ÉRUDITE
LABORATOIRE D'ÉCONOMIE
PARIS-EST

Series of ERUDITE Ph.D Dissertations

N° 01-2020

Title

Financial Inclusion, Entrepreneurship and Gender
An Empirical Assessment using Microeconomic Data

Author

Fozan FAREED

UNIVERSITÉ PARIS-EST

ÉCOLE DOCTORALE ORGANISATIONS, MARCHÉS, INSTITUTIONS

LABORATOIRE ÉRUDITE
2019-2020

Thèse pour le doctorat en Sciences Économiques

Présentée et soutenue publiquement par

FOZAN FAREED

20th May, 2020

Financial Inclusion, Entrepreneurship and Gender

An Empirical Assessment using Microeconomic Data

Under the supervision of Professor Julie LOCHARD and Professor Catherine BROS
Érudite, Université Paris-Est

Members of the Jury

Pr. William PARIENTÉ	Professor at UC Louvain	Referee
Pr. Rohini SOMANATHAN	Professor at Delhi School of Economics	Referee
Dr. Amina LAHRECHE	Deputy Division Chief, IMF	Examiner
Pr. Patrick LENAIN	Assistant Director, OECD	Examiner
Pr. Melika Ben SALEM	Professor at University Gustave Eiffel	President of the Jury

Abstract

Financial inclusion as part of the development process has gained considerable attention from policymakers worldwide. The numbers remain quite stark as 1.7 billion people worldwide, as of 2017, remain without access to basic financial services. This PhD thesis aims to empirically disentangle some of the many interrelationships between financial inclusion, entrepreneurship and gender. It consists of four chapters and relies on the use of nationally representative survey data at the individual level. The first chapter studies the effect of geographical access to microfinance on entrepreneurship and examines if having this access enables individuals to move up the economic ladder in Pakistan. The second chapter examines if financial inclusion promotes women's autonomy by generating women entrepreneurship. This chapter also constructs a comprehensive financial inclusion index to measure the state of financial inclusion in Mexico after taking into account access as well as the usage of different financial products. The third chapter explores the main drivers of financial exclusion in Pakistan after taking into consideration the need for credit and voluntary financial exclusion. Finally, the fourth chapter proposes a novel methodological approach for measuring household financial vulnerability by relying on unsupervised machine learning algorithms in the case of U.S. All these chapters use pseudo-panel datasets and rely on several methodologies to tackle endogeneity issues and concerns pertaining to selection bias. The results indicate that financial inclusion, entrepreneurship and gender are intimately related with each other. Microfinance seems to play an effective role in promoting entrepreneurship and enabling individuals to move up the economic ladder, whereas financial inclusion also seems to encourage women's autonomy by fostering entrepreneurship. The empirical results also uncover the main drivers of involuntary financial exclusion: financial illiteracy, poverty, and gender. Moreover, this thesis considers new methodological approaches to analyze household financial vulnerability and involuntary financial exclusion as an alternative to the standard line of research on these topics.

Key words: Financial inclusion, entrepreneurship, gender, microfinance, credit, savings, employment, labor market, economic ladder, poverty, economic development, panel data, instrument variables, endogeneity, selection bias, household financial vulnerability, unsupervised machine learning.

Abstract (In French)

Dans le cadre du processus de développement, l'inclusion financière a attiré une attention considérable de la part des décideurs du monde entier. Les chiffres restent assez frappants avec 1,7 milliard de personnes dans le monde n'ayant toujours pas accès aux services financiers de base. Cette thèse vise à démêler empiriquement certaines des nombreuses interrelations entre l'inclusion financière, l'entrepreneuriat et le genre. Elle se compose de quatre chapitres et repose sur l'utilisation de données longitudinales au niveau individuel. Le premier chapitre étudie l'effet de l'accès géographique à la microfinance sur l'entrepreneuriat et vérifie si cet accès permet aux individus de gravir les échelons économiques au Pakistan. Le deuxième chapitre examine si l'inclusion financière favorise l'autonomie des femmes en promouvant l'entrepreneuriat féminin. Ce chapitre construit également un indice complet d'inclusion financière pour mesurer l'état de cette dernière au Mexique après avoir pris en compte l'accès ainsi que l'utilisation de différents produits financiers. Le troisième chapitre explore les principaux facteurs de l'exclusion financière au Pakistan après avoir pris en considération le besoin de financement et l'exclusion financière volontaire. Enfin, le quatrième chapitre propose une nouvelle approche méthodologique pour mesurer la vulnérabilité financière des ménages en s'appuyant sur des algorithmes d'apprentissage automatique non supervisés aux États-Unis. Tous ces chapitres utilisent des données d'enquêtes représentatives au niveau national et s'appuient sur plusieurs méthodologies pour résoudre les problèmes d'endogénéité et de biais de sélection. Les résultats indiquent que l'inclusion financière, l’entrepreneuriat et le genre sont étroitement liés les uns aux autres. La microfinance semble jouer un rôle efficace dans la promotion de l’entrepreneuriat et dans l’ascension des échelons économiques des individus. L’inclusion financière semble, elle aussi, encourager l'autonomie des femmes en favorisant l'entrepreneuriat féminin. Les résultats empiriques révèlent également les principaux facteurs de l'exclusion financière involontaire: l'analphabétisme financier, la pauvreté et le genre. De plus, afin d’analyser la vulnérabilité financière des ménages ainsi que l’exclusion financière involontaire, cette thèse utilise de nouvelles approches méthodologiques pouvant être considérées comme une alternative aux standards de recherche sur ces thématiques.

Acknowledgements

This research would not have been possible without the support of many people. First of all, I would like to extend my gratitude to both my supervisors, Prof. Julie LOCHARD and Prof. Catherine BROS, for their constant support and guidance on research, writing, teaching, career steps and much more. I am grateful for their confidence in me since the master's program and I am extremely appreciative of their persistent encouragement throughout the three and a half years of my PhD.

I am also very honored to have a jury of esteemed professors and researchers in my PhD committee. I am grateful to Melika BEN SALEM, Amina LAHRECHE, Patrick LENAIN, William PARIENTÉ, and Rohini SOMANATHAN for taking out the time from their busy schedules and for providing me with valuable feedback on my research work.

I would like to thank the Higher Education Commission of Pakistan (HEC) and Campus France for providing me with a scholarship and for helping me out with all the logistics throughout the course of my PhD in France. I would also like to thank the Erudite lab and the O.M.I doctoral school for providing a supporting environment for research and for sponsoring my participation in international research conferences. I also thank Thanh-Ha for her prompt administrative help in my research as well as teaching related activities.

During my PhD, I got the chance to work at the International Monetary Fund (IMF) in Washington DC and the Organization for Economic Cooperation and Development (OECD) in Paris. I am very grateful for these opportunities to carry out research at these international organizations and work alongside some amazing people who have incredible skills and great work ethic. I would like to extend my appreciation to all my colleagues, especially Julien Reynaud, Patrick Lenain, Anastasia Guscina, Christoph Duenwald, and Koshy Mathai, for all the helpful discussions and guidance.

Most of my PhD research was conducted at the Erudite lab, and it goes without saying that many of its members contributed to the progress of my work. I would like to thank Eric Defebvre for helping me out with econometrics related questions. In addition, I would like to thank Neha, Redha, Juliette, Justine, Arnold, Patrick, Sylvain, Christoph, and Sandrine for all the discussions (serious and otherwise) and for all the fun times (doing sports together and trying out the best pizza places in Paris). I also wish to thank Juliette, Sumbul, Neha and Elisabeth for reading parts of my research and providing useful feedback, and last but not least, a special thanks to my dearest and close friends Simon and Jelena for their amazing friendship and for giving me some of the best memories in Paris.

I also wish to thank my family, especially my parents, who always believed in me and gave me their unconditional support. Baba and mama, this PhD would not have been possible without your encouragement and love.

Table of Contents

List of Figures

List of Figures

List of Tables

List of Acronyms

AFI	Alliance for Financial Inclusion
AIC	Akaike Information Criterion
ATM	Automated Teller Machine
CGAP	Consultative Group to Assist the Poor
CIs	Confidence Intervals
CNBV	Comisión Nacional Bancaria y de Valores (English: National Banking and Securities Commission)
CONAIF	Consejo Nacional de Inclusión Financiera (English: National Council for Financial Inclusion)
ENIF	Encuesta Nacional de Inclusión Financiera (English: National Financial Inclusion Survey)
ENOE	Encuesta Nacional de Ocupación y Empleo (English: National Survey of Occupation and Employment)
FII	Financial Inclusion Index
FII Survey	Financial Inclusion Insight Survey
FinTech	Financial Technology
GDP	Gross Domestic Product
GIS	Geographical Information System
GNI	Gross National Income
GPFI	Global Partnership for Financial Inclusion
HAC	Hierarchical Ascending Clustering
HDI	Human Development Index
HEC	Higher Education Commission
IMF	International Monetary Fund
INEGI	Instituto Nacional de Estadística y Geografía (English: National Institute of Statistics and Geography)
KPK	Khyber Pakhtunkhwa (a provinve in Pakistan)
KYC	Know Your Customer
LPM	Linear Probability Model
MDGs	Millennium Development Goals

List of Acronyms

MFIs	Microfinance Institutions
MIX	Microfinance Information Exchange
ML	Machine Learning
MLE	Maximum Likelihood Estimation
MPI	Multidimensional Poverty Index
NFIS	National Financial Inclusion Strategy
NGO	Non-governmental organization
OECD	Organization for Economic Cooperation and Development
OMI	Organisations, Marchés, Institutions (English: Organizations, Markets, Institutions)
PBS	Pakistan Bureau of Statistics
PKR	Pakistani Rupees
PMN	Pakistan Microfinance Network
POS	Point of Sale
PPI	Poverty Probability Index
RCT	Randomized Control Trial
ROSCAs	Rotating Savings and Credit Associations
SBP	State Bank of Pakistan
SCF	Survey of Consumer Finance
SDGs	Sustainable Development Goals
SECP	Securities and Exchange Commission of Pakistan
SME	Small and Medium Enterprise
UNDP	United Nations Development Programme

General Introduction

I. Background

Finance and development has emerged as a distinct field of research over the last three decades. Until the end of 1980s, textbooks on economic development did not generally discuss the role of access to finance and were not concerned about the links between finance, poverty, gender, and development (Levine, 2005). Beck and Levine (2018) point out that even though Schumpeter (1911), Goldsmith (1969) and Shaw (1973) argued about the importance of well-developed financial systems for economic growth, their views were generally peripheral to the separate studies of financial economics and economic development until the 1990s. The literature highlighting the link between finance and development took off in the 1990s when researchers started integrating financial frictions into endogenous growth models and studied the role of financial systems in economic growth (Romer, 1990; King and Levine, 1993).

Empirical evidence on the nexus between finance and economic development also began to appear in the 1990s. Research based on cross-country analysis, industry level analysis, and historical case studies has demonstrated that a well-functioning financial system can ease credit constraints that impede growth at the industry level. Furthermore, studies show that it can improve the allocation of capital, decrease transaction costs, and accelerate technological innovation (Galor and Zeira, 1993; Rajan and Zingales, 1998; Demirguc-Kunt and Maksimovic, 1998; Beck et al., 2008). There is an abundant amount of literature that argues that the relationship between financial development and economic growth is causal (Beck and Levine, 2018).

In recent times, another strand of literature that focuses on access to affordable formal financial services by unbanked households – also known as financial inclusion – has gained a lot of traction from policy makers and researchers worldwide. It is well understood that households from all income groups require tools to cope with risks, however, low-income households in developing countries are disproportionately affected by economic shocks. These shocks can include health emergencies, loss of employment, theft, natural disasters, adverse shifts in the labor market or other unforeseen emergencies. When these kinds of shocks hit, low-income households are not only more likely to be affected by them, but they are also the least prepared to cope with them. Imperfect credit markets, low rates of savings, inadequate insurance mechanisms, and inefficient payment facilities in developing economies do not generally allow low-income households a financial cushion to cope with these risks. Demirguc-Kunt et al. (2018) highlight that households in developing economies are 27% less likely to report that they could come up with funds in case of an emergency as compared to their wealthier counterparts[1]. Therefore, the provision of financial services to the unbanked low-income households is considered as an important tool to enable them to handle, mitigate, and recover from these various shocks.

Keeping in mind the possible beneficial effects of providing financial access to the unbanked, the G20 made financial inclusion one of its pillars in 2010. They established the Global Partnership for Financial Inclusion (GPFI) and endorsed the first Financial Inclusion Action Plan that promises to foster financial inclusion worldwide. By the end of 2018, 92 developing and emerging countries had signed the alliance for financial inclusion that creates financial inclusion targets and drives policy changes to meet the challenges of promoting financial inclusion. In part, this increased concentration on financial inclusion is motivated by expectations that financial inclusion can pull unbanked people out of poverty, promote job creation through entrepreneurial activity, and stimulate income equality. However, these expectations are not without controversy, and are subject to an ongoing debate and several points of criticism. The next section describes what the term financial inclusion entails, and the subsequent section critically analyses the existing literature that has studied the impact of different dimensions of financial inclusion on socio-economic development.

II. What is financial inclusion?

Financial inclusion is broadly defined as the access to and use of formal financial services by households and firms. Those without such access are considered financially excluded. It is a multidimensional concept and it is crucial to note that there is still no universally accepted definition of financial inclusion or financial exclusion. The term financial inclusion was coined in the early 1990s and it referred to the access to commercial banks in the context of liberalization of the financial sector (Demirguc-Kunt and Klapper, 2012). However, Leyshon and Thrift (1995) highlighted that a large portion of population is not served by commercial banks because they lack

[1] This response was based on a survey question asking respondents whether they would be able to come up with an amount equal to one twentieth of GNI per capita (in local currency) within the next month in case of an emergency. These results rely on a nationally representative sample from over 140 economies in 2017.

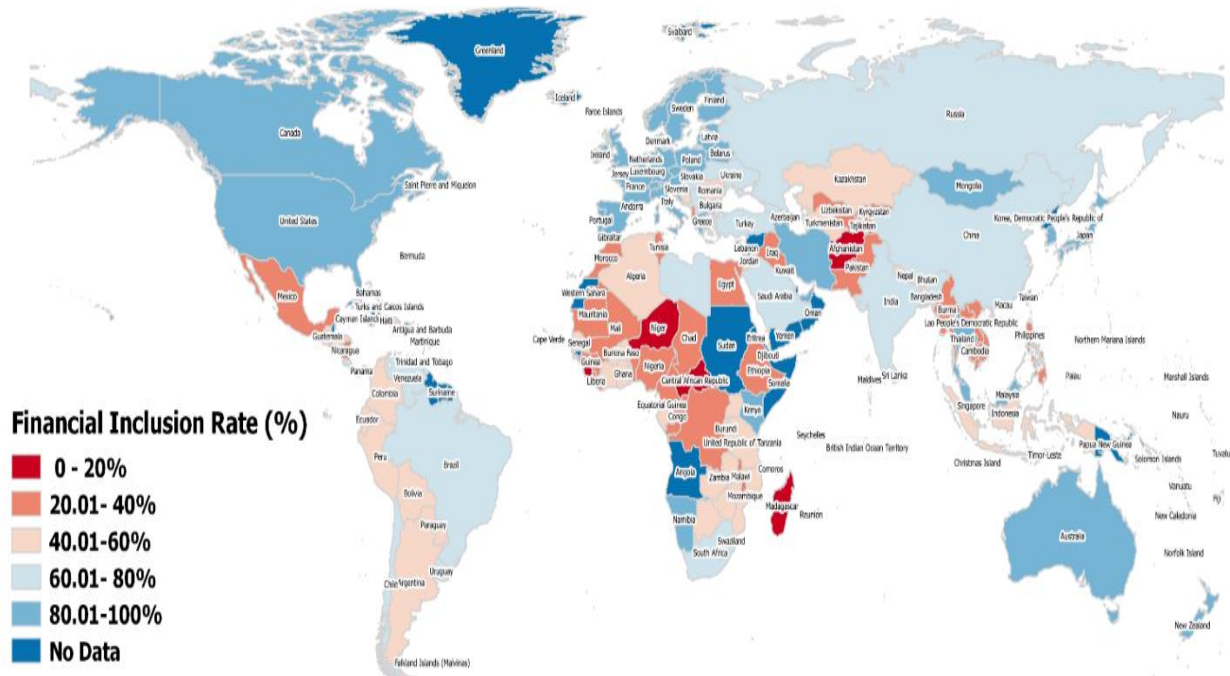
physical collateral or because of high costs associated with opening a bank account. They elaborated on the idea of financial exclusion as "*processes that serve to prevent certain social groups and individuals from gaining access to the financial system*".

Over the years, many scholars, policymakers and international organizations have established several definitions due to the evolving nature of various dimensions of financial inclusion or exclusion[2]. The varying types of financial services which come under the umbrella of financial inclusion include saving accounts, credit, insurance, and payment services. According to the definition by the World Bank, *"financial inclusion means that individuals and businesses have access to useful and affordable financial products and services that meet their needs – transactions, payments, savings, credit and insurance – delivered in a responsible and sustainable way"* (World Bank, 2014). Given the multidimensional nature of financial inclusion, the existing empirical literature has relied on different ways to measure it. Some have relied on information from the demand side, i.e. at an individual, household or firm level, whereas others have relied on information from the supply side i.e. financial institutions.

In its very basic form, financial inclusion sometimes refers to the ownership of a bank account at a formal financial institution. Having access to a bank account is considered as a first step to broader financial inclusion as it provides people with a safe place to save money and it increases their chances of utilizing other formal financial services such as credit, insurance, and payment facilities. The World Bank measures financial inclusion globally by looking at the percentage of adults in a country with a formal bank account. As of 2017, 69% of the adult population globally had a formal bank account as compared to only 51% adults in 2011. However, about half of the adult population in developing countries still remains financially excluded and about 200-250 million small and medium enterprises (SMEs) in developing countries remain underserved (World Bank, 2017). The map provided in Figure 0.1 highlights how the financial inclusion rate, measured here as the percentage of adults in a country with a formal bank account, varies across different regions in the world. These numbers go on to show that countries in Asia and Africa exhibit the lowest levels of financial inclusion.

[2] See the table in Appendix 0.1 for an overview of different definitions.

Figure 0.1: Financial inclusion rate worldwide



Source: World Bank's Global Findex database, 2017

Different dimensions of financial inclusion such as savings accounts, credit, insurance and payment facilities can help individuals as well as businesses in a variety of ways in their everyday life and in case of emergencies. The next section elaborates on the links between these different dimensions of financial inclusion and socio-economic development in detail and provides an overview of the existing literature.

III. Financial inclusion and development: What do we know?

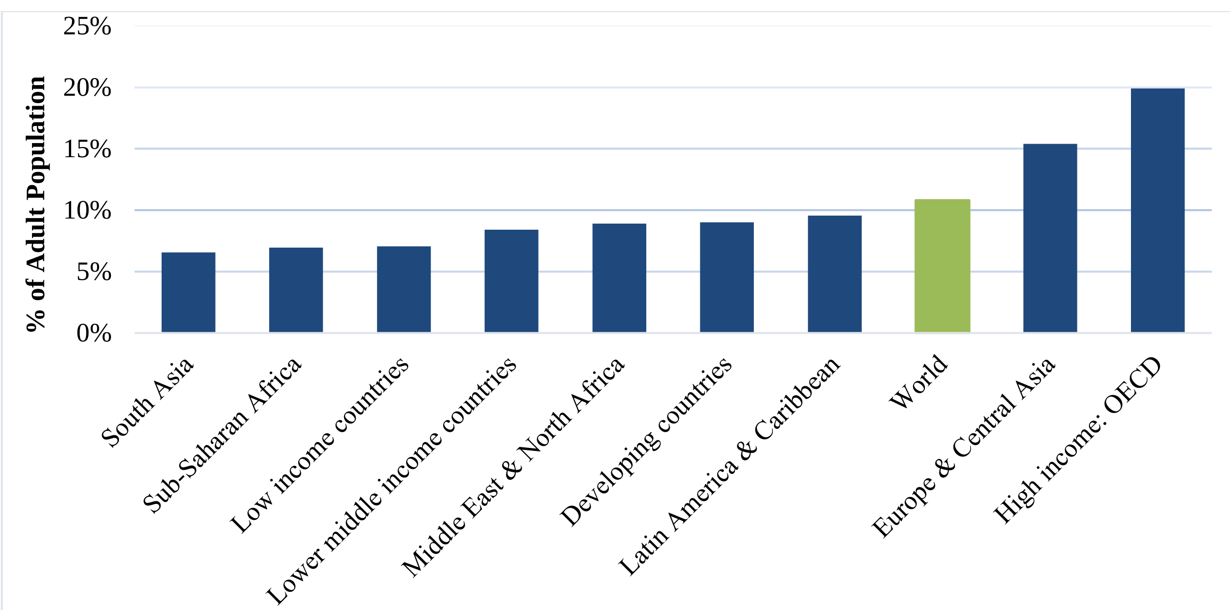
There has been an increased focus on the importance of financial inclusion, and it is generally believed that financial inclusion can play an important role in helping people, especially the poor, in improving their livelihoods and can spur economic activity. However, what does the empirical evidence have to say about this positive impact of financial inclusion on development? How does this impact vary across different dimensions of financial inclusion? What are the channels through which these relationships hold? This section tries to answer these questions and provides an overview of the literature that has critically studied the positive and negative impacts of different dimensions of financial inclusion on business outcomes, socioeconomic outcomes, and economic growth, at a microeconomic level as well as the macroeconomic level.

i. Microeconomic Level

A. Credit

Access to credit is considered particularly important, especially for low-income households, for a number of reasons. Households may want to avail credit to invest in their business or education, buy home or livestock, start a new business, organize weddings, manage unexpected emergencies, or simply to tackle various ups and downs in life. According to the World Bank's Findex database 2017, a little less than half (47%) of the world's adult population reported borrowing money in the past year. In the case of developing countries, friends and family were the most common source of borrowing, whereas, in the case of developed countries such as the OECD countries, the most common source of borrowing was a formal financial institution. Globally, about 11% of the adults (one quarter of borrowers) borrowed money from a formal financial institution. However, formal borrowing was quite low in some regions such as South Asia (6.6% of adults) and Sub-Saharan Africa (7% of adults) as highlighted in Figure 0.2. Countries in these regions mostly rely on informal sources of borrowing. A map depicting the concentration of formal borrowing around the world is provided in Appendix 0.2.

Figure 0.2: Borrowing from formal institutions across the world



Source: World Bank's Global Findex database, 2017

Accessing credit from a formal financial institution is generally considered better as compared to borrowing from informal sources such as moneylenders, market vendors, or suppliers. This can be due to various factors. First, some of the informal sources of borrowing such as moneylenders can charge exorbitantly high interest rates and prey on poor households who are in dire need of credit. The literature has generally referred to these moneylenders as “loan sharks”. A study of moneylenders conducted in rural Pakistan by Aleem (1990) showed that the average annual interest rate charged by moneylenders was about 78.5% with a standard deviation of 38.3%. This translates to the fact that interest rates ranging from 2% to as much as 150% lie in the 95% confidence interval, indicating very high variability of interest rates amongst moneylenders. Second, borrowing from a formal source might come with better credit terms as compared to informal sources. Third, when people borrow from friends and family or through other informal sources, they are restricted to funds within the community. These funds might not be sufficient to fulfil their credit needs especially if they belong to a poor community, and this can contribute towards increased inequalities and poverty traps (Demirguc-Kunt et al., 2017).

Microcredit, as a credit source, started more than three decades ago with the aim to provide credit to low income households which were previously excluded from the formal financial sector. It is generally believed that the idea of microcredit was first introduced by the Bangladeshi economist Dr. Muhammad Yunus. In 1976, Muhammad Yunus visited the village *Jobra* in Bangladesh and provided a loan worth $27 from his own pocket to 42 women to help them start a business (Yunus, 1998). He believed that these microloans could help in kickstarting a dynamic development process by helping micro entrepreneurs grow, generate income, hire more people, and reduce poverty. With the help of the Bangladesh central bank and international donors, Muhammad Yunus established the Grameen bank in 1983 and started providing microcredit to low-income households at a much larger scale. Over the next two decades, microfinance started to grow in developing countries all across the world. United Nations named 2005 as the “International Year of Microcredit”. Moreover, in 2006, Muhammad Yunus and the Grameen Bank won the Nobel Peace Prize “for their efforts through microcredit to create economic and social development from below”. Since then, microfinance has expanded rapidly, and the total microfinance clientele has grown by more than 16 times from 8 million in 1997 to about 139 million in 2017[3]. Currently,

[3] These statistics come from the Microfinance Barometer which is based on Microfinance Information Exchange dataset (2018). See http://www.convergences.org/wp-content/uploads/2018/09/BMF_2018_EN_VFINALE.pdf.

there are over 10,000 microfinance institutions all around the world providing financial services to low-income households.

There is a large body of the literature that has studied the impact of access to credit on a variety of outcomes and has found mixed results. Most of the studies which were conducted in the early 1990s and 2000s showed a significant impact of microcredit on different socio-economic outcomes, but they were mostly fueled by anecdotal evidence, descriptive statistics, or impact evaluations conducted by the microfinance institutions themselves (Hannig and Jansen, 2010). For example, a study conducted by the Grameen Bank claimed that 65% of their clients had crossed the poverty line (Grameen Bank, 2007). The recent empirical evidence conducted by independent academics has produced relatively modest results. The literature does not generally find a significant effect of microcredit on socio-economic outcomes such as poverty, health or education outcomes. However, the evidence is relatively more positive for some business-related outcomes.

Pitt and Khandker (1998) conducted one of the earliest impact evaluations of microcredit in Bangladesh by looking at microcredit offered by three microfinance institutions including the Grameen Bank. They found a positive impact of microcredit on household consumption, labor supply, assets, and school attendance rates of children. They also highlighted that the impact was more significant for female borrowers as compared to male borrowers. Moreover, a similar study conducted by Pitt, Khandker and Cartwright (2006) showed that microcredit had a significant impact on women's empowerment in Bangladesh. They showed that microcredit program improved women's decision-making power in the household, women had greater bargaining power, and also greater freedom of mobility. However, Morduch (2000) and Roodman and Morduch (2014) questioned the identification strategy of their empirical analysis and doubted the validity of their results, which led to a dynamic discussion between the authors on the validity of the early positive results on microcredit[4]. Similarly, some other studies have also shown that having access to credit has a positive effect on monthly income (Honohan and King, 2012), household income and consumption levels (Mahjabeen, 2008), household welfare (Khandker and

[4] In a reply to this criticism, Pitt and Khandker (2012) rejected the claims and called it misplaced criticism. See Pitt (2014) for a detailed discussion.

Faruqee, 2003), business creation (Bruhn and Love, 2014), and investment in human capital (Amendola et al., 2016).

In 2005, Banerjee et al. (2015) carried out the first randomized control trial of expanding access to credit in an urban market in Hyderabad, India. They partnered with one of the fastest growing microfinance organization in India named *Spandana* and selected 104 neighborhoods where the microfinance organization would have been interested in opening up a branch. They randomly selected 52 of these neighborhoods as treatment group where *Spandana* opened up a branch. The rest of the neighborhoods became the control group. The main findings of their study showed that there were no changes in health outcomes, education, or women's empowerment even after three years of intervention. However, they did find that small business investments and profits of pre-existing businesses increased. Moreover, they found that expenditure on durable goods for businesses and households increased while the expenditure on temptation goods (tea, alcohol, tobacco, gambling) went down.

Similarly, in 2006, Crépon et al. (2015) conducted one of the first randomized control trials on microcredit in a rural setting in Morocco. Their findings suggested that the take up of microfinance was surprisingly quite low (around 13%) which pointed towards a relatively low demand for microcredit. Secondly, they reported that households who had access to credit expanded their self-employment activity and their profits increased. Lastly, they also did not find any impact of having access to credit on education, health or consumption related outcomes. However, Bédécarrats et al. (2019) replicated this study using the same data and raised questions about the external validity as well as some concerns over the internal validity of their main findings.

Similar studies were conducted in Mexico (Angelucci, Karlan and Zinman, 2015), Bosnia and Herzegovina (Augsburg et al., 2015), Ethiopia (Tarozzi et al., 2015) and Mongolia (Attansio et al., 2015) and they reached parallel conclusions. Banerjee et al. (2015) summarized the findings of the six widely cited randomized control trials and concluded that access to microcredit had "a pattern of modestly positive, but not transformative effects". They concluded that whilst the businesses appeared to benefit from access to microcredit, it did not translate into broader development impacts such as poverty reduction or better education or improved health related outcomes. They also noted heterogeneity in business related outcomes across and within different RCTs, with some studies finding a notable change taking place at either tails of the distribution of business outcomes such as profits.

Some studies have also looked at the impact of improved access to credit on economic activity and have generally found positive effects. Bruhn and Love (2014) used a natural experiment in Mexico where a microfinance institution named *Banco Azteca* opened up 800 branches simultaneously. They showed that this increase in access to finance caused a 1.4% increase in overall employment, a 7.6% increase in informal businesses, and a 7% increase in income on average. Similarly, Burgess and Pande (2005) used state level panel data for India and showed that access to finance led to a significant reduction in rural poverty. There is also some evidence which suggests that flexibility in credit product design could result in improved impact (Field et al., 2010).

On the other side, there is some evidence which also alludes to the negative side of credit expansion. This strand of research highlights that microcredit can sometimes do more harm than good and can lead to increased poverty levels, exploitation of women, child labor, increased workload, and the creation of a culture of dependence which has an adverse effect on economic growth (Rooyen et al., 2012; Bateman and Chang, 2009; Copestake, 2002). The reasons cited behind this negative side of credit expansion generally include issues pertaining to over-indebtedness, exploiting nature of lending products, and non-productive use of credit. The crises of over-indebtedness and delinquency in certain microfinance markets around the world (prominently in India, Pakistan and Morocco) are a few examples which also support this view on the negative impact of credit expansion. Bateman (2010) even argues that microcredit is ineffective and it only offers an illusion for entrepreneurship creation and poverty reduction.

In summary, studies discussed in this section have highlighted that the evidence on the impact of access to credit is mixed at best. Studies have not generally found a significant effect on aggregate welfare of the households, although some small businesses seem to benefit from credit. An important point to emphasize here is that the impact of access to financial services on the labor market, especially on entrepreneurship creation, remains an understudied phenomenon.

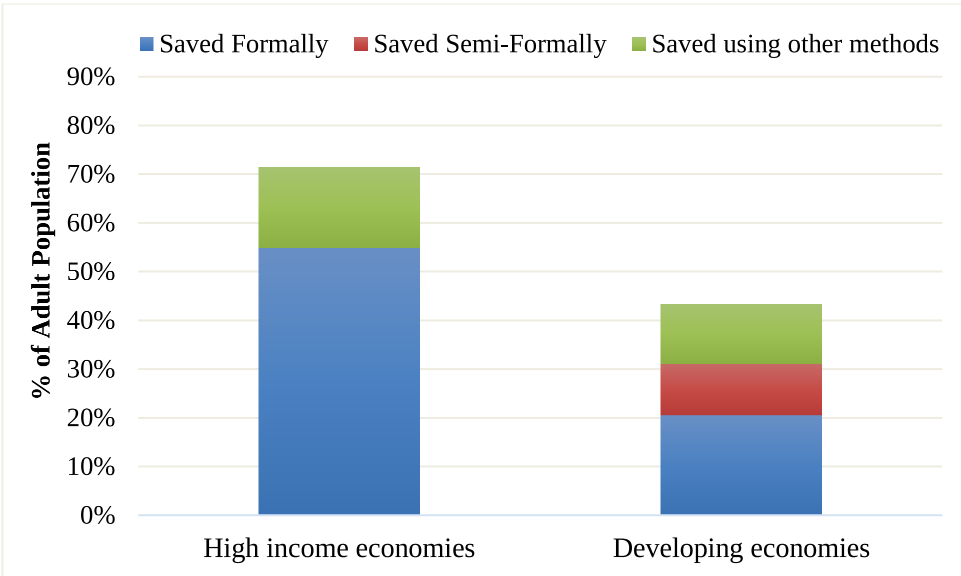
B. Savings account

Financial inclusion is not merely about credit. Another very important dimension of financial inclusion is access to savings accounts. Individuals save money in order to manage future expenses and investments. This includes, but is not limited to, saving for future investments in business or education, healthcare, old age, weddings, funerals, and other potential emergencies. Saving money at a formal financial institution can have several potential benefits as compared to saving at home or saving informally. First, saving at a formal financial institution provides safety against theft and other hazardous events such as fire or flood. Second, having savings at a formal financial institution can allow individuals to earn interest on their money which can offset the impact of increased inflation. Third, having savings at a formal financial institution rather than at home can curb impulsive spending and encourage better cash management (Karlan et al., 2017). Fourth, having money at a formal financial institution provides confidentiality and more control by making it more difficult for friends and family to access this money. This phenomenon is well documented by Anderson and Baland (2002) and Baland et al. (2011). Using case studies from Cameroon and Kenya, they highlight that women sometimes participate in Rotating Savings and Credit Associations (ROSCAs) to escape forced solidarity and to protect their savings from their husbands and family relatives. Similarly, Demirguc-Kunt et al. (2017) highlight that saving at a formal bank account can strengthen women's economic empowerment by giving them more control over the household funds.

According to World Bank's Findex database, as of 2017, about half (48.4%) of the world's adult population reported saving money in the past year. 71% of the adults in high income economies and about 43% of adults in developing countries reported saving. However, only about one quarter of the adult population (about half of the savers) around the world saved at a formal financial institution. Moreover, there are large gaps in formal savings between high income countries and developing countries. About 55% of the adult population (three quarters of savers) in high income countries saved formally, whereas only 21% of the adult population (half of savers) in developing countries saved formally.

In the case of developing countries, many people save either semi-formally or informally. Semiformal saving generally refers to saving in a savings club. One common example of these are the Rotating Savings and Credit Associations (ROSCAs). ROSCAs are generally community-owned and run by local people themselves. They are also known as *committees* and are quite common in developing countries. People who are part of a ROSCA generally pool their deposits weekly or bi-weekly and disburse the lump-sum amount to a different member each week or twoweeks. In 2017, 11% of the adult population in developing countries reported saving semi-formally in the past year, according to the World Bank's Findex database. However, one of the most common form of saving in developing countries remains informal savings (Figure 0.3). A large number of people keep their savings at home and 'under the mattress', or they save in the form of livestock or jewelry. These informal sources of savings were reported by about one third of the savers in developing countries (World Bank, 2017).

Figure 0.3: Adults saving any money in the past year (%)



Source: World Bank's Global Findex database, 2017

The results of the empirical studies on the impact of formal savings on welfare of households are generally positive. Studies which look at the impact of savings are relatively fewer as compared to the studies on credit, however, most of them have found encouraging results. A study conducted in rural Malawi by Brune et al. (2016) found a positive impact of providing bank accounts on aggregate savings. As a result of this, the authors also noticed an increase in business investment, crop output, and household expenditures. Similarly, other recent studies have found a significant impact of access to savings accounts on women empowerment and business investment. Karlan et al. (2017) studied the impact of a savings-led microfinance program in poor communities of three African countries including Ghana, Malawi and Uganda, and found a significant impact on business outcomes and women empowerment. However, they did not find any significant impact on average consumption and other livelihoods. In the Philippines, Ahsraf, Karlan, and Yin (2010) showed that having access to a commitment savings product – one that does not allow withdrawals until a specific date – led to a positive impact on women empowerment, self-reported household decision making, and purchase of durable goods in the household.

In case of Nepal, Prina (2015) showed that women household heads with a savings account were able to cope with income shocks in a better way and were able to reallocate their expenditures on education and food rather than on dowries. The author also noted a significant increase in the self-reported level of the financial situation of these women. Another study conducted in Nepal by Carvalho, Prina, and Sydnor (2013) found that women with no-fee saving accounts were less risk-averse one year later than the ones without a savings account. Dupas and Robinson (2013) conducted a similar randomized control trial in Kenya where they studied the impact of a commitment savings product. They found that having access to this account resulted in a significant increase in savings for the female market vendors, which in turn resulted in an increase in food expenditures for the households. Their results also highlighted that private expenditure went up by 13% and business investment went up by 56% for the treatment group as compared to the control group. However, the same study did not find this significant impact for male rickshaw drivers working in the same town.

These studies have highlighted that providing access to savings services at a very low cost to low-income individuals can have a positive impact on their welfare. On the contrary, there are also studies which did not find any evidence of increased savings or better business-related outcomes or better socio-economic outcomes as a result of having a bank account. Dupas et al. (2018) conducted a set of randomized control trials in Chile, Malawi, and Uganda and after following up with the households over the period of two years they concluded that having access to bank accounts had no significant impact on their level of savings or welfare. The authors highlighted that unsuitable product design or the fact that respondents in their study were relatively poorer as compared to other similar studies might be the possible explanation for their results.

In summary, studies on bank accounts have highlighted that savings facilities are a very crucial component of financial inclusion as they seem to help households smooth out consumption, build working capital, and manage cash flows more effectively. These findings are quite consistent across different studies and regions. This underlines that the importance of savings accounts over other financial services, such as credit, should not be overlooked.

C. Other financial services

Some other prominent financial services which fall under the umbrella of financial inclusion include insurance and payment facilities. Insurance is an important tool that can be used to manage shocks and mitigate risks. It is harder for poor households to rise above poverty when they are vulnerable to risks and external shocks such as death of their livestock, illness, crop failure, flooding, or drought. There is limited empirical evidence that has looked into evaluating the impact of insurance products, however, this limited literature does suggest that microinsurance can be an important instrument for mitigating risk. Although one common finding of this literature is that the demand and uptake of insurance products is extremely low in developing countries even when the insurance is offered for free (Matul et al., 2013). A field experiment conducted by Karlan et al. (2014) in Ghana found that households who obtained free microinsurance ended up adopting high risk crops, invested more in cultivation, and ended up with a higher revenue. Other studies conducted in different contexts have reached similar conclusions, but these studies do highlight that there is a need to rethink the design of the insurance product for low-income households (Cole et al., 2017).

Provision of fast and secure payment services is another feature of formal financial institutions. Households send or receive money to pay bills, receive salaries, send remittances, collect government transfers, purchase or sell goods and services. Evidence has suggested that making payments using formal accounts rather than relying on cash has many potential benefits. First, payments through accounts can reduce transaction costs and save time for both the sender and the receiver. Second, sending money through accounts increases transparency and can limit bribery and fraud (Demirguc-Kunt et al., 2017). Third, transferring money through accounts is generally safer and reduces the risk of theft and crime. Fourth, payments made through formal financial institutions can save money for governments as well as businesses.

A study conducted in Niger by Aker et al. (2016) found evidence of several benefits as a result of a social transfer program delivered through mobile money. They found that disbursing money through mobile transfers significantly saved time by decreasing the waiting and traveling time. They also showed that these mobile money transfers decreased the administrative costs of the program by about 20%. Lastly, the study showed that recipient household’s cost saving resulted in a higher diet diversity and fewer depleted assets. Similarly, another study in Kenya by Jack and

Suri (2014) considered the impact of lower transaction costs of mobile money through M-PESA on risk sharing. Their findings showed that the users of mobile money were able to successfully absorb negative income shocks without any reduction in their household consumption levels, whereas the consumption of non-users fell by about 7%. Moreover, some studies have also found an increase in the size and number of remittances due to availability of mobile banking. For example, Lee et al. (2017) documented an increase of 30% in remittances as a result of providing access to mobile banking from their field experiment in rural Bangladesh. Many microfinance institutions across the world are now providing services like digital payments and mobile banking to their customers. However, given the relative novelty of this phenomenon and due to lack of available data, it might take some time to have a better picture of how mobile money and digital payments impact the lives of people and their businesses in developing countries.

Based on this extensive literature that has argued about the positive impact of different dimensions of financial inclusion on socio-economic development, Klapper et al. (2016) claim that financial inclusion can even play an important role in reaching a number of sustainable development goals (SDGs) such as promoting inclusive and sustainable economic growth, promoting gender equality, reducing hunger and promoting food security, fostering quality education, achieving good health and well-being, and eliminating extreme poverty. However, the existing literature has presented mixed findings and the extent to which financial inclusion can aid in improving socio-economic outcomes and uplifting poor households, and the exact mechanisms through which these relationships work, remain open questions.

ii. Macroeconomic Level

Since the pioneering work on the relationship between finance and growth by Goldsmith (1969) and Shaw (1973), many economists have pointed out that financial development is good for growth. Aghion and Bolton (1997) emphasized that a lack of access to financial services can increase income inequality and poverty. Galor and Zeira (1993) illustrated this point by showing that frictions in the financial market has a negative effect on productivity, income generation, and investment in education for the poor. They show that poor households are trapped in a vicious cycle of poverty as a result of these frictions in the financial market.

The contribution of financial inclusion in economic development can run through a number of channels. First, the use of formal bank accounts is likely to increase the aggregate level of savings in the economy that in turn increases the total amount of investible resources. Dabla-Norris et al. (2015) show that an increase in financial inclusion can lead to a 0.2 to 0.6 percent increase in the investment to GDP ratio. Similarly, Aghion et al. (2009) show that this phenomenon has a positive impact on economic growth in the long run. Second, greater access to bank accounts and deposits makes the funding base of the financial institutions more resilient which is quite important for the economy especially in times of financial distress (Bayoumi and Melander, 2008). Third, an increase in the aggregate level of savings in the formal sector can have a positive effect on the supply of formal credit. This can enable financial institutions to diversify their portfolios and reach out to firms and households that were previously excluded (Cull et al., 2014). Consequently, financial inclusion can have a positive impact on economic activity. Fourth, financial inclusion can help in reducing income inequality. Demirguc-Kunt and Levine (2009) argue that financial inclusion decreases inequality by relaxing credit constraints on poor people who do not have collateral, credit history and connections. They also highlight that the easing of credit constraints increases the flow of credit to firms and people with better entrepreneurial ideas rather than to the ones with more collateral, which implies an increase in economic opportunities for those with less wealth. Fifth, higher level of financial inclusion can have a positive effect on the effectiveness of monetary policies (Mehrotra and Yetman, 2014). As an expanding number of people become part of the formal financial sector, it can allow central banks to use the interest rate as a primary policy tool (Dabla-Norris et al., 2015). Lastly, financial inclusion can help governments in targeting social spending in a more efficient and transparent way (Loukoianova et al., 2018).

The empirical literature that studies this relationship between financial inclusion and economic growth at the macroeconomic level relies on cross-country studies and is quite conclusive. This literature argues that broader access to financial services is not only positively correlated with, but also, causally related to economic growth (Rajan and Zingales, 1998; Clarke et al., 2006; Pasali, 2013; Samargandia et al., 2015). A recent study by the International Monetary Fund shows that bringing the level of financial inclusion in Pakistan to the Emerging Market's average would imply an increase in the economic growth rate by about 1.2%, after controlling for other factors (International Monetary Fund, 2017).

However, the literature at the macroeconomic level does point towards some caveats as well. A study conducted by Cecchetti and Kharroubi (2012) shows that the relationship between financial depth and economic growth is non-linear and shaped like an inverted “U”. They argue that at very high and very low levels of financial intermediation, this positive relationship between financial depth and growth disappears. Similarly, Demetriades and Law (2006) indicate that the positive relationship between financial development and growth does not hold for countries with low level of governance and weak institutional quality. Moreover, Rousseau and Wachtel (2002) show that this positive relationship does not hold for economies with high inflation rates and weak financial regulatory framework. Han and Melecky (2013) point out that further empirical research is required to better understand the lines of causation between financial inclusion and economic development.

Besides the importance of financial inclusion for development, there is another stream of the literature that focuses on the dark side of financial inclusion. This literature mainly discusses the tradeoff between financial inclusion and financial stability by examining the negative implications of credit expansion. Cecchetti and Kharroubi (2012) argue that financial development is beneficial up to a certain threshold and then it can become a drag on economic growth. Bunn and Romstom (2016) argue that the run up in households’ debt levels can even anticipate a financial crisis. Similarly, Schularick and Taylor (2012) label credit expansion as the most prominent predictor of financial instability, even more imperative than external imbalances. This strand of the literature highlights that since not everyone is creditworthy and can handle credit responsibly, a rapid increase in financial inclusion can hamper financial stability. This growing body of the literature suggests that too much or too fast finance can even plant the seeds for financial crises (Arcand et al., 2015; Mian and Sufi, 2014). The financial crisis of 2007-08 is a good example that demonstrates this issue. However, it is also crucial to realize that this vulnerability, due to excess finance, is not just a feature of financial markets in the developed economies.

Another example that highlights the negative implications of credit expansion is the microfinance crisis in India. In 2010, about one hundred borrowers of microfinance institutions in the State of Andhra Pradesh committed suicides due to issues linked with over-indebtedness and coercive recovery practices of financial institutions. As a result of this crisis, the government issued an emergency ordinance bringing the operations of microfinance institutions to a complete halt in the state. This contraction in microfinance supply had an adverse effect on the labor market and on consumption levels as highlighted by Breza and Kinnan (2018). Similarly, the microfinance industry in Bosnia, Pakistan and Morocco also went through a similar crisis where thousands of borrowers were over-indebted which had serious consequences on the wellbeing of their households and communities (Rooyen et al., 2012). There is research which has highlighted that credit expansion programs can actually be harmful and can plunge the poor households deeper into debt which has negative consequences for the economy (Rooyen et al., 2012; Bateman and Chang, 2009).

In summary, microfinance industry and the financial sector in general has had its share of challenges that have been exposed by empirical studies. These lacunae highlight the need for future research to better understand the real effects of financial inclusion so that financial products and services can be designed in a way that cater to the needs of recipients in an efficient and sustainable way.

IV. The Context of Pakistan

This thesis is comprised of four chapters. Two are based on Pakistan and one each on Mexico and the U.S. Since the majority of this research focuses on Pakistan, this section provides a broad overview of the country's characteristics and describes the landscape of access to finance in Pakistan[5].

The choice of Pakistan as the primary research setting is justified by two main factors. First, Pakistan has one of the fastest growing microfinance industries in the world and the country was ranked among the top 5 in the world for its enabling environment for financial inclusion by the Economist's Intelligence Unit. There are more than 3500 microfinance outlets in Pakistan which provide low interest rates, credit facilities not requiring collateral, and other financial services such as bank accounts, remittances and payment services. The banking sector is regulated by the State Bank of Pakistan (SBP) and Securities and Exchange Commission of Pakistan (SECP). The international community also holds significant stakes in the effectiveness of financial inclusion

[5] The country context of Mexico and U.S. is provided in their respective chapters.

initiatives in Pakistan: FINCA, Telenor, Dubai Islamic Bank and KIVA, for example, are some of the key players. Secondly, since the start of the War on Terror after the September 11 attacks in 2001, there has been a dearth of research at the microeconomic level in Pakistan, primarily due to security concerns and data constraints. However, in recent years, the situation has improved, and this thesis attempts to fill that research gap by providing some country specific policy recommendations on the role of financial inclusion in Pakistan.

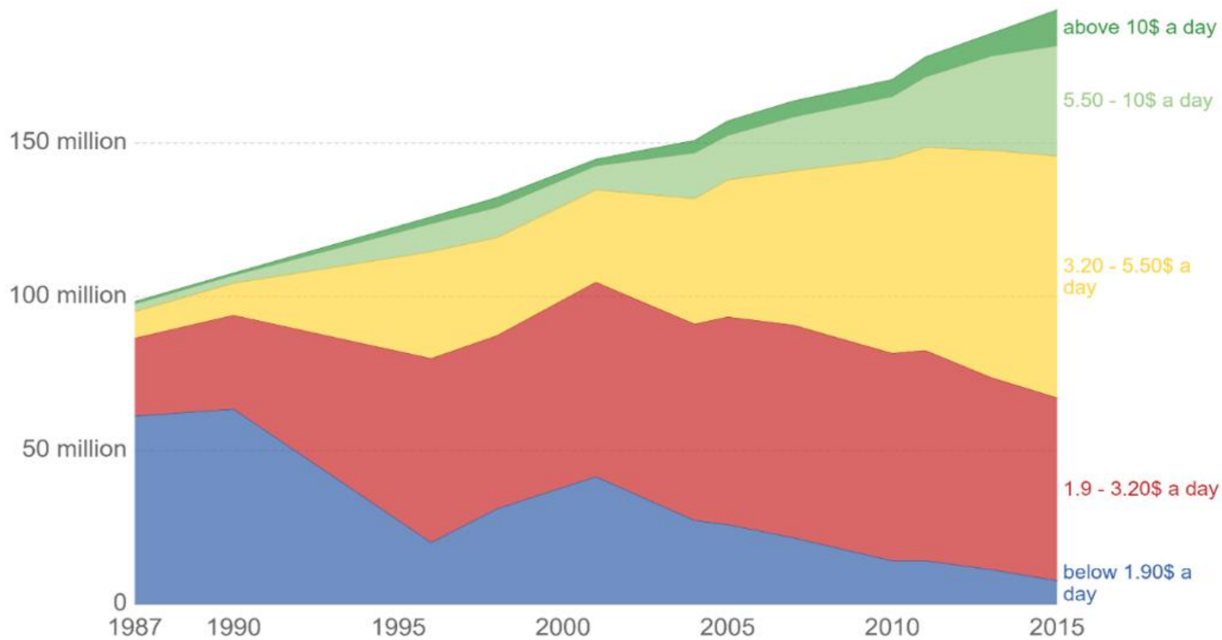
Pakistan constitutes of four provinces that are further divided into 158 districts. Pakistan is the sixth most populous country in the world with a population of over 205 million people (Pakistan Bureau of Statistics, 2017). About 39.2% of the population lives in urban areas whereas the rest resides in rural areas. The main religion is Islam with about 96.4% of the total population being Muslim (Pakistan Bureau of Statistics, 2017). The World Bank classifies Pakistan as a "Lower Middle Income" country.

The GDP growth in Pakistan averaged 4.5% between 2014 and 2017 (IMF, 2018). GDP per capita (PPP) has reached US$ 5,870 but remains below the South Asian regional average of US$ 7,770 (IMF, 2018). With regard to the labor force, the unemployment rate in Pakistan is about 5.9% and the informal sector is quite big and accounts for about 73% of the non-agricultural employment (Pakistan Bureau of Statistics, 2015). The agriculture sector in Pakistan accounts for about one fifth of the GDP but the country has seen a rapid growth in industries (textile, cement) and the services sector over the last few years.

i. Status of Poverty in Pakistan

Poverty in Pakistan has significantly decreased over the last two decades but still remains high. Extreme poverty, measured by the World Bank as people living below the international poverty line of $1.9 per person per day, has declined considerably to 4% and is below the regional average (South Asia average is 16%). Over the 2000-2015 time period, the poverty headcount in Pakistan more than halved (IMF, 2017). Figure 0.4 provides a distribution of population in Pakistan across different poverty thresholds.

Figure 0.4: Distribution of population between different poverty thresholds in Pakistan

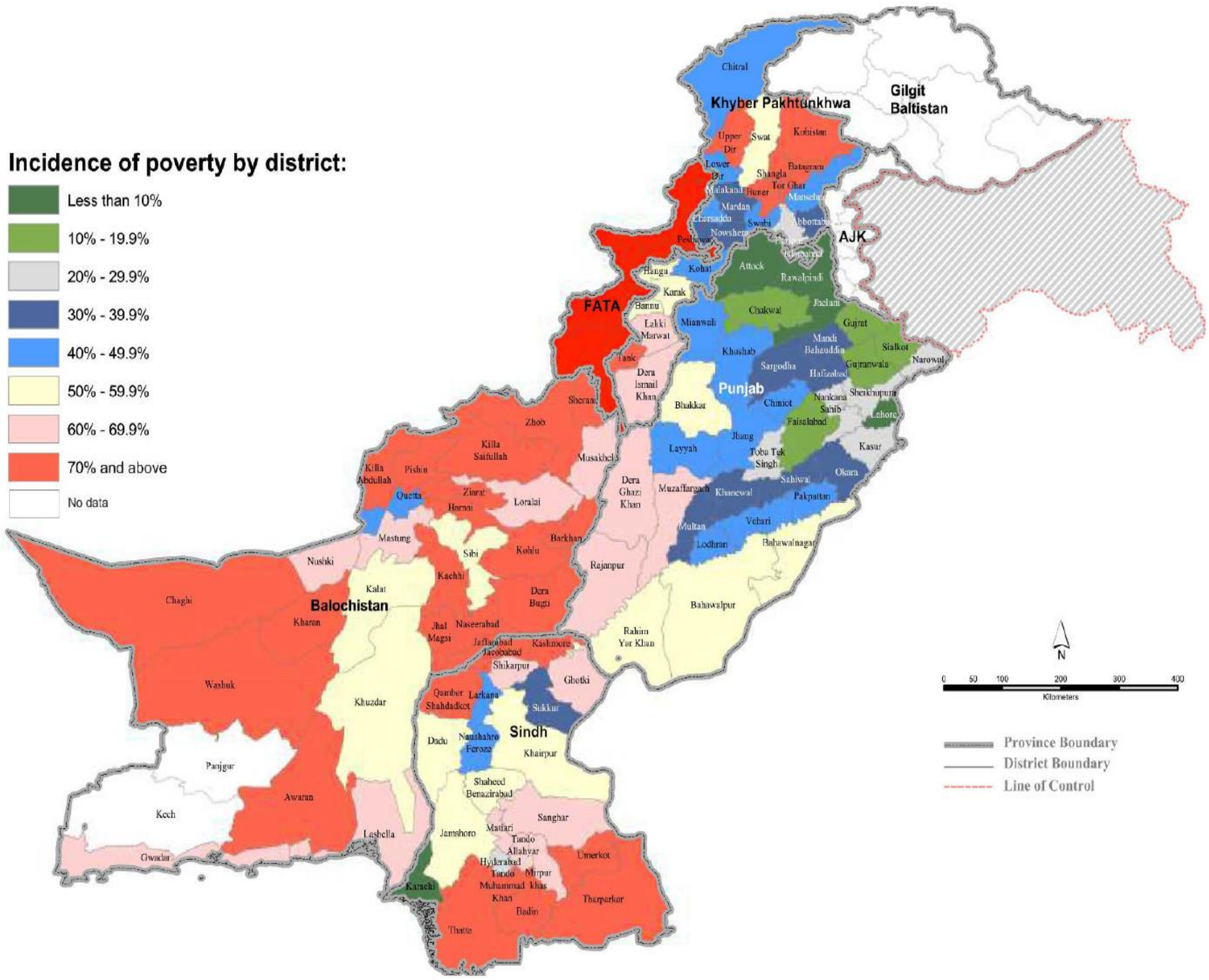


Note: Poverty thresholds are in 'international dollars' at constant 2011 PPP prices.
Source: World Bank's Poverty and Equity Data, 2015

In spite of these improvements, the multidimensional poverty index by the United Nations Development Program (UNDP) reports a higher incidence of poverty in Pakistan. This index takes into account three different dimensions of development including education, health, and the standard of living. This index is calculated using the Alkire-Foster methodology[6]. According to this index, incidence of poverty stands at 38.8% of the population, with some prominent regional discrepancies (Figure 0.5). Multidimensional poverty is higher in rural areas (54.5%) as compared to urban areas (9.4%). Moreover, amongst all provinces, *Balochistan* has the highest rate of multidimensional poverty (70%). *Balochistan* is also the biggest of the four provinces in terms of area (44% of total land), but it represents only 6% of the total population. Figure 0.5 provides a map of Pakistan highlighting the stark differences in multidimensional poverty across different regions.

[6] Multidimensional poverty takes into account 15 different indicators in three different dimensions of development: Education, health and standard of living.

Figure 0.5: Multidimensional poverty in Pakistan

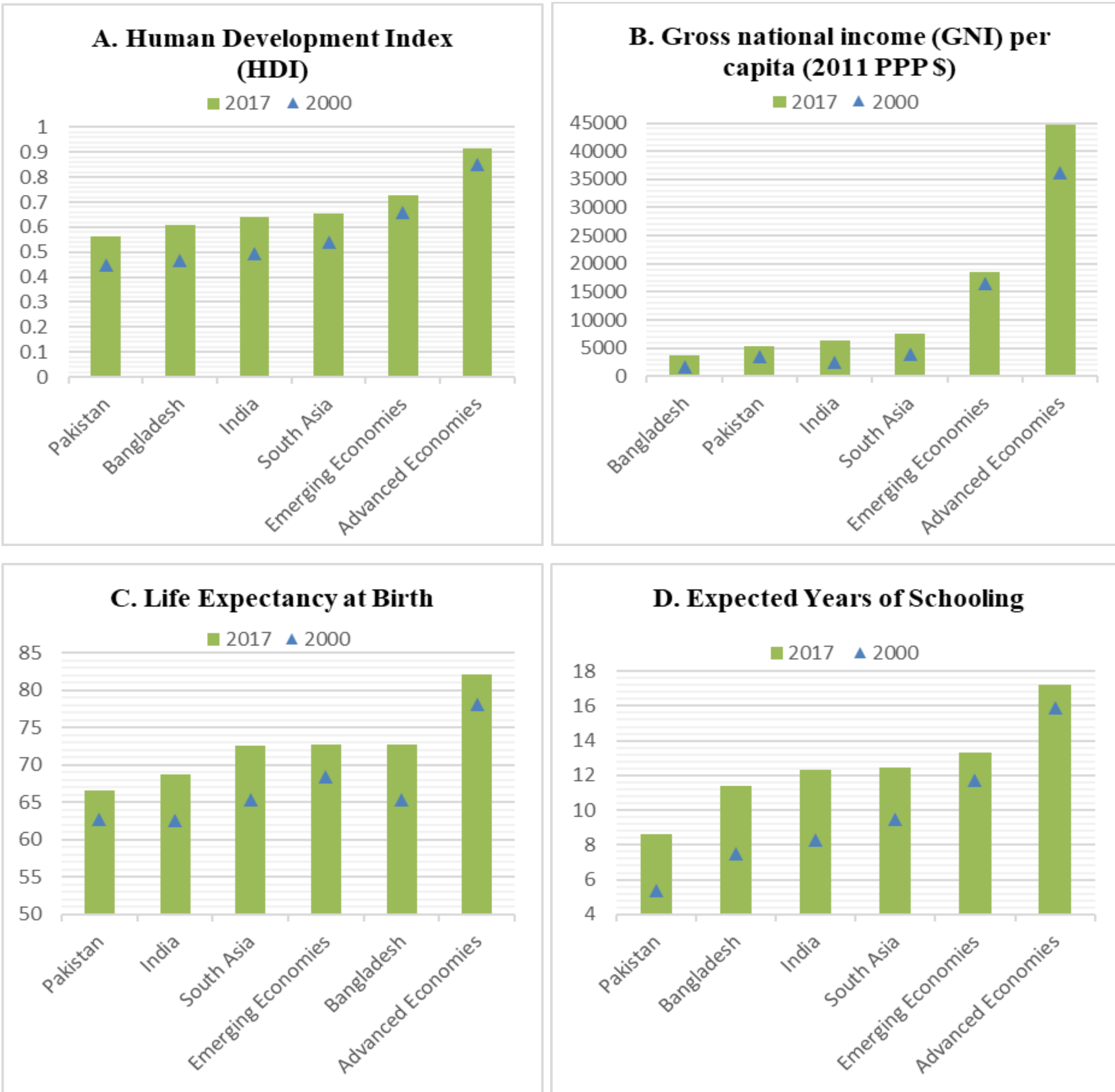


Source: UNDP, 2017

Socioeconomic outcomes in Pakistan have improved over time but remain below the regional averages. Pakistan was ranked 147th out of 189 countries on the Human Development Index (HDI)[7] by the United Nations in 2016. Moreover, the literacy rate in Pakistan is about 58% which is one of the lowest in South Asia. In terms of life expectancy at birth, expected years of schooling, and HDI, Pakistan lags behind regional comparators such as Bangladesh, India, South Asian average, and the average of emerging economies (Figure 0.6).

[7] HDI ranges from 0 to 1 where 1 reflects the highest level of human development. It is a summary measure which is computed after taking into account life expectancy at birth (health dimension), average years of schooling (education dimension) and gross national income per capita (standard of living dimension) in a country.

Figure 0.6: Socioeconomic outcomes in Pakistan as compared to regional peers

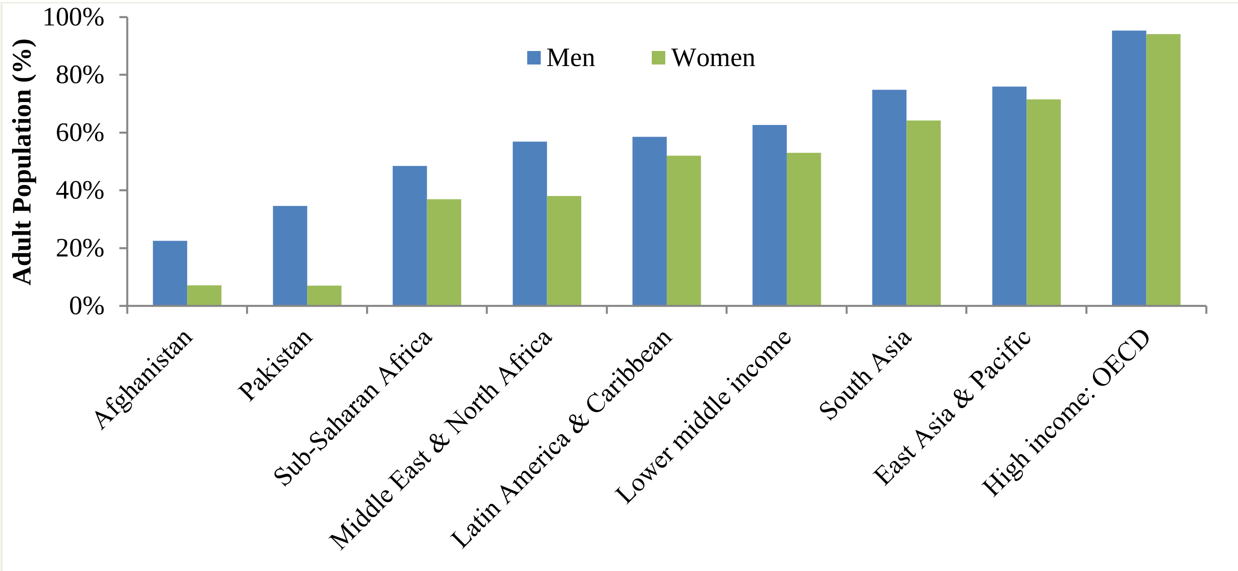


Source: World Development Indicators, 2017

ii. Access to Finance in Pakistan

Pakistan has made some improvements regarding financial inclusion over the last decade. As of 2017, 21% of the adults in Pakistan had a formal bank account as compared to 13% in 2014 and 10% in 2011 (World Bank, 2017). However, a large segment of population remains financially excluded with large gaps across gender and across urban and rural locations. Figure 0.7 provides a comparison of Pakistan vis-à-vis other regions around the world in terms of ownership of formal bank accounts. The numbers go on to show that Pakistan lags behind regional comparators in terms of formal bank account ownership as well.

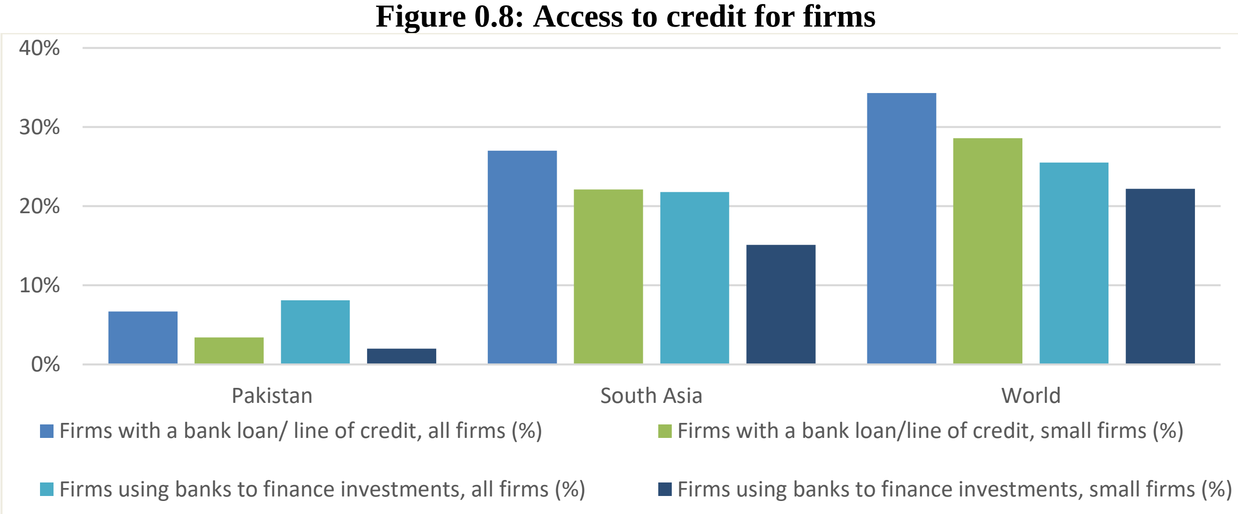
Figure 0.7: Adults with an account at a formal financial institution (2017)



Source: World Bank's Global Findex database, 2017

World Bank's *Ease of Doing Business Report* lists access to finance related issues as one of the top constraints for the private sector in Pakistan. A large number of firms, especially small firms, do not have access to formal finance in Pakistan. Only 3.4% of small firms in Pakistan reported having access to formal credit (Figure 0.8). Furthermore, Pakistan generally lags behind regional peers in terms of access to finance not only for individuals but also for firms, as highlighted in Figure 0.8.

Figure 0.8: Access to credit for firms



Source: World Bank's Enterprise Survey, 2013

The State Bank of Pakistan (SBP) launched the National Financial Inclusion Strategy (NFIS) in 2015 with the goal of expanding access to financial services to at least 50% adults by the end of year 2020. In the last few years, Pakistan has experienced a rapid growth in the banking sector. The total number of bank branches has gone up by about 45% between 2010 to 2017 (Pakistan Bureau of Statistics, 2018). As of 2018, there are a total of 53 banks in the country which operate through more than 15,000 branches all across the county (State Bank of Pakistan, 2018). Due to these increased efforts, Pakistan has been ranked among the top 5 countries in the world for its enabling environment for financial inclusion (Economist Intelligence Unit, 2015). Table 0.1 highlights some of the key statistics with respect to the access to finance in Pakistan from the supply side as well as the demand side.

Table 0.1: Access to finance statistics for Pakistan

Demand Side Statistics (Bank accounts)			**Supply Side Statistics**	
(% of adult population)	2008	2015	Number of Banks	53
Banked (A)	11%	16%	Number of Branches	15,464
Other Formal (B)	1%	7%	ATMs	14,361
Informally served (C)	32%	24%	POS terminals	49,261
Financially Served (A+B+C)	44%	47%	Branchless Banking Agents	405,571
Financially Excluded	56%	53%	Total number of borrowers	7 million plus
			Total number of accounts	53 million plus
Source: SBP's Access to Finance Survey, 2015			*Source: SBP, 2018*	

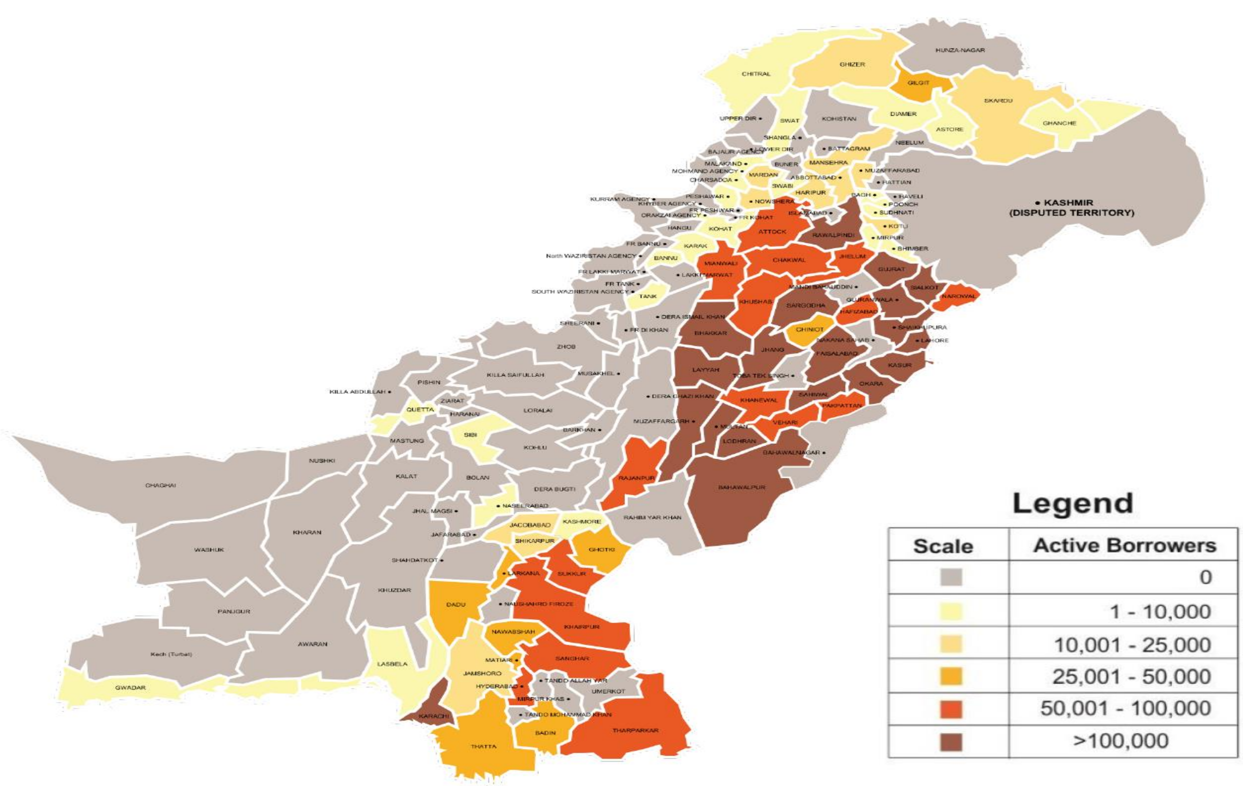
Microfinance is responsible for much of the progress in terms of financial inclusion in developing economies where institutional quality is weak and problems pertaining to limited access to formal financial services and information frictions in the credit markets are quite prevalent (Morduch, 1999). Pakistan's microfinance sector is one of the fastest growing microfinance sectors in the world. Pakistan was ranked 3rd out of 55 countries in 2013 in terms of having one of the most favorable environments for microfinance by Economist's Intelligence Unit in their report *Microscope on the Business Environment for Microfinance*. The total number of microfinance branches increased by more than 100% from 2012 to 2017 (Pakistan Microfinance Network, 2017). Moreover, the numbers of active borrowers grew by about 3 times while the number of savers grew up by about 7 times from 2012 to 2017 (Pakistan Microfinance Network, 2017). Pakistan now has more than 40 microfinance providers, which operate through a network of 3,673 microfinance branches with about 5.8 million active borrowers and 30 million saving accounts across the country (Table 0.2). About 45% of the total microfinance loans are group loans and the rest are individual loans (Pakistan Microfinance Network, 2017). The average loan size is about 48,695 Pakistani Rupees (US$ 400 approx.). It is crucial to note that there are large gaps in terms of access to microfinance across different regions in Pakistan i.e. across different provinces and districts. On average, a district in Pakistan has less than 2 microfinance branches per 100,000 adults, but the concentration of microfinance varies considerably across regions. The concentration of borrowers is relatively lower in the western regions (Balochistan province) and in the northern regions (KPK province) as compared to the provinces of Punjab and Sindh (Figure 0.9).

Table 0.2: Outreach of microfinance in Pakistan

Indicators	2012	2017	Change (in %)
Number of microfinance branches	1,739	3,673	111%
Active borrowers (Millions)	2.1	5.8	176%
Gross loan portfolio (PKR Millions)	28,845	202,699	603%
Average loan size (PKR)	21,126	48,695	130%
Number of saving accounts	3,933,496	30,984,717	688%
Value of savings (PKR million)	15,508	186,941	1105%
Average saving balance per account (PKR)	3,942	6,033	53%
Microinsurance amount (PKR Millions)	30,136	198,680	559%

Source: Pakistan Microfinance Network, 2017

Figure 0.9: Concentration of microfinance borrowers



Source: Pakistan Microfinance Network, 2017

Pakistan has made substantial improvements regarding financial inclusion over the last decade. These improvements are the result of policies and regulations focused on enhancing access to formal financial services for people who have been financially excluded. Financial inclusion for women has also been a top priority in the national financial inclusion plan and substantial efforts have been made in this regard e.g. setting up of the *First Women's Bank* and making huge investments in *Kashf Foundation* which is one of the largest microfinance institutions in the country providing financial services exclusively to women. However, the gender gap in financial inclusion in terms of formal bank accounts has increased from 16 to 28 percentage points from 2014 to 2017. According to the data from World Bank's Global Findex database, Pakistan has the fourth worst gender gap in the world with respect to bank accounts at a formal financial institution. Only Jordan (30 percentage points), Bangladesh (29 percentage points), and Turkey (29 percentage points) have a bigger gender gap[8]. Therefore, it is important to research what drives financial exclusion and classify policy measures to address these issues.

V. Outline of the Thesis

This PhD thesis is a collection of four chapters in empirical development economics. It attempts to provide a better understanding of the many interrelations between financial inclusion, entrepreneurship, and gender by relying on nationally representative pseudo panel data at the individual level. A large amount of existing evidence, summarized in the sections above, has looked into the impact of different dimensions of financial inclusion on household welfare and economic growth. However, there is limited evidence which has considered the impact of financial inclusion in its entirety, and not just credit, on the labor market. More specifically, the link between financial inclusion and entrepreneurship has not received adequate attention in the literature. The labor market is one of the channels through which financial inclusion may help reduce poverty and improve women's autonomy and therefore it is important to understand this link. So, one of the main objectives of this PhD thesis is to attempt to fill that research gap. Secondly, this thesis also aims to provide a better understanding of the drivers of financial exclusion in the context of credit after making a distinction between voluntary and involuntary financial exclusion. Voluntary

[8] A world map depicting this gender gap is provided in Appendix 0.3.

reasons are usually linked with individual preferences, culture or religious practices, whereas involuntary reasons may arise due to asymmetric information and other market failures. Given that financial inclusion has risen onto the global policy agenda, this thesis strives to provide country specific policy recommendations on how to promote financial inclusion by differentiating between the two types of financial exclusion. Lastly, keeping in mind the consequences of over-indebtedness and financial instability, this thesis adds to the standard line of research on household financial vulnerability assessment by proposing an alternative methodology using unsupervised machine learning algorithms. The four chapters in this thesis can be read independently and contain an extensive introduction. All these chapters address the concept of financial inclusion in a different manner and attempt to improve our understanding of the various concepts that revolve around the multidimensional nature of financial inclusion. The outline of the rest of the thesis is provided below where I summarize the main motivation, research questions, empirical methodology, and outcomes of each of the four chapters.

Chapter 1 is titled *"Access to microfinance and the economic ladder: Empirical evidence from Pakistan"* and aims to assess the relationship between geographical access to microfinance and entrepreneurship. Microfinance has played a vital role in promoting financial inclusion in developing countries where weak institutions along with information frictions in the credit markets limit access to financial services for large segments of the population. However, the role and impact of microfinance remains an important and debatable area of research for policymaking. Very few studies have examined the effect of access to microfinance, in its entirety and not just group microcredit, on entrepreneurship. On one hand, Yunus (2010) in his book *Building Social Businesses* argues that poor people have great entrepreneurial skills and that they remain poor because they do not have the opportunities to turn their creativity into sustainable income. He argues that providing these people with access to finance will harness their entrepreneurial spirit and will allow them to start their own businesses. Some of the recent empirical evidence does support this original idea that microfinance increases and promotes entrepreneurship (Augsburg et al., 2015; Attanasio et al., 2015). However, on the other hand, Duflo (2010) argues based on the work conducted in India that only a fraction of those who are not entrepreneurs have the will to become one and extending access to microcredit cannot be expected to convert everyone into self-employed businessmen. Bateman (2010) also claims that microfinance only offers an illusion for entrepreneurship creation and poverty reduction, and stresses that other factors such as regulations and business environment might be more important for entrepreneurship creation than access to finance. So, the main research question that is studied in this chapter is: **Does increased geographical access to microfinance institutions promote entrepreneurship and allow individuals to move up the economic ladder?** This chapter also discusses the different channels through which the relationship between microfinance and entrepreneurship operates. It relies on the two waves of the Financial Inclusion Insight (FII) Survey, from the years 2015 and 2016, to conduct this empirical analysis. This survey is a nationally representative survey of adults[9], comprised of about 12,000 observations, and is conducted by Gallup International Association under the supervision of Bill and Melinda Gates Foundation and InterMedia USA. Different maximum likelihood estimation techniques (probit and multinomial logit) are used to understand this relationship. Moreover, in order to control for endogeneity bias coming from non-random placement of microfinance institutions, bivariate probit model is used after relying on concentration of microfinance in the neighboring district as an instrumental variable (IV). Our main result is that geographical access to microfinance has a statistically and economically significant effect on entrepreneurship. Our findings show that having geographical access to microfinance can allow individuals, both men and women, to move up the economic ladder by shifting to entrepreneurship and running their own business rather than working as low paid employees, farm workers, and housewives. Econometric estimates show that having access to microfinance increases the chances of being an entrepreneur by about 4 percentage points, on average. However, we find that this impact of access to microfinance on entrepreneurship runs only through the poverty channel i.e. access to microfinance is more effective in the poorer districts. Lastly, the multinomial logit regression results show that having geographical access to microfinance does not increase the likelihood of becoming an entrepreneur for people who are either unemployed or working as farm owners.

Chapter 2 is titled *"Financial inclusion: An illusion for creating women entrepreneurship?"* and studies the interface between financial inclusion and women entrepreneurship across formal and informal work and across different economic sectors in Mexico. Financial inclusion and women

[9] Aged more than 15 years. One person per household is surveyed.

entrepreneurship concern policymakers because of their potential impact on job creation, economic growth and women empowerment. The existing literature has generated confounding views on this matter and has generally relied on a narrow definition of financial inclusion. This chapter proposes a new way of measuring financial inclusion at the municipality level in Mexico after taking into account different indicators of access and usage of financial services. The main research question which is studied in this chapter is: **Does financial inclusion promote women's autonomy by generating women entrepreneurship? How does this relationship vary across formal and informal work and across different economic sectors?** In order to conduct this empirical investigation, we build a dataset that comes from several sources including the National Survey of Occupation and Employment (ENOE for its acronym in Spanish), National Banking and Securities Commission (CNBV for its acronym in Spanish), the National Institute of Statistics and Geography (INEGI for its acronym in Spanish) and the Ease of Doing Business database by the World Bank. It is an individual-based pseudo panel dataset with more than 4 million observations. The data spans over the time period between 2009-2015 with a quarterly frequency. Limited dependent variable models are used to gauge this relationship between financial inclusion and women entrepreneurship. In order to control for possible endogeneity bias arising from simultaneity between financial inclusion and women entrepreneurship, instrument variable (IV) estimation and different time lag dynamics are used. The main findings of our research suggest that financial inclusion is positively linked with entrepreneurship, meaning that it can open up economic opportunities for women entrepreneurs. However, the positive relationship between women entrepreneurship and financial inclusion does not hold for women entrepreneurs working in rural areas in the informal sector or women working in the commerce sector, highlighting lower entry barriers, including financial, in the informal sector. The results also highlight that the probability of a woman being an entrepreneur in the informal sector is 13.5 percentage points higher as compared to her being an entrepreneur in the formal sector. Results also point towards the existence of gender disparity in the status of entrepreneurship across formal and informal work. On an average, women are about 3 percentage points less likely to be entrepreneurs in the formal sector and 11 percentage points more likely to be entrepreneurs in the informal sector, as compared to men, after taking into account other relevant individual and geographical characteristics. Lastly, the chapter classifies a number of policy measures that can be taken into consideration in order to improve the status of financial inclusion and possibly narrow down the gender gap in financial inclusion.

Chapter 3 is titled *"Pushed into the corner: An empirical study on the drivers of financial exclusion"*. In spite of an increased amount of evidence on the importance and impact of financial inclusion on different dimensions of development, little is known about the underpinnings of financial exclusion with respect to credit and the policies than can be used to alleviate it. It is generally assumed that households in developing countries, especially poor households, have a very high demand for formal credit and they remain financially constrained. Looking at raw descriptive statistics on financial exclusion supports that view. However, this assumption is at odds with the actual experiences of credit officers and bankers who frequently report issues pertaining to finding qualified microcredit clients in developing countries (Aslam and Azmat, 2012). Moreover, many methodological challenges remain in a comprehensive analytical underpinning of the drivers of financial exclusion with respect to credit. This chapter highlights these methodological challenges and tries to provide a detailed understanding of financial exclusion, at the microeconomic level, after taking into account the need for credit and after differentiating between two different types of financial exclusion: voluntary and involuntary exclusion. The main research question that is studied in this chapter is: **What are the main drivers of financial exclusion with respect to credit?** For the purpose of this empirical analysis, this chapter mainly relies on the nationally representative individual-level data from the Financial Inclusion Insight Survey of 2015 and 2016 from Pakistan[10]. The empirical analysis is based on the use of a sequential logit model. Heckman estimation is also used to address issues pertaining to selection bias. The main results highlight that the extent of involuntary financial exclusion is considerably less than what is conventionally believed once effective demand and voluntary exclusion is taken into account. Our findings highlight that the demand for credit is perhaps overrated as 38% of the adult population show no need for credit and about 25% opt for voluntary financial exclusion. Econometric estimates show that financial illiteracy and poverty are amongst the main determinants of involuntary financial exclusion. According to the results, financially literate

[10] This dataset is the same as the one being used in Chapter 1.

people are about 5.8 percentage points less likely to be financially excluded involuntarily as compared to people who are financially illiterate. This shows that the degree of financial literacy can make a big difference in reducing involuntary financial exclusion and it seems to have a clear policy message. Results also point towards the existence of gender disparity. Women are about 19.7 percentage points more likely to be involuntarily financially excluded, as compared to men, after controlling for other individual and regional characteristics. Regional characteristics also seem to be strongly associated with involuntary financial exclusion in Pakistan.

Chapter 4 is titled *"Assessing household financial vulnerability using machine learning: Empirical evidence from the U.S"*. Besides the importance of financial inclusion for development, there is another stream of the literature that focuses on the negative implications of credit expansion and highlights the potential trade-offs between financial inclusion in credit and financial stability. Cecchetti and Kharroubi (2012) have argued that financial development is beneficial up to a certain threshold and afterwards it can become a drag on economic growth. Schularick and Taylor (2012) labelled credit expansion as the most prominent predictor of financial instability, even more important than external imbalances. Similarly, Bunn and Romstom (2016) argued that the run up in households' debt levels can even anticipate a financial crisis. Since the 2007-08 global financial crisis, the concept of household financial vulnerability has gained considerable attention as it concerns policymakers due to its impact on macroeconomic indicators of financial instability for developed as well as emerging economies. The rising levels of household debt in many countries are becoming a source of concern for central banks worldwide. However, financial vulnerability is a complex multidimensional concept and large gaps still remain in the analytical underpinning of a comprehensive financial vulnerability assessment at the household level. In recent times, due to advances in digital technology, machine learning and big data have emerged as possible solutions to more accurately estimate financial conditions of households. Therefore, in this methodological chapter, a novel approach is proposed to measure households' financial vulnerability as an alternative to the standard line of research on this topic. The main research question that is being addressed in this chapter is: **What drives household financial vulnerability and how to measure it using unsupervised machine learning?** Nationally representative household data from the Survey of Consumer Finance (SCF) by the U.S. Federal Reserve, for the years 1998, 2007 and 2016, is being used to conduct this analysis. This chapter relies on a two-step empirical strategy. First, Hierarchical Ascending Clustering (HAC) and K-means clustering analysis is undertaken to identify homogenous clusters of households that are financially vulnerable. For that purpose, we use three sets of variables including the leverage ratio, debt burden ratio, and the households' income level. Afterwards, the probability of being financially vulnerable is estimated depending upon different household and geographical characteristics using a logistic regression. The main results indicate that about 28% of the households in United States are financially vulnerable as of 2016, which is 4 percentage points less as compared to 2007. This highlights that the extent of household financial vulnerability is under-represented when the conventional methodologies are used. Moreover, the results of the econometric estimates highlight that African-Americans and Hispanic Americans are more likely to be financially vulnerable than non-Hispanic white persons, after taking into account other household and regional level characteristics. Econometric estimates also highlight the existence of large gaps in household financial vulnerability across other household characteristics, such as education level, employment status, gender, marital status, and age of the household head. Empirical findings advocate that a buoyant labor market and other policy levers such as increasing educational attainment may help mitigate household financial risks. All these empirical results are obtained after accounting for imputation error and sample variability error.

It is important to emphasize once again that all four chapters in this thesis look at different perspectives of financial inclusion. The first chapter looks at geographical access to microfinance, the second chapter builds a financial inclusion index using supply-side data on access and usage of different financial services at the municipality level, the third chapters looks at financial exclusion solely from the perspective of credit by differentiating between voluntary and involuntary exclusion, and the fourth chapter mainly focuses on credit by analyzing household debt. A number of econometric techniques such as maximum likelihood estimation, and a number of non-parametrical techniques such as unsupervised machine learning algorithms, have been used to conduct this empirical research.

VI. Appendix

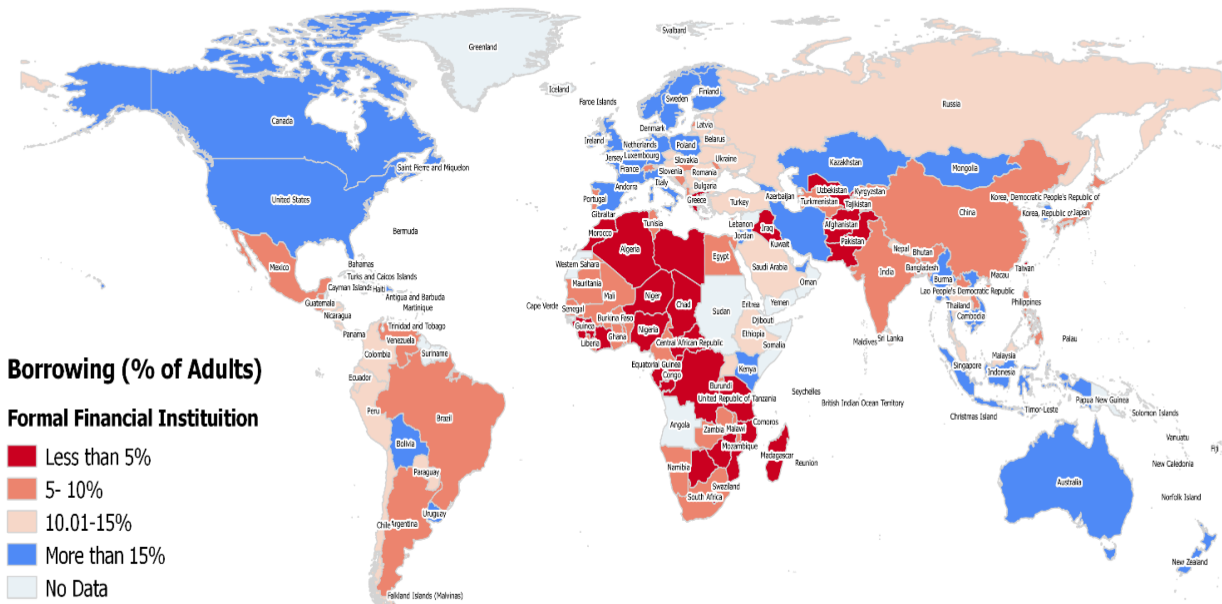
Appendix 0.1: Definitions of Financial Inclusion by International Organizations and AFI Member Countries

	Definitions of Financial Inclusion
World Bank	Financial inclusion means that individuals and businesses have access to useful and affordable financial products and services that meet their needs – transactions, payments, savings, credit and insurance – delivered in a responsible and sustainable way.
European Commission	European commission formally defined financial exclusion as a process whereby people encounter difficulties accessing and/or using financial services and products in the mainstream market that are appropriate to their needs and enable them to lead a normal social life in the society in which they belong.
The Global Partnership for Financial Inclusion (GPFI)	Financial inclusion refers to a state in which all working age adults have effective access to credit, savings, payments, and insurance from formal service providers. "Effective access" involves convenient and responsible service delivery, at a cost affordable to the customer and sustainable for the provider, with the result that financially excluded customers use formal financial services rather than existing informal options.
OECD	Financial inclusion refers to the process of promoting affordable, timely and adequate access to regulated financial products and services and broadening their use by all segments of society through the implementation of tailored existing and innovative approaches, including financial awareness and education, with a view to promote financial wellbeing as well as economic and social inclusion.
Pakistan	State Bank of Pakistan defines financial inclusion as individuals and firms can access and use a range of quality payments, savings, credit and insurance services which meet their needs with dignity and fairness.
Mexico	The National Council for Financial Inclusion (CONAIF for its acronym in Spanish) defines financial inclusion as the access and the usage of formal financial services under a proper regulation that will guarantee consumer protection and promote financial education to improve the financial capabilities of all population segments.

Source: Alliance for Financial Inclusion (AFI) [11]

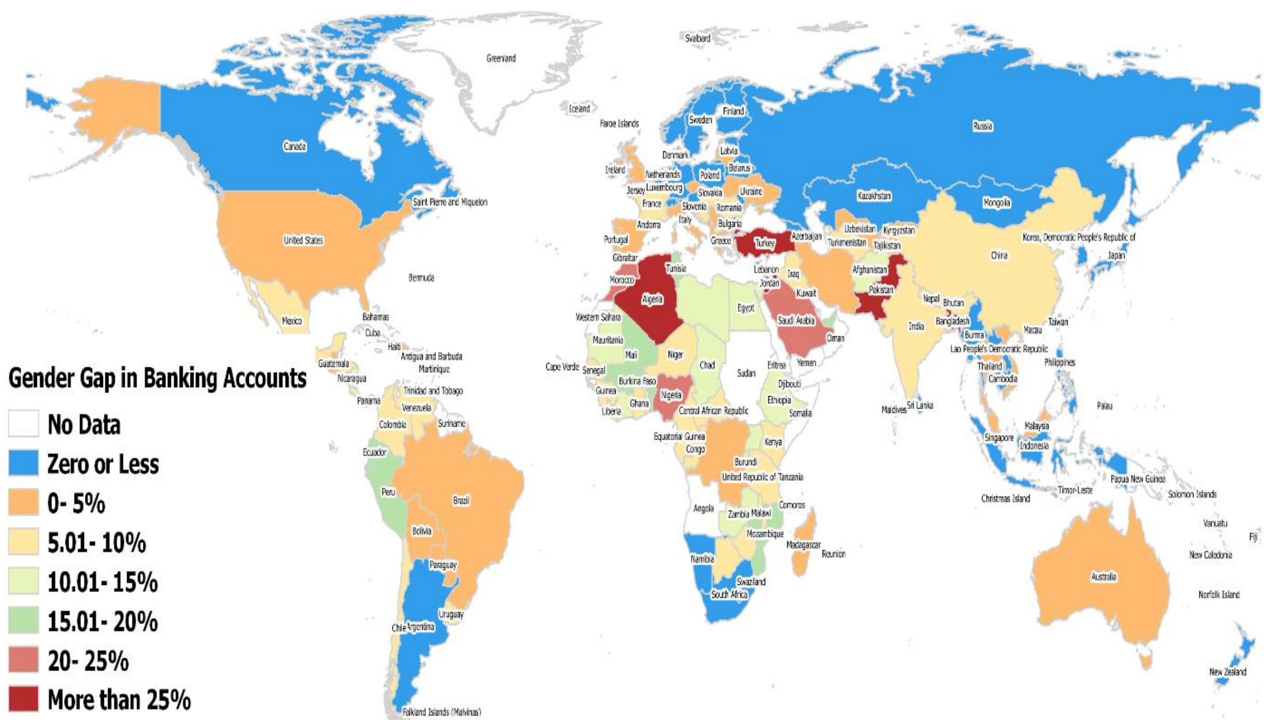
[11] These definitions come from the Alliance for Financial Inclusion (AFI). See https://www.afi-global.org/members for more details.

Appendix 0.2: Formal Borrowing Worldwide



Source: World Bank's Global Findex database, 2017

Appendix 0.3: Gender gap in ownership of bank accounts



Source: World Bank's Global Findex database, 2017

Chapter 1: Access to Microfinance and the Economic Ladder

Empirical Evidence from Pakistan

This chapter is based on a paper co-authored with Catherine BROS and Julie LOCHARD.

Abstract

A large body of the literature has focused on studying the effects of using micro-credit on various outcomes such as poverty, women's empowerment and more generally the welfare of poor households. However, few studies have examined the effect of access to microfinance in its entirety and not just as a provider of group micro-credit. This chapter uses the nationally representative pseudo-panel data from Pakistan for the years 2015 and 2016 to analyze the interface between geographical access to microfinance and entrepreneurship. Our results show that having geographical access to microfinance can allow individuals, both men and women, to move up the economic ladder by shifting to entrepreneurship and running their own business rather than working as low paid employees, farm workers, and housewives. Econometric estimates show that having access to microfinance increases the chances of being an entrepreneur by about 4 percentage points, on average. However, we find that this impact of access to microfinance on entrepreneurship runs only through the poverty channel i.e. access to microfinance is more effective in the poorer districts. Moreover, the multinomial logit regression results show that having geographical access to microfinance does not increase the likelihood of becoming an entrepreneur for people who are unemployed. The results also point towards the existence of a large gender disparity in the status of entrepreneurship. Lastly, the significance of the relationship between access to microfinance and entrepreneurship is supported by the Instrumental Variable (IV) estimation.

1.1. Introduction

The effects of micro-credit on the welfare of poor households have been extensively analyzed. Although studies have reached heterogeneous conclusions, many have brought forward the role played by micro-credit in first, loosening the financial constraint- thereby, lifting the economic prospects of the poor, and second, in buffering the impact of negative shocks. However, credit may not be the only way to either boost economic activities or absorb shocks. Other financial products such as savings accounts could also provide such services provided their value gets protected. For these reasons, the emphasis has started shifting away from the role of micro-credit to that of financial inclusion which is defined as the access to and/or usage of formal financial services by households and firms. Recent studies have started inquiring about the role of microfinance in its entirety, and not just micro-credit, in improving household's welfare in the developing world. This study contributes to this new strand of the literature by studying the relationship between access to microfinance and entrepreneurship in detail by using nationally representative individual-level data from Pakistan.

The growing importance given to financial inclusion as a tool for development is demonstrated by its inclusion in the list of key enablers for the 2030 sustainable development goals (SDGs)[12], or by the G20 featuring prominently a financial inclusion action plan in their strategy. In 2013, the World Bank postulated the 'Universal Financial Access Goal', which strives to achieve a world where everyone will have access to formal financial services by 2020. As a result of the increased focus on financial inclusion, 62% of the adult population in the world in 2014 had a bank account as compared to 51% in 2011, but around half of the adult population in the poorest households still remains financially excluded (Demirguc-Kunt et al., 2017).

This strong emphasis on financial inclusion is originally rooted in macro analyses that have established the role of financial development in either economic stability (Han and Melecky, 2013) or better governance through increased transparency (Demirguc-Kunt et al., 2017) or reduction in income inequalities (Beck at al., 2007), all factors resulting in steadier growth. The empirical

[12] See for instance Klapper et al., (2016).

literature at the macroeconomic level is quite conclusive and argues that broader access to financial services or financial development is not only positively but also causally related to economic growth (Rajan and Zingales, 1998; Claessens and Laeven, 2003; Clarke, Xu and Zou, 2006; Pasali, 2013; Samargandia et al., 2015).

From a micro-perspective the contribution of access to finance to sustained economic activity can run through various channels. It could increase the number of realized business opportunities, as potential entrepreneurs may venture into new business (Klapper et al., 2006) or actual entrepreneurs may decide to expand their own (Demirguç-Kunt and Levine, 2009). Having access to finance can also help businesses to survive economic shocks and emergencies (Claessens and Laeven, 2003). Banerjee and Newman (1993) and Galor and Zeira (1993) also argue based on a theoretical model that limited access to financial services restrict the number of potential entrepreneurs, while financial exclusion hinders investment in business growth and create poverty traps.

Microfinance institutions have played a key role in the increased amount of financial inclusion in developing countries where weak institutions along with information frictions in the credit markets limit access to financial services for a large segment of the population (Morduch, 2009; Brown, Guin and Kirschemmann, 2016). During the last two decades, microfinance has expanded rapidly in developing countries and the total microfinance clientele has grown by more than 16 times from 8 million in 1997 to 139 million in 2017[13]. However, analyses based on randomized control trials (RCTs) studying the link between microcredit and entrepreneurship have produced confounding views.

On one hand, Attanasio et al. (2015) found a positive effect of access to group loans on female entrepreneurship in Mongolia, and Augsburg et al. (2015) also found a positive effect of micro-credit on business creation in Bosnia and Herzegovina. On the other hand, Banerjee et al. (2015) and Tarozzi et al. (2015) find that micro-credit does not have a significant effect on business creation in case of India and Ethiopia. Similarly, Duflo (2010) argues based on the work conducted

[13] These statistics have been taken from the Microfinance Barometer 2018.

with an Indian micro-credit institution in Hyderabad that only a fraction of those who are not entrepreneurs have the will to become one. Thus, extending the access to micro-credit cannot be expected to turn a whole nation into self-employed business men (Duflo, 2010).

Despite the many advantages of RCTs, there are several concerns which have been highlighted in the literature with respect to the use of RCTs to answer broad questions in development such as the impact of microfinance (Deaton, 2009; Pritchett and Sandefut, 2015; Bédécarrats et al., 2017). RCTs in the field of microfinance are typically limited to studying short term and average treatment effects from one-time intervention, focusing usually on group credit provided by one specific microfinance provider[14]. Rodrik (2008) argues that one major shortcoming of the RCTs is that they might lack external validity and it is problematic to apply the findings of RCTs conducted in a particular setting to other settings. Bédécarrats et al. (2019) replicated a flagship RCT conducted in rural Morocco on the impact of microfinance by Crépon et al. (2015) and argued that RCTs may not be suitable to gauge the impact of microfinance and that resulting policy recommendations should be established with extreme diligence due to their limited scope.

It is imperative to note that becoming an entrepreneur could be a constrained choice for many people, and an increase in the proportion of entrepreneurs is not automatically welfare improving. This paper contributes to this debate by emphasizing that, although not everyone wants to become an entrepreneur, having geographical access to microfinance does aid in moving from the least earning occupation to a more income generating entrepreneurship status in the context of Pakistan. Our results show that living less than 10 kilometers away from a microfinance institution's branch, our measure of access to microfinance here[15], does increase the probability to become an entrepreneur, even after controlling for local and individual characteristics and controlling for potential endogeneity issue. However, this positive effect of access to microfinance on the likelihood to become an entrepreneur is conditional on the district's level of poverty. The higher the district's poverty rate, the more effective access to microfinance is on the probability of being

[14] By design, RCTs on the topic of microfinance are highly specific to the population, type of microfinance institution, time frame, and the microcredit product or intervention under study.

[15] As will be discussed later, this result is robust to alternative definitions of access to microfinance.

an entrepreneur. Moreover, being financially included does raise the likelihood to move from a low earning position such as being a salaried employee, farm worker, or even a housewife to a more profitable entrepreneurship status, as the proportion of poor individuals in the latter position is the lowest of the three. We conclude that access to microfinance helps climbing the economic opportunities ladder.

It should be emphasized that most studies on the link between microfinance and entrepreneurship have looked only at the use of micro-credit through group loans while we take a broader view by looking at financial inclusion through geographical access to microfinance. While micro-credit is expected to boost entrepreneurship mainly through affordable interest rates, our results show that the mere access to a formal source of credit as well as other financial services such as savings does increase entrepreneurship. This results from the fact that leaving the bulk of the market to informal money lenders is likely to result in financial exclusion as clients are locked into a specific relationship with few informal credit providers. This is best explained by the account given by Aleem (1990) of the Pakistani credit market. The author shows that informal money lenders in Pakistan incur large expenses in trying to alleviate information asymmetry about the risk profiles of the loan applicant in a context where credit histories are not documented and pooled. Thus, once these costs have been incurred for the first loan, the money lender will retain their clients and not compete against each other. This has two consequences. First, borrowers are thus locked into a relationship with a specific money lender and suspicion about the risk profile of new applicants will be high. Thus, a large portion of applications will not be considered. Second, due to the absence of competition, interest charged by these moneylenders become outrageously high. Both factors contribute to the exclusion of a large section of the population who ends up being financially excluded and thus constrained in its aspirations to set up new businesses. Our results therefore emphasize that geographical access to formal financial services through microfinance institutions matters as it restores access to the credit market and other financial services for potential entrepreneurs.

This study provides a new approach in three ways. First, it relies on nationally representative micro-data, while most studies on financial inclusion have predominantly taken a macro angle. It allows a better identification of households that would benefit from an enhanced access to microfinance. Second, it sheds some light on Pakistan that is one of the fastest growing, yet understudied market, for microfinance services. Finally, this study focuses on availability of financial services rather than usage and shows that the accessibility of institutions can make a difference. This translates into policy recommendations in terms of coverage, especially in a country such as Pakistan, where contrasts between provinces can be stark.

The rest of the chapter is structured as follows: Section 1.2 describes the data used, Section 1.3 sets out the empirical methodology, Section 1.4 presents and discusses the results. Section 1.5 concludes.

1.2. Details about the Data

The data that we use for our analysis comes from the Financial Inclusion Insight (FII) Program. The survey is conducted by Gallup International Association under the supervision of Bill and Melinda Gates Foundation and InterMedia USA. This is a nationally representative survey of adults[16]. We rely on two waves of the survey from 2015 and 2016 to conduct this analysis. The dataset is a pseudo panel i.e. respondents in both waves are different. The overall sample is proportionally to the size of population with rural and urban strata and comprised of 12,000 observations. For the purpose of this research, we exclude students and retired people from the overall sample since they are not looking for jobs and cannot be entrepreneurs. We are left with a total of 11,055 observations including both waves of the survey.

We define the term entrepreneur as someone who owns or co-owns a business. This business can operate either with the entrepreneur as the only employee or with multiple employees working for the entrepreneur. We utilize the section that provides information on basic demographics and employment status to get our main variable of entrepreneurship. The survey asks respondents to list their current status of employment and their profession. About 15% of the adults in our sample classify as entrepreneurs. These entrepreneurs mainly run micro businesses. According to our data, more than 80% of these businesses operate without any additional employees and the average number of employees in the business are 0.56. Top five professions of entrepreneurs come out to be shop owners, mechanics or electricians, tailors, beauticians, and street vendors. For the purpose

[16] Aged more than 15 years. One household member is surveyed per household.

of this research, we only consider off-farm entrepreneurship i.e. people working on the agricultural farms are not considered as entrepreneurs and are treated as a separate occupational category[17].

We use the section on access and usage of financial services to obtain the indicator of access to microfinance. The survey asks respondents whether a microfinance institution is available in a radius of 5, 10, 15 kilometers, or more, away from the respondent's home. For the purpose of this study, an individual or household is considered to have access to a microfinance institution if one is available within ten kilometers. In order to gain confidence in the robustness of our results, we replicate the estimations using the 5 and 15 kilometers ranges[18].

We also use data at the district level such as the Multidimensional Poverty Index (MPI), developed by the United Nations Development Programme (UNDP) using the Alkire- Foster methodology. This indicator considers three dimensions while measuring the district poverty rate: education, health and living standards. A score close to zero highlights that there are no poor people in that district and a score close to 100 highlights extreme poverty. We also use data from the Pakistan's Census of 2017 as far as district level population variables are concerned.

1.2.1. Descriptive Statistics

Access to microfinance is measured by whether the individual has access to a microfinance institution within a 10 kilometers radius or not. By this metric, 41% of the sample has access to a microfinance institution. Regarding the employment status, about 2% of the total individuals in the sample reported to be unemployed. 15% reported being entrepreneurs but there is a big gender gap in the status of entrepreneurship. Only 2% of the total women in our sample reported to be entrepreneurs as compared to 28% men. Furthermore, a majority of the men reported to be either

[17] Agricultural entrepreneurship is quite different from other types of entrepreneurship in terms of business dynamics. In line with the existing literature (Nagler and Naudé, 2017; Gurley-Calvez, 2009; among others), we differentiate between the two and focus on non-agricultural entrepreneurship.

[18] It is important to emphasize that certain microfinance institutions which are regulated by the State Bank of Pakistan are not legally allowed to provide financially services to households outside the 25 kilometers radius due to reasons pertaining to consumer protection mainly. See Prudential Regulations for Microfinance Institutions 2014 by the State Bank of Pakistan for more details.

working as salaried employees (42%) or working in the agriculture sector (26%), whereas, majority of the women reported being housewives.

Regarding other socio-demographic indicators, 68% of the individuals in our sample reported living in rural areas while others live in urban areas. 49% of the individuals in our sample are female. 35% of the total individuals reported having no formal education. Amongst men, about 28% reported having no formal education, and amongst women, 43% reported having no formal education. The average age of individuals in our sample is about 35.3 years and the average household size is 7.7 individuals. A summary of the descriptive statistics is provided in Table 1.1.

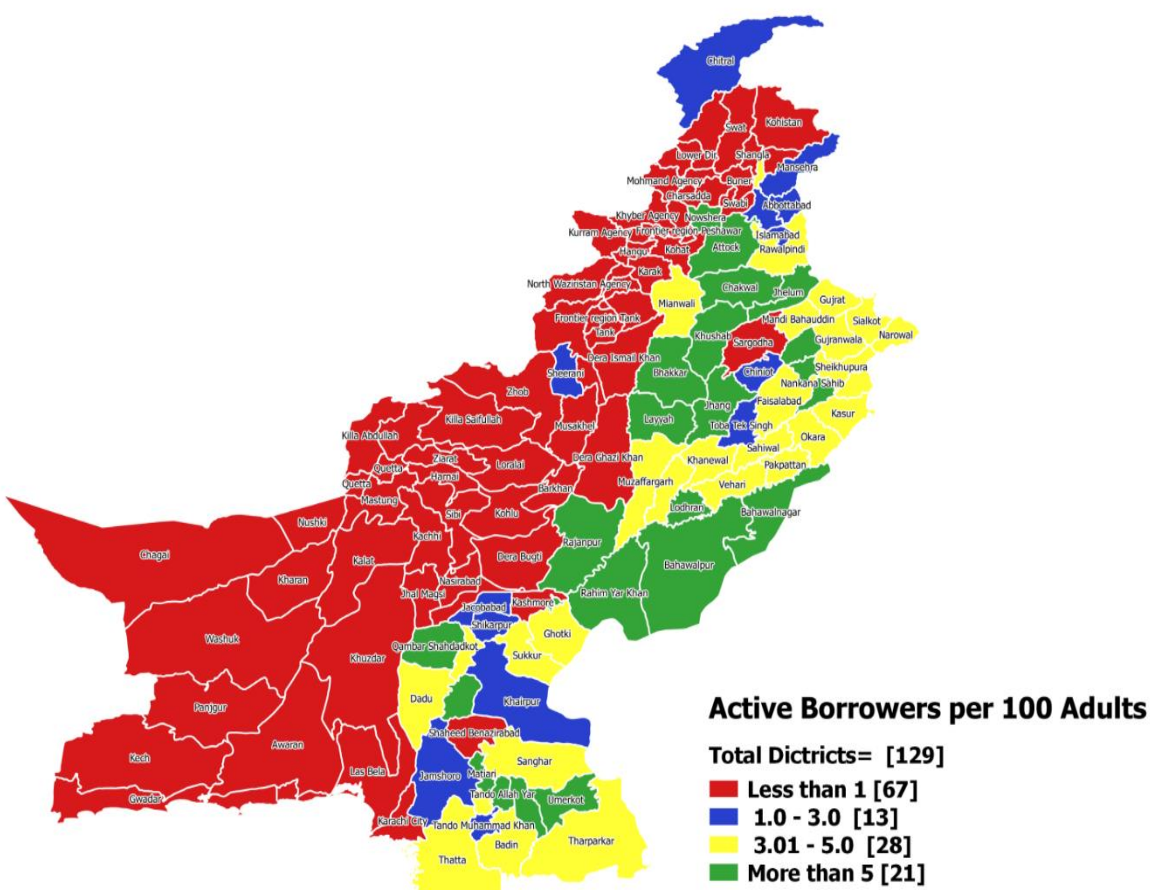
Table 1.1: Descriptive statistics

Variable Name	Description	# of Obs.	Mean	S.D.	Min	Max
Access to Microfinance Institutions (MFIs)	Dummy=1 if individual has access to a microfinance institution within 10kms, else 0	11,055	0.41	0.49	0	1
Entrepreneur Dummy	Dummy=1 if individual is an entrepreneur, else 0	11,055	0.15	0.36	0	1
Salaried Employee Dummy	Dummy=1 if individual is a salaried employee, else 0	11,055	0.23	0.44	0	1
Unemployed Dummy	Dummy=1 if individual is unemployed, else 0	11,055	0.02	0.14	0	1
Farm dummy	Dummy=1 if individual works in the agricultural sector, else 0	11,055	0.14	0.35	0	1
Housewife/ Househusband	Dummy =1 if individual is a housewife or a househusband	11,055	0.46	0.50	0	1
Female Dummy	Dummy=1 if individual is a female, else 0	11,055	0.49	0.50	0	1
Age	Age in terms of number of years	11,055	35.32	13.02	15	92
Household Size	Number of household members	11,055	7.67	3.7	1	43
Urban Dummy	Dummy=1 if individual lives in an urban area, else 0	11,055	0.34	0.47	0	1
No Formal Education Dummy	Dummy=1 if individual has no formal education, else 0	11,055	0.35	0.48	0	1
Primary Education Dummy	Dummy=1 if individual’s maximum education is till Primary, else 0	11,055	0.23	0.42	0	1
Secondary Education Dummy	Dummy=1 if individual’s maximum education is till Secondary, else 0	11,055	0.32	0.47	0	1
Tertiary Education Dummy	Dummy=1 if individual’s maximum education is tertiary level, else 0	11,055	0.09	0.29	0	1
Poverty Incidence Rate -at District Level	Percentage of people living in poverty in the district, provided by UNDP	10,904	0.35	0.23	0.03	0.97

Source: Financial Inclusion Insight (FII) Survey, 2015 and 2016

It is important to note that there are also large gaps in terms of access to microfinance across different regions in Pakistan i.e. across different provinces and districts. Pakistan has a total of four provinces which are further divided into more than 150 districts. We employ the data from Pakistan Microfinance Network (PMN) to provide insight about the outreach of microfinance institutions (MFIs). On average, a district in Pakistan has less than 2 microfinance branches per hundred thousand adults. Moreover, around half of the districts have one or less microfinance branches per hundred thousand adults. The outreach of microfinance providers is relatively lower in western regions (i.e. Balochistan province) and in the northern regions (KPK province) as compared to the provinces of Punjab and Sindh. Figure 1.1 depicts how the concentration of microfinance borrowers varies across different regions in Pakistan highlighting that the differences are quite prominent.

Figure 1.1: Concentration of microfinance borrowers in Pakistan



Source: Based on data from Pakistan Microfinance Network and PBS, 2016

1.3. Empirical Methodology

Entrepreneurship is the dependent variable and it represents the status of the worker (entrepreneur= 1, otherwise = 0). Given our interest in a binary dependent variable, we resort to a probit model. $E_{\boldsymbol{i,d,p,t}}$ represents the binary dependent variable that takes the value 1 if an individual i who lives at time t (either 2015 or 2016) in district d of province p declared being an entrepreneur and defined as follows, where $\boldsymbol{E^*_{i,d,p,t}}$ is the latent variable.

$$\boldsymbol{E^*_{i,d,p,t}} = \boldsymbol{\alpha} + \boldsymbol{B_1}(\boldsymbol{AccessMFI})_{\boldsymbol{i,d,p,t}} + \boldsymbol{B_2}(\boldsymbol{FemaleD})_{\boldsymbol{i,d,p,t}} + \boldsymbol{B_3}(\boldsymbol{Characteristics})_{\boldsymbol{i,d,p,t}} + \boldsymbol{B_4}(\boldsymbol{PovertyIndex})_{\boldsymbol{d,p,t}} + \eta_{\boldsymbol{p}} + \eta_{\boldsymbol{t}} + \boldsymbol{\varepsilon}_{\boldsymbol{i,d,p,t}} \qquad (\mathbf{1.1})$$

$$E_{\boldsymbol{i,d,p},t} = 1 \text{ if } \boldsymbol{E^*_{i,d,p,t}} > 0$$
$$E_{\boldsymbol{i,d,p},t} = 0 \text{ if } \boldsymbol{E^*_{i,d,p,t}} \leq 0$$

$AccessMFI_{i,d,p,t}$ is a binary variable equal to one if individual i has access to a Microfinance Institution within ten kilometers. $Female_{i,d,p,t}$ is a binary variable capturing potential gender gaps in entrepreneurship. $Characteristics_{i,d,p,t}$ is a set of individual level control variables such as education, age, marital status, location, and size of the household. $Poverty_{d,p,t}$ is the district's level of poverty measured as MPI defined by the UNDP. The term $\eta_{\boldsymbol{p}}$ refers to fixed effects at the province level which we include to control for provincial level heterogeneity which might affect chances of entrepreneurship. In additional regressions, we account for district unobservable characteristics by adding fixed effects at the district level ($\eta_{\boldsymbol{d}}$) instead of fixed effects at the province level. Furthermore, the term $\eta_{\boldsymbol{t}}$ denotes time fixed effects and $\boldsymbol{\varepsilon}_{\boldsymbol{i,d,p},t}$ is the error term that follows a standard normal cumulative distribution function We use sampling weights in all our regression results to compensate for unequal probabilities of selection. Lastly, we cluster the standard errors at the district level for all the regressions in order to control for possible error correlation within districts.

The existing literature shows that various determinants of entrepreneurship mainly include three categories of factors i.e. sociological, institutional, and demographic factors (Koellinger et al., 2013; Dohmen et al., 2011). Characteristics such as education level, sex, locality (rural or urban residence), ethnicity, age, marital status etc. as well as financial, political and legal institutions have been found to influence entrepreneurial behavior and we control for them in our econometric estimates. Apart from these factors, some might argue that entrepreneurs are different from the rest based on their attitudes, risk taking abilities, family background and connections, or other psychological traits. However, due to lack of data, these attributes are not discussed in our empirical analysis, but we do acknowledge that they might induce a bias in our results. Nevertheless, we run a number of tests to ensure the robustness of our results.

It is important to emphasize that the location decisions of financial institutions are driven by local aggregate variables such as population density and socioeconomic, political and commercial activity (Ruiz-Tagle and Vella, 2015). Geographical access to microfinance institutions does not reflect any aspect of individual level demand for a specific type of work. Therefore, we argue that the variable that we use to measure the access to microfinance is exogenous to the individual's decision to be an entrepreneur and does not propagate the reverse causality issue with the entrepreneurship status of the individual. Nevertheless, as a robustness check, we also account for endogeneity related to non-random placement of microfinance branches using a bivariate probit regression model (see Section 1.4.3).

It is also essential to explore the different channels through which this relationship between access to microfinance and entrepreneurship works across different employment statuses. In order to do that, we rely on a multinomial logit regression instead of a simple probit. The dependent variable in this case is $\boldsymbol{Y_{i,d,p,t}}$ i.e. the employment status for individual *i*, in district *d*, of province *p*, in year *t*. It represents the different categories of occupational status and the individual is either an entrepreneur (reference category), salaried employee in the non-agricultural sector (2), unemployed (3), farm owner (4), farm worker (5) or a housewife (6). The regression equation takes the following form where $\mathbf{Y}^*_{i,d,p,t}$ is the latent variable and the error term $\boldsymbol{\varepsilon_{i,d,p,t}}$ follows a logistic distribution. The control variables stay the same as in the case of equation 1.1.

$$\mathbf{Y}^*_{i,d,p,t} = \boldsymbol{\alpha} + \boldsymbol{C_1}(\boldsymbol{AccessMFI})_{\boldsymbol{i,d,p,t}} + \boldsymbol{C_2}(\boldsymbol{FemaleD})_{\boldsymbol{i,d,p,t}} + \boldsymbol{C_3}(\boldsymbol{Characteristics})_{\boldsymbol{i,d,p,t}} + \boldsymbol{C_4}(\boldsymbol{PovertyIndex})_{\boldsymbol{d,p,t}} + \eta_{\boldsymbol{p}} + \eta_{\boldsymbol{t}} + \boldsymbol{\varepsilon_{i,d,p,t}} \qquad (\mathbf{1.2})$$

$$\begin{aligned}
\boldsymbol{Y_{i,d,p,t}} &= 1 \quad \text{if} \quad \mathbf{Y}^*_{i,d,p,t} \leq \mu1 \\
\boldsymbol{Y_{i,d,p,t}} &= 2 \quad \text{if} \quad \mu1 < \mathbf{Y}^*_{i,d,p,t} \leq \mu2 \\
\boldsymbol{Y_{i,d,p,t}} &= 3 \quad \text{if} \quad \mu2 < \mathbf{Y}^*_{i,d,p,t} \leq \mu3 \\
\boldsymbol{Y_{i,d,p,t}} &= 4 \quad \text{if} \quad \mu3 < \mathbf{Y}^*_{i,d,p,t} \leq \mu4 \\
\boldsymbol{Y_{i,d,p,t}} &= 5 \quad \text{if} \quad \mu4 < \mathbf{Y}^*_{i,d,p,t} \leq \mu5 \\
\boldsymbol{Y_{i,d,p,t}} &= 6 \quad \text{if} \quad \mathbf{Y}^*_{i,d,p,t} > \mu5
\end{aligned}$$

1.4. Results and Discussion

1.4.1. Baseline Results

This section presents the results of the econometric analysis which aim at determining the relationship between access to microfinance and entrepreneurship, and investigating the key determinants of entrepreneurship. Table 1.2 reports the results of our baseline econometric estimates. In case of a probit model, the coefficients cannot be directly interpreted. So, in order to comment on the magnitude of the coefficients, we have calculated and reported the average marginal effects.

The empirical results show that the relationship between access to microfinance and entrepreneurship is positive and statistically significant i.e. having geographical access to microfinance institutions increases the probability of an individual being an entrepreneur. In the first column of Table 1.2, we estimate the econometric model with limited number of controls. We only include individual level controls, fixed effects at the province level, and time fixed effects. The results show a positive link between access to microfinance and entrepreneurship which comes out to be statistically significant at the 5% level. However, there is a lot of variation in terms of development level and infrastructure across districts in Pakistan which can influence entrepreneurial behavior as well as availability of microfinance institutions. Therefore, in column two, we control for the difference in urban and rural locations by including a dummy for urban location, and we also control for the development level at the district level by including poverty rate as a proxy for the level of development. The results highlight that the positive relationship between access to microfinance and entrepreneurship still holds. Having access to a microfinance institution increases the probability of being an entrepreneur by about 4 percentage points, and the result is statistically significant at a 1% level. This effect is quite meaningful in terms of magnitude given that only 15% of the adult population is classified as entrepreneurs. Moreover, the district level poverty rate variable comes out to be non-significant indicating that the poverty level of the district does not affect entrepreneurship creation once we control for province fixed effects, the urban and rural location and other individual characteristics.

Table 1.2: Econometric estimates on entrepreneurship status

Variables (Probit Model)	**(1) Entrepreneurship**	**(2) Entrepreneurship**	**(3) Entrepreneurship**
Access to Microfinance Institutions (MFIs)	0.035**	0.04***	0.032**
	(0.01)	(0.01)	(0.01)
Education Level (Base: No formal Education)			
Primary Education	0.05***	0.05***	0.05***
	(0.004)	(0.01)	(0.01)
Secondary Education	0.07***	0.06***	0.06***
	(0.01)	(0.01)	(0.01)
Higher Education	-0.004	-0.01	-0.02
	(0.02)	(0.01)	(0.01)
Age (Base: 15-24)			
25-34	0.01	0.01	0.01
	(0.01)	(0.01)	(0.01)
35-44	0.04**	0.03**	0.03**
	(0.01)	(0.01)	(0.01)
45-54	0.01	-0.001	0.003
Over 55	(0.02) -0.01	(0.01) -0.02	(0.01) -0.01
	(0.03)	(0.02)	(0.02)
Number of household members	-0.001	0.0001	0.001
	(0.0005)	(0.001)	(0.001)
Marital Status (Base: Single)			
Married	0.025*	0.03***	0.03**
	(0.01)	(0.01)	(0.01)
Divorced	0.26***	0.27***	0.21***
	(0.06)	(0.08)	(0.07)
Widowed	0.05	0.05	0.03
	(0.06)	(0.05)	(0.05)
Female Dummy	-0.29***	-0.29***	-0.30***
	(0.02)	(0.01)	(0.01)
Urban Dummy	-	0.03**	0.045***
		(0.01)	(0.01)
MPI: Poverty Rate	-	-0.05	-
		(0.03)	
Year Fixed Effects	**Yes**	**Yes**	**Yes**
Province Fixed Effects	**Yes**	**Yes**	**No**
District Fixed Effects	**No**	**No**	**Yes**
Total Observations	**11,041**	**11,041**	**10,811**
Pseudo R2	**0.2248**	**0.2308**	**0.2760**

*Note: Results of the probit regression have been reported with average marginal effects. Standard errors are in parentheses (clustered at the district level). The dependent variable is entrepreneurship status dummy (see equation 1.1). ***,** and * represent significance at 1%, 5% and 10%, respectively.*

One can also argue that there are some factors other than the poverty level at the district level which have an effect on entrepreneurial behavior e.g. ease of doing business, crime rate, or regulations amongst other factors. So, in the last column of Table 1.2, we carry out the estimation by including district level fixed effects to control for district level unobserved characteristics[19]. The results show that having access to microfinance institutions still increases the chances of being an entrepreneur by about 3.2 percentage points and the variable is statistically significant at a 5% confidence level.

The results can be explained by the fact that having access to formal financial services through microfinance institutions can reduce the start-up costs for potential entrepreneurs who either do not have access to external finance and self-finance or who are reluctant to use the expensive informal sources of finance from moneylenders and suppliers. Moreover, having access to microfinance institutions does not merely allow access to credit, but it can also provide access to other financial services such as savings, insurance, and money transfer services in most cases. This can enable existing entrepreneurs to better manage their business and handle economic shocks so that they do not have to shut down their business and take a job. This phenomenon is highlighted in the existing literature where without having access to financial services, potential entrepreneurs are trapped and forced to take a job rather than creating one themselves (Banerjee and Newman, 1993).

Education is another factor that comes out as a statistically significant determinant of entrepreneurship. Having primary or secondary level of education, as compared to no formal education, increases an individual’s chances of being an entrepreneur. However, having a higher level of education (university education or post graduate studies) does not come out to be statistically significant. A possible explanation for this can be the fact that individuals with a higher education level are more likely to have better job opportunities in the job market.

The results also highlight that women are less likely to be entrepreneurs as compared to men. The marginal effects calculated at means indicate that the probability of a woman being an entrepreneur

[19] In this case, we drop the poverty rate variable because of perfect collinearity with the district fixed effects.

is about 23 percentage points less as compared to a man. From a theoretical perspective, this corroborates with some of the existing literature. A few reasons might explain this gender differentiation in entrepreneurship. First, women are deemed more risk averse as compared to men (Dohmen et al., 2011) and thus less likely to choose entrepreneurship as a career. Second, women are considered to have a lower or less diversified level of social capital in terms of networks and business contacts as compared to men (Koellinger et al., 2013) which is likely to put them at a disadvantage when it comes to pursuing entrepreneurship as a career. A higher need for a stronger work-life balance amongst women (Gurley-Calvez et al., 2009) could be another reason for this gender differentiation.

The econometric estimates also show that residing in an urban area increases the chances of becoming an entrepreneur by 4.5 percentage points as compared to living in a rural area (column 3). Moreover, being married or divorced increases the probability of becoming an entrepreneur as compared to being unmarried. We do not find a statistically significant relationship of household size on entrepreneurship.

Another avenue through which an increased access to financial services can help entrepreneurship is by allowing them to further expand their businesses. However, in this research, we only assess the effect of access to microfinance on the formation of entrepreneurship and not the effect of microfinance on profits and growth of existing entrepreneurs.

Next, in order to explore the different channels through which this relationship between access to microfinance and entrepreneurship works across different employment statuses, we rely on a multinomial logit regression as shown in equation 1.2. The dependent variable in this case is the employment status and it is categorical. We use entrepreneurship as the reference category and compare the results for working as a salaried employee in non-agricultural sector (1), being unemployed (2), being a farm owner (3), being a farm worker (4) and being a housewife (5) with being an entrepreneur (reference category).

The results reported in Table 1.3 highlight that access to microfinance increases the chances of becoming an entrepreneur as compared to being a salaried employee in non-agricultural field (column 1). Since about 75% of the non-agricultural labor force works in the informal sector where wages are not competitive[20], this result might suggest that having access to microfinance can allow individuals to move up the economic ladder by moving away from lower paid salaried jobs and operating their own business. We further discuss this in subsequent sections.

Second, we look at unemployed individuals to see if having access to microfinance plays a role in employment generation by allowing unemployed people to start their own businesses. The econometric estimate shows that having access to microfinance does not have a significant impact on the chances of becoming an entrepreneur in this case (column 2). However, this may be due to the fact that only 2% of the individuals in the sample are unemployed.

Third, results also highlight that access to microfinance increases the chances of becoming an entrepreneur as compared to working on a farm as an employee (column 4). However, the relationship is not statistically significant for people in the agricultural sector who are farm owners (column 3). Lastly, the results also show that access to microfinance increases the probability of becoming an entrepreneur as compared to being a housewife (column 5). This shows that women who are housewives and when provided with access to financial services by the microfinance institution can move towards entrepreneurship and become a part of the labor force. This result is anticipated because some of the top microfinance institutions in Pakistan provide financial services only to women, especially in regions where they do not have other employment opportunities. Consequently, women start a micro business, mostly at their own house, which allows them to generate some additional income for the household.

These results highlight that access to microfinance can allow individuals, men and women, to move up the economic ladder by shifting to entrepreneurship and starting their own business rather than working as low paid employees, farm workers, and housewives.However, this can be corroborated by looking at the average score for poverty probability index (PPI) provided in the survey. The PPI index[21] ranks every individual on a poverty scale between 0-100. A higher PPI score reflects a better economic status. Figure 1.2 lists the average poverty probability index scores

[20] Labor force statistics from Pakistan Bureau of Statistics (2015).

[21] Poverty probability index (PPI) is calculated using the Grameen methodology. See https://www.povertyindex.org/ for further details. The indicator was initially called Progress out of Poverty Index.

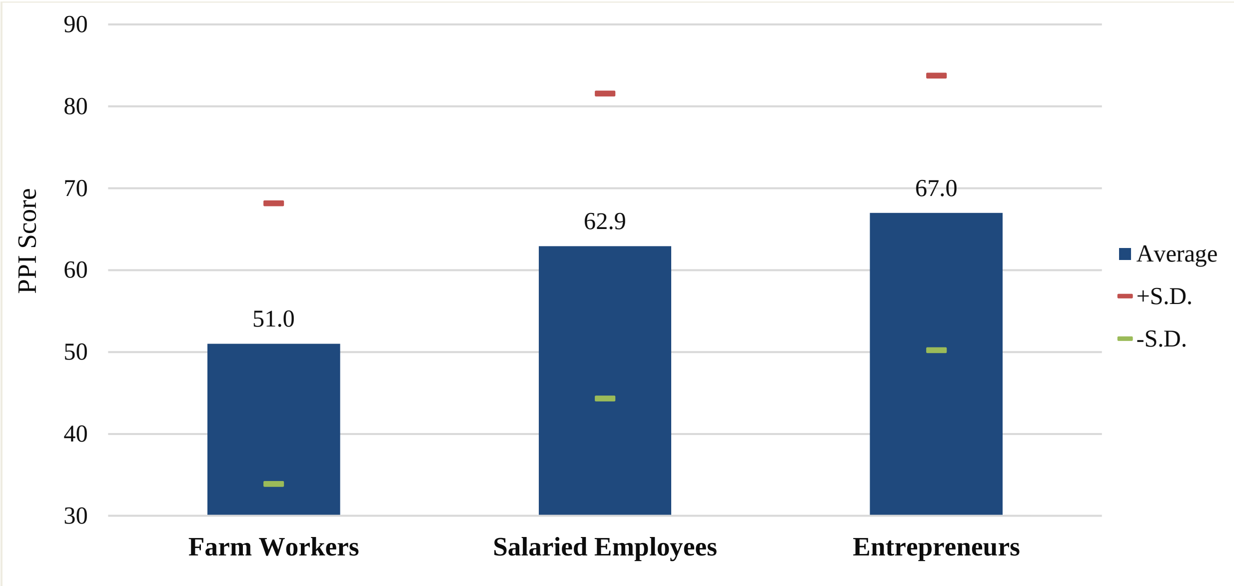
across different employment statuses and shows that entrepreneurs have a better economic status than salaried employees and farm workers, on average. The bars in red and blue highlight one standard deviation above the mean and one standard deviation below the mean respectively.

Table 1.3: Econometric estimates across employment statuses

	(1)	(2)	(3)	(4)	(5)
Variables (Multinomial Logit Model, Reference: Entrepreneurs)	**Salaried Employees (Non-agricultural)**	**Unemployed**	**Farm Owners**	**Farm Workers**	**Stay home housewives**
Access to Microfinance Institutions	-0.02**	0.01	0.00	-0.02***	-0.01**
Age (Base: 15-24)					
25-34	0.01	-0.02***	-0.02*	0.02	0.00
35-44	0.00	-0.02***	-0.02**	0.01	-0.01**
45-54	0.01	-0.02**	0.00	0.02	-0.01*
Over 55	-0.03	0.00	0.03**	0.01	0.00
Education Level (Base: No formal Education)					
Primary Education	-0.01***	-0.01**	-0.01**	-0.01***	-0.01***
Secondary Education	0.02**	-0.01**	-0.01***	-0.05***	-0.01***
Higher Education	0.11***	0.01	0.01	-0.08***	-0.04***
Marital Status (Base: Divorced/ Widowed)					
Single	0.04	0.00	-0.06	0.02	0.02*
Married	0.03	-0.03**	-0.06**	0.01	0.03**
Number of household members	0.00	0.00	0.00***	0.00	0.00
Female Dummy	-0.24	-0.03	-0.20***	-0.08**	-
Urban Dummy	0.12*	0.00**	-0.11***	-0.07***	0.02
MPI	-0.12	-0.07**	0.25***	0.11**	-0.08
Year Fixed Effects	**Yes**				
Province Fixed Effects	**Yes**				
Total Observations	**11,036**				
Pseudo R2	**0.4929**				

*Note: Results of the multinomial logit regression have been reported with average marginal effects. Standard errors are clustered at the district level. The dependent variable is employment status and the reference category is entrepreneurs (see equation 1.2). ***, ** and * represent significance at 1%, 5% and 10%, respectively.*

Figure 1.2: Average poverty probability index (PPI)



Source: Financial Inclusion Insight Survey (FII), 2015 and 2016

To further corroborate this result, we analyze this relationship between access to microfinance and entrepreneurship across different quintiles of the poverty probability index (PPI). The first quintile includes individuals that belong to the poorest households, whereas the fifth quintile includes individuals who belong to the richest households in Pakistan. The regression results have been summarized in Table 1.4.

These results indicate that individuals belonging to low income households, especially the 1st quintile, are more likely to become entrepreneurs, when compared to other employment types, once provided with access to microfinance institutions. The positive relationship between access to microfinance institutions and entrepreneurship comes out to be significant, at least at a 5% level, for the first three quintiles. This result indicates that having access to microfinance increases the probability to be an entrepreneur for individuals who are living below the poverty line or close to it[22].

[22] The third quintile also includes some households that are living below the poverty line according to the Poverty Probability Index (PPI) classification.

Table 1.4: Econometric estimates across poverty quintiles

VARIABLES	**Entrepreneurship**				
	1st Quintile (Poorest)	**2nd Quintile**	**3rd Quintile**	**4th Quintile**	**5th Quintile (Richest)**
Access to Microfinance Institutions (MFIs)	0.057***	0.038**	0.050**	0.043*	0.023
	(0.017)	(0.019)	(0.023)	(0.022)	(0.033)
Female Dummy	-0.183***	-0.309***	-0.296***	-0.354***	-0.342***
	(0.029)	(0.027)	(0.018)	(0.024)	(0.023)
Age (Base: 15-24)					
25-34	0.009	-0.020	0.007	0.045*	0.012
	(0.019)	(0.027)	(0.023)	(0.024)	(0.029)
35-44	0.052**	0.000	0.037	0.058**	0.011
	(0.024)	(0.024)	(0.028)	(0.030)	(0.036)
45-54	0.060**	-0.021	-0.003	-0.011	-0.035
	(0.024)	(0.028)	(0.032)	(0.027)	(0.034)
Over 55	0.049	-0.058*	-0.043	-0.033	-0.015
	(0.030)	(0.031)	(0.034)	(0.034)	(0.037)
Marital Status (Base: Single)					
Married	0.035*	0.039*	0.033	0.036	0.024
	(0.018)	(0.023)	(0.023)	(0.025)	(0.035)
Divorced	0.466**	0.100	0.255*	0.447***	0.138
	(0.205)	(0.123)	(0.134)	(0.167)	(0.129)
Widowed	0.189**		0.032	0.123	
	(0.083)		(0.077)	(0.147)	
Education Level (Base: No formal Education)					
Primary Education	0.055***	0.043**	0.059***	0.017	-0.008
	(0.016)	(0.018)	(0.018)	(0.021)	(0.029)
Secondary Education	0.083***	0.065***	0.051***	0.036	-0.040
	(0.019)	(0.020)	(0.019)	(0.022)	(0.029)
Higher Education	0.027	-0.029	-0.019	-0.045*	-0.117***
	(0.026)	(0.028)	(0.028)	(0.025)	(0.027)
Number of household members	0.002	0.003	0.002	-0.003	0.006
	(0.001)	(0.002)	(0.002)	(0.003)	(0.004)
MPI: Poverty Rate	0.050	-0.082*	-0.021	-0.123**	-0.027
	(0.043)	(0.046)	(0.046)	(0.052)	(0.088)
Urban Dummy	0.037	0.036**	-0.003	0.003	0.057***
	(0.024)	(0.018)	(0.022)	(0.016)	(0.022)
Provincial Fixed Effects	**Yes**				
Year Fixed Effects	**Yes**				
Observations	**10,955**				
Wald (Prob > chi2)	**0.00**				

*Note: Results of the probit regression with quintiles of the poverty probability index (PPI) have been reported with average marginal effects. Standard errors are in parentheses (clustered at the district level). The dependent variable is entrepreneurship status dummy. ***, ** and * represent significance at 1%, 5% and 10%, respectively.*

1.4.2. Instrumental Variable Estimation

We believe that the endogeneity arising from the reverse causality issue between access to finance and entrepreneurship is not probable because location decisions of financial institutions are driven by local aggregates and not individual level demand of starting a business. However, we acknowledge that the endogeneity issue due to non-random placement of microfinance branches could potentially bias our estimations. Therefore, in order to tackle this issue and test whether the relationship between access to microfinance and entrepreneurship is causal, we apply an instrument variable (IV) approach. For that purpose, we use bivariate probit regression model which is estimated using the maximum likelihood estimation. This method is equivalent to the linear two stage least squares (2SLS) approach and it is used when both the dependent and the endogenous independent variable are binary (Lollivier, 2006). Extending 2SLS to non-linear models such as this leads to inconsistent estimates and the literature refers to it as the "forbidden regression" (Wooldridge, 2000). Hence, a new specification (1.3) is implemented. It takes the form of a bivariate probit model using the main specification (1.1) as the main model and simultaneously explains access to microfinance institutions by employing an instrument variable.

$$\begin{cases} \boldsymbol{y^*_{i,d,p,t} = \alpha + D_1(AccessMFI)_{i,d,p,t} + D_2(characteristics)_{i,d,p,t}} + \eta_{\boldsymbol{p}} + \eta_{\boldsymbol{t}} + \boldsymbol{\varepsilon_{i,d,p,t}} \\ \boldsymbol{AccessMFI^*_{i,d,p,t} = \alpha' + \theta Inst'_{i,d,p,t} + D'_2(characteristics)_{i,d,p,t}} + \eta'_{\boldsymbol{p}} + \eta'_{\boldsymbol{t}} + \boldsymbol{\varepsilon'_{i,d,p,t}} \end{cases} \quad \textbf{(1.3)}$$

$$y_{i,d,p,t} = \left| \begin{array}{l} 1 \; if \; y^*_{i,d,p,t} > 0 \\ 0 \; if \; y^*_{i,d,p,t} \leq 0 \end{array} \right. \qquad AccessMFI_{i,d,p,t} = \left| \begin{array}{l} 1 \; if \; AccessMFI^*_{i,d,p,t} > 0 \\ 0 \; if \; AccessMFI^*_{i,d,p,t} \leq 0 \end{array} \right.$$

We assume that the error terms in specification (1.3) follows a bivariate normal distribution:

$$\left[\frac{\boldsymbol{\varepsilon_i}}{\boldsymbol{\varepsilon'_i}}\right] \rightarrow \mathrm{N}\left[\begin{pmatrix}0\\0\end{pmatrix}, \begin{pmatrix}1 & \rho\\ \rho & 1\end{pmatrix}\right]$$

The instrumental variable that we use here is $Inst'_{i,d,p,t}$ and it represents the concentration of microfinance branches in the closest neighboring district in time period t. A good instrument in this case has to be highly correlated with geographical access to microfinance, and it must not have any direct effect on an individual's entrepreneurship status. The assumption here is that the microfinance institutions are likely to take into consideration the concentration of microfinance and its performance in neighboring districts. We already know from the existing literature that the location decisions of financial institutions such as microfinance organizations are driven by local aggregates (Ruiz-Tagle and Vella, 2015) and they also rely on the experience and policies from the neighbors (Allen et al., 2016). Our first stage estimation results clearly show that the concentration of microfinance in the neighbor district, measured as the number of microfinance branches, is significantly related to our access to microfinance indicator at a 1% significance level (Appendix 1.2).

Regarding the exclusion restriction, the concentration of microfinance institutions in the neighbor district should have no direct effect on entrepreneurship decision of the individual. Since microfinance institutions provide financial services by focusing on social collateral rather than physical collateral, the financial products of microfinance institutions are designed in a manner that potential customers are required to be permanents residents of the locality for a certain number of years in order to exhibit strong social ties within the community[23]. For example, the eligibility criteria to apply for any type of individual or group loan from one of the biggest and the oldest microfinance organizations in Pakistan '*Khushali Microfinance'* is that the person should be a permanent resident of the locality for at least two years. Similarly, in the randomized control trail conducted by Banerjee et al. (2015) in India, the microfinance institution required clients to be living in the same location for at least one year. In the case of Pakistan, this eligibility criteria of having a permanent residence normally ranges between one to two years, depending on the microfinance institution. Moreover, according to the prudential regulations by the State Bank of Pakistan, microfinance bank branches can provide services within a limited radius of their branch location within the same district[24]. Therefore, due to these regulatory requirements and the product design of microfinance institutions, individuals cannot go to neighbor districts to apply for microfinance. Hence, we believe that the instrument variable only affects entrepreneurship through our access to microfinance indicator. Moreover, we provide some balancing tests to show that individual characteristics are not very different for people who live in districts with relatively more microfinance in the neighboring districts as compared to people living in districts with less

[23] See Aslam and Azmat (2012) to have a detailed understanding of different collateral options for microfinance clients in Pakistan.

[24] See the State Bank of Pakistan's Branch Licensing Policy http://www.sbp.org.pk/bprd/2016/C4-Annx-A.pdf.

microfinance in the neighboring districts, highlighting that the instrument seems to be exogenous[25].

In the case of a bivariate probit model, there are no specialized tests to check the validity and strength of an IV. However, we obtain the F-statistic using 2SLS estimator on linear probability model. The F statistics comes out to be 9.4 which is very close to the rule of thumb of 10 highlighting the validity of the instrument. Moreover, the results of the Kleibergen-Paap rk test and Stock-Wright LM statistic also suggest that the instrument is relevant. Since we rely only on one Instrument variable, it is not possible to directly test the overidentification restriction through Sargan's J test. However, we conduct a test to see if the error terms from the first stage are correlated with the error terms from the second stage. The test comes out to be insignificant providing some more confidence about the validity of our instrument.

The results of the bivariate probit estimation, using an instrumental variable technique, are displayed in Table 1.5. They show that having access to a microfinance institution increases the likelihood of being an entrepreneur, after controlling for other individual and regional characteristics. On average, having access to microfinance increases the likelihood of being an entrepreneur by about 13 percentage points. However, the bivariate probit model can lower the precision of estimation so the magnitude of the effect should be treated with caution[26]. Results for other control variables remain quite similar to that of the univariate probit model. To conclude, access to microfinance is still found to significantly affect entrepreneurship even after taking into account the endogeneity bias; highlighting that the relationship between geographical access to microfinance and entrepreneurship is not spurious.

[25] A table with the comparison is provided in Appendix 1.5.

[26] See Mourifié and Méango (2014) and Filippini et al. (2018) for details.

Table 1.5: Econometric estimates with bivariate probit (IV)

Variables (Bivariate Probit (IV)	**(1) Entrepreneurship**
Access to Microfinance Institutions **(Instrumented)**	0.13**
	(0.06)
Education Level (Base: No formal Education)	
Primary Education	0.05***
	(0.01)
Secondary Education	0.07***
	(0.01)
Higher Education	-0.01
	(0.01)
Age (Base: 15-24)	
25-34	0.02
	(0.01)
35-44	0.04***
	(0.01)
45-54	0.01
	(0.02)
Over 55	-0.01
	(0.02)
Number of household members	-0.001
	(0.001)
Marital Status (Base: Single)	
Married	0.02*
	(0.01)
Divorced	0.27***
	(0.08)
Widowed	0.05
	(0.05)
Female Dummy	-0.29***
	(0.01)
Urban Dummy	0.03***
	(0.01)
MPI: Poverty Rate	-0.09**
	(0.04)
Year Fixed Effects	**Yes**
Province Fixed Effects	**Yes**
Total Observations	**11,055**
Wald (Prob > chi2)	**0.00**
F-test	**9.4**
Kleibergen-Paap rk (p-value)	**0.04**
Stock-Wright LM S statistic (p-value)	**0.05**

*Note: Results of the bivariate probit regression with an IV have been reported with average marginal effects. Standard errors are in parentheses (clustered at the district level). The dependent variable is entrepreneurship status dummy (see equation 1.3). ***, ** and * represent significance at 1%, 5% and 10%, respectively.*

Although we believe that the instrument variable is exogenous, it might still be argued that a household who is living right next to the border of a district might be able to get financial services from the neighboring district if the regulations are not followed strictly. In order to take this into account, we conduct another instrumental variable estimation, as a robustness check, to provide more confidence about the validity of our results. In this case, we rely on an instrument variable which is the average of microfinance branches in the neighboring districts after excluding the closest neighbor. The results remain fairly stable and highlight that the positive relationship between access to microfinance and entrepreneurship is significant and positive (Appendix 1.4).

Overall, our results show that having access to a microfinance institution has a positive effect on entrepreneurship. The potential mechanisms that drive this effect could possibly include impacts of individual borrowing, saving decisions, as well as general equilibrium effects (Kinnan and Breza, 2018). The numbers seem to suggest that the overall take-up of financial services from microfinance institutions in Pakistan has grown rapidly over the last few years. According to the Pakistan Microfinance Network, the total number of microfinance branches have increased by more than 100%, from 1,739 in 2012 to 3,671 in 2017. Moreover, the numbers of active borrowers grew by more than 2.5 times from 2.1 million in 2012 to 5.8 million in 2017, while the number of total saving accounts grew up by about 8 times from 4 million in 2012 to 31 million in 2017 (Pakistan Microfinance Network, 2017). Furthermore, the numbers from the Financial Inclusion Insight Survey also show that the take-up of formal loans is higher in areas with a microfinance branch as compared to areas without a microfinance branch. This is especially true for the bottom half of the population. For example, people with access to a microfinance branch within 10 kilometers borrow almost twice as much as people without access to a microfinance institution. Moreover, informal borrowing from moneylenders also seems to be slightly lower, on average, in areas with a microfinance branch. About 2.0% people reported borrowing from a moneylender in areas without a microfinance branch, as compared to 1.7% people in areas with a microfinance branch.

1.4.3. Results with the interaction between poverty and microfinance

To further explore the nexus between access to microfinance, poverty and entrepreneurship, we conduct a regression with an interaction term for access to microfinance and the district poverty rate. In this case, we use a linear probability model (LPM) because in non-linear models (including logit and probit models), the interaction effect cannot be interpreted directly by looking at the sign, magnitude or the statistical significance of the coefficient of the interaction term (Ai and Norton, 2003)[27]. The results of the econometric estimates with the interaction term are provided in Table 1.6. The last column of Table 1.6 presents the results after instrumenting the access to microfinance variable. The results show that the access to microfinance variable (not interacted) is no longer significant and the poverty rate at the district level is negative and statistically significant. Whereas, the interaction term between access to microfinance and poverty rate is statistically significant and positive. This means that access to microfinance in itself does not have an impact on entrepreneurship, but the impact runs through the poverty channel. In other words, having access to microfinance increases the chances of becoming an entrepreneur in poorer regions.

Table 1.6: Results with the interaction between district poverty and microfinance

Variables (Linear Probability Model (LPM))	**(1) Entrepreneurship**	**(2) Entrepreneurship**
Interaction term (Access to MFIs and MPI)	0.13**	0.10***
	(0.05)	(0.03)
Access to Microfinance Institutions (MFIs)	-0.01	-
	(0.02)	
Access to Microfinance Institutions (MFIs) (Instrumented)	-	0.10
		(0.07)
MPI: Poverty Rate	-0.13***	-0.15**
	(0.04)	(0.04)
Education Level (Base: No formal Education)		
Primary Education	0.04***	0.04***
	(0.01)	(0.01)
Secondary Education	0.06***	0.06***
	(0.01)	(0.01)
Higher Education	-0.05***	-0.04**
	(0.02)	(0.02)
Age (Base: 15-24)		
25-34	0.00	0.01
	(0.01)	(0.01)
35-44	0.02*	0.04**
	(0.01)	(0.02)
45-54	-0.01	-0.00
	(0.01)	(0.02)
Over 55	-0.03*	-0.02

[27] Results using the linear probability model are comparable with those obtained from the Probit model. For example, the corresponding estimate on the access to microfinance institutions variable using the LPM is 0.04 (results are provided in appendix 1.1) as compared to 0.04 (column (2) in Table 1.2).

	(0.02)	(0.02)
Number of household members	0.00	-0.00
	(0.00)	(0.00)
Marital Status (Base: Single)		
Married	0.04***	0.04**
	(0.01)	(0.01)
Divorced	0.19***	0.21***
	(0.06)	(0.06)
Widowed	0.06**	0.06**
	(0.03)	(0.03)
Female Dummy	-0.27***	-0.28***
	(0.02)	(0.02)
Urban Dummy	0.03**	0.02**
	(0.01)	(0.01)
Year Fixed Effects	**Yes**	**Yes**
Province Fixed Effects	**Yes**	**Yes**
Total Observations	**11,055**	**11,055**
Fischer (Prob > F)	**0.00**	**0.00**
R Squared	**0.175**	**0.175**

*Note: Results of the linear probability model have been reported. Standard errors are in parentheses (clustered at the district level). The dependent variable is entrepreneurship status dummy. ***, ** and * represent significance at 1%, 5% and 10%, respectively.*

1.4.4. Robustness Checks

In order to assess the strength of our results, we conduct different robustness checks using the baseline regression reported earlier in Table 1.2. First, instead of using a ten kilometers range as our dependent variable of geographical access to microfinance, we use five kilometers and fifteen kilometers respectively to check if the results remain robust. The results of the econometric estimates are reported in Table 1.7. Column 1 shows the results when we use the five kilometers range and column 2 presents the results when we use the fifteen kilometers range as the measure of geographical access to microfinance. The access to microfinance variable comes out to be significant at a 5% significance level in both cases. All the control variables are not displayed in the table for ease of reading.

We conduct an additional robustness check by using the transportation time to the microfinance institution instead of using distance in kilometers. Although the fixed effects in our models control for the heterogeneity in infrastructure quality across districts, it might still be useful to look at access in terms of transportation time given the differences in quality of roads and public transportation across different areas. For that purpose, we rely on the question *“If you had to go to a microfinance institution, how much time would it take you?”*. The binary variable in this case takes the value of “1” if the time taken to reach the microfinance institution is less than one hour, and it takes the value “0” otherwise. The results are reported in Table 1.7 (Column3) and they confirm that access to microfinance increases the chances of being an entrepreneur.

Table 1.7: Robustness checks using different thresholds of access to microfinance

Variables	Entrepreneurship (1)	Entrepreneurship (2)	Entrepreneurship (3)
	Access to microfinance (5 KMs)	**Access to microfinance (15 KMs)**	**Access to microfinance (Less than 1 hour)**
Access to Microfinance Institutions	0.030** (0.01)	0.031** (0.01)	0.033** (0.01)
Other Controls	Yes	Yes	Yes
District Fixed Effects	Yes	Yes	Yes
Year Fixed Effects	Yes	Yes	Yes
Settlement-size Fixed Effects	No	No	Yes
Total Observations	10,811	10,811	10,811
Pseudo R2	0.2755	0.2758	0.2762

*Note: Results of the probit regressions have been reported with average marginal effects. Standard errors are in parentheses (clustered at the district level). The dependent variable is entrepreneurship status dummy. Control variables have been included but not displayed due to ease of reading. ***, ** and * represent significance at 1%, 5% and 10%, respectively.*

Moreover, the urban and rural classification might not be the perfect proxy to control for the type of location and population density. The densely populated areas may be attractive for microfinance institutions and also more suitable for entrepreneurial activity. Although our estimates take into account the variation in population densities across different district in Pakistan through district fixed effects, it is also important to control for this at a more granular level. For that purpose, we incorporate settlement size fixed-effects in the model. The “settlement size” variable in the dataset provides much more detailed information pertaining to the total population of the area where the household resides. This variable is categorical and divides the households into different categories with respect to the total population of the area. These different categories of settlement size include households living in an area with “Less than 2000 people”, “2,000-4,999 people”, “5,000-19,999 people”, “20,000-49,999 people”, “50,000-99,999 people”, “100,000 and above”, and the “capital city”. So, instead of using the urban and rural classification, we use settlement size fixed-effects to run the same baseline regression as in Column 3 of Table 1.2. The result confirms that access to microfinance increases the chances of being an entrepreneur (Column 1 of Table 1.8).

As a final robustness check, we run the baseline regression of Table 1.2 (Column 3) after excluding housewives from the overall sample. Since majority of the housewives exclude themselves from the labor market voluntarily, we feel that it is important to run the regression by including only those who belong to the labor market. The results show that there is still a significant and positive relationship between access to microfinance and entrepreneurship (Column 2 of Table 1.8).

Table 1.8: Additional robustness checks

Variables	**Entrepreneurship (1)**	**Entrepreneurship (2)**
	Settlement-size control	**Excluding housewives**
Access to Microfinance Institutions (10 kms)	0.034** (0.015)	0.06** (0.01)
Other Controls	Yes	Yes
District Fixed Effects	Yes	Yes
Year Fixed Effects	Yes	Yes
Settlement-size Fixed Effects	Yes	No
Total Observations	10,811	5,851
Pseudo R2	0.2772	0.1147

*Note: Results of the probit regressions have been reported with average marginal effects. Standard errors are in parentheses (clustered at the district level). The dependent variable is entrepreneurship status dummy. Control variables have been included but not displayed due to ease of reading. ***, ** and * represent significance at 1%, 5% and 10%, respectively.*

1.5. Conclusion

This chapter investigated the effect of geographical access to microfinance on the formation of entrepreneurs, hence, contributing to the research on both entrepreneurship and microfinance. Using a pseudo-panel data from Pakistan for the years 2015 and 2016, our econometric estimates found a statistically significant and positive relationship between access to microfinance and the status of entrepreneurship. The results showed that having access to microfinance can allow individuals, both men and women, to move up the economic ladder by shifting to entrepreneurship and starting their own business rather than working as low paid employees, farm workers, and housewives. Having this access to microfinance institutions is likely to lower credit costs and allow previously financially excluded individuals to have access to formal financial services. We found that having access to microfinance increases the probability of being an entrepreneur by about 4 percentage points. The result is corroborated by the instrumental variable estimation.

Our empirical results do support the initial claims put forward by Muhammad Yunus regarding the role of access to microfinance in harnessing the entrepreneurial spirit of people in developing economies. However, the magnitude of this effect of access to microfinance on entrepreneurship comes out to be modest. This can point towards two things. First, as Banerjee (2013) pointed out, credit constraints might not be the biggest constraint faced by businesses and some of the major problems may instead be factors pertaining to the quality of business environment in which they operate. Secondly, not everyone desires to become an entrepreneur and thus extending access to microfinance cannot be expected to turn a whole nation into entrepreneurs (Duflo, 2010). Nonetheless, having geographical access to microfinance, which encompasses more than group micro-credit, does play a noteworthy role in helping people move from least earning occupations to switch to entrepreneurship in the context of Pakistan. This highlights that microfinance holds real appeal, especially in the context of countries such as Pakistan where a large segment of population remains unbanked.

Another important finding is that access to microfinance in itself does not have an impact on entrepreneurship, but the impact of access to microfinance on entrepreneurship runs through the poverty channel i.e. access to microfinance increases the chances of becoming an entrepreneur in poorer regions. This reiterates the poverty alleviation power of financial inclusion and indicates that financial inclusion can be a key enabler of the sustainable development goals (SDGs). This is consistent with the empirical evidence provided by Bruhn and Love (2014) where they showed that increased access to microfinance led to an increase in micro-entrepreneurship and an increase in income for a large segment of the population in Mexico.

Moreover, our econometric estimates also point towards the existence of gender disparity in the status of entrepreneurship. On average, women are about 23 percentage points less likely to be entrepreneurs as compared to men, after taking into account other individual and district level variables. The next chapter of this thesis discusses this link with gender in much more detail. The results also highlight that residing in an urban area increases an individual's probability of being an entrepreneur as compared to residing in a rural area by 4.5 percentage points. Similarly, other factors that have an effect on the status of entrepreneurship include education level and marital status. An increase in formal education also seem to increase the probability of being an entrepreneur. Lastly, we do not find an effect of household size on entrepreneurship.

As for policy implications, the empirical investigation suggests that in order to obtain new sources of sustainable growth by boosting entrepreneurial activities, Pakistan needs to promote financial inclusion through microfinance, particularly in less developed regions. A very high level of financial exclusion constraints entrepreneurial behavior and makes it problematic for individuals and businesses to generate income. However, this step alone will not be sufficient to have the desired outcome. There is a considerable amount that needs to be done in order to improve the climate of conducting business across different districts in Pakistan, especially in the poorer regions. Pakistan has shown a clear focus on financial inclusion with the development of a financial inclusion strategy dedicated to populations without financial access. However, many of these reforms do not directly address the barriers that are faced by entrepreneurs living in poor regions. Some of the policy measures to promote entrepreneurship, especially in poor regions, could include enforcing increased access to financial services for entrepreneurs, pairing targeted financing schemes with other measures such as professional trainings, increased access to support programs and financial literacy programs, fostering a gender neutral legal framework for businesses, and improving the overall climate of doing business.

Another area of concern is the large gender disparity in the status of entrepreneurship in Pakistan. The issues faced by women entrepreneurs are similar to those faced by entrepreneurs in general and largely centered on access to finance and markets as well as the climate for doing business. Nevertheless, many characteristics of women entrepreneurs and of their enterprises differ from those of men, and therefore require specific policy interventions. The next chapter of this thesis focuses on women entrepreneurship in particular.

Acknowledgements

Comments and feedback from the participants of the 6th European Research Conference on Microfinance at the University of Paris Dauphine, 8th International Research Conference on Financial Inclusion and Microfinance at the Oslo Metropolitan University, 18th session of Institutional and Organizational Economics Academy at the Institute of Scientific Studies in Cargèse, and the conference on development economics issues for doctoral students at CERDI are very much appreciated.

1.6. Appendix

Appendix 1.1: Linear probability model (LPM) estimation of entrepreneurship status

Variables (LPM)	**(2) Entrepreneurship**
Access to Microfinance Institutions (MFIs)	0.04**
	(0.02)
Education Level (Base: No formal Education)	
Primary Education	0.04***
	(0.01)
Secondary Education	0.06***
	(0.01)
Higher Education	-0.04***
	(0.02)
Age (Base: 15-24)	
25-34	0.005
	(0.01)
35-44	0.02*
	(0.01)
45-54	-0.01
	(0.01)
Over 55	-0.035*
	(0.02)
Number of household members	0.001
	(0.01)
Marital Status (Base: Single)	
Married	0.04***
	(0.01)
Divorced	0.19***
	(0.06)
Widowed	0.06
	(0.03)
Female Dummy	-0.27***
	(0.01)
Urban Dummy	0.02*
	(0.01)
MPI: Poverty Rate	-0.12***
	(0.04)
Year Fixed Effects	**Yes**
Province Fixed Effects	**Yes**
District Fixed Effects	**No**
Total Observations	**10,904**
R-Squared	**0.1754**
Prob> F	**0.000**

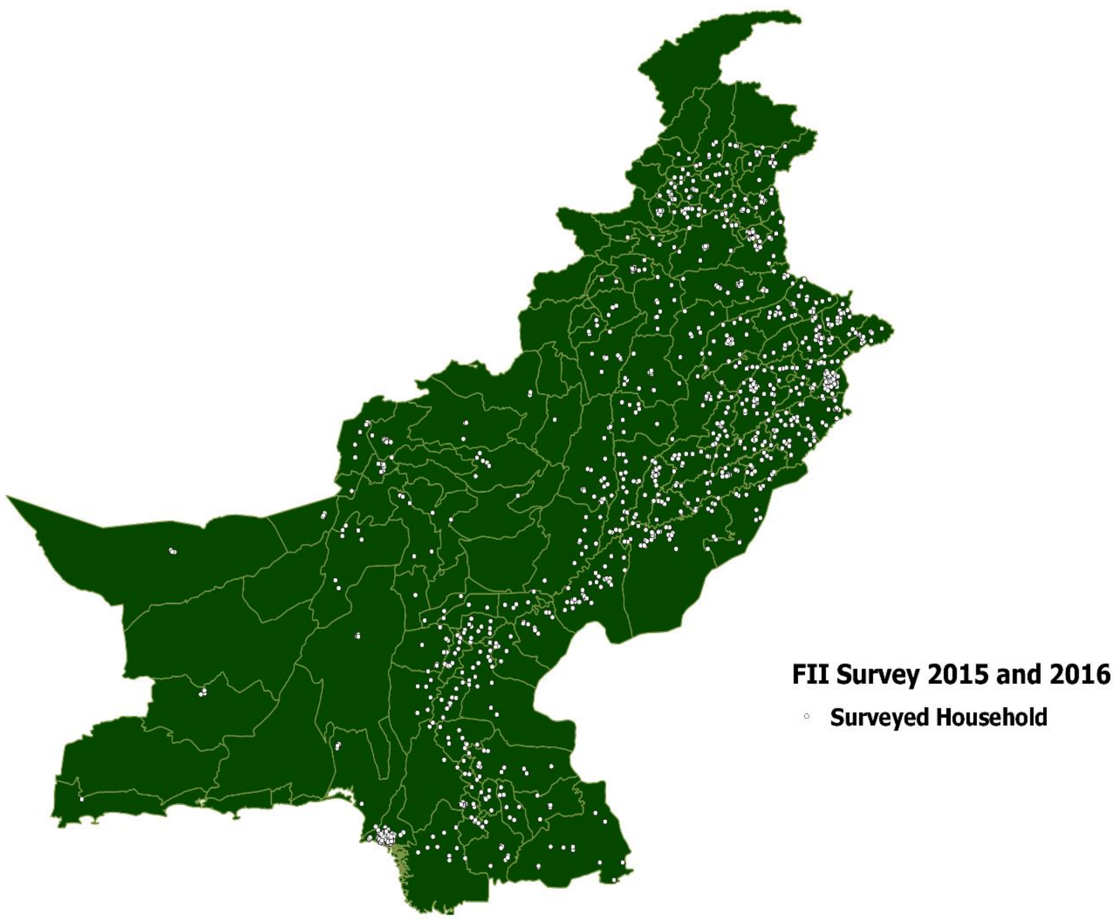
****, ** and * represent significance at 1%, 5% and 10%, respectively.*

Appendix 1.2: First-stage regression results for the access to microfinance indicator

Variables (First Stage of Bivariate Probit)	**(1)** Access to Microfinance Institutions
Neighbor District Microfinance Concentration	-0.005***
	(0.002)
Other Controls	**Yes**
Provincial Fixed Effects	**Yes**
Year Fixed Effects	**Yes**
Total Observations	**11,055**
Wald (Prob > chi2)	**0.000**
F-test	**9.4**
Kleibergen-Paap rk (p-value)	**0.04**
Stock-Wright LM S statistic (p-value)	**0.05**
*** p<0.01, ** p<0.05, * p<0.1	

Note: Controls for individual characteristics and regional characteristics have been included (not reported).

Appendix 1.3: Surveyed Households and their Locations



Source: Financial Inclusion Insight (FII) Survey, 2015 and 2016

Appendix 1.4: Second Instrument Variable Estimation Results

Variables (Bivariate Probit (IV))	**(1) Entrepreneurship**
Access to Microfinance Institutions **(Instrumented)**	0.12**
	(0.06)
Education Level (Base: No formal Education)	
Primary Education	0.05***
	(0.01)
Secondary Education	0.06***
	(0.01)
Higher Education	-0.01
	(0.02)
Age (Base: 15-24)	
25-34	0.01
	(0.01)
35-44	0.04***
	(0.01)
45-54	0.01
	(0.02)
Over 55	-0.01
	(0.02)
Number of household members	-0.00
	(0.00)
Marital Status (Base: Single)	
Married	0.02**
	(0.01)
Divorced	0.28***
	(0.08)
Widowed	0.05
	(0.05)
Female Dummy	-0.29***
	(0.01)
Urban Dummy	0.03***
	(0.01)
MPI: Poverty Rate	-0.07*
	(0.04)
Year Fixed Effects	**Yes**
Province Fixed Effects	**Yes**
Total Observations	**11,055**
Wald (Prob > chi2)	**0.00**

*Note: Results of the bivariate probit regression with an IV have been reported with average marginal effects. Standard errors are in parentheses (clustered at the district level). The dependent variable is entrepreneurship status dummy (see equation 1.3). ***, ** and * represent significance at 1%, 5% and 10%, respectively*

Appendix 1.5: Robustness tests

A. Comparing individual characteristics at different thresholds of instrument variable (IV)

	IV below the Median threshold	IV above the Median threshold	Difference
	Average	Average	
Female Dummy	.499	.484	.014
Urban Dummy	.329	.337	-.008
Education_ No formal education	.353	.356	-.003
Education_ Primary	.252	.233	.019
Education_ Secondary	.304	.32	-.016
Education_ Tertiary	.089	.09	-.001
Age	37.515	35.855	1.66***
Marital Status_ Single	.135	.152	-.018*
Marital Status_ Married	.841	.816	.025**
Marital Status_ Divorced/ Separated	.003	.004	-.001

*Note: The instrument variable (IV) is the concentration of microfinance branches in the neighboring district. Total sample is 11,055 observations. *, ** and *** denote statistical significance at 10%, 5% and 1% respectively.*

B. Comparing individual characteristics at different thresholds of the Access to Microfinance Indicator

	No MFI Access within 10km	MFI Access available within 10km	Difference
	Average	Average	
Female Dummy	.484	.505	-.02
Urban Dummy	.363	.284	.079***
Education_ No formal education	.357	.349	.009
Education_ Primary	.238	.25	-.012
Education_ Secondary	.312	.311	.001
Education_ Tertiary	.091	.087	.004
Age	36.978	36.446	.532
Marital Status_ Single	.14	.145	-.005
Marital Status_ Married	.829	.832	-.003
Marital Status_ Divorced/ Separated	.004	.002	.003
Total Observations	6616	4439	

, ** and * denote statistical significance at 10%, 5% and 1% respectively.*

Chapter 2: Financial Inclusion: An illusion for creating women entrepreneurship?

Empirical Evidence from Mexico

This chapter is based on a paper co-authored with Mabel GABRIEL, Patrick LENAIN, and Julien REYNAUD. The paper was published as an OECD Working Paper and cited as below:

Fareed, F., Gabriel, M., Lenain, P. and Reynaud, J. (2017). “Financial Inclusion and Women Entrepreneurship: Empirical Evidence from Mexico”. *OECD Economics Department Working Papers*, No. 1411, OECD Publishing, Paris.

Abstract

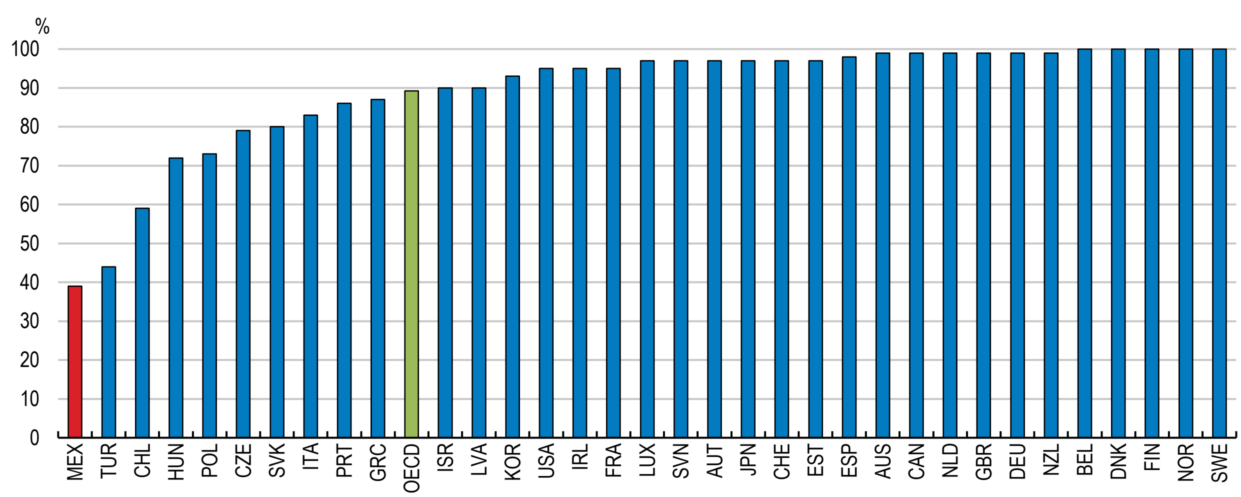
Financial inclusion and women entrepreneurship concern policymakers because of their impact on job creation, economic growth and women empowerment. Women in Mexico do engage in paid work but many of them lack opportunities to work in the formal sector. Moreover, the financial exclusion rate in Mexico is amongst the highest in comparison to its peers, affecting women in particular. This chapter uses an individual-based panel dataset over the period 2009-2015, with more than four million observations, to study the interface between financial inclusion and women entrepreneurship across different economic sectors, disentangling between informal and formal jobs. A comprehensive financial inclusion index is constructed in order to estimate the impact of financial inclusion on women entrepreneurship in Mexico. The results suggest that financial inclusion is positively linked with entrepreneurship, meaning that it can open up economic opportunities for women entrepreneurs. However, the positive relationship between women entrepreneurship and financial inclusion does not hold for women entrepreneurs working in rural areas in the informal sector or women working in the commerce sector, highlighting lower entry barriers, including financial, in the informal sector. Results also point towards the existence of gender disparity in the status of entrepreneurship across formal and informal work. On average, women are 3 percentage points less likely to be entrepreneurs in the formal sector and 11 percentage points more likely to be entrepreneurs in the informal sector, as compared to men, after taking into account other relevant individual and geographical characteristics. The positive relationship between financial inclusion and women entrepreneurship is also supported by the Instrumental Variable (IV) estimation.

2.1. Introduction

The term financial inclusion is broadly defined as the access to and the use of formal financial services by households and firms, those without such access are termed as financially excluded. Having access to formal financial services allows firms to invest and people to save and borrow, allowing them to smooth their consumption, invest in entrepreneurial venture and build capital over time, which can lead to improvement in people's livelihoods (Bauchet et al., 2011; Demirguc-Kunt and Levine, 2007). Furthermore, it can also enable individuals to successfully manage income shocks and emergencies such as illnesses or loss of employment (Rojas-Suarez, 2010). Financial inclusion is considered as a key enabler of economic growth and poverty reduction (Rajan and Zingales, 1998; Pasali, 2013; Samargandia et al., 2015) and the G20 have prominently featured financial inclusion action plan in their strategy.

Despite an increased focus on financial inclusion, around 1.7 billion adults across the globe remain unbanked and women remain financially more excluded than men, especially in developing countries, according to the global Findex database (Demirguc-Kunt et al., 2018). This is also the case in Mexico where the share of women with an account at a financial institution is the lowest amongst all members of the Organization for Economic Co-operation and Development (OECD) as shown in Figure 2.1. Only 39% of the women have a formal bank account in Mexico as compared to the OECD average of 94% and the world's average of 57% (Demirguc-Kunt et al., 2015).

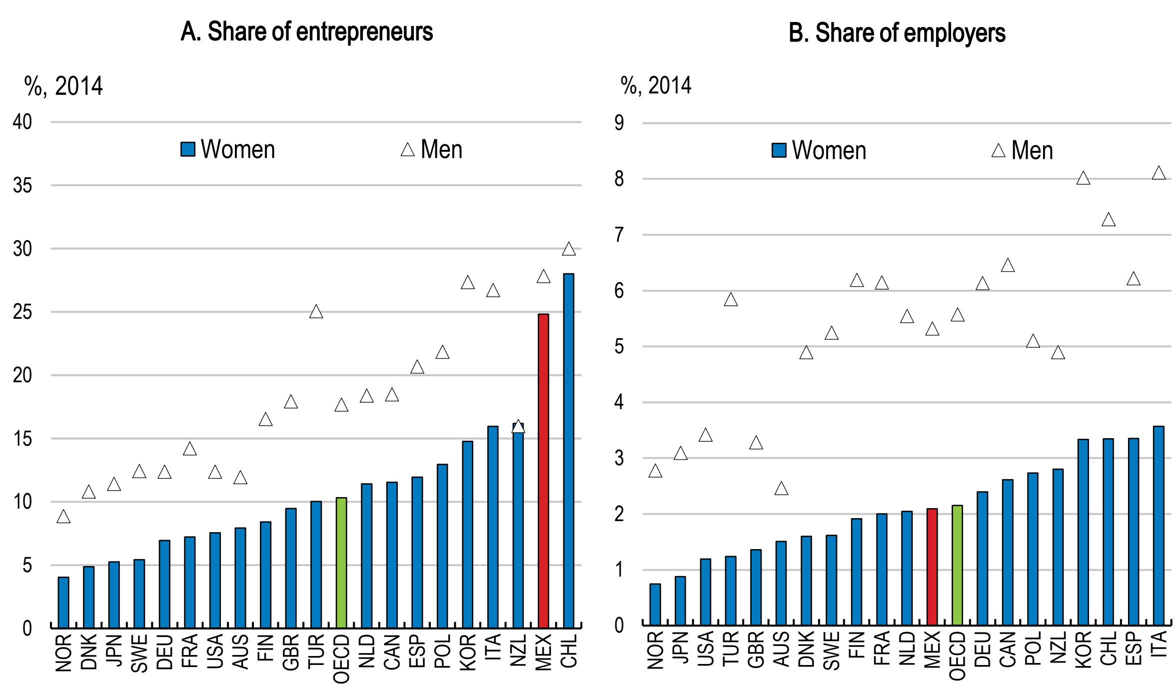
Figure 2.1: Share of adult women with an account at a financial institution



Source: World Bank's Global Findex database, 2014-15

Moreover, when women in Mexico engage in paid work, many of them work in the informal sector because they lack opportunities to work in the formal sector (Chen, 2001). Women-owned businesses could be a key source of job creation, innovation, and a way to address inequalities, given the gender gap in labor participation in Mexico (OECD, 2017). However, women in Mexico represent less than 3% of formal employers and the percentage of women entrepreneurs is also lower than men in Mexico (Figure 2.2). Although the share of women entrepreneurs is higher for Mexico as compared to other OECD countries, this is mainly because standard datasets do not distinguish between formal and informal entrepreneurship. Most of the studies on entrepreneurship in developing countries are based on male entrepreneurs with very little focus on their female counterparts and the informal sector. Beatrice (2012) argues that it is inadequate to use the findings of these studies focused primarily on men to understand the dynamics of women entrepreneurs. Many issues faced by women entrepreneurs are generally parallel to those faced by men and are largely centered on the climate for doing business and access to financial products and markets. However, these factors do not tend to influence both genders in the same manner or with the same magnitude. Many characteristics of these women run enterprises differ from those of men, and therefore require specific policy interventions.

Figure 2.2: Women represent a low share of employers and entrepreneurs



Note: Data for AUS, MEX, NZL refers to 2015
Source: OECD Entrepreneurship at a Glance 2016

In case of financial exclusion, many small and medium sized businesses and poor individuals who depend on external source of finance encounter financial constraints as a barrier to start or grow their businesses. A range of theoretical models have been used to demonstrate that the lack of access to financial services can hamper entrepreneurial activity (Beck et al., 2007; Banerjee and Newman, 1993; Galor and Zeira, 1993). Different empirical studies at the micro level have also tried to gauge the impact of financial inclusion on entrepreneurship but have found mixed results. On one hand, Black and Strahan (2002) and Angelucci, Karlan and Zinman (2015) found that an increased level of financial inclusion led to an increase in entrepreneurial activity. Bruhn and Love (2014) also showed a significant effect of increased access to finance on business creation after relying on a natural experiment in Mexico where 800 new bank branches were opened simultaneously. On the other hand, Ashraf et al. (2010) and Tarozzi et al. (2015) failed to find a meaningful effect. In fact, research by Karlan and Zinman (2011) found a negative effect of increased level of financial inclusion on various business outcomes in Philippines.

It is important to note that financial inclusion is a multi-dimensional concept, as discussed in the general introduction section, and there is still no universally accepted way of defining or measuring financial inclusion. The existing empirical literature on financial inclusion has focused mainly on the accessibility of formal financial services rather than the actual usage[28]. Notwithstanding financial inclusion's beneficial effects and the efforts by the World Bank and G20 to improve universal access to formal bank accounts, having mere access to financial services or having a bank account in itself does not guarantee that firms and households can actually use them. The existing literature typically disregards this issue and tends to mix access with the actual distribution of credit, savings, and other financial services. It is important to note that being able to manage an official bank account can be a challenge for firms and workers in the informal sector, especially women. Keeping that in mind, we propose a new way of looking at financial inclusion which combines access with the actual usage of financial services from different financial channels.

[28] A summary of the selected empirical literature is provided in Appendix 2.1 where we clearly mention the different dimensions of financial inclusion and their impact on different socio-economic indicators.

This chapter examines empirically the relationship between financial inclusion, women entrepreneurship, and informality using an individual based pseudo-panel dataset over the period 2009-2015 with more than four million observations. Three research questions are studied: First, does an increased level of financial inclusion enhance women entrepreneurship? Secondly, how do these findings vary across different economic sectors such as manufacturing, commerce, services etc., and across formal and informal work type? Lastly, what are the possible channels that influence this relationship and what are the different factors enhancing women entrepreneurship from a policy point of view? The results suggest that being financially included does help women entrepreneurship in case of Mexico. An increase in the level of financial inclusion can improve access to formal financial system and alleviate credit constraints for potential entrepreneurs. Indeed, our results show that living in a municipality with a higher level of financial inclusion does increase the probability of becoming an entrepreneur, even after controlling for individual and local characteristics. Our results also point towards the existence of a high gender disparity in the status of entrepreneurship across formal and informal work in Mexico. More specifically, women are less likely to be entrepreneurs in the formal sector and more likely to be entrepreneurs in the informal sector, as compared to men. Lastly, we find that the positive relationship between financial inclusion and entrepreneurship does not hold for women working in the commerce sector.

This chapter adds to the existing literature in three distinct ways. First, to the best of our knowledge, this is one of the earliest studies to systematically address the issue of financial inclusion by building a comprehensive financial inclusion index using within country geographical data in case of a developing economy. Most studies typically provide a cross-country analysis by using a very narrow definition of financial inclusion e.g. percentage of adults with a bank account (Anand and Kuldip, 2013). It is well recognized that cross country analyses exhibit several limitations due to their inability to account for differences in distributions which makes within-country studies more appealing especially from a policy point of view. The results of our financial inclusion index highlight the differences in the access and usage of financial services across different municipalities in Mexico over time. Secondly, this chapter contributes to the literature on understanding the link between financial inclusion and entrepreneurship. The link across different economic sectors, across urban and rural areas, and across formal and informal work are studied separately to better understand the dynamics of this linkage across various channels. The existing literature focuses only on one or a few of these channels. Finally, this chapter contributes to the literature on gender discrimination with respect to the status of financial inclusion and entrepreneurship.

The rest of the chapter is structured as follows. Section 2.2 presents an overview of the data. It also offers an overview of Mexico's financial inclusion and entrepreneurship landscape, providing stylized facts about the recent state of financial inclusion and entrepreneurship in Mexico across regions, across gender, and across economic sectors. Moreover, this section presents the details of the financial inclusion index (FII) constructed to measure the state of financial inclusion in Mexico. Section 2.3 presents the empirical methodology and Section 2.4 discusses the results of the econometric analysis aiming at examining the key drivers of women entrepreneurship and investigating its link with financial inclusion. Section 2.5 presents the details of the robustness checks. Finally, Section 2.6 concludes.

2.2. Data and the Mexican Context

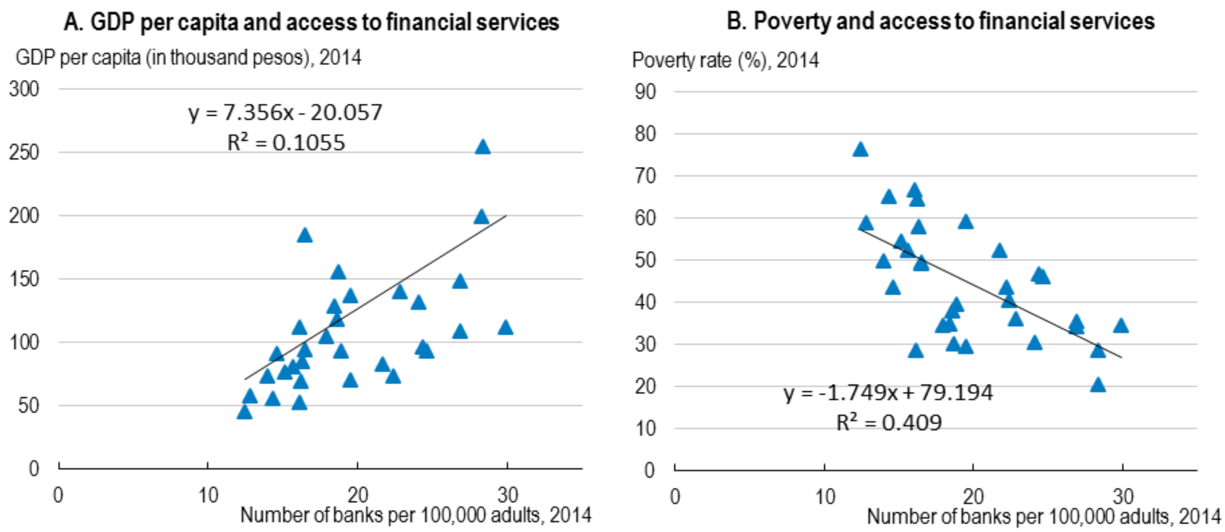
2.2.1. About the Data

In order to conduct this empirical examination, a panel dataset at the individual level is constructed from three main databases. First, data from the National Survey of Occupation and Employment (ENOE for its acronym in Spanish) is used which provides detailed information about the socio-economic characteristics and the labor conditions of individuals on a quarterly basis. This data from quarter four of 2009 to quarter three of 2015 was pooled. This is a pseudo panel because the same individuals are not followed over time. Then, this data was paired with municipality level data from the National Banking and Securities Commission (CNBV for its acronym in Spanish) which provides quarterly information on financial inclusion indicators pertaining to access and usage of financial services. Lastly, the World Bank indicators pertaining to the ease of doing business, which show the extent to which the regulatory environment is favorable for starting and operating a business at the state level, were added. The total sample contains people who fall in the active population category, i.e. people that are either working as employees, employers, entrepreneurs or working without pay and looking for jobs. In this chapter, we define an entrepreneur as someone who owns or co-owns a business. This is the same definition that is being used by the Mexican National Survey of Occupation and Employment (ENOE). In line with this definition and the existing literature, we mainly focus on off-farm entrepreneurship i.e. people working on the agricultural farms are not considered as entrepreneurs and thus treated as a separate employment category[29]. These businesses operated by entrepreneurs can operate with multiple employees or they can have the entrepreneur as the only employee. A table with summary statistics along with the description of all the variables is provided in Appendix 2.2 and 2.3.

2.2.2. Panorama of Financial Inclusion in Mexico

Access to financial services is considered as one of the key enablers of economic growth and poverty reduction (Cull, Ehrbeck, and Holle, 2014). The following scatter plots, using state level data from Mexico, show that access to finance[30] and GDP are positively correlated and access to finance and poverty are negatively correlated (Figure 2.3).

Figure 2.3: Association between access to finance, economic growth, and poverty



Source: CNBV, Coneval, and INEGI 2014

First, this section provides a detailed overview of the current state of financial inclusion in Mexico by looking at the demand and supply side of financial inclusion. Then, this section provides the details of our financial inclusion index (FII), which we have constructed considering different dimensions of financial inclusion using municipality level data.

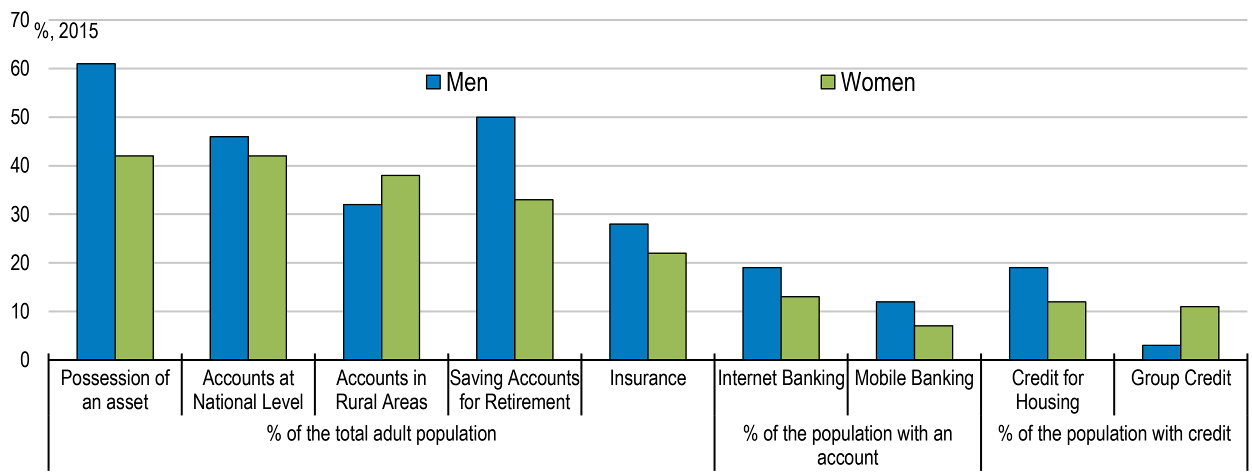
[29] This is the same classification that was used in Chapter 1. Off-farm entrepreneurship is quite different in terms of business dynamics from agricultural entrepreneurship and thus we differentiate between the two.

[30] Measured as the number of bank branches per 100,000 adults.

Stylized facts from the Demand Side

The National Financial Inclusion Survey (ENIF for its acronym in Spanish), conducted in 2012 and then in 2015, provides information on the usage of financial services, i.e. the demand side. With respect to credit, 29% of the adult population reported having a credit from a formal financial institution, whereas, about 44% of the adult population reported saving at a formal financial institution. The results of the survey highlight that there are some large gaps across gender in the usage of different financial services. At the national level, men have more bank accounts than women and they use internet banking and mobile banking more. Moreover, some striking differences are visible in the indicators pertaining to ownership of an asset, having a retirement savings account and the usage of insurance services, where women tend to be more financially excluded than men. On the other hand, women in rural areas seem to have more bank accounts as compared to men and they are also the main users of group loans (Figure 2.4). Progress in the percentage of women with a saving account in rural areas should be noted, rising from 19% of women in 2012 to 38% in 2015, while for men this increase was from 26% in 2012 to 32% in 2015. This positive increase seems to be partly linked to the government initiative *Programa Integral de Inclusión Financiera* launched in 2014 which provides financial education, credit, programmed savings, insurance and other products and services to beneficiaries of social programs, the vast majority of whom are women (CONAIF, 2016). One factor that might clarify the gender gap in insurance and savings account for retirement is the higher labor force participation of men compared to women, as in many cases, employees receive these as part of employment benefits. This represents a higher poverty risk for women at old ages.

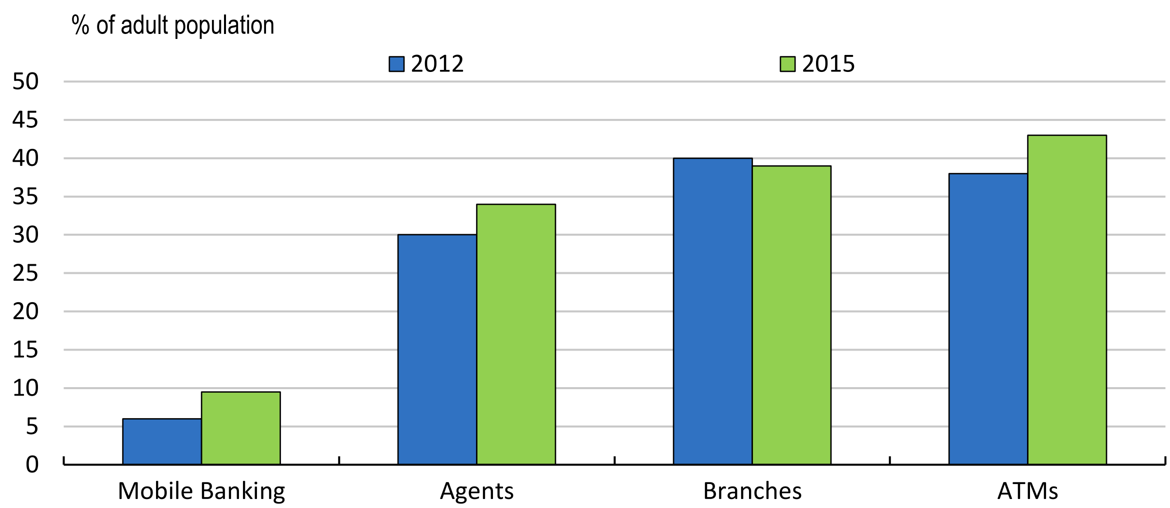
Figure 2.4: Gender gaps in financial inclusion



Source: ENIF, 2015

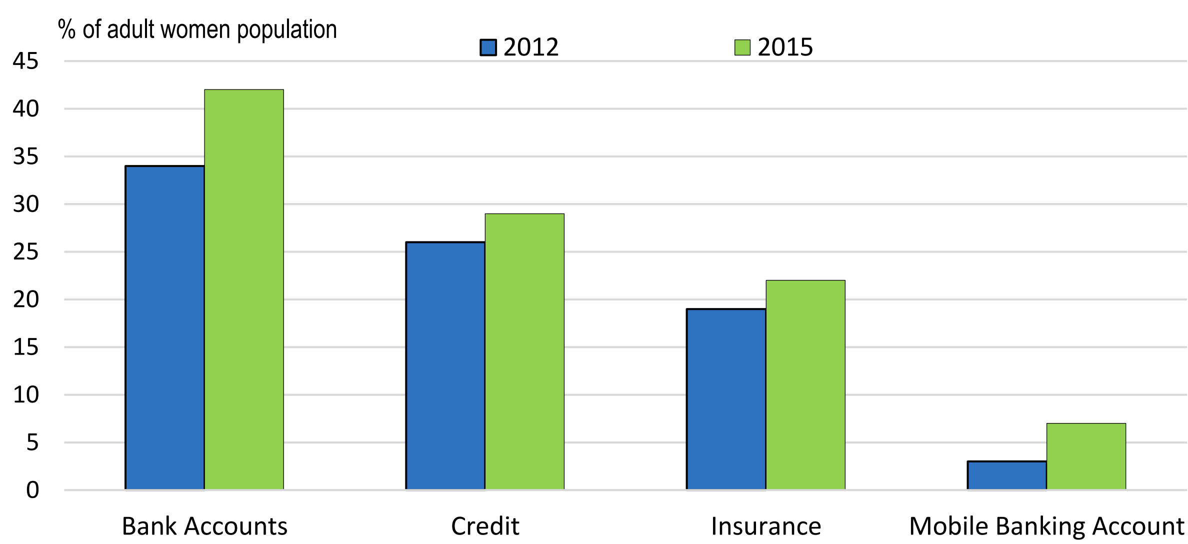
The usage of various financial access points, such as banking agents, ATMs and mobile banking, has shown a slight increase from 2012 to 2015, except from bank branches which fell down by 1%. Figure 2.5 below shows the different formal financial channels that are being used to get financial services and how the usage of these channels has changed over time. Moreover, Figure 2.6 highlights how the access to different financial services such as savings accounts, credit insurance and mobile banking has increased for women in particular.

Figure 2.5: Usage of various financial access points



Source: ENIF, 2012 & 2015

Figure 2.6: Usage of financial services for women

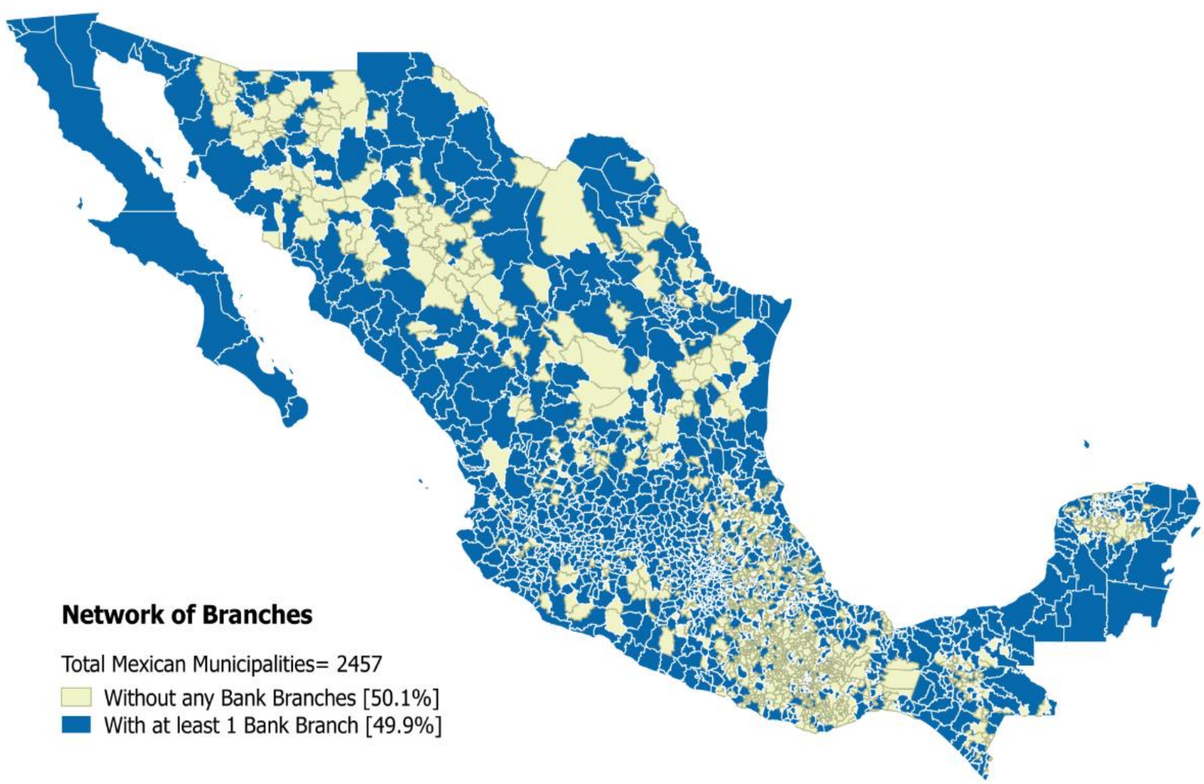


Source: ENIF, 2012 & 2015

Stylized facts from the Supply Side

Regarding the state of financial inclusion, Mexico has shown a clear focus on financial inclusion by developing a national financial inclusion body and introducing crucial financial reforms in 2009-2010 and 2014. However, despite all these actions over the past years, there exist large gaps in financial inclusion across regions. Half of the municipalities in Mexico have no bank branches[31] (as shown in Figure 2.7), while 31% of the 2,457 municipalities have no formal financial access points i.e. no bank branches, no banking agents and no automated teller machines (ATM). Figure 2.8 highlights the change in the percentage of municipalities with no formal bank branches over time depicting that the percentage of municipalities with no formal banking branch has gone down from 61% in 2009 to 50% in 2015.

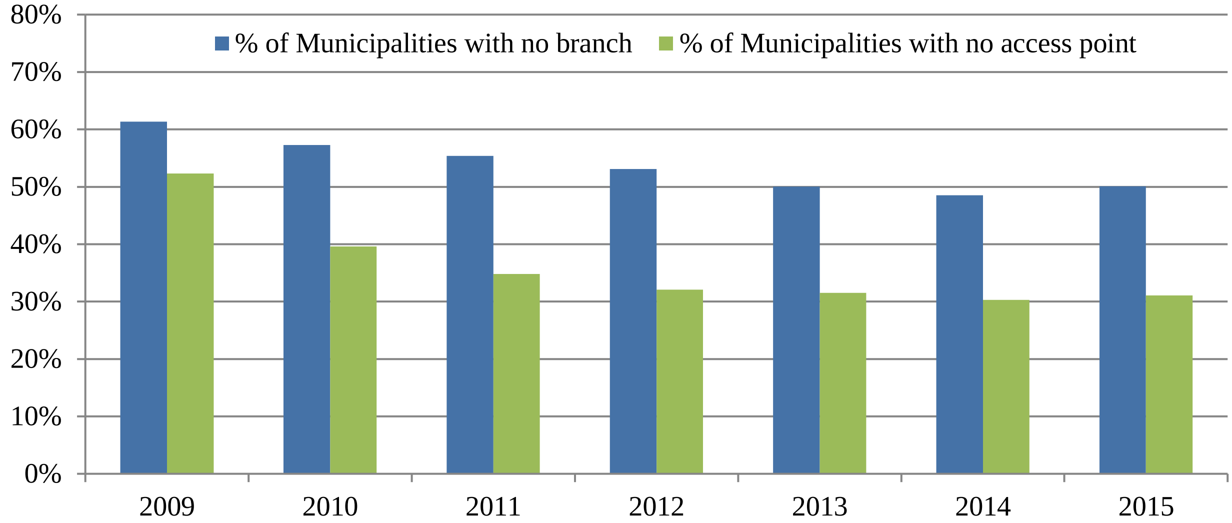
Figure 2.7: Half of the Mexican municipalities are without any bank branch



Source: CNBV, 2015

[31] This term bank branches includes branches of commercial banks, development banks and branches operated by microfinance institutions.

Figure 2.8: Financially excluded municipalities over time



Source: CNBV, 2015

2.2.3. Financial Inclusion Index (FII)

Financial inclusion is a multidimensional concept and it is important to note that there is still no universally accepted definition of financial inclusion or financial exclusion. The term financial inclusion was initially applied in the early 1990s and it was referred to as the access to bank branches in the context of liberalization of the financial sectors (Anand and Kuldip, 2013). But, over the years, many scholars, policymakers and international organizations have developed several definitions because of the evolving nature of various dimensions of financial inclusion. Various dimensions of financial inclusion include bank accounts, savings, personal and business loans, insurance, and payment facilities.

The literature on financial inclusion lacks a comprehensive indicator that can be used to gauge financial inclusion or exclusion in an economy in its entirety. Some of the indicators that are generally used in the literature either look at a) the penetration of access using a proxy measure such as ‘the number of bank accounts per capita’ or b) looking at availability and proximity using the proxy measure of ‘the number of banking units or ATMs per capita’ or c) measuring the extent and frequency of usage by looking at proxy measures such as ‘the number of loans or credit product per 10,000 adults’ (Sarma, 2008; Burgess and Pande, 2005; among others). In order to better understand the current status of financial inclusion in Mexico, we construct a dynamic Financial Inclusion Index (FII) using data from the National Banking and Securities Commission (CNBV) on the availability and usage of different financial services at the municipality level. Mexico has a total of 32 states which are further divided into 2,457 municipalities, as of 2015. CNBV provides detailed information about accessibility of financial services and the usage of different types of financial products at this very fine geographical level known as municipalities, and we exploit this information to come up with our index of financial inclusion.

The Financial Inclusion Index (FII) takes into account five different dimensions by focusing not only on the accessibility of formal financial institutions but also the actual usage of different financial products and services such as credit, saving accounts, and payment facilities. The five dimensions which make up our index include: i) accessibility of financial services, ii) depth of credit services, iii) concentration of checking accounts, iv) concentration of non-checking accounts, and v) usage of financial payment channels. Accessibility of financial services is measured by the total number of access points per 10,000 adults, including bank branches of formal financial institutions and banking agents. Depth of credit services is based on the number of credit products per 10,000 adults in each municipality; these credit products include personal loans, group loans, nominal loans, automobile loans, housing loans and consumer durables loans. Concentration of checking accounts refers to all types of checking bank accounts per 10,000 adults in each municipality while concentration of non-checking accounts refers to all types of non-checking bank accounts per 10,000 adults in each municipality. Finally, the usage of financial payment channels is measured by the number of transactions carried out using ATMs and cellular services per 10,000 adults in each municipality.

For each of the abovementioned financial inclusion dimensions, an index is created for each municipality using the following formula.

$$\mathbf{D_{im}} = \frac{\mathbf{A_{im} - Min_i}}{\mathbf{Max_i - Min_i}} \qquad \mathbf{(2.1)}$$

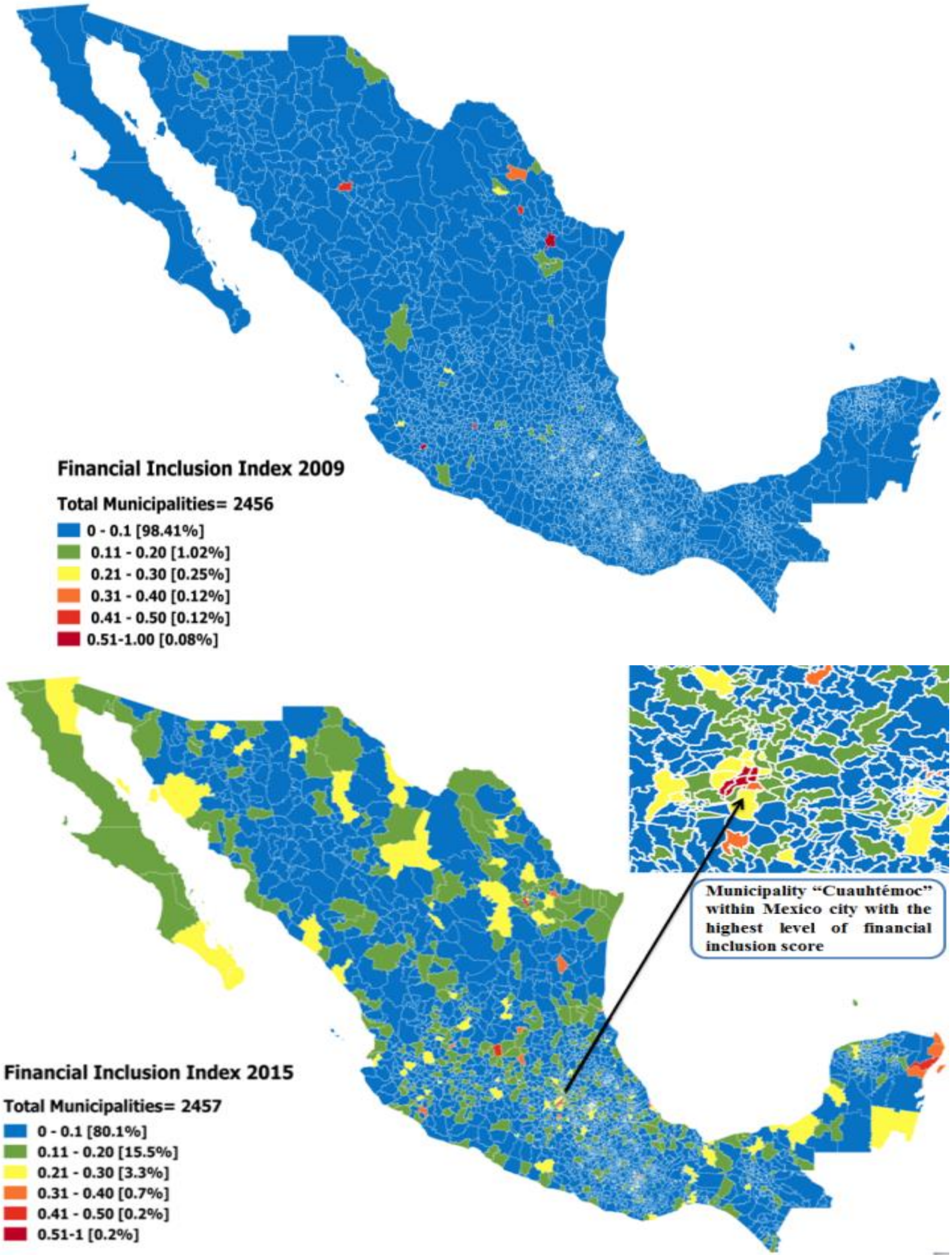
Where D_{im} is the index for indicator *i* for municipality *m*. A_{im} is the actual value of indicator *i* for municipality *m*. Min_i is the minimum value of indicator *i* and Max_i is the maximum value of indicator i. The FFI is the simple equally weighted average of the dimension indices, normalized to range between 0 and 1, where 0 refers to the lowest level of financial inclusion.

$$\mathbf{FII} = \frac{1}{n}\sum_{i=1}^{n} \mathbf{D_{im}} \qquad (2.2)$$

The allocation of weights is complex and a number of papers that have attempted to compute composite indices assign equal weights to all variables and dimensions. This is the case for those indices proposed by Sarma (2008) as well as Anand and Kuldip (2013). As a result, each normalized variable is implicitly considered as constituting a specific dimension (Amidžić et al., 2014) as indicated in equation 2.2. All these dimensions are considered as an essential part of financial inclusion and can play a role in promoting entrepreneurship. The accessibility of bank branches can allow access to financial services and it can even help non-user households through general equilibrium effects which can have an impact on the labor market (Breza and Kinnan, 2018). Similarly, a number of studies have argued that access to credit can promote entrepreneurship by reducing the business start-up costs and providing seed capital, particularly for people who do not have self-financing or external sources of financing (Klapper at al., 2006). Moreover, having access to bank accounts and savings can also help in promoting entrepreneurship through the investment-finance channel by allowing people to invest in their business ideas (Dupas and Robinson, 2013). Lastly, digital payment services can conveniently and affordably connect entrepreneurs with their suppliers, customers, employees, and new markets for businesses. Klapper (2017) highlights that these payment services can accelerate business registration and payments for business permits and licenses by significantly reducing transaction costs.

Results from FII are shown in Figure 2.9 highlighting that the state of financial inclusion has improved considerably from 2009 to 2015, although the current level of financial exclusion still remains quite high. According to the 2015 financial inclusion index, 80% of the municipalities have a score close to 0.1, meaning that financial exclusion still remains quite high. Municipality 'Cuauhtemoc', in the capital Mexico City, has the highest value of financial inclusion index highlighting that the level of financial inclusion in this municipality remains the highest as compared to the rest of the municipalities in Mexico.

Figure 2.9: Results of the Financial Inclusion Index 2009 and 2015



Note: Data in parenthesis represents the distribution of municipalities.
Source: Author's Calculations based on CNBV's data.

2.2.4. Entrepreneurship in Mexico

The term entrepreneurs is defined as self-employed people owning or co-owning a micro-firm or a small and medium enterprise (SME). This chapter uses the same definition of micro firms and SMEs that has been used by the Mexican National Survey of Occupation and Employment (ENOE) as provided in Table 2.1.

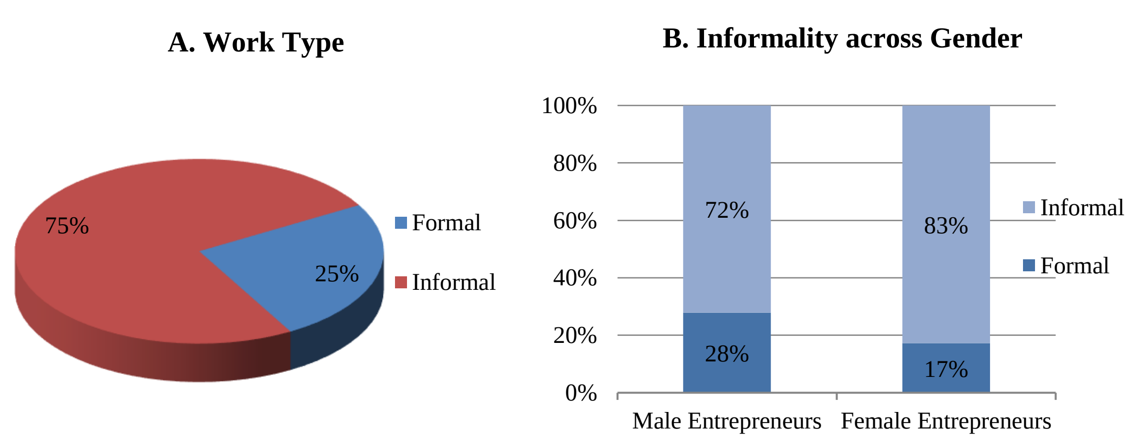
Table 2.1: Classification of firms by number of employees in selected sectors

	Industry	Commerce	Services
Micro firm	0-10	0-10	0-10
Small firm	11-50	11-30	11-50
Medium sized firm	51-250	31-100	51-100

Source: ENOE

Based on the definition of entrepreneurs, 21% of the active population in Mexico falls in the category of entrepreneurs, and women represent about 44% of those entrepreneurs, which is relatively high as compared to other OECD countries (OECD, 2016). Three out of four entrepreneurs work in the informal sector and there are more women entrepreneurs in the informal sector as compared to men entrepreneurs. Following the definition of informality by the International Labor Organization (ILO), informal businesses refer to organizations that are not registered with a local authority, do not pay taxes, and employ workers without any formal work contract and without any formal safety net in place. In the case of Mexico, about 83% of the women entrepreneurs work in the informal sector whereas 72% of the men entrepreneurs work in the informal sector (Figure 2.10).

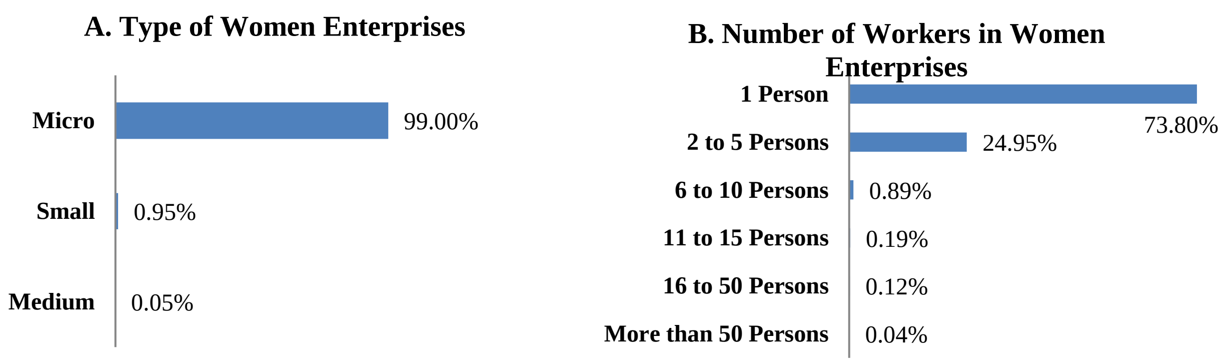
Figure 2.10: Informality and gender shares of Mexican entrepreneurs



Source: ENOE, 2015

Almost all women entrepreneurs belong to micro firms and approximately 99% of these women led enterprises have 5 or less workers (Figure 2.11). Furthermore, the average monthly income of women entrepreneurs is significantly less as compared to men entrepreneurs.

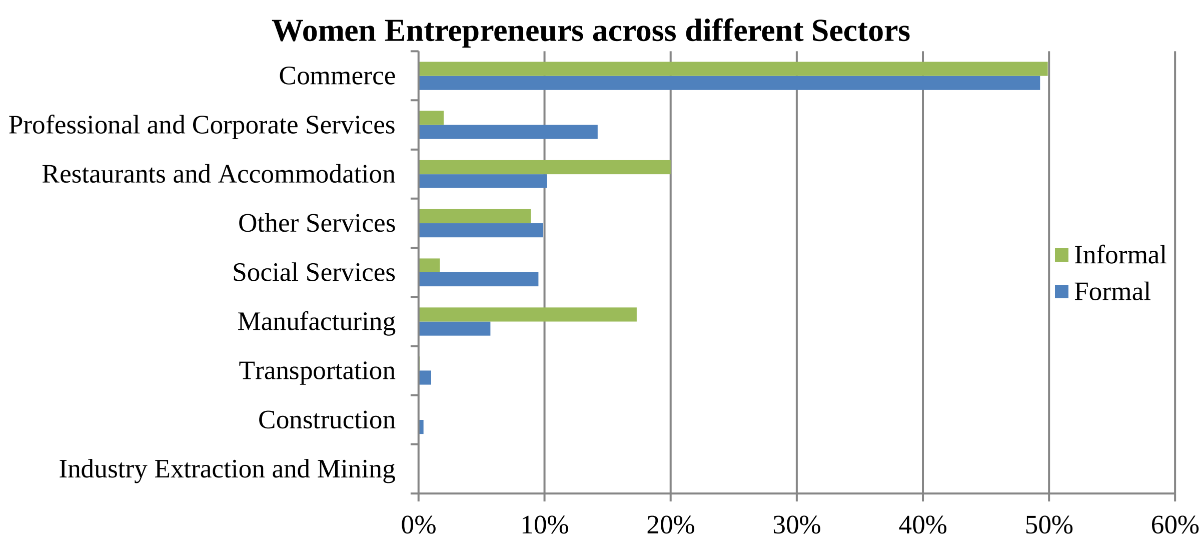
Figure 2.11: Women entrepreneurs mostly belong to micro-firms



Source: ENOE, 2015

Almost half of the women entrepreneurs work in the commerce sector. The other prominent sectors where women entrepreneurs work are restaurants and accommodations, services and manufacturing, respectively. The distribution of sectors where women entrepreneurs work informally differs from to where women work formally. This has been highlighted in Figure 2.12 below. There are more women entrepreneurs working informally in manufacturing and restaurants and accommodations sectors as compared to the formal sector. Conversely, in case of services, most of the women entrepreneurs work in the formal sector as compared to the informal sector.

Figure 2.12: Distribution of Sectors where Women Entrepreneurs Work



Source: ENOE, 2015

The concentration of women entrepreneurs also varies geographically across Mexican states, and it is quite different from that of men entrepreneurs. The state of Oaxaca, followed by Guerrero and Nayarit, has the highest percentage of women entrepreneurs and the majority of the women entrepreneurs in these areas work informally. States which are situated in the South of Mexico have a higher poverty rate and that is where informal women entrepreneurs are mostly concentrated. The maps in Figure 2.13 depict the presence of women entrepreneurs as compared to men entrepreneurs and provide an overview of the states where most of the women entrepreneur's work[32].

Figure 2.13: Share of men and women entrepreneurs across states

Note: Numbers in parenthesis show the percentage of states that belong to each category.
Source: ENOE, 2009-15

---

[32] An overview of where informal women entrepreneurs are located is provided in Appendix 2.4.

Note: Numbers in parenthesis show the percentage of states that belong to each category.
Source: ENOE, 2009-15

2.3. Empirical Methodology

This section presents the empirical methodology aimed at investigating the relationship between financial inclusion and women entrepreneurship alongside studying the key drivers of women entrepreneurship. In order to conduct this empirical analysis, a logistic regression model is used using a pseudo panel data i.e. different individuals are observed across time from 2009 to 2015. The overall sample includes people who fall in the active population category[33]. $Y_{\boldsymbol{i,m,s,t}}$ is the binary dependent variable in our main regression equation 2.3 and it represents the entrepreneurship status of the individual *i*, living in municipality *m* of state *s*, in time period *t* (entrepreneur= 1, otherwise = 0).

$$\mathbf{Y}^{*}_{\boldsymbol{i,m,s,t}} = \boldsymbol{\alpha} + \boldsymbol{B_1}(\boldsymbol{FII})_{\boldsymbol{m,s,t}} + \boldsymbol{B_2}\boldsymbol{X}_{\boldsymbol{i,m,s,t}} + \boldsymbol{B_3}\boldsymbol{M}_{\boldsymbol{m,s,t}} + \boldsymbol{B_4}\boldsymbol{D}_{\boldsymbol{s,t}} + \mu_{\boldsymbol{s}} + \mu_{\boldsymbol{t}} + \boldsymbol{\varepsilon}_{\boldsymbol{i,m,s,t}} \qquad (\mathbf{2.3})$$

$$Y_{\boldsymbol{i,m,s,t}} = 1 \text{ if } \mathbf{Y}^{*}_{\boldsymbol{i,m,s,t}} > 0$$
$$Y_{\boldsymbol{i,m,s,t}} = 0 \text{ if } \mathbf{Y}^{*}_{\boldsymbol{i,m,s,t}} \leq 0$$

[33] This includes people that are either working as employees, employers, entrepreneurs, or working without pay and looking for jobs.

$\mathbf{Y}^*_{i,m,s,t}$ is a latent variable. *FII* refers to the financial inclusion index. First, we estimate the main model with the financial inclusion index and later on we use different financial access indicators like bank branches, banking agents and POS terminals per 10,000 adults at the municipality level to establish the relationship of these indicators with entrepreneurship. $\boldsymbol{X}_{i,m,s,t}$ is the matrix of individual level characteristics, whereas $\boldsymbol{M}_{m,s,t}$ is the matrix of municipality level characteristics. $\boldsymbol{D}_{s,t}$ refers to the Ease of Doing Business indicator at the state level. The term $\mu_{\mathbf{t}}$ takes into account time fixed effects whereas fixed effects at the regional state level are absorbed by the term $\mu_{\boldsymbol{s}}$. The error term is represented by $\boldsymbol{\varepsilon}_{i,m,s,t}$ and it follows a logistic distribution. In order to control for the possible error correlation within municipalities, all regressions are estimated by clustering the standard errors at the municipality level. The main coefficient of interest which determines the link between entrepreneurship and financial inclusion index is B_1. The overall sample includes only women unless specified otherwise.

2.3.1. Explanatory Variables

In order to measure financial inclusion, the estimated Financial Inclusion Index (FII) has been used. FII ranges between 0 and 1, where 0 refers to the lowest level of financial inclusion and 1 refers to the highest level of financial inclusion. In addition to FII, we run the baseline regression model using other types of indicators for financial access as robustness tests. One of these indicators which is used to measure access to financial services is the number of bank branches of formal financial institutions per 10,000 adults in the municipality where the respondent lives. These formal financial institutions include commercial banks, development banks as well as microfinance institutions. Other variables that have been used to measure access to financial services are the 'number of banking agents per 10,000 adults' and 'POS terminal per 10,000 adults'. As explained before, an increase in the level of financial inclusion can improve the access to formal financial markets and can alleviate credit constraints, which can promote creation of new businesses.

Among the individual level characteristics in $\boldsymbol{X}_{imst}$, we have included a number of socio-economic variables that might have an effect on the entrepreneurship status of an individual. All these variables come from ENOE. The *urban* variable indicates whether the individual lives in an urban location or a rural location. The variable takes the value 1 if the individual is located in an urban area and 0 otherwise. There is a large gap in terms of development level and population density between urban and rural areas in Mexico, which can have an effect on starting a new business as indicated by Faggio and Silva (2014).

The education variable indicates the level of education of the individual. Each individual falls in one of the four categories: (1) if the individual has "primary or no education," (2) if “secondary education", (3) if "upper secondary or undergraduate education", and (4) if the person has a "masters or PhD" degree. Doms et al. (2010) show that education plays a critical role in enhancing an individual’s entrepreneurial skills and intentions. Therefore, we expect an increase in educational level to be positively related to entrepreneurship.

The variable *Marital* describes the current marital status of the individual. The categories include “in a relationship”, “separated”, “divorced”, “widowed”, “married”, “single”, and “Others/ Do not want to disclose”. The *Age* variable measures the age of the individuals in number of years. Experience is gained with age and it is likely to be positively related to entrepreneurship. In addition, the *Economic Sector* variable lists the sector of the individual where he or she works. We introduce sector fixed effects in order to control for any unobserved heterogeneity across sectors. Similarly, the *Informal* variable measures whether the individual works for the economic unit in the formal sector (0) or the informal sector (1).

Among municipality level characteristics in $\boldsymbol{M_{mst}}$, we include *employment rate* and the log of *average yearly income* in order to control for the activity level and development level of the municipality. Both these variables might affect entrepreneurship decision as well as the level of financial inclusion. In general, a more developed location might be more attractive for starting new businesses. We also speculate that the difference in regulations across states might have an effect on new business creation. Therefore, we include the ease of doing business indicator $\boldsymbol{D_{st}}$ that refers to the extent to which the regulatory environment is favorable for starting and operating a business in a state. A higher score refers to a state with the more favorable business environment.

The ease of doing business variable comes from the Ease of Doing Business Report by the World Bank and it is available only for the years 2010, 2012 and 2014[34].

Other than these characteristics, some might debate that entrepreneurs are different from non-entrepreneurs because of certain attributes like risk taking preferences, attitudes, family background, social networks, and other psychological traits. These attributes are not discussed in this chapter due to the lack of information on these characteristics. We acknowledge that this might induce a bias in our results. However, we run a number of additional tests to ensure the robustness of our results.

2.4. Results and Discussion

The main finding from the analysis shows that financial inclusion is positively linked with women entrepreneurship, meaning financial inclusion opens up economic opportunities for women entrepreneurs. The econometric estimates summarized in Table 2.2 allow us to examine how different individual and regional characteristics are related to the likelihood of being a woman entrepreneur. In case of logistic regression, the coefficients cannot be directly interpreted. So, in order to comment on the magnitude of the coefficients, we have calculated and reported the average marginal effects. Our results show that an increase in the financial inclusion index increases the probability of a woman being an entrepreneur after controlling for other individual and geographical characteristics. In the first column of Table 2.2, we estimate the econometric model with very basic specifications. We only include individual level controls such as age, education level, marital status, location type, and sector fixed effects along with time fixed effects. The results show a positive and statistically significant link between financial inclusion and women entrepreneurship. However, there is a lot of variation in terms of development level and infrastructure across states in Mexico which can influence entrepreneurial behavior. Moreover, the type of work i.e. formal or informal can also influence entrepreneurship decision. Therefore, in order to control for this heterogeneity across states, column 2 shows the result for the same regression including state fixed effects and the informal dummy. The result highlights that the

[34] This reduces the total number of observations for econometric estimates that include *business ease score* as an independent variable.

positive relationship between financial inclusion and women entrepreneurship still holds. An increase in the financial inclusion index increases the probability of a woman being an entrepreneur and the result is statistically significant at a 1% level. One can also argue that there are some time variant factors at the state level which have an effect on entrepreneurial behavior, e.g. ease of doing business and regulations amongst other factors. In order to control for that, we include the ease of doing business indicator in column 3[35]. We also add controls for the employment ratio and average yearly income at the municipality level which can influence entrepreneurship as well as the level of financial inclusion. The regression estimates still show that an increase in financial inclusion is positively associated with women entrepreneurship.

The results also indicate that the education level is a significant driver of women entrepreneurship. An increase in education increases the chances of becoming an entrepreneur as compared to having no or primary level education. Having an upper secondary education or undergraduate degree increases the probability of being an entrepreneur by about 6.5 percentage points and having a masters' or PhD increases this probability by about 16.6 percentage points, as compared to no or primary education. An increase in age is also positively linked with the probability of a woman being an entrepreneur. This result is expected because with age comes experience and it can influence the decision to become an entrepreneur. Moreover, residing in an urban area also seems to have a significant and positive effect on women entrepreneurship. Business ease score and employment rate at the municipality also come out to be statistically insignificant. However, an increase in income level at the municipality level increases the probability of a woman being an entrepreneur. The marital status has a significant effect on the entrepreneurial status as well. The results indicate that being married or being divorced increases the probability of a woman being an entrepreneur as compared to being in a relationship. Whereas, being separated, widowed or single decreases her probability of being an entrepreneur. The results also show that the probability of a woman being an entrepreneur in the informal sector is about 13.5 percentage points more as compared to being an entrepreneur in the formal sector.

[35] Please note that the information on ease of doing business variable is only available for three years between 2009-2015 (2010, 2012 and 2014) which decreases the number of observations in regressions which include this control.

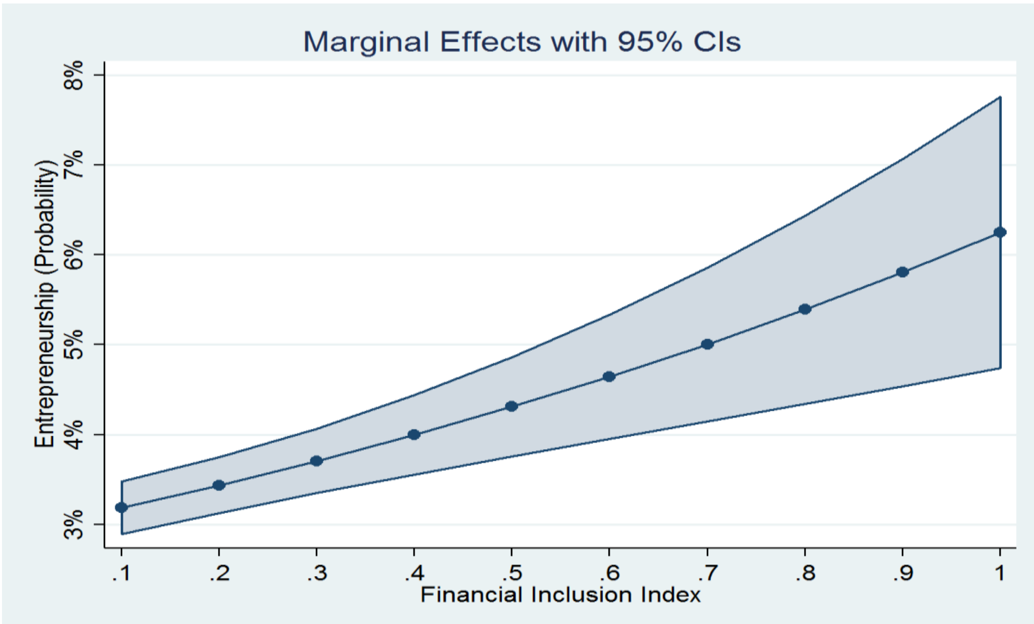
Table 2.2: Determinants of being a woman entrepreneur

VARIABLES	**Woman Entrepreneur (1)**	**Woman Entrepreneur (2)**	**Woman Entrepreneur (3)**
Financial Inclusion Index (FII)	0.021**	0.036***	0.041***
	(0.008)	(0.007)	(0.008)
Education Level (Base: Primary or No Education)			
Secondary Education	0.023***	0.022***	0.023***
	(0.001)	(0.001)	(0.001)
Upper Secondary or Undergraduate	0.048***	0.064***	0.065***
	(0.001)	(0.002)	(0.002)
Masters or PhD	0.133***	0.164***	0.166***
	(0.006)	(0.006)	(0.006)
Age	0.003***	0.003***	0.003***
	(0.000)	(0.000)	(0.000)
Urban Dummy	0.012***	0.005***	0.005**
	(0.003)	(0.002)	(0.002)
Marital Status (Base: In a Relationship)			
Separated	-0.013***	-0.011***	-0.010***
	(0.002)	(0.001)	(0.002)
Divorced	0.016***	0.019***	0.019***
	(0.003)	(0.002)	(0.003)
Widowed	-0.009***	-0.006***	-0.006***
	(0.002)	(0.002)	(0.002)
Married	0.012***	0.016***	0.016***
	(0.002)	(0.001)	(0.002)
Single	-0.041***	-0.024***	-0.024***
	(0.002)	(0.001)	(0.002)
Others	-0.023	0.014	0.067
	(0.025)	(0.025)	(0.061)
Informal Dummy		0.135***	0.135***
		(0.001)	(0.002)
Income level in Municipality (log)			0.009***
			(0.001)
Employment Rate in Municipality			-0.019
			(0.020)
Business Ease Score			0.004
			(0.003)
Sector Fixed Effects	**Yes**	**Yes**	**Yes**
State Fixed Effects	**No**	**Yes**	**Yes**
Time Fixed Effects	**Yes**	**Yes**	**Yes**
Observations	**1,518,183**	**1,518,183**	**554,659**
Wald (prob> chi2)	**0.00**	**0.00**	**0.00**
Pseudo R2	**0.60**	**0.68**	**0.68**

Note: Results of the logit regression have been reported with average marginal effects. The total sample includes only women. Standard errors are in parentheses (clustered at the municipality level). The dependent variable is the entrepreneurship status dummy (see equation 2.3). The number of observations in column (3) are less because the ease of doing business variable is only available for 3 years. ****, ** and * represent significance at 1%, 5% and 10%, respectively.*

In order to comment on the magnitude of the positive relationship between financial inclusion index and women entrepreneurship, we compute marginal effects at different thresholds of the financial inclusion index (using the regression in Column 3 of Table 2.2). The results are summarized in Figure 2.14. They highlight that an increase in the level of financial inclusion in the municipality can increase the chances of being a woman entrepreneur in the range of about 3% to 6%, after controlling for other individual and regional characteristics.

Figure 2.14: Financial Inclusion and Women Entrepreneurship - Marginal Effects



2.4.1. Estimation Results for Different Economic Sectors

Another objective of this chapter is to study the relationship between financial inclusion and women entrepreneurship across different economic sectors. In order to do that, we run the same regression model in column 3 of table 2.2 after dividing the total sample into different subsamples by economic sectors. The idea is to gauge whether an increase in financial inclusion increases the chances for women to become entrepreneurs across all economic sectors or not. The results show that the positive relationship between financial inclusion and women entrepreneurship holds for almost all the sectors, except the commerce sector. This means that for women working in manufacturing, services and other sectors except commerce, an increase in financial inclusion can allow women to start their own business and become entrepreneurs. However, in case of commerce, we get a statistically insignificant estimated coefficient. This result is in line with the existing literature which points out that the commerce sector offers fewer possibilities of developing a career as an entrepreneur, especially for women (Vejsiu, 2011). Moreover, it is important to note that women working in the commerce sector also earn less income on average and have less education as compared to women working in other sectors which might explain this result. The regression results for different economic sectors are provided in Table 2.3. All the control variables are not displayed in the table for ease of reading.

Table 2.3: Econometric estimates for different economic sectors

	Manufacturing	**Construction**	**Commerce**	**Restaurants & Accommodation**	**Transport and Communication**	**Professional, Financial and Corporate Services**	**Social Services**	**Other Services**
Financial Inclusion Index	0.05*** (0.02)	0.08** (0.03)	-0.03 (0.02)	0.04* (0.02)	0.03* (0.02)	0.03* (0.02)	0.02*** (0.01)	0.03** (0.01)
Individual Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
State Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	75,819	3,949	146,922	51,006	7,343	31,445	82,233	86,294
Pseudo R2	0.48	0.29	0.39	0.36	0.44	0.36	0.43	0.65

*Note: Results of the logit regressions, for different sub-samples have been reported with average marginal effects. Total sample includes only women. Standard errors are in parentheses (clustered at the municipality level). The dependent variable is entrepreneurship status dummy. Control variables have been included but not reported due to ease of reading. ***, ** and * represent significance at 1%, 5% and 10%, respectively.*

2.4.2. Estimation Results for Women engaged in Formal versus Informal Work

Table 2.4 examines the link between financial inclusion and women entrepreneurship by dividing the sample according to the type of location (urban or rural) and across work type (formal or informal). Financial inclusion indicators are generally positively linked with women entrepreneurship in the urban areas and in the formal sector. However, the positive relationship does not hold in case of women working in rural areas especially in the informal sector. These results can be explained by a number of factors. First, financial illiteracy might be a major reason why women working in the rural areas and informal sector do not benefit from the access and usage of financial services. Barriers to entry are lower in case of commerce and in other informal sectors as compared to the formal sector and this might be a reason why women with lower financial literacy tend to work in the informal sector and the commerce sector more. This also highlights the need for effective coaching and mentoring of women entrepreneurs, especially in rural areas and in the informal sector. Secondly, numerous documentation requirements to access financial services make the usage of formal financial services less attractive for informal and rural businesses. Other reasons might include gender biases in credit decisions and a slightly lower demand for formal financial services.

Table 2.4: Econometric estimates for work type and location type for women

	Financial Inclusion Index
Women in the Formal Sector	0.05*** (0.01)
Women in the Informal Sector	0.03** (0.01)
Women in Urban Areas	0.03*** (0.01)
Women in Rural Areas	0.03 (0.04)
Women in Rural Areas and Informal Sector	-0.01 (0.09)
Women in Urban Areas and Informal Sector	0.04*** (0.01)

*Note: Results of the logit regressions, for different sub-samples have been reported. The overall sample includes only women. Standard errors are in parentheses (clustered at the municipality level). The dependent variable is entrepreneurship status dummy. Control variables have been included but not reported due to ease of reading. ***, ** and * represent significance at 1%, 5% and 10%, respectively.*

2.4.3. Gender Gap in Formal versus Informal Entrepreneurship

In this subsection, we focus on understanding the gender gap in the entrepreneurship status across formal and informal work in Mexico. Therefore, for this analysis, our overall sample includes women as well as men. Our econometric estimates point towards a significant gender disparity in the entrepreneurship status across formal and informal sector (Table 2.5). In the formal sector, women are about 3 percentage points less likely to be entrepreneurs as compared to men. However, in the informal sector, women are about 11 percentage points more likely to be entrepreneurs as compared to men, after controlling for other individual and regional characteristics. This is an interesting finding that can be explained by a number of factors. Women in developing countries are less able to compete with men in the formal sector because they are less likely than men to own property or have a better know-how of the market or have higher level of skills and education (Ramani et al., 2013). Moreover, Georgellis and Wall (2005) argue that entrepreneurship is a closer substitute for labor market inactivity and part-time work for women than it is for men, which might also explain this gender disparity. Since a large proportion of informal women entrepreneurs work as home-based producers and street venders, another reason that might explain why women choose informal sector over the formal sector might revolve around a higher need for a stronger work-life balance amongst women (Gurley-Calvez et al., 2009). Lastly, Chen (2005) highlight that social and cultural norms are probably behind this gender disparity which assign the responsibility for social reproduction to women and discourage investment in their education and training.

Table 2.5: Gender gap in entrepreneurship across formal and informal work

	Entrepreneur (Formal Sector)	Entrepreneur (Informal Sector)
Female -Dummy	-0.03*** (0.001)	0.11*** (0.004)
Other Controls	**Yes**	**Yes**
Sector Fixed Effects	**Yes**	**Yes**
State Fixed Effects	**Yes**	**Yes**
Time Fixed Effects	**Yes**	**Yes**
Observations	**2,934,731**	**1,067,918**
Pseudo R2	**0.47**	**0.57**

*Note: Results of the logit regressions have been reported with average marginal effects. The total sample includes men and women. Standard errors are in parentheses (clustered at the municipality level). The dependent variable is entrepreneurship status dummy. All the control variables have been included but not reported due to ease of reading. ***, ** and * represent significance at 1%, 5% and 10%, respectively.*

2.4.4. Results using different Financial Access Indicators

We estimate a number of different specifications by using a range of indicators for financial access such as concentration of bank branches, POS terminals and banking agents, mainly as a robustness check. The results show that there is a positive and significant relationship between most of these indicators and entrepreneurship. These results suggest that different financial access points like bank branches, POS terminals and banking agents can be a gateway to the use of additional financial services which can allow businesses development. However, as highlighted before, this positive relationship between these financial access indicators and being an entrepreneur does not hold for women entrepreneurs working in rural areas. A summary of the results of the estimated specifications using different indicators of financial inclusion for different segments of the population is provided in Appendix 2.6.

2.4.5. Addressing Endogeneity Concerns

Ruiz-Tagle and Vella (2015) argue that the location decisions of financial institutions are driven by local aggregates such as commercial activity, population density, socioeconomic and political factors. These decisions do not reflect any aspect of individual level demand for a specific type of work and should not propagate a simultaneity issue between entrepreneurship status of an individual and our financial inclusion index that is measured at the municipality level. However, in order to make sure that the endogeneity bias is not present and that the relationship between financial inclusion and entrepreneurship is not spurious, we rely on two different methods.

First, we use the time-lagged structure for our financial inclusion variable to avoid the issue of simultaneity with entrepreneurship. It is unlikely that an individual's choice of work, such as being an entrepreneur, will affect municipality's financial inclusion level for the last year. Therefore, we re-estimate our baseline regression after using the financial inclusion indicator for the past quarter and the past year. The results have been reported in column 1 and column 2 of Table 2.6 respectively and they show that financial inclusion and women entrepreneurship are positively and significantly linked.

Secondly, we conduct an instrument variable estimation to address the potential endogeneity issue. The instrument that we use is the concentration of bank accounts, measured as the number of bank accounts per 10,000 adults, in the neighboring municipalities. In order to obtain this information, we first identify the neighboring municipalities for each one of the 2,457 municipalities using GIS mapping. Some municipalities turn out to have three or four neighbors whereas a few municipalities have even more than ten neighbors. Once we obtain this information on the neighbors, we calculate the average of bank accounts concentration in these neighboring municipalities. The assumption is that financial institutions are likely to take into consideration the concentration of financial inclusion and competition in neighboring municipalities before providing financial access in a municipality. This argument about financial institutions relying on the experience and policies from the neighbors is backed by the literature and Allen et al. (2016) rely on a similar IV to instrument for their financial inclusion indicator. Regarding the exclusion restriction, we believe that the instrument variable only affects entrepreneurship through our financial inclusion index. Our first stage estimation results clearly show that the average concentration of bank accounts in the neighboring municipalities is significantly associated with our financial inclusion index (Table 2.8). The results of the F-statistic from the first stage is also well above the commonly used cut-off of 10, indicating that the instrument is valid (Table 2.8). As we rely only on one instrument, over-identification cannot be tested in this case. The results shown in Table 2.7 highlight that the relationship between financial inclusion and women entrepreneurship remains significant. One main difference between the IV estimation and our baseline estimation is that the IV result is statistically significant at 5% as compared to the baseline model, which is significant at a 1%.

Table 2.6: Econometric estimates using time-lag for the financial inclusion index

VARIABLES	**Woman Entrepreneur (1)**	**Woman Entrepreneur (2)**
Financial Inclusion Index (t-1)	0.034***	
	(0.006)	
Financial Inclusion Index (t-4)		0.040***
		(0.007)
Education Level (Base: Primary or No Education)		
Secondary Education	0.022***	0.022***
	(0.001)	(0.001)
Upper Secondary or Undergraduate	0.064***	0.065***
	(0.002)	(0.002)
Masters or PhD	0.163***	0.165***
	(0.006)	(0.006)
Age	0.003***	0.003***
	(0.000)	(0.000)
Urban Dummy	0.006***	0.005***
	(0.002)	(0.002)
Informal Dummy	0.135***	0.135***
	(0.001)	(0.002)
Marital Status (Base: In a Relationship)		
Separated	-0.011***	-0.011***
	(0.001)	(0.002)
Divorced	0.019***	0.019***
	(0.003)	(0.003)
Widowed	-0.007***	-0.006***
	(0.002)	(0.002)
Married	0.016***	0.016***
	(0.001)	(0.001)
Single	-0.024***	-0.023***
	(0.001)	(0.001)
Others	0.015	0.028
	(0.025)	(0.028)
Income level in Municipality (log)	0.009***	0.008***
	(0.001)	(0.001)
Employment Rate in Municipality	0.003	0.002
	(0.015)	(0.016)
Sector Fixed Effects	**Yes**	**Yes**
State Fixed Effects	**Yes**	**Yes**
Time Fixed Effects	**Yes**	**Yes**
Observations	**1,454,505**	**1,264,523**
R-squared	**0.6840**	**0.6842**

*Note: Results of the logit regressions have been reported with average marginal effects. The overall sample includes all women. Standard errors are in parentheses (clustered at the municipality level). The dependent variable is entrepreneurship status dummy. ***, ** and * represent significance at 1%, 5% and 10%, respectively.*

Table 2.7: Instrument variable (IV) estimation

VARIABLES	**Woman Entrepreneur (1)**
Financial Inclusion Index (Instrumented)	0.072**
	(0.035)
Education Level (Base: Primary or No Education)	
Secondary Education	0.030***
	(0.002)
Upper Secondary or Undergraduate	0.082***
	(0.003)
Masters or PhD	0.103***
	(0.004)
Age	0.004***
	(0.000)
Urban Dummy	0.006
	(0.005)
Marital Status (Base: In a Relationship)	
Separated	-0.012***
	(0.002)
Divorced	0.011***
	(0.003)
Widowed	-0.001
	(0.002)
Married	0.017***
	(0.002)
Single	-0.018***
	(0.002)
Others	0.048
	(0.045)
Informal Dummy	0.346***
	(0.005)
Employment Rate in Municipality	0.032
	(0.027)
Income level in Municipality (log)	0.010***
	(0.001)
Sector Fixed Effects	**Yes**
State Fixed Effects	**Yes**
Year Fixed Effects	**Yes**
Observations	**1,518,183**
Pseudo R2	**0.6488**
Wald (p-value)	**0.00**

*Note: Results of the instrument variable probit regression have been reported with average marginal effects. The overall sample includes all women. Standard errors are in parentheses (clustered at the municipality level). The dependent variable is entrepreneurship status dummy. ***, ** and * represent significance at 1%, 5% and 10%, respectively.*

Table 2.8: First-stage regression results

Variables (First Stage of IV Estimation)	**(1)** Financial Inclusion Index
Average Bank Accounts Penetration in Neighborhood	0.46**
	(0.19)
Other Controls	**Yes**
State Fixed Effects	**Yes**
Year Fixed Effects	**Yes**
Total Observations	**1,518,183**
Prob > F	**0.00**
R- Squared	**0.2905**
F-test of excluded instruments	**16.07**
*** p<0.01, ** p<0.05, * p<0.1	

*Note: Results of the first stage of the instrument variable regression have been reported. Standard errors are in parentheses (clustered at the municipality level). Control variables have been included but not reported. ***, ** and * represent significance at 1%, 5% and 10%, respectively.*

2.5. Robustness Checks

In order to check for the significance of the financial inclusion indicator in our model, we conducted the 'Akaike information criterion (AIC)' test. The result of this test advocates that the models including the financial inclusion indicators fit the data better (Appendix 2.7). Moreover, since the economic profile of an individual can influence the choice to become an entrepreneur, we run the baseline regression with an additional control by including the monthly income level of the individuals in the model. We repeat the regression in Column 3 of Table 2.2 and the results suggest that the relationship between financial inclusion index and entrepreneurship remains significant (Appendix 2.8). Furthermore, we do a predictivity analysis by predicting the probability of woman being an entrepreneur using the list of independent variables in our model. Then we calculated the percentage of women whose entrepreneurship status has been predicted accurately by the model. Using different thresholds for probabilities, we find that the predictive ability of the model, i.e. the goodness of fit is very high, and accurate predictions go as high as 91% of the total sample. Lastly, in order to check for the significance of the financial inclusion indicator in our model, we also conducted the likelihood ratio test. This is a hypothesis test which compares the goodness-of-fit of two models, i.e. an unconstrained model with all the parameters and a constrained model with a fewer parameters (Casella and Berger, 2001). In our case, the constrained model is the one without the financial inclusion indicator and the unconstrained model is the one with all the explanatory variables including the financial inclusion indicator. The likelihood ratio test estimates the two models and compares the fit of one model to another. Having fewer explanatory variables most of the times makes the model fit less, but the likelihood ratio test assesses whether the observed difference in the model fit is statistically significant or not. The result of the test comes out to be significant at 1%. It highlights that the difference between these two models is significant and the model with the financial inclusion indicator fits the data significantly better.

2.6. Conclusion and Policy Recommendations

In this chapter, we proposed a financial inclusion index and investigated the key drivers of women entrepreneurship in Mexico using a large pseudo-panel of micro-level data from 2009 to 2015. Our results revealed that an increase in the level of financial inclusion is positively linked with women entrepreneurship and it can open up economic opportunities. We found that various financial access points like bank branches, POS terminals and banking agents are positively linked with women entrepreneurship. Other than financial inclusion indicators, an increase in education, age, and income level at the municipality level increase the chances of women becoming entrepreneurs. The marital status of women also has a significant effect on the probability of being an entrepreneur. Results also highlight that the probability of a woman being an entrepreneur in the informal sector is significantly higher than in the formal sector.

The results also show a clear gender gap in case of the entrepreneurship status across formal and informal work in Mexico. After controlling for individual level and municipality level characteristics, our results showed that women are less likely to be entrepreneurs in the formal sector and more likely to be entrepreneurs in the informal sector, as compared to men. Another important finding is that the positive relationship between women entrepreneurship and the financial inclusion index and other access indicators does not hold for women entrepreneurs working in the rural areas especially in the informal sector, highlighting the existence of lower entry barriers, including financial, in the informal sector.

The issues pertaining to financial inclusion and women entrepreneurship are crucial for policymakers because of the impact that they can have on job creation and women empowerment. Given that high gender gaps exist in Mexico, our results highlight that improving women financial inclusion can have an impact on job creation for women by lowering financial constraints for them. Lastly, our results also support the idea that with specific policies, targeting certain economic sectors, financial inclusion can also support formalization of women entrepreneurs in Mexico.

The literature has also highlighted that an increase in financial inclusion can improve the efficiency of public social spending which in turn can increase entrepreneurial activity (Loukoianova et al., 2018; Islam, 2015). Therefore, programs in Mexico such as *Programa Integral de Inclusión Financiera*, which promote bank accounts opening for women participating in conditional cash transfer programs, might also be able to play a role in promoting entrepreneurship.

Our analysis underlines that women face constraints that hinder their access and usage of financial services. Several of these constraints emerge from the supply side, as product design and service delivery might not be tailored according to their needs. Scarcity of granular data, segregated by gender, is another main reason that limits the underpinning of a comprehensive assessment on the financial needs of women. This asks for policy responses which can deal with these gaps. Such policy responses can be broadly classified into the following three categories.

The first category includes technology driven policy responses. Relying on technology related interventions with regard to account opening and credit expansion is a possibility worth exploring to enhance financial inclusion for women. Regarding bank accounts, although there are regulations in place regarding know your customers (KYC) and anti-money laundering norms, it is possible to come up with effective ways to deliver remote access for opening bank accounts. Some examples of this include opening bank accounts through banking agents or through tablet banking. Since women in rural areas, especially the ones who are part of the informal sector, have limited formal documentation, it is worth considering expanding the range of identification documents. Moreover, innovative digital technology can help with things such as biometrics and finger printing to reach far off rural areas. One such example of this is the use of tablet banking by the UKaid *Sakchyam Access to Finance* program in Nepal to reach out to the rural population. Similar initiatives are now taken by financial institutions in other developing countries such as India and Bangladesh. The initial findings of these case studies seem to suggest that these initiatives can help reduce the operational costs of financial institutions and can also improve the quality of customer service (Bazarbash, 2019). Similarly, with regards to credit expansion, innovative technology can be used to reduce operational cost of financial institutions. One such example of this are the financial institutions who rely on machine learning and big data, rather than conventional credit risk ratings, to provide credit to customers with little or no credit history. *Tex Financial Services* is one such example which relies on machine learning algorithms to determine an individual's financial ability in Pakistan to provide them with microcredit[36]. In summary, creating a regulatory environment that encourages technological innovation in the financial section can help boost financial inclusion especially for women.

The second category of policy responses includes institutional driven responses. The role of institutions such as central banks and credit bureaus is quite important when it comes to devising policies that foster financial inclusion. In recent years, credit bureaus in several African countries including Ghana and Uganda lowered the minimum loan amount that can be offered by financial institutions. This proved to be an important move in enhancing financial inclusion for women because women generally make use of smaller loans as compared to men (Beck, Behr and Madestam, 2015). Similarly, dedicated financial institutions that mainstream gender consideration in the provision of financial services can also help foster financial inclusion. In certain developing countries, where laws and culture tend to discriminate women from having collateral such as land, women are generally excluded from the use of formal financial sector. In order to address such challenges, certain financial institutions target their financial services (such as loans and insurance) mainly to women without requiring any type of physical collateral from them. One such example is the *Kashf Foundation*, a microfinance institution which provides financial services primarily to women. More than 90% of their clients are women and the majority of their loan officers are also women. These types of institutions also contribute towards enhancing financial inclusion for women. Moreover, designing financial products while keeping in mind the specific needs of women can go a long way in improving financial inclusion for women.

The third category of policy responses focuses on data-driven measures. Lack of granular data, disaggregated by gender, acts as a bottleneck when it comes to developing products and services that are tailored to the needs of women. Central banks around the world are increasingly recognizing this issue and are engaging with financial institutions to collect data that is disaggregated by gender. This step can be helpful in improving the status of financial inclusion for women.

[36] See https://www.tezfinancialservices.pk/ for more details.

Acknowledgements

The authors thank both the reviewers, David Halabisky and Naima Smaini, for their valuable feedback and comments. Feedback from the research seminar participants at the OECD, the Ministry of Finance and Public Credit of Mexico in Mexico City, and the 5th European Research Conference on Microfinance in England is also gratefully acknowledged.

2.7. Appendix

Appendix 2.1: Summary of Related Empirical Evidence

Study	Region	Type of Analysis	Dimension of Financial Inclusion	Development Indicators (Dependant Var.)	Effect
Angelucci, Karlan and Zinman (2015)	Mexico	Randomized Control Trial: Average Intent-to-Treat Effects with OLS Equations	Expansion of microfinance by Compartamos Banco	Well-being	+
				Business Growth	+
				Female household decision power	+
				Profit	0
				Household Income & Household Consumption	0
Augsburg et al. (2015)	Bosnia and Herzegovina	Randomized Control Trial: Comparison of means using OLS	Microcredit	Household Income	0
				Consumption	0
				Business Creation and Survival	+
				Enterprise Profits	0
				Health	0
				Education	0
				Women Empowerment	0
				Poverty	0
				Business Sales	+
				Savings	+
Attanasio et al. (2015)	Mongolia	Randomized Field Experiment: Intent to Treat Estimates	Microcredit	Food Consumption	+
				Entrepreneurship	+
				Income	0
Banerjee, Duflo, Glennerster and Kinnan (2015)	Hyderabad, India	Randomized Control Trial: Intent-to-treat (ITT) estimates	Microcredit	Health	0
				Education	0
				Women Empowerment	0
				Business Profit and Investment	+
				Business Creation	0
				Poverty	0
Samargandi et al. (2015)	52 Middle Income Countries	Pooled Mean Group Estimator in a Dynamic Heterogeneous Panel Setting (1980-2008)	Financial Development	Economic Growth	+ (Inverted U shaped)
Bruhn and Love, 2014	Mexico	Regressions using Difference in Difference strategy (2000-2004)	Financial Access	Entrepreneurial activity	+ (informal businesses for men)
				Employment	+ (only for women)
				Income levels	+
Brune et al. (2013)	Malawi	Randomized Control Trial	Commitment Savings	Business Investment	+
				Business Output	+
Dupas and Robinson (2013)	Kenya	Randomized Control Trial: Intent-to-treat (ITT) estimates	Micro Savings	Labor Supply (Number of hours worked per day)	0
				Business Investment	+
				Income	+
Karlan and Zinman (2010)	South Africa	Randomized Control Trial: Intention-to-treat (ITT) and treatment-on-the-treated (TOT) effects	Access to Credit	Well-being	+
				Food consumption	+
				Economic self-sufficiency	+
				Overall Health	+

				Level of Stress	+
				Income	+
Bianchi (2010)	Cross Country Analysis- 46 countries	Ordered Probit Regressions (Data: 1981-2000)	Financial Development	Job Satisfaction of Entrepreneurs	+
Lluss´a (2009)	Cross Country Analysis: 41 countries	Probit Regression (Data: 2001- 2004)	Financial Development	Need based Entrepreneurship	+
				Bridging Gender Gap in Entrepreneurship	0
Beck, Demirguc-Kunt and Levine (2007)	Cross Country Analysis: 72 countries	Generalized-methods-of-moments (GMM) panel estimator (1960- 2005)	Financial Development	Income Inequality	-
				Poverty	-
Clarke, Xu and Zou (2006)	Cross Country Analysis: 83 countries	OLS and 2SLS (Data: 1960-1995)	Financial Development	Income Inequality (Gini Coefficient)	-
Burgess and Pande (2005)	India	OLS Regression (State level Panel Data 1961-2000)	Financial Access (Rural Branch Expansion)	Poverty	-
				Economic Growth (Total per capita output)	+

Note: No effect is denoted by '0', positive effect by '+' and negative effect by '-'.

Appendix 2.2: Source and Coverage of Variables

Variable	**Description**	**Source & Coverage**
Entrepreneur	This is the dependent variable which is binary. It indicates whether a person is an entrepreneur (1) or not (0). The definition of entrepreneur refers to self-employment i.e. people owning or co-owning a micro-firm or a small and medium enterprise (SME). These businesses can operate with multiple employees or with the entrepreneur being the only worker in the business. This data is from ENOE.	ENOE (2009- 2015)
Age	This variable measures the age of the respondent in years.	ENOE (2009- 2015)
Average Monthly Income	Average monthly income measures the respondent's income in current Mexican pesos.	ENOE (2009- 2015)
Education Level	Education is a categorical variable which indicates the level of education of the person. It is 0 if the person has "primary or no education," 1 if "secondary education", 2 if "Upper secondary or Undergraduate", and 3 if the person has "Master or PhD".	ENOE (2009- 2015)
Gender	Gender is a binary variable with 1 referring to women and 0 referring to men.	ENOE (2009- 2015)
Economic Sector	This indicator lists the sector of the respondent where he or she is working. Sectors include "agriculture, forestry, hunting and fishing", "extractive industry and electricity", "manufacturing industry", "construction", "commerce", "restaurants and accommodation services", "transport, storage and communications mail", "professional, financial and corporate services", "social services", "various other services" and "government and international organizations".	ENOE (2009- 2015)
Informal Work	This variable measure whether the respondent works for the economic unit in the formal sector (0) or the informal sector (1).	ENOE (2009- 2015)
Urban	This variable presents the location type of the respondent. It is 1 if the respondent lives in an urban area and 0 if he or she lives in rural area.	CNBV (2009- 2015)
Marital Status	This variable describes the current marital status of the individual. The categories include "in a relationship", "separated", "divorced", "widowed", "married", "single", and "Others/ Do not want to disclose".	ENOE (2009- 2015)
Size of Enterprise	This variable indicates the total number of employees in the company where the respondent works. The categories are defined as follows: "1 person", "2-5 persons", "6-10 persons", "11-15 persons", "16-50 persons" and "more than 50 persons".	ENOE (2009- 2015)
Business Ease Score	This indicator shows the extent to which the regulatory environment is favorable for starting and operating a business in a state. This score ranges from 0 to 1. 0 refers to a state with	World Bank (2010, 2012, 2014)

	the most favorable business environment and 1 refers to a state with the least favorable business environment.	
Income level in the Municipality	This indicator provides the average yearly income of people living in that municipality in Mexican pesos to control for economic activity in the municipality.	ENOE (2009- 2015)
Employment Rate in Municipality	This indicator provides the employment rate of the municipality.	ENOE (2009- 2015)
Financial Inclusion Indicators	In order to measure financial inclusion, the estimated Financial Inclusion Index (FII) has been used along with different other types of indicators for financial access. One of the indicators which is used to measure access to financial services is the number of bank branches per 10,000 adults in the municipality where respondent lives. Other variables that have been used to measure access to finance are 'number of banking agents per 10,000 adults' and 'POS terminal per 10,000 adults'.	CNBV (2009- 2015)

Appendix 2.3: Descriptive Statistics

Variable (for women)	Observations	Mean	Standard Deviation	Min	Max
Entrepreneur Dummy	1,577,198	.2303953	.421086	0	1
Financial Inclusion Index	1,577,198	.1635192	.1092814	0	1
Branches per 10,000 adults	1,577,198	2.13317	1.164752	0	72.84768
Education: Primary or No Education Dummy	1,577,198	.2530988	.4347872	0	1
Education: Secondary Education Dummy	1,577,198	.2512792	.4337489	0	1
Education: Upper Secondary or Undergraduate Education Dummy	1,577,198	.4756999	.4994093	0	1
Education: Masters and PhD Dummy	1,577,198	.019922	.1397325	0	1
Urban Dummy	1,577,198	.8471149	.3598768	0	1
Informal Dummy	1,577,198	.27226	.4451232	0	1
Marital Status: In a relationship	1,577,198	.1182388	.3228908	0	1
Marital Status: Separated	1,577,198	.0684613	.2525359	0	1
Marital Status: Divorced	1,577,198	.0335278	.1800103	0	1
Marital Status: Widowed	1,577,198	.0482672	.2143305	0	1
Marital Status: Married	1,577,198	.3961151	.489089	0	1
Marital Status: Single	1,577,198	.3353346	.4721074	0	1
Marital Status: Others	1,577,198	.0000552	.0074269	0	1
Age	1,577,198	38.01204	13.47166	15	98
Monthly Income (log)	1,577,198	6.010921	3.597481	0	13.71015
Income level in Municipality (log)	1,577,187	10.12023	.4022077	4.945308	11.20321
Employment Rate in Municipality	1,577,198	.9504539	.0218196	.5	1
Business Ease Score	576,492	1.121549	.4994135	.0861552	2.222678

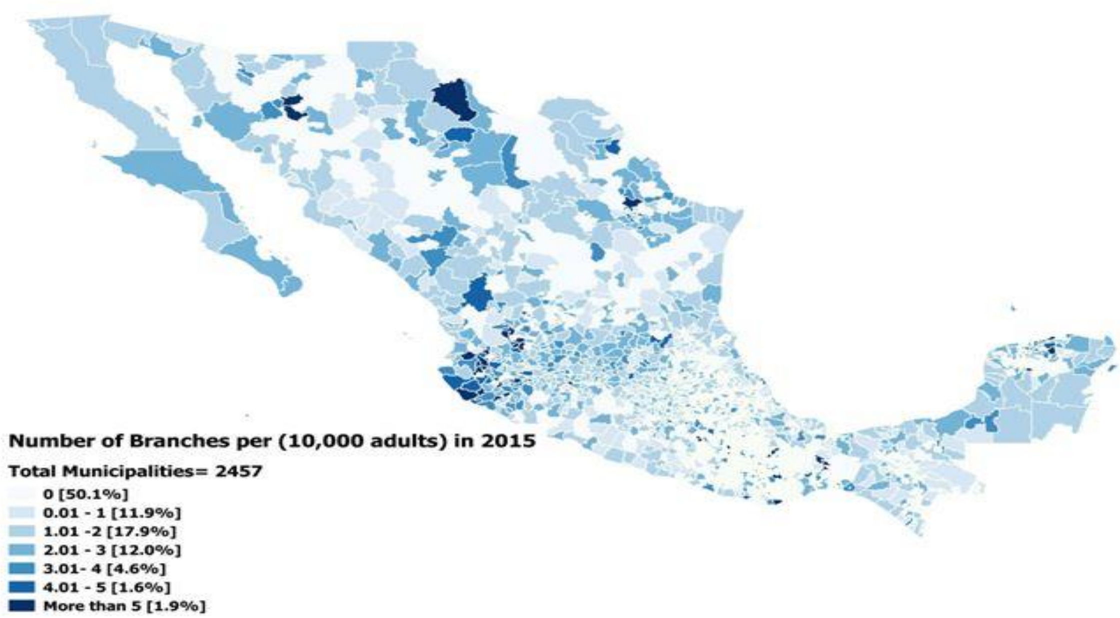
Source: ENOE

Appendix 2.4: Maps on Entrepreneurship and Financial Inclusion

Share of informal women entrepreneurs across states

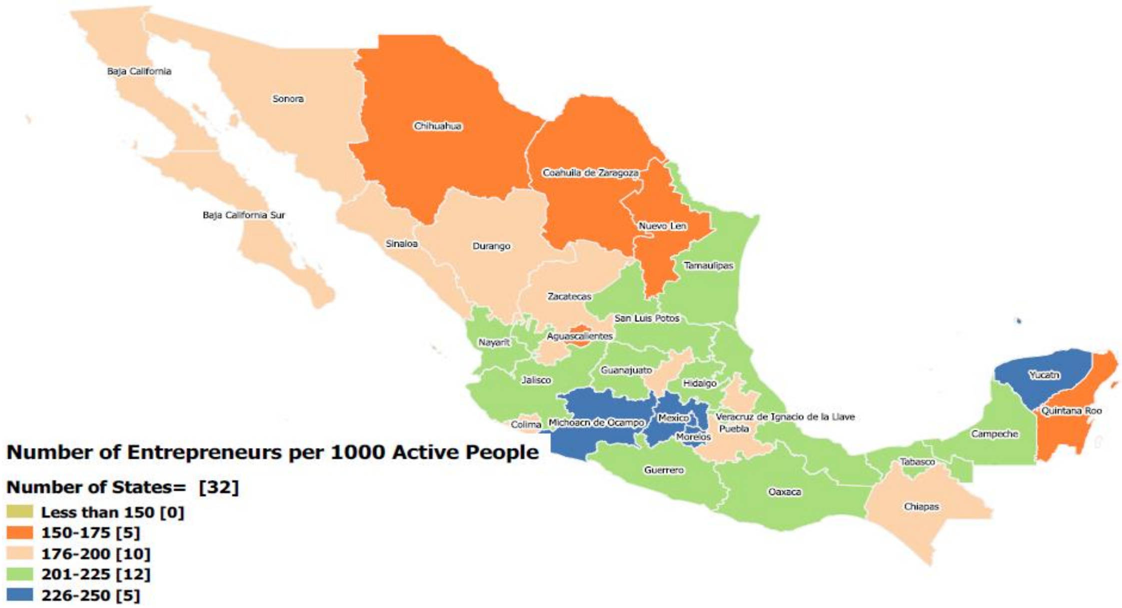
Note: Numbers in parenthesis show the percentage of states that belong to each category.
Source: ENOE, 2009-15

Concentration of bank branches across municipalities

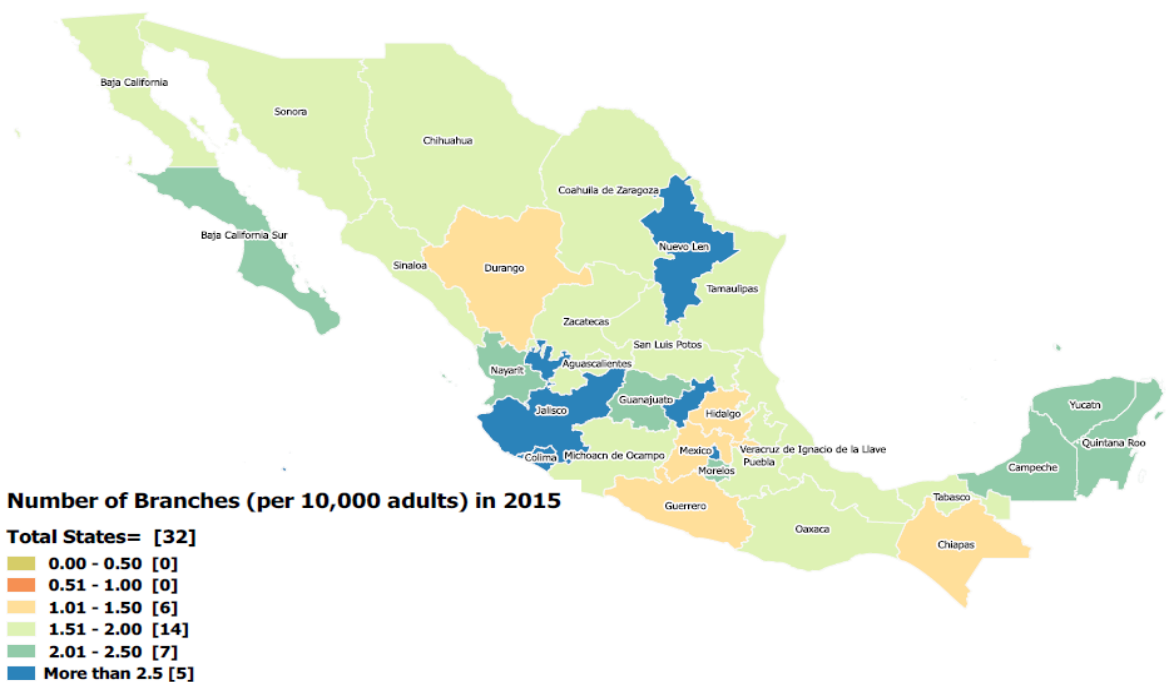


Note: Numbers in parenthesis show the percentage of municipalities that belong to each category.
Source: INEGI.

Concentration of entrepreneurs across states



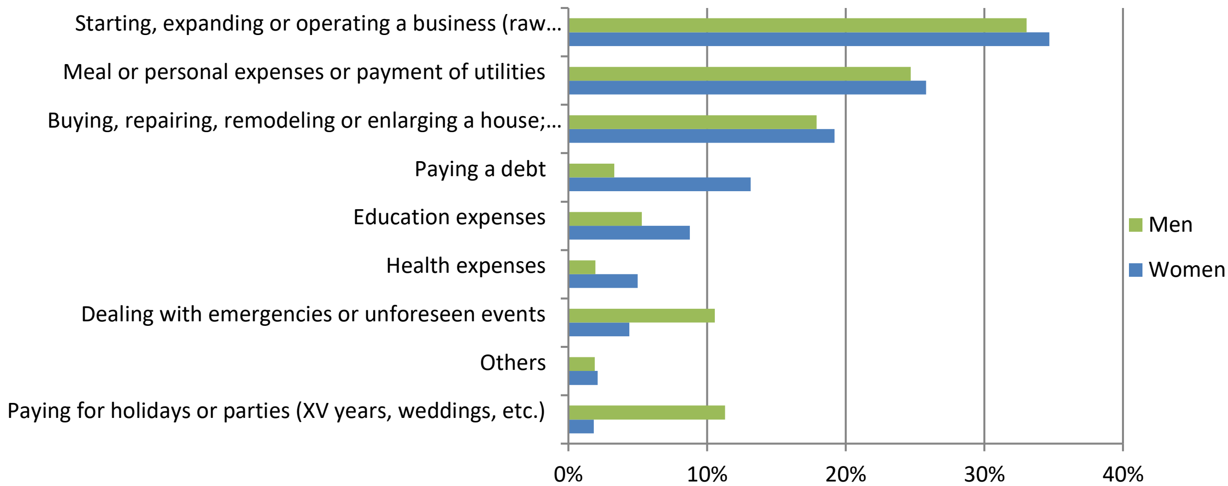
Concentration of branches across states



Note: Numbers in parenthesis show the percentage of states that belong to each category.
Source: ENOE, INEGI.

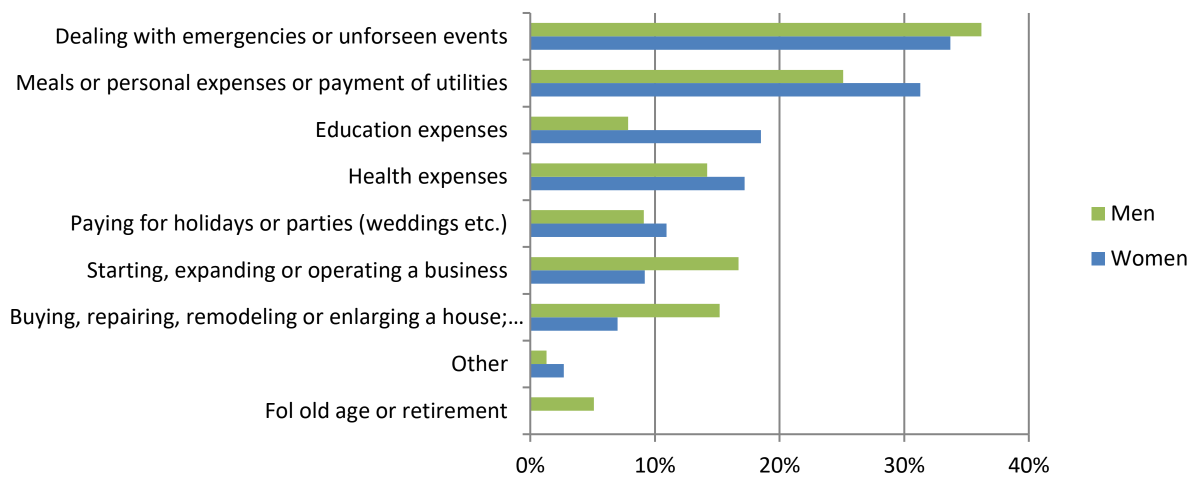
Appendix 2.5: Reasons behind using credit and saving facilitates by entrepreneurs

A. Purpose behind using credit



Source: ENIF, 2015

B. Purpose behind savings



Source: ENIF, 2015

Appendix 2.6: Results for sub-samples using different indicators for financial inclusion

	Branches/ 10,000 adults	**POS terminals/ 10,000 adults**	**Banking Agents/ 10,000 adults**
Overall Sample (Men and Women)	0.005***	0.0001***	0.004***
Women	0.0025***	0.00004***	0.0015***
Women in the Formal Sector	0.003***	0.00003***	0.001**
Women in the Informal Sector	0.002	0.0001***	0.005***
Women in Urban Areas	0.003***	0.00004***	0.002***
Women in Rural Areas	0.001	0.000002	0.0001
Men	0.0053***	0.0001***	0.004***
Men in Formal Sector	0.0045***	0.0001***	0.003***
Men in Informal Sector	0.003***	0.0001***	0.004***
Men in Urban Areas	0.008***	0.0001***	0.007***
Men in Rural Areas	0.001	0.0001	0.0002

*Note: Results of the logit regressions, for different sub-samples, have been reported with average marginal effects. Standard errors are clustered at the municipality level. The dependent variable is entrepreneurship status dummy. Control variables have been included but not reported due to ease of reading. ***, ** and * represent significance at 1%, 5% and 10%, respectively.*

Appendix 2.7: Robustness Check: Akaike information criterion (AIC)

Model	**Observations**	**ll(null)**	**ll(model)**	**df**	**AIC**
With the FII	554.659	-305977.9	-108933.1	52	217970.1
Without FII	554.659	-305977.9	-108984.8	51	218071.7

Note: The model with the financial inclusion indicators fits the data better because it has a lower AIC value.

Appendix 2.8: Robustness check with additional controls

Variables	**Women Entrepreneur (1)**
Financial Inclusion Index (FII)	0.031***
	(0.15)
Education Level (Base: Primary or No Education)	
Secondary Education	0.022***
	(0.02)
Upper Secondary or Undergraduate	0.064***
	(0.03)
Masters or PhD	0.164***
	(0.08)
Age	0.003***
	(0.0009)
Urban Dummy	0.002
	(0.04)
Marital Status (Base: In a Relationship)	
Separated	-0.01***
	(0.04)
Divorced	0.018***
	(0.05)
Widowed	-0.006**
	(0.05)
Married	0.016***
	(0.03)
Single	-0.023***
	(0.03)
Others	0.07
	(0.89)
Informal Dummy	0.135***
	(0.05)
Monthly Income (log)	0.009***
	(0.01)
Business Ease Score	0.004
	(0.05)
Employment Rate in Municipality	0.005
	(0.36)
Income level in Municipality (log)	0.009***
	(0.04)
Sector Fixed Effects	**YES**
State Fixed Effects	**YES**
Time Fixed Effects	**YES**
Observations	554,659
Pseudo R2	0.6830

****,** and * represent significance at 1%, 5% and 10%, respectively.*

Chapter 3: Pushed into the Corner: An Empirical Study on the Drivers of Financial Exclusion

Evidence from Pakistan using Demand Side Data

Abstract

The role and impact of financial exclusion in terms of credit has been an important area of research for policymaking in the recent past. It is important to note that financial exclusion can be due to voluntary as well as involuntary reasons, and the existing literature does not generally differentiate between the two and renders them under the same umbrella which can lead to biased results. This chapter differentiates between these two types and focuses on studying the main drivers of involuntary financial exclusion. This chapter relies on the use of a nationally representative survey of adults for the year 2015 and 2016 from Pakistan to conduct this study. Our findings highlight that the demand for formal credit is considerably less than what is conventionally believed as 38.1% of the adult population show no need for credit and about 24.5% opt for voluntary financial exclusion. The results of the econometric estimates show that financial illiteracy and poverty are strongly associated with involuntary financial exclusion. Results also point towards the existence of a gender disparity. According to the findings, women are more likely to be involuntarily financially excluded, as compared to men, after controlling for other individual and regional characteristics. Regional characteristics also seem to play a significant role in driving involuntary financial exclusion.

3.1. Introduction

The concept of financial inclusion (or, alternatively, financial exclusion) covers a wide range of financial services such as credit, saving accounts, insurance, and payment facilities (Demirguc-Kunt et al., 2017). An increased amount of the literature has discussed the adverse effects of financial exclusion on different dimensions of development[37], however, little is known about the underpinnings of financial exclusion from the demand side and the policies that can be used to alleviate it. This is especially true for credit. Credit does not have a universal demand, unlike bank accounts which are more likely to be universally demanded (Karlan et al., 2017). Moreover, due to concerns pertaining to financial instability, universal use of credit services is also not a policy goal. Furthermore, the co-existence of formal and informal credit markets and their interaction makes it an interesting yet complex phenomenon to understand. Therefore, keeping these points in mind, it is important to critically analyze the factors that drive financial exclusion in a developing economy with regards to credit.

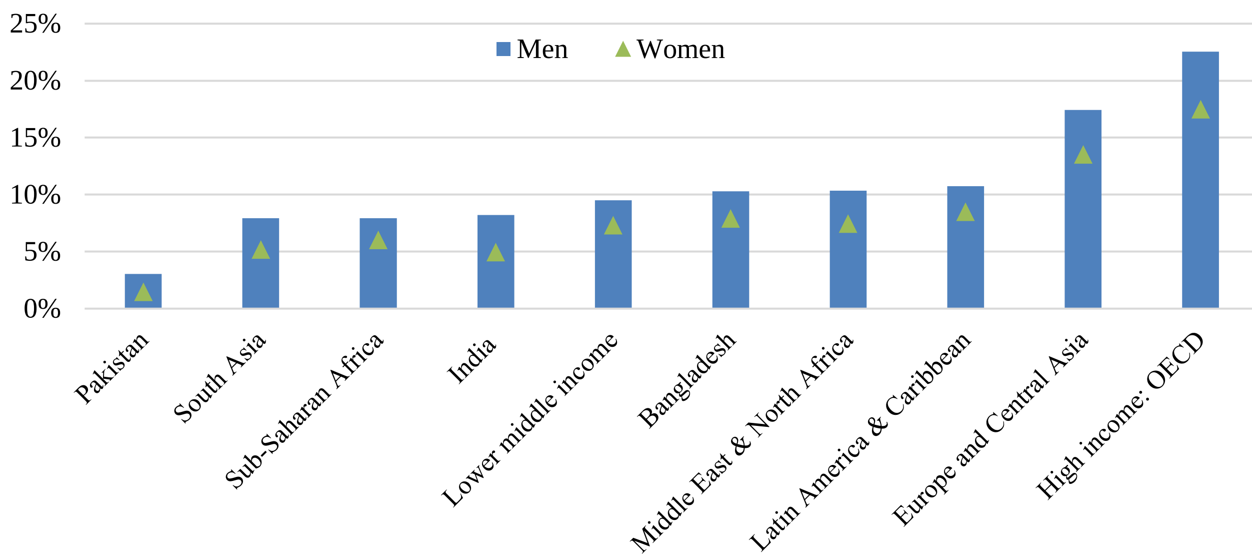
This chapter looks at financial exclusion from the perspective of credit[38]. Dev (2006) argues that credit is the most vital component of financial inclusion (or exclusion) due to a number of reasons. It provides individuals, especially poor people working as entrepreneurs and farmers, with opportunities to make investments in fixed capital and to finance their working capital. It also helps these households smooth their consumption, particularly in times of low economic activity and seasonal fluctuations in the business cycle. Moreover, it can also help with unforeseen emergencies such as illnesses and loss of employment, or in the case of events like weddings which at times can create a pressing need to borrow money. The existing literature has also shown that availability of credit is also essential for business growth (Bruhn and Love, 2014), adoption of new technologies (Beck et al., 2008), reducing income inequality and poverty (Demirguc-Kunt and Levine, 2009), and women empowerment (Pitt, Khandker and Cartwright, 2006).

[37] See Section 0.3 for a detailed review of the literature.

[38] Throughout this chapter, the term financial exclusion or inclusion refers to credit. Other financial services such as bank accounts, insurance and payment services are not analysed in this chapter.

It is generally assumed that the demand for formal credit is very high especially amongst the poor households. However, in the case of Pakistan, only about 1.2% of the adults report using formal credit, despite three decades of microfinance in the country and in spite of the fact that more than 85% of the adult population has geographical access to a formal financial institution within 10 kilometers of their household (Financial Inclusion Insight, 2016). This assumption about excessive demand is also at odds with the actual experiences of credit officers and bankers who frequently report issues pertaining to finding microcredit clients in developing countries (Aslam and Azmat, 2012; Kochar, 1997). Crépon et al. (2015) conducted a study in Morocco and showed that the take up of microfinance was quite low (around 13%). This goes on to show that there is a vast gap between the availability of formal financial services and their actual usage, highlighting the importance of having a comprehensive understanding of the main drivers of financial exclusion in terms of credit. Figure 3.1 summarizes how the usage of formal credit in Pakistan lags behind neighbors and other regions across the world. These numbers also highlight that only a limited number of people borrow from formal financial institutions even in the case of highly developed countries which have a close to perfect supply of formal financial services. A world map depicting how formal borrowing varies across the world is provided in Appendix 3.1.

Figure 3.1: Borrowing from formal financial institutions



Source: World Bank's Findex Database, 2017

The term financial exclusion refers to individuals who have a need for credit and do not get access to formal credit (Dev, 2006). However, it is important to understand that financial exclusion is a multi-dimensional concept and the reasons behind exclusion can be voluntary or involuntary (Claessens, 2006). Voluntary reasons are usually linked with individual preferences, culture or religious practices when individuals choose to voluntarily rely on sources other than formal financial institutions. On the other hand, involuntary reasons can stem from the lack of geographical access to those services or it can be conditional exclusion where a certain segment of population is excluded from using formal financial products because they are either not designed according to their needs or they are expensive and unaffordable (Mylonidis et al., 2017; World Bank, 2014). In case of involuntary financial exclusion, individuals do not have access to formal credit despite their need for it, and hence these individuals are the ones who are credit constrained unlike the ones who opt for voluntary financial exclusion. The existing empirical literature that has looked into financial exclusion from the perspective of credit has generally classified both of these types under the same umbrella, mainly due to data constraints, which can lead to biased results. From a policy perspective, focusing on involuntary financial exclusion is of significant importance since the individuals concerned lack access to formal credit in spite of their need for it.

The main objectives of this chapter are twofold. The first objective is to provide a detailed understanding of financial exclusion, at the microeconomic level, after incorporating the need for credit and after differentiating between two different types of financial exclusion: voluntary and involuntary exclusion. We recognize that the demand for credit is not universal and not everyone needs credit. Therefore, in reality, the path to financial exclusion is likely to follow a sequential process. As a first step, the individual decides whether she or he needs credit or not. As a second step, based upon the fact that the individual needs credit, there is the possibility of three different choices. First, the individual can get the credit from a formal financial institution[39.] In this case, the individual is classified as financially included. The second possibility is that the individual voluntarily self excludes herself or himself from the usage of formal source of credit due to reasons pertaining to religion, cultural preferences, or because of individual preferences to rely on friends and family or other informal sources of credit. These individuals are classified as exhibiting

[39] Formal sources of financial services here refer to commercial banks, microfinance institutions, and the government.

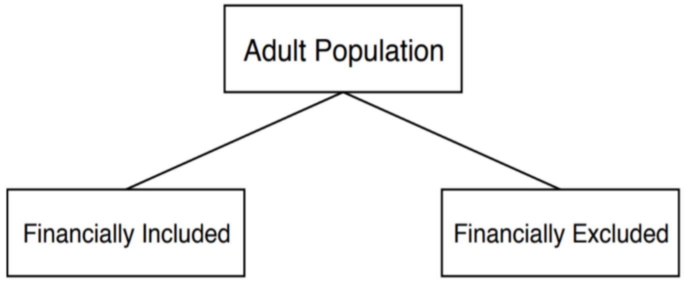
voluntary financial exclusion and are not credit constrained. The third possibility is that the individual is unable to obtain credit from a formal source due to market failures and asymmetry of information and suffers from involuntary financial exclusion. In this case the individual either remains credit constrained and borrows nothing or ends up borrowing from informal source despite its preference for formal credit. Therefore, the individuals in the third category are the ones which are facing financial exclusion as they are the ones who are credit constrained. Hence, in the second step, the three possible outcomes are: (1) financial inclusion, (2) voluntary financial exclusion, and (3) involuntary financial exclusion. The existing empirical literature has generally taken a simpler approach while analyzing financial exclusion by treating it as a one-step binary decision process i.e. if an individual or household has credit from a formal financial institution then that individual or household is treated as being financially included, otherwise it is financially excluded. However, we believe that the process consists of two transitional steps and it is imperative to differentiate between voluntary and involuntary financial exclusion from a policy perspective. Figure 3.2 provides a visual representation of the sequential decision process as compared to the conventional way of looking at financial inclusion.

The second objective is to conduct an empirical investigation to better understand the main covariates of involuntary financial exclusion by taking the selection bias issue into account. This chapter also focuses on highlighting the several methodological challenges which arise when conducting an empirical study to analyze financial exclusion.

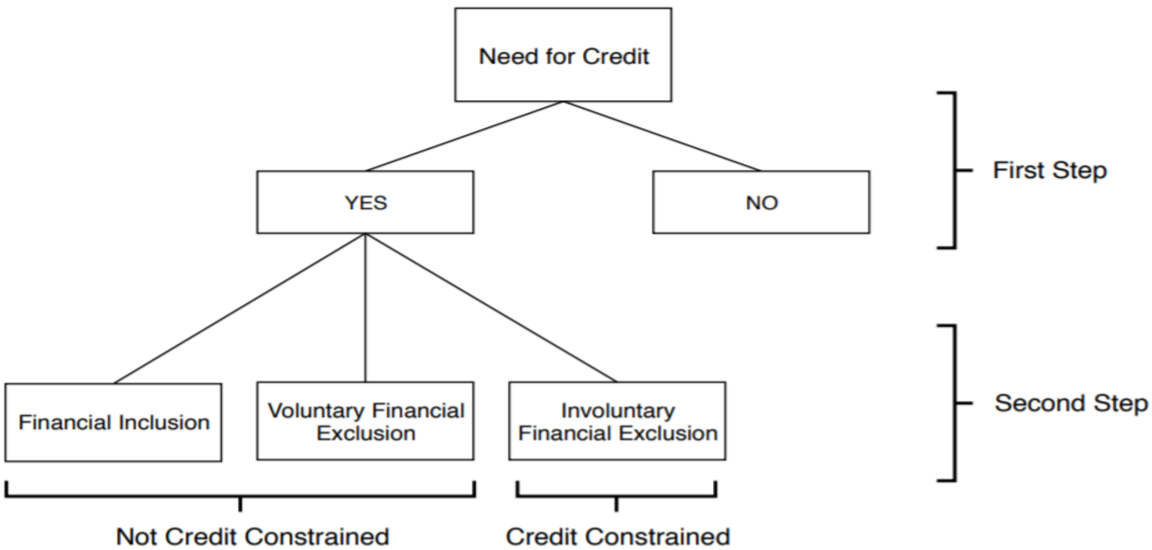
This chapter adds to the literature on financial exclusion in many ways. This is one of the few studies which systematically investigates the subject of financial exclusion by looking at voluntary and involuntary exclusion separately, at the microeconomic level, using nationally representative data. This chapter also tries to address the issue of selection bias in its empirical strategy in order to provide a better understanding of the determinants of involuntary financial exclusion. Moreover, this analysis contributes to the existing literature on the gender disparity regarding the status of financial exclusion, and the role that financial literacy cab play in alleviating financial exclusion. To the best of our knowledge, this is the first study from Pakistan which systematically analyses the determinants of credit need and involuntary financial exclusion. The findings give rise to policy recommendations that can help alleviate involuntary financial exclusion.

Figure 3.2: Two step approach towards financial exclusion

A: Conventional approach towards financial exclusion



B: Two step approach towards financial exclusion



Source: Author's conceptualizations

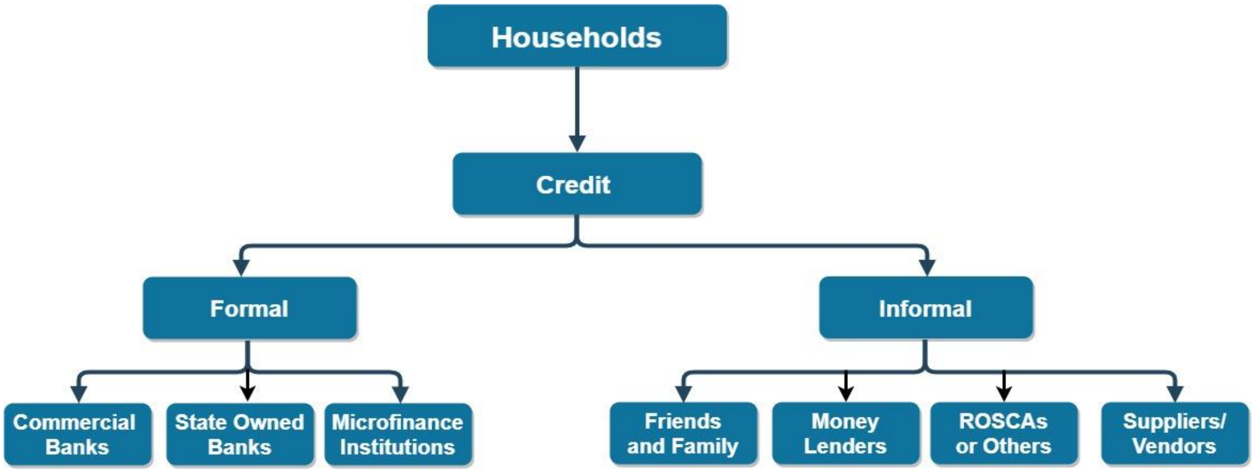
Our results show that the demand for credit is perhaps overrated as 38.1% of the total adult population expressed no need for credit and 24.5% of the adult population chose voluntary financial exclusion. Only 1.2% of the adult population reports borrowing from a formal financial institution and 36.2% turn out to be financially excluded involuntarily. We carry out an empirical investigation to gauge what are the main determinants of need for credit and involuntary financial exclusion. With regard to the need for credit, not surprisingly the results show that the need for credit is higher for marginalized people i.e. people who are poor, less educated, unemployed and looking for jobs. Concerning involuntary financial exclusion, econometric estimates indicate that financial illiteracy and poverty levels are amongst the main determinants of this exclusion. Financially literate people come out to be about 5.8 percentage points less likely to be financially excluded involuntarily as compared to people who are financially illiterate. This shows that the degree of financial literacy makes a big difference in reducing involuntary exclusion and this finding seems to have a clear policy message. Furthermore, the results also point towards the existence of a significant gender disparity. Women are about 19.7 percentage points more likely to be financially excluded involuntarily, as compared to men, after controlling for other individual and regional characteristics. This finding is in line with similar empirical evidence from other south Asian countries which also highlights that women are more credit constrained than men. For example, Ghosh and Vinod (2017) find that women are 8% less likely to be financially included in terms of credit in India as compared to men. Similarly, Malapit (2012) find that women in Philippines are 11% more credit constrained as compared to men. Lastly, our findings also highlight that regional characteristics seem to play a significant role in driving the level of involuntary financial exclusion in Pakistan.

The rest of the chapter is structured as follows. Section 3.2 provides a brief theoretical backdrop regarding the credit rationing theory. Section 3.3 then provides a detailed discussion on the definitions of our main variables of financial exclusion and how they have been measured. Section 3.4 provides a theoretical background of three main types of factors that influence financial exclusion in an economy. Section 3.5 provides details about our data and discusses our empirical methodology. Section 3.6 provides the results of the econometric analysis aiming at examining the key determinants of need for credit and involuntary financial exclusion. Section 3.7 discusses possible limitations of this work and highlights some of the main methodological challenges. Finally, Section 3.8 concludes and discusses some policy responses that can help reduce the level of involuntary financial exclusion.

3.2. Theoretical Backdrop

There are a number of formal and informal sources of credit available to individuals living in developing countries. On one hand, informal sources of credit refer to friends and family, moneylenders, suppliers, harvest borrowers and informal savings or borrowing groups such as ROSCAs. On the other hand, formal sources of credit refer to commercial banks, microfinance institutions, and state-run financial institutions that provide credit facilities. Figure 3.3 depicts the potential sources of credit that are available to individuals living in a developing country such as Pakistan.

Figure 3.3: Potential sources of credit



Source: Author's conceptualization

Stiglitz and Weiss (1981) came up with one of the first theoretical models to investigate the reasons behind credit rationing in developing economies. They argued that credit rationing happens due to two main sources of informational asymmetries: adverse selection (i.e. hidden information) and moral hazard (i.e. hidden actions). Since then, further developments have been made in the credit rationing theory by others. Ghosh et al. (2000) argued that adverse selection as the main explanation behind credit rationing is not always applicable in the context of developing countries because of the fact that informal lenders do have information on potential borrowers, especially in rural areas. Similarly, Koschar (1997) incorporated the demand of credit in the theory of credit rationing and also took into account the role of the informal sector. He argued that if a household does not borrow from a formal channel that does not always mean credit rationing. It might be because of the fact that borrowing from informal sources is actually cheaper as compared to borrowing from formal sources. This argument is also supported by the recent data and the existing literature. Despite a significant increase in the availability of formal financial institutions such as microfinance providers and state-owned banks that do not generally require physical collateral, informal sector borrowing still remains quite large. For example, in the case of Pakistan, more than one third of the adults who needed credit and who were not borrowing from a formal institution mentioned that they actually prefer borrowing from informal sources.

The existing empirical literature that has looked into what drives financial exclusion generally has not stressed enough upon two things. First, the demand for credit is not universal and not everyone needs credit. Secondly, households might actually prefer to borrow from the informal sector. The existing literature considers households who have not received any credit from formal financial institutions as credit rationed and thus financially excluded. Consequently, these papers sometimes end up overestimating the extent of financial exclusion[40].

3.3. Measuring Financial Exclusion Indicators

This section provides precise definitions of the indicators pertaining to credit need, financial inclusion, involuntary financial exclusion, and voluntary financial exclusion. The data that we use to conduct this empirical research is the same as the first chapter and comes from the Financial Inclusion Insight (FII) Program. The survey is conducted by Gallup International Association under the supervision of Bill and Melinda Gates Foundation and InterMedia USA. It is a nationally representative survey of adults (aged more than fifteen years) and we rely on two waves of the survey which were conducted in 2015 and 2016. It is a pseudo panel and individuals surveyed in the two waves are not the same. The overall sample is distributed proportionally to the size of population with rural and urban strata.

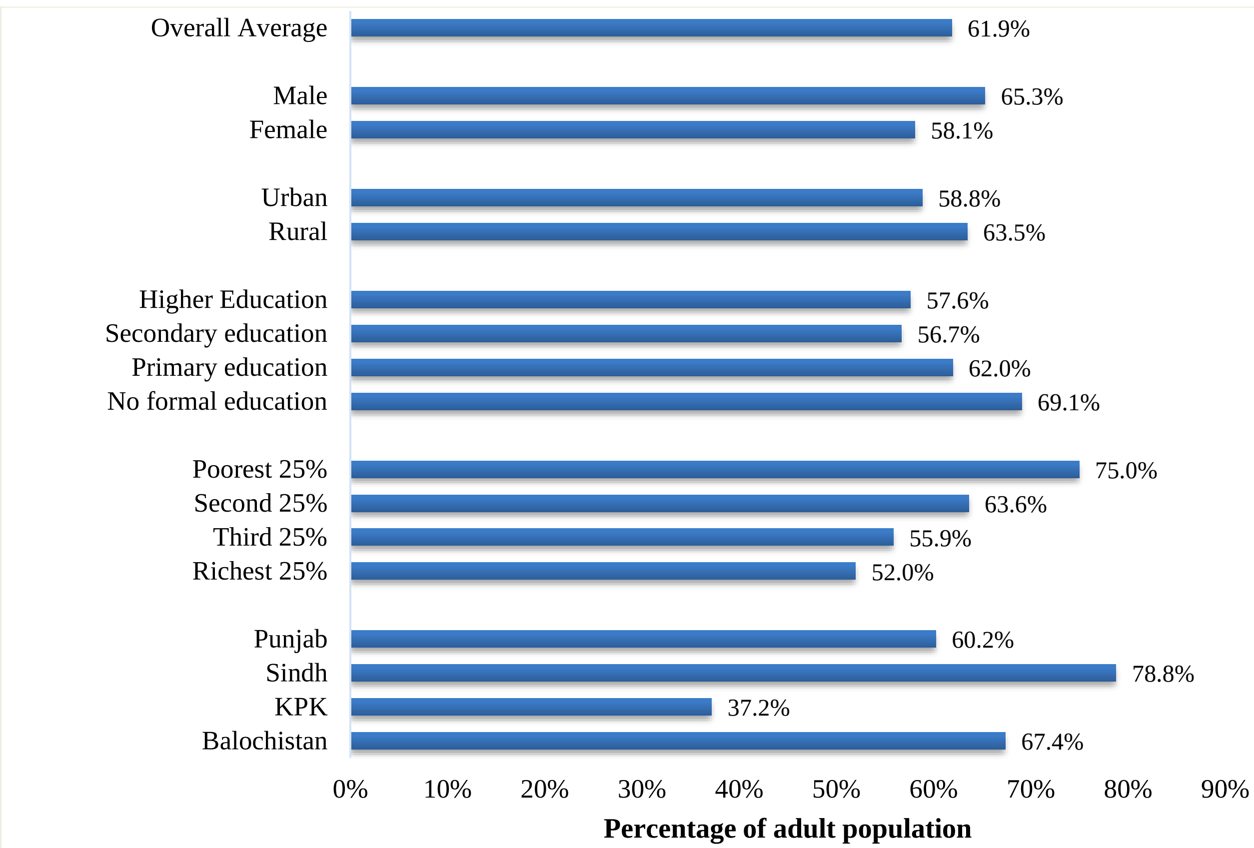
3.1. Need for Credit

The approach to analyze credit demand and usage is complex. If an individual has credit, this can be translated to the fact that the individual has needed credit, demanded credit, and uses it. If an individual does not have credit, it cannot be directly inferred that there is an issue of exclusion (Parienté, 2005). Perhaps the individual has no need for credit and therefore does not demand it. Also, a lot of individuals who are in need of credit might not apply for it maybe because they do not know who to go to or because of issues pertaining to financial illiteracy. This shows that need for credit and credit demand are two distinct concepts and need for credit comes before credit demand. Therefore, as a first step, it is important to identify if there is an effective need for credit or not. We use the question “What is the main reason why you do not borrow from a formal financial institution?” for that purpose. From the overall sample, 38.1% individuals reported *“I do not need to borrow”*, whereas the rest of the sample (61.9%) ended up exhibiting a need for credit.

[40] For a discussion on empirically investigating financial exclusion in terms of credit, see for instance Becchetti et al. (2011) and Boucher et al. (2009).

In terms of differences across individuals who exhibit a need for credit, poverty level seems to be a major one. About 75% of the individuals in poorest households (quartile 1) report a need for credit, whereas this number is 52% for individuals in the richest households (quartile 4). There is also some disparity along the gender lines. 65% of men report having a need for credit, as compared to 58% of women. This difference in the need for credit is much bigger across different regions in Pakistan. Adults in Sindh have the highest need for credit (79%) which is almost double as compared to adults in KPK who exhibit the lowest need for credit (37%). Figure 3.4 summarizes these differences in need for credit across different individual and regional characteristics. The statistical difference in individual and regional characteristics between those who need credit and those who do not is also provided in Appendix 3.2.

Figure 3.4: Need for credit across regional and individual characteristics



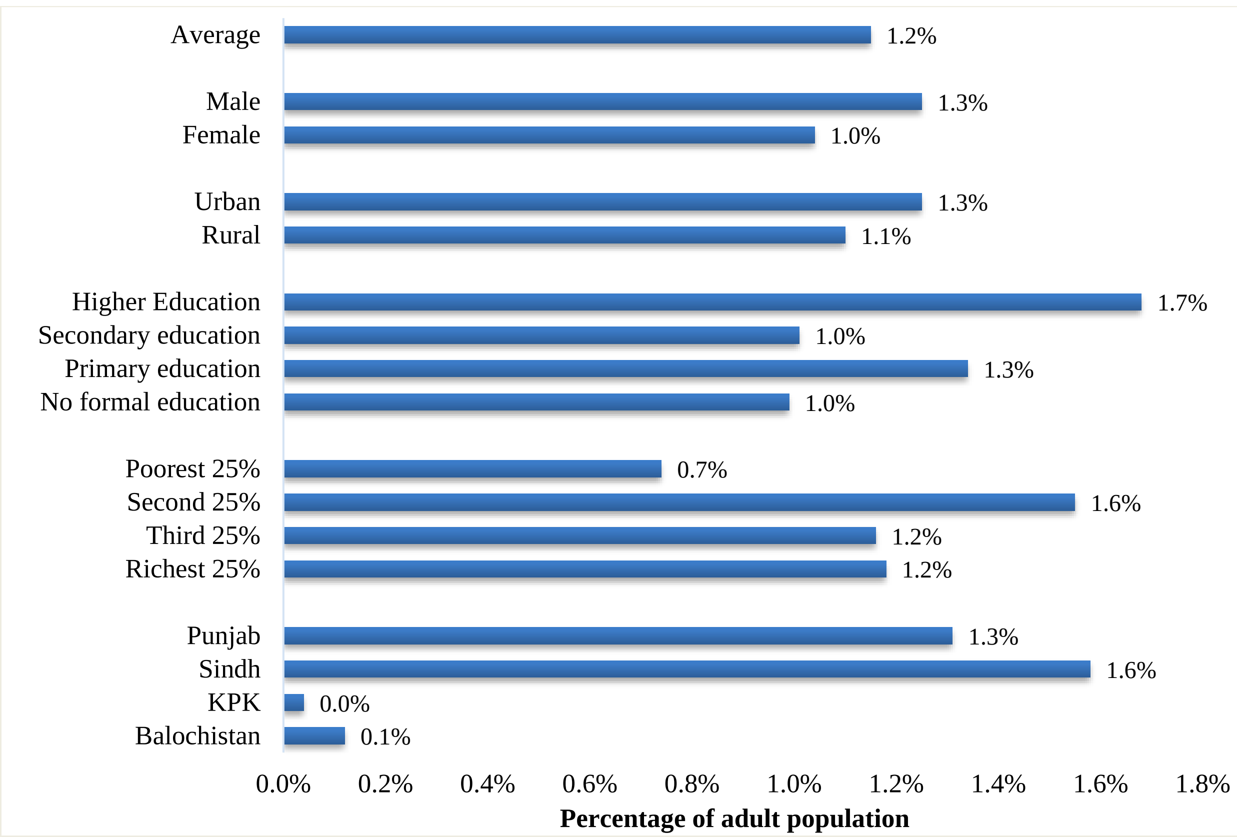
Source: Adopted from Financial Inclusion Insight (FII) Survey (2015 and 2016)

3.2. Financial Inclusion

The concept of financial inclusion refers to the use of formal financial services i.e. credit in this case. Formal financial institutions are the ones that are registered with a legal authority and in

Pakistan's case they include commercial banks, microfinance institutions and state-run banks[41]. We use the question *"Have you borrowed money/ have loans from a bank/ microfinance institution?"* to gauge financial inclusion for our analysis. If an individual has borrowed money from either of these financial institutions, that individual is considered to be financially included. On average, 1.2% of the total adult population in Pakistan reports being financially included i.e. using credit from a formal financial institution. There are some clear differences in the financial inclusion rates across regions; Provinces of Balochistan and KPK have a very low financial inclusion rate of close to zero[42]. With regard to gender and urban/ rural location, there does not seem to be much of a difference. Figure 3.5 summarizes these differences in financial inclusion across different individual and regional characteristics.

Figure 3.5: Financial inclusion across regional and individual characteristics



Source: Adopted from Financial Inclusion Insight (FII) Survey (2015 and 2016)

---

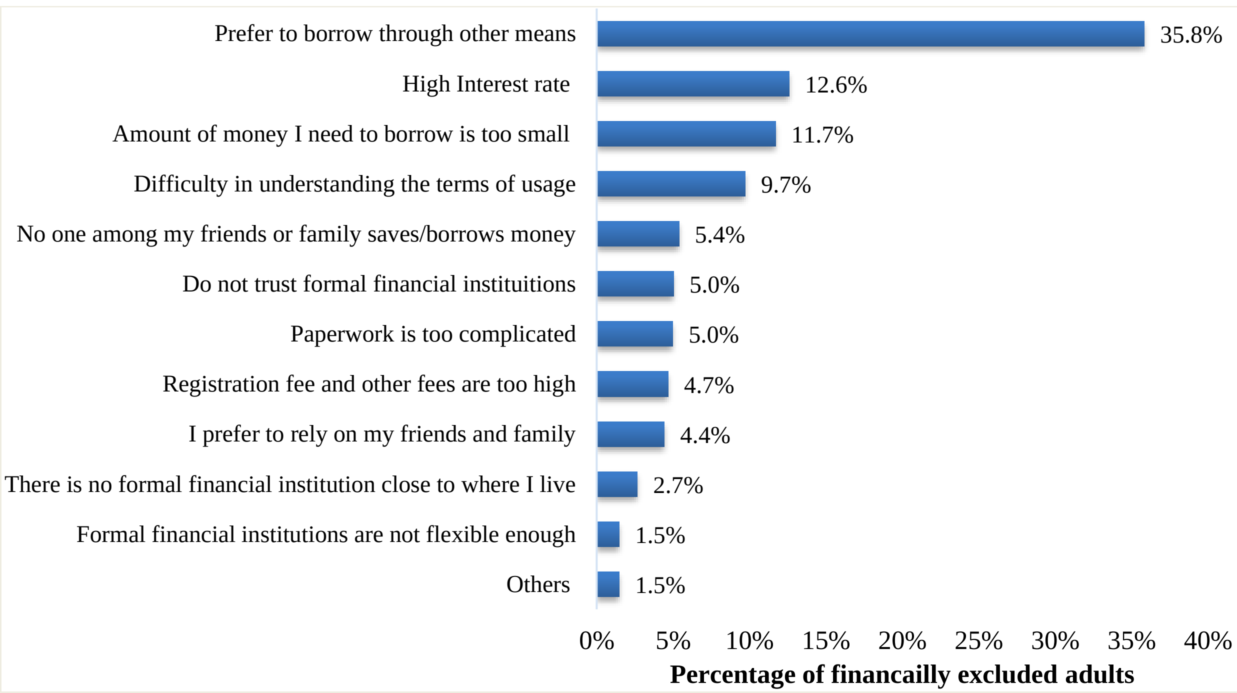
[41] Mobile banking providers are not allowed to offer credit services in Pakistan yet.

[42] It is important to note that this survey is representative at a national level but not at the provincial level. Total observations for Balochistan are quite low because Balochistan represents only 6% of the total population in Pakistan.

3.3. Financial Exclusion and Types

Financial exclusion here refers to individuals who needed credit and did not obtain it from a formal financial institution. As highlighted before, financial exclusion is a multi-dimensional concept and individuals can be financially excluded due to voluntary or involuntary reasons. However, the identification of the financial exclusion status of people who do not participate in the formal credit market is quite challenging because of the differences in the characteristic of contracts offered by different types of formal and informal lenders. In order to differentiate between the two types of financial exclusion in this case, we use the question "*What is the main reason you do not borrow from a formal financial institution*?"[43]. The results of the reported reasons behind financial exclusion are summarized in Figure 3.6.

Figure 3.6: Reasons behind financial exclusion



Source: Adopted from Financial Inclusion Insight (FII) Survey (2015 and 2016)

Involuntary Financial Exclusion

Involuntary financial exclusion refers to individuals who do not use formal credit because of different constraints that arise due to asymmetric information and other market failures (Claessens,

[43] The survey allows single response per individual i.e. one main reason behind financial exclusion.

2006). These involuntary reasons could stem from the lack of geographical access or unaffordability of financial products or issues pertaining to unsuitable product design.

The data shows that, indeed, one of the most important reasons behind financial exclusion is the high cost of borrowing. About 12.6% of the financially excluded adults reported that they do not borrow from financial institutions because *the interest rate is too high* and 4.7% of the adults reported *registration and other fees are too high* as the main reason behind not borrowing from financial institutions. Another important reason for involuntary financial exclusion appears to be the size of credit as 11.7% of financially excluded individuals reported that *the amount of money they need to borrow is too small to use such a service*. Moreover, 9.7% of individuals reported that they have *difficulties in understanding the terms and conditions of usage of financial services*. Other reported reasons behind financial exclusion, which fall under the involuntary category, include: *I do not trust formal financial institutions*, *paperwork is too complicated*, *no one among my friends or family borrows/ saves*, *there is no formal financial institution close to where I live*, and *formal financial institutions are not flexible enough*.

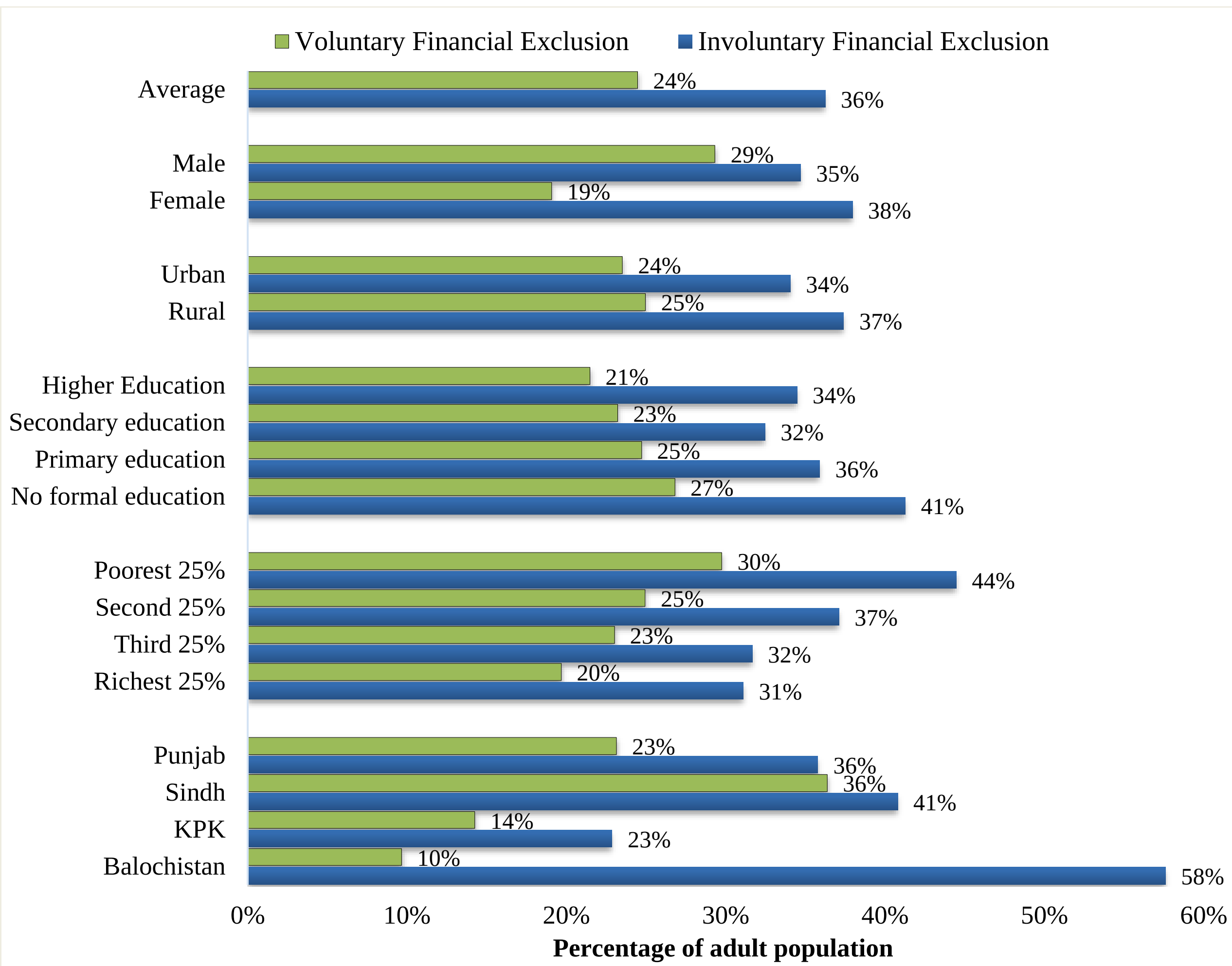
On average, 36.2% of adults in Pakistan are financially excluded involuntarily. There are some piercing differences in involuntary financial exclusion across the four regions in Pakistan. Balochistan has the highest rate of involuntary financial exclusion (58%), whereas KPK has the lowest (23%). Moreover, this difference in involuntary financial exclusion is about 3 percentage points between men and women and between urban and rural location. Furthermore, the involuntary financial exclusion rate is the highest amongst the poorest households (44%) and it is the lowest amongst the richest (31%) as shown in Figure 3.7.

Voluntary Financial Exclusion

An individual is considered to exhibit “voluntary financial exclusion” if he or she voluntarily chooses not to use formal financial services due to reasons linked with culture, religious practices, or individual preferences (Claessens, 2006). Indeed, the most cited reason behind financial exclusion in the data is the *preference to borrow money through other means* (35.8%). Moreover, 4.4% of the financially excluded people report that *I prefer to rely on my friends and family* and 1.5% choose the option “other” and mainly cite religious reasons for not borrowing from formal financial institutions. These individuals have chosen to voluntarily self-exclude themselves from the use of formal financial institutions because of their personal preferences, and therefore they fall under the category of "voluntary financial exclusion".

There are also some clear differences in voluntary financial exclusion across different individual characteristics. These differences have been summarized in Figure 3.7. For example, there is a 10 percentage points difference in voluntary financial exclusion between men and women. On average, 29% men are financially excluded voluntarily as compared to 19% women. There are some visible differences across regions as well. Adults in Sindh have the highest percentage of voluntary financial exclusion (36%) as compared to adults in Balochistan who have the lowest one (10%). Moreover, there are practically no differences in voluntary financial exclusion across urban and rural locations. These differences in terms of involuntary financial exclusion across different individual characteristics have been summarized in Figure 3.7.

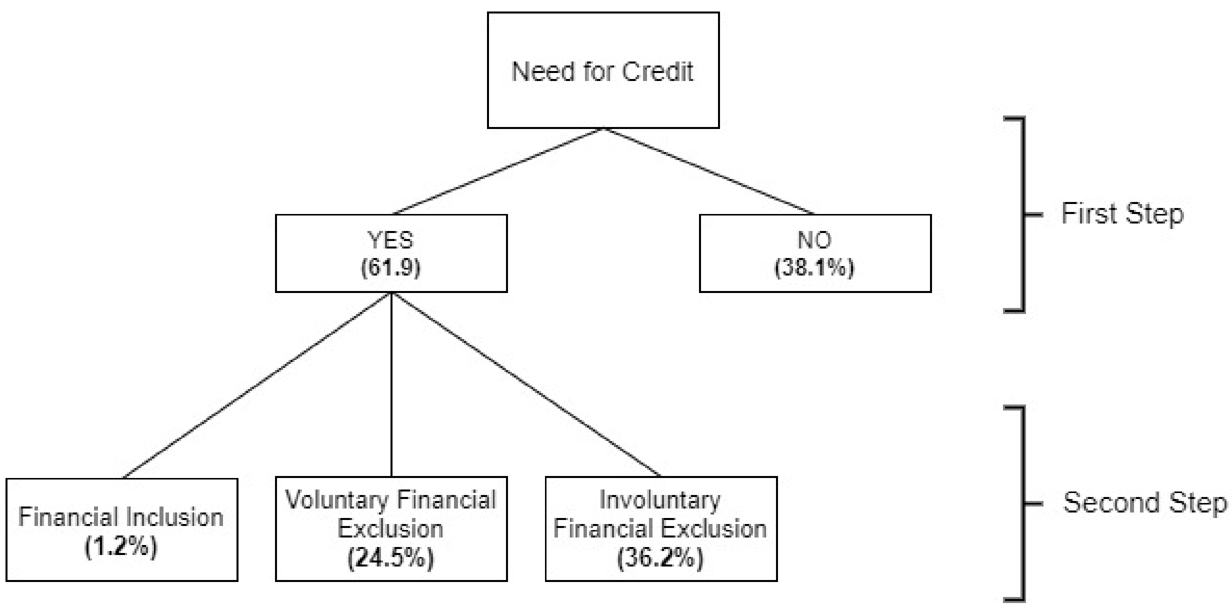
Figure 3.7: Voluntary and Involuntary Financial Exclusion



Source: Adopted from Financial Inclusion Insight (FII) Survey (2015 and 2016)

In summary, according to our classification, 61.9% of the adult population in Pakistan exhibited a need for credit. People who exhibited a need were then categorized as either financially included or financially excluded. Since financial exclusion can be due to voluntary or involuntary reasons, we differentiated between these two types to investigate in detail the determinants of both types of exclusion. The distribution of the overall population with respect to their need for credit, financial inclusion, voluntary and involuntary financial exclusion is provided in Figure 3.8. Moreover, this classification is not without its limitations and we discuss this in Section 3.7.

Figure 3.8: Need for Credit and Financial Exclusion in Pakistan



Note: The number in parenthesis represents the percentage of adult population.
Source: Authors' calculations based on Financial Inclusion Insight (FII) Survey

3.4. Drivers of Financial Exclusion

Achieving a situation in which the population has access to affordable formal financial services and be able to use them is a challenging task to achieve for developing economies. It is a complex problem to tackle as it can be a result of the interaction of several factors. These factors can be broadly organized into three categories. The first category includes different factors that define the overall environment in which the whole system of financial services operates. The second category includes the supply side factors driven by financial service providers that influence financial exclusion, whereas the third category includes the demand side characteristics which can drive financial exclusion. All these factors combine and interact with each other, to different extents, and can influence the status of financial exclusion in an economy. These factors and their possible influence have been discussed in this section.

3.1. Environment-related factors

Factors that are part of the environment include the macroeconomic context of the country, geographical characteristics, infrastructure, and the quality of institutions. All these different aspects of the environment act as exogenous conditions under which the financial system of the economy operates. The importance of these factors in driving the performance of the financial system has been widely documented in the literature (Beck and Levine, 2018).

Macroeconomic factors such as economic growth, inflation, and macroeconomic stability encourage investments and funding opportunities in new ventures. Factors related to geography and infrastructure affect the transaction costs. Having an enabling telecommunications infrastructure can help reduce these transaction costs which can have a soothing effect on financial exclusion. Quality of institutions such as an empowering regulatory framework for financial institutions can also help reduce information asymmetries and improve provision of financial services to the excluded population. Rojas-Suarez (2010) highlights that the presence of effective credit risk bureau plays an important role in reducing asymmetries in information which has a positive impact on decreasing financial exclusion. A combination of prudential and non-prudential regulations to ensure stability and solvency of financial system as well as safeguarding consumer protection and transparency of information is also considered to play a significant role in shaping the status of financial exclusion in an economy (Cull et al., 2014).

Most of these factors related to the environment are quite homogenous at a country level, at least to an extent. However, due to the stark differences amongst provinces in Pakistan, we focus on this regional heterogeneity because it can also drive financial exclusion.

3.2. Supply-side factors

The supply side factors such as proximity and affordability of financial services play a crucial role in defining the participation of people in the formal financial system. There are different types of financial institutions which operate within an economy which differ from each other in terms of their mission, target customers, business model, and the range of products and services offered. If the products are tailored to the needs of the target population and the financial institutions are willing to innovate to reach out to the unbanked customers, this can alleviate financial exclusion. On the opposite, if the products are not affordable due to their high costs, individuals are more prone to look for possible solutions outside the formal financial system.

Several studies have highlighted the importance of supply side factors in influencing financial exclusion. Allen et al. (2016) show that low fees, less documentation requirements, and closer proximity to financial institutions have a desirable impact on decreasing financial exclusion rates. Similarly, a research by Talledo (2015) highlighted that the availability and the physical distance to bank branches and service centers is directly associated with the chances of asking for credit or a savings account. Hence, the supply side factors including the characteristics of financial products and services and the convenience of reaching out to financial channels play a vital role in enhancing the use of formal financial services.

3.3. Demand-side factors

Studies which have tried to study financial exclusion (or, financial inclusion) from the supply side are more abundant as compared to the ones on demand side, mainly due to higher availability of data from the supply-side which is collected on regular basis by regulatory authorities. This supply side information is provided by the financial institutions themselves and it has some shortcomings. First, this information generally provides statistics on progress rather than needs and perceptions of the population. Secondly, commonly used proxies such as borrowers per financial institution or number of bank accounts can overestimate the concentration of financial inclusion due to multiple borrowing or having inactive accounts. Therefore, many researches have shifted their attention to understanding the influence of demand side factors on the usage of financial services. However, significant gaps continue to remain in the underpinnings of a comprehensive assessment of financial exclusion with respect to credit after taking into account need for credit and voluntary financial exclusion.

Demand side factors consist of different demographic and socioeconomic characteristics such as gender, poverty status, financial literacy, work profile, education level, age, religion, and personal preferences. The influence of these demand side factors on the status of financial exclusion is discussed in detail in Section 3.5.2.

3.5. Empirical Strategy and Data

3.5.1. Empirical Specification

In order to gauge the main determinants behind need for credit and involuntary financial exclusion, a two-step sequential process is employed by using a sequential logit model. In the first step, a logit regression is used to estimate the determinants of need for credit. The dependent variable in this case is need for credit ($\boldsymbol{N_{i,d,p,t}}$) and it is dichotomous i.e. if the individual needed credit (1) or not (0). The following regression model is used to investigate the determinants of credit need.

$$\boldsymbol{N^{*}_{i,d,p,t}} = \boldsymbol{\alpha} + \boldsymbol{B_1 X_{i,d,p,t}} + \boldsymbol{B_2 Z_{d,p,t}} + \mu_{\mathbf{p}} + \mu_{\mathbf{t}} + \boldsymbol{\varepsilon_{i,d,p,t}} \qquad \textbf{(3.1)}$$

$$\boldsymbol{N_{i,d,p,t}} = 1 \text{ if } \boldsymbol{N^{*}_{i,d,p,t}} > 0$$
$$\boldsymbol{N_{i,d,p,t}} = 0 \text{ if } \boldsymbol{N^{*}_{i,d,p,t}} \leq 0$$

$\boldsymbol{N^{*}_{i,d,p,t}}$ is a latent variable where *i* and *d* and *p* are indexed for individual, district, and province respectively, and t represents the year. $\boldsymbol{X_{i,d,p,t}}$ refers to the individual and household characteristics and $\boldsymbol{Z_{d,p,t}}$ refers to the vector of regional characteristics which basically include supply-side variables in time period t. The term $\mu_{\mathbf{p}}$ refers to fixed effects at the province level which we include to control for provincial level heterogeneity which might affect chances of credit demand. Moreover, the term $\mu_{\mathbf{t}}$ refers to year fixed effects and $\boldsymbol{\varepsilon_{i,d,p,t}}$ is the error term that is assumed to follow a logistic distribution. In order to control for possible correlation within districts, we cluster the standard errors at the district level for all the regressions. Moreover, the survey does not have an equal probability design. Therefore, we use nonresponse-adjusted sampling weights in our estimations in order to compensate for unequal probabilities of selection. In some specifications, we replace $\mu_{\mathbf{p}}$ with district fixed effects.

As a second step, we investigate the determinants of the different types of financial exclusion. Since we observe this variable of financial exclusion only for individuals who had a need for credit in the first step, we have to rely on a selection model.

$$\boldsymbol{V^{*}_{i,d,p,t}} = \boldsymbol{\alpha} + \boldsymbol{C_1 X_{i,d,p,t}} + \boldsymbol{C_2 Z_{d,p,t}} + \mu_{\mathbf{p}} + \mu_{\mathbf{t}} + \boldsymbol{\varepsilon_{i,d,p,t}} \qquad \textbf{(3.2)}$$

$$\boldsymbol{V_{i,d,p,t}} = 1 \quad \text{if} \quad \boldsymbol{V^{*}_{i,d,p,t}} < \eta 1$$
$$\boldsymbol{V_{2i,d,p,t}} = 2 \quad \text{if} \quad \eta 1 \leq \boldsymbol{V^{*}_{i,d,p,t}} < \eta 2$$
$$\boldsymbol{V_{2i,d,p,t}} = 3 \quad \text{if} \quad \boldsymbol{V^{*}_{i,d,p,t}} \geq \eta 2$$

In equation (3.2), the dependant variable ($\boldsymbol{V_{i,d,p,t}}$) takes the value 1 if the individual is financially included, 2 if the individual exhibits voluntary financial exclusion, and 3 if the individual exhibits involuntary financial exclusion. We estimate the equation (3.2) using a sequential logit regression where the second transition takes the form of a multinomial regression. As before, *i* and *d* are indexed for individuals and districts respectively, and $\boldsymbol{V^{*}_{i,d,p,t}}$ is the latent variable. $\boldsymbol{X_{i,d,p,t}}$ and $\mathbf{Z}_{\boldsymbol{d,p,t}}$ refer to individual and district characteristics respectively at time period t. The terms $\mu_{\mathbf{p}}$ and $\mu_{\mathbf{t}}$ refer to provincial fixed effects and year fixed effects, whereas $\boldsymbol{\varepsilon_{i,d,p,t}}$ is the error term that is assumed to follow a logistic distribution.

The sequential logit model is quite useful in modelling sequential decisions, however, Cameron and Heckman (1998) argue that the presence of unobserved heterogeneity might cause a bias in the sequential logit results. In order to make sure that our results are consistent and reliable, we revert to the use of a Heckman selection model to investigate the determinants of involuntary financial exclusion.

The Heckman model includes a selection function and a response function. In our case, the selection function i.e. need for credit is similar to a regular binary logistic regression as indicated in equation 3.3. Furthermore, the dependent variable in the response function (equation 3.4) is also binary. It takes the value of 1 if the individual exhibits involuntary financial exclusion (i.e. it includes individuals who are credit constrained) and 0 otherwise (i.e. individuals who are not credit constrained- this includes individuals who are financially included and also individuals who choose voluntary financial exclusion). The selection equation needs to contain an exclusion variable V which affects the selection (need for credit) but has no direct effect on the main response variable (involuntary financial exclusion). In this case, this variable V is the number of children in the household. The assumption is that a household with more number of children is likely to have a higher need for credit, whereas, it should not have any direct effect on not being able to obtain financial services from a formal financial institution and being involuntarily financially excluded. Since the dependent variables are dichotomous in both the selection equation as well as the response equation, a Heckman probit model is used.

$$\begin{cases} \boldsymbol{G^{*}_{i,d,p,t}} = \boldsymbol{\alpha}' + \boldsymbol{\theta}\mathbf{V}'_{\boldsymbol{i,d,p,t}} + \mathbf{D}'_{\mathbf{1}}\boldsymbol{X_{i,d,p,t}} + \mathbf{D}'_{\mathbf{2}}\mathbf{Z}_{\boldsymbol{d,p,t}} + \mu'_{\boldsymbol{p}} + \mu'_{\boldsymbol{t}} + \boldsymbol{\varepsilon}'_{\boldsymbol{i,d,p,t}} & (\mathbf{3.3}) \\ \boldsymbol{F^{*}_{i,d,p,t}} = \boldsymbol{\alpha} + \mathbf{D}_{\mathbf{1}}\boldsymbol{X_{i,d,p,t}} + \mathbf{D}_{\mathbf{2}}\mathbf{Z}_{\boldsymbol{d,p,t}} + \mu_{\mathbf{p}} + \mu_{\mathbf{t}} + \boldsymbol{\varepsilon_{i,d,p,t}} & (\mathbf{3.4}) \end{cases}$$

$$G_{i,d,p,t} = \begin{cases} 1 \; if \; \boldsymbol{G}^*_{i,d,p,t} > 0 \\ 0 \; if \; \boldsymbol{G}^*_{i,d,p,t} \leq 0 \end{cases} \qquad F_{i,d,p,t} = \begin{cases} 1 \; if \; \boldsymbol{F}^*_{i,d,p,t} > 0 \\ 0 \; if \; \boldsymbol{F}^*_{i,d,p,t} \leq 0 \end{cases}$$

As before, *i* and *d* and *p* are indexed for individuals, districts and provinces respectively and *t* for year. $\boldsymbol{\theta}$ refers to the coefficient of the exclusion variable, whereas $\boldsymbol{G}^*_{i,d,p,t}$ and $\boldsymbol{F}^*_{i,d,p,t}$ are the latent variables. In all of the regression estimations, standard errors are clustered at the district level and sampling weights are used in order to compensate for unequal probabilities of selection.

3.5.2. Explanatory Variables

In this section we talk about the different individual and regional characteristics that we believe might affect the need for credit as well as the status of financial exclusion. Information on these basic socioeconomic and household characteristics come from the Financial Inclusion Insight (FII) survey.

We use the Poverty Probability Index (PPI) to measure household's *economic profile*. PPI is a country specific poverty measurement tool that uses the national poverty line to rank households from 0 to 100 based on their economic status. We disaggregate the economic profile of the households into quartiles. The first quartile includes the poorest households which are the bottom 25% in terms of their economic status, whereas the fourth quartile includes the richest 25% households with respect to their economic status. We include the first quartile i.e. the poorest 25% households as the reference category in all our estimations. We expect that belonging to a poor household might have a significant effect not only on credit need, but also on the status of financial exclusion.

One of the important concepts which will be discussed in this chapter is that of financial literacy. Financial literacy is typically defined as the ability to use knowledge and skills to handle financial resources effectively (Grohmann et al., 2018). In order to measure the degree of financial literacy, we rely on the five survey items that are proposed by the Global Financial Literacy Centre[44]. These

[44] The World Bank and the Global Financial Literacy Centre conducted a representative survey of more than 1,000 adults per country in 144 countries around the globe to gauge how financial literacy rate varies across different regions and across different individual characteristics.

items ask questions on four different concepts. These concepts include inflation, risk diversification, interest rate and interest rate compounding. We follow the financial literacy indicator proposed by Klapper et al. (2015). This indicator is a dummy variable which takes the value 1 if at least three out of the four financial literacy concepts are answered correctly. These questions which make up this binary indicator have been commonly used in the literature to measure the level of financial literacy (Grohmann et al., 2018; Lusardi and Mitchell, 2014; Xu and Zia, 2012).

The question which is being used to report the understanding of interest rate compounding is "Suppose you had 100 Pakistani Rupees (PKR) in a savings account and the bank adds 10 percent per year to the account. How much money would you have in the account after five years if you did not remove any money from the account?" The possible answers to this question include "1) More than 150 PKR, 2) Exactly 150 PKR, 3) Less than 150 PKR, 4) Don't Know". About 38. 4% of the total population in the survey answered this question correctly. Similarly, for an understanding of the concept of interest, the question that is being used is "Suppose you need to borrow 100 PKR. Which is the lower amount to pay back: 105 PKR or 100 PKR plus 3 percent of 100 PKR?". About 38.6% of the adult population in the survey was able to answer this question correctly[45]. All other questions along with their response categories are provide in Appendix 3.4. The results of the financial literacy indicator show that 23.9% of the adult population in Pakistan is financially literate. About 18.35% of the respondents were not able to correctly answer any of the five questions. The results are summarized in Appendix 3.5.

The variable *female* indicates that the respondent is a woman. The existing literature has linked women with discrimination in the credit market (Ghosh and Vinod, 2017). Allen et al. (2016) found that women are less likely to be financially included mainly because they are less likely to make independent financial decisions and because they are less likely to work in comparison to men in developing economies. We therefore include the gender variable to gauge the effect of being a woman not only on different types of financial exclusion, but also on credit need, after

[45] For comparison purpose, this question was answered correctly by 50% of the adult population from the 143 countries around the world where this survey was conducted (Klapper et al., 2016).

controlling for other individual and regional characteristics. In our overall sample, about half of the respondents are female.

Regarding the employment status, each respondent falls under one of the five categories. The first category refers to individuals who work as *full-time salaried employees*. The second category refers to individuals who work as *part time salaried employees* or *as seasonal workers*. The third category refers to individuals who are *self-employed* and run their own businesses. The fourth category refers to individuals who are *unemployed and looking for a job*. The fifth and the last category included individuals who are *out of the labor force* and not looking for jobs. This category includes students, housewives and retired people. We expect that certain employment statuses might be more likely to have a need for credit and might be more likely to be financially excluded. From our overall sample, about 24% respondents work as employees, where half of them work as full-time employees and the other half works as part-time employees. Moreover, 19% of the respondents are self-employed, 2% are unemployed and looking for a job, and the remaining are out of the labor force.

The variable *urban* takes the value 1 if the individual resides in an urban location and 0 if the individual resides in a rural location. From our overall sample, 66% of the respondents live in rural areas. Generally, formal financial institutions are less prevalent in rural locations as compared to urban locations, thus we expect this variable to have a positive coefficient. Moreover, we include two variables *Age* and *Age Squared* which are measured in number of years. The average age of the respondents comes out to be about 34 years. We expect that age might have a non-linear relationship with need for credit as well as for the financial exclusion variables, and that is why we have included *Age Squared*.

The education variable indicates the highest level of education of the individual. Each individual falls in one of the four categories: (1) if the individual has "no formal education," (2) if “primary education", (3) if "secondary education", and (4) if the individual has a "above secondary education". Baland, Somanathan and Vandewalle (2019) highlight that the level of education plays a crucial role in determining an individual’s access and usage of credit. We expect an increase in educational level to be negatively related to financial exclusion. About 58% of adult population in Pakistan has primary education or less.

The marital status of the respondent is captured by the *Marital Status* variable. Each respondent can fall under one of the four categories. The first category refers to individuals who are *Single*. The second category refers to individuals who are *Married* and the third category refers to individuals who are either *divorced, separated or widowed*. The last category refers to individuals who belong to any *other* category. In the case of Pakistan, live-in relationships are forbidden by the law and that is why there is no separate option in the survey for that. We also include the *household head dummy* variable which takes the value 1 if the respondent is the household head and 0 if the respondent is not the household head. We speculate that being a household head might affect our dependent variables since household heads have more authority over financial decision making. About 31% of the respondents in our sample are household heads[46].

On top of controlling for individual and household characteristics, we also control for a number of regional characteristics which might be associated with credit need and financial exclusion. The existing literature has pointed out that factors such as asymmetric information, high transaction costs, and a lack of geographical access to financial services can hinder the use of formal financial services (Beck et al., 2007; Karlan and Morduch, 2009). We use the Multidimensional Poverty Index (MPI), developed by the United Nations Development Programme (UNDP) using the Alkire- Foster methodology, in order to proxy for the development level of the district. This indicator takes into account three dimensions while measuring the district poverty rate: education, health and living standards. A score close to zero highlights that there are no poor people in that district and a score close to 100 highlights extreme poverty. We also include variables to proxy for the accessibility of financial service providers and the supply of financial services. This includes looking at the *concentration of banks,* which is measured by the number of branches at the district level. This information comes from Pakistan Microfinance Network (PMN). The data on population numbers at the district level come from Pakistan's Census of 2017.

Pakistan is divided into four provinces which are named *Punjab, Sindh, Balochistan and Khyber Pakhtunkhwa (KPK)*. These four provinces are quite different from each other in terms of culture, law and order situation, and public infrastructure. As mentioned before, we include dummy

[46] One household member is surveyed per household.

variables for respondents belonging to one of these provinces in order to control for provincial level heterogeneity that might affect chances of credit demand and financial exclusion. We include Punjab as the reference category for all our estimations. A summary of descriptive statistics is provided in Table 3.1.

Table 3.1: Descriptive statistics

	First Stage		**Second Stage**[47]		**Range**	
	Need for Credit		**Financial Inclusion Status**			
	N=12000		**N=7504**			
Variables	**Mean**	**Std. Dev.**	**Mean**	**Std. Dev.**	**Min**	**Max**
Credit Need_Yes Dummy	0.62	0.49			0	1
Financial Inclusion Dummy			0.02	0.14	0	1
Voluntary Financial Exclusion Dummy			0.40	0.49	0	1
Involuntary Financial Exclusion Dummy			0.58	0.49	0	1
Female Dummy	0.47	0.50	0.45	0.50	0	1
Age	34.03	13.72	34.28	13.60	15	92
Financial Literacy Dummy[48]	0.239	0.43	0.242	0.43	0	1
Employment: Salaried Full Time Dummy	0.12	0.32	0.13	0.33	0	1
Employment: Salaried Part time Dummy	0.12	0.33	0.13	0.34	0	1
Employment: Self Employed Dummy	0.19	0.39	0.20	0.40	0	1
Employment: Unemployed Looking for a job Dummy	0.02	0.13	0.02	0.12	0	1
Employment: Out of Labor Force Dummy	0.56	0.50	0.52	0.50	0	1
Household Head Dummy	0.31	0.46	0.34	0.47	0	1
Rural Dummy	0.66	0.48	0.67	0.47	0	1
Household Size	7.70	3.68	7.79	3.95	1	43
PPI Score	64.26	17.97	61.68	18.46	5	100
Economic Profile: 1st Quartile	0.26	0.44	0.32	0.47	0	1
Economic Profile: 2nd Quartile	0.24	0.43	0.25	0.43	0	1
Economic Profile: 3rd Quartile	0.25	0.43	0.23	0.42	0	1
Economic Profile: 4th Quartile	0.24	0.43	0.20	0.40	0	1
Marital Status: Single Dummy	0.27	0.44	0.25	0.43	0	1
Marital Status: Married Dummy	0.71	0.46	0.72	0.45	0	1
Province: Balochistan Dummy	0.05	0.21	0.05	0.22	0	1

[47] Second stage is reached when the credit demand variable takes the value of 1 in the first step.

[48] This indicator is only available for the year 2016.

Province: KPK Dummy	0.13	0.34	0.08	0.27	0	1
Province: Punjab Dummy	0.58	0.49	0.57	0.50	0	1
Province: Sindh Dummy	0.24	0.43	0.30	0.46	0	1
Poverty Rate District (MPI)	0.34	0.23	0.36	0.24	0.03	0.97
Bank Accounts per adults- District	5.85	18.80	6.63	19.97	0	69

Source: Adopted from Financial Inclusion Insight (FII) Survey (2015 and 2016)

3.6. Results and Discussion

3.6.1. Results for Credit Need

This section presents the results of the econometric analysis which aims to analyze the determinants of credit need in detail. Table 3.2 reports the results of our baseline regression. In case of a logit model, the coefficients cannot be directly interpreted. In order to comment on the magnitude of the coefficients, we have calculated and reported the average marginal effects. In column (1) of Table 3.2, we run the regression with individual controls and provincial fixed effects. We also control for the development level and the supply of formal finance at the district level. However, it can be argued that there are some regional factors other than the development level and the supply of formal finance that may have an effect on the need for credit and financial exclusion e.g. business climate, infrastructure quality, law and order, amongst other factor. Therefore, in column (2), we run the regression with district fixed effects and by adding the financial literacy variable. The variable on financial literacy is available only for one of the two years. Thus, the overall sample gets reduced by about half when we add financial literacy in the model.

The results show that poverty levels and need for credit are intimately related with each other. Need for credit is the highest for poorest households (1st quartile i.e. households who belong to the bottom 25%) as compared to all other categories. People belonging to 3rd and 4th quartile are 7.2 and 10.9 percentage points less likely to need credit respectively as compared to people in the 1st quartile. Moreover, the district variable on multidimensional poverty index also comes out to be significant and positive suggesting that the need for credit is higher in less developed districts with higher poverty levels.

Another important determinant of credit need comes out to be the level of education. Having a higher level of education decreases the chances of having a need for credit. Having secondary education or higher education, as compared to no formal education, decreases the chances of having a need for credit by 3.9 and 6.7 percentage points respectively. Moreover, people who are unemployed and looking for jobs exhibit a higher need for credit as compared to people who are working as full-time employees. Divorced people also seem to have a higher need for credit as compared to single people. The variables on age, gender, and financial literacy come out to be insignificant.

Regional characteristics also seem to play a role. People living in rural areas tend to have a higher need for credit as compared to people living in urban areas. People living in the province of Balochistan and KPK also seem to show a lower need for credit as compared to people living in Punjab.

Table 3.2: Econometric estimates for the Need for Credit

VARIABLES	**(1) Need for Credit**	**(2) Need for Credit**
Economic Profile (Base: 1st Quartile)		
2nd Quartile	-0.069***	-0.030
	(0.018)	(0.020)
3rd Quartile	-0.137***	-0.072***
	(0.023)	(0.019)
4th Quartile (Rich)	-0.180***	-0.109***
	(0.026)	(0.023)
Female Dummy	-0.054*	-0.009
	(0.028)	(0.026)
Education Level (Base: No formal Education)		
Primary Education	-0.028	-0.019
	(0.019)	(0.020)
Secondary Education	-0.062***	-0.039*
	(0.022)	(0.021)
Higher than Secondary	-0.049*	-0.067**
	(0.027)	(0.029)
Employment Status (Base: Salaried Full Time)		
Salaried Part time	0.024	0.055
	(0.032)	(0.034)
Self Employed	0.008	-0.006
	(0.030)	(0.039)
Unemployed Looking for a job	0.047	0.113**
	(0.044)	(0.048)
Out of Labor Force	0.036	0.019
	(0.032)	(0.034)
Age	0.001	0.004
	(0.004)	(0.003)
Age Squared	-1.17e-05	-5.94e-05
	(4.12e-05)	(3.71e-05)
Urban Dummy	-0.021	-0.055**
	(0.027)	(0.022)
Household Head Dummy	-0.008	-0.003
	(0.021)	(0.025)
Marital Status (Base: Single)		

Married	0.031	0.023
	(0.020)	(0.025)
Divorced	0.095**	0.122**
	(0.041)	(0.051)
Widowed	0.105	0.067
	(0.069)	(0.065)
Financial Literacy		-0.020
		(0.020)
Multidimensional Poverty Index (MPI)	0.476**	
	(0.187)	
Concentration of Bank Branches	48.08	
	(46.37)	
Province (Base: Punjab)		
Sindh	0.121***	
	(0.032)	
KPK	-0.300***	
	(0.066)	
Balochistan	-0.156**	
	(0.069)	
District Fixed Effects	**No**	**Yes**
Year Fixed Effects	**Yes**	**No**
Observations	**12,000**	**5,650**
Wald (Prob > chi2)	**0.00**	**0.00**

*Note: Results of the first step of the sequential logit regression have been reported with average marginal effects. Standard errors are in parentheses (clustered at the district level). The dependent variable is need for credit dummy (see equation 3.1). The financial literacy variable is only available for one of the two years and the number of observations go down by about half in regressions that include financial literacy (column 2). ***, ** and * represent significance at 1%, 5% and 10%, respectively.*

3.6.2. Drivers of Involuntary Financial Exclusion

This section presents the results of the second stage of the sequential logit model that aims to explore the main determinants of involuntary financial exclusion. The total sample in this case comprises of individuals who exhibited a need for credit in the first step. The dependent variable here is comprised of three categories i.e. individuals that are either financially included or financially excluded voluntarily or financially excluded involuntarily. The definition of involuntary financial exclusion refers to individuals who cannot use formal credit because of involuntary constraints despite having a need for it[49]. Column 1 and 2 of Table 3.3 present the results of the first regression with provincial fixed effects and column 3 and 4 present the results of the second regression with district fixed effects and financial literacy variable.

The results highlight that financially illiteracy, poverty and gender are strongly associated with involuntary financial exclusion as compared to voluntary financial exclusion. In terms of the

[49] See Section 2.3.3. for more details.

economic status, people belonging to the 1st quartile (i.e. the poorest households who belong to the bottom 25%) come out to be more financially excluded involuntarily than individuals in the 2nd or 3rd quartile. This shows that despite a high need for credit amongst individuals in the 1st quartile, this group experiences the highest amount of involuntary financial exclusion. Moreover, individuals who are financially literate are less likely to be financially excluded involuntarily as compared to individuals who are financially illiterate. This can be explained by the fact that financially literate people are more informed and are likely to make better financial decisions thus leading to lower level of involuntary financial exclusion. These results for the variables on economic profile and financial literacy are statistically significant at a 10% level.

Gender also seems to play a significant role as women are more likely to be financially excluded involuntarily as compared to being financially excluded voluntarily. Being a woman increases the chances of being involuntarily financially excluded by about 16.8 percentage points as compared to being a man. This finding is in line with the literature that highlights that women in developing countries are more credit constrained as compared to men (Ghosh and Vinod, 2017; Malapit, 2012). This can be explained by the fact that financial products might not be well designed according to their needs and perhaps women are discriminated against in the credit market because they generally possess less collateral as compared to men in countries such as Pakistan.

In terms of employment status, the findings highlight that part time salaried individuals and individuals who are out of the labor force are more likely to be financially excluded involuntarily as compared to salaried employees. This can be explained by the fact that not having a stable source of income makes that individual a riskier borrower and thus potential lenders would be reluctant in providing credit to them. Another interesting finding is that regional characteristics other than poverty and supply of financial services seem to play a significant role as well. Being from the province of Balochistan increases the probability of involuntary financial exclusion and being from Sindh decreases the probability of involuntary financial exclusion as compared to being from the province of Punjab. This suggests that regional characteristics other than poverty and supply of financial services such as law and order, infrastructure and climate of doing business might have a significant influence on financial exclusion.

Table 3.3: Econometric estimates for Involuntary Financial Exclusion

	Model 1		Model 2	
VARIABLES (Base: Voluntary Financial Exclusion)	(1) Financially Included	(2) Involuntary Financial Exclusion	(3) Financially Included	(4) Involuntary Financial Exclusion
Economic Profile (Base: 1st Quartile)				
2nd Quartile	0.013***	-0.011	0.020**	-0.055*
	(0.005)	(0.025)	(0.008)	(0.030)
3rd Quartile	0.007	-0.021	0.008	-0.062*
	(0.004)	(0.028)	(0.008)	(0.034)
4th Quartile (Rich)	0.007	0.008	0.008	-0.068
	(0.005)	(0.029)	(0.008)	(0.042)
Female Dummy	0.011	0.172***	0.008	0.168***
	(0.007)	(0.047)	(0.010)	(0.051)
Employment Status (Base: Salaried Full Time)				
Salaried Part time	-0.013*	0.015	-0.011	0.074*
	(0.007)	(0.039)	(0.010)	(0.042)
Self Employed	-0.01*	0.099***	-0.001	0.015
	(0.006)	(0.037)	(0.008)	(0.047)
Unemployed Looking for a job	-0.008	0.131**	0.009	0.110
	(0.01)	(0.065)	(0.016)	(0.078)
Out of Labor Force	-0.021***	0.003	-0.017**	0.065*
	(0.006)	(0.043)	(0.008)	(0.039)
Age	0.003***	-0.009***	0.004***	-0.003
	(0.001)	(0.003)	(0.001)	(0.003)
Age Squared	-2.78e-05***	8.25e-05**	-4.83e-05***	2.43e-05
	(8.77e-06)	(3.86e-05)	(1.79e-05)	(3.91e-05)
Education Level (Base: No formal Education)				
Primary Education	0.008*	-0.037	0.003	-0.009
	(0.005)	(0.027)	(0.007)	(0.027)
Secondary Education	0.007	-0.035	0.007	-0.021
	(0.005)	(0.027)	(0.007)	(0.023)
Higher than Secondary	0.016**	0.011	0.017**	0.032
	(0.007)	(0.032)	(0.008)	(0.036)
Urban Dummy	-0.002	0.009	0.005	0.0001
	(0.006)	(0.027)	(0.008)	(0.034)
Household Head Dummy	0.003	0.020	0.004	0.026
	(0.005)	(0.030)	(0.007)	(0.028)
Marital Status (Base: Single)				
Married	0.001	0.010	0.007	0.001
	(0.006)	(0.030)	(0.007)	(0.026)
Divorced	0.004	0.061	0.011	0.075
	(0.013)	(0.051)	(0.016)	(0.054)
Widowed	-0.232***	0.182*	-1.423***	0.724***
	(0.035)	(0.109)	(0.064)	(0.097)
Financial Literacy Dummy			0.003	-0.056*
			(0.005)	(0.033)
Multidimensional Poverty Index (MPI)	-0.015	0.208		
	(0.024)	(0.195)		
Concentration of Bank Branches	-9.015	67.88		
	(10.08)	(53.70)		

Province (Base: Punjab)				
Balochistan	-0.034**	0.286***		
	(0.017)	(0.092)		
KPK	-0.050**	0.037		
	(0.019)	(0.123)		
Sindh	0.005	-0.116**		
	(0.008)	(0.050)		
District Fixed Effects	**No**		**Yes**	
Year Fixed Effects	**Yes**		**No**	
Observations	**7,504**		**3,393**	
Wald (Prob > chi2)	**0.00**		**0.00**	

*Note: Results of the second step of the sequential logit regression have been reported with average marginal effects (see equation 3.2). Standard errors are in parentheses (clustered at the district level). The dependent variable is financial inclusion status (1= financially included, 2= involuntary financial exclusion, 3= voluntary financial exclusion (baseline)). The financial literacy variable is only available for one of the two years and the number of observations go down by about half in regressions that include financial literacy (column 3 and 4). ***, ** and * represent significance at 1%, 5% and 10%, respectively.*

3.6.3. Heckman estimation to address selection bias

Although a sequential logit model is very useful in modelling sequential decisions, Cameron and Heckman (1998) highlight that the presence of unobserved heterogeneity might cause a bias in the sequential logit results. In order to make sure that our results are reliable after dealing with the selection bias issue, we rely on a Heckman selection model proposed by Heckman (1979) to investigate the determinants of involuntary financial exclusion. In our case, the selection function i.e. need for credit is similar to a regular binary logistic regression, whereas, the dependent variable in the response function is binary as well. It takes the value of 1 if the individual exhibits involuntary financial exclusion (i.e. it includes individuals who are credit constrained) and 0 otherwise (i.e. individuals who are not credit constrained- this includes individuals who are financially included and individuals who chose voluntary financial exclusion). The exclusion variable that we rely on to carry out this estimation is the number of children in the household. The assumption is that a household with more number of children is likely to have a higher need for credit, whereas, it should not have any direct effect on not being able to obtain financial services from a formal financial institution and being involuntarily financially excluded. The results of the first stage corroborate this reasoning (Appendix 3.7).

The results of the Heckman probit estimation mostly confirm the results of our earlier estimations. According to these results, financial illiteracy is a main determinant of involuntary financial exclusion. An individual who is financially literate is about 5.8 percentage points less likely to be financially excluded involuntarily as compared to an individual who is financially illiterate. Moreover, women also come out to be 19.7 percentage points more likely to be financially excluded involuntarily as compared to men. This result is statistically significant at a 1% level. District poverty also seems to be a strongly associated with involuntary financial exclusion. However, in case of the Heckman estimation results, the variable on economic profile of the individual comes out to be insignificant. Lastly, regional characteristics other than multidimensional poverty also seem to play a significant role in driving involuntary financial exclusion. Belonging to Balochistan increases the chances of being financially excluded involuntarily by about 35.3 percentage points as compared to being from Punjab, similarly, being from Sindh or KPK decreases the chances of being financially excluded involuntarily as compared to being from Punjab. The inverse mills ratio, or lambda, is also calculated and reported in Table 3.4. It comes out to be statistically significant at a 5% level, indicating that a sample selection bias problem exists which is corrected in the hackman estimation.

Table 3.4: Heckman estimation results: Drivers of involuntary financial exclusion

VARIABLES	Involuntary Financial Exclusion (1)
Economic Profile (Base: 1st Quartile)	0.003
2nd Quartile	(0.02)
	-0.025
3rd Quartile	(0.02)
	-0.035
4th Quartile (Rich)	(0.03)
	0.003
Female Dummy	0.197***
	(0.03)
Financial Literacy Dummy	-0.058**
	(0.02)
Employment Status (Base: Salaried Full Time)	
Salaried Part time	0.044
	(0.03)
Self Employed	0.017
	(0.03)
Unemployed Looking for a job	0.093
	(0.07)
Out of Labor Force	0.021
	(0.03)
Age	-0.003
	(0.003)
Age squared	1.98e-05
	(3.88e-05)
Education Level (Base: No formal Education)	
Primary Education	-0.009
	(0.02)
Secondary Education	-0.015
	(0.02)
Higher than Secondary	0.001
	(0.03)
Urban Dummy	-0.002
	(0.02)
Household Head	0.059***

	(0.02)
Marital Status (Base: Single)	
Married	0.016
	(0.02)
Divorced	0.107*
	(0.06)
Widowed	0.114
	(0.08)
Multidimensional Poverty Index (MPI)	0.235***
	(0.07)
Concentration of Bank Branches	-0.0001
	(0.00)
Province (Base: Punjab)	
Sindh	-0.127***
	(0.03)
KPK	-0.169***
	(0.02)
Balochistan	0.353***
	(0.04)
Lambda	0.24**
Observations	**6,000**
Wald (Prob > chi2)	**0.00**

*Note: Results of the Heckman probit estimation have been reported with average marginal effects. Standard errors are in parentheses (clustered at the district level). The dependent variable is involuntary financial exclusion dummy. The financial literacy variable is only available for one of the two years, so the results are based on a cross-sectional data. ***, ** and * represent significance at 1%, 5% and 10%, respectively.*

3.7. Methodological Issues and Limitations

Although we have relied on a number of techniques to offer robust results, this study is not without certain limitations. This section discusses the main limitations and highlights some of the main methodological issues.

There is an issue of inaccurate reporting of credit related information in surveys. According to the financial inclusion insight survey of 2016, 1.2% of adult population in Pakistan has a formal loan[50]. Similarly, World bank's Findex survey of 2017 shows that 2.3% of the adult population in Pakistan took a formal loan in the last year. However, the numbers from regulatory authorities (the supply side) portray a slightly different picture. According to Pakistan Microfinance statistics, there are about 5.8 million active borrowers (PMN, 2017) which suggests that about 4% of the adult population in Pakistan has a formal microcredit that they are currently utilizing and repaying. This highlights that when it comes to reporting of access to credit in surveys, there is an issue of underreporting as people do not feel comfortable talking about it. This issue of underreporting in

[50] Including microfinance organizations as well as commercial banks.

surveys is well documented in the literature (Dev, 2006) and it affects the measurement of financial inclusion. This highlights the need to have improved data quality measures in place to ensure better policies.

It is also not straightforward to differentiate between reasons of financial exclusion from the data. The question used in this survey to ask *What is the main reason you do not borrow from a formal financial institution?* only allows respondents to choose one option. However, there could be multiple reasons behind why someone does not want to use a formal financial institution for borrowing. Furthermore, people who say that they do not borrow from a formal financial institution because they can borrow through other preferred means might need further probing. This might be driven either by their personal preferences or maybe some of this is driven by unsuitable product design of formal financial institutions (although unsuitable product design is given as a separate option in the survey). Hence, it is not particularly forthright to classify reasons of not using a formal source to borrow into voluntary and involuntary categories. This underlines the need to collect more detailed data and perhaps rely on qualitative measures to have a better understanding of the reasons behind financial exclusion. Moreover, since the demand for credit from formal institutions depends on the nature of the contract, it would be interesting to dig deep into the conditions which would make people switch between voluntary and involuntary exclusion.

Another type of financial exclusion that is important to look into is due to quantity rationing i.e. people needing more credit than what they have borrowed. However, the survey does not contain any information on such individuals and the amounts that they borrow. If the individual decides not to borrow from a formal financial institution because the amount offered was less than what he or she wanted, those individuals are included in the financial exclusion category. However, since there is no separate option for quantity rationing, it is not possible to identify and sort out these individuals from the others.

Selection bias is another concern that needs to be addressed. Since we argue that financial inclusion follows a two-step sequential decision, it is important to estimate the two equations separately. We try to address this issue by relying on a Heckman estimation, but our model might suffer from a weak exclusion variable issue. However, due to data limitations, we believe it is not possible to come up with a better exclusion variable in this case. Other similar studies have also reached the same conclusion about difficulties in finding an exclusion variable (Allen et al., 2016).

Lastly, due to certain factors not captured in the survey (e.g. interest rate charged, multiple borrowing), and, therefore not included in the model, we are basically analyzing associations between variables rather than causality. As discussed during the chapter, identifying the exclusion status of people who do not participate in the credit market is very challenging. A change in the characteristics of the contract offered by the lender can allow people to switch from voluntary to involuntary exclusion or vice versa which needs to be further analyzed. As more detailed data from the demand side as well as the supply side becomes available, many other dimensions and covariates should be incorporated in the understanding of different types of financial exclusion for a more comprehensive subsequent analysis.

3.8. Conclusion and Policy Recommendations

Financial exclusion is a topic of very high interest for policy makers and hence it is important to learn about the demand side drivers of financial exclusion and to come up with national level policies to tackle this issue. In this chapter, we deviate from the conventional way of looking at financial exclusion in two ways a) by considering that the need for credit is not universal and incorporating that in the econometric strategy and 2) by dividing financial exclusion into two different types i.e. voluntary and involuntary. The involuntary financial exclusion is the stricter version of exclusion where individuals are credit constrained. These individuals have a need for credit, and they want to use a formal source of finance, but due to market failures and asymmetry of information, they are incapable of doing so. This research critically analyses the main drivers of this involuntary financial exclusion and studies its interface with poverty, financial literacy and gender. The main results suggest that financial illiteracy and poverty level are amongst the variables which are strongly associated with involuntary financial exclusion. We also find evidence that suggests that women are more likely to be financially excluded involuntarily as compared to men.

Pakistan has devised a National Financial Inclusion Strategy (NFIS) that aims to expand access to financial services to the financially excluded segment of the population. This research helps to identify policy messages which can help in alleviating the current level of involuntary financial exclusion. Our results indicate that improving the level of financial literacy makes a significant difference in reducing involuntary financial exclusion. Financially illiterate people do not possess a good understanding of basic financial concepts and they do not have the ability to make well informed decisions pertaining to financial management. Pakistan is ranked 108 out of 144 countries in terms of financial literacy rate (World Bank, 2014). Only one out of four Pakistanis are financially literate, and this is well below the world average of 37%[51]. The recently launched National Financial Literacy Program by the State Bank of Pakistan is a step in the right direction but a lot needs to be done to reach out to the financially illiterate and unbanked population in provinces of Balochistan and KPK, and especially the rural areas.

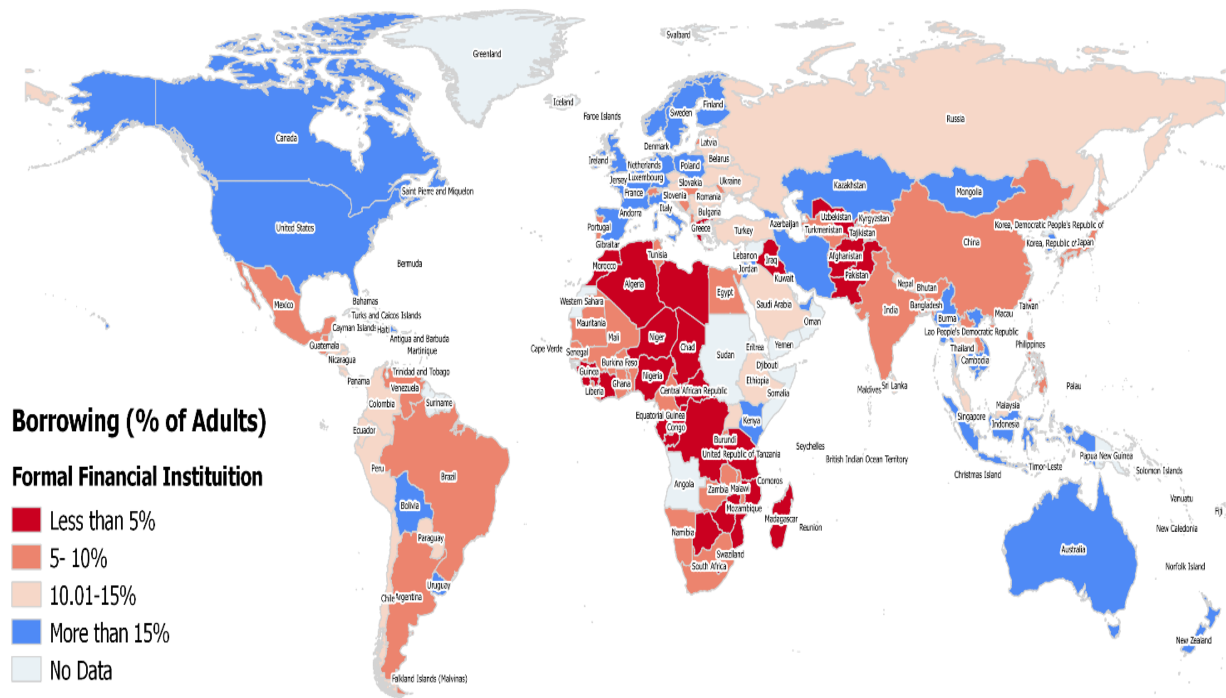
In line with the existing literature, our results also show that women encounter higher levels of involuntary financial exclusion and are more credit constrained as compared to men. These constraints hinder the use of formal credit amongst women and this asks for policy responses which can deal with this gender disparity. A large number of women highlight complicated paperwork, inappropriate product design, and outreach as the main reasons behind financial exclusion. Technology driven policy responses can play a vital role by decreasing the operational cost of financial institutions and allowing them to reach out to the financially excluded. Some examples of technology driven responses include agent banking, tablet banking and credit risk assessment based on machine learning and big data. For example, by relying on machine learning rather than conventional credit risk ratings, institutions in Pakistan such as *Tez Financial Services* are providing credit to customers with little or no credit history that can go a long way in alleviating financial exclusion especially for women. There are many potential benefits associated with adopting financial technology to extend credit outreach, and the regulatory environment should encourage the use of these avenues. However, more research still needs to be conducted to have a better understanding of potential shortcomings of these technology driven initiatives that mainly revolve around client protection, data privacy and other financial risks.

[51] The level of financial literacy is measured by the World Bank using the same methodology that we have used in our analysis. Appendix 3.6 highlights how financial literacy in Pakistan compares with the rest of the world.

Designing credit products keeping in mind the specific needs of women can also go a long way in reducing financial exclusion for women. It seems worth considering expanding the range of identification documents and easing collateral related information, especially for women working in the informal sector and residing in rural areas who do not possess this information. Dedicated financial institutions who mainstream gender consideration in their product design also help to alleviate financial exclusion. One such example of this is *Kashf Foundation*, a microfinance institution in Pakistan, which provides financial services primarily to women and designs products specific to their needs. More than 90% of their clients are women and the majority of their loan officers are also women. Their financial products are also tailor made for women borrowers. A main shortcoming in coming up with the right financial products for women is the lack of availability of disaggregated data by gender. Central banks worldwide are increasingly recognizing this issue and are engaging with financial institutions to collect data that is disaggregated by gender. This measure can also be helpful in reducing the status of financial exclusion.

3.9. Appendix

Appendix 3.1: Percentage of Formal Borrowing Worldwide



Source: World Bank's Findex Database, 2017

Appendix 3.2: Descriptive statistics

A. Need for Credit

	Credit Need (Mean)	No Credit Need (Mean)	Difference
Female	.526	.458	.069***
Urban	.366	.318	.049***
No formal Education	.283	.37	-.086***
Primary Education	.227	.223	.004
Secondary Education	.364	.305	.059***
Higher Education	.124	.1	.024***
Balochistan	.044	.053	-.009**
KPK	.238	.079	.159***
Punjab	.592	.567	.025***
Sindh	.127	.301	-.175***
1st Quantile	.183	.329	-.146***
2nd Quantile	.239	.252	-.013
3rd Quantile	.29	.227	.064***
4th Quantile (Richest)	.288	.194	.095***
Observations	**7,504**	**4496**	

*Note: This table represents the results of the t-tests. *, ** and *** denote statistical significance at 10%, 5% and 1% respectively.*

B. Financially Included

	Not Financially Included (Mean)	Financially Included (Mean)	Difference
Female	.484	.445	.04
Urban	.336	.373	-.037
No formal Education	.338	.275	.064*
Primary Education	.224	.268	-.044
Secondary Education	.328	.287	.041
Higher Education	.109	.17	-.061**
Balochistan	.051	.006	.044**
KPK	.14	.006	.134***
Punjab	.575	.647	-.072*
Sindh	.234	.34	-.106***
1st Quantile	.275	.183	.092**
2nd Quantile	.246	.321	-.074**
3rd Quantile	.25	.255	-.004
4th Quantile (Richest)	.229	.242	-.013
Observations	**7,504**	**4,496**	

*Note: This table represents the results of the t-tests. *, ** and *** denote statistical significance at 10%, 5% and 1% respectively.*

C. Voluntary Financial Exclusion

	No Voluntary Exclusion (Mean)	Voluntary Exclusion (Mean)	Difference
Female	.516	.385	.132***
Urban	.339	.326	.014
No formal Education	.331	.357	-.026***
Primary Education	.224	.226	-.002
Secondary Education	.331	.319	.012
Higher Education	.113	.098	.016***
Balochistan	.06	.018	.043***
KPK	.16	.074	.086***
Punjab	.585	.549	.035***
Sindh	.196	.358	-.163***
1st Quantile	.256	.328	-.071***
2nd Quantile	.247	.249	-.003
3rd Quantile	.256	.234	.021**
4th Quantile (Richest)	.242	.19	.052***
Observations	**9,029**	**2,971**	

*Note: This table represents the results of the t-tests. *, ** and *** denote statistical significance at 10%, 5% and 1% respectively.*

D. Involuntary Financial Exclusion

	No Involuntary Exclusion (Mean)	Involuntary Exclusion (Mean)	Difference
Female	.469	.508	-.038***
Urban	.35	.311	.041***
No formal Education	.312	.382	-.07***
Primary Education	.227	.22	.007
Secondary Education	.345	.297	.048***
Higher Education	.115	.1	.015**
Balochistan	.034	.079	-.045***
KPK	.17	.085	.085***
Punjab	.576	.575	.001
Sindh	.221	.262	-.041***
1st Quantile	.239	.334	-.095***
2nd Quantile	.245	.251	-.006
3rd Quantile	.268	.221	.047***
4th Quantile (Richest)	.249	.195	.055***
Observations	**7,620**	**4,380**	

*Note: This table represents the results of the t-tests. *, ** and *** denote statistical significance at 10%, 5% and 1% respectively.*

Appendix 3.3: Variables and their Sources

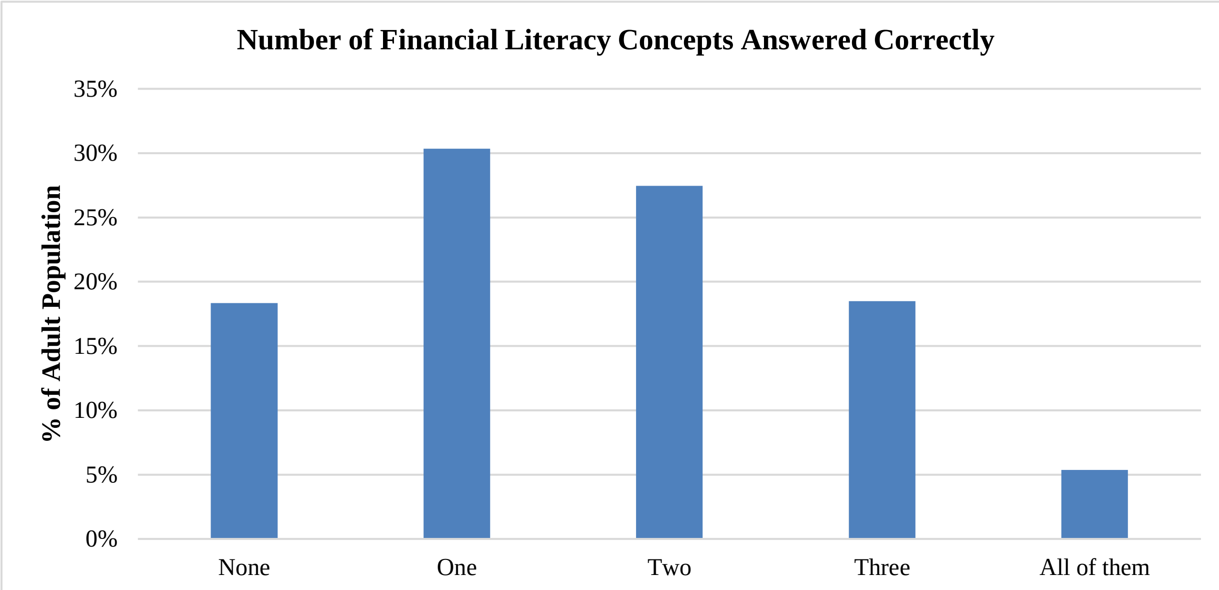
Variables	Source	Coverage
Credit Need	Financial Inclusion Insight Survey	2015 & 2016
Financial Inclusion	Financial Inclusion Insight Survey	2015 & 2016
Voluntary Financial Inclusion	Financial Inclusion Insight Survey	2015 & 2016
Involuntary Financial Inclusion	Financial Inclusion Insight Survey	2015 & 2016
Female	Financial Inclusion Insight Survey	2015 & 2016
Financial Literacy Indicator	Financial Inclusion Insight Survey	2016 Only
Age	Financial Inclusion Insight Survey	2015 & 2016
Age Squared	Financial Inclusion Insight Survey	2015 & 2016
Employment Status	Financial Inclusion Insight Survey	2015 & 2016
Household Head Dummy	Financial Inclusion Insight Survey	2015 & 2016
Rural_Dummy	Financial Inclusion Insight Survey	2015 & 2016
Household Size	Financial Inclusion Insight Survey	2015 & 2016
Economic Profile_PPI Score	Financial Inclusion Insight Survey	2015 & 2016
Marital Status	Financial Inclusion Insight Survey	2015 & 2016
Province	Financial Inclusion Insight Survey	2015 & 2016
Poverty Rate District	UNDP	2015-2016
Bank Accounts per adults	Pakistan Microfinance Network (PMN)	2015 & 2016
Concentration of bank branches: Branches per 1000 adults	Pakistan Microfinance Network (PMN)	2015 & 2016

Appendix 3.4: Details about the Financial Literacy Measurement

Concept	Question	Options
Risk Diversification	Suppose you have some money. Is it safer to put your money into one business or investment, or to put your money into multiple businesses or investments?	1=One business or investment 2=Multiple businesses or investments 99=Don't Know
Inflation	Suppose over the next 10 years the prices of the things you buy double. If your income also doubles, will you be able to buy less than you can buy today, the same as you can buy today, or more than you can buy today?	1=Less 2=The same 3=More 99= Don't Know
Interest Rate	Suppose you need to borrow 100 PKR. Which is the lower amount to pay back: 105 PKR or 100 PKR plus 3 percent of 100 PKR?	1=105PKR 2=100PKR plus 3 percent of 100PKR 99= Don't Know
Interest Rate Compounding	Suppose you put money in the bank for two years and the bank agrees to add 15 percent per year to your account. Will the bank add more money to your account the second year than it did the first year, or will it add the same amount of money both years?	1=More 2=The same 99= Don't Know
Interest Rate Compounding	Suppose you had 100 PKR. in a savings account and the bank adds 10 percent per year to the account. How much money would you have in the account after five years if you did not remove any money from the account?	1=More than 150 PKR 2=Exactly 150 PKR 3=Less than 150 PKR 99= Don't Know

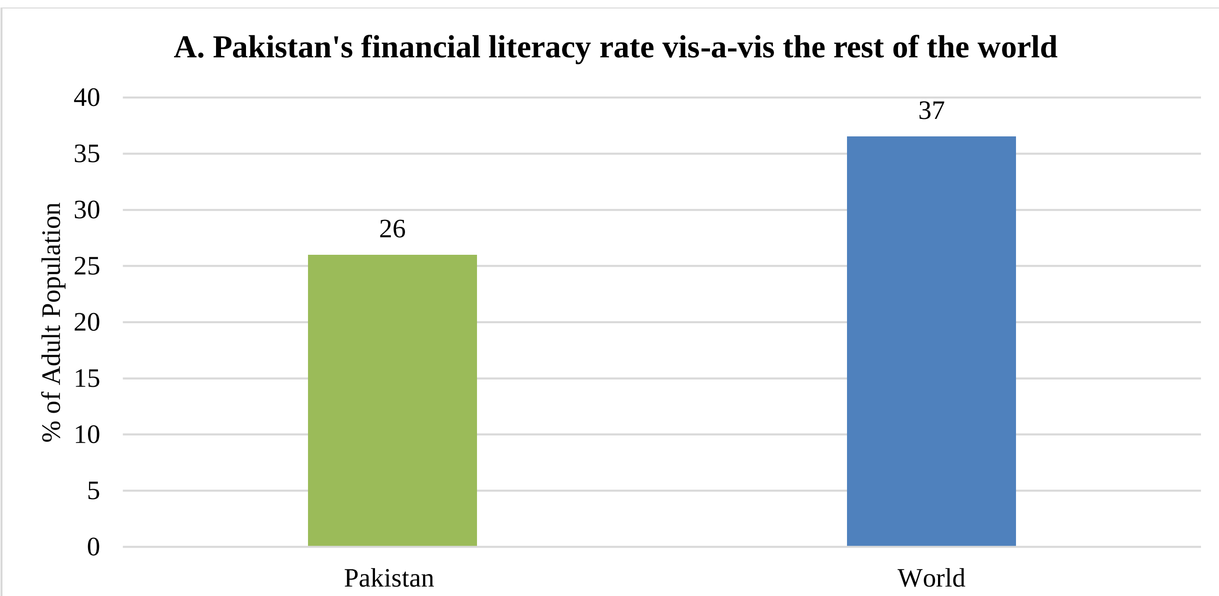
Source: Financial Inclusion Insight Survey (2016)

Appendix 3.5: Financial Literacy Scores



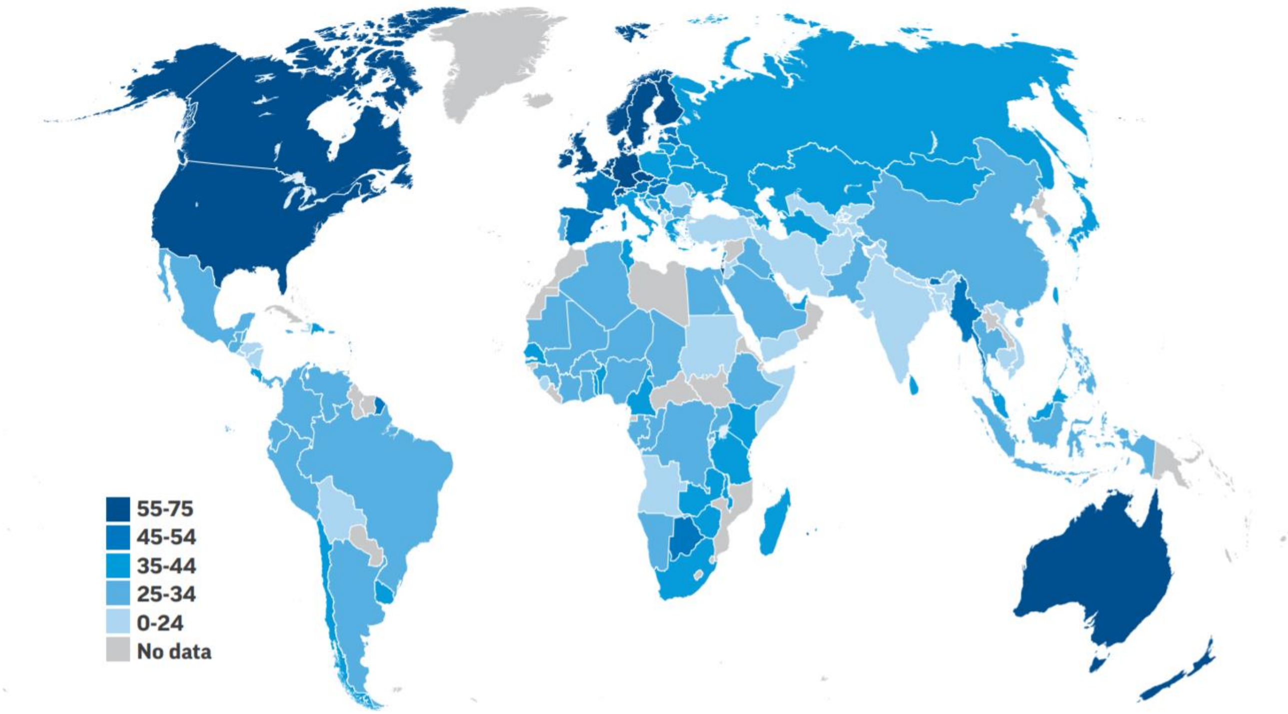
Source: Financial Inclusion Insight Survey (2016)

Appendix 3.6: Comparison of financial literacy with the rest of the world



Source: World Bank, 2014
Note: The world average is based on data from 144 countries.

B. Financial literacy rate worldwide



Source: World Bank, 2014

Appendix 3.7: First stage (selection equation) results of the Heckman estimation

VARIABLES	Need for Credit (1)
Number of children in the household	0.098***
	(0.028)
Provincial Fixed Effects	**Yes**
Observations	**6,000**
Wald (Prob > chi2)	**0.00**
*** p<0.01, ** p<0.05, * p<0.1	

Note: Controls for individual characteristics and regional characteristics have been included (not reported).

Chapter 4: Assessing Household Financial Vulnerability using Machine Learning

Empirical Evidence from the U.S.

This chapter is based on a paper co-authored with Damien AZZOPARDI, Patrick LENAIN, and Douglas SUTHERLAND.

It has been published as a chapter in an OECD book and cited as below:

Azzopardi, D., Fareed, F., Lenain, P. and Sutherland, D. (2019), "Assessing Household Financial Vulnerability: Empirical Evidence from the U.S. using Machine Learning", in *OECD Economic Survey of the United States: Key Research Findings,* OECD Publishing, Paris, 123- 145.

Abstract

Household financial vulnerability has gained considerable attention since the global financial crisis and it concerns policymakers due to its impact on macroeconomic indicators of financial instability. However, financial vulnerability is a complex multidimensional concept and large gaps remain in the underpinning of a comprehensive financial vulnerability assessment at the micro level. In this chapter, a new approach is proposed to assess financial vulnerability by employing an unsupervised machine learning technique. A two-step empirical strategy is used to conduct this analysis. First, Hierarchical Ascending Clustering (HAC) and K-means clustering analysis is undertaken to identify homogenous clusters of households, which are financially vulnerable. Afterwards, we estimate the probability of being financially vulnerable depending upon different household and geographical characteristics using a logistic regression. Data from the Survey of Consumer Finance (SCF), for the years 1998, 2007 and 2016, is used for this analysis. The empirical results show that about 28% of the households in United States are financially vulnerable as of 2016, which is 4 percentage points less as compared to 2007. The results of the econometric estimates highlight that African-Americans are 8 percentage points and Hispanic Americans are 6 percentage points more likely to be financially vulnerable than non-Hispanic white persons, after taking into account other household and regional level characteristics. Econometric estimates also highlight the existence of large gaps in household financial vulnerability across other household characteristics, such as education level, employment status, marital status, and age of the household head. Lastly, regional characteristics do not seem to have a significant effect on household's financial vulnerability as long as the net-worth of the household is taken into account. All our empirical results are obtained after accounting for the imputation error as well as the sample variability error.

4.1. Introduction

Besides the importance of financial inclusion for development, there is another stream of literature that focuses on the negative implications of credit expansion and highlights the potential trade-offs between financial inclusion in credit and financial stability (Breza and Kinnan, 2018; Arcand et al., 2015; Mian and Sufi, 2014; Cecchetti and Kharroubi, 2012). Since the financial crisis of 2007-08, policy makers have increasingly focused on monitoring the financial health of indebted households in developing as well as developed countries. One of the reasons for this interest is that financial risks and vulnerabilities of households are generally analogous to that of financial institutions (Brown et al., 2010). A delay in the payment of debt or inability of a household to meet its financial commitments affect the bank's profitability and asset quality, and can result in financial instability (Acharya et al., 2009). Moreover, an increase in the level of indebtedness makes households more sensitive to economic shocks (Michelangeli and Ramapazzi, 2016; Herceg and Nestić, 2014). Bunn and Romstom (2015) argue that the run up in households' debt can also anticipate a financial crisis. The levels of household debt are starting to rise again in many countries since the financial crisis of 2007-08, and the central banks are increasingly concerned about the risks that this might pose to the macroeconomic financial stability (Federal Reserve, 2018; Reserve Bank of Australia, 2017)[52].

Financial vulnerability is a complex and multi-dimensional phenomenon and a single metric is not sufficient to fully capture its effect. That is why the definition of household's financial vulnerability remains quite vague, and there is a lack of consensus when it comes to defining or analyzing household's financial vulnerability. In broad terms, financially vulnerable households are the ones who are likely to default on their financial commitments if an adverse event happens. Leika and Marchettini (2017) provide a detailed summary of the existing empirical research on household financial vulnerability. They argue that most of the existing literature uses a very restricted definition of financial vulnerability by focusing only on the fragility of a household with

[52] Appendix 4.5 provides a chart on the current levels of household debt (as a percentage of GDP) for a number of developed economies including the U.S. Moreover, see Section III of the General Introduction for an overview of the literature that discusses the tradeoffs between financial inclusion and financial stability.

respect to its debt commitments, which provides an incomplete picture of financially vulnerable households. According to this approach, a household is termed as vulnerable if its indebtedness level or its debt service ratio exceeds a certain threshold. This line of research measures household financial vulnerability by either looking at solvency indictors such as debt to asset ratio (Albacete and Linder, 2013; Christelis et al., 2010; among others) or by looking at liquidity indicators such as debt service to income ratio (Michelangeli and Pietrunti, 2014).

Some studies have used a slightly broader notion of household's financial vulnerability by considering certain dimensions apart from debt, e.g., household expenditures, utility bills or rent payments (Anderloni et al., 2012; Worthington, 2006). However, these indicators are often criticized for their lack of cogently defined boundaries, along with the fact that most of them do not take into account other factors, such as income and wealth levels, differences in life-cycle stages, and other economic conditions among different households.

In this chapter, we propose a new methodological approach as an alternative to the standard line of research on household financial vulnerability. The main objective of this chapter is twofold. First, to use an unsupervised machine learning technique to create profiles of homogenous households that are financially vulnerable. Second, to analyze how household characteristics and geographical characteristics affect financial vulnerability.

The availability of nationally representative micro data in the U.S. allows us to identify the types of households that are more susceptible to the risk of being financially distressed and more likely to default on their financial commitments. We use the Survey of Consumer Finance (SCF) for the years 1998, 2007 and 2016 to conduct this analysis. First, instead of using a definition of financial vulnerability that relies on imposing a certain threshold on a single debt related indicator, we use Hierarchical Ascending Clustering (HAC) and K-means Clustering to identify homogenous clusters of financially vulnerable households. For that purpose, we use three sets of variables including the leverage ratio, debt burden, and the households' income level. Leverage ratio is measured as total debt over total assets, and the debt burden is measured by the ratio of monthly repayments to monthly income. The results show that 28% of the households in the U.S. are financially vulnerable in 2016, whereas the percentage of financially vulnerable households was 32% and 30.3% in 2007 and 1998, respectively. A comparison of households which are financially vulnerable vis- à- vis households which are not financially vulnerable indicates that vulnerable households earn a significantly lower income, have a larger debt burden, save less with respect to other households, and have fewer assets.

To further contribute to this debate, we estimate the probability of being financially vulnerable depending on different household and geographical characteristics, using a logistic regression. We analyze the differences in financial vulnerability across regions, ethnicities, gender, education level and other individual characteristics. To be more specific, we empirically analyze whether certain ethnic groups are more vulnerable than others, and whether certain geographical regions are more prone to financial vulnerability. Moreover, we examine how these findings vary across urban and rural areas, and across other household characteristics, such as employment status, education level, age, and the marital status of the household head. Lastly, we analyze what the evidence suggests regarding the gender gap in financial vulnerability of households. All of our results are established after accounting for the imputation error and the sample variability error.

The main findings suggest that household characteristics such as ethnicity and race, education level, age, marital status, and working status of the household head are amongst significant determinants of household's financial vulnerability. Econometric estimates show that African-Americans and Hispanics are financially more vulnerable as compared to non-Hispanic white Americans, after controlling for other household and regional characteristics. A higher education level of the household head also appears to be statistically significant and is negatively linked with financial vulnerability. On average, having a college degree decreases the probability of being financially vulnerable by about 11 percentage points as compared to having no high school diploma. Moreover, an increase in the age of the household head decreases the chances of being financially vulnerable. Being married and living with the spouse is also negatively associated with the probability of financial vulnerability. Our results also highlight that, even after controlling for traditional household and regional characteristics, there is clear evidence that an increase in the size of monthly debt repayment raises household's financial vulnerability, and an increase in the net-worth of the household decreases financial vulnerability. Lastly, regional characteristics such as living in an urban locality do not seem to have any association with households' financial vulnerability once we take into account the net-worth of the household.

This chapter adds to the existing literature on financial vulnerability in two different ways. First, according to the best of our knowledge, this is one of the earliest studies which empirically investigates the subject of household financial vulnerability using an unsupervised machine learning technique. Most of the existing analyses are based on aggregate data and cross-country analysis which have some significant shortcomings due to their inability to account for differences in distributions (Christelis et al., 2010; Vatne, 2006). Even the within country studies have not adequately addressed this issue as they concentrate on a constrained definition of financial vulnerability, by solely looking at a debt related ratio. Second, our analysis contributes to the existing literature on financial vulnerability by analyzing how racial background, along with other household and regional characteristics, has an influence on household's financial vulnerability.

The remainder of the chapter is organized as follows: Section 4.2 briefly reviews the existing empirical literature. Section 4.3 provides a broad overview of the common machine learning techniques. Section 4.4 goes on to discuss our two-step empirical strategy. It presents the results of the unsupervised machine learning approach in detail and then presents our econometric model for the assessment of households' financial vulnerability. Section 4.5 then provides the results of the econometric analysis, aiming at examining the key drivers of financial vulnerability. Section 4.6 provides a conclusion.

4.2. Empirical Literature on Measuring Household Financial Vulnerability

There is a vast amount of empirical literature that analyses households' financial vulnerability. This empirical literature can be broadly categorized under two main lines of research. On one hand, there is empirical literature which looks at household's financial vulnerability from a "macro perspective". On the other hand, there is a large volume of research that looks at the issue of household's financial vulnerability from a "micro perspective".

The literature which adopts the macro approach uses aggregate data in order to analyze the various channels and causes of households' indebtedness growth. However, numerous studies have highlighted several limitations of using aggregate data to analyze household's financial vulnerability. There is a general consensus in the existing literature about the inability of aggregate data to account for the differences in distributions between groups (Albacete and Fessler, 2010; Dey et al., 2008, among others). It is important to note that country level indicators like average household's debt to income ratio are quite useful in detecting the fluctuation in financial vulnerability over time, across different regions and countries. However, due to a potentially large variation in financial vulnerability between groups of households, using aggregate data on debt burden offers a coarse direction on the actual household vulnerabilities.

The most recent empirical literature on financial vulnerability uses household level microdata from surveys in order to identify the profiles and distribution of household vulnerabilities. Several academics and central banks around the world have developed their own indicators of household financial vulnerability by establishing different threshold levels. The most widely used methodology in this regard is to (a) establish a benchmark indicator to identify financially vulnerable households, and then (b) test for the effect of different economic shocks, policies, and household characteristics on financial vulnerability. Some of the standard indicators which are used to measure household financial vulnerability include (i) debt to asset ratio (DTA), (ii) debt to income ratio (DTI), and (iii) debt-service to income ratio (DSTI) (Bank of England, 2016; Michelangeli and Ramapazzi, 2016; among others). This substantial amount of literature has empirically looked at the determinants of financial distress as well as household debt burden (Anderloni et al., 2012; Christelis et al., 2010). However, these approaches provide a rather restricted way of assessing financial vulnerability by establishing a threshold for a debt related indicator. It is important to note that household's financial vulnerability can also be derived by factors apart from debt. These other factors may include aspects pertaining to low income and wealth level, non-optimal money management, and economic shocks, amongst other reasons. These reasons have been largely ignored by the existing empirical literature.

Some studies have also used a subjective methodological approach to gauge financial vulnerability of a household. This approach usually involves constructing a financial vulnerability indicator based on the household's self-assessment of their financial wellbeing. One major issue with this self-assessed financial wellbeing is that it often does not correlate with financial distress (Herrala and Kauko, 2007). It is often swayed by other factors such as comparisons with reference groups.

Keeping in mind the shortcomings of the existing methodologies, we propose a new methodological approach in this chapter by using an unsupervised machine learning technique to identify financially vulnerable households. This approach provides an alternative to the standard line of research on household financial vulnerability.

4.3. Overview of Machine Learning Methods

By and large, machine learning is an application of artificial intelligence that takes advantage of high computational power of machines to run algorithms and learn from the data. Models that fall within the confines of machine learning are generally non-parametric in nature[53], as compared to econometric models that are mainly parametrical. Parametric models rely on a functional form that link the outcome variable with the regressors based on certain assumptions. Besides the main advantages of parametrical models such as their ease of interpretability, fast computational speed, less reliance on large datasets, some of their main weaknesses include their limited predictive power and specification risks (James et al., 2013). Nonparametric models, however, make the minimum amount of assumptions regarding the functional form of the relationship in order to have more flexibility to learn from the data. Their main strength is associated with their flexibility and high predictive power, whereas one main shortcoming of these model is the risk of overfitting (Jordan and Mitchell, 2015)[54].

When it comes to machine learning, it is important to distinguish between two different types: supervised and unsupervised machine learning. In case of supervised learning, the machine is trained using data which is “labelled”. In this case, the data on the outcome variable is available and the supervised algorithm learns from this “labeled” data to predict the outcomes for unforeseen data. The objective in this case is to make predictions based on modeling the relationship between different features and the main outcome variable. For example, let’s suppose a financial institution is interested in predicting the chances of default on their agricultural loan based on the characteristics of the borrower (such as income level, debt, age, and profession among other characteristics), crop attributes (type of crop, seasonality and tenure among other attributes) and relevant economic indicators (climate, inflation and interest rates among other indicators). Now,

---

[53] There are some exceptions such as Least Absolute Selection and Shrinkage Operator.

[54] See Hastie, Tibshirani and Friedman (2016) for a detailed discussion on the comparison between econometric modelling and machine learning.

if the dataset already has representative information on past borrowers on their loan repayment history and their characteristics, then the machine learning model is called a supervised model. In short, observed data on defaults here is being used to train the algorithm and then to predict future defaults. These algorithms have recently gained a lot of attention from the financial sector. Financial institutions in some countries have already started using these algorithms for the purpose of credit risk analysis and to predict financial stability of new clients. Examples of the most commonly used machine learning techniques in this regard include the random forest, gradient boosting decision trees, neural networks, and the support vector machine algorithm (Bazarbash, 2019). Some examples of financial institutions who are using these machine learning algorithms for credit risk analysis include BBVA bank and Tez Financial Services.

In case of unsupervised machine learning models, the outcome variable is not available, and the models basically deal with “unlabeled data”. Since the information on the outcome variable is not available, the goal of the algorithm in this case is to find similarities among different features and group characteristics based on a set of identified variables. In the example that we discussed earlier about predicting defaults, let’s suppose now that the data is “unlabeled” and the variable on repayments is not available. In such case, unsupervised machine learning models can cluster borrowers in different groups with similar features, based on a set of relevant characteristics that explain loan repayment and default behavior. One common form of unsupervised machine learning models include the k-means clustering analysis and the hierarchical clustering analysis. We rely on the unsupervised machine learning models in order to identify homogenous clusters of households that can be classified as financially vulnerable. The details of the techniques which we employ in our case are discussed in detail in the next section.

4.4. Empirical Strategy

The empirical strategy that we use is divided into two parts. In the first part, we use unsupervised machine learning algorithms to identify clusters of households which are financially vulnerable. Afterwards, in the second part, we estimate the probability of being financially vulnerable depending on different household and geographical characteristics by using a logistic regression.

In order to conduct this analysis, we use the Survey of Consumer Finance (SCF). It is a triennial cross-sectional survey of U.S. households conducted by the Federal Reserve. We build a pseudo-panel of nationally representative data sets for the years 1998, 2007 and 2016 with a total of about 15,000 households in the dataset. The survey provides information pertaining to the debt burden of households, their income and wealth levels, credit commitments, expenditure vulnerabilities, saving patterns, socio-demographic characteristics, and other household characteristics. The details of our empirical strategy are provided below.

4.4.1. Assessing Financial Vulnerability using Unsupervised Machine Leaning

This section details the machine learning methodology applied to discover how the U.S households are structured in terms of financial vulnerability. Unsupervised machine learning techniques, and in particular clustering algorithms, make it possible to break down set of observations into several subsets that are fairly homogeneous in their characteristics. To cope with computational complexity of the algorithms, this approach adopts a two-step procedure using hierarchical and partitioning methods.

In the first step, the Hierarchical Ascending Clustering (HAC) is run on four different random samples chosen from the overall population. This establishes the optimal number of underlying clusters ‘K’ within the U.S households. The HAC organizes the clusters into a structured tree called a dendrogram, which is the conventional way of depicting HAC results. The ascending approach, or agglomeration, starts with an observation in each class, then successively merges the two closest classes, and stops when there is only one class containing all the observations. The numbers of clusters are not required to be predefined in this algorithm. The dendrogram represents the inclusive relationships of the clusters. The classification is obtained by cutting the tree at a given level and Figure 4.1 shows how a dendrogram helps in depicting different homogenous clusters.

In the second stage, the partitioning method called the K-means algorithm is applied on the whole survey. The K-means algorithm partitions the full set of observation into K numbers of clusters. Each cluster is represented by its gravity center; and an observation belongs to the cluster whose center is the nearest. The number of clusters must be defined prior to running the algorithm. The general framework for identifying homogenous groups of financial vulnerability is summarized below:

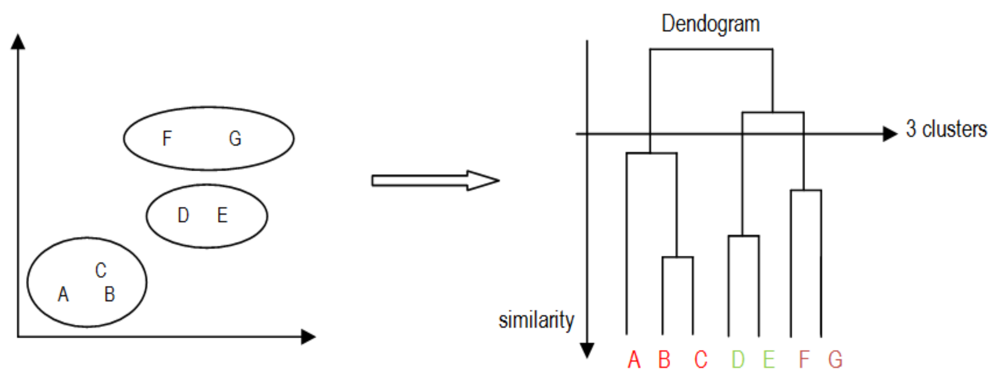
1- Selection of variables related to household’s financial vulnerability

2- Determining the optimal number of clusters 'K' using HAC algorithm

3- Partitioning overall sample into homogenous clusters using K-Means algorithm

4- Characterizing the clusters

Figure 4.1: Selection of clusters on a dendrogram



Selection of Variables

Financial vulnerability is a complex phenomenon and a single metric is not enough to fully gauge its effect. Therefore, instead of relying on a definition of financial vulnerability that requires using a certain threshold level for one dimension of financial vulnerability, we use three different variables and rely on the unsupervised machine learning technique to classify clusters of financially vulnerable households. These variables have been chosen considering the existing literature that identifies the pertinence of certain variables in explaining financial vulnerability. First, we use the ratio of monthly repayments to monthly income to reflect the burden of debt repayments for a household. A higher ratio indicates a higher debt burden which increases the chances of default on loan commitments. Second, we use the leverage ratio measured as total debt over total assets. It is a measure of solvency and financial flexibility of a household. Third, we use household income to reflect the needs and behaviors of the household since a household in the top quintile has different financial needs and different consumption patterns than the one in the bottom quintile. To normalize the distribution of these variables and to reduce the distance with the outliers, we do a logarithmic transformation of the income variable and we cap the ratios to two.

A summary of these variables across time is provided in Table 4.1. The level of debt burden remained constant from 1998 to 2007 at 18%, but it went down to 15% in 2016. The leverage ratio steadily rose from 34% in 1998 to 37% in 2007 and then to 41% in 2016. The average level of total debt has also increased more than the value of assets. A correlation matrix along with the scatter plot matrix for these variables is provided in Appendix 4.2.

Table 4.1: Summary statistics for debt burden, leverage ratio and income

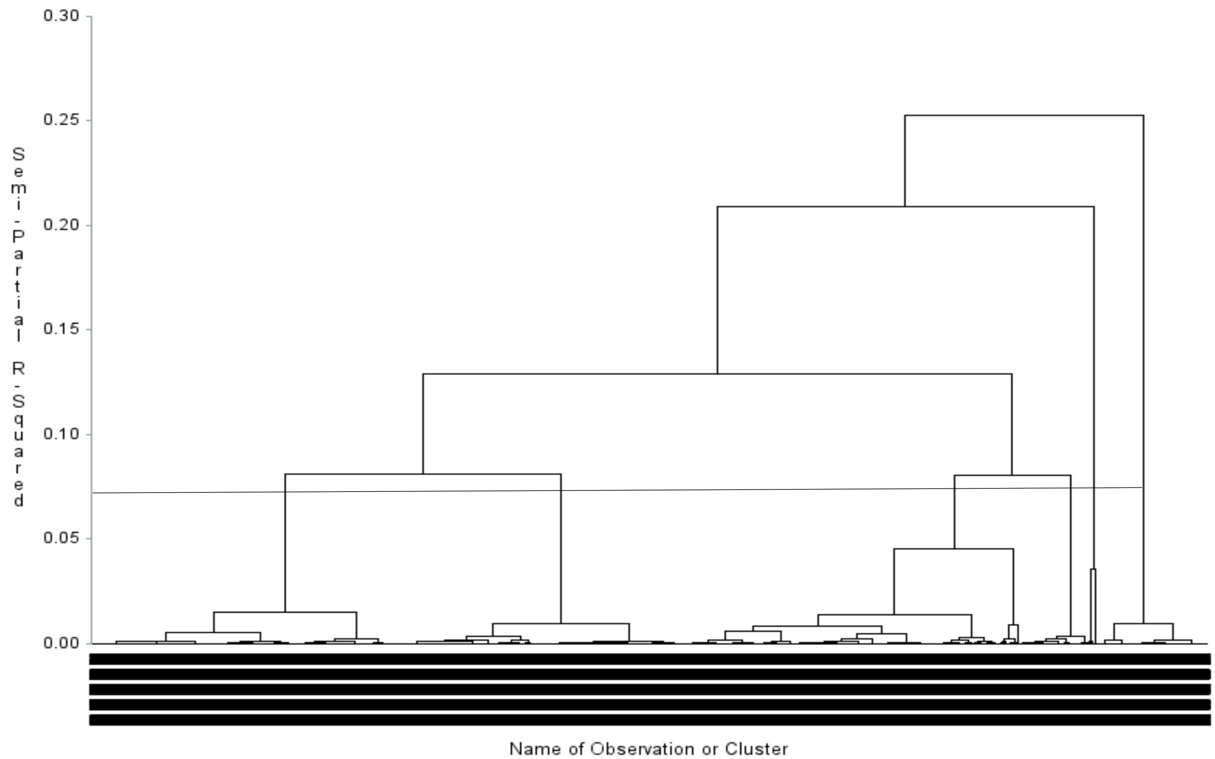
Variable	N	Year	Mean	Std. Dev	Minimum	Maximum
Debt Burden	31240	2016	0.15	12.9	0	2
Leverage Ratio	31240	2016	0.41	33.7	0	2
Household Income	31240	2016	10.90	74.0	0	19.5
Debt Burden	22085	2007	0.18	17.6	0	2
Leverage Ratio	22085	2007	0.37	33.7	0	2
Household Income	22085	2007	10.84	86.9	0	19.2
Debt Burden	21525	1998	0.18	17.5	0	2
Leverage Ratio	21525	1998	0.34	31.7	0	2
Household Income	21525	1998	10.61	108.9	0	19.4

Source: Survey of Consumer Finance

Determining the Optimal Number of Clusters

In order to find the underlying number of clusters amongst the overall population, a hierarchical ascending clustering (HAC) was conducted on four different random samples of 8000 households. The benefit of this technique is the tree-based representation of the observations on a figure known as the *dendrogram* which is particular easy to interpret. Figure 4.2 represents the results of the hierarchical ascending clustering for the year 2016. The results show a clear separation of six clusters on the dendrogram with the semipartial R-squared (SPRSQ) represented by the vertical line. The semipartial R-squared (SPRSQ) is a measure of the homogeneity of merged clusters. So, SPRSQ is basically considered as the loss of homogeneity due to combining two groups or clusters to form a new group or cluster. The value of SPRSQ should be small implying that, after merging, members of the two groups are homogenous. One drawback of the HAC is the minimum algorithm complexity in O (N^2), N being the number of observations which are too high for computational purposes (Murtagh and Contreras, 2012). Therefore, this is one of the reasons why the K-means algorithm is preferred for the full sample exercise, and HAC is used to determine the number of clusters.

Figure 4.2: Dendrogram from hierarchical ascending clustering (HAC)



Source: Authors' calculations based on Survey of Consumer Finance

K-means Clustering

In order to characterize all the households into different clusters, we carried out K-means clustering. The main benefit of the K-means algorithm is its algorithmic simplicity. However, it requires that the number of clusters are known prior to running the algorithm. Using the total number of clusters as six from HAC, the K-means algorithm carries out the following steps to classify the overall sample into clusters.

1. Conduct an initial cluster assignment for all the observations by randomly assigning a number from 1 to K to all the observations in the data

2. Keep the iteration process going until the cluster assignments stop changing:

   a. Compute the centroid for all of the K clusters. The kth cluster centroid is the vector of the p feature means for the observations in the kth cluster.

b. Based on the closest centroid, assign each observation according to the centroid. Euclidean distance is used to define the closest centroid.

Characterizing the Clusters

Hierarchical ascending clustering (HAC) and K-means clustering allowed us to classify all the households into six different clusters. The clusters significantly vary from each other with respect to different characteristics whereas the households within these clusters are homogenous. The main characteristics of each cluster are summarized below.

- Cluster 1: It is the smallest group with only 0.1% of the total households in it. This cluster contains households with almost no income, household head being out of the labor force, and over representation of single household heads.
- Cluster 2: The size of this cluster is small too and it contains only 0.4% of the total households. It represents households who are over indebted, have income below the median level, with debt payment to monthly income greater than 1. There is an over representation of self-employed people, households which have been turned down for credit applications, 55-64 year-olds, white non-Hispanics, and female headed households.
- Cluster 3: This cluster contains about 18% of the U.S. households. It represents households with high debt to monthly income ratio (an average of 0.40). This population is likely to have financial difficulties.
- Cluster 4: This cluster covers wealthy families. About 38% of the households belong to this cluster. Households in this cluster have high income and relatively lower debt to monthly income ratio. There is an over representation of savers, managers, male headed households, and white non-Hispanics households.
- Cluster 5: About 34% of the U.S. households belong to this category. This cluster contains households with low income, an over representation of households with no credit card balance, very low debt to asset ratio, and mostly retired and more than 55 years old individuals.
- Cluster 6: This cluster mainly represents households which have quite high debt to assets ratio and relatively low income. They have late repayments history, were denied credit

in the past year, with an over representation of 35 years-old, female household heads, African-Americans, not married, and no children. About 9.4% of the households belong to this cluster. A summary of all the clusters is provided in Table 4.2.

Table 4.2: Classification of clusters

Cluster Number	Percentage of Households	Brief Description of the Cluster Characteristics
1	0.1%	No income, household head not working
2	0.4%	Over indebted, low income
3	18.1%	Over indebted, debt payment to monthly income averages 0.4
4	38.0%	High income, low debt repayment to monthly income ratio, saves money
5	33.9%	Low debt no asset ratio, low income, over representation of retired people
6	9.3%	Over indebted, high debt to assets ratio, over representation of African-Americans.

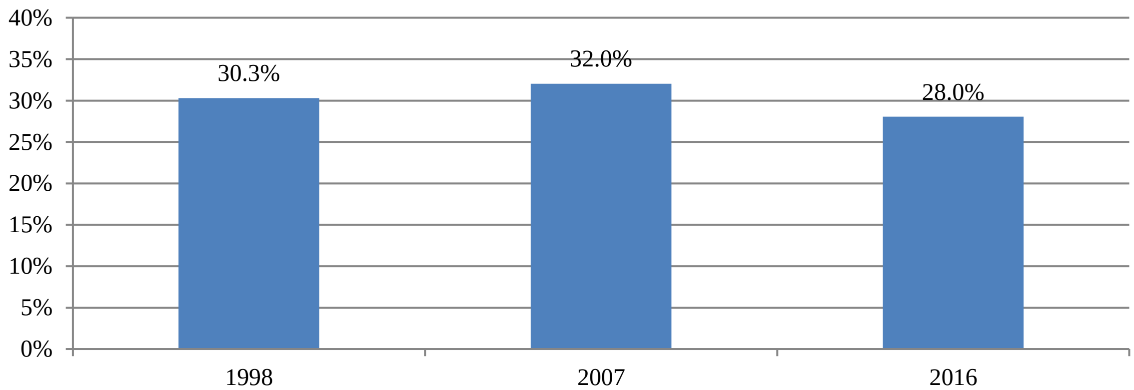
The clustering was run with the same center of clusters on the Survey of Consumer Finance for the years 2007 and 1998. The composition of the different clusters looks relatively stable across the three surveys. There is a peak of financial vulnerable households in 2007, mostly coming from cluster 3. The distribution of clusters across time is shown in Table 4.3.

Table 4.3: Distribution of clusters across time

	Year 2016	Year 2007	Year 1998
Cluster 1	0.1%	0.1%	0.9%
Cluster 2	0.4%	0.7%	1.1%
Cluster 3	18.1%	24.8%	21.8%
Cluster 4	38.0%	35.2%	33.2%
Cluster 5	33.9%	32.7%	36.5%
Cluster 6	9.4%	6.4%	6.4%

We use this clustering analysis to divide the households into two distinct categories i.e. households that are financially vulnerable (1) and otherwise (0). The households which belong to Clusters 1 2, 3, and 6 are the ones which are financially more vulnerable because of their relatively high levels of indebtedness, low income levels, and less savings. This implies that for the year 1998, 30.3% of the households were financially vulnerable. The share of financially vulnerable households increased by about 2 percentage points in 2007, but it came down by about 4 percentage points in 2016. As of 2016, 28% of the households in the U.S. can be classified as financially vulnerable (Figure 4.3).

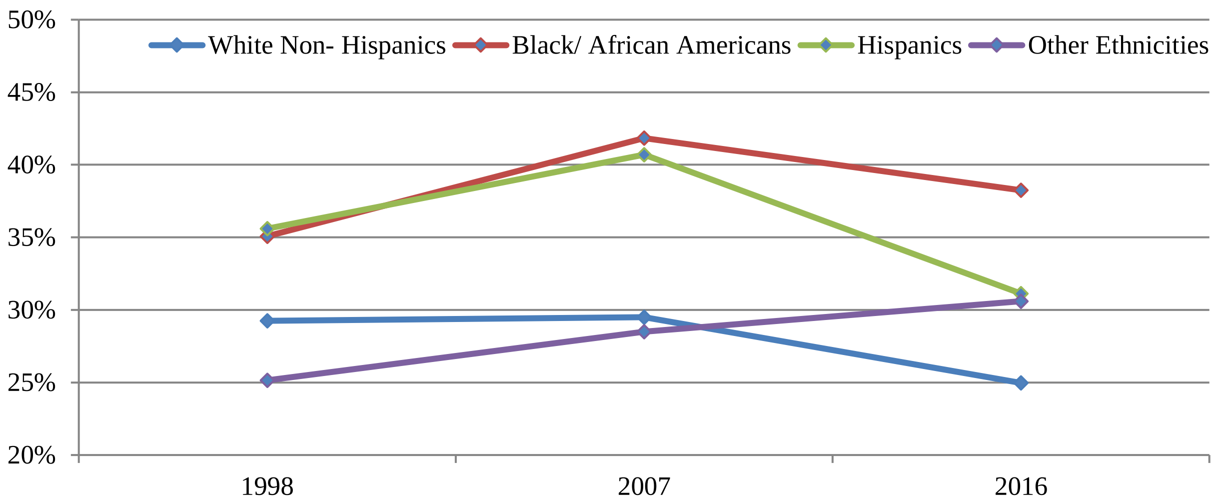
Figure 4.3: Percentage of Financially Vulnerable Households



Source: Authors' calculations based on Survey of Consumer Finance

Race and ethnicity is associated with financial vulnerability in the United States. African-American households have the highest proportion of financially vulnerable households (38.2 percent), followed by Hispanics (31.1 percent), other ethnicities which include Asians, native Hawaiian, Alaska native (30.6 percent), and white non-Hispanics (25 percent). The level of financial vulnerability decreased for all ethnicities from 2007 to 2016 except for households that belong to the other category (Figure 4.4).

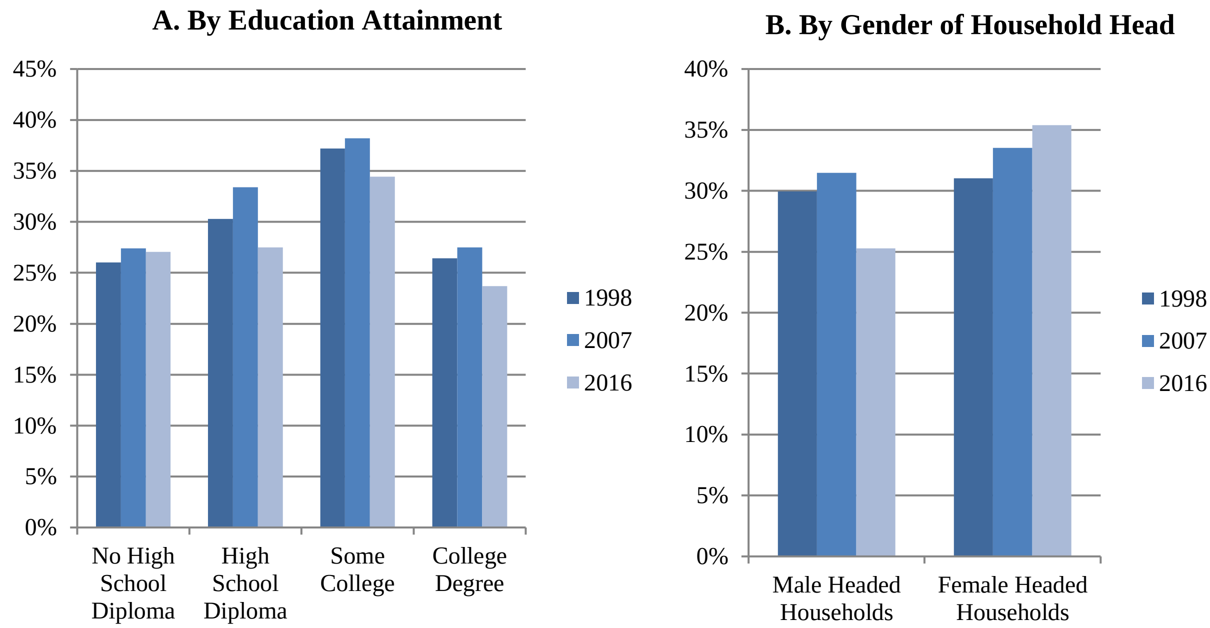
Figure 4.4: Financial Vulnerability by Race and Ethnicity



Source: Authors' calculations based on Survey of Consumer Finance

Household's financial vulnerability also varies across the education attainment level of the household head and the gender of the household head. Households where the household head has a college degree are less financially vulnerable as compared to households where the education level of the household head is lower (Figure 4.5). Moreover, female headed households are more financially vulnerable as compared to male headed households. However, it is important to note that household head is always considered to be male in case of a male/ female couple. This is the way household head is characterized in the survey of consumer finance due to historical reasons.

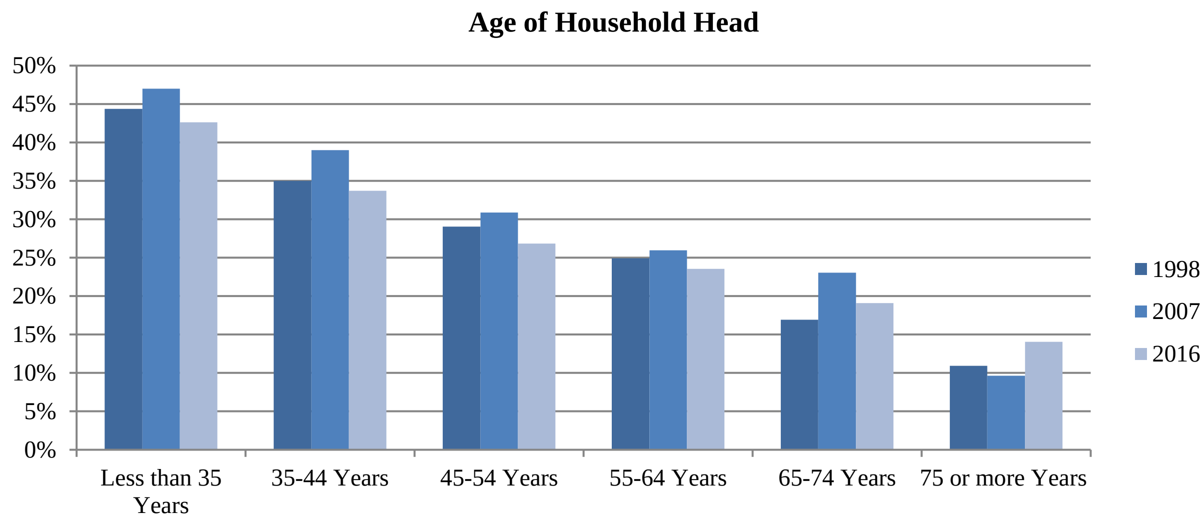
Figure 4.5: Household financial vulnerability by education attainment and gender



Source: Authors' calculations based on Survey of Consumer Finance

Household's financial vulnerability also varies across different age groups of the household head as highlighted in Figure 4.6. The financial vulnerability of households decreases with the increase in the age of the household head. Household heads within the age bracket of less than 35 years are the most financially vulnerable ones (42.6 percent) followed by household heads within the age bracket 35-44 years (33.7 percent) and household heads within the age bracket 45-54 years (26.9 percent).

Figure 4.6: Financial vulnerability is lower for older household heads



Source: Authors' calculations based on Survey of Consumer Finance.

A comparison of households which are vulnerable vis-à-vis households that are not financially vulnerable highlights that vulnerable households earn higher income, have less debt burden, are older in age, save more as compared to other households (Table 4.4). About 44% of the financially vulnerable households have a student debt as compared to 14% of households that are not financially vulnerable. With regard to the working status of the household head, household financial vulnerability is the lowest amongst self-employed household heads (20 percent), followed by people who work for someone (31.4 percent). A detailed comparison of financially vulnerable households with other households is provided in Table 4.4. Another important thing to highlight is that household financial vulnerability is underreported when conventional measures of financial vulnerability are used. For example, using the methodologies proposed by Michelangeli and Rampazzi (2016) and Bank of Italy (2012), about 25% of households in the U.S. come out to be financially vulnerable as of 2016[55].

[55] According to them, a household is considered vulnerable if its debt service-to-income ratio exceeds 30% and its income is below the median of the population.

Table 4.4: Profile of financially vulnerable households with respect to other households

Financially vulnerable households vis-à-vis households that are not financially vulnerable				
	High Financial Vulnerability		**Low or No Financial Vulnerability**	
	28%		72%	
	Median	Standard Deviation	Median	Standard Deviation
Age	44	16.2	54	17.3
Total Household Income (US $)	40,505	33,552.4	62,783	531,919.0
Value of Total Household Debt (US $)	78,900	209,137.7	9,000	239,374.4
Ratio of Total Debt to Total Assets	0.80	809.7	0.06	0.24
Log (Monthly Loan Payment)	6.8	2.1	5.5	3.2
Number of Children	0	1.2	0	1.1
Ratio of Monthly Debt Payments to Monthly Income	0.29	2.3	0.05	0.1
	Percentage		Percentage	
Self Employed/ Partnership	25.1%		74.9%	
Employee	31.4%		68.6%	
Retired/ Disabled/ Homemaker/ Student	20%		80%	
% of Household with a Debt	100%		68.2%	
% of Households with Savings (i.e. spending less than Income)	44.3%		59.7%	
% of Households with Student Debt	43.9%		13.80%	

Source: Survey of Consumer Finance

4.4.2. Econometric Model: Estimating key drivers of financial vulnerability

The section develops a household financial vulnerability model to analyze the main drivers of financial vulnerability. The binary dependent variable is $Y_{h,r,t}$ and it represents the status of household's financial vulnerability i.e. if the household is financially vulnerable (1) and otherwise (0). We use a logistic regression to conduct this analysis. $\mathbf{Y}^*_{h,r,t}$ is the latent variable where *h* and *r* and *t* are indexed for household, region, and year, respectively.

$$\mathbf{Y}^*_{h,r,t} = \boldsymbol{\alpha} + \boldsymbol{B_1}(\boldsymbol{HH\ Characteristics})_{\mathbf{h,r,t}} + \boldsymbol{B_2}(\mathbf{HCharacteristics})_{\mathbf{h,r,t}} + \boldsymbol{B_3}(\boldsymbol{Location\ type})_{\mathbf{h,r,t}} + \mu_{\mathbf{r}} + \mu_{\mathbf{t}} + \boldsymbol{\varepsilon}_{\mathbf{h,r,t}}) \quad (4.1)$$

$$Y_{h,r,t} = 1 \text{ if } \mathbf{Y}^*_{h,r,t} > 0$$
$$Y_{h,r,t} = 0 \text{ if } \mathbf{Y}^*_{h,r,t} \leq 0$$

HH Characteristics refer to the variables pertaining to the household head's level of education, gender, marital status, occupation status and age. *HCharacteristics* refer to controls pertaining to size of the household, value of debt payments, ethnicity, family structure of the household, and net-worth of the household. Location type refers to whether the household lives in a rural area or an urban area. Fixed effects at the regional level $\mu_{\mathbf{r}}$ are also included to control for regional level heterogeneity which might affect chances of financial vulnerability. These regional fixed effects control whether the household lives in northeast, or north central, or south, or west part of the country[56]. Furthermore, the term $\mu_{\mathbf{t}}$ denotes time fixed effects since we have a pseudo panel for years 1998, 2007 and 2016. The error term is represented by $\varepsilon_{\mathrm{h,r,t}}$ and follows a logistic distribution. Lastly, we cluster the standard errors at the regional level for all the regressions in order to control for possible error correlation within regions.

Survey of consumer finance uses a multiple imputation methodology to approximate the distribution of missing data. These imputations are stored as five successive replicates i.e. five implicates for each observation. Moreover, the Survey of Consumer Finance does not have an equal probability sampling design. So, we use the nonresponse-adjusted sampling weights in order to compensate for unequal probabilities of selection. Furthermore, in order to get results with accurate standard errors in SCF, it is important to take into account the imputation error as well as the sample variability error. We take both of these aspects into account in our empirical results. We estimate sampling variability in our regression models by using a set of bootstrap replicate weights by following the methodology proposed by Pence (2015)[57]. Results of the econometric estimates along with a detailed discussion on these results are provided in the next section.

[56] Information on the regional location of the household as well as the urban/ rural classification is not available in the publicly available data. The authors are thankful to Kevin Moore from the Board of Governors of the Federal Reserve System for helping us run the econometric estimates using regional controls.

[57] The "scfcombo" Stata module is used that "calculates and combines the imputation uncertainty and bootstrapped standard errors for estimation commands run on the Survey of Consumer Finances (SCF)".

4.5. Results and Discussion

This section presents the results of our econometric analysis that aims at determining the main determinants of household's financial vulnerability in United States. Table 4.5 summarizes the results of the econometrics estimations. Summary statistics for all the control variables are provided in Appendix 4.1.

In the first column of Table 4.5, we estimate the econometric model with quite basic specifications. We only include household level controls and time fixed effects, but no regional characteristics. The results show that various household characteristics such as ethnicity and race, education level, work status and age, are associated with household's financial vulnerability. However, regional characteristics can also play a key role in influencing household's financial vulnerability. Therefore, in column (2), we control for the difference in urban and rural locations by including an urban dummy. We also include regional fixed effects to control for regional level unobserved heterogeneity. In column (3), we also control for the net worth of the household as well as the value of monthly debt repayment. The results still come out to be significant for various household characteristics. However, the regional variables do not seem to be significant once we control for household net worth in the model. Lastly, in order to estimate the magnitude of coefficients, we run a linear probability model and report the results in Table 4.6. The "scfcombo" module by Pence (2015) that is used to correct for the imputation error and the sample variability error does not support the estimation of marginal effects after maximum likelihood estimations. Therefore, in order to get the correct standards errors along with an idea of the magnitude of coefficients, we run a linear probability model. The results highlight the importance of socioeconomic and demographic characteristics such as ethnicity, age, marital status, education attainment, working status and household's net-worth in explaining household's financial vulnerability.

The econometric estimates highlight that ethnicity is a significant determinant of household's financial vulnerability. African-Americans and Hispanics are financially more vulnerable as compared to non-Hispanic white Americans. On average, African-Americans are 8 percentage points more likely to be financially vulnerable whereas Hispanics are 6 percentage points more likely to be financially vulnerable, as compared to white non-Hispanics, after controlling for other household and regional characteristics. This finding is line with the numbers which suggest that

African-Americans and Hispanics have a lower rate of homeownership as compared to non-Hispanic white Americans, and the lack of capital can force them to incur more credit card debt.

Education level of the household head also comes out to be statistically significant. An increase in the education level is negatively related with the probability of financial vulnerability. On average, having a college degree decreases the probability of being financially vulnerable by 11 percentage points as compared to having no high school diploma. This finding seems to suggest that individuals with higher education have a better understanding of financial planning and the importance of savings. Moreover, the age of the household head is also significant, and it is negatively associated with financial vulnerability. An increase in the age of the household head decreases the chances of being financially vulnerable.

The results also provide strong evidence that an increase in the size of monthly debt repayment raises a household's financial vulnerability, even after controlling for traditional household and regional characteristics. Furthermore, living in an urban locality does not seem to have an impact on households' financial vulnerability once we take into account the net worth of the household. The net worth of a household comes out to be statistically significant as well and it is negatively associated with households' financial vulnerability. An increase of 1% in the household's net-worth decreases the probability of being financially vulnerable by about 0.47 percentage points, on average. Hence, an increase in financial and real assets reduces financial vulnerability.

The working status of the household head also comes out to be statistically significant. Household heads who are self-employed are more likely to be financially vulnerable as compared to household heads who are working as employees. Some reasons behind this might be the lack of diversification in assets and income, especially in the case of small businesses and start-ups. According to the U.S. Bureau of Labor Statistics, as of 2017, about 30% of small businesses fail before reaching the end of their second year, potentially leaving these households financially vulnerable. Moreover, household heads who are unemployed, retired, studying, homemaker or disabled are also more likely to be financially vulnerable as compared to household heads who are working as employees. Amongst other socioeconomic determinants, numbers of children in the family do not play any role in explaining households' financial vulnerability. However, marital status of the household head plays a significant role in impacting the financial vulnerability level. Being not married or not living with a partner raises the level of financial vulnerability.

Table 4.5: Results for the determinants of household financial vulnerability

Variables	**(1) Financial Vulnerability**	**(2) Financial Vulnerability**	**(3) Financial Vulnerability**
Race/ Ethnicity (Base: White Non- Hispanics)			
African-Americans	0.29***	0.56***	0.45***
	(0.06)	(0.06)	(0.06)
Hispanic	0.10	0.36***	0.27***
	(0.07)	(0.08)	(0.08)
Others	-0.13	-0.06	-0.04
	(0.10)	(0.11)	(0.11)
Age of HH (Base: Less than 35)			
35-44	-0.35***	-0.83***	-0.72***
	(0.05)	(0.06)	(0.06)
45-54	-0.64***	-1.25***	-1.01***
	(0.05)	(0.05)	(0.06)
55-64	-0.81***	-1.43***	-1.05***
	(0.06)	(0.07)	(0.08)
65-74	-1.07***	-1.56***	-1.19***
	(0.07)	(0.08)	(0.09)
75 or more	-1.70***	-1.70***	-1.28***
	(0.10)	(0.10)	(0.12)
Education Level of HH (Base: No High School Diploma)			
High School Diploma	0.11*	-0.31***	-0.26***
	(0.06)	(0.06)	(0.07)
Some College	0.29***	-0.37***	-0.24***
	(0.06)	(0.07)	(0.02)
College Degree	-0.13**	-1.05***	-0.68***
	(0.06)	(0.07)	(0.02)
Occupation Status of HH (Base: Employee)			
Self Employed/ Partnership	-0.05	-0.06	0.39***
	(0.06)	(0.07)	(0.08)
Out of labor force	-0.04	0.63***	0.70***
	(0.06)	(0.07)	(0.07)
Not Working	0.10	0.97***	1.06***
	(0.07)	(0.08)	(0.09)
Urban Dummy		-0.12**	-0.03
		(0.05)	(0.05)
# of children in the household	0.06***	0.005	0.02
	(0.02)	(0.02)	(0.02)
Marital Status of HH (Base: Married/ living with partner)			
Neither married nor living with partner	0.78***	1.01***	0.89***
	(0.04)	(0.05)	(0.05)
Log (Monthly Debt Payment)			0.56***
			(0.01)
Log (Total Household Net Worth)			-28.8***
			(2.94)
Regional Fixed Effects	**No**	**Yes**	**Yes**
Time Fixed Effects	**Yes**	**Yes**	**Yes**
Total Observations	**74,850**	**74,850**	**74,850**
Pseudo R2	**0.0543**	**0.2235**	**0.2591**

*Note: Results of the logit regression have been reported. Standard errors are in parentheses (clustered at the regional level). The dependent variable is households' financial vulnerability dummy (see equation 4.1). ***, ** and * represent significance at 1%, 5% and 10%, respectively*

Table 4.6: Results for the determinants of household financial vulnerability with LPM

Variables	**(4)** **Financial Vulnerability**
Race/ Ethnicity (Base: White Non- Hispanics)	
African-Americans	0.08 ***
	(0.01)
Hispanic	0.06***
	(0.01)
Others	-0.01
	(0.02)
Age of HH (Base: Less than 35)	
35-44	-0.13***
	(0.01)
45-54	-0.20***
	(0.01)
55-64	-0.21***
	(0.01)
65-74	-0.21***
	(0.01)
75 or more	-0.22***
	(0.02)
Education Level of HH (Base: No High School Diploma)	
High School Diploma	-0.03**
	(0.07)
Some College	-0.02**
	(0.02)
College Degree	-0.11***
	(0.02)
Occupation Status of HH (Base: Employee)	
Self Employed/ Partnership	0.02*
	(0.01)
Out of labor force	0.08***
	(0.01)
Not Working	0.13***
	(0.01)
Urban Dummy	-0.02*
	(0.01)
# of children in the household	-0.0004
	(0.003)
Marital Status of HH (Base: Married/ living with partner)	
Neither married nor living with partner	0.13***
	(0.01)
Log (Monthly Debt Payment)	0.07***
	(0.001)
Log (Total Household Net Worth)	-0.47***
	(0.05)
Regional Fixed Effects	**Yes**
Time Fixed Effects	**Yes**
Total Observations	**74,850**
R-squared	**0.22**

*Note: Results of the linear probability model have been reported. Standard errors are in parentheses (clustered at the regional level). The dependent variable is households' financial vulnerability dummy (see equation 4.1). ***, ** and * represent significance at 1%, 5% and 10%, respectively.*

4.6. Conclusion

This chapter proposed a new methodological approach of assessing household financial vulnerability by employing unsupervised machine learning. Hierarchical ascending clustering (HAC) and K-means clustering was used to identify homogenous clusters of households within the nationally representative sample of households in the United States. We analyzed the characteristics of these clusters and identified households with high financial vulnerability. The results show that about 28% of the households in 2016 can be classified as financially vulnerable. The percentage of financially vulnerable households comes out to be higher in 1998 and 2007. Moreover, these numbers suggest that household financial vulnerability is underreported when conventional measures of financial vulnerability are being used.

To better understand the main determinants of household financial vulnerability in the United States, this chapter then analyzed the probability of being financially vulnerable based on household and regional characteristics. At the household level, characteristics such as ethnicity, age, marital status, educational attainment, number of children, working status, value of monthly debt repayments and net-worth are considered, while regional level controls included whether the household resides in an urban location or a rural location along with regional fixed effects.

In line with the theoretical predictions, the econometric estimates confirm the importance of socioeconomic and demographic characteristics such as ethnicity, age, marital status and working status on household's financial vulnerability. Results highlight that African-Americans and Hispanics are financially more vulnerable than non-Hispanic white Americans. In the case of educational attainment, the results show a strong and negative association with household's financial vulnerability. Regarding the age of the household head, results indicate that there is a strong and negative association with financial vulnerability as well. In the case of working status of the household head, the results highlight that self-employed, retired and people who are not working are likely to be more financially vulnerable as compared to people who work as employees. These findings highlight that a buoyant labor market and other policy levers such as increasing educational attainment may help mitigate household financial risks and vulnerabilities.

Lastly, there is evidence suggesting that net-worth of the household and the size of monthly debt repayments affect household's financial vulnerability. Regarding net worth of the household, results show a significant and negative relationship. An increase in the net-worth of the households decreases the probability of being financially vulnerable. Furthermore, an increase in the value of monthly debt repayment raises the probability of being financially vulnerable. Whereas, living in an urban locality does not seem to have a significant impact on financial vulnerability after controlling for the net-worth of the household into account.

Acknowledgements

This chapter has benefitted from comments by Sebastian Schich from the OECD Directorate of Financial and Enterprise Affairs and Hervé Kieffel of Stanford University. The co-operation of Kevin Moore from the Board of Governors of the Federal Reserve System is gratefully acknowledged. This chapter was prepared as a technical background paper for the *OECD Economic Survey of United States 2018*.

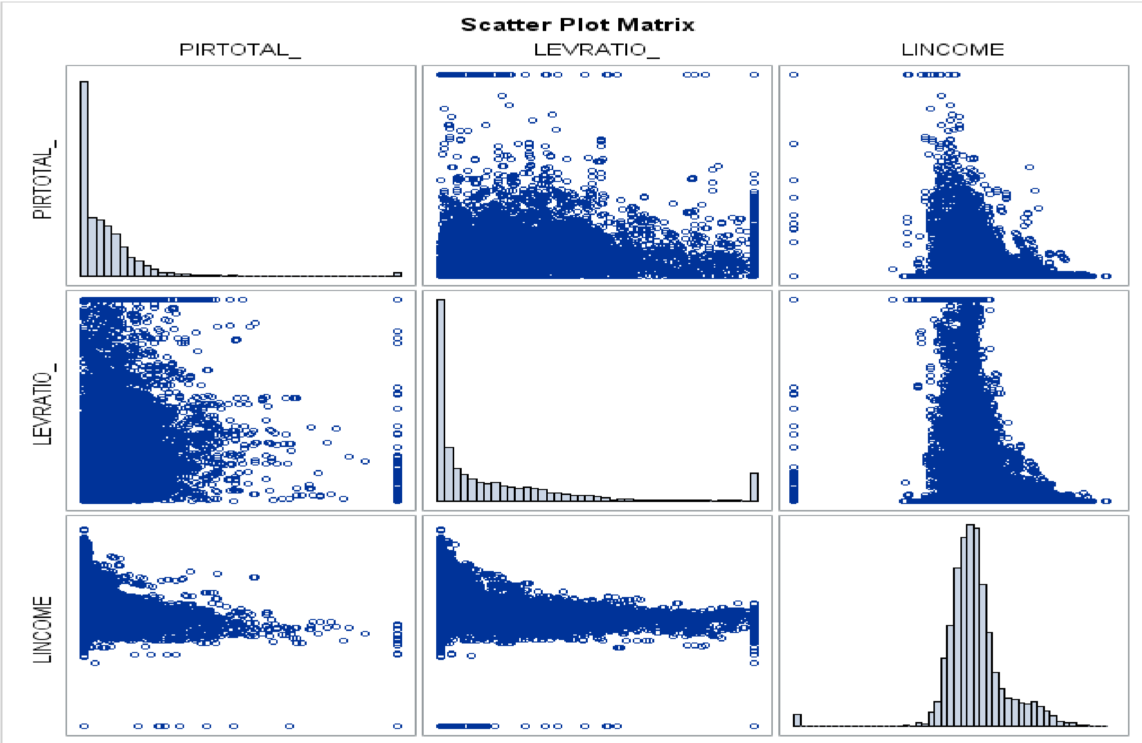
4.7. Appendix

Appendix 4.1: Descriptive Statistics

Variable	No. of Observations	Mean	Std. Dev.	Min	Max
Financial Vulnerability_ Dummy	74,850	0.300507	0.458482	0	1
Monthly Debt Payment (Log)	74,850	4.801251	3.158039	0	13.22045
Total Household Net Worth (Log)	74,850	16.94653	0.074511	10.26876	21.26919
Gender_ Male Dummy	74,850	0.723286	0.447377	0	1
White Non- Hispanics_ Dummy	74,850	0.729047	0.444455	0	1
African-Americans_ Dummy	74,850	0.135617	0.342383	0	1
Hispanics_ Dummy	74,850	0.094565	0.292615	0	1
Other Ethnicities _ Dummy	74,850	0.040772	0.197763	0	1
No High School Diploma_ Dummy	74,850	0.140907	0.347928	0	1
High School Diploma_ Dummy	74,850	0.300423	0.458445	0	1
Some College_ Dummy	74,850	0.256414	0.436656	0	1
College Degree_ Dummy	74,850	0.302256	0.459239	0	1
Employee_ Dummy	74,850	0.58359	0.492967	0	1
Self Employed/ Partnership_ Dummy	74,850	0.107667	0.309961	0	1
Retired/ Disabled/ Student/ Homemaker_ Dummy	74,850	0.259052	0.438117	0	1
Not Working_ Dummy	74,850	0.049692	0.21731	0	1
Number of children in the household	74,850	0.80332	1.134475	0	10
Married/ living with partner_ Dummy	74,850	0.579671	0.493615	0	1
Age Less than 35 Years_ Dummy	74,850	0.216244	0.411685	0	1
Age 35-44_ Dummy	74,850	0.196786	0.397572	0	1
Age 45-54_ Dummy	74,850	0.194136	0.395537	0	1
Age 55-64_ Dummy	74,850	0.165251	0.37141	0	1
Age 65-74_ Dummy	74,850	0.120523	0.325574	0	1
Age 75 or more_ Dummy	74,850	0.10706	0.309192	0	1

Source: Survey of Consumer Finance

Appendix 4.2: Correlation Matrix[58]



Source: Authors' calculations using Survey of Consumer Finance (SCF)

	Debt Burden (PIRTOTAL)	**Leverage Ratio (LEVRATIO)**	**Household Income (LINCOME)**
Debt Burden (PIRTOTAL)	1	0.22	-0.19
Leverage Ratio (LEVRATIO)	0.22	1	-0.13
Household Income (LINCOME)	-0.19	-0.13	1

Source: Survey of Consumer Finance (SCF)

[58] *PIRTOTAL* represents debt burden, LEVRATIO represents leverage ratio, *LINCOME* represents household income.

Appendix 4.3: Classification of Clusters and Distribution of Debt Burden, Leverage Ratio and Household Income

Cluster = 1

Variable	Cluster mean	Overall mean	Cluster size	Test value	Probability of test value	---------
LINCOME	0	10.84515	82	-1,33	0.0920	*|

Cluster = 2

Variable	Cluster mean	Overall mean	Cluster size	Test value	Probability of test value	---------
PIRTOTAL_	1.86443	0.14815	263	2,16	0.0154	|**
PIRTOTAL	8.76196	0.17838	263	1,76	0.0390	|**
LINCOME	5.14604	10.84515	263	-1,25	0.1049	|

Cluster = 3

Variable	Cluster mean	Overall mean	Cluster size	Test value	Probability of test value	---------
PIRTOTAL_	0.40112	0.14815	4496	1,42	0.0783	|*

Cluster = 4

Variable	Cluster mean	Overall mean	Cluster size	Test value	Probability of test value	---------
INCCAT	4.63798	3.10008	15199	2,64	0.0041	|***
LINCOME	11.78958	10.84515	15199	2,20	0.0141	|**

Cluster = 5

Variable	Cluster mean	Overall mean	Cluster size	Test value	Probability of test value	---------
LEVRATIO_	0.08132	0.40718	8596	-1,05	0.1460	|
HDEBT	0.47254	0.77116	8596	-1,22	0.1114	|
INCCAT	1.88800	3.10008	8596	-1,32	0.0939	*|

Cluster = 6

Variable	Cluster mean	Overall mean	Cluster size	Test value	Probability of test value	---------
LEVRATIO_	1.73730	0.40718	2604	2,11	0.0176	|**

Appendix 4.4: Basic Statistics about United States

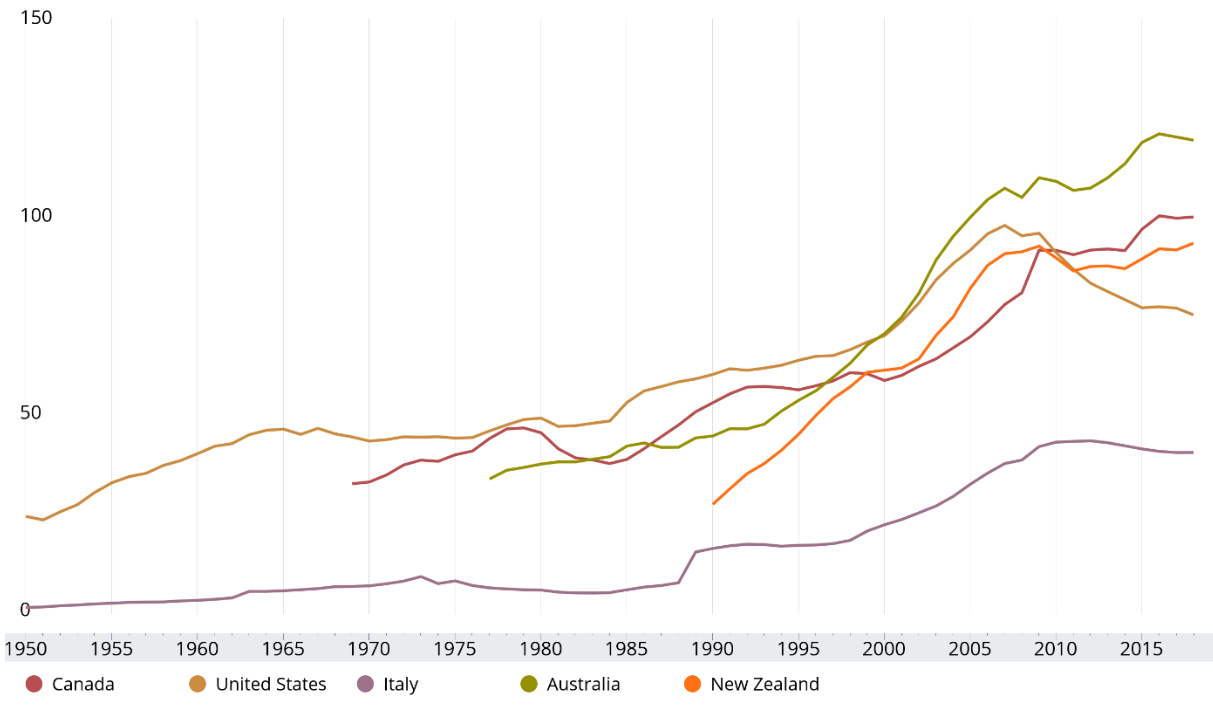
BASIC STATISTICS OF UNITED STATES

(Data refer to 2017 or latest available. Numbers in parentheses refer to the OECD average)*

LAND, PEOPLE AND ELECTORAL CYCLE					
Population (million)	323.13		Population density per km²	35.3	(37.5)
Under 15 (%)	19.0	(18.0)	Life expectancy (years)	78.8	(80.7)
Over 65 (%)	15.0	(16.5)	Men	76.3	(78.1)
Foreign-born (%)	13.5		Women	81.2	(83.3)
Latest 5-year average growth (%)	0.7	(0.6)	Latest general election	November 2016	
ECONOMY					
Gross domestic product (GDP)			Value added shares (%)		
In current prices (billion USD)	19,390.6		Primary sector	0.9	(2.4)
Latest 5-year average real growth (%)	2.2	(2.1)	Industry including construction	19.1	(26.7)
Per capita (000 USD PPP)	59.5	(43.8)	Services	80.0	(70.9)
GENERAL GOVERNMENT					
Per cent of GDP					
Expenditure	37.7	(40.6)	Gross financial debt	105.4	(110.2)
Revenue	34.1	(37.7)	Net financial debt	80.3	(71.2)
EXTERNAL ACCOUNTS					
Exchange rate EUR per USD	0.9		Main exports (% of total merchandise exports)		
PPP exchange rate (USA = 1)	1.0		Machinery and transport equipment	34.9	
In per cent of GDP			Chemicals and related products, n.e.s.	13.6	
Exports of goods and services	12.1	(29.1)	Commodities and transactions, n.e.s.	12.5	
Imports of goods and services	15.0	(28.8)	Main imports (% of total merchandise imports)		
Current account balance	-2.40	(0.34)	Machinery and transport equipment	43.2	
Net international investment position	-40.5		Miscellaneous manufactured articles	16.9	
			Manufactured goods	10.8	
LABOUR MARKET, SKILLS AND INNOVATION					
Employment rate for 15-64 year-olds (%)	69.4	(67.0)	Unemployment rate, Labour Force Survey (age 15 and over) (%)	4.9	(6.5)
Men	74.8	(74.8)	Youth (age 15-24, %)	10.4	(12.9)
Women	64.0	(59.4)	Long-term unemployed (1 year and over, %)	0.8	(2.0)
Participation rate for 15-64 year-olds (%)	73.0	(71.7)	Tertiary educational attainment 25-64 year-olds (%)	45.7	(35.7)
Average hours worked per year	1 783	(1763)	Gross domestic expenditure on R&D (% of GDP)	2.7	(2.3)
ENVIRONMENT					
Total primary energy supply per capita (toe)	6.7	(4.1)	CO_2 emissions from fuel combustion per capita (tonnes)	15.5	(9.2)
Renewables (%)	7.1	(9.7)	Water abstractions per capita (m³)	1582	(804)
Fine particulate matter concentration ($PM_{2.5}$, µg/m³)	9.2	(14.9)	Municipal waste per capita (kilogrammes)	738	(523)
SOCIETY					
Income inequality (Gini coefficient)	0.39	(0.32)	Education outcomes (PISA score, 2015)		
Relative poverty rate (%)	16.8	(11.6)	Reading	497	(493)
Median disposable household income (000 USD PPP)	32.1	(23.0)	Mathematics	470	(490)
Public and private spending (% of GDP)			Science	496	(493)
Health care, current expenditure	17.2	(9.1)	Share of women in parliament (%)	19.2	(28.7)
Pensions	7.0	(9.1)	Net official development assistance (% of GNI)	0.18	(0.39)
Education (primary, secondary, post sec. non tertiary)	3.6	(3.6)			

Source: OECD's Economic Survey of United States, 2017

Appendix 4.5: Household debt as a percentage of GDP



Source: IMF's Global Debt Database, 2019

General Conclusion

I. Summary and Policy Recommendations

This thesis is centred around questions of access to institutional credit and other financial instruments and the impact of this access on entrepreneurship. It adds to the existing research on financial inclusion that supports the view that financial inclusion can play an important role in the global development agenda. On one hand, the global economy is expected to be more than USD 85 trillion (in nominal terms); on the other hand, 31% of the world's adult population remains without access to very basic financial instruments and services- most basic one being a bank account[59]. It is in a time like this that the focus on financial inclusion and microfinance becomes substantially significant. This thesis underlines that improving financial inclusion and access to microfinance holds a real appeal in the context of developing economies where majority of the population remains financially underserved.

The existing evidence on the role of expanding access to financial services has been mixed and has not normally studied the impact of financial inclusion in its entirety. This thesis has attempted to fill that research gap. To further this motive, this thesis combines four studies to provide a comprehensive analysis on several contemporary issues pertaining to the link between financial inclusion, entrepreneurship and gender. Moreover, given the methodological challenges highlighted in the existing literature, this thesis also proposes new methodological approaches to analyze household financial vulnerability and involuntary financial exclusion. The main findings of this research give rise to many important public policy considerations which are discussed in this section.

This thesis uncovered the relationship between geographical access to microfinance and entrepreneurship. The main findings highlight that microfinance has a positive effect on entrepreneurship and it can allow people to move up the economic ladder by shifting to entrepreneurship rather than working as low paid employees, farm workers, and housewives. The results suggest that having geographical access to microfinance increases the probability of becoming an entrepreneur by about 4 percentage points. Having access to a formal financial

[59] These numbers are as of 2017 and come from the World Bank's World Development Indicators and Findex database.

institution such as a microfinance branch is likely to lower credit constraints and it can allow individuals to use other financial services such as payment services and bank accounts. However, the results show that the impact of access to microfinance on entrepreneurship is conditional on the poverty channel i.e. access to microfinance increases the chances of becoming an entrepreneur in poor regions. From a policy perspective, this finding suggests that improving the accessibility of microfinance in Pakistan can have a meaningful effect on fostering entrepreneurship and this relationship will work best if microfinance branches are targeted in poor regions. This goes on to show that microfinance holds meaningful appeal in the context of countries like Pakistan where a large segment of population remains financially excluded.

The impact of financial inclusion on women entrepreneurship is also studied in detail. A comprehensive indicator to measure the state of financial inclusion in Mexico is composed and we gauge its effect on women entrepreneurship. The results highlight that financial inclusion has a positive effect on women entrepreneurship and it can open up economic opportunities for them, including in the formal sector. The results also uncover the existence of a significant gender disparity in the status of entrepreneurship across formal and informal work in Mexico. The main findings show that the probability of a woman being an entrepreneur in the informal sector is higher as compared to her being an entrepreneur in the formal sector. Given the existence of relatively high gender gaps in Mexico, our results indicate that improving women's financial inclusion can have a meaningful impact on job creation for women. Our results also support the idea that with specific policies, targeting some economic sectors in rural areas, financial inclusion can also support formalization of women entrepreneurs in Mexico.

The findings also highlight that there are several supply side constraints that hinder access and usage of financial products and services for large segments of the society, especially women. Some policy responses which can help improve the status of financial inclusion and possibly narrow down the gender gap are discussed here. These policy responses can be broadly classified into three categories: i) technology driven responses ii) institutional driven responses, and iii) data driven responses.

First, with regard to technology driven responses, one possible way is to have less strict account opening and credit application requirements. In recent years, advances in digital technology have allowed FinTech to emerge as a potential solution to reduce operational costs of financial institutions and foster financial inclusion. In the case of bank accounts, even though there are certain regulations regarding know your customers (KYC) and anti-money laundering rules, it is possible to use creative ways to deliver access to no-frill bank accounts using mechanisms such as banking agents and tablet banking. People who work in the informal sector, especially women, have limited formal documentation, and it is worth considering expanding the range of identification documents needed to open up bank accounts. Similarly, innovative digital technology can help with requirements such as biometrics and finger printing to reach far off rural areas. Some of these innovative ways are already being piloted and launched in certain countries. For example, the UKaid Sakchyam Access to Finance program in Nepal and a start-up called Vaya in India are making use of tablet banking to reach out to unbanked people in rural areas. Similarly, regarding access to credit, big data and machine learning have emerged as a possible solution to reduce the operational costs of expanding credit. Some financial institutions across the world are already starting to rely on machine learning based credit assessment tools to gauge financial ability of potential clients. One of the many such examples is a microfinance provider in Pakistan named *Tez Financial Services* which relies on machine learning rather than conventional microcredit risk assessments tools to assess credit worthiness of clients who are mostly unbanked and don’t have any credit history. Therefore, creating a regulatory environment which encourages technological innovation without threatening financial stability can help boost financial inclusion. However, more in-depth research is required to better understand the potential risks and benefits associated with these technological innovations.

Second, regarding institutional driven responses, the role of central banks and credit bureaus can be quite important. For example, credit bureaus in Ghana and Uganda lowered the minimum loan amount that can be offered by financial institutions. This proved to be an important move in enhancing financial inclusion for women because women generally make use of smaller loans as compared to men. Similarly, promoting dedicated financial institutions that mainstream gender consideration in the provision of financial services can also help foster financial inclusion. Organizations such as Banco Santender Santiago in Chile and Kashf Foundation in Pakistan have embedded gender consideration in their loan extension strategy by easing collateral requirements, hiring female credit officers, and designing products that fit the needs of women.

Third, regarding data driven responses, one possible solution is better management of data. Unavailability of granular data, disaggregated by gender, is a major blockade when it comes to developing products and services that are tailored to the needs of women. Central banks around the world are increasingly recognizing this issue and are engaging with financial institutions to collect data that is disaggregated by gender. For example, central bank in Zambia was able to convince financial institutions to collate gender disaggregated data in order to target their financial products and services in a better way. All these steps can go a long way in improving the state of financial inclusion in developing countries.

This thesis also highlights that there are several methodological challenges when it comes to analyzing the factors that drive financial exclusion. The results provide a better understanding of financial exclusion in Pakistan from the perspective of credit by taking into account the demand for credit as well as voluntary financial exclusion. The findings highlight that the demand for credit is perhaps overrated. This is in line with the actual experiencers of credit officers and bankers who frequently report issues pertaining to finding microcredit clients in developing countries. In case of Pakistan, 38.1% of the adult population shows no need for credit, 1.2% are financially included, 24.5% exhibit voluntary financial exclusion whereas the remaining 36.2% exhibit involuntary financial exclusion. According to the empirical results, financially illiteracy, poverty, and gender are amongst the main determinants of involuntary financial exclusion. From a policy perspective, the findings clearly indicate that improving the level of financial literacy makes a significant difference in reducing involuntary financial exclusion. Pakistan ranks 108 out of 144 countries in terms of financial literacy rate and a lot needs to be done to improve financial literacy in less developed regions and in rural areas. The results also indicate that women are more credit constrained than men and policy measures highlighted earlier need to be considered in order to overcome this gap and foster financial inclusion.

Another important aspect that this thesis highlights is the issue pertaining to the measurement of financial inclusion. In order to have good policies in place, one should be able to measure different aspects of financial inclusion properly. However, due to over-reporting in the supply side data and under-reporting in the demand side data, mainly with respect to credit and bank accounts, accurate measurement of financial inclusion becomes quite a big challenge. This highlights the need to have cogent definitions in place along with better data quality management to ensure well-designed and better targeted policies.

The last chapter of the thesis is a methodological one that deals with the adverse consequences of credit expansion and discusses the role of household financial vulnerability. Given the shortcomings of the existing empirical approaches that try to gauge the extent of household financial vulnerability, a new methodological approach is proposed by relying on unsupervised machine learning. The results indicate that the extent of household financial vulnerability is underrepresented when conventional methodologies are used. Our results show that about 28% of the households in the United States can be classified as financially vulnerable as of 2016, which is 4% less as compared to 2007. The findings also show that African-Americans are 8 percentage points and Hispanic Americans are 6 percentage points more likely to be financially vulnerable than non-Hispanic white persons, after taking into account other household and regional level characteristics. From a policy perspective, a buoyant labor market can help households reduce financial vulnerability through increasing employment. However, financial vulnerability is multifaceted and other policy levers, such as increasing educational attainment, may also mitigate household risks.

Overall, this thesis highlights that the labor market is one of the channels through which financial inclusion can help generate new economic opportunities and improve women's autonomy. These results support the commitment made by the G20s to expand financial inclusion as a key aspect of the 2030 Agenda for Sustainable Development. The results also highlight that increasing financial inclusion would require working on gender gaps, improving the levels of financial literacy, and reaching out to marginalized poor communities. However, it is crucial to keep the tradeoff between financial inclusion and stability in mind and support sustainable financial inclusion without threating the financial stability of the economy.

II. Limitations and Prospects for Future Research

This thesis has attempted to provide a better understanding of the relationships between financial inclusion, entrepreneurship and gender by relying on different parametrical and non-parametrical techniques to offer robust results that have a wide range of policy implications for policymakers, financial institutions, and governments. However, this research is not without its limitation, and they have been discussed in this section along with some of the prospects for future research.

With regard to entrepreneurship status, there are a large number of people who choose to be entrepreneurs not because they want to take advantage of a business opportunity and become one, but because they do not have any other employment opportunities. This form of entrepreneurship can be termed as “necessity” driven, as opposed to the other form that is “opportunity” driven. Therefore, it might be important to differentiate between the two different types especially from a policy perspective. However, in the case of this thesis, the data does not allow us to make the distinction between these two types of entrepreneurships. It would be interesting to make this distinction and study how microfinance and financial inclusion have an impact on both these types of entrepreneurship. It would also be interesting to study how these two types of entrepreneurship interact with gender, economic sectors, locality (urban and rural areas), and across formal and informal work. Moreover, it would have also been interesting to explore how these entrepreneurs and new businesses perform over time after their creation. However, all the chapters in this thesis rely on pseudo panel datasets and the same individuals are not followed over time that restricts this type of analysis.

In case of Chapter 1, the variable of geographical access to microfinance institutions does not allow us to differentiate between the different types of microfinance institutions in the area. Microfinance institutions can be broadly classified into two different categories i.e. for-profit institutions (microfinance banks) and non-profit institutions (mainly NGOs). They differ from each other in terms of their client outreach, product portfolio, and financial performance. Due to data constraints, it is not possible for us to differentiate between these two types in the survey. Moreover, due to the categorical nature of the variable, we had to rely on a 10km radius as the distance threshold for our baseline model. However, we use different distance thresholds to run a set of robustness checks and we find that the main results remain stable throughout all the specifications. Nonetheless, for future research, it would be interesting to see how the impact varies across different types of microfinance institutions.

In case of Chapter 2, although we added as many different dimensions as possible to get a comprehensive picture of the overall level of financial inclusion at the municipality level, but we acknowledge that some factors might influence financial inclusion more than the others. Nevertheless, the robustness of our financial inclusion index has been checked using both access-based and usage-based indicators in our econometric estimates. Still, as more data–both demand and supply side– becomes available, some other dimensions and covariates can be incorporated in the understanding of financial inclusion and household financial vulnerability for a more comprehensive subsequent analysis.

Furthermore, despite our best efforts to deal with the issues concerning endogeneity, it is still possible that some of our results may be biased because of the use of weak instruments. It is challenging to find good instrument variables, and this is particularly a big challenge in the case of Pakistan where data constraints are very pertinent. For example, the last census in Pakistan took place more than two decades ago in 1998[60]. Scarcity of data on some very basic socio-economic indicators at the geographical level makes it very challenging to find strong instruments. Moreover, due to the unavailability of true panel micro datasets, addressing some of these concerns becomes more difficult. Despite all these challenges, our econometric estimates rely on a number of techniques to address these issues. We conduct a wide range of statistical tests and they show that our core findings are robust.

In case of Chapter 3, due to certain factors not captured in the survey and because of measurement issues, we are basically analyzing associations between involuntary financial exclusion and covariates rather than causality. As more data from the demand side as well as the supply side become available, many other factors and methodological issues should be taken into account for a better understanding of involuntary financial exclusion. This chapter also highlights the need to collect more data and even conduct qualitative surveys to have a better understanding of the main drivers of financial exclusion.

The empirical evidence in the case of Chapter 4 is based on the survey data from the United States. It was not possible to conduct a similar study for Pakistan or Mexico mainly because, to the best of our knowledge, there are no nationally representative datasets available that provide information

[60] An attempt to conduct the census took place in 2017, but due to politicization, the findings have not been made public.

on the amount of borrowing as well as the net-worth of the household assets for these countries. However, household financial vulnerability is a topic that affects developing as well as developed countries and our proposed unsupervised machine learning approach can be applied to other countries as well.

Finally, considering the recent developments in the usage of FinTech to reach out to the unbanked population in developing countries, future research can focus on the role of technology and how it affects financial inclusion and the wellbeing of businesses and households. Although there are many potential benefits associated with the usage of financial technology to extend outreach of financial services, but some key challenges also need to be addressed which could cause issues pertaining to consumer protection, data privacy, and other ethical concerns. Due to the relative newness of emerging phenomenon such as tablet banking and credit risk assessments based on machine learning, it might take some time to develop an unequivocal understanding of the impact that they have on the financial and social performance of financial institutions in developing economies.

References

Acharya, V., Philippon, T., Richardson, M., and Roubini, N. (2009). "The Financial Crisis of 2007-2009: Causes and Remedies". *Financial Markets, Institutions & Instruments,* 18 (2), 89–137.

Aghion, P., and Bolton, P. (1997). "A theory of trickle-down growth and development". *The Review of Economic Studies*, 64(2), 151-172.

Aghion, P., Bacchetta, P., Ranciere, R., and Rogoff, K. (2009). "Exchange rate volatility and productivity growth: The role of financial development". *Journal of Monetary Economics*, 56(4), 494-513.

Ai, C. and Norton, E. (2003). "Interaction Terms in Logit and Probit Model". *Economic Letters* 80, 123- 129.

Aker, J. C., Boumnijel, R., McClelland, A., and Tierney, N. (2016). "Payment mechanisms and antipoverty programs: Evidence from a mobile money cash transfer experiment in Niger". *Economic Development and Cultural Change*, 65(1), 1-37.

Albacete, N., and Fessler, P. (2010). "Stress testing Austrian Households". *Financial Stability Report*, 19, 72-91.

Albacete, N., and Lindner, P. (2013). "Household Vulnerability in Austria–A Microeconomic Analysis based on the Household Finance and consumption survey". *Financial Stability Report*, 25, 57-73.

Aleem, I. (1990). "Imperfect information, screening, and the costs of informal lending: A study of a rural credit market in Pakistan". *The World Bank Economic Review,* 4(3), 329-349.

Allen, F., Demirguc-Kunt, A., Klapper, L. and Martinez-Peria, M.S. (2016), "The Foundations of Financial Inclusion: Understanding Ownership and Use of Formal Accounts". *Journal of Financial Intermediation*, 27, 1–30.

Amendola, A., Boccia, M., Mele, G., and Sensini, L. (2016). "Financial access and household welfare: evidence from Mauritania". World Bank Policy Research. Working Paper No. 7533.

Amidžić, G., Massara, A., and Mialou, A. (2014). "Assessing Countries' Financial Inclusion Standing—A New Composite Index". IMF Working Paper, WP/14/36.

Anand, K. and Kuldip, C. (2013). "A Theoretical and Quantitative Analysis of Financial Inclusion and Economic Growth". *Management and Labor Studies 38*.1-2: 103-133.

Anderloni, L., Bacchiocchi, E., and Vandone, D. (2012). "Household financial vulnerability: An empirical analysis". *Research in Economics*, 66(3), 284-296.

References

Anderson, S., and Baland, J. M. (2002). "The economics of roscas and intrahousehold resource allocation". *The Quarterly Journal of Economics*, 117(3), 963-995.

Angelucci, M., Karlan, D., and Zinman, J. (2015). "Microcredit Impacts: Evidence from a Randomized Microcredit Program Placement Experiment by Compartamos Banco". *American Economic Journal: Applied Economics*, 7(1): 151-82.

Arcand, J. L., Berkes, E., and Panizza, U. (2015). “Too much finance?”. *Journal of Economic Growth*, 20(2), 105-148.

Ashraf, N., Karlan, D., and Yin, W. (2010). “Female empowerment: Impact of a commitment savings product in the Philippines”. *World Development*, 38(3), 333-344.

Aslam, A. and Azmat, N. (2012). “A Study of Collateral Options for Microfinance Loans in Pakistan”. Pakistan Microfinance Network.

Attanasio, O., Augsburg, B., De Haas, R., Fitzsimons, E., and Harmgart, H. (2015). “The impacts of microfinance: Evidence from joint-liability lending in Mongolia”. *American Economic Journal: Applied Economics*, 7(1), 90-122.

Augsburg, B., De Haas, R., Harmgart, H., and Meghir, C. (2015). “The impacts of microcredit: Evidence from Bosnia and Herzegovina”. *American Economic Journal: Applied Economics*, 7(1), 183-203.

Ayyagari, M., Demirguc-Kunt, A., and Makismovic, V. (2012). “Financing of Firms in Developing Countries: Lessons from Research”. World Bank Policy Research. Working Paper No. o36, Washington, D.C.: World Bank.

Baland, J. M., Somanathan, R., and Vandewalle, L. (2019). “Socially disadvantaged groups and microfinance in India”. *Economic Development and Cultural Change*, 67(3), 537-569.

Baland, J. M., Guirkinger, C., and Mali, C. (2011). "Pretending to be poor: Borrowing to escape forced solidarity in Cameroon". *Economic Development and Cultural Change*, 60(1), 1-16.

Banerjee, A. (2013). “Microcredit under the microscope: What have we learned in the past two decades, and what do we need to know?”. *Annual Review of Economics,* 5, 487–519.

Banerjee, A. and Newman, A. (1993). “Occupational Choice and the Process of Development”. *Journal of Political Economy*, 101, (2), 274-98.

Banerjee, A., Duflo, E., Glennerster, R., and Kinnan, C. (2015). “The miracle of microfinance? Evidence from a randomized evaluation”. *American Economic Journal: Applied Economics*, 7(1), 22-53.

References

Banerjee, A., Karlan, D., and Zinman, J. (2015). "Six randomized evaluations of microcredit: Introduction and further steps". *American Economic Journal: Applied Economics*, 7(1), 1-21.

Bank of England (2016). "Financial Stability Report 39". July.

Bank of Italy (2012). "Financial Stability Report 04". April.

Bateman, M. (2010). "Why doesn't microfinance work?: The destructive rise of local neoliberalism". Zed Books Ltd.

Bateman, M., and Chang, H. J. (2009). "The microfinance illusion". Available at SSRN 2385174.

Bauchet, J., Marshall, C., Thomas, J., and Yalouris, A. (2011). "Latest Findings from Randomized Evaluations of Microfinance". Financial Access Initiative, Innovations for Poverty Action, and Abdul Latif Jameel Poverty Action Lab. Forum 2. Washington, D.C.: CGAP

Bayoumi, M. T., and Melander, O. (2008). "Credit Matters: Empirical Evidence on US Macro-Financial Linkages". IMF Working Paper No. 08/169. International Monetary Fund.

Bazarbash, M. (2019). "FinTech in Financial Inclusion: Machine Learning Applications in Assessing Credit Risk". International Monetary Fund.

Beatrice, A. (2013). "Why Women Enter into Entrepreneurship? An Emerging Conceptual Framework Based on The Peruvian Case". *Journal of Women's Entrepreneurship and Education.* 3(4), 43-63.

Becchetti, L., Garcia, M. M., and Trovato, G. (2011). "Credit rationing and credit view: Empirical evidence from an ethical bank in Italy". *Journal of Money, Credit and Banking*, 43(6), 1217-1245.

Beck, T., and Levine, R. (Eds.). (2018). "Handbook of Finance and Development". Edward Elgar Publishing.

Beck, T., Behr, P., and Madestam, A. (2015). "Sex and credit: Is there a gender bias in lending?". *European Banking Center Discussion Paper,* No 2012-017.

Beck, T., Demirguc-Kunt, A., and Levine, R. (2007). "Finance, Inequality and the Poor". *Journal of Economic Growth,* 12 (1): 27–49.

Beck, T., Demirguc-Kunt, A., and Peria, M. (2007). "Reaching out: Access to and use of banking services across countries". *Journal of Financial Economics,* 85(1), 234–266.

Beck, T., Demirguc-Kunt, A., Laeven, L., and Levine, R. (2008). "Finance, firm size, and growth". *Journal of Money, Credit and Banking*, 40(7), 1379-1405.

Bédécarrats, F., Guérin, I., and Roubaud F. (2017). "All That Glitters Is Not Gold. The Political Economy of Randomized Evaluations in Development". *Development and Change.*

Bédécarrats, F., Guérin, I., Morvant-Roux, S., and Roubaud, F. (2019). "Estimating microcredit impact with low take-up, contamination and inconsistent data. A replication study of Crépon, Devoto, Duflo, and Parienté (American Economic Journal: Applied Economics, 2015)". *International Journal for Re-Views in Empirical Economics* (IREE), 3(2019-3), 1-22.

Bianchi, M. (2010). "Financial Development, Entrepreneurship, and Job Satisfaction". *Review of Economics and Statistics*, 94(1): 273–286.

Black, S. and Strahan, P. (2002). "Entrepreneurship and bank credit availability". *Journal of Finance* 57, 2807–2833.

Boucher, S. R., Guirkinger, C., and Trivelli, C. (2009). "Direct elicitation of credit constraints: Conceptual and practical issues with an application to Peruvian agriculture". *Economic Development and Cultural Change*, 57(4), 609-640.

Breza, E., and Kinnan, C. (2018). "Measuring the equilibrium impacts of credit: Evidence from the Indian microfinance crisis". National Bureau of Economic Research. Working Paper No. 24329.

Brown, M., Guin, B. and Kirschenmann, K. (2016). "Microfinance Banks and Financial Inclusion". *Review of Finance*, 20 (3), 907–946.

Brown, M., Haughwout, A., Lee, D., and Van der Klaauw, W. (2010). "The Financial Crisis at the Kitchen Table: Trends in Household Debt and Credit", Federal Reserve Bank of New York Staff Reports, 480.

Bruhn, M. and Love. I. (2014). "The Real Impact of Improved Access to Finance: Evidence from Mexico". *Journal of Finance*, 69 (3): 1347-1376.

Brune, L., Giné, X., Goldberg, J., and Yang, D. (2013). "Commitments to Save: A Field Experiment in Rural Malawi". *World Bank Policy Research Working Paper*, No. 5748. Washington, D.C.: World Bank.

Brune, L., Giné, X., Goldberg, J., and Yang, D. (2016). "Facilitating savings for agriculture: Field experimental evidence from Malawi". *Economic Development and Cultural Change*, 64(2), 187-220.

Bunn, P., and Rostom, M. (2016). "Household debt and spending in the United Kingdom". The UK Economy in the Long Expansion and its Aftermath, 244.

Bunn, P., Johnson, R., Le Roux, J. and Rostom, M. (2015). “Household debt and spending in the United Kingdom”. Bank of England Staff Working Paper No. 554.

Burgess, R. and Pande, R. (2005). “Do Rural Banks Matter? Evidence from the Indian Social Banking Experiment”. *The American Economic Review*, 95(3): 780–95.

Cameron, S. V., and Heckman, J. J. (1998). “Life cycle schooling and dynamic selection bias: Models and evidence for five cohorts of American males”. *The Journal of Political Economy*. 106 (2), 262–333.

Carvalho, L., Prina, S., and Sydnor, J. (2013). “The effects of savings on risk-taking and intertemporal choice behavior: Evidence from a randomized experiment”. Working Paper.

Casella, G. and Berger, R. (2001). “Statistical Inference: Second Edition”. *Duxbury Advanced Series.*

Cecchetti, S. G., and Kharroubi, E. (2012). “Reassessing the impact of finance on growth”. BIS Working Paper 381. Basel: BIS.

Chakrabarty, K. C. (2012). “Financial inclusion: issues in measurement and analysis”. In Keynote address, IFC workshop on financial inclusion indicators, Kuala Lumpur, November.

Chen, M. A. (2001). “Women in the informal sector: A global picture, the global movement”. *SAIS Review*, 21(1), 71–82.

Christelis, D., T. Jappelli., O. Paccagnella. and Weber, G. (2010). “Income, wealth and financial fragility in Europe”. *Journal of European Social Policy,* 19 (4), 359–376.

Claessens, S. (2006). “Access to financial services: a review of the issues and public policy objectives”. *The World Bank Research Observer,* 21(2), 207-240.

Claessens, S. and Laeven, L. (2003). “Financial Development, Property Rights, and Growth”. *The Journal of Finance*, 58: 2401–2436.

Clarke, G. R., Xu, L. C., and Zou, H. F. (2006). “Finance and income inequality: what do the data tell us?”. *Southern Economic Journal*, 578-596.

Cole, S., Giné, X., and Vickery, J. (2017). “How does risk management influence production decisions? Evidence from a field experiment”. *The Review of Financial Studies*, 30(6), 1935-1970.

Cole, S., Sampson, T., and Zia, B. (2011). Price or knowledge? “What drives demand for financial services in emerging markets?”. *Journal of Finance,* 66(6), 1933–1967.

CONAIF (2016). “Política Nacional de Inclusión Financiera, Junio de 2016”. Ciudad de México.

Copestake, J. (2002). “Inequality and the polarizing impact of microcredit: Evidence from Zambia's copperbelt”. *Journal of International Development*, 14(6), 743-755.

Crépon, B., Devoto, F., Duflo, E., and Parienté. W. (2015). “Estimating the Impact of Microcredit on Those Who Take It Up: Evidence from a Randomized Experiment in Morocco”. *American Economic Journal: Applied Economics*, 7(1), 123–150.

Cull, R., Demirguc-Kunt, A ., and Morduch, J. (2009). "Microfinance Meets the Market". *Journal of Economic Perspectives*, 23 (1): 167-92

Cull, R., Ehrbeck, T., and Holle, N. (2014). “Financial Inclusion and Development: Recent Impact Evidence”. CGAP Focus Note 92. Washington, D.C.

Dabla-Norris, M. E., Deng, Y., Ivanova, A., Karpowicz, M. I., Unsal, D. F., VanLeemput, E., and Wong, J. (2015). “Financial inclusion: Zooming in on Latin America”. IMF Working Paper No. 15/206. International Monetary Fund.

Deaton, A. (2009). “Instruments of Development: Randomization in the Tropics, and the Search for the Elusive Keys to Economic Development”. NBER Working Paper No. 14690.

Demetriades, P., and Hook Law, S. (2006). “Finance, institutions and economic development”. *International Journal of Finance and Economics*, 11(3), 245-260.

Demirguc-Kunt, A., and Klapper, L. (2012). “Measuring financial inclusion: The global findex database”. The World Bank.

Demirguc-Kunt, A., and Levine, R. (2009). “Finance and inequality: Theory and evidence”. *Annual Review of Financial Economics*, 1(1), 287-318.

Demirguc-Kunt, A., and Maksimovic, V. (1998). “Law, finance, and firm growth”. *The Journal of Finance*, 53(6), 2107-2137.

Demirguc-Kunt, A., Klapper, L. and Singer, D. (2017). “Financial Inclusion and Inclusive Growth: A Review of Recent Empirical Evidence”. *World Bank Policy Research Working Paper*, No. 8040. Washington, D.C.: World Bank, April.

Demirguc-Kunt, A., Klapper, L. Singer, D., Ansar, S., and Hess, J. (2018). “The Global Findex Database 2017: Measuring Financial Inclusion and the Fintech Revolution”. Washington, DC: World Bank.

Demirguc-Kunt, A., L. Klapper, D. Singer and P. Van Oudheusden. (2015). “The Global Findex Database 2014: Measuring Financial Inclusion Around the World”. World Bank Group Policy Research, Working Paper 7255.

References

Dev, S. M. (2006). “Financial inclusion: Issues and challenges”. *Economic and Political Weekly*, 4310-4313.

Dey, S., R. Djoudad, and Terajima, Y. (2008). “A Tool for Assessing Financial Vulnerabilities in the Household Sector” *International Journal of Microsimulation*, 8(1). 97-109.

Dohmen, T., Falk, A., Huffman, D., Sunde, U., Schupp, J. and Wagner, G. (2011). “Individual Risk Attitudes: Measurement, Determinants, and Behavioural Consequences”. *Journal of the European Economic Association*, 9: 522-550.

Doms, M. Lewis, E., and Robb, A (2010). “Local Labor Market Endowments, New Business Characteristics, and Performance”. *Journal of Urban Economics*, 67, 61-77.

Duflo, E. (2010). “Lutter contre la pauvreté: la politique de l’auitonomie”. *La Republique Des Idees*. Seuil.

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

Dupas, P., Karlan, D., Robinson, J., and Ubfal, D. (2018). “Banking the Unbanked? Evidence from three countries”. *American Economic Journal: Applied Economics*, 10(2), 257-97.

Duvendack M, Palmer-Jones R, Copestake JG, Hooper L, Loke Y, Rao N (2011).” What is the evidence of the impact of microfinance on the well-being of poor people?”. London: EPPI-Centre, Social Science Research Unit, Institute of Education, University of London.

Economist Intelligence Unit (2015). “Global microscope 2015: The enabling environment for financial inclusion”.

Faggio, G., and Silva, O. (2014). “Self-employment and entrepreneurship in urban and rural labor markets”. *Journal of Urban Economics*, 84, 67-85.

Federal Reserve (2018). “Financial Stability Report”. November.

Field, E., Pande, R., Papp, J., and Rigol, N. (2013). “Does the classic microfinance model discourage entrepreneurship among the poor? Experimental evidence from India”. *American Economic Review*, 103(6), 2196-2226.

Filippini, M., Greene, W. H., Kumar, N., and Martinez-Cruz, A. L. (2018). “A note on the different interpretation of the correlation parameters in the Bivariate Probit and the Recursive Bivariate Probit”. *Economics Letters*, 167, 104-107.

Galor, O., and Zeira, J. (1993). "Income distribution and macroeconomics". *The Review of Economic Studies*, 60(1), 35-52.

Georgellis, Y., and Wall, H. J. (2005). "Gender differences in self-employment". *International Review of Applied Economics*, 19(3), 321-342.

Ghosh, P., Mookherjee, D., and Ray, D. (2000). "Credit rationing in developing countries: an overview of the theory". *Readings in the theory of economic development*, 383-401.

Ghosh, S and Vinod, D (2017). "What Constrains Financial Inclusion for Women? Evidence from Indian Micro data". *World Development*, Vol. 92. 60–81

Goldsmith, R. (1969). "Financial structure and development". New Haven: Yale University Press.

Grameen Bank. (2007). "Annual Report 2007". Dhaka.

Grohmann,, A., Klühs, T., and Menkhoff, L. (2018). "Does financial literacy improve financial inclusion? Cross country evidence". *World Development*, 111, 84-96.

Gurley-Calvez, T., Biehl, A. and Harper, K. (2009). "Time-use patterns and women entrepreneurs". *American Economic Review*, 99: 139-144.

Han, R., and Melecky, M. (2013). "Financial inclusion for financial stability: access to bank deposits and the growth of deposits in the global financial crisis". World Bank Policy Research Working Paper 6577.

Han, R., and Melecky, M. (2013). "Financial Inclusion for Financial Stability: Access to Bank Deposits and the Growth of Deposits in the Global Financial Crisis". *World Bank Policy Research Working Paper*, No. 6577. Washington, D.C.: World Bank.

Hannig, A., and Jansen, S. (2010). "Financial inclusion and financial stability: Current policy issues". ADBI Working Paper No. 259.

Hastie, T., Tibshirani, R., and Friedman, J. (2016). "The elements of statistical learning". New York: Springer.

Heckman, J. (1979). "Sample Selection Bias as a Specification Error". *Econometrica*, 47(1), 153-161.

Herceg, I., and Nestić, D. (2014). "A New Cluster-Based Financial Vulnerability Indicator and Its Application to Household Stress Testing in Croatia". *Emerging Markets Finance and Trade*, 50:5, 60-77.

Herrala, R., and Kauko, K. (2007). "Household Loan Loss Risk in Finland- Estimations and Simulations with Micro Data". Research Discussion Paper, Bank of Finland, Helsinki.

Honohan, P., and King, M. (2012). “Cause and effect of financial access: cross-country evidence from the Finscope surveys”. Banking the world: Empirical foundations of financial inclusion, 45-84.

Islam, A. (2015). "Entrepreneurship and the allocation of government spending under imperfect markets". *World Development,* 70, 108-121.

International Monetary Fund (2017). “Pakistan 2017: Article IV Country Report. IMF Staff Country Reports”. Washington DC.

International Monetary Fund (2017). “Regional Economic Outlook: Middle East and Central Asia”. Washington DC.

Jack, W., and Suri, T. (2014). “Risk sharing and transactions costs: Evidence from Kenya's mobile money revolution”. *American Economic Review,* 104(1), 183-223.

James, G., Witten, D., Hastie, T., and Tibshirani, R. (2013). “An introduction to statistical learning”. Vol. 112. New York: Springer.

Jordan, M. I., and Mitchell, T. M. (2015). “Machine learning: Trends, perspectives, and prospects”. *Science,* 349(6245), 255-260.

Karlan, D. and Zinman, J. (2010). “Expanding Credit Access: Using Randomized Supply Decisions to Estimate the Impacts”. *The Review of Financial Studies*, Vol. 23 (1): 433–64.

Karlan, D. and Zinman, J. (2012). “List of Randomization for Sensitive Behavior: An application for measuring use of loan proceeds”. *Journal of Development Economics,* 98 (1): 71-75.

Karlan, D., and Morduch, J. (2009). “Access to finance”. In D. Rodrik and M. Rosenzweig (Eds.). Handbook of development economics. 4703–4784. Amsterdam: Elsevier.

Karlan, D., Osei, R., Osei-Akoto, I., and Udry, C. (2014). “Agricultural decisions after relaxing credit and risk constraints”. *The Quarterly Journal of Economics,* 129(2), 597-652.

Karlan, D., Savonitto, B., Thuysbaert, B., and Udry, C. (2017). “Impact of savings groups on the lives of the poor”. *Proceedings of the National Academy of Sciences*, 114(12), 3079-3084.

Khandker, S., and Faruqee, R. R. (2003). “The Impact of Farm Credit in Pakistan”. *Agricultural Economics*, 28 (3), 197-213.

King, R. G., and Levine, R. (1993). “Finance and growth: Schumpeter might be right”. *The Quarterly Journal of Economics*, 108(3), 717-737.

Klapper, L. (2017). "How digital payments can benefit entrepreneurs". IZA World of Labor.

Klapper, L., El-Zoghbi, M., and Hess, J. (2016). "Achieving the sustainable development goals: The role of financial inclusion". Washington, DC: CGAP.

Klapper, L., Laeven, L. and Rajan, R. (2006). "Entry Regulations as a Barrier to Entrepreneurship". *Journal of Financial Economics*, 82, 2995-3015.

Klapper, L., Lusardi, A., and Oudheusden, P. (2015). "Financial literacy around the world: Insights from the standards and poor's ratings service global financial literacy survey". https://www.finlit.mhfi.com.

Klapper, L., M. El-Zoghbi, and Hess, J. (2016). "Achieving the Sustainable Development Goals: The Role of Financial Inclusion". CGAP.

Kochar, A. (1997). "An empirical investigation of rationing constraints in rural credit markets in India". *Journal of Development Economics*, 53(2), 339-371.

Koellinger, P., Minniti, M. and Schade. C. (2013). "Gender differences in entrepreneurial propensity". *Oxford Bulletin of Economics and Statistics*, 75: 213-234.

Lee, J. N., Morduch, J., Ravindran, S., Shonchoyand, A. S., and Zamany, H. (2017). "Poverty and Migration in the Digital Age: Experimental Evidence on Mobile Banking in Bangladesh". IGC Working Paper C-89233-BGD-1.

Leika, M., and Marchettini, D. (2017). "A Generalized Framework for the Assessment of Household Financial Vulnerability" IMF Working Paper. No. 17/ 228.

Levine, R. (2005). "Finance and growth: theory and evidence". Handbook of economic growth, 1, 865-934.

Leyshon, A., and Thrift, N. (1995). "Geographies of financial exclusion: financial abandonment in Britain and the United States". *Transactions of the Institute of British Geographers*, 312-341.

Llussá, F. (2009). "Financial Development, Gender and Entrepreneurship". Massachusetts Institute of Technology Engineering Systems Division, Cambridge, MA: MIT

Loukoianova, M. E., Yang, Y., Guo, M. S., Hunter, M. L., Jahan, M. S., Jamaludin, M. F., and Schauer, J. (2018). "Financial Inclusion in Asia-Pacific". International Monetary Fund.

Lusardi, A., and Mitchell, O. S. (2014). "The economic importance of financial literacy: Theory and evidence". *Journal of Economic Literature*, 52(1), 5–44.

Mahjabeen, R. (2008). "Microfinancing in Bangladesh: Impact on households, consumption and welfare". *Journal of Policy Modeling*, 30(6), 1083-1092.

Malapit, H. J. L. (2012), “Are women more likely to be credit constrained? Evidence from low-income urban households in the Philippines”. *Feminist Economics*, 18(3), 81–108.

Matul, M., Dalal, A., De Bock, O., and Gelade, W. (2013). “Why People Do Not Buy Microinsurance and What We Can Do About It”. International Labour Office: Microinsurance Paper 20.

Mehrotra, A. N., and Yetman, J. (2014). “Financial inclusion and optimal monetary policy”. BIS Working Paper No. 476, Bank for International Settlements, Basel, Switzerland.

Mian, A., and Sufi, A. (2015). “House of debt: How they (and you) caused the Great Recession, and how we can prevent it from happening again”. University of Chicago Press.

Michelangeli, V., and Pietrunti, M. (2014). “A Microsimulation Model to evaluate Italian Households’ Financial Vulnerability”. *International Journal of Microsimulation*, 7(3), 53-79.

Michelangeli, V., and Rampazzi, C. (2016). “Indicators of financial vulnerability: A household level study”. Banca d’Italia: Questioni di Economia e Finanza.

Morduch, J. (1998). “Does microfinance really help the poor?: New evidence from flagship programs in Bangladesh”. Research Program in Development Studies, Woodrow School of Public and International Affairs.

Morduch, J. (1999). “The Microfinance Promise”. *Journal of Economic Literature*, 37 (4), 1569–1614.

Morduch, J. (2000). “The microfinance schism”. *World Development*, 28(4), 617-629.

Mourifié, I., and Méango, R. (2014). “A note on the identification in two equations probit model with dummy endogenous regressor”. *Economics Letters*, 125(3), 360-363.

Murtagh, F., and Contreras, P. (2012). “Algorithms for hierarchical clustering: an overview”. *Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery*, 2(1), 86-97.

Mylonidis, N., Chletsos, M., and Barbagianni, V. (2017). “Financial exclusion in the USA: Looking beyond demographics”. *Journal of Financial Stability*. JFS 575.

Nagler, P., and Naudé, W. (2017). “Non-farm entrepreneurship in rural sub-Saharan Africa: New empirical evidence”. *Food Policy*, 67, 175-191.

OECD (2016). “Entrepreneurship at a Glance 2016”. OECD Publishing. Paris,

Pakistan Bureau of Statistics (2015). “Labour Force Survey 2014-15: Thirty Third Issue”. Government of Pakistan.

Pakistan Bureau of Statistics (2017). “Population Census 2017”. Retrieved from http://www.pbscensus.gov.pk/.

Pakistan Microfinance Network (2016). “MicroWatch: A Quarterly Update on Microfinance Outreach in Pakistan”. Issue 42. Quarter 4: October to December 2016. Islamabad: Pakistan.

Pakistan Microfinance Network (2017). “MicroWatch: A Quarterly Update on Microfinance Outreach in Pakistan”. Issue 46. Quarter 4: October to December 2017. Islamabad: Pakistan.

Parienté, W. (2005). “Financial Demand and Access to Credit in Low-Income Areas of Rio de Janeiro, Brazil”. Pesquisa & Debate. Revista do Programa de Estudos Pós-Graduados em Economia Política, 16 (2 (28)), 269-308.

Pasali, S. (2013). “Where Is the Cheese? Synthesizing a Giant Literature on Causes and Consequences of Financial Sector Development”. World Bank Policy Research. Working Paper No. 6655. Washington D.C.: World Bank, October.

Pence, K. (2015). "SCFCOMBO: Stata module to estimate errors using the Survey of Consumer Finances". Statistical Software Components S458017, Boston College Department of Economics, revised 04 Oct 2015.

Pitt, M. M. (2014). “Response to: The Impact of Microcredit on the Poor in Bangladesh: Revisiting the Evidence”. *Journal of Development Studies*, 50:4, 605-610.

Pitt, M. M. and Khandker, S. R. (2012). “Replicating replication: Due diligence in Roodman and Morduch’s replication of Pitt and Khandker (1998)”. Working Paper 6273, World Bank: Washington, DC.

Pitt, M. M., and Khandker, S. R. (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.

Pitt, M. M., Khandker, S. R., and Cartwright, J. (2006). “Empowering women with micro finance: Evidence from Bangladesh”. *Economic Development and Cultural Change*, 54(4), 791-831.

Prina, S. (2015). “Banking the poor via savings accounts: Evidence from a field experiment”. *Journal of Development Economics*, 115, 16-31.

Pritchett, L. and J. Sandefur (2015). “Learning from Experiments when Context Matters”. *American Economic Review*. 105(5): 471–75.

Rajan, G. R., and Zingales, L. (1998). “Financial Dependence and Growth”. *The American Economic Review*, Vol. 88, No. 3. pp. 559-586.

Rajan, R. G., and Zingales, L. (1998). “Power in a Theory of the Firm”. *The Quarterly Journal of Economics*, 113(2), 387-432.

Ramani, S. V., Thutupalli, A., Medovarszki, T., Chattopadhyay, S., and Ravichandran, V. (2013). “Women in the informal economy: Experiments in governance from emerging countries”. United Nations United Nations University Policy Brief, (5).

Reserve Bank of Australia (2017). “Financial Stability Review”. April.

Rodrik, D. (2008). “The New Development Economics: We Shall Experiment, But How Shall We Learn?”. Harvard Kennedy School Faculty Research Working Paper Series, 08-055.

Rojas-Suarez, L. (2010). “Access to Financial Services in Emerging Powers: Facts, Obstacles and Policy Implications”. OECD Development Centre Publishing.

Romer, P. M. (1990). “Endogenous technological change”. *Journal of Political Economy*, 98(5, Part 2), S71-S102.

Roodman, D. and Morduch, J. (2011). “Comment on Pitt’s responses to Roodman and Morduch”. Center for Global Development: Washington, DC.

Roodman, D., and Morduch, J. (2014). “The impact of microcredit on the poor in Bangladesh: Revisiting the evidence”. *Journal of Development Studies*, 50(4), 583-604.

Rousseau, P. L., and Wachtel, P. (2002). “Inflation thresholds and the finance–growth nexus”. *Journal of International Money and Finance*, 21(6), 777-793.

Ruiz-Tagle, J., and Vella, F. (2015). “Borrowing Constraints and Credit Demand in a Developing Economy”. *Journal of Applied Econometrics*, 31(5), 865–891

Samargandia, N., Fidrmuca, J., and Ghosha, S. (2015). “Is the Relationship Between Financial Development and Economic Growth Monotonic? Evidence from a Sample of Middle-Income Countries”. *World Development*. Vol 68. 66–81.

Sarma, M. (2008). “Index of Financial Inclusion”. Indian Council for Research on International Economic Relations. Working Paper 2015, August.

Schularick, M., and Taylor, A. M. (2012). “Credit booms gone bust: Monetary policy, leverage cycles, and financial crises, 1870-2008”. *American Economic Review*, 102(2), 1029-61.

Schumpeter, J. A. (1911). “The theory of economic development: An inquiry into profits, capital, credit, interest, and the business cycle”. Oxford University Press, Oxford (Reprinted, 1969).

Shaw, E. S. (1973). “Financial deepening in economic development”. Oxford University Press, New York.

Stiglitz, J., Weiss, A. (1981). “Credit Rationing in Markets with Imperfect Information”. *The American Economic Review*, Vol. 71- 3, pp. 393-410.

Talledo, J. (2015). "Access to and use of financial services: Evidence from Peru". SBS Working Paper Series. WP 03/2015.

Tarozzi, A., Desai, J., and Johnson, K. (2015). “The impacts of microcredit: Evidence from Ethiopia”. *American Economic Journal: Applied Economics*, 7(1), 54-89.

UNDP (2016), “Multidimensional Poverty in Pakistan” Planning Commission of Pakistan, Islamabad.

Van Rooyen, C., Stewart, R., and De Wet, T. (2012). “The impact of microfinance in sub-Saharan Africa: a systematic review of the evidence”. *World Development*, 40(11), 2249-2262.

Vatne, B. H. (2006). “How large are the financial margins of Norwegian households? An analysis of micro data for the period 1987–2004”. Norges Bank Economic Bulletin 4/06, 77: 173-180.

Vejsiu, A. (2011). “Incentives to the self-employment decisions in Sweden”. *International Review of Applied Economics*. Vol. 25. 379-403.

Wooldridge, J. M. (2010). “Econometric analysis of cross section and panel data”. MIT press.

World Bank (2014). “Global financial development report 2014: Financial inclusion (Vol. 2)”. World Bank Publications.

World Bank. (2014), World Development Report 2014. Washington, D.C.: World Bank.

World Bank. (2017). “The Global Findex Database 2017”. Washington DC.

Worthington, C. (2006). “Debt as a source of financial stress in Australian households”. *International Journal of Consumer Studies*, 30(1). 2-15.

Xu, L., and Zia, B. (2012). “Financial literacy around the world”. World Bank Policy Research Paper, No. 6107.

Yunus, M. (1998). “Banker to the Poor”. Penguin Books India.

Yunus, M. (1998). “Poverty alleviation: Is economics any help? Lessons from the Grameen Bank experience”. *Journal of International Affairs*, 47-65.

Yunus, M. (2010). “Building social business: The new kind of capitalism that serves humanity's most pressing needs”. PublicAffairs.