefficient is the average of that coefficient over M regressions [Rubin, 1987]. The resulting standard error consists of two parts: the average within- imputation (average across M regressions) and the between-imputation (difference across the M regressions):

$$SE(b) = \sqrt{\sum_{m=1}^{M} \frac{SE^2(b_b)}{M} + \frac{M+1}{M} \sum_{m=1}^{M} \frac{b_m - \bar{b}}{M-1}}. \qquad \forall \ m = 1, ..., M \tag{A.1}$$

Rubin showed that MI can be efficient even with a small number of imputations (m=3 or 5). Although multiple imputation (MI) is a very attractive technique, it places heavy demands on computers, even when using quite advanced softwares available on the internet[1]. Since the ERHS has many missing values on different variables, multiple imputation could not be performed. Hence, we have used *hotdeck* imputation.

[1]For example, *AMELIA II* is a relatively easy program to perform multiple imputation. It is available on Gary King's website.

MALAWI

Rural Microfinance Institutions in Malawi

In Malawi there are numerous rural credit programs developed both at the national and district level. The FMHFS contains data on four credit programs: i) the Malawi Rural Finance Company (MRFC), a state-owned and nationwide agricultural credit program; ii) the Malawi Mudzi Fund (MMF), a credit system similar to the Grameen Bank; iii) the Malawi Union of Savings and Credit Cooperatives (MUSCCO), a union of savings and credit associations; and iv) the Promotion of Micro-Enterprises for Rural Women (PMERW), a micro-credit program in support of income generating activities and targeted at women. All these credit programs, except for MUSCCO, are based on group lending. A more detailed analysis of the programs can be found in the IFPRI report by Diagne and Zeller (2001) from which most of the following description has been taken.

Malawi Rural Finance Company (MRFC)

The MRFC is a program funded by the World Bank. It was created after the failure of the SACA, a department of the Ministry of Agriculture that used to provide seasonal agricultural loans to smallholder farmers. Unlike the SACA, the MRFC is not dependant on the Government of Malawi and it operates under commercial principles.

The MRFC provides in-kind seasonal agricultural loans for fertilizers, seeds and pesticides for hybrid maize and tobacco. Also, it offers short-term (two-year) and medium-term (five-year) loans for farm equipment. The targeted people are jointly liable groups of 5-10 smallholder farmers.

The MRFC also offers two saving deposits to its borrowers: ordinary and contract savings accounts. With contract savings account, clients can choose the amount and timing of deposits. For honouring commitments, they can either get a bonus or earn a collateral-free loan limit.

Malawi Mudzi Fund (MMF)

The MMF was created in 1987 as a pilot credit program and was funded by the World Bank and by the International Fund for Agricultural Development (IFAD). It was designed as a replica of the Grameen Bank in Bangladesh.

The targets of the MMF were poor rural households with less than one hectare of land. It provided loans for non-farm income-generating activities in two districts of Malawi (Chiradzulu and Mangochi) during a pilot phase of five years (1990-1995) after which it was absorbed by the MRFC. Group members were individually and jointly responsible for the repayment of all loans. Most of the MMF loans were given for the sale of products (such as fish, maize, beans etc.) and other small-scale trading activities. As a consequence of high default rate among male borrowers, after two years the MMF concentrated its lending on women only.

Malawi Union of Savings and Credit Cooperatives (MUSCCO)

MUSCCO is a federation of locally based savings and credit cooperatives (SACCOs). It was financially supported by the United States Agency for International Development and created in 1980. It provides credit and saving options to low income people who do not have access to commercial banks. After having failed its original attempt to target the relatively better-off farmers in rural areas, in 1985 MUSCCO refocused its activities in urban areas.

MUSCCO members in Dowa (selected in the household survey) are relatively poor farmers who obtain loans for seasonal agricultural inputs such as fertilizers and seeds.

Promotion of Microenterprises for Rural Women (PMERW) Credit Program

The PMERW credit program was financially supported by the German Agency for Technical Cooperation (GTZ). It was started in 1986 by the Ministry of Women and Children's Affairs and Community Services (MOWCACS). The original program used to target rural poor women with less than half hectare of land in rural centres in Dedza, Mangochi, Nkhotakota and Rumphi. However, it failed because of its poor structure, management and operational problems. In 1991 with the help of a Kenyan NGO, the Undugu Society, it was designed as a group-based credit program (PMERW1).

The PMERW1 is a revolving fund operated by MOWCACS that gives two-year loans of approximately 70US$ to saving and credit clubs made of 10-15 poor entrepreneurial women who have completed training courses. In order to be eligible for the loan, the saving and credit club must have at least 60 percent of the loan deposited in a post office saving account. The MK 1,000 loan is distributed in turn to half of the club's members in smaller loans of two months' maturity not exceeding MK 300. The annual interest rate is of 30 percent. Only after the first half of the members has repaid the loan, the other half can receive their loan. This method elicits peer pressure within the group. Also, each member must have two guarantors within the group and MK 20 of savings before getting the loan. Individual loans are given for non-farm income-generating activities. After a period of two years, the full loan of MK 1,000 should be reimbursed and it should have generated enough funds to allow self-finance. Then the ministry can proceed to finance other newly formed groups.

In 1993 it was started a new program, the PMERW2. This program was financed by the MOWACACS/GTZ in cooperation with the Commercial Bank of Malawi (CBM). The PMERW2 targets women groups of 5-10 who are skilled in business activities. The structure is similar to the saving and credit clubs except that individual members can borrow up to MK 1,000 and they can receive loans directly from the CBM. Credit

Figure A.2: Surveyed sites in Malawi

members are selected among those who have excellent credit and business management skills. The loans given to credit groups by CBM are guaranteed up to 70 percent by a MOWCACS/GTZ fund maintained in an account at CBM.

Appendix B

Derivation of models

Chapter 3

Derivation of Principal components:

Principal components analysis was originally introduced by Pearson (1901) and independently by Hotelling (1933). It is a statistical technique that reduces the dimensionality of data by linearly transforming a set of correlated variables into a smaller set of uncorrelated variables. Principal components analysis is used to describe the variation in multivariate data in terms of fewer dimensions. Also, it can be used in regression analysis to address multicollinearity problems or to detect outliers.

The basic idea is quite simple: if p original variables are correlated then we can linearly transform them into a smaller subset of j components derived in a decreasing order of importance. The first principal component accounts for most of the variation in the original data. The second component accounts for most of the remaining variation subject to being uncorrelated to the first component and so on.

Algebraic derivation

This derivation has been extended from Cox (2005). Let $x = (x_1, x_2, \ldots, x_p)^{'}$ be a p dimensional random vector which can be linearly transformed by $y = a_1x_1 + a_2x_2 + \ldots + a_px_p$. The weights a_p can be represented by the vector $a = (a_1, a_2, \ldots, a_p)^{'}$ and the derived set of variables y_p are denoted by $y = (y_1, y_2, \ldots, y_p)$. Principal components linearly transforms x to y such that:

$$
\begin{array}{lll}
(a)\ \mathrm{y}_j = a_{1j}x_1 + a_{2j}x_2 + \ldots + a_{pj}x_p & \forall & j = 1, \ldots, p \\
(b)\ \text{cov}\ (y_j, y_k) = 0 & \forall & j \neq k \\
(c)\ \text{var}\ (y_1) \geq\ \text{var}\ (y_2) \geq \ldots \geq\ \text{var}\ (y_p) & &
\end{array}
$$

More formally, suppose that x has a mean of μ and that the covariance matrix is given by Σ. The first principal component y_1 can be written as:

$$y_1 = a_{11}x_1 + a_{12}x_2 + ... + a_{1p}x_p = \sum_{i=1}^{p} a_{1i}x_i = a_1^{'}x \tag{B3-1}$$

Note that $var(y_1) = var({a_1}'x) = {a_1}'\Sigma a_1 = V$.

The first principal component y_1 can be found by choosing a_1 such that:

$$\begin{aligned} &\max_{a_1} {a_1}'\Sigma a_1 \\ &\text{s.t. } {a_1}'a_1 = 1 \end{aligned} \tag{B3-2}$$

The constraint is necessary to prevent the variance to be unbundled and thus to increase up to infinite.

Setting up a standard langrangean with multiplier λ:

$$L = {a_1}'\Sigma a_1 - \lambda({a_1}'a_1 - 1)$$

$$f.o.c.$$

$$\frac{\partial L}{\partial a_1} = 2\Sigma a_1 - 2\lambda {a_1}' = 0 \tag{B3-3}$$

$$\frac{\partial L}{\partial \lambda} = {a_1}'a_1 \tag{B3-4}$$

Hence, from (B3-3):

$$(\Sigma - \lambda I)a_1 = 0 \tag{B3-5}$$

Since Σ is an n n variance-covariance matrix and λ is a scalar, it can be shown that:

(a) $(\Sigma - \lambda I)$ is a singular matrix because $\det(\Sigma - \lambda I) = 0$

(b) $(\Sigma - \lambda I)a_1 = 0$ for some nonzero eigenvector a_1

Thus the equation:

$\mid \Sigma - \lambda I \mid = 0$

has a solution if and only if λ is an eigenvalue of Σ. Since Σ is a positive-semidefinite matrix, it has p eigenvalues which are all nonnegative. Let the eigenvalues be λ_1,λ_2,λ_3,...,λ_p such that $\lambda_1 > \lambda_2 > \lambda_3 > \ldots > \lambda_p$. From equation (B3-5):

$$\begin{aligned}
\mathrm{var}(y_1) = \mathrm{var}(a_1{}'x) =& a_1{}'\Sigma a_1 \\
=& a_1{}'\lambda I a_1 \\
=& \lambda a_1{}' a_1 \\
=& \lambda
\end{aligned}$$

Hence, since y_1 has the highest variance, λ_1 is chosen to be the largest eigenvalue.

The same logic applies to the second principal component with the additional constraint that it has not to be correlated with the first component.

$$\begin{aligned}
\mathrm{cov}(y_1, y_2) =& \mathrm{cov}(a_2{}'x, a_1{}'x) \\
=& E[a_2{}'(x-\mu)a_1{}'(x-\mu)] \\
=& E[a_2{}'(x-\mu)(x-\mu)'a_1] \\
=& a_2{}'E[(x-\mu)(x-\mu)']a_1 \\
=& a_2{}'\Sigma a_1 \\
=& a_2{}'\lambda I a_1 \\
=& \lambda_1 a_2{}' a_1
\end{aligned}$$

$$=0 \qquad \text{if and only if } {a_2}' a_1 = 0$$

The second principal component y_2 can be found by choosing a_2 such that:

$$\max_{a_2} {a_2}' \Sigma a_2$$
$$\text{s.t. } {a_2}' a_2 = 1$$
$${a_2}' a_1 = 0$$

Setting up a standard langrangean with multipliers λ and γ:

$$L = {a_2}' \Sigma a_2 - \lambda({a_2}' a_2 - 1) - \gamma({a_2}' a_1 - 0)$$

f.o.c.

$$\frac{\partial L}{\partial a_2} = 2(\Sigma - \lambda I) a_2 - \gamma a_1 = 0 \tag{B3-6}$$

$$\frac{\partial L}{\partial \lambda} = {a_2}' a_2 = 0 \tag{B3-7}$$

$$\frac{\partial L}{\partial \gamma} = {a_2}' a_1 = 0 \tag{B3-8}$$

Multiplying both sides by ${a_1}'$ in (B3-6):

$$2{a_1}'(\Sigma - \lambda I) a_2 - \gamma {a_1}' a_1 = 0 \tag{B3-9}$$

and rearranging, we get:

$$2{a_1}' \Sigma a_2 - 2\lambda I {a_1}' a_2 - \gamma {a_1}' a_1 = 0$$

Note that $a_1{}'\Sigma a_2 = \mathrm{cov}(y_1, y_2) = 0; a_1{}'a_2 = 0; a_1{}'a_1 = 1$ hence, $\gamma = 0$ and (B3-9) reduces to:

$$(\Sigma - \lambda I)a_2 = 0$$

Again, this equation has a solution for any nonzero eigenvector a_2 if and only if λ is an eigenvalue of Σ.

The variance of the second component can be written as:

$$\begin{aligned}\mathrm{var}(y_2) = \mathrm{var}(a_2{}'x) =& a_2{}'\Sigma a_2 \\ =& a_2{}'\lambda I a_2 \\ =& \lambda a_2{}'a_2 \\ =& \lambda\end{aligned}$$

λ cannot be chosen to be equal to λ_1 because if it were, $a_2 = a_1$ would violate the constraint $a_2{}'a_1 = 0$. Hence, λ_2 is chosen to be the second largest eigenvalue. The same process can be repeated for all p components giving the following equations:

$$\begin{aligned}y_1 &= a_{11}x_1 + a_{21}x_2 + \cdots + a_{p1}x_p \\ y_2 &= a_{12}x_2 + a_{22}x_2 + \cdots + a_{p3}x_p \\ y_3 &= a_{13}x_3 + a_{23}x_3 + \cdots + a_{p3}x_p \\ \vdots &\quad \vdots \quad\quad \vdots \quad\quad\quad \vdots \quad\quad\quad \vdots \\ y_j &= a_{1j}x_3 + \cdots + a_{pj}x_p\end{aligned}$$

which can be written in matrix form as:

$$y = A^{'}x$$

where y is a vector of principal components, $A^{'}$ is a p p matrix of latent vectors and x is a column vector of original variables. The correspondent matrix of variance-covariance can be written as:

$$\Omega = \begin{bmatrix} \mathrm{var}(y_1) & \mathrm{cov}(y_1, y_2) & \dots & \mathrm{cov}(y_1, y_j) \\ \mathrm{cov}(y_2, y_1) & \mathrm{var}(y_1) & \dots & \mathrm{cov}(y_2, y_j) \\ \vdots & \vdots & \vdots & \vdots \\ \mathrm{cov}(y_i, y_j) & \dots & \dots & \mathrm{var}(y_j) \end{bmatrix}$$
$$= \begin{bmatrix} \lambda_1 & 0 & \dots & 0 \\ 0 & \lambda_2 & \dots & 0 \\ 0 & 0 & \lambda_3 & \dots \\ \vdots & \vdots & \vdots & \vdots \\ 0 & \dots & \dots & \lambda_p \end{bmatrix}$$

where:

$$\mathrm{var}(y) = \Lambda = \mathrm{diag}(\lambda_1, \lambda_2, \lambda_3, \dots, \lambda_p)$$

which can be written as:

$$\mathrm{var}(y) = \Lambda = A^{'}\Sigma A$$

since A is orthogonal to $A^{-1} = A^{'}$:

$$\Sigma = A\Lambda A^{'}$$

It can be shown that the sum of the variances of principal components is equal to the sum of the variances of the original variables:

$$\begin{aligned}\sum_{j=1}^{p}\text{var}(y_j) = \sum_{j=1}^{p}\lambda_j =& tr(\Lambda)\\ =& tr(A^{'}\Sigma A)\\ =& tr(\Sigma A A^{'})\\ =& tr(\Sigma)\end{aligned}$$

Note that the variance-covariance matrix is a square matrix and hence the trace can be written as the sum of its diagonal elements. Therefore:

$$\sum_{j=1}^{p}\text{var}(y_j) = \sum_{j=1}^{p}\text{var}(x_j)$$

The jth principal component accounts for $\frac{\lambda_j}{\sum_{i=1}^{p}\lambda_i}$ proportion of the total variation in the data $\sum_{j=1}^{p}\text{var}(x_j)$. Similarly, the first k principal components account for $\frac{\sum_{i=1}^{k}\lambda_i}{\sum_{i=1}^{p}\lambda_i}$ of total variation.

Looking at the correlation between principal components and original variables should help to detect which components are more important. Assume that x is a standardized variable with zero mean and variance equal to one. Also, suppose that $y = A^{'}x$ or $Ay = x$. The covariance between y_j and x_i is given by:

$$\begin{aligned}\text{cov}(y_j, x_i) =& \text{cov}(y_j, \sum_{k=1}^{p}a_{ik}y_k)\\ =& \sum_{k=1}^{p}a_{ik}\text{cov}(y_j, y_k)\end{aligned}$$

$=a_{ij}\text{var}(y_j)$ since the components are orthogonal to each other

$=a_{ij}\lambda_j$

Since the standard deviation of y_j is $\sqrt{\lambda}$:

$$\text{corr}(y_j, x_i) = a_{ij}\frac{\lambda_j}{\sqrt{\lambda_j}}$$

In a matrix form:

$$\text{corr}(y, x) = A\Lambda^{1/2}$$

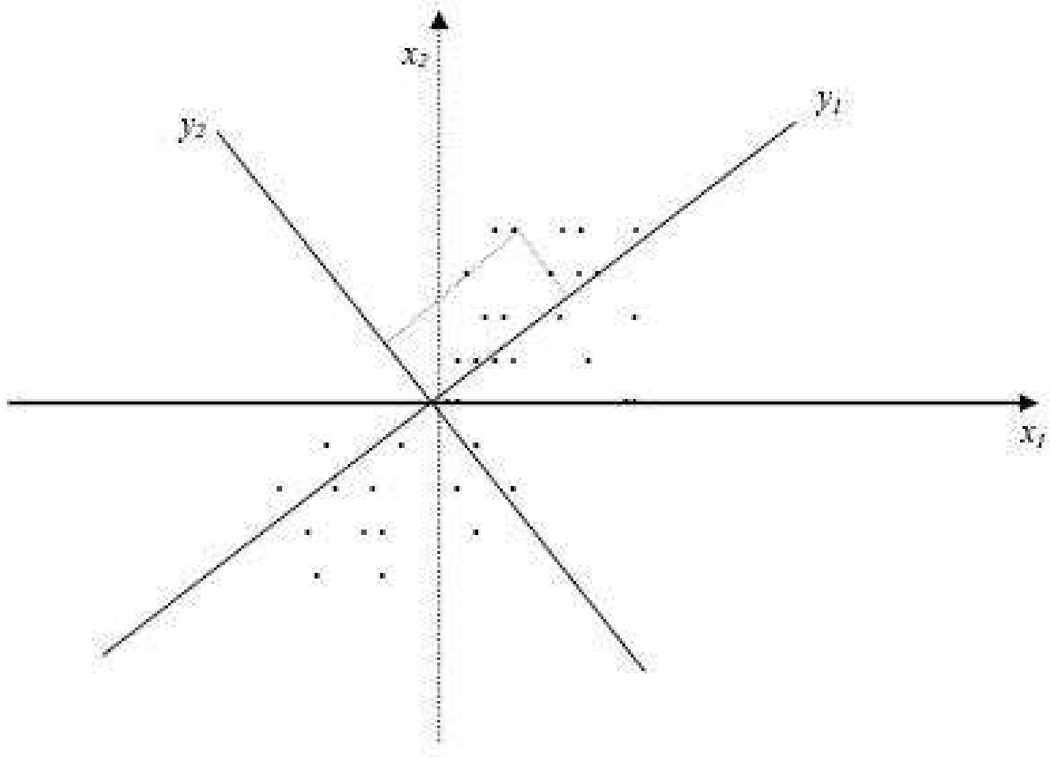
$$\text{where } \Lambda^{1/2} = \text{diag}(\sqrt{\lambda_1}, \ldots, \sqrt{\lambda_1})$$

Geometrical interpretation

It is perhaps more intuitive to look at the geometrical interpretation of principal components. In order to keep the exposition legible, consider only two dimensions. Assume that a sample of observations is characterised by two standardized variables x_1 and x_2 with a certain correlation (i.e. 0.80). Principal components can be seen as an orthogonal rotation of the orthogonal set of axes represented by the original variables. The first principal component, y_1, is a new coordinate axis oriented in a direction to maximize the variation of the projections of the points on the new coordinate axis.

The projection of the data points onto the second principal component, y_2, gives the maximum variance possible for the projected points with the additional constraint that y_2 is orthogonal to y_1. An example is given in figure B3-1.

FIGURE B3-1: Principal components



Source: Dunteman (1989).

The first problem with the principal components arises when the original variables are measured differently (i.e. local currency, kilograms etc.). If a set of multivariate data where the variables $x_1, x_2, x_3, \ldots, x_p$ are completely different is used to derive the principal components from the covariance matrix the results will depend on the different measures (as the variances will differ). Hence, in this case one can either derive the components from a correlation matrix or standardize the variables to have unit variance and zero mean [Joliffe, 2002]. The second problem is that principal components analysis relies on the normality assumption.

Application to the ERHS

We use a correlation matrix to create principal components for variables like assets (i.e. equipment, house and other assets), value of livestock, land size, number of plots, quantity of harvested crops, food and non food expenditure which are measured in different units. By looking at table B3-1, the nine variables seem to be positively correlated. For example, non-food expenditure and house assets have a correlation of 0.37. Hence, if we had to use these indicators among other factors which may affect the choice of borrowing, it would be better to use fewer component scores rather than all the nine variables.

TABLE B3-1: Correlation matrix for some asset and expenditure indicators

	Equip.	House assets	Other assets	Crops	Land size	N. of plots	Livestock	Exp.	Food exp.
Equipment^a	1.000								
House assets^a	0.335	1.000							
Other assets^a	0.275	0.214	1.000						
Crops^b	0.274	0.152	0.092	1.000					
Land size^c	0.199	0.047	0.104	0.227	1.000				
N. of plots	0.209	0.109	0.067	0.266	0.758	1.000			
Livestock^a	0.384	0.201	0.345	0.058	0.376	0.289	1.000		
Expenditure^a	0.400	0.369	0.246	0.172	0.128	0.070	0.373	1.000	
Food expenditure^a	0.315	0.236	0.214	0.067	0.180	0.169	0.309	0.358	1.000

Note:^{a}value in birr;^{b}quantity in kilograms;^{c}size in hectares. Source: own calculation from ERHS.

A similar approach[1] has been taken to create socio-economic indicators in Ethiopia by using the Demographic Health Survey. The eigenvectors (also called latent vectors) and the corresponding eigenvalues (also called latent roots) are presented in table B3-2. The correlations of the variables with the principal components, called *component loadings*, are obtained by multiplying each eigenvector with the square root of the associated eigenvalue. Component loadings are shown in table B3-3.

The first component has the largest variance of 2.961 (table B3-2) which accounts for $\left(\frac{2.961}{9} * 100\right)$ or 32.9% of the variance of the nine variables (table B3-3). The remaining components account for less variance ranging from 17% for the second principal component to 3% for the last smallest component.

[1] "Constructing socio-economic status indices: how to use principal components analysis" [Vyas and Kumaranayake, 2006].

TABLE B3-2: Latent vectors and latent roots from the correlation matrix

Principal components:	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**
Equipment	0.396	-0.180	0.193	0.052	-0.162	-0.328	-0.735	-0.309	0.045
House assets	0.291	-0.312	0.300	-0.212	0.759	0.227	-0.053	0.210	0.095
Other assets	0.279	-0.246	-0.251	0.755	0.050	0.409	0.097	-0.224	-0.016
Crops	0.231	0.160	0.750	0.264	-0.356	0.094	0.196	0.338	0.020
Land size	0.347	0.558	-0.135	-0.046	0.099	0.002	0.101	-0.163	0.708
N. of plots	0.333	0.572	-0.014	-0.087	0.184	0.127	-0.033	-0.181	-0.686
Livestock	0.399	-0.025	-0.417	0.135	-0.001	-0.437	0.029	0.669	-0.095
Expenditure	0.361	-0.334	0.066	-0.226	-0.089	0.356	0.625	-0.413	-0.086
Food exp.	0.326	-0.185	-0.224	-0.484	-0.465	0.576	-0.083	0.138	0.018
Latent root (variance)	2.961	1.552	1.018	0.828	0.716	0.659	0.577	0.463	0.227

Source: own calculation from ERHS.

How many components should be retained?

As mentioned previously, there are many criteria to decide which components should be retained and the choice between them is quite arbitrary. Nevertheless, it is useful to analyse either the latent vectors in table B3-2 or the principal component loadings in table B3-3 (since the two vectors are proportional to each other).

TABLE B3-3: Principal component loadings and percent of explained variance

Principal components:	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**
Equipment	0.681	-0.224	0.195	0.047	-0.137	-0.266	-0.558	-0.210	0.021
House assets	0.500	-0.388	0.303	-0.192	-0.642	0.184	-0.041	0.143	0.045
Other assets	0.480	-0.306	-0.253	0.686	0.043	0.332	0.074	-0.152	-0.007
Crops	0.398	0.199	0.757	0.240	-0.301	0.076	0.149	0.230	0.010
Land size	0.598	0.695	-0.136	-0.042	0.084	0.002	0.076	-0.111	0.338
N. of plots	0.574	0.712	-0.014	-0.079	0.156	0.103	-0.025	-0.124	-0.327
Livestock	0.687	-0.031	-0.421	0.123	-0.001	-0.354	0.022	0.455	-0.045
Expenditure	0.621	-0.416	0.067	-0.205	-0.075	-0.289	0.475	-0.281	-0.041
Food exp.	0.561	-0.231	-0.226	-0.441	-0.394	0.467	-0.063	0.094	0.009
% variance explained individually	32.9	17.2	11.3	9.2	8	7.3	6.4	5.1	2.5
% variance explained cumulatively	32.9	50.1	61.4	70.6	78.6	85.9	92.3	97.5	100

Source: own calculation from ERHS.

The first component has large correlation with all nine variables. Therefore, it can be interpreted as an overall wealth measure. The correlations are of about the same magnitude and they are all positive. This type of first component is usually called the *size factor*. The second principal component has high positive correlations (or large weights) with number of plots and plot size, and high negative correlations (or large weights) with non-food expenditure and value of house assets. Consequently, we can interpret the second component as a measure of contrast between farm assets and non-farm assets. The third component could be interpreted as a difference between quantity of harvested crops and value of livestock. Usually, the first components are more easily interpretable and explain most of the variance in the data.

The sum of squares of all loadings on a particular component is equal to the latent root (variance) corresponding to that component. By examining the sum of squares of the loadings for each row of the principal component loading matrix, it is possible to see how much variance for that variable is accounted for by the retained principal components. For example, the proportion of variance in equipment explained by the first three components is $0.681^2 + \left(-0.224^2 + 0.195^2\right)$ or 0.552 as shown in table B3-4.

TABLE B3-4: Proportion of variance accounted for, by first three principal components

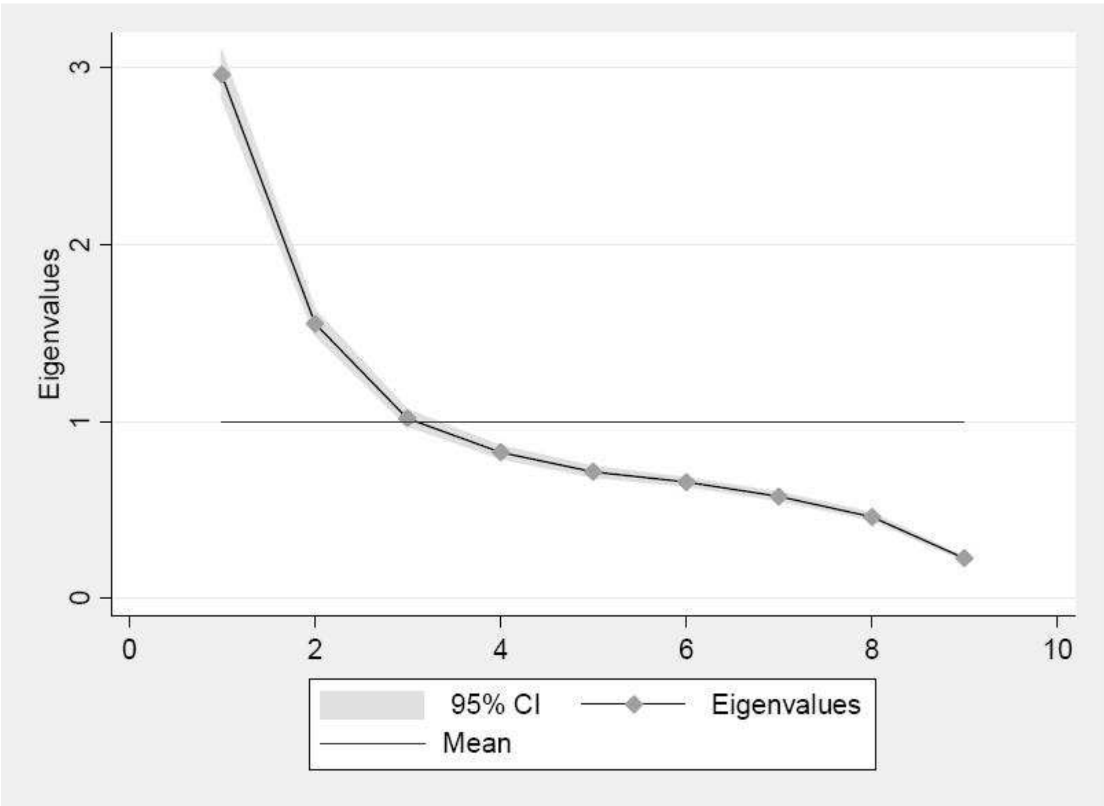
	Proportion of variance accounted for
Equipment	0.552
House assets	0.492
Other assets	0.388
Crops	0.771
Land size	0.859
N. of plots	0.837
Livestock	0.65
Expenditure	0.563
Food expenditure	0.419

Source: own calculation from ERHS.

In order to adequately represent all variables, the first three components have been retained so that the proportion of variance explained for each variable is approximately 40 percent. Indeed, the proportion of variance explained by the first three components ranges between 39 percent (other assets) and 86 percent (land size).

The scree plot of the latent roots against the latent vectors is another useful tool to decide how many components should be retained. Figure B3-2 shows a scree plot together with the mean and heteroskedastic bootstrap confidence intervals. Since principal components have been derived from a correlation matrix, the mean eigenvalue is one.

FIGURE B3-2: Scree plot of principal components



Source: own calculation from ERHS. Note: heteroskedastic bootstrap confidence intervals.

By using Cattell's (1966) scree criterion, a steep slope is evident from the first to the third latent roots and a straight line can be fitted from the third through the last component. Cattell's criterion would suggest retaining only three components. There is an obvious trade-off between interpretability of the components and adequate variance explained by the retained components. Although some variables may be under-represented in terms of explained variance[2], we have retained three components in order to gain in interpretability and reduced dimensions.

The *component scores* can be calculated for each household. Suppose that these scores can be placed in a matrix Y, so that the rth row of Y contains p component scores for the rth household. In appendix B, A has been defined as the matrix of eigenvectors, all component scores are calculated by $Y = XA$ or if the variables have been standardized:

$$Y = (X - 1\bar{x})\, A$$

for the rth household:

$$y_r = A^{'}\,(x - \bar{x})$$

component scores are showed in table B3-5.

TABLE B3-5: Components scores

Principal components:	**1**	**2**	**3**
Equipment	0.396	-0.180	0.193
House assets	0.291	-0.312	0.301
Other assets	0.279	-0.246	-0.251
Crops	0.231	0.160	0.750
Land size	0.347	0.558	-0.135
N. of plots	0.333	0.572	-0.014
Livestock	0.399	-0.025	-0.417
Expenditure	0.361	-0.334	0.066
Food expenditure	0.326	-0.186	-0.224

Source: own calculation from ERHS. Note: N. obs. 4,102.

[2]Note that the first three variables belong to a larger aggregate which can be interpreted as assets.

Empirical models

Univariate Logit model: the odds ratios

The conditional probability of a Logit model can be written as:

$$\Pr(I_i = 1 \mid x_i) = \frac{\exp({x_i}'\beta)}{1 + \exp({x_i}'\beta)}$$

where the marginal effect is:

$$\frac{\partial \Pr(I_i = 1 \mid x_i)}{\partial x_{ij}} = \frac{\exp({x_i}'\beta)}{[1 + \exp({x_i}'\beta)]^2} \times \beta_j$$

Because it is a nonlinear function it is quite difficult to interpret. So let's consider the odds ratios:

$$\begin{aligned}
\Omega =& \frac{\Pr(I_i = 1 \mid x_i)}{1 - \Pr(I_i = 1 \mid x_i)} \\
=& \frac{\Pr(I_i = 1 \mid x_i)}{\Pr(I_i = 0 \mid x_i)} \\
=& \frac{\dfrac{\exp({x_i}'\beta)}{1 + \exp({x_i}'\beta)}}{1 - \dfrac{\exp({x_i}'\beta)}{1 + \exp({x_i}'\beta)}} \\
=& \exp({x_i}'\beta)
\end{aligned}$$

So, the log of the odds ratios:

$$\begin{aligned}
\ln(\Omega) =& \ln\left\{\frac{\Pr(I_i = 1 \mid x_i)}{1 - \Pr(I_i = 1 \mid x_i)}\right\} \\
=& \ln[\exp({x_i}'\beta)] \\
=& {x_i}'\beta
\end{aligned}$$

That is, for a change in a given regressor x_i , we expect the logit to change by β_j holding other variables constant.

Since:

$$\begin{aligned}\Omega(x_i) =& \exp(x_i{}'\beta)\\ =& \exp(\beta_0)\times\exp(x_{i1}{}'\beta_1)\ldots\times\exp(x_{ik}{}'\beta_k)\ldots\times\exp(x_{ij}{}'\beta_j)\end{aligned}$$

By adding one:

$$\begin{aligned}\Omega(x_i, x_{ij}+1) =& \exp(\beta_0)\times\exp(x_{i1}{}'\beta_1)\ldots\times\exp[\beta_k(x_{ik}{}'+1)]\ldots\times\exp(x_{ij}{}'\beta_j)\\ =& \exp(\beta_0)\times\exp(x_{i1}{}'\beta_1)\ldots\times\exp(x_{ik}{}'\beta_k)\exp(\beta_k)\ldots\times\exp(x_{ij}{}'\beta_j)\end{aligned}$$

Then, the odds ratio becomes:

$$\begin{aligned}\frac{\Omega(x_i, x_{ij}+1)}{\Omega(x_i, x_{ij})} =& \frac{\exp(\beta_0)\times\exp(x_{i1}{}'\beta_1)\ldots\times\exp(x_{ik}{}'\beta_k)\exp(\beta_k)\ldots\times\exp(x_{ij}{}'\beta_j)}{\exp(\beta_0)\times\exp(x_{i1}{}'\beta_1)\ldots\times\exp(x_{ik}{}'\beta_k)\ldots\times\exp(x_{ij}{}'\beta_j)}\\ =& \exp(\beta_k)\end{aligned}$$

For a unit change in a given regressor x_i, the odds are expected to change by $\exp(\beta_k)$ holding other variables constant. So, when $\exp(\beta_k) > 1$ the odds are $\exp(\beta_k)$ times larger and when $\exp(\beta_k) < 1$ the odds are $\exp(\beta_k)$ smaller.

Chow-type test of structural change

Suppose that the groups are represented by the following equations:

$$\Pr(y = 1 \mid x) = \alpha_1 + \beta_1 x + u_i \qquad \text{if there are equbs in PA} \qquad \text{(B3-10)}$$

$$\Pr(y = 1 \mid x) = \alpha_2 + \beta_2 x + u_i \qquad \text{if there are no equbs in PA} \tag{B3-11}$$

The Chow-type test is used to test the hypothesis that $\beta_1 = \beta_2$. First, $\hat{\beta}$ is obtained form a restricted regression where β is assumed to be identical across groups:

$$\Pr(y = 1 \mid x) = \alpha + \beta x + u_i \tag{B3-12}$$

Then, the residual sum of squares of the restricted regression is compared to the residual sum of squares from the unrestricted regressions where β is allowed to vary across groups. The test statistic if given by:

$$\mathrm{F}(k+1, n_1 + n_2 - 2(k+1)) = \frac{\frac{(\text{ rrss-urss })}{(k+1)}}{\frac{\text{urss}}{(n_1 + n_2 - 2(k+1))}}$$

Another way of testing the same hypothesis is by pooling data to convert multiple equations into a single equation:

$$\Pr(y = 1 \mid x_1) = \alpha_3 + \beta_3 x + \alpha_3{}' \text{ equb} + \beta_3{}' \text{ equb} + u_i \tag{B3-13}$$

which is identical to equations (B3-10) and (B3-11):

$$\Pr(y = 1 \mid x) = (\alpha_3 + \alpha_3{}') + (\beta_3 + \beta_3{}')x + u_i \qquad \text{if equb=1} \tag{B3-14}$$

$$\Pr(y = 1 \mid x) = \alpha_3 + \beta_3 x + u_i \qquad \text{if equb=0} \tag{B3-15}$$

where:

$$(\alpha_3 + \alpha_3{}') = \alpha_1$$

$$(\beta_3 + \beta_3{}') = \beta_1$$

$$\alpha_3 = \alpha_2$$

$$\beta_3 = \beta_2$$

An F-test of $\alpha_3{}' = 0$ and $\beta_3{}' = 0$ would test whether the pooled equation (B3-12) (restricted model) is identical to equation (B3-13) (unrestricted model). This leads to exactly the same result as the Chow-type test.

Heckman Selectivity model

FIML

Suppose to have the following model:

$$y_i{}^* = x_i{}'\beta + u_i \qquad \forall i = 1, \ldots, N \tag{B3-16}$$

Representing the amount borrowed from informal sources by household i. In addition, define the latent relationship for the observability of y_i as:

$$I_i{}^* = z_i{}'\gamma + v_i \tag{B3-17}$$

where the error terms u_i and v_i have a bivariate normal distribution with covariance $\mathrm{cov}(u_i, v_i) = \sigma_{uv}$ and:

$$I_i = 1(I_i{}^* > 0) \tag{B3-18}$$

Hence, the observability criterion for the selectivity model is:

$$y_i = y_i{}^*.1(I_i{}^* > 0) \tag{B3-19}$$

For the censored observations, the probability of observing $y_i = 0$ is given by:

$$\begin{aligned}\Pr(y_i = 0) = &\ \Pr({I_i}^* < 0 \mid z_i)\\ = &\ \Pr({v_i}^* < {z_i}'\gamma)\\ = &\ \Phi(-{z_i}'\gamma)\\ = &\ 1 - \Phi({z_i}'\gamma)\end{aligned} \tag{B3-20}$$

For the uncensored observations, the conditional density can be decomposed as follows:

$$f(y_i \mid x_i, z_i, I_i = 1) = \Pr({I_i}^* > 0 \mid y_i) f(y_i) \tag{B3-21}$$

Because u_i and v_i have a bivariate normal distribution, the conditional probability can be written as:

$$\begin{aligned}\Pr({I_i}^* > 0 \mid y_i) = &\ \Pr(v_i > -{z_i}'\gamma \mid y_i)\\ = &\ \Phi\left(\frac{{z_i}'\gamma + \dfrac{\sigma_{uv}}{\sigma_u}(y_i - {x_i}'\beta)}{\sqrt{1 - \left(\dfrac{\sigma_{uv}}{\sigma_u}\right)^2}}\right)\end{aligned} \tag{B3-22}$$

Hence, the likelihood function can be derived as:

$$L = \prod_{y_i=0} \Pr(I_i = 0 \mid z_i) \prod_{y_i>0} f(y_i \mid x_i, z_i, I_i = 1)$$

Substituting equations (B3-20 to B3-22):

$$L = \prod_{y_i=0} \left[1 - \Phi({z_i}'\gamma)\right] \prod_{y_i>0} \Phi\left(\frac{{z_i}'\gamma + \dfrac{\sigma_{uv}}{\sigma_u}(y_i - {x_i}'\beta)}{\sqrt{1 - \left(\dfrac{\sigma_{uv}}{\sigma_u}\right)^2}}\right) \times \frac{1}{\sigma_u}\phi\left(\frac{y_i - {x_i}'\beta}{\sigma_u}\right) \tag{B3-23}$$

This model is mentioned in the second and fourth chapter. We did not use it in the second chapter because the Wald test did not allow us to reject the hypothesis of independent equations.

Two-step Heckman model

Since FIML models can be computationally heavy, Heckman proposed a two-step estimator. Like in the truncated models the conditional expectation can be written as:

$$E(y_i \mid y_i > 0) = x_i{}' \beta + E(u_i \mid y_i > 0)$$

considering the selection process in equation (B3-17):

$$\begin{aligned} E(y_i \mid y_i > 0) &= E(y_i \mid I_i{}^* > 0) \\ &= x_i{}' \beta + E(y_i \mid I_i{}^* > 0) \\ &= x_i{}' \beta + E(u_i \mid v_i > -z_i{}' \gamma) \\ &= x_i{}' \beta + \left(\frac{\sigma_{uv}}{\sigma_u} \frac{\phi(z_i{}' \gamma)}{\Phi(z_i{}' \gamma)} \right) \\ &= x_i{}' \beta + \left(\frac{\sigma_{uv}}{\sigma_u} \lambda(z_i{}' \gamma) \right) \end{aligned}$$

where $\lambda(z_i{}' \gamma)$ is the hazard rate or Mills' ratio. The two-stage method estimates the entire sample by probit and gets $\hat{\gamma}$, then uses these estimates to construct the hazard rate. Finally, it regresses y_i on x_i and the hazard rate by OLS for the sub-sample of non-censored observations.

Chapter 4

Choice-based sampling weights

Suppose that $j = 1, 2, ..., J$ is the set of alternatives that defines the stratification[3] of the sample. In the MRFMHFS, $J = 4$ indicate four mutually exclusive choices: membership only into the MRFC[4], membership into a second programme, past membership and non membership. Exogenous sampling is characterized by the following likelihood function [Amemiya, 1985]:

$$L_E = \prod_{i=1}^{N} P\left(j_i|x_i, \beta\right) g(x) \tag{B4-24}$$

where $i = 1, 2, ..., N$ indicates the sample of N households, x is the exogenous variable and *g(x)* is the density function according to which the researcher draws *x*.

Manski and Lerman (1977) showed that using the above defined likelihood function in a choice-based sample would produce inconsistent estimates. Hence, they proposed a weighted likelihood function[5] (WMLE). In this method the likelihood function is given by:

$$L_C = \prod_{i=1}^{N} P\left(j_i|x_i, \beta\right) f(x_i) \frac{H(j_i)}{Q(j_i|\beta_0)} \tag{B4-25}$$

where $f(x_i)$ is the true density of x. Note that if $g(x) = f(x)$ and $H(j) = Q(j|\beta_0)$ [Amemiya, 1985], then B4-25 becomes the standard likelihood function under random sampling:

$$L_R = \prod_{i=1}^{N} P\left(j_i|x_i, \beta\right) f(x_i) \tag{B4-26}$$

[3]As mentioned in chapter 3, the sample has been stratified according to programme membership status.

[4]The credit programmes are described in Appendix A.

[5]Manski and Lerman (1977) showed the WMLE to be consistent and asymptotically distributed with a normal distribution.

Manski and McFadden (1981) proposed another estimator that only requires the knowledge of Q and of the sampling distribution H. Amemiya and Vuong (1987) showed that the Manski-McFadden estimator (MME) is asymptotically more efficient than the Manski-Lerman weighted maximum likelihood estimator (WMLE). The MME is defined to be the value of β that maximizes:

$$L_C = \prod_{i=1}^{N} \frac{P\left(j_i|x_i,\beta\right)\dfrac{H(j_i)}{Q(j_i)}}{\sum_{j=1}^{J} P\left(j|x_i,\beta\right)\dfrac{H(j)}{Q(j)}} \tag{B4-27}$$

Cosslett (1981) showed that a more efficient estimator can be obtained by replacing H and Q with the sample and population frequencies where $Q\left(j_i|\beta_0\right) = \dfrac{N_j}{N}$ is known and represents the decision-making population selecting the jth alternative. $H(j_i) = \dfrac{n_j}{n}$ is the choice-based sampling ratio; N_j is the size of the population defined by programme j and n_j is the size of the sample stratum; n and N are the total sample and population sizes, respectively. This is exactly the probability weight used in the MRFMHFS where the population frequencies have been obtained by the village census conducted prior to the survey.

The CIA with propensity scores

Lechner (1999a) showed a generalization of the CIA given by:

$$\begin{aligned} &Q^k \coprod T^k | b^k(X) = b^k(x), \qquad \forall\ x \in \chi, \\ &if\ \ E\left[P^k(x)|b^k(x)\right] = P^k(x), \qquad 0 < P^k(x) < 1, k = m, l \end{aligned} \tag{4.4}$$

Lechner (1999a) also showed that the average effect[6] can be written as follows:

$$\begin{aligned}
\gamma_0^{m,l} =& E\left(Q^m - Q^l\right) = E\left(Q^m|T=m\right)P\left(T=m\right) + E\left(Q^m|T\neq m\right)P\left(T\neq m\right)\\
&- E\left(Q^l|T=l\right)P\left(T=l\right) + E\left(Q^l|T\neq l\right)P\left(T\neq l\right)\\
=& E\left(Q^m|T=m\right)P\left(T=m\right) + E_{p^m(X)}\left[E\left(Q^m|P^m(X),T=m\right)|T\neq m\right]P\left(T\neq m\right)\\
&- E\left(Q^l|T=l\right)P\left(T=l\right) + E_{p^l(X)}\left[E\left(Q^l|P^l(X),T=l\right)|T\neq l\right]P\left(T\neq l\right)
\end{aligned} \tag{4.5a}$$

and:

$$\begin{aligned}
\vartheta_0^{m,l} =& E\left[Q^m - Q^l|T=m\right] = E\left[Q^m|T=m\right] - E\left[Q^l|T=m\right]\\
=& E\left[Q^m|T=m\right] - E_{p^{l|ml}(X)}\left[E\left(Q^l|P^{l|ml}(X),T=l\right)|T=m\right]\\
&\text{where } P^{l|ml}(x) = P^{l|ml}\left(T=l|T\in l,m,X=x\right) = \frac{P^l(x)}{P^l(x)+P^m(x)}
\end{aligned} \tag{4.5b}$$

Independence of Irrelevant Alternatives (IIA)

Hausman and McFadden (1984) proposed a Hausman-type test to check whether the IIA property is violated. In order to understand this property let's consider an example.

Suppose a consumer initially has to choose between two transport modes, a car and a red bus, with equal probability 0.5 so that the ratio between the two choices is one. Then, a third alternative, a blue bus, is added. The new alternative is irrelevant because the consumer is assumed to be indifferent between the colours of the two buses, hence, the consumer will choose between them with equal probability. But the IIA implies that the probability of each transport mode is 1/3 therefore the probability of choosing car would fall from 1/2 to 1/3 which is unreasonable.

[6] The effects apply the law of iterated expectations: $E(Q) = E_X[E(Q|X)]$.

In more formal terms, since the MNL implies that:

$$\Pr(y_i = m \mid x_i) = \frac{\exp(x^{'}\beta_m)}{\sum_{j=1}^{3}\exp(x^{'}\beta_j)} \qquad \forall \text{ m} = 1, 2, 3$$

So, the ratio between the probabilities of any two choices:

$$\frac{\Pr(y_i = 1 \mid x_i)}{\Pr(y_i = 2 \mid x_i)} = \frac{\exp(x^{'}\beta_1)}{\exp(x^{'}\beta_2)}$$

is independent of the probability of any other outcome. As mentioned above, adding a new alternative (i.e. blue bus) leaves this ratio unchanged.

Small-Hsiao test

The Small-Hsiao test is similar to the split sample Chow test, but uses the log-likelihood values rather than the residual sum of squares. The test is constructed as follows: estimate the model over the full set of outcomes and obtain the log-likelihood value L_f. This model is (1). Arbitrarily exclude one of the categories and re-estimate the model and obtain the log-likelihood value L_c. This is model (2). Then, artificially stack the data for model (1) above the data for model (2) and re-estimate using the expanded set of observations. Define the resultant log-likelihood value as L_s. The Small-Hsiao test is computed as a likelihood ratio test:

$$\text{Small-Hsiao} = -2[L_s - (L_f + L_c)] \sim \chi^2_{(q)}$$

where q is the number of parameters estimated in the given category.

Chapter 5

Tobit model

Consider a model where the demand[7] by household i is a function of the credit limit and of other household's characteristics which we denote for simplicity with the variable x_i:

$$y_i{}^* = x_i{}' \beta + u_i \qquad \text{with } u_i \sim N(0, \sigma^2)$$

We cannot observe y_i^*, but only y_i which is censored on the left. In other words, we observe the demand only of those who apply for the loan, for those who do not apply we observe zero:

$$y_i = \max\{0, y_i^*\}$$

Firstly, let us derive the probability that we observe a zero demand (i.e. our variable is censored):

$$\begin{aligned}
\Pr(y_i = 0 \mid x_i) &= \Pr(y_i^* \leq 0 \mid x_i) \\
&= \Pr(u_i \leq -x_i' \beta) \\
&= \Phi(-z_i) \\
&= 1 - \Phi(z_i)
\end{aligned}$$

where the standardised variable $z_i = \dfrac{x_i' \beta}{\sigma}$.

Secondly, the density function for uncensored observations can be written as:

$$f(u_i \mid x_i) = \frac{1}{\sigma} \phi \left(\frac{u_i}{\sigma} \right)$$

[7]Here we use the variable y to denote demand but it is the same as D in chapter 5.

$$= \frac{1}{\sigma}\phi\left(\frac{y_i - x_i^{'}\beta}{\sigma}\right)$$

Hence, the likelihood function can be written as:

$$l_i(x_i;\beta,\sigma) = 1.(y_i = 0)\ \ln[1 - \Phi\left(\frac{x_i^{'}\beta}{\sigma}\right)] + 1.(y_i > 0)\ \ln\left[\frac{1}{\sigma}\phi\left(\frac{y_i - x_i^{'}\beta}{\sigma}\right)\right]$$

After setting d equal to one when $y = 0$ and equal to zero otherwise, we can write the sample likelihood function as:

$$L_N(\beta,\sigma) = \sum_{i=1}^{N}\left\{d_i\ln\left[1 - \Phi\left(\frac{x_i^{'}\beta}{\sigma}\right)\right] + (1 - d_i)\left[\ln\phi\left(\frac{y_i - x_i^{'}\beta}{\sigma}\right) - \ln\sigma\right]\right\}$$

The first order condition with respect to β is cumbersome to derive:

$$\frac{\partial L_N}{\partial\beta} = \sum_{i=1}^{N}\left\{-\frac{\frac{x_i^{'}}{\sigma}\phi\left(\frac{x_i^{'}\beta}{\sigma}\right)d_i}{1 - \Phi\left(\frac{x_i^{'}\beta}{\sigma}\right)} + \frac{\frac{x_i^{'}}{\sigma}\phi^{'}\left(\frac{y_i - x_i^{'}\beta}{\sigma}\right)(1 - d_i)}{\phi\left(\frac{y_i - x_i^{'}\beta}{\sigma}\right)}\right\}$$

$$= \sum_{i=1}^{N}\left\{(1 - d_i)\left[\frac{y_i - x_i^{'}\beta}{\sigma^2}\right]x_i - d_i\frac{\phi\left(\frac{x_i^{'}\beta}{\sigma}\right)}{1 - \Phi\left(\frac{x_i^{'}\beta}{\sigma}\right)}x_i\right\}$$

Marginal effects

In order to calculate the marginal effects, let us write the expectations for the uncensored and censored dependent variable.

$$E(y_i \mid x_i, y_i > 0) = x_i^{'}\beta + E(u_i \mid x_i, y_i > 0)$$

$$= x_i^{'}\beta + E(u_i \mid u_i > -x_i^{'}\beta)$$

$$= x_i^{'}\beta + \frac{\sigma\phi(-z_i)}{1 - \Phi(-z_i)}$$

$$
\begin{aligned}
&= x_i'\beta + \sigma\frac{\phi(z_i)}{\Phi(z_i)} \\
&= x_i'\beta + \sigma\lambda(z_i) \\
&= x_i'\beta + \sigma\lambda\left(\frac{x_i'\beta}{\sigma}\right)
\end{aligned}
$$

and the expected value of the observed demand is given by:

$$
\begin{aligned}
E(y_i \mid x_i) &= \ \mathrm{P}(y_i = 0 \mid x_i)0 + \ \mathrm{P}(y_i > 0 \mid x_i)E(y_i \mid x_i, y_i > 0) \\
&= \Phi\left(\frac{x_i'\beta}{\sigma}\right)\left[x_i'\beta + \sigma\lambda\left(\frac{x_i'\beta}{\sigma}\right)\right] \\
&= \Phi\left(\frac{x_i'\beta}{\sigma}\right)\left[x_i'\beta + \sigma\frac{\phi\left(\frac{x_i'\beta}{\sigma}\right)}{\Phi\left(\frac{x_i'\beta}{\sigma}\right)}\right] \\
&= \Phi\left(\frac{x_i'\beta}{\sigma}\right)x_i'\beta + \sigma\phi\left(\frac{x_i'\beta}{\sigma}\right)
\end{aligned}
$$

Hence the marginal effects are given by:

$$
\begin{aligned}
\frac{\partial(E(y_i \mid x_i, y_i > 0)}{\partial x_i} &= \beta + \sigma\frac{\partial\lambda}{\partial x_i} \\
&= \beta + \left\{1 - \left[\frac{x_i'\beta}{\sigma}\right]\lambda\left(\frac{x_i'\beta}{\sigma}\right) - \left[\lambda\left(\frac{x_i'\beta}{\sigma}\right)\right]^2\right\}
\end{aligned}
$$

and:

$$
\begin{aligned}
\frac{\partial(E(y_i \mid x_i)}{\partial x_i} &= \beta\Phi\left(\frac{x_i'\beta}{\sigma}\right) + (x_i'\beta)\phi\left(\frac{x_i'\beta}{\sigma}\right) - \frac{x_i'\beta}{\sigma}\phi\left(\frac{x_i'\beta}{\sigma}\right)\sigma \\
&= \beta\Phi\left(\frac{x_i'\beta}{\sigma}\right) \\
&= \beta\ \mathrm{P}(y_i > 0 \mid x_i)
\end{aligned}
$$

Quantile censored regression models

From equation 5.14 in chapter five, we can write that $\text{Med}\left(D_i^{k*} \mid x_i\right) = x_i\beta$, so that the median of D^{k*} is linear in x. Notice that for any non-decreasing function $g\left(D_i^k\right)$, the $\text{Med}\left[g\left(D_i^k\right)\right]$
$= g\left[\text{Med}\left(D_k^i\right)\right]$ [Wooldridge, 2002]. Then, given that $D_i^k = \max\left(0, D_i^{k*}\right)$ is a non decreasing function, we can write:

$$\text{Med}\left(D_i^k \mid x_i\right) = \max\left[0, \text{Med}\left(D_i^{k*} \mid x_i\right)\right] = \max(0, x_i\beta)$$

where *k=formal, informal* as described above. The estimator suggested by Powell (1984) for the quantile censored model can be derived as:

$$\min_{\beta} \sum_{i=1}^{N} \left| D_i^k - \max(0, x_i\beta) \right|$$

which can be equivalently written as follows [Pagan and Ullah, 1999]:

$$\min_{\beta} \sum_{i=1}^{N} I(x_i\beta > 0) \left| D_i^k - \max(0, x_i\beta) \right|$$

since as a consequence of censoring, for observations $x_i \leq 0$, $\max(0, x_i\beta) = 0$ and $\left|D_i^k - \max(0, x_i\beta)\right| = \left|D_i^k\right|$ is not a function of β. Hence, we can minimize $\sum \left|D_i^k - x_i\beta\right|$ using only observations for which $x_i\beta > 0$.

Proof: marginal effect of credit limit when only its expected value is observed

This section follows Diagne (1999) by showing that the marginal effects can be estimated even when we observe only the expected value of the credit limit.

The marginal effect in a standard regression is the quantity $\frac{\partial f(L)}{\partial L}$ where L is the credit limit. This quantity is different from $\frac{\partial E[f(L)]}{\partial L}$ which is the change in the expected value of the dependent variable with respect to a random variable L.

Suppose that the actual demand of household i depends on the expected value of the credit limit as stated in chapter 5. That is:

$$\mathrm{Q}^* = f(L)$$

where:

$$f(L) = \beta EL$$

The coefficient β measures the marginal effect and it is the coefficient to be estimated in the regression.

Note that because the expectation operator is a linear function:

$$\frac{\partial f(L)}{\partial L} = \beta E$$

For all random variables h:

$$\frac{\partial f(L)}{\partial L(h)} = \beta Eh$$

So, for a marginal change in L of size dL:

$$\frac{\partial f(L)}{\partial L(dL)} = \beta EdL$$

Indeed, β is equal to $\frac{\partial f(L)}{\partial L}$ and it can be interpreted as the expected change in Q^* following an expected change in L.

Appendix C

Additional results

CHAPTER 3

Table C3-1: Overview of the Peasant Associations

P.A.	Location	Wealth	Background	Mean Rainfall (mm)	Main crops	Infrastructure
Adado	*Separ*	Mixed	Rich coffee producing area	1417	Coffee, Enset	The nearest big town is Dila (23Km) on dry weather road
Adele Kebe	*Oromiya*	Rich	Highland site. Drought in 1985/1986	748	Millet, maize, coffee, chat	PA linked to Dire Dawa, Alemaya (7 Km) and Haar by road
Azr Deboa	*Separ*	Mixed; Migration dependent	Densely populated. Long tradition of substantial seasonal and temporary migration	1509	Enset, coffee, maize, teff, sorghum	An all weather road links Azr Deboa to Durame (4Km) and Hosaina
Debre Birhan	*Amhara*	Usually self-supporting	Highland site. Near town.	919	Teff, barley, beans	4 PAs in the vicinity of Debre Birhan
Dinki	*Amhara*	Vulnerable to famine	Badly affected from famine 1984/1985; not easily accessible even thought near Debre Berhan.	1664	Millet, teff	*Wereda* capital
Dooma	*Separ*	Vulnerable to famine	Resettlement area (1985); Semi-arid; droughts in 1985,1989,1990; remote.	1150	Enset, maize	Nearest town of Wacha is 20 minutes walk
Gara Godo	*Separ*	Vulnerable to famine	Densely packed enset-farming area. Famine in 1983/1984. Malaria in mid 1988.	1245	Barley, enset	Densely populated area
Geblen	*Tigray*	Vulnerable to famine	Poor and vulnerable area; used to be quite wealthy.	504	Cereals	3 hours walk from Adigrat
Haresaw	*Tigray*	Vulnerable to famine	Poor and vulnerable area.	558	Cereals	There is a dry weather road to Atsbi(1 and 1/2 hours on foot)
Imdibir	*Separ*	Migration	Densely populated enset area	2205	Enset, chat, coffee,	The PA is on the all

		dependent			maize	weather road between Hosaina and Wolkite
Korodegaga	*Oromiya*	Vulnerable to famine	Poor cropping area in neighbourhood of rich valley.	874	Cereals	PA linked to Dheera by dirt road and Awash Malkaasa by a raft over the Awash river
Shumsha	*Amhara*	Vulnerable to famine	Poor area in neighbourhood of airport near Lalibela.	654	Cereals	The PA is on a dry weather road from Lalibela to Woldia, near airport and new all weather road
Sirbana G.	*Oromiya*	Rich	Near Debre Zeit. Rich area. Much targeted by agricultural policy. Cereal, ox-plough system.	672	Teff	PA next to main road to Debre Zeit (1 hour walk)
Turufe Ketchema	*Oromiya*	Rich	Near Shashemene. Ox-plough, rich cereal area. Highlands.	812	Wheat, barley, teff, potatoes	PA linked to Shashemene by 10 Km all weather road and 2-4 Km dry weather road
Yetmen	*Amhara*	Rich	Near Bichema. Ox-plough, cereal farming system of highlands.	1241	Teff, wheat and beans	PA linked to Bichena (15Km) and Dejen all weather road and 2-4 (17Km) by all weather road

Source: Community survey ERHS, Webb and von Braun (1994), Bevan and Pankhurst (1996). Note: Peasant Association is the smallest unit of aggregation in Ethiopia, an administrative unit of one or a small number of villages.

Table C3-2: Timing of the surveys and activities

Survey site	**Main Harvest**	Survey Round			
	Time of interview	*Round 1 (1994)*	*Round 2 (1994-1995)*	*Round 3 (1995)*	*Round 4 (1997)*
Haresaw	October-November	June-July	January	March	June
Geblen	October-November	June-July	January	March	June
Dinki	December	March-April	November	January	October,November
Debre Berhan	November-December	March-April	October	March	June-August
Yetmen	November-December	March-April	October	March	September,October
Shumsha	October-December	June-July	December-January	May	October,November
Sirbana Godeti	November-December	March-April	November	March	June,July
Adele Kebe	November-December	May-June	October	April	October,November
Koro-Degaga	October-November	May-June	November-December	May-June	June,July
Turufe Ketchema	December	March-April	September-October	March-April	September, October
Imdibir	October-December	March-April	October	March	June, July
Azr Deboa	October-November	March-April	September-October	March	September, October
Adado	December-January	March-April	January	March	June, July
Gara Godo	August-December	March-May	October	March	June, July
Doma	September-December	April-May	December-January	May-June	November

Source: Community survey ERHS and Bevan and Pankhurst (1996).

TABLE C3-3: Fisher Index by Peasant Association

Peasant Associations	**Fisher Index (1997 base year)**		
	Round 1: 1994 (Jan-Mar)	**Round 2: 1994 (Aug-Oct)**	**Round 3: 1995 (Jan-Mar)**
Haresaw	0.60	0.75	0.61
Geblen	0.34	0.29	0.38
Dinki	1.35	2.21	1.38
Yetmen	0.43	0.64	0.50
Shumsha	0.83	1.25	1.58
Sirbana Godeti	0.87	1.89	1.30
Adele Kebe	1.32	1.31	1.42
Korodegaga	0.54	0.20	0.69
Turufe Ketchema	1.45	1.27	1.12
Imdibir	0.77	1.28	0.88
Azr Deboa	0.45	0.82	0.35
Adado	0.55	0.81	0.82
Gara Godo	0.82	1.77	0.89
Dooma	0.80	1.27	0.84
Debre Birhan	0.96	1.35	0.83

Source: own calculation from ERHS.

TABLE C3-4: Health of children by region

child < 7 with a lot of difficulty to ...	**Tigray**	**Amhara**	**Oromiya**	**Separ**
stand up	5.4	18.3	51.5	24.9
sweep the floor	3.8	20.6	54.1	21.5
walk for 5 Km.	4.1	23.6	49.0	23.3
carry 20 Lt. of water for 20 Mt.	3.1	31.6	37.2	28.1
hoe a field for a morning	2.8	32.6	38.0	26.6

Source: Own calculation from ERHS. Note:% of borrowing households having at least one child who is not healthy.

TABLE C3-5: Health of households' members by region

	N. of ill HH members		**N. of days of no work**		**Expenditure in medicines**	
Regions	*N. obs.*	*Mean (std. dev.)*	*N. obs.*	*Mean (std. dev.)*	*N. obs.*	*Mean (std. dev.)*
Tigray	81	1.4 (1)	66	19.3 (10)	76	4.1 (10)
Amhara	211	1.5 (1)	190	11.6 (8)	209	8.8 (28)
Oromiya	392	1.8 (1)	344	10.4 (9)	375	16.9 (36)
Separ	540	1.9 (1)	495	11.1 (8)	534	7.7 (17)

Source: Own calculation from ERHS. All values refer to a period of four weeks. Expenditure in birr (1 birr=0.1143 U.S.$), deflated by using the Fisher Index (1997 base year) and square root of household size.

Table C3-6: Multinomial regressions by loan type

Pr(source of credit...)	**Formal only**	**Informal only**
	All loans	
Land size at t-1 (ha)	1.31	0.91
	*(0.15)***	*(0.09)*
round 2	0.57	0.90
	(0.96)	*(0.82)*
round 4	1.40	0.06
	(1.76)	*(0.05)*
N. obs.	1585	
pseudo-R^2	0.17	
	Production loans	
Land size at t-1 (ha)	1.48	0.95
	*(0.24)***	*(0.14)*
round 3	2.26	9.21
	(0.80)	*(2.40)*
N. obs.	431	
pseudo-R^2	0.11	
	Other loans	
Land size at t-1 (ha)	1.24	1.04
	(0.29)	*(0.18)*
round 4	0.49	0.11
	(0.67)	*(0.06)*
N. obs.	788	
pseudo-R^2	0.09	

Source: own calculation from ERHS. Note: odds ratio displayed and robust std. errors in ().
***$p < 0.01$,**$p < 0.05$, *$p < 0.1$. Reference outcome: both formal and informal credit.

Table C3-7: Logit models - Standard and principal components regression

Pr(informal)	**Model I: std. regression**	**Model II: pca regression**
hh characteristics:		
age head	0.97	0.96
	(0.07)	*(0.06)*
age head squared	1.00	1.00
	(0.00)	*(0.00)*
hh size	1.03	1.05
	(0.28)	*(0.27)*
hh size squared	1.01	1.01
	(0.02)	*(0.02)*
female head	0.77	0.74
	(0.29)	*(0.28)*
n. children 0-5	0.90	0.91
	(0.12)	*(0.11)*
n. children 6-10	0.97	0.97
	(0.24)	*(0.21)*
n. children 11-17	0.73*	0.76
	(0.13)	*(0.13)*
head schooling	0.53*	0.55*
	(0.20)	*(0.20)*
head ethnic minority	2.78	2.16
	(3.72)	*(2.76)*
assets and expenditure:		
assets & exp. (pc1)	-	0.88
		(0.10)
assets & exp. (pc2)	-	0.90
		(0.11)
assets & exp. (pc3)	-	0.70***
		(0.06)
equipment	1.00	-
	(0.00)	
house assets	1.00	-
	(0.00)	
other assets	1.00	-
	(0.00)	
non-food expenditure	1.00	-
	(0.00)	
food expenditure	1.00	-
	(0.00)	
land size	0.79***	-
	(0.06)	
n. plots	1.08	-
	(0.06)	
harvested crops	1.00***	-
	(0.00)	
livestock	1.00	-
	(0.00)	
HH shocks:		
household only	1.57	1.56
	(0.59)	*(0.57)*
land lost for disputes with rel.	0.02**	0.02**
	(0.04)	*(0.04)*
ill husband	0.39**	0.36**
	(0.18)	*(0.17)*
son left	0.24***	0.24***
	(0.10)	*(0.11)*
diseases	0.70	0.63*

	(0.17)	*(0.15)*
destruction of house	0.22***	0.19***
	(0.10)	*(0.09)*
Haresaw	0.01***	0.01***
	(0.00)	*(0.00)*
Geblen	0.004***	0.004***
	(0.00)	*(0.00)*
Dinki	0.05*	0.06*
	(0.08)	*(0.08)*
Yetmen	0.50	0.54
	(0.42)	*(0.37)*
Adele Kebe	0.29	0.32
	(0.22)	*(0.22)*
Korodegaga	0.08***	0.07***
	(0.06)	*(0.04)*
Imdibir	0.26*	0.39
	(0.19)	*(0.25)*
Dooma	0.19**	0.16***
	(0.15)	*(0.10)*
Debre Birhan	0.12**	0.12***
	(0.10)	*(0.08)*
round 1	0.06**	0.06**
	(0.08)	*(0.09)*
round 2	0.27	0.31
	(0.42)	*(0.49)*
round 4	0.06**	0.06**
	(0.07)	*(0.08)*
Observations	1,144	1,144
pseudo-R^2	0.36	0.35

Source: own calculation from ERHS. Note: coefficients in e^{β} form. Standard errors adjusted for within cluster correlation and inclusion of PCs in (). Land size in hectares, assets and expenditure in local currency deflated by using Fisher index, quantity of crops in kilograms. $^{***}p < 0.01$,$^{**}p < 0.05$,$^{*}p < 0.1$

Table C3-8: Logit models - Test for structural change

Pr(informal)	**Model I: Equbs in PA**	**Model II: No Equbs in PA**
hh characteristics:		
age head	1.01	0.91
	(0.11)	*(0.08)*
age head squared	1.00	1.00
	(0.00)	*(0.00)*
hh size	0.78	1.22
	(0.29)	*(0.43)*
hh size squared	1.02	0.99
	(0.02)	*(0.02)*
female head	0.80	0.60
	(0.42)	*(0.37)*
n. children	0.75	1.05
	(0.14)	*(0.25)*
head schooling	0.82	0.18**
	(0.41)	*(0.15)*
PCs of hh assets:		
assets & exp. (pc1)	0.88	0.79
	(0.14)	*(0.28)*
assets & exp. (pc2)	0.93	0.54
	(0.19)	*(0.30)*
assets & exp. (pc3)	0.80	0.85
	(0.19)	*(0.32)*
Shocks:		
household only	1.56	1.56
	(0.77)	*(0.79)*
land lost for disputes with rel.	0.03**	-
	(0.05)	
ill husband	0.49	0.24**
	(0.28)	*(0.16)*
son left	-	0.09***
		(0.08)
diseases	0.44	0.61
	(0.24)	*(0.56)*
house destroyed	0.11**	0.42
	(0.09)	*(0.66)*
Haresaw	-	0.03***
		(0.02)
Geblen	-	0.02***
		(0.02)
Yetmen	0.94	-
	(1.63)	
Adele Kebe	0.51	-
	(0.69)	
Korodegaga	0.08**	-
	(0.10)	
Imdibir	0.48	-
	(0.67)	
Dooma	0.27	-
	(0.36)	
Debre Birhan	0.08*	-
	(0.12)	
round 1	0.22	0.0004***
	(0.24)	*(0.00)*
round 2	1.23	0.0003***
	(1.84)	*(0.00)*
round 4	0.06***	0.0001***

	(0.06)	*(0.00)*
N. of observations	916	218
pseudo-R^2	0.26	0.32
Likelihood ratio test	LR - χ^2	$Prob. > \chi^2$
	29.57	$(0.02^{**})^{\dagger}$

Source: own calculation from ERHS. Note: coefficients in e^{β} form. Standard errors adjusted for within cluster correlation in (). $^{***}p < 0.01$,$^{**}p < 0.05$,$^{*}p < 0.1$ † p-value in ().

TABLE C3-9: Selectivity models - FIML (PA has Equbs)

	Model I		Model II	
	Pr(Informal)	**Log(informal credit)**	**Pr(Informal)**	**Log(informal credit)**
hh characteristics:				
age head	0.01	0.02	0.01	0.03
	(0.02)	*(0.01)*	*(0.02)*	*(0.01)**
age head squared	-0.0002	-0.0002	-0.0002	-0.0003
	(0.00)	*(0.00)**	*(0.00)*	*(0.00)**
hh size	0.32	0.03	0.35	0.03
	*(0.05)****	*(0.03)*	*(0.05)****	*(0.04)*
hh size squared	-0.004	-0.001	-0.01	-0.002
	*(0.00)***	*(0.00)*	*(0.00)***	*(0.00)*
female head	0.06	-0.03	0.10	0.04
	(0.11)	*(0.06)*	*(0.13)*	*(0.10)*
number of children	-0.19	-0.001	-0.19	0.03
	*(0.04)****	*(0.02)*	*(0.04)****	*(0.03)*
head schooling	0.97	-0.001	1.05	0.01
	*(0.11)****	*(0.06)*	*(0.13)****	*(0.09)*
head ethnic minority	0.23	-	0.28	-
	*(0.12)**		*(0.14)***	
bank (lagged)	-	-	-	-0.38
				(0.45)
NGO (lagged)	-	-	-	1.16
				*(0.63)**
PCs of hh assets:				
assets & exp. (pc1)	-	0.17	-	0.17
		*(0.01)****		*(0.02)****
assets & exp. (pc2)	-	-0.08	-	-0.06
		*(0.02)****		*(0.03)**
assets & exp. (pc3)	-	0.02	-	0.07
		(0.02)		*(0.04)**
shocks:				
household only	0.44	-	0.52	-
	*(0.09)****		*(0.11)****	
land slide	-	0.59	-	0.73
		*(0.26)***		*(0.33)*
harvest diseases	-	-0.07	-	-0.21
		(0.05)		*(0.07)*
land taken by cooperative	-	-0.06	-	0.87
		(0.52)		*(0.90)*
head imprisoned	-	0.30	-	0.87
		(0.52)		*(0.91)*
assets resettlements	-	-0.33	-	-1.54
		(0.64)		*(0.90)**
banditry		-1.39		(1.68)
		(0.90)		*(0.91)**
PA characteristics:				
n. villages in PA	0.09	-	0.10	-
	*(0.01)****		*(0.01)****	
dist. nearest bank all weather road	0.01	-	0.01	-
	*(0.00)****		*(0.00)****	
n. of agricultural offices in PA	0.26	-	0.14	-
	*(0.10)***		*(0.12)*	
irrigated land (ha)	0.001	-	0.001	-
	*(0.00)****		*(0.00)****	
rain fed land (ha)	0.001	-	0.002	-
	*(0.00)****		*(0.00)****	
south	-	0.17	-	-0.16

		*(0.07)***		*(0.12)*
round 2	-	-0.87	-	-0.20
		*(0.06)****		*(0.09)***
round 3	-	-0.65	-	0.01
		*(0.07)****		*(0.09)*
round 4	-	-0.61	-	-
		*(0.07)****		
constant	-3.20	5.00	-3.88	4.41
	*(0.46)****	*(0.25)****	*(0.56)****	*(0.39)****
N. Obs	1,940		1,063	
LR test of ind. equs.	(0.20†)		(0.06†)*	

Source: own calculation from ERHS. Standard errors in parenthesis. †p-value ***$p < 0.01$,**$p < 0.05$, *$p < 0.1$

Table C3-10: Selectivity models - FIML (PA has no Equbs)

	Model I		Model II	
	Pr(Informal)	**Log(informal credit)**	**Pr(Informal)**	**Log(informal credit)**
hh characteristics:				
age head	-0.02	0.01	0.02	0.05
	(0.02)	*(0.03)*	*(0.04)*	*(0.12)*
age head squared	0.0002	-0.00004	-0.0001	-0.0004
	(0.00)	*(0.00)*	*(0.00)*	*(0.00)*
hh size	-0.22	-0.08	-0.13	-0.48
	*(0.08)****	*(0.11)*	*(0.12)*	*(0.25)**
hh size squared	-0.001	0.002	-0.004	0.02
	(0.00)	*(0.01)*	*(0.01)*	*(0.02)*
female head	-0.18	-0.38	-0.30	-0.35
	(0.14)	*(0.21)**	*(0.22)*	*(0.55)*
number of children	0.19	0.13	0.15	0.24
	*(0.05)****	*(0.08)*	*(0.07)***	*(0.19)*
head schooling	-1.42	0.19	-1.46	-0.66
	*(0.18)****	*(0.33)*	*(0.32)****	*(0.85)*
head ethnic minority	-1.20	-	-1.10	-
	*(0.26)****		*(0.40)****	
NGO (lagged)	-	-	-	-0.51
				(0.44)
PCs of hh assets:				
assets & exp. (pc1)	-	0.32	-	0.62
		*(0.12)****		*(0.38)*
assets & exp. (pc2)	-	-0.28	-	-0.55
		*(0.16)**		*(0.31)**
assets & exp. (pc3)	-	0.24	-	0.58
		*(0.15)***		*(0.30)*
shocks:				
household only	-0.62	-	-0.60	-
	*(0.12)****		*(0.19)****	
land slide	-	0.67	-	0.62
		(0.59)		*(0.82)*
harvest diseases	-	-0.05	-	-0.11
		(0.28)		*(0.79)*
land taken by cooperative	-	0.26	-	-
		(0.81)		
PA characteristics:				
n. villages in PA	-0.18	-	-0.17	-
	*(0.03)****		*(0.05)****	
dist. nearest bank all weather road	-0.06	-	-0.06	-
	*(0.01)****		*(0.02)****	
n. of agricultural offices in PA	0.02	-	-0.11	-
	(0.16)		*(0.23)*	
irrigated land (ha)	-0.01	-	-0.01	-
	*(0.00)****		*(0.00)****	
rain fed land (ha)	-0.004	-	-0.004	-
	*(0.00)****		*(0.00)****	
Haresaw	-	0.54	-	-
		(0.22)**		
Geblen	-	0.93	-	0.73
		*(0.38)***		*(0.71)*
round 2	-	-0.89	-	0.70
		*(0.16)****		*(0.39)**
round 3	-	-0.85	-	0.71
		*(0.30)***		*(0.61)*
round 4	-	-1.15	-	-

		*(0.17)****		
constant	5.12	5.37	-3.16	5.36
	*(0.78)****	*(0.75)****	*(1.24)***	*(3.73)*
N. Obs	2,219		2,103	
LR test of ind. equs.	(0.80†)		(0.74†)	

Source: own calculation from ERHS. Standard errors in parenthesis. †p-value ***$p < 0.01$,**$p < 0.05$, *$p < 0.1$

Table C3-11: Selectivity models - 2 Step estimation (PA has no Equbs)

Log(informal credit)	**Model I** 1^{st}*stage*	2^{nd}*stage*	**Model II** 1^{st}*stage*	2^{nd}*stage*
hh characteristics:				
age head	-0.02	0.01	0.02	-0.06
	(0.02)	*(0.03)*	*(0.04)*	*(0.11)*
age head squared	0.0002	-0.00004	-0.0001	-0.001
	(0.00)	*(0.00)*	*(0.00)*	*(0.00)*
hh size	-0.22	-0.08	-0.13	-0.45
	*(0.08)****	*(0.11)*	*(0.12)*	*(0.25)**
hh size squared	-0.001	0.002	-0.004	0.01
	(0.00)	*(0.01)*	*(0.01)*	*(0.02)*
female head	-0.18	-0.38	-0.29	-0.30
	(0.14)	*(0.21)**	*(0.22)*	*(0.53)*
number of children	0.19	0.13	0.15	0.21
	*(0.05)****	*(0.09)*	*(0.07)***	*(0.19)*
head schooling	-1.43	0.20	-1.46	-0.50
	*(0.18)****	*(0.33)*	*(0.32)****	*(0.92)*
head ethnic minority	-1.20	-	-1.07	-
	*(0.26)****		*(0.38)****	
NGO (lagged)	-	-	-	-0.64
				(0.45)
PCs of hh assets:				
assets & exp. (pc1)	-	0.32	-	0.60
		*(0.11)****		*(0.35)**
assets & exp. (pc2)	-	-0.28	-	-0.46
		*(0.16)**		*(0.33)*
assets & exp. (pc3)	-	0.24	-	0.48
		(0.15)		*(0.33)*
shocks:				
household only	-0.62	-	-0.60	-
	*(0.12)****		*(0.18)****	
land slide	-	0.67	-	0.64
		(0.59)		*(0.79)*
harvest diseases	-	-0.05	-	0.002
		(0.28)		*(0.78)*
land taken by cooperative	-	0.25	-	-
		(0.81)		
PA characteristics:				
n. villages in PA	-0.18	-	-0.16	-
	*(0.03)****		*(0.05)****	
dist. nearest bank all weather road	-0.06	-	-0.06	-
	*(0.01)****		*(0.02)****	
n. of agricultural offices in PA	0.02	-	-0.12	-
	(0.16)		*(0.23)*	
irrigated land (ha)	-0.01	-	-0.01	-
	*(0.00)****		*(0.00)****	
rain fed land (ha)	-0.004	-	-0.004	-
	*(0.00)****		*(0.00)****	
Haresaw	-	0.54	-	0.37
		*(0.22)***		*(0.73)*
Geblen	-	0.93	-	1.13
		*(0.39)***		*(1.14)*
round 2	-	-0.89	-	0.60
		*(0.16)****		*(0.39)*
round 3	-	-0.85	-	0.64
		*(0.30)****		*(0.58)*
round 4	-	-1.15	-	-
		*(0.17)****		

constant	5.12	5.37	3.09	5.29
	*(0.78)****	*(0.75)****	*(1.22)***	*(3.36)*
Mills ratio	-0.07		-0.32	
	(0.26)		*(0.61)*	
N. Obs	2,219		2,103	

Source: own calculation from ERHS. Standard errors in parenthesis. †p-value
***$p < 0.01$,**$p < 0.05$,*$p < 0.1$

TABLE C3-12: Endogenous switching regression

	Model I		Model II	
	Pr(Pa has Equbs)	**Log(informal credit)**	**Pr(Pa has no Equbs)**	**Log(informal credit)**
hh characteristics:				
age head	0.03	0.03	0.02	0.32
	(0.02)	*(0.02)*	*(0.03)*	*(0.16)**
age head squared	-0.0003	-0.0003	-0.0001	-0.003
	*(0.00)**	*(0.00)**	*(0.00)*	*(0.00)**
hh size	0.47	0.02	-0.18	-0.21
	*(0.06)****	*(0.04)*	*(0.09)**	*(0.45)*
hh size squared	-0.01	-0.002	-0.001	0.002
	*(0.00)***	*(0.00)*	*(0.00)*	*(0.02)*
female head	0.13	0.06	-0.38	0.19
	(0.12)	*(0.11)*	*(0.19)***	*(0.83)*
number of children	-0.26	0.01	0.15	0.14
	*(0.04)****	*(0.03)*	*(0.06)***	*(0.31)*
head schooling	1.59	0.01	-1.36	-0.46
	*(0.14)****	*(0.10)*	*(0.27)****	*(1.44)*
head ethnic minority	1.66	-	-0.99	-
	*(0.22)****		*(0.35)****	
bank (lagged)	-	-0.30	-	-
		(0.52)		
NGO (lagged)	-	1.23	-	-1.90
		*(0.73)**		*(0.71)****
PCs of hh assets:				
assets & exp. (pc1)	-	0.16	-	0.38
		*(0.02)****		*(0.49)*
assets & exp. (pc2)	-	-0.06	-	-0.55
		(0.04)		*(0.59)*
assets & exp. (pc3)	-	0.06	-	0.41
		(0.04)		*(0.44)*
shocks:				
household only	0.70	-	-0.55	-
	*(0.11)****		*(0.16)****	
land slide	-	0.70	-	1.93
		*(0.38)**		*(1.40)*
harvest diseases	-	-0.29	-	0.25
		*(0.09)**		*(1.38)*
land taken by cooperative	-	0.99	-	-
		(1.05)		
head imprisoned	-	0.90	-	-
		(1.05)		
assets resettlement	-	-1.65	-	-
		(1.06)		
banditry	-	-1.69	-	-
		(1.06)		
PA characteristics:				
n. villages in PA	0.26	-	-0.27	-
	*(0.03)****		*(0.06)****	
dist. nearest bank	0.07	-	-0.06	-
all weather road	*(0.01)****		*(0.02)****	
n. of agricultural offices in PA	-0.39	-	0.09	-
	*(0.15)***		*(0.22)*	
irrigated land (ha)	0.01	-	-0.01	-
	*(0.00)****		*(0.00)****	
rain fed land (ha)	0.004	-	-0.01	-
	*(0.00)****		*(0.00)****	
south	-	-0.24	-	-

		*(0.13)**		
Haresaw	-	-	-	-0.99
				(0.89)
Geblen	-	-	-	0.27
				(1.83)
round 2	-	-0.07	-	1.42
		(0.10)		*(0.62)***
round 3	-	0.16	-	2.05
		(0.11)		*(0.96)***
constant	-8.38	4.50	3.85	-2.81
	*(0.67)****	*(0.46)****	*(1.17)****	*(4.30)*
N. Obs	1,612		4,149	

Source: own calculation from ERHS. Standard errors in parenthesis. $^{***}p < 0.01$,$^{**}p < 0.05$,$^{*}p < 0.1$

CHAPTER 4

TABLE C4-1: Summary of literature on crowding-out and crowding-in

Country and segment of population	Year	GDP per capita 2000 US$	Transfer responsiveness to income†	Source
			Evidence on crowding-out:	
Bangladesh	1998-1999	346		McKernan, Pitt and Moskowitz
Gifts			-0.25(women)	(2005)
Informal loans			-0.31(men)	
Burkina Faso				Kazianga (2006)
Rural	1994	211	-0.153	
	1998	225	-0.132	
Urban	1994		-0.194	
	1998		-0.244	
India	1981-1982	234	-0.0209	Rosenzweig (1988)
(six villages in semi-arid tropics)				
Indonesia	1993	730	-0.494	Raut and Tran (2005)
(parents receiving from non-cores children)				
Jamaica	1989	2894	-0.25	Clarke and Wallsten (2003)
(remittances post Hurricane Gilbert)				
Mexico	1998	5513	between	Albarran and Attanasio (2002)
(poor, rural area)			-0.23 & -1.59	
Mexico	1998	5513	-	Attanasio and Rios-Rull (2000)
Peru	1985-1986	2188	High Income	Cox, Eser and Jimenez (1998)
			-0.013	
Philippines	1988	882		Cox, Hansen and Jimenez (2004)
Rural			Low inc. -0.4	
			High inc. -0.03	
Urban			Low inc. -0.39	
			High inc. -0.01	

Polland	1987	3053	-0.054	Cox, Jimenez and Okrasa (1997)
Russia	1994-2000	1591	-0.1 (elderly hh only)	Kuhn and Stillman (2004)
South Africa	1994	2846	Earned income: Below poverty level: -0.07 Public pensions: Below poverty level: -0.09	Maitra and Ray (2003)
South Africa *(remittances going to pensioners in Venda) province - low income*	1989	3131	Women: -0.30 Men: -0.26 (responsiveness of remittances to pension increase between 1989 and 1992)	Jensen (2004)
United States	1988	27362	-0.013	Schoeni (1997)
United States	1968-1984	-	-0.256/-0.682	Rosenzweig and Wolpin (1994)
			Evidence on crowding-in:	
Botswana *(remittances)*	1978-1979	918	0.011 (elasticity)	Lucas and Stark (1985)
China *(rural hh cross-China)*	1988	347	0.011	Secondi (1997)
Dominican Republic *(receipt of remittances by farm households in Dominican Sierra)*	1994	1712	0.09	de la Briere *et al.* (2002)
Indonesia *(Exchange with children)*	1993	730	0.132 (elasticity)	Frankenberg, Lillard and Willis (2002)
Peru	1985-1986	2188	Low Income 0.140	Cox, Eser and Jimenez (1998)
South Africa	1994	2846	Earned income: Above poverty level: 0.00 Public pensions:	Maitra and Ray (2003)

			Above poverty level: 0.04	
Thailand *(rural areas)*	2001	1980	-	Kaboski and Townsend (2006)

Source:Handbook of Development Economics (2008).[†]Answer the question, if income increases by 1 unit, by how many units do private transfer inflows increase or decrease?

TABLE C4-2: Series of logit models of participation (second definition)

Pr(participation in ...)	**MRFC vs. Past member**	**2nd program vs. Past member**
households characteristics		
household size	5.88	1.17
	*(4.96)***	*(0.36)*
household size squared	0.88	1.00
	*(0.06)**	*(0.02)*
age head	0.68	1.00
	*(0.10)****	*(0.01)*
age head squared	1.00	-
	*(0.00)****	
female head†	-	0.88
		(0.40)
n. of children 6-10	0.75	0.86
	(0.21)	*(0.18)*
n. sick days[1]	-	0.95
		(0.04)
education and occupation of HH head		
msce certificate†	0.76	0.24
	(1.31)	*(0.24)*
occupation in agriculture†	2.32	-
	(1.69)	
contract labourer†	-	0.08
		(0.10)
households assets		
land size (ha)	1.05	0.93
	(0.10)	*(0.07)*
share of land owned by spouse (%)	1.01	1.00
	*(0.01)***	*(0.00)*
n. of gifts	1.35	-
	(0.76)	
community characteristics		
total n. of households	1.00	1.00
	*(0.00)***	*(0.00)*
electricity†	3.86	-
	*(2.03)****	
distance to government office (Km)	1.06	1.02
	*(0.03)**	*(0.03)*
distance to credit office (Km)	0.98	1.04
	(0.02)	*(0.04)*
Mangochi†	15.57	9.97
	*(21.26)***	*(8.41)****
Nkhotakota†	2.54	6.10
	(2.49)	*(4.04)****
Rumphi†	1.69	2.17
	(1.71)	*(1.54)*
round 2†	0.01	0.11
	*(0.01)****	*(0.04)****
round 3†	0.31	0.03
	*(0.16)***	*(0.02)****
N. of obs.	311	480
Pseudo-R^2	0.34	0.33

Source: own calculation from MRFMHFS. Note: odds ratios displayed and std. errors in (). Robust standard errors. Weighted regression.†dummy variables.[1] month prior to interview.***$p < 0.01$,**$p < 0.05$,

TABLE C4-3: Selected characteristics by treatment groups (second definition) prior to matching

	(I) MRFC vs. Past members		**(II) 2nd Program vs. Past members**		**Group comparisons (I)**		**(II)**	
	treated	*untreated*	*treated*	*untreated*	**t-stat.**	% $\vert bias \vert$	**t-stat.**	% $\vert bias \vert$
households characteristics:								
household size	5.46	5.86	6.17	6.37	-1.44	19.6	-0.87	8
household size squared	33.31	39.20	-	-	-1.56	22.00		
age head	49.86	45.66	45.380	45.83	2.37**	30.20	-0.4	3.7
age head squared	2698	2257.30	-	-	2.48**	49.7		
female head†	-	-	0.33	0.26	-	-	1.63	15.00
n. of children 6-10	0.91	1.07	1.15	1.13	-1.35	18.40	0.15	1.40
n. of days sick[1]	-	-	1.67	2.02	-	-	-0.87	8.20
education of households head:								
msce certificate†	0.01	0.03	0.03	0.04	-0.66	9.40	-0.99	9.40
professional training†	0.15	0.14	-	-	0.26	3.40	-	-
occupation in agriculture†	0.90	0.82	-	-	1.62	22.40	-	-
contract labourer†	-	-	0.01	0.02	-	-	-0.67	6.40
households assets:								
land size (ha)	2.30	2.34	1.99	2.44	-0.16	2.40	-2.04**	19.8
share of land owned by spouse (%)	32.99	25.32	14.22	12.70	1.42	18.20	0.53	4.90
n. of gifts	0.19	0.06	-	-	2.86***	30.70	-	-
characteristics of the community:								
total n. of households	164.77	244.24	486.23	400.38	-2.18**	31.5	2.20**	20.20
electricity†	0.13	0.22	-	-	-1.82*	25.10	-	-
distance to the government (Km)	22.93	17.47	9.88	9.97	2.92***	37.40	-0.12	1.10
distance to the credit office (Km)	9.87	8.22	2.89	3.24	0.89	11.40	-0.78	7.20

Source: own calculation from MRFMHFS.†dummy variables.[1] month prior to interview.***$p < 0.01$,**$p < 0.05$,*$p < 0.1$.

TABLE C4-4: Selected characteristics (second definition) by treatment groups after matching (when bias prior to matching> 10%)

	(I) MRFC vs. Past members		**(II) 2nd Program vs. Past members**		**Group comparisons (I)**		**(II)**	
	treated	*controls*	*treated*	*controls*	**t-stat.**	**% reduc. |*bias*|**	**t-stat.**	**% reduc. |*bias*|**
households characteristics:								
household size	5.46	5.62	-	-	-0.56	58.40	-	-
household size squared	33.31	35.09	-	-	-0.51	69.80		
age head	49.86	48.01	-	-	0.83	56.10	-	-
age head squared	2698.00	2476.10	-	-	0.97	31.20		
female head†	-	-	0.33	0.35	-	-	-0.53	62.90
n. of children 6-10	0.91	0.88	1.15	1.13	0.20	83.80	-	-
n. of days sick[1]	-	-	1.67	2.02	-	-	-	-
education & occupation of HHs head								
occupation in agriculture†	0.90	0.91	-	-	-0.27	83.50	-	-
households assets								
land size (ha)	-	-	1.99	1.84	-	-	1.07	66.40
share of land owned by spouse (%)	32.99	32.35	-	-	0.09	91.60	-	-
n. of gifts	0.19	0.10	-	-	1.28	29.80	-	-
characteristics of the community:								
total n. of households	164.77	185.58	486.23	487.07	-0.67	73.80	-0.02	99.00
electricity†	0.13	0.13	-	-	0.00	100.00	-	-
distance to the government (Km)	22.93	21.41	-	-	0.65	72.10	-	-
distance to the credit office (Km)	9.87	10.31	-	-	-0.18	73.70	-	-

Source: own calculation from MRFMHFS.†dummy variables.[1] month prior to interview.***$p < 0.01$,**$p < 0.05$,*$p < 0.1$.

Table C4-5: Selected characteristics by treatment groups prior to matching (MNL model)

	(I) MRFC vs. Past members		(II) 2nd Program vs. Past members		Group comparisons (I)		Group comparisons (II)	
	treated	*controls*	*treated*	*controls*	**t-stat.**	% $\|bias\|$	**t-stat.**	% $\|bias\|$
households characteristics:								
household size	5.73	5.79	6.40	5.79	-0.28	3.10	2.68***	26.30
age head	49.16	44.48	46.29	44.48	3.06***	34.60	21.60	15.00
n. of children 6-10	0.92	1.13	1.15	1.13	-2.09**	23.70	0.20	1.90
education & occupation of HHs head:								
msce certificate†	0.01	0.03	0.03	0.03	-1.09	12.40	-0.03	0.30
households assets								
land size (ha)	2.27	2.39	2.11	2.39	-0.44	5.00	-1.33	11.20
share of land owned by spouse (%)	32.73	21.78	13.91	21.78	2.36**	26.70	-2.48**	22.50
characteristics of the community								
total n. of households	201.72	244.64	466.95	244.64	-1.35	15.30	5.94***	60.60
distance to the government (Km)	22.50	15.45	10.14	15.45	4.41***	49.90	-5.95***	49.60
distance to the credit office (Km)	10.95	6.62	3.00	6.62	2.72***	30.70	-5.10***	40.20

Source: own calculation from MRFMHFS.†dummy variables.[1] month prior to interview.***$p < 0.01$,**$p < 0.05$.

TABLE C4-6: Selected characteristics by treatment groups after matching in MNL model (when bias prior to matching> 10%)

	(I) MRFC vs. Past members		**(II) 2nd Program vs. Past members**		**Group comparisons**			
					(I)		**(II)**	
	treated	*controls*	*treated*	*controls*	**t-stat.**	**% reduc.** $\lvert bias \rvert$	**t-stat.**	**% reduc.** $\lvert bias \rvert$
households characteristics:								
household size	-	-	6.40	6.05	-	-	2.10**	41.50
age head	49.16	47.44	46.29	43.50	1.11	63.30	3.37***	54.10
n. of children 6-10	0.92	0.93	-	-	-0.14	93.80	-	-
education & occupation of HHs head:								
msce certificate†	0.01	0.01	-	-	0.00	100.00	-	-
households assets								
land size (ha)	-	-	2.11	1.89	-	-	1.80*	24.10
share of land owned by spouse (%)	32.73	30.42	13.91	9.41	0.46	78.90	2.18**	42.90
characteristics of the community								
total n. of households	201.72	211.60	466.95	437.50	-0.31	77.00	0.95	86.80
distance to the government (Km)	22.50	21.04	10.14	10.13	0.88	79.30	0.02	99.80
distance to the credit office (Km)	10.95	10.92	3.00	2.68	0.01	99.40	1.00	91.40

Source: own calculation from MRFMHFS.†dummy variables.[1] month prior to interview.***$p < 0.01$,**$p < 0.05$,*$p < 0.1$.

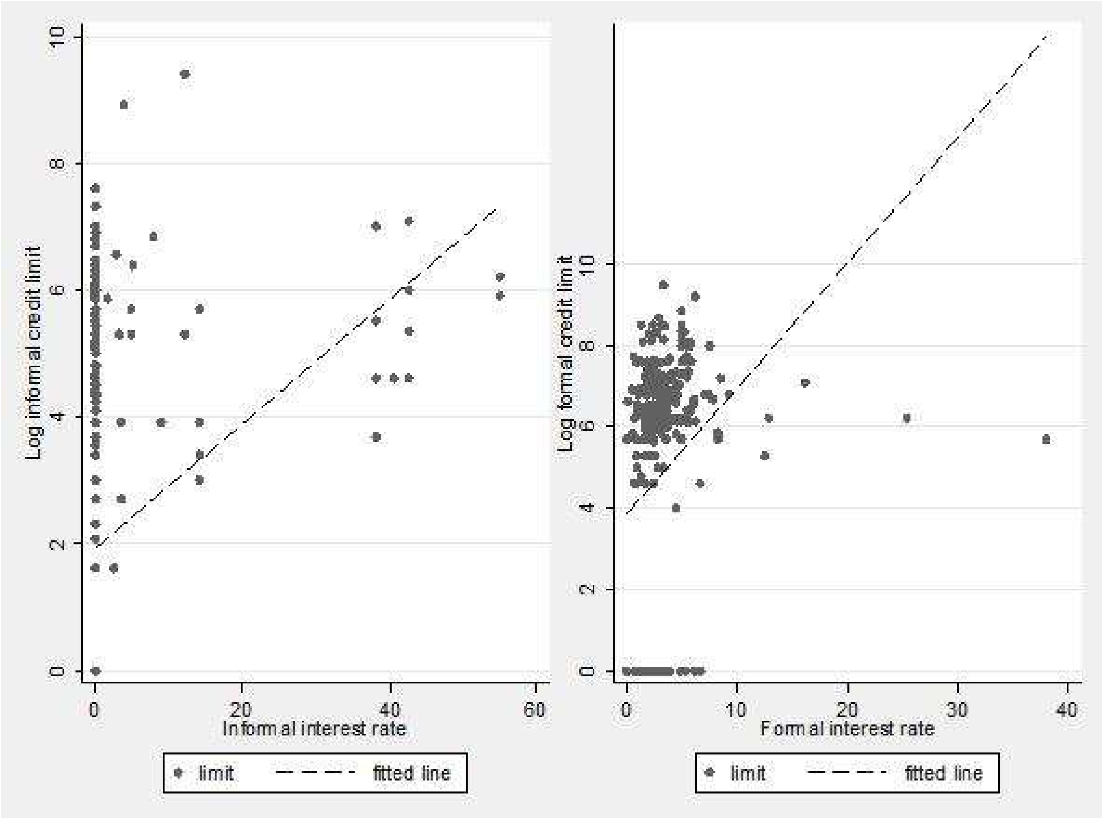
TABLE C4-7: Polichotomous selection model (OLS model - 2^{nd} stage Heckman)

Amount of credit borrowed from informal lenders	**Coefficients**	**t-stat.**
households characteristics		
household size	16.28	1.44
age head	1.21	1.62
female head†	-51.5	-1.33
n. of children 6-10	-30.72	-1.71*
n. sick days[1]	-1.98	-0.86
education & occupation of HH head		
msce certificate†	-180.3	-0.84
occupation in agriculture†	27.78	2.03**
contract labourer†	-40.81	-0.84
households assets		
land size (ha)	6.28	1.24
share of land owned by spouse (%)	-0.29	-1.64
n. of gifts	-30.32	-1.14
lambda (MRFC)	185.58	1.41
lambda (2nd Program)	91.17	1.02
Mangochi†	536.52	2.05**
Nkhotakota†	14.55	1.18
Rumphi†	16.30	0.58
round 2†	13.51	0.90
round 3†	22.35	0.92
constant	-43.74	-1.27
N. of obs.	**94**	
R^2	**0.53**	

Source: own calculation from MRFMHFS. Robust standard errors.
Weighted regression.†dummy variables.[1] month prior to interview.** $p < 0.05$, * $p < 0.1$.

CHAPTER 5

FIGURE C5-1: Scatter plot of interest rate versus credit limit by source



Source: Own calculation based on MRFMHFS.

TABLE C5-1: Costs of borrowing (including zero costs)

Type of costs:	**Mean**	**Min.**	**Max.**	**St. Dev.**
Interest rates from:				
MRFC programme	0.65 (450)	0	6.21	1.54
More than one programme	2.81 (210)	0	38.04	3.59
Informal lenders	1.44 (450)	0	55.02	6.96

Source: own calculation from MRFMHFS. Loan costs include: travel, fees, no or low wage etc.% monthly interest rate. Number of respondents in brackets. Note: all values in local currency, 15 Malawian Kwachas (MK)=1 US$. Malawi's per capita GNP is US$ 170 (approx.2,550 MK). World Bank, 1997.

Table C5-2: Instrumental variables tobit - 2^{nd} stage regressions

	MODEL I: (Log Informal credit)	**MODEL II: (Log Formal credit)**
hh characteristics:		
hh size	0.002	-0.04
	(0.08)	*(0.12)*
age head	0.01	0.01
	(0.01)	*(0.01)*
female head	-0.28	0.45
	(0.51)	*(0.79)*
n. children 0-15	0.10	0.24
	(0.10)	*(0.15)*
head can't read/write†	-2.13	-0.61
english	*(0.69)****	*(1.23)*
spouse employed in small trade	0.63	0.09
	*(0.32)**	*(0.44)*
Assets, expenditure and shocks:		
share of assets held (%)	-0.00	-0.01
as land	*(0.00)*	*(0.01)**
size of land (ha)	0.07	0.003
planted with crops	*(0.08)*	*(0.12)*
share of livestock	0.0002	0.0002
owned by spouse	*(0.00)*	*(0.01)*
food expenditure (MK)	0.003	-0.01
	(0.00)	*(0.00)*
non food expenditure (MK)	0.00	0.001
	(0.00)	*(0.00)****
number of negative shocks	-0.11	0.08
	(0.10)	*(0.13)*
Costs of loans:		
formal interest rate (%)	0.08	0.04
	(0.05)	*(0.07)*
informal interest rate (%)	0.001	0.06
	(0.01)	(0.02)***
formal loan costs	-0.02	0.01
	(0.04)	*(0.06)*
informal loan costs	0.004	0.03
	(0.02)	*(0.03)*
% Christians in the same village	0.01	-0.001
	(0.01)	*(0.01)*
% n. of tubewells	-0.14	-0.25
for drinking water	*(0.16)*	*(0.24)*
South†	0.08	0.45
	(0.22)	*(0.33)*
round 2†	-0.28	-0.36
	(0.41)	*(0.47)*
round 3†	-1.78	-0.68
	*(0.42)****	*(0.66)*
Informal credit limit (MK)	0.98	-0.60
	*(0.11)****	*(0.19)****

MRFC credit limit (MK)	-0.35	0.40
	*(0.21)**	*(0.29)*
2nd program credit limit (MK)	-0.33	0.30
	*(0.18)**	*(0.25)*
N. Obs.	259	259

Source: own calculation from MRFMHFS. †dummy variables.***$p < 0.01$,**$p < 0.05$,*$p < 0.1$.Marginal effects and standard errors displayed.

TABLE C5-3: Instrumental variables tobit - 1^{st} stage regressions

	L^I	L^{MRFC}	$L^{2nd}prog.$
hh characteristics:			
hh size	-0.12	-0.09	0.20
	(0.14)	*(0.16)*	*(0.17)*
age head	-0.001	0.01	0.03
	(0.01)	*(0.01)*	*(0.02)**
female head†	1.95	-1.64	-0.30
	*(0.78)***	*(0.87)**	*(0.95)*
n. children 0-15	0.19	-0.004	0.09
	(0.18)	*(0.20)*	*(0.22)*
head can't read/write†	-1.09	-1.97	1.37
english	*(1.38)*	*(1.54)*	*(0.81)*
spouse employed in small trade†	-0.44	-1.13	2.37
	(0.40)	*(0.45)***	*(0.49)****
Assets, expenditure and shocks:			
share of assets held (%)	-0.003	-0.001	-0.01
as land	*(0.01)*	*(0.01)*	*(0.01)*
size of land (ha)	-0.08	0.06	0.11
planted with crops	*(0.14)*	*(0.15)*	*(0.16)*
share of livestock	-0.002	0.01	-0.01
owned by spouse	*(0.01)*	*(0.01)*	*(0.01)*
food expenditure (MK)	-0.01	-0.001	0.002
	(0.01)	*(0.01)*	*(0.01)*
non food expenditure (MK)	0.001	-0.0002	-0.004
	*(0.00)***	*(0.00)*	*(0.00)*
number of negative shocks	0.08	-0.24	-0.17
	(0.12)	*(0.13)**	*(0.15)*
Costs of loans:			
formal interest rate (%)	-0.06	0.11	0.14
	(0.05)	*(0.05)***	*(0.06)****
informal interest rate (%)	0.10	-0.01	-0.01
	*(0.02)****	*(0.02)*	*(0.09)*
formal loan costs	-0.07	0.06	-0.12
	(0.07)	*(0.08)*	*(0.09)*
informal loan costs	0.08	-0.02	0.0004
	*(0.03)***	*(0.04)*	*(0.04)*
% Christians in the same village	-0.01	0.01	-0.01
	(0.01)	*(0.01)*	*(0.02)*
% n. of tubewells	0.09	-0.46	-0.01
for drinking water	*(0.31)*	*(0.34)*	*(0.01)*
South†	-1.19	2.66	-2.66
	*(0.61)**	*(0.41)****	*(0.44)****
round 2†	-1.37	-0.20	1.02
	*(0.45)****	*(0.50)*	*(0.54)**
round 3†	-1.23	1.64	0.15
	*(0.44)****	*(0.49)****	*(0.54)*
community characteristics:			
number of churches†	-0.11	-0.26	0.22

	(0.09)	*(0.10)****	*(0.11)***
number of private†	2.07	1.06	-1.68
clinics	*(0.61)****	*(0.67)*	*(0.74)***
number of NGOs	1.01	-0.17	0.43
	(0.68)	*(0.75)*	*(0.82)*
n. of shops	-0.06	-0.11	0.17
	*(0.03)**	*(0.04)****	*(0.04)****
n. of members in farms	-0.01	0.03	-0.003
clubs	*(0.01)*	*(0.01)****	*(0.01)*
hungry season (February)†	1.23	-3.32	3.69
	*(0.53)***	*(0.59)****	*(0.65)****
Constant	3.06	0.44	0.49
	*(1.09)****	*(1.79)*	*(1.32)*
N. Obs.		**259**	

Source: own calculation from MRFMHFS. †dummy variables.***$p < 0.01$,**$p < 0.05$,*$p < 0.1$.

Note: first stage regressions are the same for formal and informal credit.The dependent variables are in log form.

TABLE C5-4: Tobit regression - Marginal effects of informal credit

Log(informal credit)	$E(D \mid x)$	$E(D \mid x, D > 0)$
hh characteristics:		
hh size	0.028	0.022
age head	0.001	0.001
female head	-0.114	-0.092
n. children 0-15	0.011	0.008
head primary education†	-0.205	-0.164
spouse employed in small trade†	0.038	0.030
Assets, expenditure and shocks:		
land size (ha)	-0.017	-0.013
land share owned by spouse (%)	0.0003	0.0003
land share in total assets (%)	0.006	0.005
value of assets (MK)	0.00001	0.000001
food expenditure (MK)	0.003	0.003
non food expenditure (MK)	-0.00003	-0.00002
number of negative shocks	-0.105	-0.084
Costs of loans:		
formal interest rate (%)	0.009	0.007
informal interest rate (%)	-0.001	-0.001
formal loan costs	-0.0004	-0.0003
informal loan costs	-0.001	-0.001
% Christians in the same village	0.004	0.003
South†	0.106	0.085
round 2†	-0.227	-0.182
log informal credit limit (MK)	0.459	0.367
log MRFC credit limit (MK)	-0.051	-0.041
log 2nd program credit limit (MK)	0.001	0.001
N. Obs.	284	

Source: own calculation from MRFMHFS. Marginal effects displayed. †dummy variables.

TABLE C5-5: Tobit regression - Marginal effects of formal credit

Log(formal credit)	$E(D \mid x)$	$E(D \mid x, D > 0)$
hh characteristics:		
hh size	-0.056	-0.046
age head	0.016	0.013
female head	-0.621	-0.504
n. children 0-15	0.238	0.193
head primary education†	-0.509	-0.413
spouse employed in small trade†	-0.019	-0.015
Assets, expenditure and shocks:		
land size (ha)	-0.098	-0.079
land share owned by spouse (%)	-0.0002	-0.0001
land share in total assets (%)	-0.002	-0.001
value of assets (MK)	0.0001	0.0001
food expenditure (MK)	0.0004	0.0003
non food expenditure (MK)	0.001	0.001
number of negative shocks	0.001	0.001
Costs of loans:		
formal interest rate (%)	0.074	0.060
informal interest rate (%)	0.023	0.018
formal loan costs	-0.021	-0.017
informal loan costs	0.0003	0.0003
% Christians in the same village	0.003	0.0003
South†	0.626	0.508
round 2†	0.249	0.202
round 3†	0.218	0.177
log informal credit limit (MK)	-0.194	-0.157
log MRFC credit limit (MK)	0.227	0.184
log 2nd program credit limit (MK)	0.241	0.195
N. Obs.	284	

Source: own calculation from MRFMHFS. Marginal effects displayed.†dummy variables.

TABLE C5-6: Instrumental variables tobit - 2^{nd} stage regression (drop assets & exp.)

	Log(informal credit)	Log(formal credit)
hh characteristics:		
hh size	0.02	0.03
	(0.13)	*(0.11)*
age head	0.01	0.001
	(0.01)	*(0.01)*
female head†	0.21	-0.96
	(0.81)	*(0.80)*
n. children 0-15	0.11	0.10
	(0.16)	*(0.14)*
head primary education†	-0.42	-0.47
	(0.33)	*(0.28)*
spouse employed in small trade†	-0.05	0.17
	(0.42)	*(0.37)*
Land and shocks:		
land size (ha)	0.01	-0.02
	(0.06)	*(0.06)*
land share owned by spouse (%)	0.004	-0.002
	(0.00)	*(0.00)*
land share in total assets (%)	0.01	-0.01
	(0.01)	*(0.01)*
number of negative shocks	-0.23	0.06
	*(0.13)**	*(0.11)*
Costs of loans:		
formal interest rate (%)	0.05	0.03
	(0.06)	*(0.05)*
informal interest rate (%)	0.02	0.03
	(0.02)	*(0.02)*
formal loan costs	0.01	-0.004
	(0.06)	*(0.06)*
informal loan costs	0.01	0.02
	(0.03)	*(0.03)*
% Christians in the same village	0.02	-0.001
	*(0.01)***	*(0.01)*
South†	0.21	0.30
	(0.33)	*(0.32)*
round 2†	-0.83	0.04
	(0.60)	*(0.45)*
round 3†	-6.62	-0.80
	*(0.39)****	*(0.47)**
log informal credit limit (MK)	0.95	-0.27
	*(0.19)****	*(0.17)*
log MRFC credit limit (MK)	-0.45	0.56
	*(0.22)***	*(0.17)****
log 2nd program credit limit (MK)	-0.08	0.45
	(0.17)	*(0.14)****
N. Obs.	262	262

Source: own calculation from MRFMHFS.†dummy variables.***$p < 0.01$,**$p < 0.05$,*$p < 0.1$.Marginal effects and standard errors displayed.

TABLE C5-7: Instrumental variables tobit - 1^{st} stage regressions (drop assets & exp.)

	L^I	L^{MRFC}	$L^{2nd}prog.$
hh characteristics:			
hh size	-0.05	-0.10	0.30
	(0.13)	*(0.15)*	*(0.16)**
age head	-0.01	0.01	0.03
	(0.01)	*(0.01)*	*(0.01)**
female head†	1.59	-0.65	-1.30
	(0.86)	*(0.98)*	*(1.08)*
n. children 0-15	0.03	-0.02	0.06
	(0.16)	*(0.19)*	*(0.20)*
head primary education†	0.11	0.59	-0.47
	(0.34)	*(0.39)*	*(0.42)*
spouse employed in small trade†	-0.66	-0.89	2.27
	*(0.39)**	*(0.44)*	*(0.48)****
Land and shocks:			
land size (ha)	0.16	0.01	0.03
	*(0.07)***	*(0.08)*	*(0.08)*
land share owned by spouse (%)	-0.004	0.01	0.001
	(0.00)	*(0.00)*	*(0.01)*
land share in total assets (%)	-0.004	-0.001	-0.01
	(0.01)	*(0.01)*	*(0.01)*
number of negative shocks	0.12	-0.18	-0.10
	(0.12)	*(0.14)*	*(0.16)*
Costs of loans:			
formal interest rate (%)	-0.05	0.09	0.13
	(0.04)	*(0.05)**	*(0.06)***
informal interest rate (%)	0.09	-0.01	-0.02
	*(0.02)****	*(0.02)*	*(0.03)*
formal loan costs	-0.04	0.04	-0.11
	(0.07)	*(0.08)*	*(0.09)*
informal loan costs	0.08	-0.02	-0.003
	*(0.03)***	*(0.04)*	*(0.04)*
% Christians in the same village	-0.002	0.01	-0.01
	(0.01)	*(0.02)*	*(0.02)*
South†	-0.59	2.50	-1.74
	(0.58)	*(0.66)****	*(0.72)***
round 2†	-1.44	-0.28	0.82
	*(0.43)****	*(0.49)*	*(0.54)*
round 3†	-1.46	1.61	-0.11
	*(0.41)****	*(0.47)****	*(0.51)*
community characteristics:			
number of deep tube wells	0.49	-0.52	-0.06
	*(0.18)****	*(0.21)***	*(0.23)*
farm clubs†	-0.05	1.15	0.42
	(0.52)	*(0.59)**	*(0.65)*
other clubs†	-0.30	-0.85	1.90
	(0.43)	*(0.49)**	*(0.54)****
avg. price of maize in October	-0.32	-1.46	0.13

	(0.72)	*(0.82)**	*(0.90)*
distance to commercial bank	0.04	-0.01	-0.0002
	*(0.01)****	*(0.01)*	*(0.01)*
n. of members in farms clubs	-0.01	0.001	0.01
	(0.02)	*(0.02)*	*(0.03)*
n. of households	-0.002	0.0004	0.0004
	*(0.00)****	*(0.00)*	*(0.00)*
n. of HHs with land > 5 acres	-0.01	0.02	1.97
	(0.01)	*(0.01)*	*(0.93)***
n. of moneylenders	-0.02	0.31	0.13
	(0.33)	*(0.38)*	*(0.41)*
hungry season (February)†	2.25	-2.69	2.22
	*(0.74)****	*(0.84)****	*(0.96)***
Constant	2.04	0.52	-0.42
	(1.51)	*(1.72)*	*(1.87)*
N. Obs.		**262**	

Source: own calculation from MRFMHFS. †dummy variables.***$p < 0.01$,**$p < 0.05$,*$p < 0.1$.

Note: first stage regressions are the same for formal and informal credit.The dependent variables are in log form.

Table C5-8: Instrumental variables - 2^{nd} stage regressions

	MODEL I: Log(informal credit)	**MODEL II: Log(formal credit)**
hh characteristics:		
hh size	-0.03	-0.03
	(0.06)	*(0.10)*
age head	0.01	0.01
	(0.01)	*(0.01)*
female head	0.15	-0.63
	(0.38)	*(0.70)*
n. children 0-15	0.11	0.17
	(0.07)	*(0.13)*
head primary education†	-0.18	-0.39
	(0.14)	*(0.25)*
spouse employed in small trade	0.08	-0.03
	(0.18)	*(0.33)*
Assets, expenditure and shocks:		
land size (ha)	0.01	-0.04
	(0.03)	*(0.06)*
land share owned by spouse (%)	0.001	-0.001
	(0.00)	*(0.00)*
land share in total assets (%)	0.01	-0.01
	(0.01)	*(0.01)*
value of assets (MK)	0.0003	0.00004
	(0.00)	*(0.00)*
food expenditure (MK)	0.003	0.0003
	*(0.00)**	*(0.00)*
non food expenditure (MK)	0.00	0.001
	(0.00)	*(0.00)****
number of negative shocks	-0.07	0.07
	(0.05)	*(0.10)*
Costs of loans:		
formal interest rate (%)	-0.02	-0.06
	(0.03)	*(0.05)*
informal interest rate (%)	0.01	0.02
	(0.01)	(0.02)
formal loan costs	-0.02	-0.04
	(0.03)	*(0.06)*
informal loan costs	0.01	0.02
	(0.01)	*(0.03)*
% Christians in the same village	0.01	0.001
	(0.00)	*(0.00)*
South†	0.13	0.34
	(0.16)	*(0.29)*
round 2†	-0.14	-0.13
	(0.21)	*(0.38)*
round 3†	0.04	-0.47
	(0.24)	*(0.43)*
log informal credit limit (MK)	0.76	-0.47
	*(0.09)****	*(0.17)****

log MRFC credit limit (MK)	-0.18	0.36
	*(0.10)**	*(0.19)**
log 2nd program credit limit (MK)	-0.08	0.32
	(0.09)	*(0.16)***
Constant	-0.68	4.42
	(0.47)	*(0.86)****
N. Obs.	256	256

Source: own calculation from MRFMHFS. †dummy variables.***$p < 0.01$,**$p < 0.05$,*$p < 0.1$. standard errors displayed.

Table C5-9: Selectivity models - 2 Step estimation

	MODEL I: FIML		MODEL II: 2 step estimation	
	Pr(Informal)	**Log(informal credit)**	**Pr(formal)**	**Log(formal credit)**
hh characteristics:				
hh size	0.07	0.33	0.19	0.08
	(0.05)	*(0.17)**	*(0.06)****	*(0.11)*
hh size squared	-	-0.02	-	-0.01
		(0.02)		*(0.01)*
age head	-0.01	-0.06	0.003	0.02
	*(0.00)***	*(0.05)*	*(0.00)*	*(0.03)*
age head squared	-	0.001	-	-0.0002
		(0.00)		*(0.00)*
female head†	-1.39	-0.32	-1.99	-1.05
	*(0.22)****	*(1.73)*	*(0.28)****	*(0.45)***
n. children 0-15	-0.06	0.16	-0.08	0.09
	(0.06)	*(0.19)*	*(0.07)*	*(0.06)*
head can read and write†	-0.07	0.40	-0.05	0.05
	(0.13)	*(0.21)**	*(0.14)*	*(0.12)*
spouse does household work†	-	0.42	-	0.18
		*(0.20)***		*(0.14)*
head employed in agriculture†	-	0.49	-	0.15
		*(0.25)***		*(0.12)*
Assets, expenditure and shocks:				
land size (ha)	-	0.05	-	0.02
		*(0.03)**		*(0.03)*
land share in total assets (%)	-	0.001	-	-0.003
		(0.00)		*(0.00)*
value of assets (MK)	-	0.0001	-	0.0001
		(0.00)		*(0.00)****
food expenditure (MK)	-	0.01	-	0.001
		*(0.00)**		*(0.00)*
non food expenditure (MK)	-	0.0001	-	0.0004
		(0.00)		*(0.00)****
number of negative shocks	0.18	-	-0.06	-
	*(0.06)****		*(0.05)*	
% people in trad. religion in village	-	-0.02	-	-0.01
		*(0.01)***		*(0.01)***
South†	0.21	-0.21	-0.08	0.41
	(0.26)	*(0.19)*	*(0.22)*	*(0.12)****
round 2†	-0.82	-0.41	-1.99	-0.13
	*(0.12)****	*(0.76)*	*(0.14)****	*(0.35)*
round 3†	-1.67	0.40	-1.92	-0.02
	*(0.18)****	*(1.67)*	*(0.14)****	*(0.31)*
log informal credit limit (MK)	-	0.41	-	-0.06
		*(0.14)****		*(0.02)****
log MRFC credit limit (MK)	-	-0.12	-	0.11
		*(0.06)**		*(0.02)****
log 2nd program credit limit (MK)	-	0.06	-	0.09
		(0.04)		*(0.02)****
Village characteristics:				
number of deep tube wells	0.22	-	0.04	-
	(0.14)		*(0.09)*	
electricity†	0.59	-	-0.11	-
	*(0.22)****		*(0.26)*	
farm clubs†	0.004	-	0.73	-
	(0.22)		*(0.23)****	
traditional healers†	-0.42	-	0.36	-

	*(0.14)****		*(0.16)***	
price of maize (july)	-0.98	-	0.45	-
	*(0.39)***		*(0.33)*	
distance to credit office (Km)	0.02	-	-0.02	-
	(0.02)		*(0.01)***	
distance to comm. bank (Km)	0.01	-	-0.02	-
	(0.01)		*(0.01)****	
n. of clubs memb.	-0.01	-	-0.01	-
	(0.01)		*(0.01)*	
n. of households	0.001	-	-0.001	-
	(0.00)		*(0.00)*	
n. of HHs with land btw	-0.01	-	0.004	-
3-4.99 acres	*(0.00)***		*(0.00)*	
hungry season (February)†	1.11	-	-1.20	-
	*(0.30)****		*(0.33)****	
n. of moneylenders in the	0.21	-	0.26	-
village	*(0.09)****		*(0.09)****	
constant	0.31	0.81	-0.35	4.38
	(0.34)	*(2.25)*	*(0.40)*	*(0.71)****
N. Obs	961		946	
Mills ratio	-		0.31	
			(0.27)	
LR test of ind. equs.	**(0.61‡)**		-	

Source: own calculation from MRFMHFS. Coefficients displayed and standard errors in parenthesis.
‡p-value.†dummy variable.***$p < 0.01$,**$p < 0.05$,*$p < 0.1$

Bibliography

[1] A. Adera. Instituting effective linkages between formal and informal financial sector in Africa: A proposal. *Savings and Development*, 1:5–22, 1995.

[2] S. R. Aiyagari. Uninsured idiosyncratic risk and aggregate saving. *The Quarterly Journal of Economics*, 109(3):659–684, 1994.

[3] P. Albarran and O. P. Attanasio. Do private transfers crowd out private transfers? Evidence from a randomized experiment in Mexico. Technical report, Discussion Paper N. 2002-6. World Institute for Development Economics, Helsinki, Finland, 2002.

[4] A. Alderman and C. Paxson. Do the poor insure? A synthesis of the literature on risk sharing institutions in developing countries. Technical report, Discussion Paper 164, Princeton N.J., U.S.A.: Princeton University, Woodrow Wilson School of Public and International Affairs., 1992.

[5] I. Aleem. *The economics of rural organization; theory, practice, and policy*, chapter Imperfect Information, Screening, and the Cost of informal Lending: A Study of a Rural Credit Market in Pakistan, pages 131–154. Oxford University Press, Oxford, 1993.

[6] J. Altonji, F. Hayashi, and L. Kotlikoff. Is the extended family altruistically linked? Direct tests using micro data. *American Economic Review*, 82 (5):11771198, 1992.

[7] S. Altug and R.A. Miller. Household choices in equilibrium. *Econometrica*, 58 (3):543–570, 1990.

[8] T. Amemiya. *Advanced Econometrics*. Cambridge, Mass., U.S.A, 1985.

[9] T. Amemiya and Q. H. Vuong. A comparison of two consistent estimators in the choice-based sampling qualitative response model. *Econometrica*, 55 (3):699–702, 1987.

[10] S. Ardener. The comparative study of rotating credit associations. *The Journal of Royal Anthropological Institute of Great Britain and Ireland*, 94(2):201–229, 1964.

[11] D. Aredo. The informal and semi-formal financial sectors in Ethiopia: A study of the iqqub, iddir and savings and credit co-operatives,. Technical report, AERC Research Paper 21, 1993.

[12] B. Armendariz and J. Morduch. *The Economics of Microfinance.* The MIT Press Cambridge, Massachusetts London, England, 2005.

[13] E. Aryeetey and M. Hyuha. *Economic reform in Sub-Saharan Africa*, chapter The Informal Financial Sector and Markets in Africa: An Empirical study. A World Bank Symposium, Washington D.C., 1991.

[14] E. Aryeetey and C. Udry. The characteristics of informal financial markets in Africa. In *Conference of African Economic Research Consortium, Kenya.*, 1995.

[15] O. Ashenfelter. Estimating the effect of training program on earning. *Review of Economics and Statistics*, 60:47–57, 1978.

[16] N. ASHRAF, D. KARLAN, and W. YIN. Tying odysseus to the mast: Evidence from a commitment savings product in the philippines. Technical report, Yale University CENTER DISCUSSION PAPER NO. 917, 2005.

[17] O. Attanasio and S.J. Davis. Relative wage movements and the distribution of consumption. *The Journal of Political Economy*, 104 (6):1227–1262, 1996.

[18] O. Attanasio, P. K. Goldberg, and E. Kyriazidou. Credit constraints in the market for consumer durables: evidence from micro data on car loans. *International Economic Review*, 49 (2):401–436, 2008.

[19] O. Attanasio and G. Weber. Consumption growth and excess sensitivity to income: evidence from U.S. micro data. Technical report, University College London and Stanford University, 1992.

[20] O. P. Attanasio and J. V. Rios-Rull. Consumption smoothing and extended families. Technical report, University College London, mimeo, 2000.

[21] P.C. Austin and M. M. Mandami. A comparison of propensity score methods: a case-study estimating the effectiveness of post-AMI statin use. *Statistics in Medicine*, 25 (12):2084–2106, 2006.

[22] J. P. Azam, B. Biais, M. Dia, and C. Maurel. Informal and formal credit markets and credit rationing in Cote dIvoire. *Oxford Review of Economic Policy*, 17 N. 4:520–534, 2001.

[23] A. Badhuri. A study in agricultural backwardness under semi-feudalism. *Economic Journal*, 83:120–137, 1973.

[24] B.M. Balk. Decompositions of Fisher indexes. *Economic Letters*, 82:107–113, 2004.

[25] A. Banerjee and E. Duflo. The nature of credit constraints: evidence from and Indian bank. Technical report, mimeo, Department of Economics, MIT, 2001.

[26] J. Banks, R. Blundell, and A. Brugiavini. Risk pooling, precautionary saving and consumption growth. *Review of Economic Studies*, 68:757 779, 2001.

[27] P. Bardhan and C. Udry. *Development Microeconomics*. Oxford University Press, 1999.

[28] B. Barham, S. Boucher, and M. Carter. Credit constraints, credit unions, and small scaleproducers in Guatemala. *World Development*, 24 (5):792805, 1996.

[29] R. J. Barro. Are government bonds net wealth? *Journal of Political Economy*, 82:1095 1117, 1974.

[30] M. Barslund and F. Tarp. Rural credit in Vietnam. Technical report, Department of Economics University of Copenhagen Discussion Papers 2, 2006.

[31] K. Basu. *Rural Credit Markets: The Structure of Interest Rates, Exploitation, and Efficiency.* Oxford: Oxford University Press, 1989.

[32] M. M. Baydas, Z. Bahloul, and D. W. Adams. Informal finance in Egypt: banks within banks. *World Development*, 23(4):651–661, 1995.

[33] G. S. Becker. A theory of social interactions. *Journal of Political Economy*, 82:1063–1093, 1974.

[34] C. Bell. Interactions between institutional and informal credit agencies in rural India. *The World Bank Economic Reviwe*, 4(3):297–327, 1990.

[35] C. Bell, T. N. Srinivasan, and C. Udry. Rationing, spillover, and interlinking in credit markets: The case of rural Punjab:. *Oxford Economic Papers*, 49(4):557–585, 1997.

[36] B. S. Bernanke. Permanent income, liquidity, and expenditure on automobiles: Evidence from panel data. *Quarterly Journal of Economics*, 99:587–614, 1984.

[37] D. Bernheim, A. Schleifer, and L. Summers. The strategic bequest motive. *Journal of Political Economy*, 93 (6):10451076, 1985.

[38] G. Bertola, R. Disney, and C. Grant. *The economics of consumer credit.* The MIT Press, 2006.

[39] T. Besley. How do market failures justify interventions in rural credit markets? *The World Bank Research Observer*, 9(1):27–47, 1994.

[40] T. Besley and S. Coate. Group lending, repayment incentives and social collateral. *Journal of Development Economics*, 46:1–18, 1995.

[41] T. Besley, S. Coate, and G. Loury. The economics of rotating savings and credit associations. *The American Economic Review*, 83 (4):792–810, 1993.

[42] T. Besley and A. R. Levenson. The role of informal finance household capital accumulation: evidence from Taiwan. *The Economic Journal*, 106:39–59, 1996.

[43] H. Bester. Screening vs. rationing in credit markets with imperfect information. *The American Economic Review*, 75(4):850–855, 1985.

[44] P. Bevan and A. Pankhurst. A social analysis of fifteen rural economies in Ethiopia. Technical report, Centre for study of African Economies, University of Oxford and Department of Sociology and Social Anthropology, University of Addis Ababa, 1996.

[45] H. Binswanger and S. Khandker. The impact of formal finance on the rural economy of india. *Journal of Development Studies*, 32 (2):234–262, 1995.

[46] H. P. Binswanger and M. Rosenzweig. Wealth, weather risk and the composition and profitability of agricultural investments. *Economic Journal*, 103:56–78, 1986.

[47] R. Blundell and L. Pistaferri. Income volatility and household consumption: The impact of food assistance programs. *Journal of Human Resources*, 38:1032–1050, 2003.

[48] F. Bokosi and A. Khalil-Edriss. *Determinants and Characteristics of Household Demand for Agricultural Credits in Malawi*, chapter Vol. II. In Integrated Micro-Credits, Micro-Enterprises and Market Reforms in Subsistence Economy: Experiences from Malawi., 2003.

[49] N. Bose and R. Cothren. Asymmetric information and loan contracts in a neoclassical growth model. *Journal of Money, Credit, and Banking*, 29:423 439, 1997.

[50] P. Bose. Formal-informal interaction in rural credit markets. *Journal of Development Economics*, 66 N. 2:265–280, 1998.

[51] S. Boucher and C. Guirkinger. Risk, wealth and sectoral choice in rural credit markets. *American Journal of Agricultural Economics*, 89 (4):991–1004, 2007.

[52] A. Braverman and L.J. Guasch. Rural credit markets and institutions in developing countries: Lessons for policy analysis from practice and modern theory. *World Development*, 14:1253–1267, 1986.

[53] T. Brodaty, B. Crepon, and D. Fougere. *Econometric evaluation of labour markets policies*, chapter Using matching estimators to evaluate alternative youth employment programs: evidence from France , 1986-1988. Heidelberg: Physica-Vertlag, 2001.

[54] A. Bryson, R. Dorsett, and S. Purdon. The use of propensity score matching in the evaluation of active labour market policies. Technical report, Policy Studies Institute and National Centre for Social Research Working Paper N.4, 2002.

[55] R. Burgess and R. Pande. Do rural banks matter? evidence from the indian social banking experiment. *American Economic Review*, vol. 95(3):780–795, 2005.

[56] M. Caliendo, R. Hujer, and S. Thomsen. The employment effects of job creation schemes in Germany a microeconometric evaluation. Technical report, Discussion Paper N. 1512, IZA, Bonn, 2005.

[57] M. Caliendo and S. Kopeinig. Some practical guidance for the implementation of propensity score matching. Technical report, IZA Discussion Paper N. 1588, 2005.

[58] J. Campbell and N. G. Mankiw. Permanent income, current income, and consumption. *Journal of Business and Economics Statistics,*, 8:265279, 1990.

[59] B. Capeau and S. Dercon. Prices, local measurement units and subsistence consumption in rural surveys: an econometric approach with an application to Ethiopia. Technical report, Centre for Study of African Economies Series WPS/1998-10, 1998.

[60] S. B. Carpenter and R. T. Jensen. Household participation in formal and informal saving mechanisms: evidence from Pakistan. *Review of Development Economics*, 6:314–328, 2002.

[61] C. D. Carroll. Buffer stock saving and the permanent income hypothesis. Technical report, Washington, D.C.: Board of governors of the Federal Reserve System., 1991.

[62] E. Castronova and P. Hagstrom. The demand for credit cards: evidence from the survey of consumer finances. *Economic Inquiry*, 42 (2):304–318, 2004.

[63] R. B. Cattell. The scree test for the number of factors. *Multivariate Behavioral Research*, 1:629–637, 1966.

[64] C. Chipeta and M.L.C. Mkandawire. Financial integration and development in Sub-Saharan Africa: Case study of the informal financial sector in Malawi. Technical report, Report Prepared for the Overseas Development Institute, London., 1994.

[65] I. Chung. Market choice and effective demand for credit: The roles of borrower transaction costs and rationing constraints. *Journal of Economic Development*, 20 (2):2344, 1995.

[66] G. Clarke and S. J. Wallsten. Do remittances act like insurance? Evidence from a natural disaster in Jamaica. Technical report, Mimeo Development Research Group. The World Bank, 2003.

[67] J.H. Cochrane. A simple test of consumption insurance. *The Journal of Political Economy*, 99 (5):957–976, 1991.

[68] R. H. Cole and L. Mishler. *Consumer and commercial credit management*. McGraw-Hill College, 1998.

[69] J. Coleman. Social capital in the creation of human capital. *American Journal of Sociology*, 94 (Supplement):S95–S120, 1988.

[70] J. Conning. Mixing and matching loans: Credit rationing and spillover in a rural credit market in Chile. Technical report, mimeo Williams College, 2001.

[71] S. Cosslett. Maximum likelihood estimator for choice-based sampling. *Econometrica*, 49 (5):1289–1316, 1981.

[72] D. Cox. Motives for private income transfers. *Journal of Political Economy*, 95 (3):508 546, 1987.

[73] D. Cox, Z. Eser, and E. Jimenez. Motives for private transfers over the life cycle: an analysitical framework and evidence for Peru. *Journal of Development Economics*, 55:57–80, 1998.

[74] D. Cox and M. Fafchamps. Extended family and kinship networks: economic insights and evolutionary directions. In *Handbook of Development Economics Vol. 4*. Elsevier B. V., 2008.

[75] D. Cox, B. E. Hansen, and E. Jimenez. How responsive are private transfers to income? Evidence from a laissez-faire economy. *Journal of Public Economics*, 88:2193–2219, 2004.

[76] D. Cox and E. Jimenez. Social security and private transfers in developing countries: the case of Peru. *The World Bank Economic Review*, 6(1):155–169, 1992.

[77] D. Cox, E. Jimenez, and W. Okrasa. Family safety nets and economic transition: a study of worker households in Poland. *Review of Income and Wealth*, 43:191–209, 1997.

[78] D. Cox and M. R. Rank. Inter-vivos transfers and intergenerational exchange. *Review of Economics and Statistics*, 74(2):305314, 1992.

[79] T. F. Cox. *An introduction to multivariate data analysis.* London: Hodder Arnold., 2005.

[80] R. B. Jr. D'Agostino. Propensity score methods for bias reduction in the comparison of a treatment to a non-randomized control group. *Statistics in Medicine*, 17:2265–2281, 1998.

[81] I. Dasgupta, R. Disney, and A. Kedir. Why do households use banks and roscas simultaneously? Theory and evidence from urban Ethiopia. In *Paper presented at the CSAE 2007 conference in Oxford, available at http://www.csae.ox.ac.uk/conferences/2007-EDiA-LaWBiDC/print-prog-details.htm*, 2007.

[82] D. de Meza and D. C. Webb. Too much investment: a problem of asymmetric information. *Quarterly Journal of Economics*, 102:281–292, 1987.

[83] D. de Meza and D. C. Webb. Credit rationing something's gotta give. *Economica*, 73:563–578, 2005.

[84] A. Deaton. *Understanding consumption.* Oxford: Oxford University Press, 1992.

[85] A. Deaton. *The Analysis of Household Surveys: A Microeconometric Approach to Development Policy.* The World Bank, 1997.

[86] R. H. Dehejia and S. Wahba. Propensity score matching methods for nonexperimental causal studies. *The Review of Economics and Statistics*, 84 (1):151–161, 2002.

[87] B. deLaBriere, E. Sadoulet, A. DeJanvry, and S. Lambert. The roles of destination, gender, and household composition in explaining remittances: an analysis for the dominican sierra. *Journal of Development Economics*, 68:309–328, 2002.

[88] S. Dercon. Income risk, coping strategies, and safety nets. *The World Bank Research Observer*, 17 (2):141–166, 2002.

[89] S. Dercon and P. Krishnan. In sickness and in health: Risk sharing within households in rural Ethiopia. *Journal of Political Economy*, 108(4):688–727, 2000.

[90] A. Diagne. Determinants of household access to and participation in formal and informal credit markets in Malawi. Technical report, International Food Policy Research Institute, Washington D.C., FCND Discussion Paper N. 67, 1999.

[91] A. Diagne, M. Zeller, and M. Sharma. Empirical measurement of households access to credit and credit constraints in developing countries: methodological issues and evidence. Technical report, International Food Policy Research Institute, Washington D.C., FCND Discussion Paper N. 90, 2000.

[92] J. J. Diaz and S. Handa. An assessment of propensity score matching as a nonexperimental impact estimator. Evidence from Mexico's PROGRESA program. *Journal of Human Resources*, 41(2):319–345, 2006.

[93] J. DiNardo and J. Tobias. Nonparametric density and regression estimation. *Journal of Economic Perspectives*, 15 (4):11–28, 2001.

[94] R. Disney, A. Kedir., and A. McKay. Price deflators and food poverty in urban Ethiopia. 2004.

[95] R. Dorsett. The new deal for young people: relative effectiveness of the options in reducing male unemployment. Technical report, PSI Discussion Paper N. 7, 2001.

[96] G. H. Dunteman. Principal components analysis. newbury park. In *Quantitative applications in social sciences Vol. 69*. Newbury Park, Sage University papers series, 1989.

[97] P. B. Duong and Y. Izumida. Rural development finance in Vietnam: a microeconometric analysis of household surveys. *World Development*, 30 N. 2:319–335, 2002.

[98] S. Durlauf and M. Fafchamps. *Handbook of Economic Growth, Volume 1B, Philippe Aghion and Stephen Durlauf (eds.), pp. 1637-99, North Holland, 2005*, chapter Social Capital. North Holland, 2005.

[99] E. Engen and J. Gruber. Unemployment insurance and precautionary savings. *Journal of Monetary Economics*, 47:545–579, 2001.

[100] M. Eswaran and A. Kotwal. Credit as insurance in agrarian economies. *Journal of Development Economics*, 31:37–53, 1989.

[101] M. Fafchamps. Trade credit in Zimbabwean manufacturing. *World Development*, 25 N. 5:795–815, 1997.

[102] M. Fafchamps. Rural poverty, risk, and development. Technical report, Report submitted for the Food and Agriculture Organization, 1999.

[103] M. Fafchamps. Ethnicity and credit in African manufacturing. *Journal of Development Economics*, 61:205–235, 2000.

[104] M. Fafchamps. Development and social capital. *Journal of Development Studies*, 42(7):1180–1198, 2006.

[105] M. Fafchamps. *Handbook of Social Economics,*, chapter Risk Sharing between Households. Jess Benhabib, forthcoming.

[106] M. Fafchamps and F. Gubert. Risk sharing and network formation. *American Economic Review Papers and Proceedings*, 97 (2):75–79, 2007.

[107] M. Fafchamps and S. Lund. Risk-sharing networks in rural Philippines. *Journal of Development Economics*, 71 N. 2:261–287, 2003.

[108] S. Fay, E. Hurst, and M. White. The consumer bankruptcy decision. *American Economic Review*, 92:706–718, 2002.

[109] G. Feder, L. J. Lau, J. Y. Lin J. Y., and X. Luo. The relationship between credit and productivity in Chinese agriculture: a microeconomic model of disequilibrium. *American Journal of Agricultural Economics*, 72 (5):1151–1157, 1990.

[110] B. Fitzenberger and H. Prey. Evaluating public sector sponsored training in East Germany. *Oxford Economic Papers*, 52:497–520, 2000.

[111] E. Frankenberg, L. Lillard, and R. J. Willis. Patters of intergenerational transfers in southeast Asia. *Journal of Marriage and the Family*, 64:627–641, 2002.

[112] M. Friedman. A theory of the consumption function. In *National Bureau of Economic Research*. Princeton, N. J., 1957.

[113] M. Frlich, A. Heshmati, and M. Lechner. A microeconometric evaluation of rehabilitation of long-term sickness in Sweden. *Journal of Applied Econometrics*, 19:375–396, 2004.

[114] S. Gaechter and B. Herrmann. Reciprocity, culture, and human cooperation: previous insights and a new cross-cultural experiment. *Philosophical transactions of the Royal Society B - Biological Sciences*, 364:791–806, 2009.

[115] C. Geertz. The rotating credit association: a middle rung in development. *Economic Development and Cultural Change*, 1(3), 1962.

[116] M. Ghatak. Group lending, local information and peer selection. *Journal of Development Economics*, 60:27–50, 1999.

[117] M. Ghatak and T. W. Guinnane. The economics of lending with joint liability: theory and practice. *Journal of Development Economics*, 60:195–228, 1999.

[118] P. Gosh, D. Mookherjee, and D. Ray. *A Reader in Development Economics*, chapter Credit rationing in developing countries: an overview of the theory. London: Blackwell, 2000, 1999.

[119] C. Grant. Estimating credit constraints among US household. *Oxford Economic Papers*, 59 (4):583–605, 2007.

[120] W. Greene. *Econometric Analysis.* Pearson Education International, 2003.

[121] F. Grimard. Household consumption smoothing through ethnic ties: evidence from Cote d'Ivoire. *Journal of Development Economics*, 53, 1997.

[122] D. Gross and N. S. Souleles. Do liquidity constraints and interest rates matter for consumer behaviour? Evidence from credit card data. *Quarterly Journal of Economics*, 117 (1):149–185, 2002.

[123] J. Gruber and A. Yelowitz. Public health insurance and private savings. *Journal of Political Economy*, 107 (6):1249–1274, 1999.

[124] L. Guiso, L. Pistaferri, and F. Schivardi. Insurance within the firm. *Journal of Political Economy*, 113 (5):1054–1087, 2005.

[125] L. Guiso, P. Sapienza, and L. Zingales. The role of social capital in financial development. *American Economic Review of Development Economics*, 94 (3):526–556, 2004.

[126] R. E. Hall. Stochastic implications of the life cycle-permanent income hypothesis: Theory and evidence. *Journal of Political Economy*, 86:971–987, 1978.

[127] R. E. Hall and F. S. Mishkin. The sensitivity of consumption to transitory income: Estimates from panel data on households. *Econometrica*, 50:461–81, 1982.

[128] C. Harris and D. Laibson. Dynamic choices of hyperbolic consumers. *Econometrica*, 69, (4):935 – 957, 2001.

[129] J. Hausman and D. McFadden. A specification test for the multinomial logit model. *Econometrica*, 52:1219–1240, 1984.

[130] F. Hayashi. *Tests for liquidity constraints: a critical survey and some new observations*, chapter Advances in Econometrics. Cambridge: Cambridge University Press, 1987.

[131] F. Hayashi. Is the japanese extended family altruistically linked? A test based on Engle curves. *Journal of Political Economy*, 103 (3):661674, 1995.

[132] J. Heckman. Sample selection bias as a specification error. *Econometrica*, 47:153161, 1979.

[133] J. J. Heckman, H. Ichimura, and P. E. Todd. Matching as an econometric evaluation estimator: evidence from evaluating a job training program. *Review of Economic Studies*, 64:605–654., 1997.

[134] J. J. Heckman, R. J. LaLonde, and J. A. Smith. *Handbook of Labor Economics, Vol. 3 A*, chapter The economics and econometrics of active labor market programs, pages 1865–2097. Amsterdam: Elsevier Science, 1999.

[135] J. J. Heckman and J. A. Smith. The pre-program earnings dip and the determinants of participation in a social program: implications for simple program evaluation strategies. *Economics Journal*, 108:313–348, 1999.

[136] J. J. Heckman and P. Todd. A note on adapting propensity score matching and selection models to choice based samples. 2005.

[137] F. Heidhues. Rural finance markets an important tool to fight poverty. *Quarterly Journal of International Agriculture*, 34(2):105–108, 1995.

[138] G. Helmke and S. Levitsky. Informal institutions and comparative politics: a research agenda. Technical report, South Bend, IN: Kellog Institute for International Studies, Notre Dame University, 2003.

[139] J. Hoddinott, S. Dercon, and P. Krishnan. Networks and informal mutual support in 15 Ethiopian villages. Technical report, CSAE Working Paper, 2005.

[140] K. Hoff and J. E. Stiglitz. Introduction: Imperfect information and rural credit marketspuzzles and policy perspectives. *The World Bank Economic Review*, 4 N. 3:235–250, 1990.

[141] H. Hotelling. Analysis of a complex of statistical variables into principal components. *Journal of Educational Psychology*, 24:417–441, 1933.

[142] G. Ibrahim, A. Kedir, and S. Torres. Household-level credit constraints in urban Ethiopia. Technical report, University of Leicester, Working Paper N. 07/03, 2007.

[143] G. W. Imbens. The role of propensity score in estimating dose-response functions. *Biometrika*, 87 (3):706–710, 2000.

[144] G. Impavido. Credit rationing, group lending and optimal group size. *Annals of Public and Cooperative Economics*, 69 N. 2:243–260, 1998.

[145] T. Ishikawa. Imperfection in the capital market and the institutional arrangement of inheritance. *Review of Economic Studies*, 41:383–404, 1974.

[146] D. M. Jaffee and T. Russell. Imperfect information, uncertainty and credit rationing. *Quarterly Journal of Economics*, 90:651–666, 1976.

[147] T. Jappelli. Who is credit constrained in the U.S. economy? *Quarterly Journal of Economics*, 1:219–234, 1990.

[148] T. Jappelli and M. Pagano. Consumption and capital market imperfections: an international comparison. *American Economic Review*, 79(5):1088–1105, 1989.

[149] T. Jappelli and M. Pagano. *Public credit information: a European perspective. In credit reporting Sustems and in the International economy*, chapter 2, pages 81–114. Cambridge, Mass.: The MIT Press, 2003.

[150] R. T. Jensen. Do private transfers displace the benefits of public transfers? evidence from South Africa. *Journal of Public Economics*, 88:89–112, 2003.

[151] I. T. Jolliffe. *Principal component analysis*. New York: Springer series in statistics, 2002.

[152] J. P. Kaboski and R. M Townsend. The impacts of credit on village economies. Technical report, University of Chicago, 2006. 2006.

[153] H. Kazianga. Motives for household private transfers in Burkina Faso. *Journal of Development Economics*, 79:73–117, 2006.

[154] A.M. Kedir. Estimation of own- and cross- price elasticities using unit values: econometric issues and evidence from urban Ethiopia. *Journal of African Economies*, 14 (1):1–20, 2005.

[155] M. S. Kimball. Precautionary saving in the small and in the large. *Econometrica*, 58(1):53–73, 1990.

[156] A. Kimhi. Estimation of an endogenous switching regression model with discrete dependent variables: Monte-carlo analysis and empirical application of three estimators. *Empirical Economics*, 24 N.2:225–241, 1999.

[157] M.A. King. Capital market imperfections and the consumption function. *Scandinavian Journal of Economics*, 88:59–80, 1986.

[158] A. Kochar. An empirical investigation of rationing constraints in rural credit markets in India. *Journal of Development Economics*, 53:339–371, 1997.

[159] P. Krusell and A. Smith. Consumption-savings decisions with quasi-geometric discounting", econometrica, 71, 365-375. *Econometrica*, 71:365–375, 2003.

[160] R. Kuhn and S. Stillman. Understanding intrahousehold transfers in a transition economy: evidence from Russia. *Economic Development and Cultural Change*, 53:131–156, 2004.

[161] E. LaFerrara. Kin groups and reciprocity: A model of credit transactions in Ghana. *American Economic Review*, 3 N. 5:1730–1751, 2003.

[162] D. Laibson. Golden eggs and hyperbolic discounting. *Quarterly Journal of Economics*, 62:443–479, 1997.

[163] L. Larsson. Evaluation of swedish youth labour market programmes. Technical report, Discussion paper 2000(1) Office for Labour Market Policy Evaluation, Uppsala, 2000.

[164] M. Lechner. Identification and estimation of causal effects of multiple treatments under the conditional independence assumption. Technical report, Discussion paper 9908, University of St. Gallen, 1999a.

[165] M. Lechner. Earnings and employment effects of continuous off-the-job training in East Germany after unification. *Journal of Business Economic Statistics*, 17:74–90, 1999b.

[166] M. Lechner. An evaluation of public sector sponsored continuous vocational training programs in East Germany. *Journal of Human Resources*, 35:347–375, 2000.

[167] M. Lechner. Programme heterogeneity and propensity score matching: an application to the evaluation of active labour market policies. Technical report, Review of Economic Statistics, unpublished, 2001. Review of Economic Statistics, unpublished.

[168] L. F. Lee. Generalized econometric models with selectivity. *Econometrica*, 51:507–512, 1983.

[169] L.F. Lee. Unionism and wage rates: a simultaneous equation model with qualitative and limited dependent variables. *International Economic Review*, 19:415–433, 1978.

[170] Y. Lee, W.L. Parish, and R.J. Willis. Sons, daughters, and intergenerational support in Taiwan. *American Journal of Sociology*, 99(4):10101041, 1994.

[171] R. Lensink and E. Sterken. Asymmetric information, option to wait to invest and the optimal level of investment. *Journal of Public Economics*, 79:365374, 2001.

[172] R. Lensink and E. Sterken. The option to wait to invest and equilibrium credit rationing. *Journal of Money, Credit and Banking*, 31:221–225, 2002.

[173] C. Levi-Strauss. *The elementary structures of kinship.* London: Eyre and Spottis-woode, 1949.

[174] R. E. Lucas and O. Stark. Motivations to remit: evidence from Botswana. *Journal of Political Economy*, 93:901–918, 1985.

[175] E. M. BERTRAND, S. MULLAINATHAN and SHAFIR. A behavioral economics view of poverty. *American Economic Review Papers and Proceedings*, pages 419–423, 2004.

[176] G. S. Maddala. *Limited Dependent and qualitative variables in econometrics.* Cambridge University Press, 1983.

[177] A. Madestam. Developing credit markets. Technical report, Economic Research Institute, Stockhol School of Economics, 2005.

[178] P. Maitra and R. Ray. The effect of transfers on household expenditure patterns and poverty in South Africa. *Journal of Development Economics*, 71:23–49, 2003.

[179] A. Maki. Liquidity constraints: a cross-section analysis of the housing purchase behavior of Japanese households. *The review of economic and statistics*, 75 (3):429–437, 1993.

[180] A. Manca and P.C. Austin. Using propensity score methods to analyse individual patient level cost effectiveness data from observational studies. Technical report, Health, Econometrics and Data Group, The University of York, Working Paper 08/20, 2008.

[181] C. F. Manski and S. R. Lerman. The estimation of choice probabilities from choice based samples. *Econometrica*, 45 (8):1977–1988, 1977.

[182] C.F. Manski and D. McFadden. *In Structural analysis of discrete data with econometric applications*, chapter Alternative estimators and sample designs for discrete choice analysis. Cambridge, Mass., U.S.A.: Massachusetts Institute of Technology Press, 1981.

[183] G. Mansuri. Credit layering in informal financial markets. *Journal of Development Economics*, 84 (2):715–730, 2007.

[184] D. Marchesi. Giustizia civile e sistema economico. In *La competitivita' dell'Italia*. O. M. Petracca, Milan: Il sole 24 Ore, 146-172.

[185] A. Mauri. The role of financial intermediation in the mobilization and allocation of household savings in developing countries: Interlinks between organized and informal circuits, the case of Ethiopia,. Technical report, Paper Presented at the International Experts' Meeting on "Domestic Savings Mobilization through Formal and Informal Circuits: Comparative Experiences in Asian and African Developing Countries", Honolulu, 2-4 June., 1987.

[186] M. Mauss. *The Gift: forms and functions of exchange in archaic societies*. London: Routledge, 1990, 1925.

[187] D. McFadden. Qualitative response models. In *Handbook of Econometrics*. In Z. Griliches and M. D. Intrilligator (eds.), Amsterdam: North-Holland, 1983.

[188] K. McGarry and R.F. Schoeni. Transfer behavior in the health and retirement study: Measurement and the redistribution of resources within the family. *Journal of Human Resources*, Supplement:184226, 1995a.

[189] K. McGarry and R.F. Schoeni. Transfer behavior within the family: Results from the asset and health dynamics survey. Technical report, National Bureau of Economic Research Working Paper no. 5099. Cambridge, Mass.: National Bureau of Economic Research, 1995b.

[190] S. M. McKernan, M. Pitt, and M. Moskowitz. Use of the formal and informal financial sectors: does gender matter? Empirical evidence from rural Bangladesh. Technical report, World Bank Policy Research Working Paper, 3491, 2005.

[191] M. S. Mohieldin and P. W. Wright. Formal and informal credit markets in Egypt. *Economic Development and Cultural Change*, 48 N. 3:657–670, 2000.

[192] J. Morduch. Income smoothing and consumption smoothing. *The Journal of Economic Perspectives*, 9(3):103–114, 1995.

[193] J. Morduch. Does microfinance really help the poor? New evidence from flagship programs in Bangladesh. Technical report, Unpublished Working paper. Available at http://www.wws.princeton.edu/ rpds/macarthur/downloads/avgimp 6.pdf, 1998. Unpublished Working paper. Available at http://www.wws.princeton.edu/ rpds/macarthur/-downloads/avgimp 6.pdf.

[194] J. Morduch. Between the state and the market: Can informal insurance patch the safety net? *World Bank Research Observer, Oxford University Press*, 12(2):187–207, 1999.

[195] J. Morduch. The microfinance schism. *World Development*, 28 (4):617–629, 2000.

[196] A. Munnell, G. M. B. Tootell, L. E. Browne, and J. McNeaney. Mortgage lending in Boston: Interpreting HMDA data. *American Economic Review*, 25:25–53, 1996.

[197] K.M. Murphy and R. H. Topel. Estimation and inference in two-step econometric models. *Journal of Business & Economic Statistics*, 3 (4):370–379, 1985.

[198] D. W. Mushinski. An analysis of offer functions of banks and credit unions in Guatemala. *Journal of Development Studies*, 36 N.2:88–112, 1999.

[199] G. Nagarajan, R.L. Meyer, and L.J. Hushak. Segmentation in the informal credit markets: the case of Philippines. *Agricultural Economics*, 12:1717–1781, 1995.

[200] J. S. Nankumba. An evaluation of supervised and unsupervised agricultural credit systems in Lilongwe district of Malawi. Technical report, University of Malawi, 1980.

[201] S. Navajas, J. Conning, and C. Gonzalez-Vega. Lending technologies, competition and consolidation in the market for microfinance in Bolivia. *Journal of International Development*, 15:747–770, 2003.

[202] G. I. Nwanna. Financial accessibility and rural sector development. *Savings and Development*, 19(4):453–491, 1995.

[203] E. Ostrom. *Understanding Institutional Diversity*. Princeton, NJ: Princeton University Press, 2005.

[204] A. J. Padilla and M. Pagano. Endogenous communication among lenders and entrepreneurial incentives. *The Review of Financial Studies*, 10:205–236, 1997.

[205] A. J. Padilla and M. Pagano. Sharing default information as a borrower discipline device. *European Economic Review*, 44:1951–1980, 2000.

[206] A. Pagan and A. Ullah. *Nonparametric econometrics*. Cambridge University Press, Cambridge, 1999.

[207] A.R. Pagan and F. Vella. Diagnostic tests for models based on individual data: a survey. *Journal of Applied Econometrics*, 4:S29–S59, 1989.

[208] M. Pagano and T. Jappelli. Information sharing in credit markets. *Journal of Finance*, 48:1693–1718, 1993.

[209] M. Pagano and T. Jappelli. Information sharing, lending and defaults: Cross-country evidence. Technical report, CEPR Discussion Papers 2184, 1999.

[210] S. Pal. Household sectoral choice and effective demand for rural credit in India. *Applied Economics*, 34(14):1743–1755, 2002.

[211] A. Pankhurst and D. H. Mariam. The iddir in Ethiopia: historical development, social function and potential role in hiv/aids prevention and control. *Northeast African Studies*, 7(2):35–58, 2000.

[212] C. Paul, W. M. Mason, D. McCaffrey, and S. A. Fox. What should we do about missing data? Technical report, California Center for Population research. UCLA, 2003.

[213] C. Paxson. Using weather variability to estimate the response of savings to transitory income in Thailand. *American Economic Review*, 82(1):15–33, 1992.

[214] K. Pearson. On lines and planes of closest fit to systems of points in space. *Philosophical Magazine*, 2 N. 6:559572, 1901.

[215] S. Pejovich. The effect of the interaction of formal and informal institutions on social stability and economic development. *Markets and Morality*, 2 (2):164–181, 1999.

[216] C.A. Pissarides. Liquidity considerations in the theory of consumption. *Quarterly Journal of Economics*, 92:279–296, 1978.

[217] M. Pitt and S. Khandker. The impact of group-based credit programs on poor households in Bangladesh: Does the gender of participants matter? *Journal of Political Economy*, 106 N. 5:958–995, 1998.

[218] J.-P. Platteau. Solidarity norms and institutions in agrarian societies : Static and dynamic considerations. In *in S. Kolm, J. Mercier-Ythier, and G. Varet (eds.), Handbook on Gift-Giving, Reciprocity and Altruism*. Amsterdam: North-Holland and Elsevier, 2004.

[219] J.-P. Platteau and A. Abraham. An inquiry into quasi-credit contracts: the role of reciprocal credit and interlinked deals in small-scale fishing communities. *Journal of Development Studies*, 23:461 – 490, 1987.

[220] J. L. Powell. Least absolute deviations estimation for the censored regression model. *Journal of Econometrics*, 25:303–325, 1984.

[221] R. Putnam. The prosperous community social capital and public life. *American Prospect*, 13:35–42, 1993.

[222] M. Raturi and A. V. Swamy. Explaining ethnic differentials in credit market outcomes in Zimbabwe. *Economic Development and Cultural Change*, 47 N. 3:585–604, 1999.

[223] L. K. Raut and L. H. Tran. Parental human capital investment and old age transfers from children: is it a loan contract or reciprocity for indonesian families? *Journal of Development Economics*, 77:389–414, 2005.

[224] S. Ravi. Borrowing behaviour of rural households. Technical report, Working Paper, Department of Economics, New York University, 2003.

[225] D. Ray. *Development economics*. Princeton University Press, 1997.

[226] R. Reeser, R. Biden, S. Hoben, and T. B. Hobgood. Malawi Union of Savings and Credit Cooperative Organisations: an impact evaluation. Technical report, Development Alternatives Inc. and Institute for Development Anthropology. Report Submitted to USAID, Africa Bureau, 1989.

[227] L. J. Robinson and A. A. Schmid. *Handbook of Behavioral Economics*, chapter Interpersonal relationships and preferences: Evidence and implications., pages 347–360. Elsevier, 1988.

[228] D. Rodrik, A. Subramanian, and F. Trebbi. Institutions rule: The primacy of institutions over geography and integration in economic development. *Journal of Economic Growth*, 9:131–165, 2004.

[229] W. H. Rogers. Regression standard errors in clustered samples. Technical report, Stata Technical Bulletin Reprints Vol. 3, pp. 88-94, 1993.

[230] P. R. Rosenbaum and D. B. Rubin. The central role of the propensity score in observational studies for causal effects. *Biometrika*, 70:41–50, 1983.

[231] M. Rosenzweig and H. P. Binswanger. Wealth, weather risk and the composition and profitability of agricultural investments. *Economic Journal*, 103 (416):56–78, 1993.

[232] M. Rosenzweig and K. Wolpin. Credit market constraints, consumption smoothing and the accumulation of durable production assets in low-income countries. *Journal of Political Economy*, 101 (2):223–244, 1993.

[233] M. Rosenzweig and K. Wolpin. Parental and public transfers to young women and their children. *American Economic Review*, 84(5):1195–1212, 1994.

[234] M. R. Rosenzweig. Risk, implicit contracts and the family in rural areas of low income countries. *Economic Journal*, 98:1148–1170, 1988.

[235] D.P. Rothwell. Use of varying seasonal weights in price index construction. *Journal of the American Statistical Association*, 53 (281):66–77, 1958.

[236] A. D. Roy. Some thoughts on the distribution of earnings. *Oxford Economic Papers*, 3:135–146, 1951.

[237] D. B. Rubin. Estimating causal effects of treatments in randomized and nonrandomized studies. *Journal of Educational Psychology*, 66:688–701, 1974.

[238] D.B. Rubin. Bias reduction using Mahalanobis-metric matching. *Biometrics*, 36 (2):293–298, 1980.

[239] D.B. Rubin. *Multiple imputation for non-response in surveys.* John Wiley & Son: New York, 1987.

[240] H. Ruthenberg. *Farming systems in the tropics.* Oxford: Clarendon Press, 1971.

[241] M. Sahlins. *Stone-age Economics.* Chicago: Aldine Atheton, 1972.

[242] R. F. Schoeni. Private interhousehold transfers in rural China: are families altruistic? *Journal of Development Studies*, 33:487–511, 1997.

[243] G. Secondi. Private monetary transfers in rural China: Are families altruistic? *Journal of Development Studies*, 33(4):487511, 1997.

[244] J. Shea. Instrument relevance in mutlivariate linear models: a simple measure. *NBER Technical Working Paper*, 193, 1996.

[245] B. Sianesi. An evaluation of the active labour market programmes in Sweden. *Review of Economics and Statistics*, 86 (1):133–155, 2004.

[246] B. Silverman. *Density estimation for statistics and data analysis.* Chapman & Hall, London, 1986.

[247] J. Smith and P. Todd. Does matching overcome lalondes critique of nonexperimental estimators? *Journal of Econometrics*, 125(1-2):305–353, 2005.

[248] R. Smith and R. Blundell. An exogeneity test for a simultaneous equation tobit model with an application to labor supply. *Econometrica*, 54:679–685, 1986.

[249] A. Soyibo. Financial linkage and development in Sub-Saharan Africa: A study of informal finance in Nigeria. Technical report, Report of a Study prepared for the Overseas Development Institute, London, 1994.

[250] J. Stiglitz. Peer monitoring and credit markets. *World Bank Economic Review*, 4 (3):351–366, 1990.

[251] J. E. Stiglitz and A. Weiss. Credit rationing in markets with imperfect information. *The American Economic Review*, 71(3):393–409, 1981.

[252] R. Thaler and S. Bernartzi. Save more tomorrow: Using behavorial economics in increase employee savings. *Journal of Political Economy*, 112:164–187, 2004.

[253] J.P. Thomas and T. Worrall. Gift-giving, quasi-credit and reciprocity. Technical report, Discussion Paper Series, Department of Economics, University of St. Andrews, 2000.

[254] R. Townsend. Risk and insurance in village India. *Econometrica*, 62:539–591, 1994.

[255] C. Udry. Credit markets in Northern Nigeria: credit as insurance in a rural economy. *The World Bank Economic Review*, 4 N. 3:251–269, 1990.

[256] C. Udry. Risk and insurance in a rural credit market: an empirical investigation in Northern Nigeria. *Review of Economic Studies*, 61, 1994.

[257] T. van Bastelaer. Imperfect information, social capital and the poors access to credit. Technical report, IRIS Center Working Paper No. 234, 2000.

[258] K. Vinha. A primer on propensity score matching estimators. Technical report, CEDE 2006-13, 2006.

[259] S. Vyas and L. Kumaranayake. Constructing socio-economic status indices: how to use principal components analysis. Technical report, Published by Oxford University Press in association with The London School of Hygiene and Tropical Medicine., 2006.

[260] R. L. Williams. A note on robust variance estimation for cluster-correlated data. *Econometrica*, 56:645646, 2000.

[261] O. E. Williamson. The new institutional economics: taking stock, looking ahead. *Journal of economic literature*, 38 (2):595–613, 2000.

[262] M. Woolcock and D. Narayan. Social capital: implications for development theory, research and policy. *The World Bank Research Observer*, 15(2):225–249, 2000.

[263] J. M. Wooldridge. *Econometric analysis of cross section and panel data*. The MIT Press, 2002.

[264] B. Wydick. The effect of microenterprise lending on child schooling in Guatemala. *Economic Development and Cultural Change*, 47 (4):853–869, 1999.

[265] S. Zeldes. Optimal consumption with stochastic income: Deviations from certainty equivalence. *Quarterly Journal of Economics*, 104(2):275298, 1989.

[266] S. Zeldes. Consumption and liquidity constraints: an empirical investigation. *Journal of Political Economy*, 97:305–346, 1989b.

[267] M. Zeller. Determinants of credit rationing: A study of informal lenders and formal groups in madagascar. *World Development*, 22 (12):18951907, 1994.

[268] M. Zeller and M. Sharma. Rural finance and poverty alleviation. Technical report, International Food Policy Research Institute, 1998.