WORKING PAPER 181/2019

# Household Choice of Financial Borrowing and Its Source: Multinomial Probit Model with Selection

**Kanika Rana**
**Brinda Viswanathan**


**MADRAS SCHOOL OF ECONOMICS**
**Gandhi Mandapam Road**
**Chennai 600 025**
**India**

**May 2019**

# *Household Choice of Financial Borrowing and Its Source: Multinomial Probit Model with Selection*

**Kanika Rana**
**Corresponding Author**
**Consultant, Ernst and Young, Gurgaon**
aqf16kanika@mse.ac.in

**and**

**Brinda Viswanathan**
**Professor, Madras School of Economics, Chennai, India**
brinda@mse.ac.in

**WORKING PAPER 181/2019**

**May 2019**

**Price : Rs. 35**

**MADRAS SCHOOL OF ECONOMICS**
**Gandhi Mandapam Road**
**Chennai600 025**
**India**

**Phone: 2230 0304/2230 0307/2235 2157**
**Fax: 2235 4847/2235 2155**
**Email : info@mse.ac.in**
**Website: www.mse.ac.in**

# Household Choice of Financial Borrowing and Its Source: Multinomial Probit Model with Selection

**Kanika Rana and Brinda Viswanathan**

## Abstract

*In developing countries, the economically disempowered, borrow from multiple sources and also have multiple borrowings notwithstanding that some may be unable to access any form of credit. To ensure a greater amount of financial inclusion, it becomes necessary to understand what determines the choice between alternative loan source combinations while taking into account that borrowers may have distinct characteristics from non-borrowers.*

*Access to formal credit sources, are elusive for the disadvantaged due to different demand and supply side perceptions. Microfinance institutions (MFI) play an intermediate role having some attributes of the informal network and some similar to formal institutions. This study uses an observational data set for 2011-12 to analyse the role of socio-economic-demographic characteristics in the household's choice for different types of loan sources. In particular, the extensive nature of data allows us to study the mediating role played by MFI through its linkages with formal and informal sources.*

*The results of Multinomial Probit with Heckman selection, to account for non-borrowing households, reveal that where institutional sources are still a preferred option for the relatively advantaged section of the population, presence of microfinance loans in combination with other loan sources has contributed in ensuring greater equity in credit access to all. However, women headed households or dalit households with lesser opportunities of networking are less likely to take credit from formal sources*

**Key words:** *Household credit and sources, formal and informal institution, microfinance institutions, multinomial probit, Heckman selection*

**JEL Codes:** *C35, E51, G21*

# Acknowledgement

*The following submission is a revised version of a part of the first author's master's thesis submitted to MSE. We thank Sowmya Dhanraj and Sabuj Mondal for their useful comments and suggestions and the seminar participants at MSE seminar retreat (1-2 February, 2019) and 5th International conference on South Asia Economic development at SAU University (21-22, February 2019) for their valuable comments.*

**Kanika Rana**
**Brinda Viswanathan**

# INTRODUCTION

Our worlds within this world are demarcated by deep seated crevices of inequality. Individuals have no control over being born either to arms of luxury or to an abject state of poverty. A state of deprivation lulls them into a vicious, unending circle of despondency where they are often reminded of their inadequacies, both social and economic, disempowering them to fight this imbalance.

It is of paramount importance that such individuals are enabled to permeate through these layers and break the shackles of poverty and inequality. Empowerment and the different ways to achieve it, therefore, emerge to be a widely debated issue across the world. Along with having an external aspect to empowerment, relating to health, education, literacy etc., there are internal extensions to the concept too. These are associated with having an absolute control over one's life where choices are not evasive and opportunities are plenty. Any improvement in either of the two spheres pulls the individual a little closer to a worthy, and much coveted lifestyle that is at par with others. Therefore, there is not only a need for having substantive freedom across social, economic, cultural and political structures, but also individual freedom to choose and achieve different outcomes (Sen, 1988). Of these different dimensions, economic empowerment is truly enabling, as it has spillover effects in different areas of human life.

In a developing country like India, several productive activities are still largely informal be it agricultural or non-agricultural and hence household credit could be used for consumption or production or both. Consumption needs are also culturally determined for instance for a planned auspicious event or a sudden inauspicious event while production activities could suddenly take a different turn when there is a calamity either a natural disaster or the demise of the only breadwinner of the family. Thereby households depend on multiple sources of credit

with multiple borrowings. For a socially and economically diverse country like India it becomes important to study the determinants of not just credit access but how the determinants differ across types of credit sources. This will enable a better understanding of equity in access to credit in general.

Credit facilities, more often than not, are elusive for the economically disempowered due to different demand side and supply side perceptions. For example, banking institutions and other formal credit giving institutions, are unable to serve the poor. These institutions have a conventional bottom line to measure their financial performance that relates solely to achieving higher levels of profitability. Lending out to economically poor individuals would mean addition of "high risk" and "high cost" customers in their portfolios. Lack of interest towards such borrowers, who need credit but do not have the "desirable characteristics" to get credit, builds a gap in economic and overall well-being in the society[1].

Theoretical and empirical studies analyze the impact of credit, from (1) efficiency and/or (2) equity. Efficiency analysis relates to understanding the impact that the acquisition of loans from these sources has had on different aspects of livelihood, and equity analysis relates to understanding the ability to acquire these funds from a source at an affordable rate and at the desired time. Our study falls in the latter group where we try to see how equitable access to credit is when the households reveal credit source options. Hermes et.al (2008) describe efficiency as the ability of the organizations to provide credit in a financially sustainable manner. This relates more to the efficient means of providing credit to the households but in developing countries this may be in conflict with equity considerations wherein some households may

[1] We use the words credit, (financial) borrowing and loan interchangeably to convey similar meanings of access to finance other than one's own savings or liquidating one's own asset for the purpose of

be excluded access by these formal organizations. The high transaction cost associated with extending loans to the excluded section of the population might not be the only cause for this inequity. A study done in Ethiopia revealed that among other factors corruption was a statistically significant determinant of both access to credit and credit source choice of SME entrepreneurs (Fufa 2016)[2]. Therefore, emphasizing the need to strengthen the regulatory framework and its implementation to address both the efficiency and equity considerations.

In developing countries, as formal organizations limit their credit availability to the lower income groups and these households in order to avoid a compromise on their investment and consumption needs borrow from multiple sources. Moreover, the alternative sources of credit are not bound by formal contracts. Such informal sources like moneylenders, friends, and relatives satisfy the credit needs but in many instances, try to capitalize on this opportunity, causing indebtedness and further lowering of their economic well-being.

These two kinds of arrangements in the credit markets –formal and informal-can be compared using different metrics, e.g. ease of access, cost of acquiring funds, repayment schedules, interest rates, tenure of loans etc. Even though formal institutions as compared to informal networks charge a lower rate of interest, there is often a lack in the ease of acquiring credit from these sources. Moreover, formal institutions tend to have higher transaction costs and lower reach to the population. Therefore, even though they deliver credit at a more reasonable rate, the extent of their services is quite often compromised (Conning and Udry 2005).

A specially reengineered institutional structure to bring the expectations of the lenders in line with the groups that they intend to

[2] This study analyses it from the perspective of small and medium enterprises for Ethiopia and the credit sources are divided into Formal sources, informal sources and NGO credit program

lend to came into existence in late 1970s. Institutionalization of microfinance institutions meant a targeted approach to acquire the excluded borrowers and provide them with a more accessible channel of funds with better protection of their rights.

Sources of credit often classify under three broad headings- formal sources, informal lending networks, microfinance institutions (Pradhan 2013). The choice of credit source largely involves four levels of decision making[3] (1) Decision to borrow/ not borrow (2) Decision to apply to a particular source of lending (3) Decision of the source to approve the loan (4) Decision to finally borrow from any one or multiple sources that approved the loan. Note that at (3) there may not be sanctioning of entire loan amount. In such a case at level (4) one may bargain for a higher amount more likely from a less formal source to augment the lower amount approved from one of sources (usually, formal). Alternatively (or alongside), in (3) if the cost of borrowing is higher from one source, then even if the full loan amount is approved, it may be preferred by the borrower to borrow a lesser amount from this more costly source and complement it with borrowing from other sources. This leads to multiple sources of borrowing. Premise of the study is that the differences in the socio-economic-demographic characteristics of the household will affect the revealed preference for a combination of sources over another.

Given the difficulty in collecting information on borrowings from several sources, most large-scale surveys collect such information relating to purpose of loan, interest rates etc. only on largest loan borrowed. Nevertheless, they tend to be rich in information relating to access of different loan sources.

[3] It is usually the individuals who borrow but often, secondary data sources provide the information at the household level where different members may have borrowed from different sources and hence we refer to the borrower as the household in this study.

Using the IHDS data for 2011-12, this paper seeks to understand the determinants of decision revealed in step (4) above, while factoring in any selection bias that may arise due to the decision made in step (1).[4] This dataset also allows us to focus on the decision of the households to borrow from microfinance institutions as a separate category.

Several authors have studied the differences in characteristics of the household borrowing across different sources (Langer, 2009; Mitra and Venkatachalam, 2018; Chakraborty and Gupta, 2017). However, most of these studies use only primary survey data of a small region of the country while this study uses a large secondary data that is representative of the country's social and demographic diversities. This paper adds to the existing literature by not only looking at the access to a particular loan source but attempts to model how the households are opting for different combinations of these sources. It adopts a far more exhaustive categorization of the sources of loans that incorporates household's borrowings from multiple sources. Adequate attention to model the decision to borrow from multiple loan sources (in general) with microfinance (in particular) as an important credit source is allowed for in this analysis. It also accounts for the non-borrowing households in the dataset using a selection model. For a policy inference, the study focuses on the commonly classified types of credit markets-formal, informal and semi-formal (or microfinance).

Organization of the rest of the paper is as follows: The next section provides a brief background of the study which is followed by Section 2 on description of the data used for analysis. Section 3 on methodology elaborates the modeling technique arising out of the strength and limitations of the data available. This is followed by

[4] Note that most data sets give information relating to (1) and (4) but in this dataset, we have information relating to (2) and (3). We are overlooking steps (2) and (3) in this study because of the complicated structure of analysis as, on the one hand it has to be modeled as a multiple hurdle model and on the other hand there are very few supply side variables which could lead to an incorrect assessment of the determinants of the step-wise decision process.

presenting the main findings of the econometric analysis based on the multinomial probit model for choice of loan source while accounting for the selection of only those households, which report borrowing. The last section concludes the discussion while summarizing the finding of the study.

## LITERATURE REVIEW

### Credit Access and Household Well-Being

In 2010, poverty rate was one-half of that in 1990, as the world consciously worked towards its goal of poverty reduction. In spite of this achievement, a large number of individuals are still stuck in the cobwebs of destitution (Beck 2015). Close to 1.4 million individuals live on less than one dollar a day, and are constantly struggling from social and financial exclusion (Beck, 2015; Alimukhamedova, 2013). Dynamic Poverty, commonly known as vulnerability to poverty, adds another dimension to the abject state of misery that these individuals have to live in. Lack of stability due to unevenness in consumption patterns further perpetuates their state by straining their capabilities (Ingves 2005).

Empirical evidence has clearly shown that poverty cannot be defined as a deprivation in a single dimension of income or consumption and involves multiple dimension among which one way to characterize that would be through the aspects of natural, physical, social, and financial capital. Access to finance has been compared with access to basic needs such as water, education, and health services (Peachey and Roe, 2004; Campero and Kaiser 2013). Theoretical models usually envisage unidirectional linkages with access to credit helps individuals acquire physical capital that they can turn into economic rewards to pull themselves out of this life of dearth (Ray 2007). Institutional credit for instance, enhances production, productivity, profits and consumption and households with limited access to financial assets or those from socially disadvantaged groups are discriminated in this formal credit market

(Mitra and Venkatachalam, 2018). This shows that there are interlinkages across different forms of capital and missing out on or the other leads to a poverty trap.

In a globalized world, integrated through increased exposure to media, the aspirations and needs of an average citizen have grown. Desire to acquire better social status is increasingly being dictated by ownership of material goods, providing children with quality education, accessing healthcare facilities etc. This raises the pressure on earnings of the household. Since the income and asset levels are often unable to afford these facilities, there is rising demand for credit across households (Sarap and Venkatanarayana 2016).

Credit markets in developing nations do not behave competitively and tend to have a dual structure within them (Chakraborty and Gupta, 2017). This duality arises due to formal institutions that provide credit at a lower cost but less inclusive while informal lending networks have broadly reverse characteristic. The formal and informal markets are not completely segmented but there is a continuum and microfinance institutions for instance straddle somewhere in between as semi-formal institutions.

### *Sources of Household Credit*

The low-income households are usually occupationally diverse within and between households and this could be a strength as diversity spreads the risk while this may also be a limitation due to segmentation of the market and a single type of credit supplier may not find it easy to operate in this market. Same also holds for credit needs for varied consumption needs.

Formal institutions like banks (commercial banks, cooperative banks, government programs etc.) are primarily engaged in accepting deposits from public for active credit operations. Accessibility and affordability is essential to acquire credit from a formal source that

enables borrowers to improve their income and consumption levels. For instance, considering the agricultural sector, it is essential for buying the raw materials and inputs during the sowing season (Mitra and Venkatachalam, 2018). A small farmer may not be able to generate enough revenue to invest again for next season. The low household income would also mean that they cannot save enough and deposit in a bank and which could then be used as collateral for borrowing. Such individuals are unable to access conventional financial products like bank accounts, credit, remittances and payment services, financial advisory services, insurance facilities etc. (Singh 2010). World Bank estimates that because of the inability of about 87 percent low income households in securing a loan from a formal source, most individuals end up not borrowing at all or borrowing from informal sources of credit where interest rate range is about 48 percent -120 percent (Beck, 2015; Singh 2010).

The financial exclusion of such households is a result of the presence of asymmetric information in credit markets, where the financial intermediaries are unable to assess the customer's credit worthiness or bankability. This may materialize in two forms (a) inaccurately predicting the viability of investments using the borrowed amount (b) inability in assessing the repayment potential of the borrower (Armendariz and Morduch, 2010; Langer, 2009).

Poorer borrowers on the other perceive that the manner of assessing their creditworthiness by the formal sector lenders may dissuade them from applying and a further lower approval rate leaves them in a minority. This in turn makes the formal sector lenders not finding it profitable to invest their resources to learn about these customers. When it comes to making these services available, there are certain parameters and principles on which banking service providers operate. Their prime motive is to achieve long-term sustainability and high profitability. Due to these parameters and principles, the

conventional financial institutions have not been able to successfully target and serve the poor (Khandker, 1998; Beck 2015). High transaction costs associated with such lending leads to charging of higher rates of interests by these institutions and if they are high enough then individuals will not be able to pay back loans, thereby moving these services farther away from this segment of the population (Khandker 1998).

However, the exclusion of the poor households from the system is not just a supply side issue but it also has demand side barriers to it. Lack of financial literacy, or unavailability of formal documentation to be submitted with loan application, and a lack of inclination to seek formal credit i.e. self-selection (Beck 2015) are supply side barriers. Even though increasing the penetration of formal finance services across geographical locations and socio-economic boundaries resolves the supply side issues, demand side issues also require attention while discussing availability of these services. Moreover, individuals are voluntarily excluded from the formal credit market if they are not aware of the service, or assume a priori rejection (Campero and Kaiser 2013).

Restricted access to institutional sources of credit and its inability to permeate different segments of our society makes households reach out to other available alternatives. One of the major markets is the non-institutional sources or informal networks of lending. According to Sarap and Venkatanarayana (2016), among the non-institutional sources of credit, there are three main important players i.e. shopkeepers, moneylenders and employers but seem to exclude friends and relatives perhaps because they give interest-free loans. They conclude from an analysis of all-India secondary source data and a primary level study that the dependence of labor households for credit on non-institutional sources has not reduced but has been increasing in India.

In the financial landscape, money lenders are often viewed as monopolist (exploitative) that draw profits from their poor client base that has lesser options to acquire capital by offering interest rates that are well above the rates in formal banking sector (Armendariz and Morduch 2010). Individuals, as a result, tend to borrow from these sources out of desperation and they further add that informal networks have this monopoly power because their potential competitors are lacking the necessary information about the local markets and are therefore unable to make a break into it.

The dependency on non-institutional sources of credit is considerably high in rural India. In a village economy, strengthening of unbalanced social relationships (landlord/laborers) takes place due to the existence of informal credit markets, among other things. The higher interest rates induce dependency of laborers and catch them in a cycle of perpetual bondage (Sarap and Venkatanarayana 2016). Researchers have found that in places where there is availability of access to formal loans, informal sources still remain a popular choice due to ease of acquiring these loans (Pal 2002). Banerjee and Duflo (2007), observe a similar instance in Hyderabad wherein low-income households borrow largely from non-institutional sources, in spite of having access to formal sources of credit.

In sum, the formal sources are able to provide credit at a lower cost but there is higher transaction cost and therefore a greater plausibility of exclusion associated with this source of borrowing, whereas informal sources charge a higher interest but is able to provide credit with a larger reach. Often, households are also borrowing from a combination of both formal and informal source. The dynamic between the two markets and the roles they play in serving different segments of the population, sets the stage for the 'need' for a reengineered system. Microfinance institutions for instance play an intermediate role attempting to draw from the strengths of informal networks but operationally similar

to formal sources. However, the quantum of loan serviced or the manner in which the joint liability of the group materializes and the inability of the state to regulate has not brought in the kind of widespread success envisaged by this type of institution.

**Aspects of Equity and Efficiency**

The implications of efficiency and equity are different in case of formal and informal markets. Formal credit markets efficiently provide credit to the individuals as the rates charged by them are competitive, however, they are unable to reach out to a socio-economically disadvantaged section of the population. In the Indian context, Mitra and Venkatachalam (2018) observes that not only there is a small bias against scheduled tribes (ST) in loan approval decisions but the loan applications received from this socio-economically underprivileged group was statistically under-represented. The loan applications and not so much the loan disbursals seem to follow a stereotypical class/social hierarchy, even after accounting for the credit histories of the individuals.

Considering the informal markets, Armendariz and Morduch (2010) contends that these markets have a greater reach and but their monopolistic nature induces inefficiency in the market. Microfinance institutions are arguably able to reach out to the excluded section of the population, as well as, provide a channel of funds at reasonable interest rates. This therefore is able to increase efficiency and equity in the credit markets.

**Credit from Microfinance Institutions**

Introduction and institutionalization of modern techniques of microfinance by Dr. Mohammad Yunus, Professor at Chittagong University, Bangladesh in 1970s was the innovation of the century. He reengineered the then known credit market to overcome the barriers that were, up until then, preventing conventional banks from reaching out to

certain segments of our society and was restricting this segment to the informal sector.

The designing of these institutions kept in mind the inaccessibility of formal market for a certain section of the population. To cater to the low-income households, the operational structure had a viable repayment schedule, reasonable interest rates, penalty for default, and ease of procedure, among other things. Not only this but, for microfinance to be able to penetrate across geographic locations and reach out to the lowest sections of the society, there was also a need for the targeted population to be more accepting of microfinance activities (Armendariz and Morduch 2010). There is no single success parameter in this market. One of the targeted goals of microfinance institutions is to reach out to the financially weaker sections of the society and empower individuals to escape from the strangles of poverty, while maintaining their financial and operational viability. Balancing financial sustainability and reaching to poorest of poor is difficult and high-risk task for MFI (CGAP 1996).

Today, a borrower broadly chooses from the following when it has to borrow from the credit markets:

1. Formal credit sources
2. Informal Lending Networks
3. Microfinance Institutions
4. Combination of any of the above three[5]

To summarize, household's ability to acquire debt from a particular source of credit and transform the same into higher profits and stable consumption patterns is largely dependent on the financial services that these sources make available to them (Conning and Udry 2005). If these are limited, households may have to forego valuable investment

[5] Borrower may resort to multiple sources if there is credit rationing and the full loan amount is not sanctioned. The need for credit may not necessarily be fulfilled through a single source of borrowing

and income-generating activities and suffer the consequences of volatile consumption. To compensate for any rationing or constraints in provisioning of credit from either of them may make the household borrow from different combinations of these sources.

The socio-economic characteristics, demographic indicators, networking abilities etc. among other things may dictate the final choice of sources of credit. Different authors have tried to study and model this choice between different options.

**Decision to Borrow and Decision on Types of Loan Sources**

This section tries to review the literature, to understand how the determinants (or household characteristics) vary between the decision to borrow (Morduch, 1999; Pitt and Khandker, 1998) and then the decision on loan source(s).

*Household's decision to borrow*: At this level, the household makes a decision on whether it would like to borrow from any source or would not like to borrow at all or has the option of using its savings when it has adequate savings.

In a study on rural Vietnam Nguyen (2007) specified participation in a credit program as a function of household characteristics including gender of household head, age of household head, number of household members, highest level of education in household, agricultural work, value of house and land holding size; and of commune characteristics including distance from commune to the nearest government banks.

Within these variables, it was observed empirically that household size and agricultural work determined the financial activity rather than education or distance from the commune to the nearest bank branch. Moreover, education level has inverse U effect on borrowing possibility implying that the Household heads having university degree or

never attending school did not borrow much. Women also showed lower participation in credit activities because man as head of household significantly determined the borrowing probability. Further, number of members in the household had a large and significant effect on credit participation.

Swain (2007) suggests that farm production activities, socioeconomic characteristics of households, current and expected output prices and wages affect credit demand. Here, socioeconomic characteristics included age of the head of the household, family size, dependency ratio and primary activity of household head. Since primary data collected from villages of Puri, in the state of Odisha, was studied, the proportion of irrigated area was also a significant determinant of demand for loan. Additionally, the distance from cooperatives and the number of matriculates in household are also important determinants.

***Households' decision to borrow from a particular source*****:** Earlier discussion on credit markets of India, different options for the households are available if it decides to borrow. Between different sources of credit, there will be variability in interest rates so while making the choice between sources, cost of capital or the rate of interest and is an important consideration. Though the total loan amount may have been decided *a priori* as per the 'need' but the household may split the total amount across sources and the combination of loan amount, interest rate and repayment duration become important variables that determine the loan sources. Therefore, the decision making process may not be that easy to model as a sequential set of actions resulting in an observable outcome of multiple loan sources. Diagne (1999) suggest households are not discouraged in their borrowing decisions by further increases in the formal interest rate and/or the transaction costs associated with getting formal credit. However, Banerjee and Duflo (2007) observe that in sources of credit like informal lending networks

and microfinance institutions, if the interest rates are kept low then it may lead to a larger positive impact.

Tang *et. al.* (2017) finds that borrowings from formal sources rise with the increase in education level and the borrowing from informal sources either decreases or remains unaffected. Swain and Floro (2012), finds that SHG households had younger household members, with less non-land wealth, higher food consumption levels per capita per month and were residing in villages with easier access to public transport, primary health care centers, but were further from banking institutions compared to non-SHG households.

Household characteristic such as caste can have a say in accessing of the bank credit (Mitra and Venkatachalam, 2018). Langer (2009) found out that prior presence of social capital is advantageous to participating in microfinance programs but is not a precondition for such households to come together but they were relatively poorer, and largely from rural areas.

**Summarizing, the findings across these studies:**
Firstly the framework of analysis usually considers choosing from among one of the three or two credit sources:

(1) Formal sources and informal sources (Aarti and Chakraboraty, 2017; Diagne, 1999)
(2) Formal sources, Informal sources and Microfinance (Togba 2009)
(3) Microfinance Institutions, other sources of borrowing, and not borrowing (Langer 2009)

These studies have not considered the possibilities of some or all sources co-existing and hence how determinants vary when borrowing from multiple loan sources which are both formal and informal. By doing so the assumptions that presence of one excludes the other sources may

not capture the ground reality. Further the results for determinants of households' choice of credit source would be expectedly affected by this. As people self-select themselves into combinations of these categories, choice of multiple loan sources could play an important role in ensuring access that is more equitable across different sections of the population.

Our intention is to study the determinants of the choice and their differences between combinations of these sources. We will also be accounting for the differences between a borrowing and non-borrowing households through selection models. This study on determinants of the choice of combination of different source, accounts for accessibility and reach of credit sources. Since the main objective of microfinance is to increase the reach of credit facilities, special attention to MFI as a loan source choice is given. It is believed that adequate attention has not been given in the empirical literature to modeling the decision to borrow from multiple loan sources (in general) with microfinance (in particular) as an important credit source.

## DATA DESCRIPTION

The India Human Development Survey (IHDS) is a nationally representative, multi-topic Survey of 41,554 households in 1503 villages and 971 urban neighborhoods across India collecting data in two years 2004-05 and 2011-12 referred as IHDS-1 (Desai et. al 2009) and IHDS-2 respectively (Vanneman et.al 2017). This panel dataset has re-interview rate of about 83 percent, and this study has explored the panel feature of the data set in only a limited manner. IHDS-II (2011-12) has more probing questions particularly the data on choice of loan source is relatively more exhaustive and diverse.

The debt segment of the survey is based on several questions on household borrowing in the previous 5 years from the date of survey. Around 45.3 percent households have not borrowed from any credit

sources[6]. Households are known to borrow from multiple sources because of various demand and supply side considerations. To assess this overlap, the question "How many loans has the individual taken in the past 5 years" is analyzed in Table 1. This table records the share of households that respond to borrowing from the particular source as 'Yes'.

**Table 1: Percentage of Total Households Borrowing From Different Sources**

| S.No | Loan Source | Share (percent) |
|---|---|---|
| 1 | Bank/Govt. Program/Other Credit Program | 20.7 |
| 2 | MFI/Community Group/NGO/SHG | 9.9 |
| 3 | Money Lender | 16.5 |
| 4 | Employer | 3.2 |
| 5 | Relatives/Friends | 26.9 |
| 6 | Others | 1.6 |

**Note:** It covers 55 percent of borrowing households in the dataset and the remaining 45 percent households are non-borrowers however as the same household may have responded 'yes' to borrowing from one or more than one of the categories with the total exceeding 55 percent due to double counting.

**Source:** Author's own estimation using household level data IHDS (Round 2)

An analysis of the usage pattern of loan amounts and other related characteristics borrowed from different sources is more relevant but information on loan amount is available only for the largest loan. Thus focusing only on the nature of variation in largest loan amount borrowed and its determinants might not be representative of the total borrowings by the households. As shown in Table 1 there is more details available on the type of institutions for credit source and this study therefore, aims to understand how different demand side barriers (e.g. financial status, social status etc.) affect a household's choice of credit source from various available options of sources.

6 In NSS 70th Round (2013) survey, incidence of indebtedness is observed in 53.8 percent households. This implies that 46.2 percenthouseholds have not reported incidence of indebtedness with indebtedness defined access to credit.

## METHODOLOGY

### Model Features

A vast literature has attempted to model the impact of loans from a particular (single) source on household characteristics through the loan amount/interest rate equation, along with incorporating the determinants of this choice of source. Quoc (2012) identifies the determinants of decision to borrow and the amount borrowed by using a double hurdle model (DHM) (as constructed by Cragg (1971)) and the Heckman selection model. Swain (2007) investigates the demand and supply of credit in rural credit markets of Puri, Odisha, India, using type 3 Tobit models, which correct for sample selection bias and endogeneity of rate of interest and size of the loan. Mitra and Venkatachalam (2018) uses logit models to investigate access and the existence of taste based discrimination in India against the underprivileged castes. The unobserved credit histories are simulated with various distributions to account for that feature while assessing the role of discrimination.

Studies that analyse access to loan sources demand usually focus on households that are borrowing from single sources like either formal sources or informal sources or microfinance sources (Chakraborty and Gupta, 2017; Diagne, 1999; Togba 2009). Bhattacharya and Rajeev (2011) interpret access to credit as incidence of indebtedness to conclude on the distinct socio-economic characteristics between the households accessing formal credit versus informal credit. They continue further to indicate the SHG-bank linkage can be an effective means to bring them closer to formal sources of credit.

While analyzing household level loan source data, there are two key points that need to be taken into consideration:

(A) It is necessary to take into account the incidences of multiple borrowing. This study hypothesizes that borrowing from multiple loan sources and the choice between different combinations of

loan sources will be reflected by the socio-economic-demographic differences between households.

(B) Furthermore, there may be a non-random selection of households into borrowing households as the factors that affect the decision to borrow may be different from the factors that affect the decision to borrow from a particular loan source.[7] Household characteristics like dependency ratio, household size etc. may not affect the decision to apply for a particular loan source but are likely to affect the decision to borrow.[8] Omission of non-borrowers will not allow us to account for these selection issues.

Earlier studies in the Indian context have rarely analyzed these two features together: understanding the differences in determinants for multiple loan sources after accounting for possibly a non-random selection of borrowers. Following points show the decision of choice of source of credit through the existence of different levels of barriers/hurdles:

(1) Decision to borrow/not borrow
(2) Decision to apply to a particular source
(3) Decision of the source to approve the loan[9]
(4) Decision to finally borrow from any one or multiple sources that approved the loan

[7] While testing multiple models, the sign and the significance of the estimates of loan source choice model after omitting the non-borrowers from the dataset were different from the ones obtained if non-borrowers were included in the econometric model. This may indicate a plausible omission bias due to non-consideration of the non-borrowers.

[8] The household is more likely to report borrowing if the household size is greater (as per Table 3)

[9] Appendix (1) gives the table for rejection percentage by different loan source providers. The rejection percentage in formal categories is lower than the others. A reason for this may be that there is scheme based application in in formal sectors, resulting in a better approval rate. The categories of loan sources in the table might have some overlap that we were unable to acc

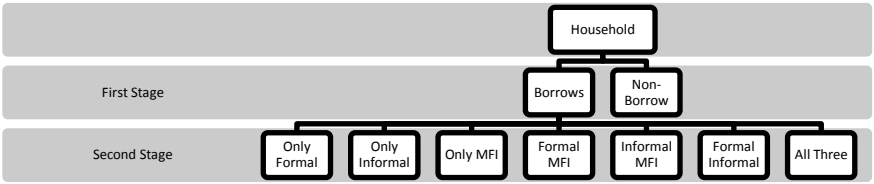
Non-borrowers could be characteristically different from the borrowing households, which may explain why they are choosing to be non-borrowers in the first place. After deciding to borrow, the household then makes a choice between the available sources of credit. The information on this choice is only available for the households that decided to borrow in the first place (Wynand and van Praag, 1981; Greene, 2009). This is a clear example of incidental truncation where the basic features of both groups (borrows and non-borrowers) will differ and will have to be accounted for. Step (4) is therefore dependent on the previous steps. We assume that when a household decides to apply to particular source step (2), the approval of loan happens in step (3) and the household will borrow from that source. Modeling step (2) and step (3) is out of the scope of this study as the complexity of the analysis increases. We will model step (4) after incorporating any selection bias that may arise due to step (1).

As seen in the previous section, 45.3 percent of our dataset consists of non-borrowers. To validate whether the nature of variables in borrowing and non-borrowing categories are significantly different from each other, we conduct a test of proportions and test of means. Table 2 shows that these two type households differ from each other significantly across all the characteristics.

The self-selection into the sample, therefore, creates a type of omitted variable problem manifesting from the fact that selection of households into the group of households that are borrowing may be a non-random event (Certo *et. al.* 2016). Borrowers are bound to be different from non-borrowers and the non-randomness in selection into the borrowing group might be a source of selection bias in the data (Greene 2009).To account for this, we use two-step Heckman selection model and the outcome equation (second stage) is modelled as nominal polychotomous variable with the selection equation as the binary variable

as is usually considered in the selection model and shown in Figure 1 below.

**Figure 1: Categorical Dependent Variable With Heckman Selection**




**Source:** Author's own categorization

**Empirical Strategy: Multinomial Probit Model with Selection**

As indicated in the Table 1, there are about 45 percent non-borrowers in the last five years. While analyzing the determinants of source of loan, we may also have to consider if the borrowers are non-randomly selected and some of the estimated coefficients would be biased in the model for choice of loan source. To account for this we estimate the multinomial probit model with selection.

The selection equation:

$$Z_i^* = W_i\gamma + u_i\,, \qquad (1a)$$

$Z_i^* = U_i^b - U_i^n$ Is taken as the difference in utility from borrowing ($U_i^b$) as against non-borrowing ($U_i^n$). If this difference is positive then the household borrows. As utility and hence $Z_i^*$ are unobserved (latent) variable, what is revealed is only the borrower or non-borrower status. That is,

$$Z_i \begin{cases} = 1 \; if \; Z_i^* > 0 \\ = 0 \quad Z_i^* < 0 \end{cases} \tag{1b}$$

The first stage of the model estimates the revealed preference to borrow ($Z_i$=1) or not ($Z_i$=0) using the probit model. In this case the model becomes:

$$Z_i = \Phi(W_i\gamma) + u_i, \tag{1c}$$

where $\Phi(.)$ is the CDF of the standard normal variable.

The outcome equation is a multinomial probit model and since it does not have a closed form solution we would express it in terms of the probability of choosing a loan source j.

As in Greene (2012), the structural equation for the MNP model with 7 choices is[10]

$$U_{ij} = \mathbf{X_i}\boldsymbol{\beta_j} + \varepsilon_{ij}, \; j=1,2,...7, \; [\varepsilon_{i1}, \varepsilon_{i2}, \varepsilon_{i3}....\varepsilon_{i7}] \sim N(\mathbf{0},\boldsymbol{\Sigma}).... \tag{2a}$$

Since this choice is conditional on the households choosing to borrow, so $\varepsilon_{ij}$ in (2a) is correlated with the $u_i$ in (1a) and hence (1c) for some or all j.

That is,

$$\begin{aligned} & E(U_{ij} \mid \boldsymbol{X_i}, Z_i^* > 0) \\ & = E(U_{ij} \mid \boldsymbol{X_i}, Z_i = 1) \\ & = E(\boldsymbol{X_i\beta_j} + \varepsilon_{ij} \mid \boldsymbol{X_i}, \boldsymbol{W_i\gamma} + u_i > 0] \\ & = \boldsymbol{X_i\beta_j} + E(\varepsilon_{ij} \mid u_i > -\boldsymbol{W_i\gamma}] \end{aligned} \tag{3a}$$

[10] https://www.stata.com/statalist/archive/2012-03/msg00191.html

Here $E(\varepsilon_{ij})=0 \ \forall$ j in (3a) if $\varepsilon_{ij}$ and $u_i$ are uncorrelated for each j. If $\varepsilon_{ij}$ and $u_i$ are correlated for each j, then there is an omitted variable bias and we need to account for $E(\varepsilon_{ij} \mid u_i > -\boldsymbol{W_i\gamma}]$ in 3(a).

$$E(U_{ij} \mid \boldsymbol{X_i}, Z_i^* > 0) = \boldsymbol{X_i\beta_j} + E(\varepsilon_{ij} \mid u_i > -\boldsymbol{W_i\gamma}]$$

$$E(U_{ij} \mid \boldsymbol{X_i}, Z_i^* > 0) = \boldsymbol{X_i\beta_j} + \rho_j\sigma_j \left[ \frac{\phi\left(\frac{\boldsymbol{W_i\gamma}}{\sigma_u}\right)}{\Phi\left(\frac{\boldsymbol{W_i\gamma}}{\sigma_u}\right)} \right] \quad (3b)$$

Here, $\rho_j$ is the correlation between $\varepsilon_{ij}$ and $u_i$ and $\sigma_j$ is the variance of $\varepsilon_{ij}$ for the $j^{th}$ outcome.

This result in (3b) above is from Theorem 19.2, Greene (2012)

$$= \boldsymbol{X_i\beta_j} + \rho_j\sigma_j \left[ \frac{\phi\left(\frac{\boldsymbol{W_i\gamma}}{\sigma_u}\right)}{\Phi\left(\frac{\boldsymbol{W_i\gamma}}{\sigma_u}\right)} \right] \quad (3c)$$

$$= \boldsymbol{X_i\beta_j} + \theta_j\lambda_i(\alpha) \quad (3d)$$

where $\theta_j = \rho_j\sigma_j \quad and \ \lambda_i(\alpha) = \left[ \frac{\phi\left(\frac{\boldsymbol{W_i\gamma}}{\sigma_u}\right)}{\Phi\left(\frac{\boldsymbol{W_i\gamma}}{\sigma_u}\right)} \right]$

where, $\sigma_u$ is the variance of $u_i$ in equation in (1c).

Thus (2a) is modified as

$$U_{ij} = \boldsymbol{X_i \beta_j} + \theta_j \hat{\lambda}_i(\alpha) + \varepsilon_{ij} \quad (4a)$$

$$U_{ij} = \boldsymbol{X_{i\lambda} \beta_{j\lambda}} + \varepsilon_{ij} \quad (4b)$$

$\boldsymbol{X}_{i\lambda}$ matrix now includes the additional column $\hat{\lambda}_i(\alpha)$ known as the Inverse Mills Ratio (imr), estimated from the probit model from estimating the first stage estimation equation (1c) and $\boldsymbol{\beta_{j\lambda}}$ has an additional coefficient $\theta_j$ to be estimated in equation (4b).

The omission of $\lambda_i(\alpha)$ imr term, (which captures the correlation between $\varepsilon_{ij}$ in (4a) with $u_i$ in (1c)) from the 2nd stage outcome equation (4a/4b) would bias some of estimated $\boldsymbol{\beta_j}$ coefficients. Thus, if estimated coefficient of imr ($\theta_j$ in 4a) is statistically significant then the (sample of) borrowers are not randomly selected.

In equation (2a) $U_{ij}$ is unobserved and what we observe is the outcome of choice for different mutually exclusive sources of borrowing:

$$Y_{ij} = \begin{cases} 1 \quad if \ \max(U_{i1}, U_{i2} .... U_{iJ}) = U_{i1} \\ 2 \quad if \ \max(U_{i1}, U_{i2} .... U_{iJ}) = U_{i2} \\ \\ J \quad if \ \max(U_{i1}, U_{i2} .... U_{iJ}) = U_{ij} \end{cases}$$

i=1,2,….N are the N observations in the sample. In this study J=7 outcomes represented as:

$$Y_i = \begin{cases} 1 & \text{Only Formal} \\ 2 & \text{Only Informal} \\ 3 & \text{Only MFI} \\ 4 & \text{MFI and Formal} \\ 5 & \text{MFI and Informal} \\ 6 & \text{Formal and Informal} \\ 7 & \text{All three} \end{cases}$$

Since $Y_{ij}$ is polychotomous discrete variable so the model specification is the multinomial logit or the multinomial probit model. However, the polychotomous model now also accounts for selection bias. So the distribution of error term that accounts for the correlation between the [$\varepsilon_{i1}$, $\varepsilon_{i2}$, $\varepsilon_{i3}$....$\varepsilon_{i7}$] of the polychotomous outcome model and the $u_i$ of the selection model is the multivariate normal distribution: [$\varepsilon_{i1}$, $\varepsilon_{i2}$, $\varepsilon_{i3}$....$\varepsilon_{i7}$, $u_i$]~N($\mathbf{0}$, $\boldsymbol{\Sigma}_\lambda$),

The maximum likelihood estimator maximizes

$$\ln L(\boldsymbol{\beta}_\lambda, \boldsymbol{\Sigma}_\lambda) = \sum_{i=1}^{N}\sum_{j=1}^{m} y_{ij} \ln p_{ij} \qquad (5a)$$

The vector $\boldsymbol{\beta}_\lambda$ includes parameters for all the explanatory variables in each of the outcomes j, including the imr $\theta_j$ parameter that varies for each of the outcome separately.

The $p_{ij}$ for j=q, any $q^{th}$ choice is defined as Prob[$choice_{iq}$] = $p_{iq}$

$p_{iq}$ = Prob[$(U_{iq}>U_{ij})$, j=1,2…J, j≠q]

$p_{iq}$ = Prob[

$$\mathbf{X_{i\lambda}\beta}_{q\lambda} + \varepsilon_{iq} > \mathbf{X_{i\lambda}\beta}_{1\lambda} + \varepsilon_{i1}, \ldots\ldots \mathbf{X_{i\lambda}\beta}_{q\lambda} + \varepsilon_{iq} > \mathbf{X_{i\lambda}\beta}_{J\lambda} + \varepsilon_{iJ}]$$

$p_{iq}$ = Prob[ $\varepsilon_{i1} - \varepsilon_{iq} < \mathbf{X}_{i\lambda}(\boldsymbol{\beta}_{1\lambda} - \boldsymbol{\beta}_{q\lambda})$.,....., $\varepsilon_{iJ} - \varepsilon_{iq} < \mathbf{X}_{i\lambda}(\boldsymbol{\beta}_{J\lambda} - \boldsymbol{\beta}_{q\lambda})$.]

$$p_{iq} = \int_{-\infty}^{V_{i1}-V_{iq}} \cdots \int_{-\infty}^{V_{iJ}-V_{iq}} f(\eta_{i1}\eta_{i2}..\eta_{iJ})d\eta_{i1}\eta_{i2}..\eta_{iJ}\ ,..... \qquad (5b)$$

where $\eta_{ij}=\varepsilon_{ij}-\varepsilon_{iq}$, $j\neq q$ $V_{ij}=\mathbf{X}_{i\lambda}\boldsymbol{\beta}_{j\lambda}$ and $f(\eta_{i1},\eta_{i2....}\ \eta_{iJ})$ is the joint probability density function (PDF) of $\eta_{i1},\eta_{i2},....\eta_{iJ}$. In this case, the PDF is a multivariate normal distribution and after substituting in the log likelihood function (5a) for each of the $p_{ij}$, the maximization is by numerical methods. To avoid the identification problem of the multinomial probit model, the set of $\boldsymbol{\beta}_{j\lambda}$ for one of the outcomes is restricted to zero. Furthermore, explanatory variables **W** and **X** in the selection and outcome equations respectively should have at least one variable that is not common. These two vectors of variables are discussed below.

## INDEPENDENT VARIABLES

The independent variables (**W**) in selection equation (decision to borrow or not) consist of:

- *General Characteristics*: (1) Dependency Ratio (2) Household size (3) Highest level of education of the household (4) Gender of the head of the household (5) Age of the head of the household (6) Main income source
- *Social*: (1) Major Social Event in the Household (2) Caste/religion (3) Region
- *Economic*: (1) Saving Index (3) Quintiles based on asset index (4) BPL household

The independent variables used to explain outcome equation (final outcome of borrowing from a particular source) consist of:

- *General Characteristics*: (1) Gender and age of head of household (2) Highest level of education of the household (3) Main Income Source (4) Region

- *Economic Status*: (1) Quintiles based on asset Index (2) If the household is Below Poverty line
- *Social Status*: (1) Caste/religion groups (2) Confidence in institutions (2) Horizontal Networks (3) Vertical Networks

Thus, following variables are used only in the selection equation (exclusion restriction) and not in the outcome equation: dependency ratio (2) household size (3) major event (4) savings index. The paper assumes that they affect the decision of the household to borrow but not its choice of credit source alternatives.

## RESULTS AND DISCUSSION

Appendix A gives a detailed description on the construction of all variables used in the analysis. Table 2 of descriptive statistics gives the population shares of households across different independent and dependent variables that have been used in the model. It further differentiates the total population into share of borrowing and non-borrowing households.

The head of the household is usually the primary decision making unit[11]. Characteristics like age, gender of household head may affect the decision of whether the household needs to borrow in the first place and if so, the choice of credit source made by the household (Nguyen 2007, Tang et al 2017). Age of the household head is used as a determinant for the decision to borrow and the decision to borrow from a specific source (Nguyen 2007). Highest level of education in the household, is a determinant of the level of financial awareness of the household, risk taking potential etc. (Tang et al, 2017).

[11] In the gender relations segment of the survey questionnaire, we observe that on aspects like purchasing of durable goods, land and property, how much to spend on social functions and marriages, the husband or senior male in the family, are the primary decision making unit and are usually the household head.

It is hypothesized that gender of the household head could be a mediating factor for access or denial of the loan. Both demand and supply side perceptions about the vulnerability of a household with a female head may result in the choice of certain types of loans, perhaps less of the formal type. It is likely that younger households will more readily adopt sources of credit like microfinance services etc. Education is added as a dummy variable to try and corroborate Nguyen's (2007) findings, who found that there was non-linearity associated with years of education with higher non-borrowers among household head with no education or university degree and larger number of borrowers in between these two categories. Further, non-borrowing households have a higher share of households with women heads, older heads and lesser proportion of households with highest level of education being less than primary education.

Income source determines stability of income level and therefore becomes an indirect determinant of credit. It also affects the need for business loans as in some professions there is a greater need for credit. Therefore, to account for both of these in the model, income source is added as a variable, both as a determinant of 'need to borrow' and 'choice of loan source'. There is a clear overrepresentation of borrowing households in income sources such as agricultural and allied services, agricultural laborers and non-agricultural laborers when we compare it to the population averages (Table 2). A larger proportion of non-borrowing households fall into salaried, business classes and others.

Quintiles based on asset index[12], added to control for the standard of living of the household, may affect both stages of decision-making. Where income is a flow variable, assets owned by the household is a stock variable. A household in a higher asset class will more likely

[12] See appendix table A.1 for details. Land holding is not incorporated in the index model as it can have different implications for rural and urban households; for instance, a landless agricultural laborer may not be the same as a landless urban businessperson.

borrow from formal sources of credit. A reason may be that they are more likely to be given a loan from a formal institution of credit.[13] Criteria to classify household as below poverty line/above poverty line is added in both the stages as this variable will indicate the need of the household to borrow as well as the sources it finally borrows from.

IHDS-II has rich information on social networks of the individuals. Questions covered under this aspect is on memberships of household in professional and social groups, their trust and confidence in institutions, political connections and activities.

Social status and social capital have linkages in the credit market and as discussed earlier, may dictate the choice of loan source of the household. Caste influences both the institution's decision to approve the loan, as well as household's decision to apply for loan from a particular loan source category (Langer, 2009; Mitra and Venkatachalam, 2018). Table 2 shows that the OBC and Dalit households are overrepresented in the borrowing households as compared to their overall population shares, possibly indicating a higher incidence of need to borrow in these communities.

To model the effect of faith in institutions, social connections and political connections, that forms the social capital of a household, three indices namely, (1) Confidence in institutions index (2) Horizontal networks index (3) Vertical networks index are constructed (Langer, 2009; Fufa, 2016; Unnikrishnan, 2016). As seen in the literature, strengthened institutional structures will affect the access to credit and credit source choice (Fufa, 2016). Confidence in institution index is constructed on the premise that there is a greater likelihood of the household to borrow from a formal structure of credit if it has higher

[13] The underlying endogeneity due to addition of this variable may affect the estimated values. However, since the asset index is stock variable that is representative of accumulated wealth of a household, the availability of higher loans may not directly impact the level of accumulation as this variable is dependent on the incomes of the previous period. The durable assets included in the index will be used for several periods and therefore don't reflect the current period's impact.

level of confidence and faith in the institutions and lower crime rate in the area.

Horizontal networking index is constructed using the variables on level of participation in community activities, neighborhood groups for socializing and local recreational institutions and clubs. Both the indices (constructed using Principal component analysis) are categorized into terciles and then aggregated as share of households in different terciles in the locality (primary sampling unit, PSU[14]) of the household (more details are in Appendix A). The variables in horizontal networking are community based and are a reflection of the strengthened horizontal ties of the individuals. The addition of these networking variables is based on the premise that horizontally networked household are more likely to secure a loan from microfinance institution as these institutions are built on the idea of 'group-based lending'.[15]

There is a need to differentiate between horizontal and vertical networks as the variables included in vertical network uses questions on having acquaintances in medical field, education field and government. These are mere acquaintances, 'useful' connections and they are not reflective of the solidarity and the coordination, which is what the horizontal networks try to capture (Unnikrishnan, 2016). In our study, we have added the vertical index as a household level variable and have not aggregated it at PSU level. This is because we believe that the vertical connections will be more of a household level variable and will not be affected by the interaction of households within the area (PSU). This variable, therefore, is not aggregated at a PSU level and is taken at the household level. Even though the level of confidence in institutions is not different for borrowers and non-borrowers but a larger number of households are borrowers if they have higher levels of horizontal networking and medium levels of vertical networking.

[14] Villages and urban blocks are the PSUs.
[15] This is as opposed to Unnikrishnnan (2016) who added it as a household level.

Variables such as dependency ratio and household size contribute to the decision on whether the household borrows or does not borrow. A larger household size means more members and therefore a greater need for income/credit. Dependency ratio is the number of children, homemakers and senior citizen (non-earning members of the household) as a proportion of the total members of the household. A greater dependency ratio will mean a greater demand for resource and perhaps a greater need for cash flow or liquidity constrained due to fewer earning members in the household. Keeping everything else constant, a household with higher dependency ratio when compared against the one with lower dependency ratio might be more in need of credit. These variables are therefore added as an exclusion restriction in stage 1 (Figure 1) of analysis.

Inclusion of savings index – the sum of total savings instruments reported by the households (see Appendix Table A)- in the stage 1 model as an explanatory variable helps us to see whether households using more number of saving instruments are less likely to borrow. The average number of savings instruments is about 1.26 compared to 1.36 for borrowing households and 1.15 for non-borrowers. There is no data on the amount of savings or investments so the number of instruments serves as a qualitative representation of a household's ability to save. A large number of households are non-borrowers and ability to save may serve as an identifying variable for selection equation. The savings capacity represents a household's option to dis-save and thereby reduces the need for credit or for fewer loans or to the contrary capture the creditworthiness of a household to qualify for loan from a formal source of credit or act as a collateral for informal source of credit. Therefore, we could expect the savings index to be either positive or negative in the selection equation depending on which aspect is dominant or both the effects could nullify each other at the aggregate.

## Table 2: Descriptive Statistics: Mean Values of the Variables

| VARIABLES | TOTAL (N=42,152) | BORROWER (N=22,630) | NON BORROWER (N=19,522) | Equality of proportions | Standard Error |
|---|---|---|---|---|---|
| **CHARACTERISTICS OF HOUSEHOLD HEAD** | | | | | |
| *Highest level of education of the household* | | | | | |
| Not iterate | 0.62 | 0.65 | 0.68 | | |
| Primary | 0.12 | 0.11 | 0.12 | -0.07*** | 0.0048 |
| Middle | 0.16 | 0.16 | 0.17 | 0.005* | 0.0032 |
| Secondary | 0.06 | 0.05 | 0.08 | 0.018*** | 0.003 |
| Higher secondary | 0.02 | 0.01 | 0.02 | 0.02*** | 0.002 |
| Post higher secondary | 0.02 | 0.01 | 0.03 | 0.009*** | 0.001 |
| Female Headed Household (Female ==1) | 0.15 | 0.13 | 0.17 | 0.04*** | 0.003 |
| Age of the head of Household *(yr)* | 49.68 | 48.65 | 50.96 | 1.87*** | 0.13 |
| **MAIN INCOME SOURCE** | | | | | |
| Agriculture and allied | 0.26 | 0.29 | 0.21 | -0.10*** | 0.004 |
| Agricultural labour | 0.12 | 0.13 | 0.09 | -0.02*** | 0.003 |
| Non-Agricultural labour | 0.24 | 0.25 | 0.22 | -0.017*** | 0.004 |
| Artisans | 0.02 | 0.02 | 0.017 | 0.002** | 0.0012 |
| Business | 0.12 | 0.11 | 0.12 | 0.01*** | 0.003 |
| Salaried | 0.17 | 0.14 | 0.2 | 0.07*** | 0.003 |
| Others | 0.08 | 0.06 | 0.11 | 0.05*** | 0.002 |
| **ECONOMIC STATUS** | | | | | |
| ***Quintiles (based on asset index)**** | | | | | |
| Quintile 1 | 0.21 | 0.22 | 0.18 | -0.04*** | 0.003 |
| Quintile 2 | 0.20 | 0.22 | 0.18 | -0.03*** | 0.003 |
| Quintile 3 | 0.20 | 0.22 | 0.18 | -0.03*** | 0.003 |
| Quintile 4 | 0.20 | 0.19 | 0.21 | 0.02*** | 0.004 |
| Quintile 5 | 0.19 | 0.15 | 0.24 | 0.09*** | 0.004 |
| ***Below Poverty Line Households*** | 0.18 | 0.16 | 0.19 | 0.018*** | 0.003 |
| **SOCIAL STATUS** | | | | | |
| ***Caste/Religion Groups*** | | | | | |
| Brahmin | 0.05 | 0.04 | 0.05 | 0.01*** | 0.002 |
| Forward caste | 0.15 | 0.14 | 0.17 | 0.04*** | 0.003 |
| Other Backward Castes (OBC) | 0.36 | 0.39 | 0.31 | -0.08*** | 0.004 |
| Dalit | 0.22 | 0.24 | 0.2 | -0.02*** | 0.004 |
| Adivasi | 0.08 | 0.06 | 0.11 | 0.03*** | 0.002 |
| Muslim | 0.11 | 0.11 | 0.11 | 0.01*** | 0.003 |
| Christian, Sikh, Jain | 0.02 | 0.02 | 0.02 | 0.002* | 0.001 |

*(Contd… Table 2)*

*(Contd… Table 2)*

| | | | | | |
|---|---|---|---|---|---|
| Terciles of Horizontal Networking Index *(aggregated at PSU Level) | | | | | |
| Low | 0.44 | 0.41 | 0.45 | 0.03*** | 0.004 |
| Medium | 0.26 | 0.28 | 0.21 | -0.06*** | 0.004 |
| High | 0.31 | 0.29 | 0.32 | 0.03*** | 0.004 |
| Terciles of Confidence in Institutions Index* (aggregated at PSU Level) | | | | | |
| Low | 0.33 | 0.33 | 0.33 | 0.001 | 0.004 |
| Medium | 0.38 | 0.39 | 0.27 | -0.017*** | 0.004 |
| High | 0.28 | 0.27 | 0.28 | 0.01*** | 0.004 |
| Terciles of Vertical Networking Index*(Household level) | | | | | |
| Low | 0.33 | 0.32 | 0.34 | 0.04*** | 0.004 |
| Medium | 0.33 | 0.35 | 0.3 | -0.05*** | 0.004 |
| High | 0.33 | 0.31 | 0.35 | 0.04*** | 0.004 |
| **EXCLUSION RESTRICTIONS#/CONTROL VARIABLES** | | | | | |
| Household Characteristics: | | | | | |
| Dependency Ratio | 0.40 | 0.39 | 0.4 | 0.01*** | 0.004 |
| Household Size (no of persons) | 4.76 | 4.96 | 4.5 | -0.41*** | 0.02 |
| Economic: | | | | | |
| Savings Portfolio (Count) | 1.25 | 1.35 | 1.14 | -0.21*** | 0.01 |
| Social Event: | | | | | |
| Family Event (Count) | 1.1 | 1.26 | 0.91 | -0.35*** | 0.01 |
| Region: | | | | | |
| Rural | 0.7 | 0.75 | 0.64 | -0.11*** | 0.004 |
| Urban Nonmetro | 0.23 | 0.2 | 0.27 | 0.07*** | 0.004 |
| Metro | 0.06 | 0.04 | 0.08 | 0.03*** | 0.002 |
| **CATEGORIES OF BORROWERS/NON-BORROWERS** | | | | | |
| Only Formal | 0.27 | 0.2 | 0 | | |
| Only Informal | 0.04 | 0.48 | 0 | | |
| Only MFI | 0.01 | 0.07 | 0 | | |
| MFI-Formal | 0.04 | 0.02 | 0 | | |
| MFI-Informal | 0.06 | 0.06 | 0 | | |
| Formal-Informal | 0.02 | 0.11 | 0 | | |
| All three | 0.45 | 0.04 | 0 | | |
| Non borrowing | 0.45 | 0 | 1 | | |

**Note:** *All these indices are estimated using Principal Component Analysis by using several indicators capturing different aspects of these indices. These are discussed in detail in Appendix A. # these are variables used only in the selection equation and not in the outcome equation.

**Source:** Author's own estimation using household level data IHDS (Round 2).

Regional dummy variables for rural, urban metro and non-metro regions of urban areas are incorporated to account for broad structural differences in economic development.[16] In Table 2, it is also observed that around 17 percent of the borrowing households have MFI as one of the sources of borrowing.

**Probit (First Stage) Estimates: Decision to Borrow/ Not Borrow**

Table 3 presents the probit estimates for first stage selection model based on equation (1c) for the decision to borrow or not. This decision is explained by household level variables like dependency ratio (where dependents are both elderly and the children), savings level, and main income source of the household, level of wealth (estimated through the asset class), if the household is below poverty line etc. Some general characteristics of the head of the household like gender, age of the head also contribute to explain this decision.

**General Characteristics Of Household**

Highest level of education in the household is divided into five distinct categories, namely, primary school, middle school, secondary, higher secondary. The rest of the households are classified under not literate. As compared to not literate households, households belonging to all the other education classes are less likely to borrow. There is not a clear non-linear U shaped trend as was observed in Nguyen (2007) perhaps because we use highest education level of the household rather than the education of the household head.

Women heads are less likely to borrow, after other factors are accounted for as is observed from unadjusted values in Table 1. With around 15 percent of women headed households, they have a lower share (13 percent) among borrowers but higher (17 percent) among non-

[16] Addition of district dummies could account for the fixed effects across different states but when they were added, there was an issue with convergence of the model. Therefore, addition of the 'region' variable was done to take care of the fixed effects.

borrowers. This corroborates with Nguyen (2007) that male as the head of the household significantly determines the probability of borrowing. As household heads grow older, they are more likely to borrow and the negative sign associated with the squared terms indicates that this likelihood falls for the elderly. In Togba (2009) age of household head has a linear specification with a result that younger households are more likely to borrow relative to older household. Inclusion of the squared age reveals that there was a likely omitted variable bias and is perhaps more closer to reality in this cultural context where parents support children for a much longer time.[17] Hence, very young households may have limited need to borrow and middle age has higher needs with adequate working age to repay the loan which would diminish with age. This makes an inverted-U that corroborates with the finding of Ramadorai (2017) that finds a similar relationship between age and participation rate in the debt market.

All the households whose main source of income is not agriculture and allied activities are less likely to borrow. Uncertainty in output will make the income stream more prone to volatilities and seasonality and hence they may be in a larger need for credit to smoothen consumption. Further, large number of cultivator households own very small land area and with viability of agriculture declining, the investment for different activities have to be met with by borrowing frequently. These two features combined with their larger share in total households could be resulting in this finding.

**Economic Variables**

A household below the poverty line is less likely to borrow indicating its lower creditworthiness perceived both from the demand and supply side. As for wealth status, the bottom three asset classes are no different in their likelihood to borrow, while the top two asset classes are less likely

[17] This is in contrast to the conventional finding of the theory of optimal decision along the lifecycle which says that debt will have a larger prevalence in younger households.

to borrow compared to other three classes with a statistically significant and negative coefficient increasing in magnitude as well. A higher score on the savings index implies that the household has invested in more number of saving avenues (like chit funds, SHG, gold, bank savings etc) and is more likely to resort to borrowing as it might be using these savings as collateral to secure a greater amount of loan. Overall, the perception of limited creditworthiness perhaps both by the borrower and lender at the lower end while a large income and savings at the upper end of the socio-economic hierarchy reduces the probability of borrowing.

While OBC's, Dalit and minority religious communities such as Christian, Sikh and Jain's are more likely to be borrowers, Adivasis and Muslims are less likely to borrow for their credit needs.

A household of a bigger size will have larger consumption needs (including social expenditure) and hence there is a greater likelihood of borrowing after controlling for other factors and in particular the economic status. This finding is similar to Nguyen's But contrary to the finding of Swain (2007); the marginal effect of dependency ratio in our model is negative and significant, implying that after controlling for all the other factors, a household with a higher dependency ratio is less likely to borrow. This negative sign that fewer households with higher dependency ratio may be approaching a lender could be a result of either supply side perception that a household with a larger share of dependents may face fewer loan approval[18] and/or perceiving a bias regarding creditworthiness of themselves.

Index of major social events in the past 5 years is constructed[19] and this variable has a positive sign implying that accounting for everything else socio-cultural aspects that help maintain a social dignity create a demand for credit.

18 Here, loan approval encompasses both formal and informal sector. It is noticed in the dataset that this information on loan approval is available for all loan types.

[19] The construction of the variable is given in Appendix A

**Table 3: Estimates of Marginal Effects for Probit Model: Decision to Borrow from Any Source**

| VARIABLES | Coefficients | Std. Error |
|---|---|---|
| **CHARACTERISTICS OF HOUSEHOLD HEAD** | | |
| ***Education (Not literate)*** | | |
| Primary | -0.038*** | (0.010) |
| Middle | -0.02** | (0.0098) |
| Secondary | -0.054*** | (0.014) |
| High secondary | -0.076*** | (0.021) |
| Post Hi Secondary | -0.12*** | (0.024) |
| Gender | -0.04*** | (0.010) |
| Age | 0.010*** | (0.0016) |
| Age Square | -0.0001*** | (0.00001) |
| **MAIN INCOME SOURCE (BASE: AGRICULTURE and ALLIED)** | | |
| Agriculture Labour | 0.013 | (0.012) |
| Non Agriculture Labour | -0.011 | (0.010) |
| Artisans | -0.037 | (0.025) |
| Business | -0.038*** | (0.012) |
| Salaried | -0.094*** | (0.011) |
| Others | -0.11*** | (0.015) |
| **ECONOMIC STATUS** | | |
| Asset Class 2 | -0.028** | (0.011) |
| Asset Class 3 | -0.02* | (0.011) |
| Asset Class 4 | -0.079*** | (0.011) |
| Asset Class 5 | -0.13*** | (0.012) |
| ***Below Poverty Line Households*** | -0.07*** | (0.009) |
| **SOCIAL STATUS (Base: Forward caste)** | | |
| Brahmin | -0.016 | (0.016) |
| OBC | 0.043*** | (0.015) |
| Dalit | 0.028* | (0.016) |
| Adivasi | -0.19*** | (0.017) |
| Muslim | -0.041** | (0.017) |
| Christian, Sikh, Jain | 0.050** | (0.022) |
| Dependency Ratio | 0.01 | (0.014) |
| Household Size | 0.017*** | (0.001) |
| Savings Portfolio (Count) | 0.037*** | (0.002) |
| Family Events (Count) | 0.05*** | (0.003) |
| **REGION** | | |
| Urban Non metro | -0.023*** | (0.008) |
| Urban Metro | -0.11*** | (0.012) |
| Observations | 39943 | |

**Note:** Standard errors in parentheses *** p<0.01, ** p<0.05, * p<0.1, Pseudo R2 0.13, N = 39670, Wald Chi(52) = 4183.11

**Source:** Author's own estimation using household level data IHDS (Round 2)

### *Multinomial Probit (Second Stage) Estimates: Decision to Borrow from Different Types of Sources*

Table 4 presents the estimates of the multinomial probit model specified in equation (5a). As mentioned in the data section, there are seven mutually exclusive categories for loan sources arising from formal, informal and MFI sources and their combinations. It is observed that correction for the selection bias is justified with the inverse mills ratio (see equations 3d and 4a) across all categories being statistically significant. The coefficient of the mills ratio is positive in Only-Formal and Only-MFI, therefore we say that in these, a positive selection bias is noted. Estimating the model without selection parameter would imply an upward bias in the estimated coefficients for these two categories. The coefficient of mills ratio for MFI-Formal, MFI-Informal, Formal-Informal and All three sources is negative, therefore we can say that there is a negative selection bias in this case, implying that if the selection wasn't accounted for, then the coefficients associated with these categories would show a downward bias.

There are characteristics of the household, general, economic and social that affects the final choice of source of credit. The estimated coefficients reported in Table 3 are marginal effects of the multinomial probit model. If the estimated effect of an independent variable ($X_i$) is statistically significant and positive (negative) then it is interpreted as keeping all other variables at their mean value, a one unit increase in $X_i$ increases (decreases) the probability of $Y_i = j$ relative to $Y_i = k$ by $\beta_i$ units. The aim is to analyze how different sources of credit are being accessed by different social and economic groups. Here, the base category (k) is the informal source.

The variables in the analysis are either discriminatory (gender of household head, caste, religion or economic status) or empowering (education, horizontal networks etc.). Some variables like age, confidence

in institutions, vertical networks etc. are included as control variables in the discussion. All these are discussed under three broad headings.

**General Characteristics of Household**

According to our analysis, it is seen that women headed households, or with any other main income source except agriculture and allied services, or located in urban metro or urban non-metro are less likely to prefer formal sources and combinations of it over Only-informal sources. In Langer (2009), there was an insignificant effect of education in determining borrowing from microcredit institutions. In another study, gender, education and occupation had no significant effects on access to formal and informal credit while comparing these against each other (Diagne 1999).

We believe that since the categorization followed in this paper is more exhaustive, we are able to see how these factors are interacting with these choices of sources in the presence of MFI. Women heads relative to male heads are significantly more likely to borrow from a combination of MFI and informal source than only informal. This indicates a 'natural' affinity of women towards microfinance institutions as well as a possible need to borrow from more than one source of borrowing. Households with younger heads are more likely to borrow from categories that have MFI as one of the options and less likely to borrow from formal sources and their combinations. This is in contrast to the result obtained by Togba (2009), who finds through the three category multinomial logit model of loan source choice that the probability of a household participating in a bank credit program is lower in case of an older head.

Households with highest level of education as primary and middle school show a preference to borrow from only formal, MFI or informal sources. However, households with highest levels of education being post hi-secondary, show a clear preference to borrow from Only-formal

sources. A greater likelihood of different income sources when compared against agriculture and allied services towards Only-MFI and MFI-informal indicates that households having them as their main source of income may borrow from both a combination of MFI and informal and Only-MFI sources. At the same time, they are less likely to borrow from Only-formal institutions when compared against Only-informal institutions. This shows that where banks are a popular loan source for extending agricultural credit, other occupations find it easier to reach out to them, thereby, facilitating investment.

Through this, we can see that Only MFI and MFI-Informal and Only-Informal categories are catering to households with similar characteristics. The groups of characteristics that prefer Only-MFI and MFI-Informal categories are in most cases showing a lesser preference to the formal source choices, thereby indicating that MFI's are probably playing the role of serving the individuals that aren't being absorbed by the formal institutions and are being served by informal networks. Rural households, with male heads, older heads, and with agriculture and income as their main income source are more likely to access formal sources of credit.

**Economic Status**

Langer (2009) found that households accessing microcredit services are more likely to be poor. The households lying below poverty line are significantly less likely to borrow from MFI-informal and all three sources as compared to Only-informal sources. This shows a preference of BPL households to borrow from informal sources and also that none of the formal sources (either formal banks or MFI) are being able to penetrate into that level of poverty. This is also indicative of these households' lower likelihood of borrowing from multiple sources of credit.

Diagne (2009) concluded that the composition of household assets is a more important determinant of household access to formal

credit. Composite index of household assets was created to study its relationship with the sources of borrowing. Comparing different quintiles of asset index, to the lowest asset quintile, we see that the households are more likely to borrow from Only-formal, MFI-Formal, Formal-Informal and All three as compared to Only-Informal sources as we move from Quintile 2-5, showing the relevance of asset level in accessing loans from a formal sources of credit. Poor households often find themselves to own less or no assets and therefore show a lesser tendency to borrow from this source.

Comparing different quintiles of asset index to the lowest asset quintile, we see that the likelihood of borrowing from Only-MFI and MFI-informal follows a U-shaped pattern- lowest asset classes and highest asset classes are less likely to borrow from these categories. This indicates that the microfinance institutions are more likely to serve households belonging in the middle asset levels, therefore still being unable to capture the households in the lowest asset quintile. The poorest household in the lowest asset quintile is still more likely to borrow from Only-Informal sources of credit.

**Social Status**

The result relating to Muslim and Dalit group in our study stands out; when compared with forward castes are significantly less likely to borrow from Only-formal sources relative to Only-Informal. This can be a consequence of Mitra and Venkatachalam (2018) findings of significant inter-caste difference between application rates and evidence of discrimination against Scheduled Tribe borrowers at the approval stage. Through our categorization, we observe that these caste groups show a likelihood towards borrowing from MFI-formal and MFI-Informal when compared against Only-Informal sources. This shows that even though these groups find themselves restricted in terms of access to formal credit services, they show relaxed constraints when these formal credit services are combined with MFI. A possible reason for this may be that

successfully it is easier for them to acquire loans from microfinance institutions. It also shows a greater tendency of these caste groups to borrow from multiple sources of credit as their need is perhaps unfulfilled by one loan source.

Positive and increasing marginal effect associated with networking variables in MFI are seen when we move from medium to high level of horizontal networking, thereby indicating that if a household is from an area which has higher networking prospects then, it is more likely to avail loans from microfinance institutions. Having said this, there may be an interaction effect between caste and networking variables as a reason for popularity of microfinance institutions may have a link to closely-knit networks of caste groups.

Keeping everything else constant, a higher confidence in institutions shows increased likelihood of borrowing from formal sources of credit and their combinations. The marginal probability is higher as we move to areas where households share higher levels of confidence, as was found by Togba, (2009). Microfinance institutions (with informal sources) is filling the gap created due to inability of people to borrow from formal credit sources. People along with borrowing from only-informal sources, in some cases, show a preference for borrowing from both MFI-Informal and Only-MFI.

## Table 4: Results of Stage 2: Multinomial Probit (Marginal Effects)

| | (1) | (3) | (4) | (5) | (6) | (7) |
|---|---|---|---|---|---|---|
| **VARIABLES** | **Only Formal** | **Only MFI** | **MFI formal** | **MFI informal** | **Formal Informal** | **All three** |
| **CHARACTERISTICS OF HOUSEHOLD HEAD** | | | | | | |
| ***Highest level of education of (Base category: Not literate)*** | | | | | | |
| Primary | 0.057*** | 0.02*** | 0.01** | -0.018*** | 0.001 | -0.0004 |
| | (0.011) | (0.008) | (0.004) | (0.006) | (0.009) | (0.005) |
| Middle | 0.05*** | 0.006 | 0.006 | -0.01*** | -0.01 | -0.008** |
| | (0.009) | (0.007) | (0.004) | (0.005) | (0.007) | (0.004) |
| Secondary | 0.065*** | 0.02* | 0.0002 | -0.01 | -0.01* | -0.0003 |
| | (0.014) | (0.01) | (0.005) | (0.01) | (0.01) | (0.009) |
| High secondary | 0.037 | 0.0012 | -0.01*** | -0.003 | 0.003 | 0.02 |
| | (0.025) | (0.01) | (0.003) | (0.01) | (0.02) | (0.02) |
| Post Hi Secondary | 0.10*** | -0.005 | -0.0078 | -0.02 | 0.07* | 0.002 |
| | (0.037) | (0.02) | (0.007) | (0.01) | (0.03) | (0.01) |
| Gender | -0.058*** | -0.007 | -0.007** | 0.02*** | -0.01** | 0.002 |
| | (0.010) | (0.007) | (0.003) | (0.007) | (0.007) | (0.005) |
| Age | 0.0072*** | -0.0003 | 0.001** | -0.001 | -0.0004 | 0.0008 |
| | (0.0018) | (0.001) | (0.0008) | (0.001) | (0.001) | (0.0009) |
| Age Square | -0.000047*** | 0.0000002 | -0.000001* | 0.000008 | 0.00001 | -0.000008 |
| | (0.00001) | (0.000013) | (0.000008) | (0.00001) | (0.00001) | (0.000009) |
| **MAIN INCOME SOURCE (BASE: AGRICULTURE and ALLIED)** | | | | | | |
| Agriculture Labor | -0.12*** | 0.035*** | 0.002 | 0.08*** | -0.054*** | 0.0043 |
| | (0.014) | (0.008) | (0.005) | (0.009) | (0.009) | (0.0066) |
| Non Agriculture Labor | -0.13*** | 0.017*** | -0.0001 | 0.02*** | -0.069*** | -0.020*** |
| | (0.011) | (0.006) | (0.004) | (0.005) | (0.0084) | (0.0044) |
| Artisans | -0.11*** | 0.030 | -0.008 | 0.07*** | -0.038* | -0.019* |
| | (0.02) | (0.022) | (0.007) | (0.02) | (0.020) | (0.011) |
| Business | -0.09*** | 0.024*** | 0.002 | 0.0092 | -0.044*** | -0.017*** |
| | (0.012) | (0.0092) | (0.004) | (0.0065) | (0.010) | (0.005) |
| Salaried | -0.06*** | 0.013 | 0.004 | 0.014** | -0.020* | -0.006 |
| | (0.01) | (0.0089) | (0.004) | (0.006) | (0.010) | (0.006) |
| Others | -0.06*** | -0.00008 | 0.006 | 0.01 | -0.039** | -0.01 |
| | (0.01) | (0.0114) | (0.00735) | (0.01) | (0.015) | (0.007) |
| **ECONOMIC STATUS:** | | | | | | |
| ***Asset Index Quintiles (Base: Quintile 1)*** | | | | | | |
| Quintile 2 | 0.018 | -0.0046 | 0.0043 | -0.009 | 0.014* | 0.009*** |
| | (0.01) | (0.0079) | (0.0037) | (0.0058) | (0.007) | (0.003) |
| Quintile 3 | 0.046*** | -0.018** | 0.012*** | 0.017*** | 0.029*** | 0.036*** |
| | (0.011) | (0.0076) | (0.0037) | (0.0065) | (0.007) | (0.004) |

*(contd.. Table 4)*

| | | | | | | |
|---|---|---|---|---|---|---|
| Quintile 4 | 0.092*** | -0.018** | 0.016*** | 0.017** | 0.059*** | 0.032*** |
| | (0.012) | (0.0086) | (0.0043) | (0.0076) | (0.010) | (0.005) |
| Quintile 5 | 0.19*** | -0.030*** | 0.026*** | 0.0029 | 0.07*** | 0.04*** |
| | (0.015) | (0.0096) | (0.006) | (0.0085) | (0.011) | (0.008) |
| ***Is the household below poverty line?*** | -0.017 | -0.0019 | -0.0016 | -0.011** | -0.0044 | -0.01** |
| | (0.011) | (0.0066) | (0.0046) | (0.0057) | (0.0085) | (0.0049) |
| **SOCIAL STATUS** | | | | | | |
| ***Castes (Base: Forward castes)*** | | | | | | |
| Brahmin | 0.025 | 0.016 | 0.017*** | 0.027*** | 0.007 | 0.035*** |
| | (0.017) | (0.016) | (0.0053) | (0.0091) | (0.015) | (0.006) |
| OBC | -0.00014 | 0.010 | 0.014*** | 0.026*** | -0.0075 | 0.02*** |
| | (0.015) | (0.015) | (0.0046) | (0.007) | (0.014) | (0.004) |
| Dalit | -0.027 | 0.037** | 0.012** | 0.03*** | -0.020 | 0.019*** |
| | (0.016) | (0.015) | (0.005) | (0.008) | (0.015) | (0.0051) |
| Adivasi | -0.021 | 0.0057 | 0.021** | 0.067*** | 0.060*** | 0.041*** |
| | (0.022) | (0.017) | (0.009) | (0.014) | (0.02) | (0.009) |
| Muslim | -0.036** | 0.0074 | 0.013* | 0.023** | -0.002 | 0.028*** |
| | (0.017) | (0.015) | (0.006) | (0.009) | (0.016) | (0.007) |
| Christian, Sikh, Jain | 0.059** | 0.033 | 0.03*** | 0.018 | -0.021 | 0.018* |
| | (0.024) | (0.021) | (0.01) | (0.012) | (0.018) | (0.01) |
| ***Confidence in Institutions (PSU Share) (Base: Low)*** | | | | | | |
| Share of confidence (Medium) | 0.056*** | 0.039*** | 0.021*** | 0.01 | -0.0005 | 0.002 |
| | (0.018) | (0.013) | (0.007) | (0.011) | (0.01) | (0.008) |
| Share of confidence (High) | 0.076*** | 0.048*** | 0.032*** | 0.015 | 0.011 | 0.011 |
| | (0.018) | (0.011) | (0.006) | (0.01) | (0.015) | (0.0078) |
| ***Horizontal Network (PSU Share) (Base: Low)*** | | | | | | |
| Share of Horrizontal Networking (Medium) | -0.061*** | -0.050*** | 0.01** | 0.006 | 0.060*** | 0.04*** |
| | (0.017) | (0.012) | (0.007) | (0.009) | (0.012) | (0.006) |
| Share of Horizontal Networking (High) | 0.019 | 0.036*** | 0.019*** | 0.04*** | 0.031** | 0.05*** |
| | (0.014) | (0.009) | (0.005) | (0.008) | (0.013) | (0.006) |
| ***Vertical Network (PSU Share) (Base: Low)*** | | | | | | |
| Share of Vertical Networking (Medium) | 0.0005 | -0.027*** | -0.0006 | 0.005 | 0.01** | 0.002 |
| | (0.009) | (0.005) | (0.003) | (0.004) | (0.007) | (0.004) |
| Share of Vertical Networking (High) | 0.051*** | -0.02*** | 0.005 | -0.008 | 0.02*** | -0.001 |

*(contd.. Table 4)*

| | | | | | | |
|---|---|---|---|---|---|---|
| | (0.009) | (0.006) | (0.003) | (0.005) | (0.007) | (0.004) |
| **REGION (BASE: RURAL)** | | | | | | |
| Urban Non Metro | -0.024*** | 0.008 | 0.0015 | 0.042*** | 0.00076 | 0.002 |
| | (0.008) | (0.006) | (0.003) | (0.005) | (0.0075) | (0.004) |
| Urban Metro | -0.06*** | -0.01* | -0.023*** | 0.049*** | -0.017 | -0.003 |
| | (0.016) | (0.011) | (0.008) | (0.0096) | (0.015) | (0.009) |
| Inverse mills Ratio from Probit | 0.157* | 0.16*** | -0.065** | -0.28*** | -0.62*** | -0.28*** |
| | (0.080) | (0.05) | (0.03) | (0.04) | (0.067) | (0.037) |
| Observations | 21,753 | 21,753 | 21,753 | 21,753 | 21,753 | 21,753 |

**Note:** Standard errors in parentheses; *** p<0.01, ** p<0.05, * p<0.1; Number of observations =22496; Wald chi2(174)= 4934.21; Log pseudolikelihood=-1.78E+08

**Source:** Author's own estimation using household level data IHDS (Round 2)

## CONCLUSION

Households borrow from one or more of the credit source options available in the market due to existing supply and demand side restrictions. The choice of different loan source combinations is often dependent on the varying socio-economic-demographic characteristics of the households. Here, households with similar characteristics often find themselves self-selecting into these similar combinations of loan sources. To ensure a greater amount of financial inclusion, it becomes necessary to address what determines the choice between these alternative combinations.

Using the Indian Human Development Survey (2011-12), we create an exhaustive categorization of different combination of options of loan sources available to a household. Given this detailed categorization of options, we get more clarity on the mediating role played by microfinance institutions through its linkages with informal and formal sources as compared to earlier studies, where comparison are made only between formal and informal sources.

After accounting for the fact that prior to source choice there is a choice made by the household on whether to borrow or not, we observe

that, where household size and the happening of any major family event may affect the decision to borrow, variables like confidence in institutions, level of networking, asset level etc. may be more likely to affect the decision to borrow from type of credit sources. This paper then, for the first time, uses IHDS 2011-12 dataset to model these two levels of decision making through a Multinomial Probit with Heckman selection.

Through our model, we find that a household is more likely to borrow from one or more credit sources if among other things, it has a greater household size and has had a larger number of family events in the past 5 years. A women headed household, or a household with an older head is more likely to be a non-borrower.

If it borrows then, where women headed households are significantly less likely to borrow from Only-formal sources, a household with an older head is significantly more likely to borrow from them. Moreover, household with women heads or younger heads prefer Only-MFI and MFI-informal sources. In addition, among borrowers, an increase in highest level of education in the household makes it more likely to borrow from Only-Formal sources as against Only-Informal sources. It is interesting to note that households with advantageous characteristics prefer to borrow from Only-Formal, MFI-Formal and Formal-Informal. The ones that did not prefer these sources to Only-Informal sources preferred to borrow from Only-MFI and MFI-Informal. This implies that where their access to formal sources and their combinations is constrained, microfinance institutions and informal sources are a preferred alternative for credit. This shows the mediating role played by MFIs in combination with formal and informal sources.

Controlling for everything else, we observe that by allowing the choice of MFI and its combination, the less economically disadvantaged rather than the least disadvantaged households are able to access credit.

This inability of microfinance institutions, specifically instituted to target the least economically disadvantaged, results in the poorest of the poor still being out of the system. They continue to either not borrow or borrow from the informal lending networks.

Among social status variables, the result relating to Dalit households and horizontal networking variables stand out. Dalit households relative to forward castes showed a preference for categories with MFI as a loan source option. In addition, households belonging in regions with better networking opportunities showed a greater likelihood towards borrowing from combinations that had MFI as one of its categories. This may be an indication of a possible interaction effect between caste and networking variables where closely-knit networks of certain communities are able to avail microfinance loans.

The results are indicative of MFI being an emerging alternative for these households who are unable to access formal credit markets. Either MFI is the only source of borrowing for households or it appears in combination with Formal or Informal sources. Here, households borrowing from MFI-Formal sources closely resemble those borrowing from Only-formal sources and the households borrowing from MFI-Informal sources resembles those borrowing from Only-Informal sources.

Households may not shift from Only-Informal sources of credit to Only-formal sources but they are more likely to switch to a combination with microfinance institutions. This paper concludes that the socially and economically disadvantaged sections, often constrained in their borrowing from formal sources, show a preference of borrowing from MFI-Informal sources instead. Presence of MFI as an option along with formal and informal credit sources, in a way, is increasing the inclusiveness in the financial market by ensuring a more equitable distribution of credit.

The analysis here overlooks the loan application and approval decisions. Modelling and incorporating them in our research will make the analysis more exhaustive. Inclusion of these will enable us to understand the role played by demand side and supply side factors in the final choice. One of the ways to do the same is to restructure the model and use another specification of the dependent variable. However, the available options of models require alternative specific variables for each stage of modeling but the dataset does not provide such information.

It is important to mention that our study discusses access to loans and the determinants of the choice of loan source, but it does not analyze the impact and implication of taking credit from these different sources. IHDS-II makes the information on interest rate, purpose of loan and loan amount available for only the largest loan source. Using this set of question would not have allowed us to derive meaningful results while incorporating the multiple loan source categorization. Extension to this work may be through generation of a dataset, specifically designed to understand how different sources of loans affects a household. Conducting a micro-survey as an extension to this research is a possible way to overcome this limitation.

## APPENDIX A

### Table A.1: Construction of independent variable

| Variable | Construction/Description |
|---|---|
| **GENERAL CHARACTERISTICS** | |
| Dependency Ratio | Dependency ratio is constructed as the proportion of children (<15 years) and senior citizen (> 60 years) of the total number of members of the household. |
| Household Size | Household size is the total number of individuals in the household. |
| Highest level of education in household | Highest level of education in household is highest level of education attained by a member of the household. It is constructed as a dummy variable with five categories as (1) Not literate (2) Primary (3) Middle (4) Secondary (5) Hi-secondary (5) Post Hi-secondary<br>*Base category: Not literate* |
| Gender | Gender is a dummy variable with 1 indicative of the household being a woman headed household and 0 being a male head household.<br>*Base category: Male Head* |
| Age | Age is the age of the head of the household and is taken as a continuous variable |
| Main Income Source | Main Income Source are categorized as (1) Agriculture and Allied (2) Agriculture Labor (3) Non Agricultural Labor (4) Artisans (5) Business (6) Salaried (7) Others.<br>*Base Category: Agriculture and Allied Services* |
| Region | The area that the household is located in also forms for an important reference point as to which category of loan source will it be able to acquire. This variable has three categories (i) Rural areas (ii) Urban Non Metro (iii) Urban Metro.<br>*Base Category- Rural areas.* |
| **MAJOR EVENTS:** | |
| Family Event | Social event is a constructed as the count of number of events from zero (no major event) to seven (all major events took place listed) based on the |

| | |
|---|---|
| | question on whether any major social event took place in the household in the past 5 years or not like marriage, death, drought etc which would have needed additional financial resources. |
| **ECONOMIC INDICATORS:** | |
| Saving Index (values 0-9) | Savings Index is sum of the different avenues for savings that the household has invested in as collected in the survey. These include Buying property, Expanding property Mutual funds, fixed deposits, bank savings, Chit funds, Post office accounts, pension funds, and gold and jewelry. There was a separate section of questions on investment and binary response of 'yes' and 'no' was added to arrive at the final level of savings. |
| Asset Classes | Asset index uses 23 dummy variables answering questions about whether the household owns assets like cycle/bicycle, sewing machine, generator, car, air conditioner from round-I and merged with the data.<br>Two additional variables are also created that relates to the condition of the household (pukka or kachcha roof walls). In terms of seeking credit, household asset is an important criterion that the formal institutions look into. This index, created using PCA, divides into 5 quintiles where each quintile is a dummy variable.<br>*Base category- lowest asset quintile* |
| Is the household below poverty line? | If the per capita expenditure of the household is below the Tendulkar committee poverty line. If yes, then the household is poor. *Base category- Non-poor household.* |
| **SOCIAL INDICATORS:** | |
| Caste/Religion categories | It uses the eight-category classification of caste and religion, that includes, (i) Brahmin (ii) Forward Castes (iii) OBC (iv) Dalit (v) Adivasi (vi) Muslim (vii) |

| | |
|---|---|
| | Christian (v) Others.<br>*Base Category- Forward Castes* |
| Horizontal Networking Index | Questions on memberships in Mahila Mandal, youth/sports/reading groups/union/business groups, SHG, Credit/savings groups, religious/social groups, dev prog/NGO, and agricultural milk cooperative are taken and an index is created using PCA. Three different groups for this index are namely, Lowest 33 percent, Middle 33 percent, highest 33 percent. For each class, we generate a share at PSU level. This share represents the likelihood of the household belonging to a particular PSU, to lie in that particular class of Horizontal Networks. This merges with the household dataset to get the classes for each household.<br>*Base category- Lowest share tercile.* |
| Vertical Networking Index | Questions on if the household has any acquaintance in village panchayat/nagarpallika, ward committee/medical field/government field and if it has attended a public meeting called by the village panchayat, nagarpallika and ward committee are taken and an index is create using PCA. The index is then divided into terciles and the resulting three categories are representative of low, medium and high shares.*Base category: Lowest 33 percent share* |
| Confidence in Institutions | To construct this variable, we take 19 questions that relate to the faith that households have in different institutions such as banks, government offices, local crime level of the region etc. An index is created using PCA. (Same methodology as Horizontal networking index above).<br>*Base category: Lowest Share tercile* |

# REFERENCES


Armendáriz, B., and J. Morduch, (2010), "The Economics of Microfinance", *MIT press.*

Banerjee, A. V., and E.Duflo (2007), "The Economic Lives of the Poor", *Journal of economic perspectives*, 21(1), 141-168.

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

Bhattacharjee, Manojit and Meenakshi Rajeev (2011), "Repayment of Short Term Loans in The Formal Credit Market: The Role Of Accessibility To Credit From Informal Sources", *ISEC Working Paper 273.*

Beck, T. (2015), "Microfinance–A Critical Literature Survey", *World Bank Independent Evaluation Group Working Paper*, 4.

Cameron, A. C. and P. K. Trivedi, (2010), "Microeconometrics Using Stata (Vol. 2)", *College Station, TX: Stata press*.

Campero, A., and Kaiser, K. (2013). "Access to Credit: Awareness and Use of Formal and Informal Credit Institutions" (No. 2013-07). *Working Papers*, 2013.

Certo, S. T., J. R. Busenbark, H. S. Woo and M. Semadeni (2016), "Sample Selection Bias and Heckman Models in Strategic Management Research", *Strategic Management Journal*, 37(13), 2639-2657.

Chakraborty, T., and Gupta, A. (2017). "Loan Repayment Behaviour of Farmers: Analysing Indian Households". *IIT Kanupur Working paper.*

Conning, J., and C. Udry (2005), "Rural Financial Markets in Developing Countries, Economic Growth Centre", *Yale University, and Centre Discussion Paper,* (914).

Cragg, J. G. (1971), "Some Statistical Models for Limited Dependent Variables with Application to the Demand for Durable Goods", *Econometrica* (pre-1986), 39(5), 829.

Desai, S., A. Dubey, B. L. Joshi, M. Sen, A. Shariff and R. Vanneman, (2009), "India Human Development Survey: Design and Data Quality", *IHDS technical paper,* 1.

Diagne, A. (1999), "Determinants of Household Access to and Participation in Formal and Informal Credit Markets in Malawi", *(No. 583-2016-39679, p. 1).*

Duy, V. Q., M. D'Haese, J. Lemba, and L. D'Haese (2012), "Determinants of Household Access to Formal Credit in the Rural Areas of the Mekong Delta, Vietnam." *African and Asian studies*, 11(3), 261-287.

Fufa Fikadu Goshu (2016), "Determinants of Access to Credit and Credit Source Choice by Micro, Small and Medium Enterprises in Nekemte, Ethiopia", *International Journal of African and Asian Studies,* Vol.28. Accessed 20-10-2018.

Greene, W. (2009), "Discrete Choice Modeling" *In Palgrave Handbook of econometrics,* (pp. 473-556). *Palgrave Macmillan*, London.

Heckman, J. J. (1977). "Sample Selection Bias as a Specification Error (with an application to the estimation of labor supply functions)"

Hermes, N. and R. Lensink (2007), "The Empirics of Microfinance: What Do We Know?", *The Economic Journal,* 117(517), F1-F10.

Ingves, S. (2005), "Microfinance: A View from the Fund", *International Monetary Fund (IMF).*

Long, J. S., and J. Freese (2005), "Regression Models for Categorical Outcomes Using Stata",

Khandker, S. R. (1998), "Fighting Poverty with Microcredit: Experience in Bangladesh", *Oxford University Press.*

Langer, A. (2009), "Socioeconomic Status and Social Capital Levels of Microcredit Program Participants in India", *Retrieved from: https://conservancy.umn.edu/bitstream/handle/11299/51478/Langer, percent20Anna.pdf?sequence=1*

Morduch, J. (1999), "The Microfinance Promise", *Journal of Economic Literature*, *37*(4), 1569-1614.

Mitra Kumar, S. and R. Venkatachalam (2018), "Caste and Credit: A Woeful Tale?", *The Journal of Development Studies*, 1-18.

Mosley, P., and B. Helms, (1996), "Financial Sustainability, Targeting the Poorest, and Income Impact: Are there Trade-Offs For Micro-finance Institutions?", *Focus*; No. 5.

Nguyen, C. H. (2007), "Determinants of Credit Participation and its Impact on household consumption: Evidence from rural Vietnam", *Retrieved from: https://ideas.repec.org/p/hwe/certdp/0703.html*

Pal, S. (2002), "Household Sectoral Choice and Effective Demand for Rural Credit in India", *Applied Economics*, 34(14), 1743-1755.

Peachey, S. and A. Roe (2004), "Access to Finance. A Study for the World Savings Banks Institute", *Retrieved from: http://www.arabic.microfinancegateway.org/sites/default/files/mfg-en-paper-access-to-finance-2004.pdf*

Pitt, M. M. and S. R. Khandker (1998), "The Impact of Group-Based Credit Programs on Poor Households in Bangladesh: Does the gender of participants matter?", *Journal of Political Economy,* 106(5), 958-996 .

Pradhan, N. (2013), "Persistence of Informal Credit in Rural India: Evidence from 'All-India Debt and Investment Survey and beyond".

Ray, D. (2007), "Introduction to Development Theory", *Journal of Economic Theory,* 137(1), 1-10.

Ramodarai T (2017), "Indian Household Finance Report of the Household Finance Committee", *RBI.*

Sarap, K., and M. Venkatanarayana (2016), "Some Aspects of Dynamics in Rural Labour Market: Interface between Credit and Labour Markets", *Retrieved from: https://pdfs.semanticscholar.org/28a7/1bc1ea8acc2d1108c92ac4ce16a44193fb19.pdf*

Sen, A. (1988), "The Concept of Development", *Handbook of Development Economics*, 1, 9-26.

Singh, P. (2010), "Understanding the Structure of Micro Finance Institutions in India and Suggesting a Regulatory Framework", *Finance and Accounting department, IIM Lucknow*.

Swain, R. B. (2007), "The Demand and Supply of Credit for Households", *Applied Economics*, 39(21), 2681-2692.

Swain, R. B. and M. Floro (2012), "Assessing the Effect of Microfinance on Vulnerability and Poverty among Low Income Households", *Journal of Development Studies,* 48(5), 605-618.

Tang, S. and S. Guo, (2017), "Formal and Informal Credit Markets and Rural Credit Demand in China", *Industrial Economics System and Industrial Security Engineering* (IEIS'2017), *2017 4th International Conference on* (pp. 1-7). IEEE.

Togba, E. L. (2009), "Microfinance, Social Capital and Households Access to Credit: Evidence from Côte d'Ivoire, Georgia Institute of Technology", *Retrieved from: https://smartech.gatech.edu/bitstream/handle/1853/36658/1252001241_LT_1.pdf?sequence=1andisAllowed=y*

Unnikrishnan, V. (2016), "How Well Targeted are Social Assistance Programs in India-A Case Study of Indira Gandhi Old Age National Pension Scheme", *Retrieved from: https://www.isid.ac.in/~epu/acegd2016/papers/VidhyaUnnikrishnan.pdf*

Vanneman, R., S. Desai, and J. Noon (2018), "Social Capital in India: Networks, Organizations and Confidence", *American Sociological Association.*

Van de Ven, W. P. and B.M. Van Praag (1981), "The Demand for Deductibles in Private Health Insurance: A Probit Model with Sample Selection", *Journal of Econometrics*, 17(2), 229-252.

# *MSE Monographs*

* Monograph 34/2015
  Farm Production Diversity, Household Dietary Diversity and Women's BMI: A Study of Rural Indian Farm Households
  *Brinda Viswanathan*

* Monograph 35/2016
  Valuation of Coastal and Marine Ecosystem Services in India: Macro Assessment
  *K. S. Kavi Kumar, Lavanya Ravikanth Anneboina, Ramachandra Bhatta, P. Naren, Megha Nath, Abhijit Sharan, Pranab Mukhopadhyay, Santadas Ghosh, Vanessa da Costa and Sulochana Pednekar*

* Monograph 36/2017
  Underlying Drivers of India's Potential Growth
  *C.Rangarajan and D.K. Srivastava*

* Monograph 37/2018
  India: The Need for Good Macro Policies *(4th Dr. Raja J. Chelliah Memorial Lecture)*
  *Ashok K. Lahiri*

* Monograph 38/2018
  Finances of Tamil Nadu Government
  *K R Shanmugam*

* Monograph 39/2018
  Growth Dynamics of Tamil Nadu Economy
  *K R Shanmugam*

* Monograph 40/2018
  Goods and Services Tax: Revenue Implications and RNR for Tamil Nadu
  *D.K. Srivastava, K.R. Shanmugam*

* Monograph 41/2018
  Medium Term Macro Econometric Model of the Indian Economy
  *D.K. Srivastava, K.R. Shanmugam*

* Monograph 42/2018
  A Macro-Econometric Model of the Indian Economy Based on Quarterly Data
  *D.K. Srivastava*

* Monograph 43/2019
  The Evolving GST
  *Indira Rajaraman*

## *MSE Working Papers*
## Recent Issues

* Working Paper 172/2018
  Sustainability and Efficiency of Microfinance Institutions in South Asia
  Brijesh C. Purohit, S. Saravanan
* Working Paper 173/2018
  Corporate Governance Practices in India
  Ekta Selarka
* Working Paper 174/2018
  Current Status of Mangroves in India: Benefits, Rising Threats Policy and Suggestions for the Way Forward
  Samyuktha Ashokkumar, Zareena Begum Irfan
* Working Paper 175/2018
  Current Status of Coral Reefs in India: Importance, Rising Threats and Policies for its Conservation and Management
  Srihitha Baswapoor, Zareena Begum Irfan
* Working Paper 176/2018
  Family Structure, Education and Women's Employment in Rural India
  Sowmya Dhanaraj , Vidya Mahambare
* Working Paper 177/2018
  Household Income Dynamics and Investment in Children: Evidence from India
  Sowmya Dhanaraj, Smit Gade, Christy Mariya Paul
* Working Paper 178/2019
  Information and Communication Technology Diffusion and Financial Inclusion: An Interstate Analysis for India
  Amrita Chatterjee, Simontini Das
* Working Paper 179/2019
  Social Performance Of Microfinance Institutions in South Asian Region: A Case Study of India, Bangladesh, Nepal and Sri Lanka
  S. Saravanan, Brijesh C. Purohit
* Working Paper 180/2019
  Farmer's Perception on Soil Erosion in Rainfed Watershed Areas of Telangana, India
  Dayakar Peddi and Kavi Kumar

---

* Working papers are downloadable from MSE website http://www.mse.ac.in
$ Restricted circulation