Evaluating the economic impacts of rural banking: experimental evidence from southern India

Erica Field, Duke University

Rohini Pande, Harvard University

Grantee Final Report

Accepted by 3ie: May 2018

International Initiative for Impact Evaluation

Note to readers

This impact evaluation has been submitted in partial fulfilment of the requirements of grant OW3.1011 awarded under Open Window 3. This version is being published online as it was received. A copy-edited and formatted version will be available in the 3ie Impact Evaluation Report Series in the near future. All content is the sole responsibility of the authors and does not represent the opinions of 3ie, its donors or its board of commissioners. Any errors and omissions are the sole responsibility of the authors. All affiliations of the authors listed in the title page are those that were in effect at the time the report was accepted. Any comments or queries should be directed to the corresponding author, Rohini Pande at rohini_pande@hks.harvard.edu

Suggested citation: Field, E and Pande, R, 2018. *Evaluating the economic impacts of rural banking: experimental evidence from southern India, 3ie Grantee Final Report.* New Delhi: International Initiative for Impact Evaluation (3ie)

Acknowledgements

Executive Summary

Three decades of sustained growth has contributed to a halving of Indian poverty rates. Yet, one in every four Indians is still classified as being extremely poor and lives on less than $1.90 a day (World Bank, 2013).[1] Further, income inequality in India is fast rising with limited changes in the well-being of many poor rural households. How can public policy in India best respond to the economic needs of its poor rural citizens?

Improved access to formal financial sources has long been considered a critical element of policy responses directed at the rural poor. A large theoretical economics literature suggests that increasing financial access has the potential to enable individuals to exit poverty by altering their production and employment choices, and by helping them retain productive assets when income shocks hit. Quasi-experimental evidence from India's social banking experiment suggests that such policies can reduce aggregate poverty (Burgess and Pande, 2005). However, less is known about the channels of influence and whether these channels remain operative in today's vastly richer India. One may, for instance, argue that the remaining poor populations have demographic and economic characteristics that imply that they are less able to benefit from formal institutions and need specific grant programs (like, say, the ultra-poor program). Equally, it is unclear whether the general equilibrium effects associated with increased banking – which could include changing social networks and altered presence of informal lenders – help or harm the poorest households.

Over the last few decades, the emergence of microcredit has also revamped the financial landscape for the poor. The relatively high cost of enabling brick-and-mortar banking access has led many to question the value of continued investment in rural banking, and indeed the last few decades have seen the private sector focus more on using Grameen-style microcredit to reach underserved individuals who mostly rely on informal (and more expensive) lenders. Reflecting this, experimental evaluations of financial access for the poor have focussed on evaluating Grameen-style microcredit, largely in urban populations that have relatively good access to credit. Results from this recent body of work suggest positive – but not transformative – effects of microcredit on the lives of the urban poor. There are to date no experimental evaluations of financial access (we will often refer to this as microfinance) in relatively unbanked settings where liquidity constraints are likely to be the most binding for a wide range of investment choices. In addition to the urban focus, it is also important to note that these studies focus on the impacts of offering specific financial products on specific household outcomes, and have yet to demonstrate the effect of financial access *as a whole* on household poverty. Hence, there are still gaps in our understanding of the effect of expanding financial access more broadly on the multiple dimensions of poverty, income, self-employment, and overall well-being.

In collaboration with IFMR Rural Channels under IFMR Trust, we designed and set-up a randomized controlled trial (RCT) to study the effects of the expansion of a rural branch banking model in Tamil Nadu, India. The partner, as a Non-banking Financial Company and Business Correspondent, uses the financial services delivery model called 'Kshetriya Grameen Financial Services' (KGFS) to provide a range of financial services. It also provides tailored financial advice through local brick and mortar village branches, thus representing an alternative to the standard microfinance movement in India, which has

---

[1] Expressed in 2011 PPP.

focused primarily on microcredit. The type of financial products offered by KGFS, along with the large data collection effort carried out by the research team, allows our study to be the first to evaluate the impact of increased access to financial services as a whole. In addition, our focus on rural areas allows us to provide valuable insights on this model of expanding financial access to remote rural communities. Starting in 2009, we identified 101 service areas over three districts – Ariyalur, Pudukkottai, and Thanjavur – from which we formed 50 service area pairs[2]. We then randomly assigned service areas to treatment and control groups status within each pair. Each service area was also assigned a branch location, and KGFS opened branches in treatment group service areas at the time of assignment, while expansion into control group areas occurred no less than 24 months later. The average service area of a bank branch spanned a radius of 3-5 km from its assigned branch location and covered approximately 10,000 people, or 10 villages. More than 4,000 households were then randomly selected across all service areas to be included in the main component of the study. A separate survey was conducted on about 19,000 households in order to create detailed village social network maps.

KGFS began opening branches in treatment service areas in 2010. We conducted surveys to a sample of households in each new service area, as the pair entered the study. Our baseline surveys (paralleling branch opening) occurred between 2010 and 2014, and endline surveys were administered between 2013 and 2016, i.e 18 to 24 months after branch opening. We obtained information on financial access (borrowing and saving), economic activity, shocks and well-being. This report presents the core comparisons across treatment and control groups for our main outcomes of interest. Living in an area where KGFS expanded increases households' likelihood of participating in formal banking. Compared to the control group, at endline, treated households are more likely to have formal outstanding loans, have a larger number of formal loans, and borrow more from formal lenders. They also report higher saving amounts. Households in treated areas are also less likely to borrow from informal sources such as moneylenders and financiers. Importantly, these household level changes are paralleled in network changes: households in treated villages report lower borrowing capacity both from moneylenders and from individuals living both inside and outside their village than control group households. Greater access to formal finance enables households to benefit from greater economic opportunities: our intervention increases households' likelihood of being self-employed and raises business income. Treated households are also more likely to use formal loans for business purposes. Consistent with a large body of theoretical research, we find that formal financial access promotes entrepreneurship and encourages households to take on riskier – but more profitable – activities. This, in turn, has a significant, positive effect on business income and on overall households' income.

All in all, our initial results suggest that expanding access to formal financial products and services to rural households not only crowds out informal borrowing, but also has a positive impact on saving, on their business activities, and on their ability to cope financially with health shocks. We also find a positive treatment effect on wages. This result is consistent with the hypothesis that the poor shift from farming to self-employment, or that the poor diversify their activities by starting a business. The main takeaway of this report is that, in our study, increasing access to formal financial services seems to positively impact poor households through income stabilization and increased financial security.

---

[2] One service area "pair" is a triplet, containing one treatment area and two control areas. Our 49 pairs and 1 triplet "pair" total 101 total service areas covered, with 50 treatment areas and 51 control areas.

Contents

List of Figures and Tables

Abbreviations and Acronyms

ATE	average treatment effect
NBFC	non-banking financial corporation
CMF	Centre for Micro Finance
DiD	difference in differences
GIC	general insurance corporation
ICT	information and communication technologies
IFMR	Institute for Financial Management and Research
IMRB	International Market Research Bureau
ITT	intention to treat
JLG	Joint-Liability Group
KGFS	Kshetriya Grameen Financial Services
LEAD	Leveraging Evidence for Access and Development
LFI	large financial institution
LIC	life insurance corporation
MDE	minimum detectable effect
MFI	micro finance institution
NGO	non-governmental organizations
NREGA	National Rural Employment Guarantee Act
PAC	primary agricultural cooperative
PFSPL	Pudhuaaru Financial Services Private Limited
PMJDY	Pradhan Mantri Jan Dhan Yojna
RBI	Reserve Bank of India
RCT	Randomized Controlled Trial
ROSCA	rotating savings and credit association
SBU	strategic business unit
SHG	self-help group
SD	standard deviation
SE	standard error
SNM	social network mapping

1. Introduction

Sustained economic growth has played a critical role in lowering poverty in India but the benefits of growth remain unevenly distributed. According to the World Bank (2013), almost one quarter of the Indian population continue to live below $1.90 a day.[3] Can better access to finance help India's rural poor better benefit from economic growth and widen the economy's structural transformation?

A large body of theoretical work in economics suggests that the provision of financial products to underserved individuals (also, often termed microfinance) can play a critical role in helping poor households alter their production and employment choices. This, in turn, can enable a virtuous cycle where they lift themselves out of economic marginalization by increasing the security of their assets, helping them absorb economic shocks, and allowing them to borrow their way to a higher socio-economic status (Aghion and Bolton, 1997; Banerjee and Newman, 1993; Banerjee, 2004).

Early policy enthusiasm for this theory of change was reflected in India's large social banking experiment, and Burgess and Pande (2005) provide quasi-experimental evidence on its poverty impact. They also show that increased density of rural banks increased rural credit and savings. Alongside, they find some reduced form evidence of structural transformation, but the use of aggregated data limits their ability to examine mechanisms. Post-liberalization (starting 1991) Indian banks were given much more freedom in terms of branch placement, and Indian banks responded by increasing bank branch density largely in urban India. In our baseline data from rural India we see that the average rural household has limited access to formal banking: indeed, 40% of our sample report not having any formal loan at the beginning of the study, and almost 20% do not save in a formal saving account.

Instead, for two and half decades after economic liberalization Indian policymakers and the private banking sector alike placed weight on using Grameen-style microcredit expansion to provide the poor financial access. This was a common trend in much of the developing world and this was also reflected in academic research. An increasing number of experimental studies conducted since the early 2000s focused on estimating the impact of providing microcredit to poor households. Karlan and Zinman (2010), Banerjee et al. (2015), and Crépon et al. (2015), among others, study the effect of facilitating access to microloans in South Africa, India, and Morocco, respectively. Though these studies are suggestive of a positive impact of microcredit on business expansion and employment, they find only small effects on other outcomes. This can also be partly explained by the modest take-up rates of microloans (Banerjee, Karlan and Zinman, 2015). Randomized evaluations of micro-insurance (weather-indexed) products also find very low take-up rates (Cole, Stein and Tobacman, 2014). In light of this, the experience that other members from the same social network had with the product (Karlan et al., 2014) and innovative contract features (Casaburi and Willis, 2016) seem to positively affect individuals' demand for index-insurance. Perhaps unsurprisingly, given that simpler design of financial products could explain higher take-up rates, recent evaluations of savings products show more promising results. Providing saving accounts has beneficial effects on business investment and income (Dupas and Robinson, 2013) and on households' ability to cope with adverse shocks (Prina, 2015), but the positive effect of access to savings technologies can be heavily undermined by intra-household

---

[3] Expressed in 2011 PPP.

pressures (Schaner, 2015). At the same time, it is also the case that take-up is key – as shown by Dupas et al. (2017).

The advantage of experimental studies that directly examine impacts on household outcomes is that they are able to take a closer look at channels of influence. However, a key limitation of studies focusing on the provision of stand-alone financial products is that they fail to demonstrate the effects of broad financial access *as a whole* on households' poverty. A closer look at the related theoretical literature suggests that the relevant policy question is unlikely to be whether to introduce a single savings product or approve a particular type of loan, but rather relates to availability of a suite of financial products. Put differently, financial services taken as a whole – whether they are savings accounts, insurance products, collateral-based loans or joint liability group loans – provide a formal mechanism for shifting income from one state of the world to another. Non-experimental studies of more "holistic" programs of financial inclusion have found large effects on households' welfare (Burgess and Pande, 2005; Kaboski and Townsend, 2006; Bruhn and Love, 2014). However, these studies are often unable to examine specific pathways in detail and, also, the banking or microfinance programs being studied often occur contemporaneous with other policy changes.

Against this background, our India-focused study was designed to provide experimental evidence on the key development question of interest: *What is the impact of increased access to financial services as a whole?* To the best of our knowledge, there are no other examples of randomized bank branch placement at scale. Our study encompasses 50 KGFS branches, covering 850 villages and a population of more than 25,000 surveyed households.

The two papers closest in nature to our focus are the non-experimental evaluation of the Indian Social Banking Experiment done by Burgess and Pande (2005) and the experimental evaluation of Spandana in India run by Banerjee et al. (2015). Context-wise, although Burgess and Pande (2005) also focus on rural banks, the current financial landscape and options in India differ greatly from the government rural banks built two decades ago that were analysed in their study. We complement the focus of that study by providing granular evidence on the channels of influence, and by examining outcomes at the household level. Our approach of randomizing the area of operation of a financial provider is similar to the experimental methodology adopted by Banerjee et al. (2015). However we differ in the nature of our intervention, as we focus on rural areas where alternative sources of credit are less common. Thus, the evidence from this evaluation provides valuable and innovative inputs for evaluating banking models specifically aimed at expanding financial access to remote rural communities.
Finally, our experimental set up allows us to speak to general-equilibrium effects, by looking at the impact of expanding credit supply (and, more broadly) financial access on social networks, on the presence of informal lenders, and on wages (Burgess and Pande, 2005; Breza and Kinnan, 2016).

Our research was designed in collaboration with IFMR Rural Channels under IFMR Trust and builds on the expansion of a large rural financial institutional model in Tamil Nadu, South India, starting in March of 2010. The partner, as a Non-banking Financial Company and Business Correspondent, uses the financial services delivery model called 'Kshetriya Grameen Financial Services' (KGFS) to provide a range of financial products spanning loans, savings, and insurance. It also provides tailored financial advice through local village branches, in order to effectively reach individuals in financially marginalized

rural communities. Overall, KGFS represents an alternative to the standard microfinance movement in India, which has focused primarily on microcredit.

We see the key innovation of our study being the 'at-scale' nature of our experimental intervention – this is key to study the impact of financial access taken as a whole, thus capturing general equilibrium effects. One related innovation is that our data collection endeavour not only gathered information on households' characteristics and behaviour, but also mapped financial and social networks in the villages under study. Our final data represent one of the largest complete social network mappings in India.

This report uses data collected from 2010 to 2016 on a sample of 4,160 households to assess the impact of expanding financial access on households' poverty. We also evaluate the impact of expanding financial inclusion on village-level outcomes such as the presence of informal lending sources (moneylenders and financiers), as well as on social and financial networks. For this last dimension, we also collected social network information for 19,183 households in the study, representing the entire population in 204 villages.

2. Study Context

In 1991, India launched a large program of economic liberalization. The period until the early 2000s largely saw a reduction in financial sector regulation vis-à-vis servicing the rural economy. However, since the early 2000s, there has been a heightened regulatory focus on providing financial services to the poor, with the belief that formal financial services allow the poor to develop income-generating activities and improve their ability to cope with shocks.

Specifically, during the period between 2010 and 2016, the Reserve Bank of India (RBI) took several steps to accelerate financial inclusion and increase access to banking services. Under the two phases of the Financial Inclusion Plan (FIP) implemented during 2010-2013 and 2013-2016, the Government ordered banks to adopt a structured and planned approach to financial inclusion. This consisted of extending branch networks into rural areas in order to bring banking within the reach of the masses, as well as various forms of ICT-based models, including banking through business correspondents.

The two phases of the FIP were then integrated with the "Pradhan Mantri Jan Dhan Yojna" (PMJDY) program, one of the biggest government sponsored financial inclusion programs, which aimed to bring access to basic financial services to every household in India.

The expansion of KGFS in rural areas of Tamil Nadu, which started in 2010, and whose impact is our object of investigation, can be seen as part of this attempt from formal financial institutions to promote financial access among the poorest. As previously mentioned, KGFS is a group of Strategic Business Units (SBUs) under an Indian non-banking financial company called Pudhuaaru Financial Services Private Limited (PFSPL). KGFS' stated mission is to "*maximize the financial well-being of every individual and every enterprise in remote rural India by providing complete financial services.*" In line with this goal, during its expansion, KGFS explicitly targeted villages with low access to banking services. Indeed, a key requirement in branch site selection was that the service area contained neither private banks nor more than one state-run bank.

3. Timeline

Baseline data collection started in September 2010 and finished in September 2014.[4] In total, 4,066 households living in 50 pairs of services areas were interviewed at baseline for the main household component: 17 pairs were surveyed from September 2010 to March 2011; 26 pairs were surveyed between October 2012 and August 2013; 7 pairs were surveyed between July 2014 and September 2014.

Endline data collection started in March 2013 and was completed in December 2016: 8 pairs were surveyed from March 2013 to May 2013; 34 pairs between February 2015 and August 2016. Finally, 8 pairs were administered endline questionnaires from September 2016 to December 2016.

4. Theory of change, Intervention and Research Questions

4.1 THEORY OF CHANGE

Our theory of change is shown in Figure 1. The opening of a new KGFS branch in a certain service area increases – or, in some cases, introduces for the first time – the availability of formal financial products in that area. This is especially true since the expansion of KGFS takes place in rural areas, which are less served or not served at all by other formal financial institutions. Indeed, according to the 2011 Census, only 54.4% of rural Indian households avail banking services, compared 67.8% of urban population.[5]

Instead, informal lenders represent the main financial service providers operating in these areas.[6] It follows that the expansion of KGFS offers rural households the opportunity to access a new range of formal financial products. The products offered by KGFS are cheaper than those offered by informal lenders. To this end, an in-depth study of rural markets in Tamil Nadu[7] shows that financiers set an average annualized interest rate of 54%, whereas MFIs in the same area levy 25% in interest fees.

Once KGFS' loans become available, we see KGFS' expansion having two main types of effects: i) at the household level; ii) at the village level. Households living in service areas where KGFS has expanded should increase their formal financial activity (borrowing and saving) and decrease their reliance on informal lenders. This has an immediate effect on households' ability to cope with shocks – indeed, when an unexpected event happens, households can now rely on cheaper sources of borrowing which were not previously available. This, in turn, should have an effect on households' level of psychological distress: a better ability to cope with shocks should translate into better psychological well-being.

---

[4] The temporal gaps between survey rounds for the baseline should be accounted for the unexpected delay of KGFS in opening new branches. This was in part due to the outbreak of the microfinance crisis in 2010.

[5] Department of Financial Services, Government of India, retrieved from: https://rbidocs.rbi.org.in/rdocs/Speeches/PDFs/MFI101213FS.pdf

[6] See NIC report *(http://planningcommission.nic.in/reports/genrep/rep_fr/cfsr_all.pdf*)

[7] *http://www.ifmrlead.org/wp-content/uploads/2016/04/KGFS_Brief2_Financiers.pdf*

At the same time, formal credit being cheaper than informal one, we should observe, from an occupation point of view, an increase in the use of formal loans for productive purposes. This should translate into higher investment in riskier but also higher-return investments and activities, such as self-employment. This should have a positive effect on business income and, potentially, on households' income. It follows that households' wealth and asset ownership should also increase, contributing in a positive way to households' well-being.

From a village-level perspective, the entry of a new formal financial institution in a certain area should increase the competition among financial services providers. Assuming that there is a much larger presence of informal than formal lenders, the former should be crowded out by the latter. When this happens, informal lenders should adjust the prices or the credit terms of informal financial products in order to be competitive in the new financial landscape.[8] In addition, the presence of a formal financial institution in a village should reduce the frequency of informal financial transfers within social networks as they are replaced by formal financial transactions.

At the same time, for the same reasons explained earlier, once a formal financial institution enters in a village, it is likely to expect an increase in self-employment activities benefiting from cheaper, formal credit.

Figure 1: Theory of change

---
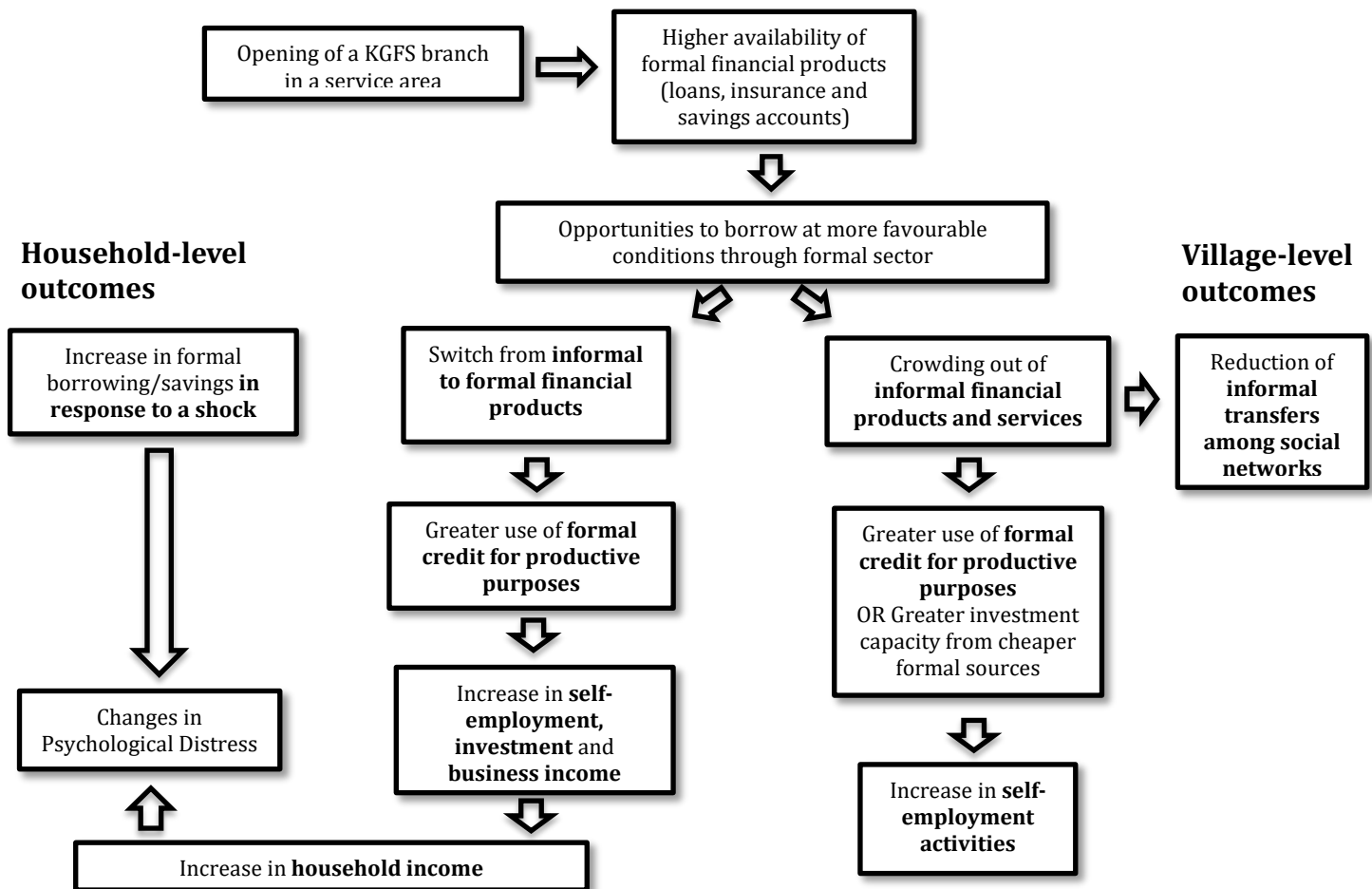
[8] Preliminary results from looking in greater detail at the behavior of informal lenders indeed suggest that they tend to adjust their behavior with respect to the entry of KGFS by increasing flexibility in their contract terms.

4.2 INTERVENTION[9]

The intervention involved providing a complete suite of formal financial services to rural populations living in Tamil Nadu, by the implementation partner of the study, KGFS. Each KGFS institution is designed to be a regional institution serving a specific territory with distinct geographic and socio-economic characteristics. Each branch is considered as a separate business unit and roughly serves a population of 10,000 individuals and 2,000 households. Each branch has on average 2-3 'wealth managers' who perform all administrative tasks and service customers. Each of these managers is a local resident of the area who has deep knowledge of his respective areas.

KGFS model operates on three broad principles: i) focused geographic commitment and complete population coverage, ii) client wealth management approach, and iii) access to a broad range of formal financial services. Such model makes KGFS stand out from other financial institutions that serve the poor and low-income households in rural and remote areas.

A first key component of the KGFS expansion is to enrol the population that resides in the service area of the KGFS branch. Eligible customers must be between 18 and 58 years old and must reside in the service area of the respective KGFS branch. A customer is considered enrolled if her or his details are entered into the KGFS database. As a first step, the individual details of the customer are collected and Know Your Customers (KYC) norms are adhered to by collecting unique identity information details. The second part of the enrolment process relates to gathering the customer's household information. This includes information on household income, expenditures, assets and liabilities. This information is used to generate a financial well-being report of the customer (at the branch level), which is then used to provide financial advice for each client. Based on the customers' financial report, the wealth managers offer financial products to their customers that will be well suited for their profile.

Products are grouped into four broad categories that correspond to clients' needs and objectives.
- Plan: Financial products that help people manage short-term liquidity needs. These include savings, mutual funds, short-term loans, payment services, jewel loans, joint liability loans, emergency loans, etc.
- Grow: Financial products that help households to increase income or reduce expenses. These include business working capital loans, education loans, livestock loans, housing loans, etc.
- Protect: Financial products that mitigate risks and include many types of insurance. These include several types of insurance policies.
- Diversify: These include long-term investment instruments such as pension schemes and gold investment schemes, etc.

KGFS initiated branch openings in March 2010. Once opened, each branch offered a suite of financial products (according to customer needs and profile), as mentioned above.

---

[9] This section heavily borrows from the CGAP report titled '*The pursuit of Complete Financial Inclusion- The KGFS Model in India*' authored by Bindu Ananth, Gregory Chen and Stephen Rasmussen. This section also draws from some of recent visits to KGFS branch by the research team to understand the working of the KGFS model.

Starting in 2009, in advance of the branch openings, we worked with KGFS to identify 101 service areas over the three districts of Ariyalur, Pudukkottai, and Thanjavur, from which we formed 50 service area pairs[10]. Our field team then randomly assigned treatment and control group statuses within each pair. Randomization within matched pairs provided a natural framework for simultaneous surveying, and also minimized imbalance in underlying characteristics across treatment and control service areas by imposing spatial symmetry[11] on the treatment and control areas.

Each service area was assigned a branch location: KGFS branches were opened in treatment group service areas at the time of assignment, while expansion into control group areas occurred 24 months later. The average service area spanned a radius of 3-5 km from its assigned branch location.

A total of 4,066 households were then randomly selected across all service areas to be included in the main component of the study. A separate survey was conducted on an additional 19,183 households to create detailed village network mappings as well.

4.3 RESEARCH QUESTIONS

Our research design and extensive data collection allow us to address the following research questions:
• How does the borrowing and saving behaviour of households change?
• What is the effect on income and employment outcomes for household members?
• How are investments in agricultural and non-agricultural activities influenced?
• Are households better able to deal with shocks as measured by responses to health shocks?
In what follows, we address each of these questions in detail.

5. Programme Implementation

5.1 INTERVENTION IMPLEMENTATION

IFMR Lead worked closely with the IFMR Trust and IFMR Rural Channels Services, the implementing partners, to finalize the design of the study and monitor the timely implementation of the intervention. Bank branch openings occurred in three phases during 2010 and 2015, and researchers at IFMR LEAD were in constant engagement with the partners to discuss any challenges related to implementation and potential solutions for the same. Eventually of the 50 treatment branches that were to be opened, only 48 could be opened due to logistical challenges faced on the field.[12]

5.2 EVALUATION

---

[10] One service area "pair" is a triplet, containing one treatment area and two control areas. Our 49 pairs and 1 triplet "pair" total 101 total service areas covered, with 50 treatment areas and 51 control areas.
[11] Pairs were formed based on a minimum distance criterion between service area branch locations. Thus, spatial symmetry exists between the treatment and control groups by design.
[12] The service areas for the two branches that were not opened are included in the study and have been classified as 'Intent to Treat'.

5.2.1 Study sample

The details of the sampling strategy are provided in Appendix A. Table 27 (Appendix F) shows descriptive statistics of the study sample from the baseline surveys, which were administered starting in 2009. For a comprehensive description of the variables and how they were constructed, refer to Appendix E.

A total of 4,066 individuals were surveyed at baseline for the main household component. On average, households at the beginning of the study consisted of 4.52 members (of which 3.14 were the members older than 18 years old). 72% of these households had a male head. Household's heads were on average 46.68 years old, with an average of 7.49 years of education. The average distance from a household's place of residency to the closest KGFS branch was 2.18 km.

In terms of occupation, 16% of the households reported being self-employed or owning a business at baseline, with an average business income of Rs. 2435.54 in their most recent 30 days of business activity. Conversely, 63% of the households were employed in non-household wage labor or services in the last 7 days, averaging a weekly wage labor income of Rs. 841.[13] As for agricultural labor specifically, 45% of households included in the sample reported farming in the previous season. This is not surprising: the intervention under study specifically targeted rural areas. That said, only 55% of the households in the sample own the land they farm – and 43% live below the poverty line – when we look at households' wealth and properties.

As for financial access characteristics, the average numbers of formal and informal outstanding loans were 1.18 and 1.86, respectively. Over the 24 months preceding baseline the interview, households had borrowed on average Rs. 46,566.11 from formal sources and Rs. 41,326.46 from informal sources. The average probability of having any saving account (formal or informal) at baseline was 84%. The average amount saved in any savings account was Rs. 5,424.46. Only 5% of the households in the study reported having given out a loan to friends and relatives.

We also examined whether, in the past twelve months, households experienced any shocks. 39% of our sample reported having experienced any type of income shock, while 21% declared having been hit by a serious injury/illness over the same time span.

Table 26 (Appendix F) shows baseline randomization checks in terms of: demographics, type of facilities, and financial sector characteristics at the service area level (Panel A); demographics and main outcome variables for the main Household sample (Panel B); the Social Network Mapping sample described at the village level (Panel C).

Panel A shows that no imbalances are detected in terms of demographics, type of facilities, and financial sector characteristics at the service area level.

When we focus on the Household sample, as shown in Panel B, out of 22 indicators tested at the household level, we find significant differences between the control and treatment groups on five measures only. Three are only weakly significant at the 10% level: distance to the nearest branch (with

---

[13] All rupee amounts, including total weekly wage labor income and total business income in the last 30 days, have been topcoded to three standard deviations.

0.15 kms less for the treated group compared to the control group), informal borrowed amounts (only - 8.5% less in treatment group), and the probability that the household has experienced a shock in the last twelve months (with the difference at approximately 5%, i.e. economically very small).

Other statistically significant differences we find between treatment and control are: the age of the household's head, the total borrowed amount from formal sources, and the number of informal loans taken out. Though statistically significant, the difference in household's head age appears to be of less than one year. Regardless, due to this imbalance, we utilize age of household's head as a control in the later analysis. The difference in total borrowed amounts from formal sources is economically small (9%). Finally, the difference in numbers of outstanding informal loans is also small (7%).

Among the three measures we compare across treatment and control in the Social Network Mapping sample (Panel C), only one is weakly significant across the two groups: number of surveyed households in a village. Again, the difference is small in terms of magnitude (10%).

5.2.2 Randomization

Within each pair, one site was randomly selected for branch opening, giving us a total of 50 service area pairs[14] across three districts.[15] In order to avoid any contamination, expansion in the control branch in each pair occurred two years after the treatment branch was opened. The average service area of a branch spanned 3-5 km from the branch office and encompassed an average of 10 villages and 2,400 households. Villages serviced by a single branch were typically well-connected by roads and bus routes. Below we first describe randomization and then surveying. The selection of potential branch sites and randomization across them proceeded as follows:

I. **Geographic survey (GPS Survey)**. In conjunction with the bank, a global positioning system (GPS) based population survey was conducted to determine all relevant political, administrative, and social boundaries. Patterns of business activity, road connectivity, and land availability were also assessed.

II. **Nomination of branch sites**. Once all feasible branch locations in the district had been designated – using information from the GPS survey – sites for branch location were nominated, such that for each branch, a service area could be constructed to keep a population of 10,000 individuals within a rough three-kilometer radius. The bank's primary goal in this process was to ensure that no pocket was left unserved. All nominated sites were reviewed by bank's infrastructure staff and signed off at the level of IFMR Rural Finance's president.

III. **Nomination of service areas**. In conjunction with the bank, the research team then nominated units of population to be 'mapped' to each branch site, such that access on foot or by road was easy and also intuitive for the population served by each branch. Service areas were defined down to the street level, with the intermediate units being political, administrative, and social villages.

IV. **Matching of branch sites**. The unit of randomization in this intervention is the area served by a single bank branch. Yet, some issues for causal inference are confounding factors and clustering

---

[14] 101 service areas are covered, due to one triplet with two control areas.
[15] The timing of the intervention was agreed with KGFS, with strict monitoring from the research team's end, thus ensuring that the branches were opened as quickly as possible.

correlation. More specifically, we were faced with two problems: seasonality and geographic correlation in outcomes that are likely to bias the results when randomizing at the service area level. For example, given the variation, and seasonal nature of farming patterns at highly localized levels, it is important to ensure that the differences between treatment and control are minimized as much as possible, in order to prevent bias in the results. The solution to these problems was to use Edmond's algorithm for minimum distance matching to construct pairs of service areas. This matching for treatment and control allowed the study to overcome issues in seasonality and geographic correlation in outcomes by minimizing differences between paired branches[16]. It also improved balance across treatment and control villages on observed and unobserved factors, and provided a strong service-area-level control variable. For several 2001 census village outcomes (including caste composition, number of primary schools, water facilities and proportion of irrigated land), we find that controlling for pair fixed effects explains roughly 70% of the variance.[17]

V. **Randomization of access**: One service area in each pair was then randomly selected to receive a bank branch first (treatment area). Once the assignment was complete, the bank infrastructure staff attempted to locate premises on the designated and agreed branch site, succeeding about 90% of the time. When suitable premises were not available, the infrastructure staff searched the service area for a nearest substitute and proposed an alternate location to the originally nominated site. The research team employed a detailed system of checks to ensure that such changes did no comprise the design of the study, or the integrity of the randomization. Following the opening of the branch in the treated area, expansion in the other area (control) was delayed for 36 months. Bank employees were not informed about the study or whether their branch is a study branch or not. Treatment and control areas of the same pair were surveyed simultaneously. Surveyors were also not informed of the treatment status of villages, and were rotated across treatment and control.

Further details on randomization are provided in Appendix A.

5.2.3 Data

This report uses data from two different survey components (see Appendix B for details):
- *Household survey:* Several members of each randomly selected household are interviewed in each study village. Data collected includes sources and uses of income, including business and cultivation activities; financial literacy and activity in loans, savings, and insurance (formal and informal); health; well-being; social capital; female empowerment; and household structure. Importantly, data collected at the household level is aggregated at the village level to address the impact of financial access along dimensions such as wage rates.
  - *Social network mapping survey*: The full social network mapping (SNM) survey was administered in a subset of villages from control and treatment service areas. The sample was composed by 102 treatment villages and 102 control villages. Within a selected village, we asked all households to name their contacts within and outside of their village. The exhaustive census we collected at the village level prior to surveying enables us to map social connections within

---

[16] The match assignment exploits geographic autocorrelation to explain fixed and time-variant factors with a geographic component.
[17] See Section 5.2.4 Data Challenges for more details on the implementation of pair wise matching methods for confounding factors and cluster correlation.

each village. Information on outside contacts cannot be mapped since households can name households living in villages not included in our sample. The households were surveyed at baseline (prior the opening of the bank branch) and at endline (18 to 24 months after the opening of the branch).

5.2.4 Data Challenges

Our methodology allows us to address several common hindrances to causal attribution in microfinance evaluations:

i.   Confounding factors in the analysis have been addressed primarily through the use of pair-wise matching methods at the service area level. These methods ensure that, within the randomization, the influence of observable confounding factors balance across our treatment and control groups.

ii. Selection bias at the branch level has been addressed primarily by randomization. Further details can be found in Appendix A.

iii. Direct spillovers may have propagated from the treatment to the control, as control group residents sought financial services. However, because KGFS enforces strict residency norms for customers, and because CMF uses listings of control group residents to check for accidental enrollment by nearby branch area residents, direct spillovers have been limited in practice. Observed accidental enrollment was addressed early in the intervention, and has been accounted for in subsequent analysis.

iv. Indirect spillovers through social networks or through market channels were anticipated, and their measurement is central to our evaluation. Externalities and spillovers at the individual level are evaluated through the comparison of baseline to endline outcomes, as shown in this initial set of results on social networks. In future analysis, we will also study spillovers at the group level.[18]

v.   Contamination of the control group represents a challenge in the evaluations of microfinance program (Banerjee et. al., 2015). Thus, our analysis needs to take into account the penetration of potential competitors of KGFS. In addition, to limit expectation effects or pressure on KGFS and field staff, the intervention was double-blind: KGFS management and infrastructure staff were informed of the randomization scheme and of the evaluation design, but branch staff and local villagers were not.

vi. Unreliable survey responses were addressed in different ways, depending on the outcome of interest: particularly, where financial variables are concerned, we will also make extensive use of transactions data collected in real time by KGFS.[19]

vii. Cluster correlation is mitigated with an innovative pair-wise matching of treatment and control service areas. Proximate service areas often share access to facilities, have similar resource endowments, and can be expected to face common economic and political shocks. Pairing improves the precision of impact estimates by mitigating cluster correlation, at no added cost in data collection.

---

[18] In doing so, we also plan to employ a panel survey of financial welfare run jointly by Yale and CMF, to detect time trends in rural villages elsewhere in Tamil Nadu—and (cautiously) attribute residual changes in our sample to the KGFS intervention.
[19] The evaluators developed timely consistency checks and scrutiny processes to avoid this type of errors.

6. Impact results – Household Level

In what follows, we discuss the main first-stage results of the intervention. Our main specification models the effect of the randomized treatment, which consists in increased access to formal finance through the opening of a KGFS bank branch in a service area. We therefore estimate the following model using endline data:

$$Y_{ik} = \alpha_0 + \alpha_1\, T_k + \delta_{pk} + \varepsilon_{ik}$$

Where *i* indexes the individual or household and *k* indexes her service area. Y_{ik} is a given outcome (e.g. extensive and intensive margin of borrowing from formal/informal sources; extensive and intensive margin of savings; income and employment) for individual or household *i* in service area *k*. T_k is the service area treatment dummy, such that α_1 gives the Intent-To-Treat (ITT) effect. δ_{pk} are pair fixed effects[20] and ε_{ik} is the idiosyncratic error term. We cluster standard errors at the level of randomization, i.e. at the service area level.

In a second specification, we also include a vector X_{ik} of household-level controls, measured at baseline: distance to the nearest bank branch, age of the household's head, years of education of the household's head, caste, religion, and land ownership.

We refer to Appendix D – Methods for details on the specification used. Appendix E provides a comprehensive description of the variables included in the analysis. Additional tables are shown in Appendix F.[21]

6.1 FIRST STAGE EFFECTS: IMPACT ON BORROWING, SAVING AND TRANSFERS WITHIN SOCIAL NETWORK

Figure 2 looks at households' probability of having either outstanding formal (Figure 2 left) or informal loans (Figure 2 right).[22] It compares treated and control households, both at baseline and at endline. Interestingly, the two sub-figures are symmetric but opposite signed: on the one hand, the probability of taking out formal outstanding loans increases from baseline to endline (Figure 2 left). On the other hand, the probability of taking out informal outstanding loans decreases from baseline to endline, as if compensating for the increase in formal loans (Figure 2 right).

---

[20] These are added to take into account the use of pair-wise matching described earlier in assigning randomization.

[21] Since baseline checks show that there are no imbalances across treatment and control group in the main outcome variables we analyze, we report here estimates for endline only. We must notice that including baseline variables appear to increase noise in the estimates, potentially because of different ways subjects report information across survey rounds. Therefore, Difference-in-Difference estimates are not shown here. Still, they are available upon request.

[22] We classify as formal sources of borrowing private banks, NGOs/MFIs, Nationalised Banks, PACs/Cooperative Banks and Non-Banking Financial Corporations. Conversely, friends, neighbors and relatives; shopkeepers; employers; moneylenders; pawnbrokers; SHGs; landlords; ROSCAs; chitfunds; financiers; religious trust are classified as informal sources of borrowing.

Figure 2: Probability of Borrowing from Formal and Informal Sources
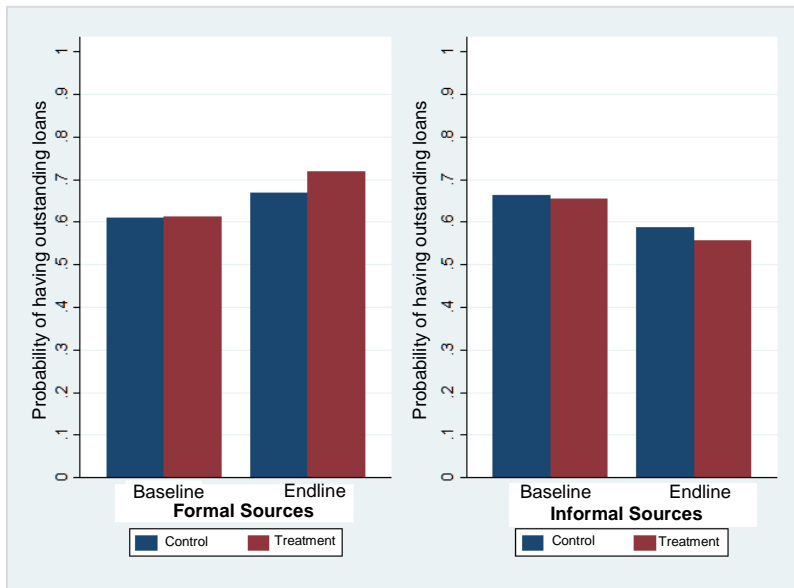
We also focus on the differences between treated and control households. At endline, the likelihood of borrowing from formal sources is higher for treated than for control households (Figure 2 left). Conversely, treated households appear less likely to borrow from informal sources than control (Figure 2 right). All in all, results from Figure 2 are suggestive of a shift from informal to formal sources of borrowing in treated service areas as a result of the expansion of KGFS.

We test these results more formally by estimating first stage impact on formal and informal borrowing. Results are shown in Table 1: Panel A and Panel B display estimates without and with household controls[23], respectively, for endline only.[24] Households in treated service areas are 5 p.p. more likely to report formal outstanding loans (Column 1, Panel A). At the same time, Column 2, Panel A shows that treated households are 4 p.p. less likely to have outstanding informal loans in the same period. Both coefficients are statistically different from zero.

Taken together, these results suggest that the presence of KGFS in a service area positively affects households' likelihood of having access to formal financial services, while it negatively affects their reliance on informal lenders.

Our intervention appears to have a differential impact across treatment and control groups not only at the extensive margin of borrowing (probability of borrowing from formal and informal sources), but also at the intensive margin (amount borrowed from formal and informal sources). Column 3 of Panel A, Table 1 shows that, at endline, the number of formal loans borrowed by treated households is 14% larger than in the control group; conversely, the number of informal loans by treated households at endline is 10% smaller than in the control group (Column 4, Panel A). These results can also be seen in Figure 3 below.

---

[23] Controls in Panel B include: age of the household's head, education (in years) of the household's head, caste, religion, distance to branch, land ownership. Further details can be found in the Pre-Analysis Plan, Appendix D.

[24] As we mentioned in footnote 22, tables do not include Diff-in-Diff estimates. Yet, these results are available upon request.

Figure 3: Number of Loans Borrowed From Formal and Informal Sources
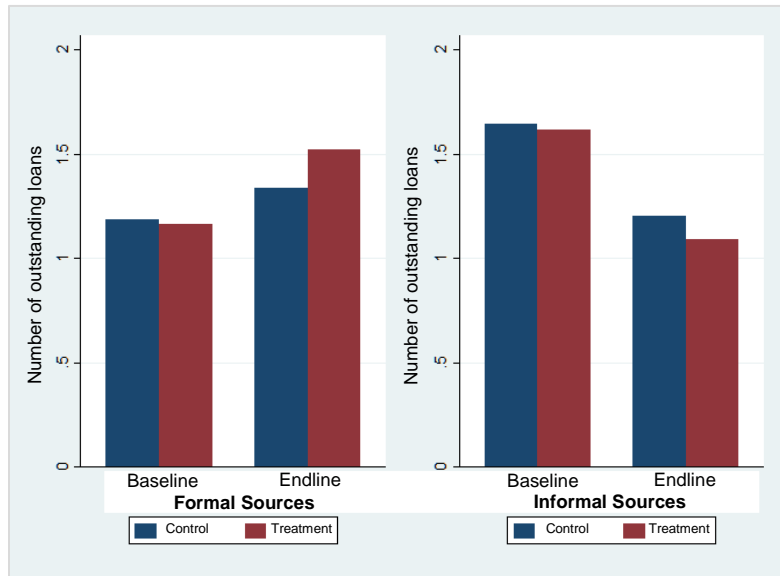
We also look at the total amount households borrowed in the last 24 months.[25] In line with previous results, Figure 4 shows that treated households have taken out a larger amount of debt from formal sources than households in control group (Figure 4 left). On the contrary, the reliance on informal credit appears lower for treated households than for control households (Figure 4 right). Results from Figure 4 are also confirmed by Columns 5 and 6 of Table 1: at endline, treated subjects borrow on average Rs 7,038.67 more than the control group (12% more) from formal sources (Panel A), while they borrow on average Rs 4,308.81 less than the control group (11% less) from informal sources (Panel A).

Finally, we study whether there is a statistically significant shift from informal to formal credit, as our results so far suggest. Column 7 of Panel A, Table 1 indicates that the amount of formal credit over total credit households borrow at endline is significantly larger for treated than for control households. This once again confirms that, in treated service areas, households have been more likely to substitute informal credit with formal credit.

Results shown in Panel A of Table 1 are also robust after controlling for household characteristics, as displayed in Panel B of the same table.

---

[25] This is computed as the sum of the principal amounts of all loans that were taken in the previous 24 months, whether they were still outstanding at the moment of the survey or had been repaid in the previous 12 months.

Table 1:First-stage Effects on Formal and Informal Loans

	P(Formal Loan Outstnd)	P(Informal Loan Outstnd)	Nr of Outstnd Formal loans	Nr of Outstnd Informal loans	Formal Borrowed Amt	Informal Borrowed Amt	Share of Formal Borrowed Amt
Panel A: Without Controls							
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Treated	0.05	-0.04	0.19	-0.15	7038.7	-4308.81	0.07
	(0.01)***	(0.01)***	(0.04)***	(0.04)***	(2262.3) **	(1667.9)*	(0.01) ***
Control Dep Var Mean	0.67	0.62	1.34	1.45	56443.38	39770.96	0.54
N	4158	4158	4156	4158	4156	4156	3645
Panel B: With Controls							
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Treated	0.05	-0.04	0.20	-0.15	7206.5	-4080.5	0.07
	(0.01)***	(0.01)***	(0.03)***	(0.04)***	(2238.9)**	(1673.0)*	(0.01)***
Control Dep Var Mean	0.67	0.62	1.34	1.45	56443.38	39770.96	0.54
N	4158	4158	4156	4158	4156	4156	3645

Note: ***, ** and * indicate significance at the 1%, 5%, and 10% levels respectively. Panel A reports the OLS coefficient estimates (standard errors) associated with regressing each column heading dependent variable on the treatment dummy Treated, using endline data only. Panel B reports the OLS coefficient estimates (standard errors) associated with regressing each column heading dependent variable on the treatment dummy Treated, using endline data only and controls at the household level. Household level controls are: age of the household's head, education (in years) of the household's head, caste, religion, distance to branch, land ownership. All regressions include pair fixed effects and survey round fixed effects (three rounds at endline). Standard errors are clustered at the service area level. All Rs. amounts are top-coded at 3 standard deviations. Refer to the appendix for variable definitions.

Figure 4: Total Borrowed Amounts from Formal and Informal Sources
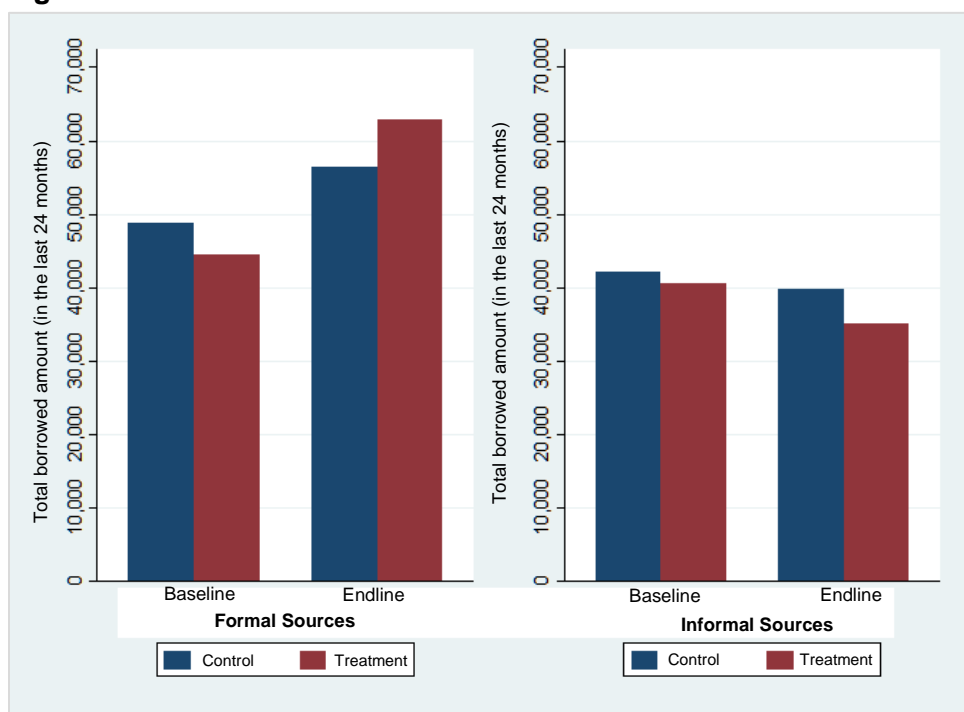
Table 2 complements results from Table 1 by looking at whether households substituted informal credit with formal credit for specific purposes/loan usage. Panel 1A and Panel 1B of Table 2 focus on informal loans. They show that treated households, compared to the control group, are less likely to borrow from informal sources for house repairs, for weddings and ceremonies, and for education-related expenses. The magnitude of this reduction is quite large, at 12%, 29% and 32% for house repairs, weddings and education respectively (Panel 1A). Panel 2A and Panel 2B of Table 2 focus on formal borrowing. Compared to control households, and in line with the Theory of Change outlined in Section 4.1, treated households are more likely to borrow from formal sources for farming and business investment (Column 1) and health-related expenses (Column 6). Results are robust after including household controls in Panel 1B.

The three most frequent reasons for borrowing from formal sources[26] are to make upgrades or repair houses, land or buildings (27% of the sample); to purchase day-to-day items for the household (20%); and to invest in farming and business (19%). While both house repairs and upgrades and daily purchases are also reported as main reasons to borrow from informal sources (by 17% and 24% of the sample, respectively), the third most frequent reason households report borrowing from informal sources[27] are weddings (15%).

---

[26] Tables not shown
[27] Tables not shown

Table 2: First-stage Effect on Formal and Informal Borrowed Amounts

	Farming & Business Investment	House and Land Repair	Weddings	Day-to-Day Expenses	Education Related	Health Related
Panel 1A: Informal Borrowing Amounts, without Household Controls						
	(1)	(2)	(3)	(4)	(5)	(6)
Treated	-430.47	-1262.50	-2162.50	12.08	-714.5	62.18
	(520.54)	(630.19)**	(565.22)***	(102.32)	(233.71)***	(151.14)
N	4160	4160	4160	4160	4160	4160
Control Mean	4992.650	10435.220	7539.290	2206.490	2236.210	2296.140
Panel 1B: Informal Borrowing Amounts, with Household Controls						
	(1)	(2)	(3)	(4)	(5)	(6)
Treated	-486.36	-1089.8	-2300.8	28.38	-702.4	73.33
	(513.2)	(610.5)*	(587.1)***	(103.8)	(240.2)***	(149.90)
N	4160	4160	4160	4160	4160	4160
Control Mean	4992.650	10435.220	7539.290	2206.490	2236.210	2296.140
Panel 2A: Formal Borrowing Amounts, without Household Controls						
	(1)	(2)	(3)	(4)	(5)	(6)
Treated	3110.14	949.94	-364.90	145.58	-88.90	397.41
	(1383.64)**	(954.25)	(492.43)	(195.17)	(313.79)	(205.82)*
N	4160	4160	4160	4160	4160	4160
Control Mean	13354.840	17415.210	4855.450	3546.830	3682.720	1769.300
Panel 2B: Formal Borrowing Amounts, with Household Controls						
	(1)	(2)	(3)	(4)	(5)	(6)
Treated	3232.5	1058.2	-334.6	112.9	-71.35	364.8
	(1360.8)**	(983.8)	(512.7)	(193.3)	(304.3)	(207.3)*
N	4160	4160	4160	4160	4160	4160
Control Mean	13354.840	17415.210	4855.450	3546.830	3682.720	1769.300

Note: ***, ** and * indicate significance at the 1%, 5%, and 10% levels respectively. Panel A reports the OLS coefficient estimates (standard errors) associated with regressing each column heading dependent variable on the treatment dummy Treated, using endline data only. Panel B reports the OLS coefficient estimates (standard errors) associated with regressing each column heading dependent variable on the treatment dummy Treated, using endline data only and controls at the household level. Household level controls are: age of the household's head, education (in years) of the household's head, caste, religion, distance to branch, land ownership. All regressions include pair fixed effects and survey round fixed effects (three rounds at endline). Standard errors are clustered at the service area level. All Rs. amounts are top-coded at 3 standard deviations. Refer to the appendix for variable definitions.

Table 3 complements the results shown in Table 2 by studying treatment effects on total borrowing amounts. Findings from Table 2 could in fact be explained by treated households resorting to cheaper – and larger – loans to a higher extent for at least some of these types of expenditures. Hence, one needs to understand if the trends observed for either formal or informal loans are not driven by a general trend in aggregated loans. A way of testing this hypothesis is to precisely look

at households' total borrowing – from formal and informal sources – across these loan usage categories. For instance, we do not find an increase in total borrowing for health-related expenses (Column 6), despite an increase in formal borrowing for this loan usage category in Table 2. This suggests that households use formal credit to higher extent for health-related expenses (Column 6). On the contrary, households seem to reduce total borrowing for wedding purposes (Column 3). We also find suggestive evidence that households increase total borrowing in farming and business investments (Column 1, Panel B).

All in all, results so far confirm our theory of change: by penetrating into rural areas, KGFS should crowd-out loans from informal lenders and informal transfers among social networks. In particular, among informal lenders, moneylenders and financiers should be the most negatively affected by KGFS, as they usually are the most active lenders in rural villages before the expansion of formal financial services providers.

Table 3: First-stage Effect on Total Borrowed Amounts

	Farming & Business Investment	House and Land Repair	Weddings	Day-to-Day Expenses	Education Related	Health Related
Panel A: Without Controls						
	(1)	(2)	(3)	(4)	(5)	(6)
Treated	2789.6	144.14	-2619.8	89.12	-645.6	194.85
	(1714.0)	(1229.19)	(851.09)***	(262.20)	(507.59)	(336.74)
Control Dep Var Mean	18992.50	28683.37	13195.51	6010.77	6324.86	4646.70
N	4160	4160	4160	4160	4160	4160
Panel B: With Controls						
	(1)	(2)	(3)	(4)	(5)	(6)
Treated	2871.8	408.6	-2711.3	78.21	-598.6	169.2
	(1670.4)*	(1273.7)	(867.0)***	(261.3)	(503.8)	(339.94)
Control Dep Var Mean	18992.50	28683.37	13195.51	6010.77	6324.86	4646.70
N	4160	4160	4160	4160	4160	4160

Note: ***, ** and * indicate significance at the 1%, 5%, and 10% levels respectively. Panel A reports the OLS coefficient estimates (standard errors) associated with regressing each column heading dependent variable on the treatment dummy Treated, using endline data only. Panel B reports the OLS coefficient estimates (standard errors) associated with regressing each column heading dependent variable on the treatment dummy Treated, using endline data only and controls at the household level. Household level controls are: age of the household's head, education (in years) of the household's head, caste, religion, distance to branch, land ownership. All regressions include pair fixed effects and survey round fixed effects (three rounds at endline). Standard errors are clustered at the service area level. All Rs. amounts are top-coded at 3 standard deviations. Refer to the appendix for variable definitions.

We test this hypothesis in greater detail in Table 4 (changes in households' reliance on moneylenders and financiers), and Table 5 and 6 (changes in households' reliance on informal transfers, either inside or outside the village). Column 1 of Table 4 shows treatment effects for the extensive margin of borrowing from moneylenders and financiers. Taken together, these two categories account for about one third of the informal loans households took out at baseline.[28] In line with our predictions, we find that treated households are 4 p.p. less likely than control to take out loans from moneylenders and financiers at endline. Similar treatment effects can be found at the extensive margin of borrowing: the number of outstanding loans from moneylenders and financiers is 14% lower in treatment than control group (Column 2, Panel A); similarly, the total amount borrowed from these two informal lenders is 12% lower for treated households at endline (Column 2, Panel A). Similar results are found when we include households' controls, in Panel B.

Table 4. First-stage Effect on Moneylender and Financier Loans

	P(Moneyl. And Fin. Loan Outstnd)	Nr of Moneyl. And Fin. Outstnd loans	Moneyl. And Fin. Borrowed Amt
Panel A: Without Household Controls			
	(1)	(2)	(3)
Treated	-0.04	-0.09	-2369.5
	(0.01)***	(0.03)**	(1518.5)*
N	4158	4158	4158
Control Mean	0.360	0.630	18380.380
Panel B: Household Controls			
	(1)	(2)	(3)
Treated	-0.04	-0.09	-2185.3
	(0.01)***	(0.04)**	(1504.6)
N	4158	4158	4158
Control Mean	0.360	0.630	18380.380

Note: ***, ** and * indicate significance at the 1%, 5%, and 10% levels respectively. Panel A reports the OLS coefficient estimates (standard errors) associated with regressing each column heading dependent variable on the treatment dummy Treated, using endline data only. Panel B reports the OLS coefficient estimates (standard errors) associated with regressing each column heading dependent variable on the treatment dummy Treated, using endline data only and controls at the household level. Household level controls are: age of the household's head, education (in years) of the household's head, caste, religion, distance to branch, land ownership. All regressions include pair fixed effects and survey round fixed effects (three rounds at endline). Standard errors are clustered at the service area level. All Rs. amounts are top-coded at 3 standard deviations. Refer to the appendix for variable definitions.

---

[28] The sum of the share of loans borrowed from moneylenders and financiers on the total of informal loans is 33.4%. Loans from friends, neighbours and relatives represent 35.6% of the total informal loans.

As already mentioned, expanding formal financial access should also have an effect on informal loans within households' social networks. In line with this hypothesis, we test the impact of KGFS expansion on households' likelihood to borrow from contacts that live either inside or outside the village. Table 5, Panel 1 displays first-stage results for inside-village contacts, excluding moneylenders: compared to control, treated households rely on a significantly smaller number of inside contacts (-5%); in addition, households in treated service areas appear less likely to be able to rely on inside contacts for business purposes (-13%). Moreover, they actually borrow a significantly smaller amount of credit from inside contacts (-13%), showing a lower reliance on informal transfers. This result is particularly relevant as one of the objectives of this study was to precisely measure indirect spill-over effects resulting from the expansion of KGFS.

Table 5 Panel 2 replicates the same analysis as Panel 1, but for contacts outside the village. Similar to the previous case, we observe that treated households borrow significantly less from outside contacts (-12.5%).

Table 6 adds to Tables 4 and 5 as it focuses on households' borrowing from moneylenders. Panel 1 (top-coded values) shows in particular that treated households are less likely to be able to resort to moneylenders both for emergency and business purposes. In addition, at endline they report borrowing 11% less credit from moneylenders than households in control service areas. Results from Panel 2 of Table 6 (non-topcoded values) confirm results shown in Panel 1.

Finally, we look at the impact of KGFS expansion on households' ability to save, both formally and informally. First-stage results are shown in Table 7 (without and with controls, in Panel A and Panel B, respectively). At endline, treated households save significantly larger amounts in their saving account than control (+22%), as shown in Column 2, Panel A. This can also be seen from Figure 5.[29] We also look at treatment effects both at the intensive and at the extensive margin of informal loans given out by the household (Column 3 and 4, Panel A): treated households are significantly more likely to give out loans, and they also give out a significantly larger amount of informal loans than control households, at endline. Results shown in Panel A are again robust after controlling for household characteristics (Panel B).

---

[29] We notice that we find an effect at the intensive margin of saving but not at the extensive margin. This absence of the latter is mainly due to the large initial penetration level of savings, with 84% of the households declaring having a saving account at baseline. It is important to notice that KGFS does not take saving deposits directly. In fact, KGFS has a partnership with a formal financial institution, a commercial bank, in order to collect saving deposits. Moreover, KGFS strongly emphasizes the importance of saving to its customers, notably through the well-being report produced for each client. Hence, the positive treatment effect we find on saving amounts can be reasonably attributed to the expansion of KGFS.

Figure 5: Total Saved Amounts in Any Account
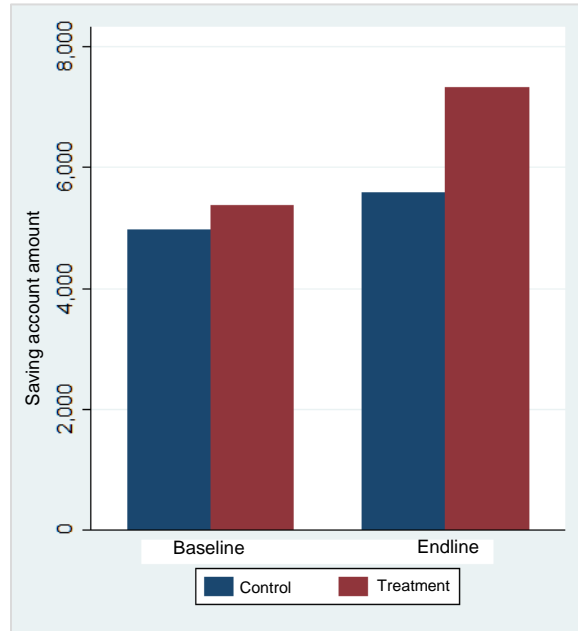
Table 5. First-stage Effect on Inside and Outside Village Contacts

	Any Contact	Nr. Contacts	Emergency Borrowing Capacity	Business Borrowing Capacity	Actual Borrowed Amt
Panel 1A: Inside Village Contacts, without Household Controls					
	(1)	(2)	(3)	(4)	(5)
Treatment	0.00	-0.14	-4161.67	-5007.40	-1345.68
	(0.01)	(0.07)**	(2540.21)	(2966.36)*	(716.49)*
N	19183	19183	19183	19183	19183
Control Mean	0.890	3.010	31628.230	36999.360	10640.930
Panel 1B: Inside Village Contacts, with Household Controls					
	(1)	(2)	(3)	(4)	(5)
Treatment	-0.00	-0.13	-4213.21	-5111.35	-1257.64
	(0.01)	(0.07)*	(2878.36)	(3348.93)	(817.36)
N	19183	19183	19183	19183	19183
Control Mean	0.890	3.010	31628.230	36999.360	10640.930
Panel 2A: Outside Village Contacts, without Household Controls					
	(1)	(2)	(3)	(4)	(5)
Treatment	0.00	-0.01	-2364.22	-2919.26	-1787.65
	(0.01)	(0.03)	(2393.35)	(2482.53)	(815.68)**
N	19183	19183	19183	19183	19183
Control Mean	0.520	0.910	36834.360	41323.780	14331.680
Panel 2B: Outside Village Contacts, with Household Controls					
	(1)	(2)	(3)	(4)	(5)
Treatment	-0.01	-0.01	-2547.06	-3358.01	-1792.29
	(0.01)	(0.03)	(2772.87)	(2893.89)	(892.60)**
N	19183	19183	19183	19183	19183
Control Mean	0.520	0.910	36834.360	41323.780	14331.680

Note: ***, ** and * indicate significance at the 1%, 5%, and 10% levels respectively. Panel 1 and 2 report the OLS coefficient estimates (standard errors) associated with regressing each column heading dependent variable on the treatment dummy Treated. Panel 1B and Panel 2B include household controls, which are: distance to the nearest branch, age of the household's head, years of education of household's head, caste, religion and land ownership. All regressions include pair fixed effects and round specific fixed effects. Standard errors are clustered at the service area level. All Rs. amounts are top-coded at 3 standard deviations. Refer to the appendix for variable descriptions.

Table 6. First-Stage Effect at the Household Level on Moneylender Contacts

	Any contact	Nr. Contacts	Emergency Borrowing Capacity	Business Borrowing Capacity	Actual Browwed Amt
Panel 1A: Money Lender Contacts Top-Coded, without Household Controls					
	(1)	(2)	(3)	(4)	(5)
Treatment	-0.01	-0.02	-1353.86	-1375.01	-566.80
	(0.02)	(0.03)	(741.66)*	(783.60)*	(391.71)
N	19183	19183	19183	19183	19183
Control Mean	0.270	0.430	10277.810	10664.730	5316.670
Panel 1B: Money Lender Contacts Top-Coded, with Household Controls					
	(1)	(2)	(3)	(4)	(5)
Treatment	-0.01	-0.02	-1464.23	-1475.63	-628.37
	(0.02)	(0.03)	(763.64)*	(820.66)*	(397.85)
N	19183	19183	19183	19183	19183
Control Mean	0.270	0.430	10277.810	10664.730	5316.670
Panel 2A: Money Lender Contacts ot Top-Coded, without Household Controls					
	(1)	(2)	(3)	(4)	(5)
Treatment	-0.01	-0.02	-2834.62	-3025.02	-1409.18
	(0.02)	(0.03)	(1000.47)***	(1025.07)***	(583.19)**
N	19183	19183	19183	19183	19183
Control Mean	0.270	0.430	12966.560	13619.220	6931.490
Panel 2B: Money Lender Contacts Not Top-Coded,with Household Controls					
	(1)	(2)	(3)	(4)	(5)
Treatment	-0.01	-0.02	-3049.56	-3359.02	-1532.63
	(0.02)	(0.03)	(1054.11)***	(1074.84)***	(608.09)**
N	19183	19183	19183	19183	19183
Control Mean	0.270	0.430	12966.560	13619.220	6931.490

Note: ***, ** and * indicate significance at the 1%, 5%, and 10% levels respectively. Panel 1 and 2 report the OLS coefficient estimates (standard errors) associated with regressing each column heading dependent variable on the treatment dummy Treated. Panel 1B and Panel 2B include household controls, which are: age of the household's head, education (in years) of the household's head, caste, religion, distance to the nearest branch, and land ownership. All regressions include pair fixed effects and round specific fixed effects. Standard errors are clustered at the service area level. All Rs. amounts are top-coded at 3 standard deviations. Refer to the appendix for variable definitions.

Table 7. First-Stage Effect on Formal and Informal Savings

	P(Any Saving Account)	Saving Account Amount	P(Giving out loans)	Nr of Informal loans given out	Informal loans Given Out Amount	P(Active Insurance)
	(1)	(2)	(3)	(4)	(5)	(6)
Treated	-0.003	1181.174	0.006	0.011	215.579	-0.002
	(0.005)	(592.741)**	(0.003)**	(0.005)**	(246.735)	(0.009)
Control Dep Var Mean	0.950	5442.870	0.010	0.020	522.080	0.800
N	4160	4159	4158	4158	4158	4160
	(1)	(2)	(4)	(5)	(6)	(7)
Treated	-0.003	1393.2	0.006	0.010	160.8	-0.002
	(0.005)	(687.535)**	(0.003)**	(0.005)**	(243.750)	(0.009)
Control Dep Var Mean	0.950	5442.870	0.010	0.020	522.080	0.800
N	4160	4159	4158	4158	4158	4160

Note: ***, ** and * indicate significance at the 1%, 5%, and 10% levels respectively. Panel A reports the OLS coefficient estimates (standard errors) associated with regressing each column heading dependent variable on the treatment dummy Treated, using endline data only. Panel B reports the OLS coefficient estimates (standard errors) associated with regressing each column heading dependent variable on the treatment dummy Treated, using endline data only and controls at the household level. Household level controls are: distance to the nearest bank branch, age of household's head, years of education of household's head, caste, religion and land ownership. All regressions include pair fixed effects and survey round fixed effects (three rounds at endline). Standard errors are clustered at the service area level. All Rs. amounts are top-coded at 3 standard deviations. Refer to the appendix for variable definitions.

6.2 REAL ECONOMY: EFFECTS ON INCOME AND STRUCTURAL CHANGES

The theory of change behind the studied intervention is that increased formal financial access has a positive impact on households' welfare. One of the channels is that formal financial access should push households towards more entrepreneurial (characterised by higher risk, but also higher return) activities. If this prediction is confirmed, we should observe treated households being more likely to be engaged in entrepreneurial activities at endline. In a similar spirit, we should expect an increase in the income of treated *vis-à-vis* control households, particularly deriving from business activities. Table 8, Panel A reports treatment effects across a number of income and employment variables, which include wage labour income (we distinguish between governmental programs, namely NREGA, and non-governmental wage labour), business income, and the household's self-reported income – which is also used to assess whether the household lives below the poverty line.

We also include an indicator of whether the household is self-employed, as well as number of employees and business investment.[30] In line with our hypotheses, treated households are 2 p.p. more likely than control to be self-employed at endline (Column 2). The negative coefficient shown in Column 1, given its weak significance and the definition of the outcome variable (see Appendix E), may also suggest that treated households are changing the nature of their involvement in agriculture. At the same time, Column 3 shows that, at endline, treated households are significantly more likely to hire employees for their business than control (+17%). Similarly, we find both business income and business investment to be about 20% higher for treated households compared to control, in Column 8 and 9, respectively.

Consistent with this increase in business income and investment, we find that household income increases by 10% in treatment compared to control group (Column 4). Moreover, the share of households living below the poverty line decreases by 2 p.p., in treatment versus control group. Results from Panel B confirm findings from Panel A.

---

[30] The "self-employment" indicator excludes farming activity. Business investment (expressed in log) represents the total of the value of the equipment they purchased and the cost of maintenance/repair for property & equipment in the past 12 months (where property and equipment is used for business activity only).

Table 8. Impact on Employment and Income Composition

	P(Farming)	P(Self-Employed)	P(Hires Employees)	Log Household Income (30d), top coded	P(Below Poverty Line	Log Governmental Wage Labour Income, top coded	Log Non-Governmental Wage labor Income, top coded	Log Business Income (30d), top coded	Log business Investment (12m), top coded
Panel A: Without Controls									
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Treated	-0.016	0.0215	0.0114	0.103	-0.0198	-0.0116	0.0312	0.198	0.181
	(0.017)	(0.008)**	(0.005)**	(0.053)*	(0.010)*	(0.060)	(0.065)	(0.077)**	(0.054)***
Control Dep Var Mean	0.440	0.140	0.070	8.160	0.310	0.640	1.660	1.270	0.760
N	4157	4160	4160	4158	4160	4160	4160	4160	4160
Panel B: With Controls									
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Treated	-0.020	0.022	0.011	0.0987	-0.0201	-0.029	0.051	0.208	0.194
	(0.0116)*	(0.008)***	(0.005)**	(0.051)*	(0.010)*	(0.061)	(0.064)	(0.075)***	(0.054)***
Control Dep Var Mean	0.44	0.140	0.070	7.920	0.310	0.640	1.660	1.270	0.760
N	4157	4160	4160	4158	4160	4160	4160	4160	4160

Note: ***, ** and * indicate significance at the 1%, 5%, and 10% levels respectively. For each heading dependent variable, we report the OLS coefficient estimate (standard errors) associated with regressing the dependent variable on the treatment dummy Treated, using endline data only. Panel A reports the OLS coefficient estimates (standard errors) associated with regressing each column heading dependent variable on the treatment dummy Treated, using endline data only. Panel B reports the OLS coefficient estimates (standard errors) associated with regressing each column heading dependent variable on the treatment dummy Treated, using endline data only and controls at the household level. Household level controls are: distance to the nearest bank branch, age of household's head, years of education of household's head, caste, religion and land ownership. All regressions include pair fixed effects and survey round fixed effects (3 rounds at endline). Standard errors are clustered at the service area level. All Rs. amounts are top-coded at 3 standard deviations. Refer to the appendix for variable definitions.

All in all, estimates shown in Table 8 convincingly show that access to formal financial services and products led to a positive, structural change on households' employment, thus generating an increase in business income and overall household income – point estimates remain unchanged, both in significance and magnitude of coefficients, also after including household controls.

Table 9. Changes in Wages from Non-Household Employment

	Total daily wage, per daily payment cycle	Total monthly wage, per monthly payment cycle
Panel A: Without Controls		
	(1)	(2)
Treated	10.39	-13.01
	(4.17)**	(56.89)
Control Dep Var Mean	126.90	950.70
Control % Earns Daily Wage	0.50	0.50
Control % Earns Monthly Wage	0.16	0.16
Control % Earns Wage	0.61	0.61
N	4157	4156
Panel B: With Controls		
	(2)	(4)
Treated	9.52	-16.80
	(4.03)**	(54.94)
Control Dep Var Mean	126.67	950.70
Control % Earns Daily Wage	0.50	0.50
Control % Earns Monthly Wage	0.16	0.16
Control % Earns Wage	0.61	0.61
N	4157	4156

Note: ***, ** and * indicate significance at the 1%, 5%, and 10% levels respectively. For each heading dependent variable, we report the OLS coefficient estimate (standard errors) associated with regressing the dependent variable on the treatment dummy Treated, using endline data only. Panel A reports the OLS coefficient estimates (standard errors) associated with regressing each column heading dependent variable on the treatment dummy Treated, using endline data only. Panel B reports the OLS coefficient estimates (standard errors) associated with regressing each column heading dependent variable on the treatment dummy Treated, using endline data only and controls at the household level. Household level controls are: distance to the nearest bank branch, age of household's head, years of education of household's head, caste, religion, and land ownership. All regressions include pair fixed effects and survey round fixed effects (3 rounds at endline). Standard errors are clustered at the service area level. Refer to the appendix for variable definitions.

Finally, we look at whether our intervention also had an impact on daily wages in our study areas. The underlying intuition is that the shift towards self-employment activities observed in Table 8, should have affected households' labour supply in other activities (i.e. wage labour and farming), and, therefore, daily wages. Consistent with this hypothesis, we find evidence in Table 9 that daily wages significantly increased for treatment households as compared to control. It may therefore be said that Table 9 provides the first suggestive evidence of the presence of general equilibrium effects, and positive externalities, in addition to the effects at the household level. Further analysis is needed in this direction to consolidate our results

6.3 SHOCKS

We also study the impact of KGFS on treated households' ability to smooth consumption. Our theory of change predicts that increased access to formal financial services provides households with better means to borrow from formal sources in times of emergency, and to support themselves during the lean season (for farmers) or periods of slow employment. On the surface, we can explore this most directly by examining how the treatment affects borrowing behaviour in situations where a household is suddenly in financial need – or, in other words, when a household experiences an income shock.

Respondents were asked questions on shocks their household experienced in the last 12 months that directly impacted their income, farming, and health, and that were likely to indirectly affect other factors as well. We collected data on the following shocks: loss of wage employment, deaths of household money earners, serious illnesses or injuries, and severe weather conditions (floods, droughts, crop diseases, etc.). The most commonly reported shock was illness, with 21% of baseline respondents reporting having a household member experience a serious illness or injury in the last 12 months that prevented the affected person from performing normal activities. Thus, in Table 10, we examine treatment effects on the same formal and informal borrowing outcomes as in Table 1 for households that experienced illness shocks.

Panel A of Table 10 shows the correlation between illness shock occurrence and borrowing behaviour over endline respondents only. Experiencing an illness shock is positively associated with a household's likelihood of borrowing at all; their number of outstanding loans; and their borrowed amounts, from both formal and informal sources. This illustrates what we would expect, especially if households must cover the cost of the injury or illness in full, or if they lose wage employment income if the ill household member is an income earner. This, combined with our theory of change, suggests that households that still borrow following illness shocks should be able to borrow less and substitute away from informal borrowing.

Given that illness shocks are associated with more borrowing in general, Panels B and C (the latter with household controls) show whether the treatment changes these patterns for households that experienced both the intervention treatment and an illness shock (*Treated * Shock)*. Results suggest that treated households borrow significantly fewer loans from informal sources (Column 4).

Combined, our results suggest that, even in the wake of an adverse shock, treated households are able to rely less on informal sources, borrow less than their control group counterparts do, and still cope financially.

Table 10. Changes in Borrowing Following Illness Shocks

	P(Formal Loan Outstnd)	P(Informal Loan Outstnd)	Nr of Outstnd Formal loans	Nr of Outstnd Informal loans	Formal Borrowed Amt	Informal Borrowed Amt	Formal Share of Borrowed Amt
Panel A: OLS, Shock							
Shock	0.067	0.078	0.138	0.310	8469.017	5902.653	-0.0086
	(0.015)***	(0.015)***	(0.052)***	(0.052)***	(3534.912)**	(1777.974)***	(0.013)
N	4158	4158	4156	4158	4156	4156	3645
Mean for people w\o shock	0.67	0.62	1.34	1.45	56443.4	39771.0	0.54
Panel B: OLS, Without Household Controls							
Treated x Shock	0.0179	0.0243	0.0307	-0.15	3807.327	-6944.673	0.0215
	(0.028)	(0.028)	(0.0997)	(0.104)	(6334.60)	(3397.210)**	(0.025)
Treated	0.041	-0.048	0.178	-0.0941	5453.712	-1651.861	0.0618
	(0.017)**	(0.0156)***	(0.0538)***	(0.0520)*	(3471.597)	(1926.513)	(0.013)***
Shock	0.057	0.066	0.121	0.388	6444.718	9489.767	-0.0207
	(0.022)***	(0.0165)***	(0.0736)	(0.0672)***	(4974.242)	(2325.523)***	(0.019)
N	4158	4158	4158	4158	4156	4156	3645
Control Mean	0.67	0.62	1.35	1.54	56443.38	39770.96	0.54
Panel C: OLS, With Household Controls							
Treated x Shock	0.00471	0.0127	-0.00901	-0.195	3713.3	-7549.8	0.0216
	(0.028)	(0.028)	(0.100)	(0.106)*	(6027.301)	(3420.091)**	(0.025)
Treated	0.0463	-0.0438	0.193	-0.0737	5663.9	-1181.3	0.0612
	(0.017)***	(0.016)***	(0.053)***	(0.058)	(3376.783)*	(1924.882)	(0.0135)***
Shock	0.0533	0.0663	0.102	0.392	5408.8	9087.6	-0.0208
	(0.021)**	(0.0165)***	(0.0724)	(0.068)***	(4905.057)	(2315.0)***	(0.018)
N	4158	4158	4158	4158	4156	4156	3645
Control Mean	0.67	0.62	1.37	1.54	56443.38	39770.96	0.54

Note: ***, ** and * indicate significance at the 1%, 5%, and 10% levels respectively. Shock is a dummy variable equal to 1 if a household has experienced an illness shock at endline, defined as a serious injury or illness that has prevented household members from participating in normal activities in the last 12 months. Panel A reports the OLS coefficient estimates (standard errors) associated with regressing each column heading dependent variable on the Shock dummy using endline data only. Panels B and C report the OLS coefficient estimates (standard errors) associated with regressing each column heading dependent variable on the treatment dummy Treated, using endline data only. Panel C regressions additionally include household level controls: distance to the nearest bank branch, the age, and years of education of the head of the household, caste, religion and landownership. All regressions include pair fixed effects and survey round fixed effects (3 rounds at endline). Standard errors are clustered at the service area level. All Rs. amounts are top-coded at 3 standard deviations. Refer to the appendix for variable definitions.

6.4 WELL-BEING

We finally look at whether expanding financial access has an effect on households' well-being. To this end, we elicit a set of measures of households' life perception (Table 11) and psychological distress (Table 12).[31] Both tables report results excluding and including household controls, in Panel A and B, respectively. We do not find significant impacts of financial access on households' happiness or future life perception. Yet, it seems to have an impact on how individuals see their past life. In a similar spirit, we do not find an impact of the provision of formal financial services on households' psychological distress. The absence of concrete findings here could be due to competing effects between our well-being measures that we have yet to identify.

Table 11. Impact on Life Perception

	Happiness Scale	Perception of Current Life	Better Perception of Life now than 5 years ago	Better Perception of Life in 5 years
Panel A: Without Controls				
	(1)	(2)	(3)	(4)
Treated	-0.001	-0.024	-0.027	-0.001
	(0.013)	(0.042)	(0.012)**	(0.005)
Control Dep Var Mean	2.84	4.750	0.530	0.940
N	4158	4156	4156	4147
Panel B: With Controls				
	(1)	(2)	(3)	(4)
Treated	-0.001	-0.0175	-0.0278	0.001
	(0.0125)	(0.0405)	(0.0114)**	(0.005)
Control Dep Var Mean	2.840	4.750	0.530	0.940
N	4158	4156	4156	4147

Note: ***, ** and * indicate significance at the 1%, 5%, and 10% levels respectively. Panel A reports the OLS coefficient estimates (standard errors) associated with regressing each column heading dependent variable on the treatment dummy Treated, using endline data only. Panel B reports the OLS coefficient estimates (standard errors) associated with regressing each column heading dependent variable on the treatment dummy Treated, using endline data only and controls at the household level. Household level controls are: distance to the nearest bank branch, age of household's head, years of education of household's head, caste, religion, and land ownership. All regressions include pair fixed effects and survey round fixed effects (three rounds at endline). Standard errors are clustered at the service area level. Refer to the appendix for variable definitions.

---

[31] Again, these variables are described in detail in Appendix E.

Table 12. Impact on Psychological Distress

	Self-Assessed Health	Felt Nervous	Felt Hopeless	Felt Restless	Felt Depressed	Felt everything was an effort	Felt worthless	K6 Scale
Panel A: Without Controls								
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Treated	-0.073	0.006	-0.036	-0.004	-0.032	0.034	0.005	-0.019
	(0.046)	(0.023)	(0.022)*	(0.020)	(0.028)	(0.029)	(0.021)	(0.081)
Control Dep Var Mean	6.25	2.730	2.340	2.640	2.560	2.930	1.780	14.980
N	3743	3742	3741	3739	3741	3742	3741	3736
Panel B: With Controls								
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Treated	-0.080	0.005	-0.033	-0.004	-0.029	0.038	0.039	0.011
	(0.047)*	(0.023)	(0.021)	(0.021)	(0.029)	(0.030)	(0.022)	(0.084)
Control Dep Var Mean	6.25	2.730	2.340	2.640	2.560	2.930	1.780	14.980
N	3743	3742	3741	3739	3741	3742	3741	3736

Note: ***, ** and * indicate significance at the 1%, 5%, and 10% levels respectively. Panel A reports the OLS coefficient estimates (standard errors) associated with regressing each column heading dependent variable on the treatment dummy Treated, using endline data only. Panel B reports the OLS coefficient estimates (standard errors) associated with regressing each column heading dependent variable on the treatment dummy Treated, using endline data only and controls at the household level. Household level controls are: distance to the nearest bank branch, age of household's head, years of education of household's head, caste, religion and land ownership. All regressions include pair fixed effects and survey round fixed effects (three rounds at endline). Standard errors are clustered at the service area level. Refer to the appendix for variable definitions.

7. Cost Effectiveness

Given the complex suite of products offered by KGFS, carrying out a comprehensive cost-effectiveness analysis of its model would require, at present, too many assumptions on take-up rates, prices and costs of the intervention. Still, because KGFS is a private entity, one useful metric we can look at to understand if KGFS' expansion has been cost effective is the institution's profitability.

The KGFS model falls under Pudhuaaru Financial Services Private Limited (PFSPL) which is a wholly owned subsidiary under the IFMR RURAL CHANNELS AND SERVICES PRIVATE LIMITED (IRCS).[32] Based on IRCS's latest annual report (2015-16) available online, the company grew its business in almost all its units profitably, except for a few branches located in the North and East India, however mostly because of extraneous reasons. Although such branches negatively impacted the overall profitability of IRCS in the current financial year, the financial statement emphasizes the exceptionality of such

---

32 IRCS and PFSPL is a for-profit entity and governed under Section 134 (5) of the Companies Act, 2013.

event. Moreover, we must highlight that the KGFS branches pertaining to our study, which are located in the state of Tamil Nadu, South India, recorded profits, instead. In addition, according to same report, IRCS (and KGFS) is expected to grow in a sustainable manner in the next financial year. Such positive signs of profitability are also confirmed by the fact that the institution is also considering opening new branches.

In light of this, we can conclude that KGFS's expansion looks promising in terms profitability and sustainability. As more information will become available we plan to refine our cost-effectiveness analysis to make it more granular, both product- and geography-wise.

8. CONCLUSIONS

In this report, we present the results from an innovative Randomized Controlled Trial that is used to study the effects of the expansion of a rural branch banking model in Tamil Nadu, India. The partner, as a Non-Banking Financial Company and Business Correspondent, uses the financial services delivery model called Kshetriya Grameen Financial Services (KGFS) to provide a range of financial services. It also provides tailored financial advice through local brick and mortar village branches, thus representing an alternative to the standard microfinance movement in India, which has focused primarily on microcredit.

The type of financial products offered by KGFS, along with the large data collection effort carried out by the research team, allows our study to be the first to evaluate the impact of increased access to financial services as a whole. In addition, our focus on rural areas allows us to provide valuable insights on this model of expanding financial access to remote rural communities.

Our results show that living in an area where KGFS expanded increases households' likelihood of participating in formal banking. Compared to the control group, at endline, treated households are significantly more likely to have formal outstanding loans, have a larger number of formal loans, and borrow more from formal lenders. They also report higher savings. In addition, households in treated areas are less likely to borrow from informal sources such as moneylenders and financiers. Importantly, these household level changes are paralleled in network changes: households in treated villages report lower borrowing capacity both from moneylenders and from individuals living both inside and outside their village than control group households.

Greater access to formal finance enables households to benefit from greater economic opportunities: our intervention increases households' likelihood of being self-employed and raises both business and household's overall income. Consistent with a large body of theoretical research, we find that formal financial access promotes entrepreneurship and encourages households to take on riskier – but more profitable – activities. With more

households becoming self-employed and diversifying their income-earning activities, we also observe daily wages increase.

All in all, our initial results suggest that expanding access to formal financial products and services to rural households not only crowds out informal borrowing, but also has a positive impact on saving, on households' business activities, and on their ability to cope financially with health shocks. Increasing access to formal financial services seems to positively impact poor households through income stabilization and increased financial security.

From a policy perspective, the results provide strong evidence on the links between financial inclusion and overall socio-economic well-being of the households. We see that access to formal financial services to unbanked and underbanked population increases the usage of financial services as depicted in the form of increased likelihood to save and borrow from formal sources, leading to positive externalities on the overall welfare of households.

Financial inclusion has been identified as a key enabler by the World Bank in achieving the Sustainable Development Goals.[33] Findings from our study are consistent with this narrative, as greater financial inclusion contributes to diversification of income sources and increased ability to cope with shocks.

Policy discussions around the impact of inclusive finance have largely centered around the impact of microfinance programs that focus on providing access to microcredit to poor and low-income households. Banerjee et al. (2015), in their assessment of six studies spanning across six countries and four continents, find the impact of microcredit to be 'modestly positive, but not transformative' on the welfare indicators of the study households. However, financial inclusion in its holistic form goes much beyond mere access to credit, by encompassing a whole suite of financial products and services that are offered responsibly and sustainably, in a well-regulated environment. In this context, our study offers an extremely policy-relevant perspective as it not just evaluates the impact of access to credit but looks at the overall impact of access to formal financial services in the form of a rural banking model. Our implementation partner, KGFS, provides a whole range of products and services such as savings, loans, insurance, remittances and other investment products. The key feature of the KGFS model lies in its ability to adopt a more "hands-on" approach with its clients due to geographical proximity to the population it serves and the level of interaction it maintains with its clients. This allows KGFS to know the customer beyond the RBI regulated-KYC norm and advising them on their financial decisions.

Recognizing the potential of financial inclusion in improving the lives of the poor and low-income households, the Indian Government has taken several steps over the last few decades towards accelerating formal financial services for the financially excluded

---

[33] http://www.worldbank.org/en/topic/financialinclusion

population. The Pradhan Mantri Jan Dhayn Yojana (PMJDY)[34] program, one of the largest financial inclusion programs in the world, is the most recent example showcasing Government's effort in this direction. However, despite sustained efforts, 47% of Indians remain excluded from the formal financial system, and 43% of the total banks accounts remain dormant, implying significant barriers to their usage[35].

The results from this study indicate that a rural banking model (like KGFS) that reaches geographically remote areas along with maintaining a high level of human touch with its client base has the ability to surmount the various barriers to the sustained usage of formal financial services.

From a policy perspective, our findings suggest that more efforts should be devoted towards accelerating the access to formal financial services among the last-mile consumer. Given the links between financial inclusion and economic and social security, more attention should be devoted to understand the key levers that influence the usage of formal financial services among low-income households, and to propose relevant solutions that remove these barriers. The KGFS model provides one solution to the long pending question on the true impact of inclusive formal financial services on some of the most vulnerable populations of our society.

References

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

Ananth, B., Chen, G., Rasmussen, S., May 2012, The Pursuit of Complete Financial Inclusion The KGFS Model in India, No. 4, May 2012, CGAP Report.

Bandiera, O., Burgess, R., Das, N., Gulesci, S., Rasul, I., Sulaiman, M. (2017) "Labor Markets and Poverty in Village Economies", *The Quarterly Journal of Economics,* Volume 132, Issue 2, 1 May 2017, Pages 811–870.

Banerjee, A.V. (2004) "Contracting Constraints, Credit Markets and Economic Development," in *Mathias Dewatripont, Lars P. Hansen and Steven Turnovsky, eds., Advances in economics and econometrics: Theory and applications*. Eighth World Congress (Vol. 3). Cambridge: Cambridge University Press, pp. 1-46.

Banerjee, A.V. and A.F. Newman (1993). "Occupational choice and the process of development. *Journal of Political Economy*, *101*(2), pp.274-298.

---

[34] https://www.pmjdy.gov.in/
[35] http://datatopics.worldbank.org/financialinclusion/country/india

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

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

Breza, E., and C. Kinnan (2016). "Measuring the equilibrium impacts of credit: evidence from the Indian microfinance crisis". *Manuscript, Northwestern University-*

Bruhn, M. and I. Love (2014). The real impact of improved access to finance: Evidence from Mexico. *The Journal of Finance*, *69*(3), pp.1347-1376.

Burgess, R. and R. Pande (2005). "Do Rural Banks Matter? Evidence from the Indian Social Banking Experiment.", *The American Economic Review* Vol. 95, No. 3, pp. 780-795

Casaburi, L. and J. Willis (2016). "Time vs. State in Insurance: Experimental Evidence from Contract Farming in Kenya", *working paper*

Casey, K., Glennerster, R., Miguel, E., (2012). "Reshaping Institutions: Evidence on Aid Impacts Using a Preanalysis Plan", *The Quarterly Journal of Economics*, Volume 127, Issue 4, 1 November 2012, Pages 1755–1812,

Cole, S., D. Stein, and J. Tobacman (2014). Dynamics of demand for index insurance: Evidence from a long-run field experiment. *The American Economic Review*, *104*(5), pp.284-290.

Crépon, B., F. Devoto, E. Duflo, and W. Parienté (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), pp.123-150.

Dupas, P., and J. Robinson (2013). "Savings constraints and microenterprise development: Evidence from a field experiment in Kenya." *American Economic Journal: Applied Economics, 5*(1), pp.163-192.

Dupas, P., A. Keats, A. and J. Robinson (2017). "The Effect of Savings Accounts on Interpersonal Financial Relationships: Evidence from a Field Experiment in Rural Kenya." *The Economic Journal, forthcoming*.

Government of India, Planning Commission, 2009, A Hundred of Small Steps, Report of the Committee on Financial Sector Reforms.

Kaboski, J. and R. Townsend (2005). "Policies and impact: An analysis of village- level microfinance institutions", *Journal of the European Economic Association* 3, 1-50.

Karlan, D. and J. Zinman, (2010). "Expanding credit access: Using randomized supply decisions to estimate the impacts." *The Review of Financial Studies,* Vol. 23, No. 1, pp. 433-464

Karlan, D., R. Osei, I. Osei-Akoto, and C. Udry (2014). "Agricultural Decisions after Relaxing Credit and Risk Constraints." *Quarterly Journal of Economics*,129 (2): 597-652

Kling JR, Liebman JB, Katz LF. "Experimental Analysis of Neighborhood Effects.", *Econometrica* 2007;75 (1) :83-119.

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

Schaner, S. (2015). "Do Opposites Detract? Intrahousehold Preference Heterogeneity and Inefficient Strategic Savings." *American Economic Journal: Applied Economics* 7 (2): 135-174.

Taniparti, N. IFMR LEAD, April 2016, Expensive but Worthwhile: Resolute Informal Financier Markets, Policy Brief #02.

Appendix A: Sample Selection and Randomization

We use data from 50 service area pairs[36] across three districts. The average service area of a branch spanned 3-5 km from the branch office and encompassed an average of 10 villages and 2,400 households. Villages serviced by a single branch were typically well-connected by roads and bus routes. Below we first describe randomization and then surveying.

The selection of potential branch sites and randomization across them proceeded as follows: In conjunction with the bank, potential location sites were identified using a global position system (GPS)-based population survey which determined relevant political, administrative and social boundaries. Once all feasible branch locations in the district had been designated, we used Edmond's algorithm for minimum distance matching to construct pairs of service areas. This matching for treatment and control allows the study to overcome issues in seasonality and geographic correlation in outcomes by minimizing differences between paired branches. It also improves balance across treatment and control villages on observed and unobserved factors, and provides a strong service-area-level control variable. For several 2001 census village outcomes (including caste

---

[36] 101 service areas are covered, due to one triplet with two control areas.

composition, number of primary schools, water facilities and proportion of irrigated land), we find that controlling for pair fixed effects explains roughly 70% of the variance. One service area in each pair was then randomly selected to receive a bank branch first (treatment area). Expansion in the other area (control) was delayed for 36 months. Bank employees were not informed about the study or whether their branch is a study branch or not. Treatment and control areas of the same pair will be surveyed simultaneously. Surveyors do not know the treatment status of villages and are rotated across treatment and control.

The opening of bank branches happened in three rounds due to operational constraints following the Indian microfinance crisis in late 2010. We will account for this fact in the empirical analysis and may, at times, restrict the analysis to certain rounds only, for example when looking at the influence of the Pradhan Mantri Jan Dhan Yojana (PMJDY) scheme which started in 2014.

Appendix B: Data

Survey instruments were the primary means of data collection for this evaluation. For the first part of the study covering 8 service area pairs, survey instruments were implemented by Abacus Field, the survey arm of International Market Research Bureau (IMRB) with projects throughout India. The initial data collection utilized paper-based survey instruments. Data for the rest of the study was collected digitally using Survey CTO, a survey and data collection application. The survey instruments used for different components of the study are described below.

- Household listing: households within the intervention population are enumerated at baseline. The surveyors will take a roster of household members as well as contact information that can be used to track and to find households. This listing will serve as a sampling instrument, and as a compliance instrument.
- Village informant survey: political and social leaders are interviewed in each study village. This instrument collects information on land use, agricultural and non-agricultural wage rates, and migration.
- Household survey: several members of each randomly selected household are interviewed in each study village. Data are collected regarding sources and uses of income, including business and cultivation activities; financial literacy and activity in loans, savings, and insurance (formal and informal); health; well-being; social capital; female empowerment; and household structure. Importantly, data collected at the household level will be aggregated up to address the impact of financial access along dimensions such as employment rates.
- Social Networking Mapping survey: the head and the spouse of all households in sampled villages are interviewed. It makes use of a census which assigns a unique identifier to each individual in a village. This survey collects data on leisure and borrowing contacts of respondents within the village and identifying them by the unique identifier. It also collects data on the possible borrowings from these contacts as well

as any borrowing contact they might have outside the village (who we cannot identify). This survey aims to capture the changes in the social network patterns at the village level following introduction of formal financial access.

Appendix C: Power calculations

C.1 INTRODUCTION

These power calculations are based on baseline and midline data from the KGFS impact evaluation. Our estimates are discussed below. We define the treatment as being the provision of access to finance over a geographic area associated with each treatment-group KGFS branch, and thus estimate the average treatment effect (ATE) over the study population.

Overall, these power calculations suggest that a change of about 10% of the mean value is detectable for the majority of outcomes, even without fully accounting for the pair-wise matching of service areas. However, some outcomes are detectable only at above a change of 20%. When taking the matching strategy into account, all outcomes can be detected at or below a change of 20% of the control mean. As the later analysis will show, decreasing the number of pairs damages the study's ability to detect key outcomes – especially for those on the margin – but outcomes, generally, remain at similar levels of detectability down until 60 pairs or below. These power calculations assume that the sample is clustered at the service area level, but that a simple random sampling method is used to draw the sample from within the service area. They have been calculated for roughly 50 potential outcomes, although results are only presented for 18 of these outcomes in the table below. Overall, the evidence from using empirical inputs suggests that the study will be capable of detecting changes from the intervention.

C.2 OUTCOMES – THRESHOLDS FOR ECONOMIC AND STATISTICAL SIGNIFICANCE

The threshold for significance in the evaluation of social programs is typically the point at which a program becomes cost-effective relative to other programs. But in this case, we are evaluating a self-funding banking intervention: any benefit that derives from the intervention is 'free' in the sense that it does not consume assets that could be devoted to another program.

Table 13 presents the minimum detectable effect (MDE) as a percentage of the control mean for outcomes across several categories. Because the cost of this intervention is zero, we define a borderline significant effect as MDE equal to 5% of the control mean for

economic outcomes, and MDE equal to 2.5% of the control mean for health outcomes. Outcomes of successful programs are typically larger than these thresholds. The design of this evaluation ensures that, if health outcomes are not medically significant, positive impacts can be ruled out; economic outcomes will be detected at about 20% of the control mean, a threshold regularly met or exceeded in similar evaluations of banking and microfinance programs.

The difference between the two types of MDES has to do with whether we control for pair-wise matching in calculating the intra-cluster correlation. Worst-case MDES are calculated using ρ without pair demeaning. These are very conservative estimates, as they don't take into account the more sophisticated methods of controlling for pair-wise matching. Best-case MDES are calculated using ρ that takes into account pair demeaning. The difference between the two values represents, generally, the extent to which the pair-wise matching strategy will likely improve the power of any final analysis. In some cases – such as for value of business inputs, and the total area cultivated for farming – these differences are quite substantial.

Table 13. Minimum Detectable Effect Size as % of Mean

Category	Outcome	Mean	SD	MDES	MDES (pair demeaned)
Consumption	Total Consumption	933	961	12.48%	7.90%
	Amount Spent on School	2125	5628	24.39%	20.70%
Employment	Self-Employment	0.41	0.49	12.74%	10.08%
	Migration of HH Members	0.14	0.34	20.56%	20.02%
	Value of Business Inputs	4455	11741	40.01%	20.52%
Farming	Total Area Cultivated	11.3	22.4	41.83%	19.96%
	Total Farm Costs	11574	13887	11.91%	9.32%
	Fertilizer Expenditure	3115	3617	9.98%	8.82%
	Own Livestock	0.57	0.49	11.11%	6.43%
Financial Access	Loan Amount Outstanding	90339	134342	10.88%	10.88%
	Have Insurance	0.780	0.415	8.32%	4.11%
Other	Average Child Education	5.612	2.123	3.68%	2.67%
	Own Buildings	0.948	0.222	1.96%	1.82%
Well-Being/Health	Happiness Ladder	4.490	2.547	5.71%	4.92%
	People Visited in Gramum	16.61	27.94	13.24%	13.24%
	Days Missed Due to Health	2.06	4.26	24.99%	20.16%

Appendix D: Pre-Analysis Plan, Study Design, and Methods

In what follows, we report the Pre-Analysis plan for the study.

D.1 INTRODUCTION

Previous studies of microfinance, which focused on microcredit alone, have not found large effects on poverty alleviation (see for example, Banerjee et al.,2014) or households' ability to cope with shocks. In contrast, non-experimental evaluations of rural bank branches in India report significant poverty reduction (for example, Burgess and Pande, 2003). The intervention we implement provides a unique opportunity to undertake rigorous experimental evaluation of the impacts of bank branch expansion in rural areas at both the household and village level. Importantly, we will evaluate a financial service delivery model that uses bank branches in villages to provide a full range of credit, saving, and insurance services to entire communities. Our evaluation is also unique in the breadth of its outcome measurement and covers asset ownership, business investment, farming, migration, trust relationships, and health outcomes. An in-depth understanding of the impact of using rural bank branches to provide comprehensive financial services will inform policy on financial inclusion both in India and globally.

Our implementing partner is a large financial institution in rural South India (referred to as LFI below). It comprises of a group of Indian non-banking financial companies with the mission to "maximize the financial well-being of every individual and every enterprise in remote rural India by providing complete financial services". LFI offers financial products spanning loans, savings, and insurance in addition to tailored financial advice through local village branches, in order to effectively reach individuals in financially marginalized rural communities. The LFI model is an alternative to the standard microfinance movement in India, which has focused primarily on microcredit.

The intervention to be evaluated is being implemented by a company founded by the LFI which operates, amongst other areas, in our study districts Thanjavur, Ariyalur and Puddukottai in Tamil Nadu. The base of the intervention is the expansion of bank infrastructure across villages. The construction and ongoing operation of the 50 bank branch locations that will constitute the intervention's treatment group is funded by the LFI.

In this pre-analysis plan, we will focus on outcomes in six areas: Financial Access, Income and Wealth, Consumption Smoothing, Human Capital Investment, Labor Market Outcomes and Female Empowerment. Additional survey components connected to this set of surveys look at Social Network, Health and Farming Technology Outcomes.

The structure of this pre-analysis plan is as follows: section 2 gives an overview of the experimental design, sampling and the key data sources used. Section 3 specifies the main regression model, how standard errors will be computed and which basic and extended controls will be used. Section 4 will lay out the various hypotheses we aim to test while section 5 deals with the analysis of treatment heterogeneity. We do not exclude the possibility of running further analyses for the final paper, but will make clear which estimations were specified in the plan and which were not (cf. Casey, Glennerster and Miguel, 2012).

D.2 STUDY DESIGN

D.2.1 Sampling

We will use data from 50 service area pairs[37] across three districts. The average service area spans several villages in a radius of 3-5 km from the branch and covers a population of an estimated 10,000 people. Below we first describe randomization and then surveying.

The selection of potential branch sites and randomization across them proceeded as follows: In conjunction with the bank, potential location sites were identified using a global position system (GPS)-based population survey which determined relevant political, administrative and social boundaries. Once all feasible branch locations in the district had been designated, we used Edmond's algorithm for minimum distance matching to construct pairs of service areas. This matching for treatment and control allows the study to overcome issues in seasonality and geographic correlation in outcomes by minimizing differences between paired branches. It also improves balance across treatment and control villages on observed and unobserved factors, and provides a strong service-area-level control variable. For several 2001 census village outcomes (including caste composition, number of primary schools, water facilities and proportion of irrigated land), we find that controlling for pair fixed effects explains roughly 70% of the variance. One service area in each pair was then randomly selected to receive a bank branch first (treatment area). Expansion in the other area (control) will be delayed for 36 months. Bank employees are not informed about the study or whether their branch is a study branch or not. We will assess the impact of increased formal financial access by comparing treatment areas with control areas two years after branch opening in treatment areas. Treatment and control areas of the same pair will be surveyed simultaneously. Surveyors do not know the treatment status of villages and are rotated across treatment and control.

The opening of bank branches happened in three rounds due to operational constraints following the Indian microfinance crisis in late 2010. We will account for this fact in the empirical analysis and may, at times, restrict the analysis to certain rounds only, for example when looking at the influence of the Pradhan Mantri Jan Dhan Yojana (PMJDY) scheme which started in 2014.

D.2.2. Selection of households within service areas

In each service area, a total of 46 households were selected for inclusion in the household survey. The selection of households generally followed a two-stage design to account for clustering of households in villages, while ensuring that the sample is representative of the chosen service areas. The first stage employed a probability proportional to size (PPS) sampling of villages within service areas. That is, villages were drawn to be included in the sample according to their relative population size. Additionally, the center village with the intended branch location was included. Each service area was allocated 46 baselines which were divided evenly into portions, and villages were drawn to be included in the sample according to their relative population size. Additionally, the center village with the intended branch location was always included in the baseline selection.

In stage two, listing was conducted with a 5-household skip in all villages sampled during stage one, collecting residential addresses and information for identification purposes, such as names and occupations of household members. We dropped all households that did not include a woman between the ages of 18 and 55. We then randomly selected the number of households in each village that had been determined in stage one.

D.2.3 Key data sources

<u>*Baseline surveys*</u> (prior to the intervention, starting in September 2010):

The baseline household survey occurred in each pair prior to the opening of the treatment branch. Attrition has been below 5% and not differential across treatment and control group. After an initial

---

[37] 101 service areas are covered, due to one triplet with two control areas.

wave of 8 branch openings in 2010, branch openings only continued at the start of 2013 and lasted until September 2015 due to operational constraints and delay in permission to open following the microfinance crisis in late 2010. Delay on the side of branch openings resulted in a further, shorter break in branch openings. Overall, 8 service areas were opened during round one, 34 during round two and 8 during round three. In total, 4,391 households were approached over the three rounds of baseline surveys.

Besides the baseline household survey, we conducted a short survey to the Panchayat head about basic village characteristics.

The Household survey itself consists of two parts with several sections each. For the first part, the head of the household or their spouse is interviewed, while for the second section part, we interview the female spouse. The second part is administered to women as it contains questions on female empowerment and child health amongst others.

Endline surveys (18-24 months after branch opening in treatment areas, starting in September 2013. Scheduled to run until mid-2017):

In the endline household survey, the same data is collected from the respondents interviewed at baseline, with adjustments for relevant external changes (e.g. the introduction of Aadhaar identification cards or voter ID cards that led to additional questions) or to clarify meaning.

LFI Customer Management System data

We will also use administrative LFI client data to augment the quality of self-reported survey measures. Monthly data on enrollment (registering in a branch), product take-up, frequency of loan renewal and other outside loans taken can be used to give a more detailed view of development of treatment over time.

D.3 METHODS

D.3.1 Estimation of treatment effects

Our main specification will model the effect of the randomized treatment, which is increased access to formal finance through the opening of a LFI bank branch in a service area. Given the use of pair-wise matching described earlier in assigning randomization, we will control for pair-wise fixed effects. Drawing on the endline household survey, we will thus estimate the following model:

$$Y_{ik} = \propto_0 + \propto_1 T_k + \propto_2 S_k + \delta_{pk} + \varepsilon_{ik}$$

In the regression above, *i* indexes the individual or household and *k* indexes her service area. Y_{ik} is a given outcome (e.g. amount of formal savings for instance) for individual or household *i* in service area *k*. T_k is the service area treatment dummy, such that $\propto_1$ gives the intent-to-treat effect. S_k are survey round dummies, δ_{pk} are pair fixed effects and ε_{ik} is the idiosyncratic error term.

Standard errors will be clustered at the level of randomization, i.e. at the service area level. In case we can draw on a subset of pairs/clusters only, we will compute wild bootstrapped standard errors clustered at the service area level following Cameron et al. (2009).

Additionally, we will draw on both baseline and endline household survey data and estimate intent-to-treat (ITT) effects in a difference-in-differences (DID) framework, using the following specification that considers changes over time in our panel:

$$Y_{ik\theta} = \propto_{0\theta} + \propto_1 T_k + \propto_2 S_k + \propto_3 \theta_k + \beta\theta_k * T_k + \delta_{pk} + \varepsilon_{ik}$$

where θ is the endline/baseline indicator and β the DID coefficient. Otherwise notation follows the basic cross-section regression above. We will additionally consider a specification using a vector of controls and pair fixed effects:

$$Y_{ik\theta} = \propto_{0\theta} + \propto_1 T_k + \propto_2 S_k + \propto_3 \theta_k + \beta\theta * T_k + \delta_{pk} + X_{ik\theta} + \varepsilon_{ik}$$

Individual, household and service area level controls
Besides estimating the model as described above, we will estimate it controlling, additionally, for several baseline characteristics in order to improve the precision of our estimates. In order to ensure that we do not obtain spurious results based on the inclusion or exclusion of controls, we will conduct a robustness analysis.
The basic set of controls at the individual and village level will be:
Basic demographic controls: {age, education, land ownership, caste, religion}
Basic village level controls: {village size, distance to branch village (real – for treated village; or hypothetical – for control villages)}
To explain heterogeneity in treatment results, we will employ an extended set of controls, namely:
Extended demographic controls: {risk aversion, experience of income shocks or major health shocks, business ownership, number of children, gender of respondent}
Extended village level controls: {road connectivity, distance to next factory, cost of travel to next city, number of financiers, number of moneylenders, average (agricultural and non-agricultural) wage level, share of land irrigated in village}
If any of these control variables has an equal to or larger than 95 percent share of uniform answers at baseline, we will only include them if we have reason to believe that they have a strong connection to the specific outcome indicator and mark this exception in the analysis.

D.3.2 Procedure for accounting for attrition, non-response, questions with limited variation and extreme values
Survey attrition
If a sampled household is not found after at least three attempts, it is not replaced and counted towards general attrition. Rates are calculated as the share of all non-surveyed households (whether not found or non-consenting) of the initially sampled households.
We will then test if attrition is differential across treatment and control areas. If attrition is found to be related to treatment status at the 5 percent significance level, we will employ a bounding method to obtain ranges on our treatment estimates which are robust to this attrition.
Missing data from non-response to individual questions
No imputation for missing data from item non-response at follow-up will be performed. We will check whether item non-response is correlated with treatment status following the same procedures as for survey attrition, and if it is, we will construct bounds for our treatment estimates that are robust to this. In the regression analysis, we will replace missing observations by 0 and generate an associated dummy.
Questions with limited variation
Questions for which 95 percent of observations have the same value within the relevant sample will be omitted from the analysis and will not be included in any indicators or hypothesis tests, in order to limit noise in the analysis. Should this omission rule result in the exclusion of all relevant variables for an indicator, we will not calculate the indicator.
Extreme values or outliers
Extreme values or outliers are identified using the three-sigma rule (also called the 68-95-99.7 rule). We top code variables at 3 standard deviations, meaning that for values outside an interval defined

as +/- 3 standard deviations of the mean, values are set at the upper or lower bound of the interval. When trimmed, values outside this interval are set to missing. Another alternative method is to top code variables at the 99th percentile. We use both methods in order to ensure the robustness of our results. In the latter method, we top code the top 1% of the distribution, meaning that values at the top are set to the value of the 99th percentile. When trimmed, values at the top of the distribution are set to missing. These usual statistical procedures are used in order to ensure that outliers do not drive the results.

D.3.3 Procedure for dealing with multiple outcomes
We will aim to account for the effects of multiple, correlated outcomes by grouping our outcome measures into domains, based on the idea that items within a domain are measuring an underlying common factor. Our six domains are detailed below. Then we will sign the outcomes within each domain, so that the hypothesized effects go in the same direction, and take a standardized treatment effect within that group (compare Kling et al., 2007, Finkelstein et al., 2010).

D.4 HYPOTHESES

We will group our hypotheses into six broader categories for which we will try to account for effects of multiple hypothesis testing (see section 3.3 above):

1. *Financial Access*
      1.1 Increase in formal financial activity
      1.2 Change in informal borrowing

2. *Income and Wealth*
      2.1 Increase in riskier but higher return activities and asset investments
      2.2 Increase in savings and wealth
      2.3 Change in diversification of financial activity

3. *Consumption smoothing*
      3.1 Increase in expenditure on durable goods
      3.2 Change in expenditure on non-durable goods
      3.3 Increase in formal borrowing or savings used in response to a shock
      3.4 Change in informal borrowing or savings used in response to a shock
      3.5 Smaller consumption decrease and volatility in response to a shock

4. *Educational Investment*
      4.1 Increase in expenditure on education

5. *Labor Market outcomes*
      5.1 Increase in entrepreneurship/ business ownership
      5.2 Reduction in unemployment
      5.3 Change in permanent migration
      5.4 Change in seasonal migration

6. *Female empowerment*
      6.1 Increase in women reporting having a source of income
      6.2 Increase in women reporting being able to make joint or individual decision in financial household issues
      6.3 Increase in women reporting being able to make joint or individual decision in non-financial household issues
      6.4 Increase in subjective well-being of female household members

D.5 OUTCOME RESPONSE TO TREATMENT HETEROGENEITY

A key interest of the study is to find out mechanisms of impact. We therefore aim to identify the channels through which household and village-level outcomes change for different groups, which will help improve the design of products and services for specific household types. Using the specification outlined above, we will examine treatment effects for the following subgroups:

- By household type: households with landholding, female-headed households, households with school-aged children, households with a literate household's head, income quintile analysis.
- By occupation: farmers, wage laborers, business owners.
- By treatment status (treatment-on-treated) or likelihood of take-up as predicted by baseline variables at endline.

Appendix E: Variable Definitions

This appendix defines all variables used throughout the main report that were constructed based on data collected in our study.

Table 14. Variable Definitions for Baseline Descriptive Variables

Variable	Definition
PANEL B **Demographics**	
Head of Household Characteristics	Age, gender, and highest level of education.
Distance to Branch (km)	Distance (in km) between household house and associated KGFS branch location: for treatment villages, this is the actual branch location; for control, this is the hypothetical branch location determined by KGFS.
Occupation & Income	
Employed in Wage Labor (last 7 days)	Dummy variable equal to 1 if the household answered yes to: "Did any member in the household work for wage as labor/service in non-household owned employment in the last 7 days in this town panchayat/ village panchayat?"
Below Poverty Line	Dummy variable equal to 1 if the household's self-reported monthly income falls below Rs. 880 times the number of household members. There are fewer observations in the table because it was not collected in Baseline I. 3 components: (1) Self-reported income: "How much rupees, in total, did household members earn in the last 30 days from all income-generating activities?", (2) Number of household members: "How many members are there in your household?" (3) Rural Tamil Nadu poverty line of Rs. 880 (monthly per capita), determined by the Suresh Tendulkar Panel Recommendations in 2011-2012.
Self-Employed or Owns a Business	Dummy variable equal to 1 if the household answered yes to: "Is any member of the household currently self-employed or the owner of a business which excludes any sort of farming or animal-husbandry?"
Earns Farming Income	Dummy variable equal to 1 if the household answered yes to: "Was your household involved in farming any crop in the past 12 months?" "Involvement" can include farming for sustenance, hiring laborers, etc.

Wage Labor Income (weekly)	For those who reported wage employment, 2 components: i) Cash wage per payment cycle (including in-kind payment, converted to Rs.) ii) Frequency of payment (daily, weekly, fortnightly, monthly, "daily" converted to "weekly" by multiplying by 8 hours per workday, and "fortnightly" and "monthly" were divided by 2 and 4, respectively). Equal to 0 for households that reported no wage employment in the last 7 days.
Business Income (30 days)	For those who reported self-employment/owned business, calculation uses 2 components: (1) Rental income: "Over the most recent 30 days of operations what was the total rental value in rupees of equipment (tools, machines, animals etc) leased from this [business] activity?" (2) Sales income: "What is the estimated value of all sales of finished good and services over the most recent 30 days of this [business] activity?" Equal to 0 for households that reported no self-employment or owned business.
Farming Income	For those who reported farming, calculated as the sum of 4 components as reported for the last farming season: i) Seasonal unsold crops, ii) Seasonal sold crops, iii) Perennial unsold crops, iv) Perennial sold crops Equal to 0 for households that reported no farming in the last 12 months.

Wealth & Properties

Owns Land	Dummy variable equal to 1 if households report "owning or currently holding the right to use plots of land", excluding land on which any buildings the household owns/use stand.
Cultivated Area (acres)	Self-reported amount of cultivable land owned. Equal to 0 if households report owning no land, or if household owns land but no cultivable land.

Formal & Informal Borrowing

Number of Outstanding Formal Loans	Equal to 0 if households report having no loans, or if household has loans but no formal loans. A loan is defined as "formal" if it is taken from a: Private Bank, NGO/MFI, Nationalized Bank, PAC/Co-operative Bank, Self Help Group (SHG), Non-Banking Financial Corporation (NBFC).
Formal Borrowed Amount (last 24m.)	Total borrowed amount of all formal loans taken by all households in the last 24 months. Equal to 0 if no loans or no formal loans.
Number of Outstanding Informal Loans	Self-reported number of loans taken from informal sources at the time of the survey. Equal to 0 if households report having no loans, or if household has loans but no informal loans. A loan is defined as being "informal" if it is taken from a: Friend/Neighbor/Relative, Shopkeeper, Employer, Moneylender, Pawnbroker, Landlord, Rotating Savings Group (ROSCA), Chitfund, Financier, or Religious Trust.
Informal Borrowed Amount (last 24 months)	Total borrowed amount of all loans taken from informal sources in the last 24 months. Equal to 0 for households that reported no loans or no informal loans.

Savings & Insurance

Probability of Having Any Savings	Dummy variable equal to 1 for households that reported a non-zero amount in their savings account(s) with a bank or financial institution. Equal to 0 if no savings accounts, or savings accounts with no savings.
Amount in Savings Account(s)	Self-reported total amount of money saved in all savings accounts with banks or financial institutions: "What is the total amount of money that you have saved in all accounts?" Equal to 0 for households that report having no savings accounts.

Probability of Giving Out Any Loans	Dummy variable equal to 1 for households that answer yes to: "Has your household given out loans to individuals not in your household, which are currently outstanding?"	

Shocks Experienced in the last 12 months

Any Shock	Dummy variable equal to 1 for households that reported having experienced at least once of the following events "unexpectedly during the course of the last 12 months": Death of a money earner, Heavy rain/flood, Drought, Pest infestation or crop diseases which destroyed more than 1/4 of the crop, Serious injury/illness that kept the affected household member from doing normal activities.
Serious Injury or Illness Shock	Dummy variable equal to 1 for households that reported having the event of a "serious injury/illness that kept the affected household member from doing normal activities" at least once in the last 12 months.

PANEL C
SNM

Number of households (Census)	Number of households in each village as enumerated in SNM Census Survey.
Number of surveyed households (SNM)	Number of households actually surveyed within each village recorded in SNM Census.
Distance to the bank branch (km)	Distance in km from the center of each village to the associated KGFS branch location (for treatment villages, this is the actual branch location; for control, this is the hypothetical branch location identified by KGFS).

Note: Panel B variables are from the Household Survey component of our study. Fluctuations in observation counts are due to missing values. Panel C are from the Social Network Mapping (SNM) survey component of our study, and are aggregated at the village level. The fluctuation in observation counts is due to missing GPS data. Tables 26 uses these variables.

Table 15. Variable Definitions for Formal & Informal Loans

Variable	Column	Definition
P(Formal Loan Outstnd), P(Informal Loan Outstnd), Nr of Outstnd Formal Loans, Nr of Outstnd Informal Loans, Formal Borrowed Amt, Informal Borrowed Amt	1-6	cf Table 14.
Formal Share of Borrowed Amt	7	Total formal amount borrowed in the last 24 months divided by total formal and informal amount borrowed in the last 24 months, i.e. the variable used in Column (5) divided by the sum of those used in Columns (5) and (6) for each household. The observation count falls in Column (7) only. This is because only households with non-zero borrowing are included, to allow for the share to be computed.

Note: Tables that concern these variables are Tables 1 and 10.

Table 16. Variable Definitions for Reasons for Borrowing

Variable	Column	Definition
Farming & Business Investment	1	Total principal amount borrowed in currently outstanding loans for which the reported primary purpose was "general farming or business inputs (seeds, fertilizer, livestock, raw materials, inventory)", "farming or business equipment (plow, sewing machine, rickshaw)", or "paying employees [for business or farming labor]".
House & Land Repair	2	Total principal amount borrowed in currently outstanding loans for which the reported primary purpose was "repairing or upgrading a house, land, or buildings".
Weddings	3	Total principal amount borrowed in currently outstanding loans for which the reported primary purpose was "weddings, festivals, or functions".
Day-to-Day Expenses	4	Total principal amount borrowed in currently outstanding loans for which the reported primary purpose was "buying food, clothing, or other day-to-day household items".
Education Related	5	Total principal amount borrowed in currently outstanding loans for which the reported primary purpose was "education related".
Health Related	6	Total principal amount borrowed in currently outstanding loans for which the reported primary purpose was "covering health expenses".

Note: These variables are used in Tables 2, and 3

Table 17. Variable Definitions for Borrowing from Moneylenders

Variable	Column	Definition
P(Moneyl. and Fin. Loan Outstnd)	1	Probability of having an outstanding loan at the time of the survey from a moneylender or financier. Dummy equal to 1 if a loan is held, and equal to 0 if the household only has loans from other sources or has no loans.
Nr of Moneyl. and Fin. Outstnd Loans	2	Total number of outstanding loans held by the household at the time of the survey from moneylenders or financiers.
Moneyl. and Fin. Borrowed Amount	3	Total borrowed amount of all loans taken from moneylenders or financiers in the last 24 months. Equal to 0 if the household had no reported loans or none from moneylenders/financiers.

Note: These variables are used in Table 4.

Table 18. Variable Definitions for SNM Contacts

Variable	Column	Definition
Any Contact	1	Dummy variable equal to 1 if households reported any contacts they could borrow from their inside village/outside village/moneylender contacts in the Social Network Mapping component survey.
Nr. Contacts	2	Total number of inside village/outside village/moneylender village contacts households reported having.

Emergency Borrowing Capacity	3	Total Rs. amount that households report they could borrow in the event of an emergency from inside village/outside village/moneylender contacts.
Business Borrowing Capacity	4	Total Rs. amount that households report they could borrow from inside village/outside village/moneylender contacts if they were going to start a business or expand an existing business in the last 12 months.
Actual Borrowed Amt	5	Total Rs. amount households have actually borrowed from inside village/outside village/moneylender contacts in the last 12 months.

Note: These variables are used in Tables 5 and 6.

Table 19. Variable Definitions for Savings & Insurance

Variable	Column	Definition
P(Any Saving Account), Saving Account Amt, P(Giving out loans)	1, 2, 3	Columns 1, 2, 3 can be matched one-to-one to variables defined in Table 14.
Nr of Informal Loans Given Out	4	Total number of loans given out by the household informally that are currently outstanding at the time of the survey. This includes loans to: friends, relatives outside the household, neighbors, shopkeepers, money guards, tenants, and employees.
Informal Loans Given Out Amount	5	Total Rs. amount that households report having currently lent out.
P(Active Insurance)	6	Dummy variable equal to 1 if any household member has insurance that is currently active. This includes government provided crop insurance, Life Insurance Corporation (LIC), General Insurance Corporation (GIC), etc.

Note: These variables are used in Table 7.

Table 20. Variable Definitions for Employment & Income Composition

Variable	Column	Definition
HH Farms, HH Self Employed, Self-Reported Income (log), Below Poverty Line, Business Income (30 days) (log)	1, 2, 4, 5, 8	Columns 1, 2, 4, 5, and 8 can be matched one-to-one to variables defined in Table 14. Log income is calculated as log(income + 1).

P(Business Hires Employees)	3	Dummy variable equal to 1 if the household business hired employees to work in the business. Note that this is as per the definition of "self-employment/business" specified in Table 14 - that is, it excludes farming and animal-husbandry. Equal to 0 if household did not hire any employees in the most recent 30 days, or if the household is not self-employed or does not own a business.
Governmental Wage Labor Income (log)	6	Log of the household's daily wage labor income subsidized by the government, namely from NREGA. For each member of the household who has worked in wage labor in the past 7 days, respondents report the cash and in-kind wage paid in various time units (hourly, daily, weekly, etc.), which are then converted to a log daily wage using log(wage + 1).
Non-Governmental Wage Labor Income	7	Log of the household's daily wage labor income, excluding farming, government work, and government subsidized work (NREGA). For each member of the household who has worked in wage labor in the past 7 days, respondents report the cash and in-kind wage paid in various time units (hourly, daily, weekly, etc.), which are then converted to a log daily wage using log (wage + 1). Primary labor types include: non-farm jobs (unskilled), private formal salary jobs, electricians, drivers, and woodworkers.

Note: These variables are used in Table 8.

Table 21. Variable Definitions for Wages from-Non Household Employment

Variable	Column	Definition
Log total daily wage, per daily payment cycle	1, 2, 3	Total daily wage, per daily payment cycle. Sum of daily cash and inkind wages per household, earned from working for non-household owned employment. Unconditional value, i.e. zero for households who either did not work for non-household owned employment, or, who worked for non-household owned employment that only paid employees on non-daily payment cycles.
Log total monthly wage, per daily payment cycle	4,5,6	Total monthly wage, per monthly payment cycle. Sum of monthly cash and inkind wages per household, earned from working for non-household owned employment. Unconditional value, i.e. zero for households who either did not work for non-household owned employment, or, who worked for non-household owned employment that only paid employees on non-monthly payment cycles.

Note: These variables are used in Table 9.

Table 22. Variable Definitions for Life Perception

Variable	Column	Definition

Variable	Column	Definition
Happiness Scale	1	Respondents are asked to consider how they feel at the moment and what they hope for their future, and are then asked: "Keeping everything in mind, tell us about yourself overall: are you very happy, quite happy, not very happy, or not happy at all?" The variable is coded as (1) if very happy, (2) if quite happy, (3) if not very happy, and (4) if not happy at all.
Perception of Current Life	2	Equal to a number from 1 to 10 representing each household respondents' current perception of life, where 10 is the best possible life, and 1 is the worst. Respondents are presented with a ladder with 10 steps, increasing from 1 at the bottom to 10 at the top, and are told that the top of the ladder represents the best possible life for them, while the bottom represents the worst possible life. They are then asked: "On which step of the ladder do you personally feel you stand at this time, assuming that the higher the step, the better you feel about life, and the lower the step, the worse you feel about it? Which step comes to the closest to the way you feel?"
Better Perception of Life 5 years ago	3	Equal to 1 if the respondent feels they had a better life 5 years ago than at the time of the survey; 0 if the respondent feels their life is the same as it was 5 years ago; -1 if the respondent feels they had a worse life 5 years ago than at the time of the survey. This is measured by asking the respondent the same question as in Column 2, but with respect to their life 5 years ago as opposed to at the time of the survey. Their responses between the two questions are then compared to determine this variable.
Better Perception of Life in 5 years	4	Equal to 1 if the respondent feels they will have a better life 5 years in the future than at the time of the survey; 0 if the respondent feels their life will be the same; -1 if the respondent feels their life will be worse in 5 years than at the time of the survey. This is measured by asking the respondent the same question as in Column 2, but with respect to their life 5 years in the future as opposed to at the time of the survey. Their responses between the two questions are then compared to determine this variable.

Note: These variables are used in Table 11.

Table 23. Variable Definitions for Psychological Distress

Variable	Column	Definition
Self-Assessed Health	1	Equal to a number from 1 to 10 representing the Household Survey Part 2 respondents' self-reported health at the time of the survey. The respondent is presented with a picture of a ladder, and is asked: "if the top rung of the ladder represents very good health and the bottom step represents very bad health, where would you place yourself?". Note that the Part 2 respondent is a female spouse within the household, and answers questions on female empowerment, health, social networks, and risk aversion, while the Part 1 respondent answers all else.

Felt Nervous, Hopeless, Restless, Depressed, Everything Was an Effort, Worthless	2-7	Columns 2-7 correspond to questions asked about how respondents felt over the past 30 days. The variables are equal to: 1 if the respondent reported feeling nervous/hopeless/restless/depressed/that everything was an effort/worthless none of the time; 2 if a little; 3 if some of the time; 4 if most of the time; and 5 if all of the time.
K6 Scale	8	The Kessler Psychological Distress Scale (K6) measure, calculated for each household by summing the values reported in Columns 2-7.

Note: These variables are used in Table 12.

Table 24. Variable Definitions for Household Controls

Variable	Definition
Distance to branch (km)	Distance in km from each household house to the associated KGFS branch location (for treatment villages, this is the actual branch location; for control, this is the hypothetical branch location identified by KGFS).
Head of Household Age	Age of the head of household, identified by the respondent.
Head of Household Years of Education	Number of years of education completed by the head of household. Respondents are asked what the highest class is that the head of the household completed, which is then converted into years: no class passed/Anganwadi/Balwadi is 0 years, and all other specified classes (Class 1-5, 6, 7, etc.) are converted by adding 2 to the class number. For example, Class 6 equates to 8 years.
Household Caste	Reported caste of the household, of the following categories: Forward Caste, Backward Caste, Most Backward Caste, Scheduled Caste, Scheduled Tribe, Other Backward Caste. Included in tables concerning the Household Survey Sample only, since caste was not asked in the SNM survey.
Household Religion	Reported religion of the household, of the following categories: Hindu, Christian, Muslim, Jain, Sikh, Buddhist, Zoroastrian (Parsi), Atheist. Included in tables concerning the Household Survey Sample only, since religion was not asked in the SNM Survey.
Land ownership	Reported land ownership of the household.

Note: These variables are utilized in all tables that report using household controls.

Appendix F: Additional Tables

Table 25. Baseline Randomization Checks

	Control Mean [SD] [1]	Treatment Mean difference (SE) [2]	Nr. Observations [3]
Panel A: Census 2001 - Service Area Level			
Demographics			
Total Households	2603.65 [1047.03]	94.90 (162.92)	101
Total Population	11381.76 [4475.15]	433.33 (717.92)	101
Scheduled Caste Population	2370.73 [1511.61]	238.74 (241.42)	101
Scheduled Tribe Population	52.31 [136.73]	-16.00 (15.56)	101
Facilities			
Number of Primary Schools	9.37 [4.12]	0.32 (0.76)	101
Number of Primary Health Centers	0.47 [0.64]	0.05 (0.14)	101
Financial Sector			
Has banking facility	0.98 [0.91]	-0.21 (0.15)	101
Number of Commercial Banks	0.55 [0.67]	-0.04 (0.11)	101
Number of Co-operative Commercial Banks	0.67 [0.89]	-0.09 (0.14)	101
Has credit societies	1.55 [1.06]	-0.05 (0.19)	101
Panel B: Main Household Sample at Baseline - Household Level			
Demographics			
Head of Household: Age	46.36 [12.43]	0.672 ** (0.30)	4066
Head of Household: Male	0.73 [0.44]	-0.014 (0.01)	4066
Head of Household: Years of education	7.46 [4.70]	-0.036 (0.13)	4066
Distance to Branch	2.26 [1.62]	-0.150* (0.09)	4066
Occupation & Income			
Below Poverty Line	0.42 [0.49]	0.01 (0.01)	2727
Employed in Wage Labor (last 7 days)	0.64 [0.48]	-0.01 (0.01)	4065
Self-Employed or Owns Business	0.17 [0.37]	-0.00 (0.01)	4066
Earns Farming Income	0.46 [0.50]	-0.03 (0.02)	4060
Wage Labor Income (weekly)	843.22	-2.46	4066

	Control Mean [SD] [1]	Treatment Mean difference (SE) [2]	Nr. Observations [3]
	[1830.30]	(38.39)	
Business Income (30 days)	2366.10	94.63	4066
	[12983.18]	(381.92)	
Farming Income (last season)	10137.34	-1155.30	4066
	[23517.84]	(743.07)	
Wealth & Properties			
Owns Land	0.55	-0.00	4064
	[0.50]	(0.02)	
Cultivated Area (acres)	0.97	0.07	4030
	[1.88]	(0.05)	
Formal & Informal Borrowing			
Number of Outstanding Formal Loans	1.19	-0.03	4048
	[1.51]	(0.04)	
Formal Borrowed Amount (last 24 months)	48764.87	-4544.17**	4040
	[95532.46]	(2175.77)	
Number of Outstanding Informal Loans	1.89	-0.15**	4047
	[2.06]	(0.06)	
Informal Borrowed Amount (last 24 months)	42103.90	-1790.36	4040
	[68094.41]	(1490.7)	
Saving & Insurance			
Probability of Having Any Savings Acct	0.84	0.00	4066
	[0.37]	(0.01)	
Amount in Savings Account(s)	5131.81	568.68	3960
	[17022.92]	(563.99)	
Probability of Giving Out Any Loans	0.05	0.00	4064
	[0.22]	(0.001)	
Shocks Experienced in the Last 12 Months			
Any Shock	0.40	-0.02*	4066
	[0.49]	(0.01)	
Serious Injury or Illness Shock	0.21	-0.01	4066
	[0.41]	(0.01)	
Daily and Monthly Wages			
Total Daily Labor Wages	99.31	2.53	4055
	(142.01)	(3.47)	
Total Monthly Labor Wages	621.09	26.13	4053
	(2001.57)	(52.22)	
Panel C: Social Network Sample - Village Level			
SNM			
Number of households (Census)	119.04	-10.43	191
	[55.12]	(6.41)	
Number of surveyed households (SNM)	110.26	-11.54*	191
	[53.02]	(6.18)	
Distance to the bank branch, *kms*	2.31	-0.10	189
	[1.32]	(0.11)	

Note: ***, ** and * indicate significance at the 1%, 5%, and 10% levels respectively. Column [1] reports variable means of the control group, with standard errors in parentheses. Column [2] reports the OLS coefficient estimates associated with regressing each outcome on the treatment dummy (a dummy variable equal to 1 if the service area of the observation was in the treated group). Column [3] reports the number of observations. Panel A contains outcomes at the service area level from the 2001 Census of India. Panel B contains outcomes at the household level from our household survey. Panel C contains outcomes at the village level for villages included in the Social Network Mapping component. All regressions include pair fixed effects. Standard errors are clustered at the service area level. All Rs. amounts are top-coded at 3 standard deviations. Refer to the data appendix for variable definitions**.**

Table 26. Baseline Descriptive Statistics

	N [1]	Mean (SD) [2]
Population		
No of household members	4066	4.52 (1.87)
No of household members 18yo+	4066	3.14 (1.39)
Demographics		
Head of Household: Age	4066	46.68 (12.60)
Head of Household: Male	4066	0.72 (0.45)
Head of Household: Years of education	4066	7.49 (4.84)
Distance to Branch	4066	2.18 (1.61)
Occupation & Income		
Below Poverty Line	2727	0.43 (0.49)
Employed in Wage Labor (last 7 days)	4065	0.63 (0.48)
Self-Employed or Owns Business	4066	0.16 (0.37)
Earns Farming Income	4060	0.45 (0.50)
Wage Labor Income (weekly)	4066	841.00 (1804.47)
Business Income (30 days)	4066	2435.54 (13783.49)
Farming Income (last season)	4066	9651.90 (23244.58)
Wealth & Properties		
Owns Land	4064	0.55 (0.50)
Cultivated Area (acres)	4030	1.00 (2.02)
Formal & Informal Borrowing		
Number of Outstanding Formal Loans	4048	1.18 (1.48)
Formal Borrowed Amount (last 24 months)	4040	46566.11 (89655.46)
Number of Outstanding Informal Loans	4047	1.86 (2.07)
Informal Borrowed Amount (last 24 months)	4040	41326.46 (67145.76)
Saving & Insurance		
Probability of Having Any Savings Acct	4066	0.84 (0.37)
Amount in Savings Account(s)	3960	5424.46 (20403.18)
Probability of Giving Out Any Loans	4064	0.05 (0.22)
Shocks Experienced in the Last 12 Months		
Any Shock	4066	0.39 (0.49)
Serious Injury or Illness Shock	4066	0.21 (0.40)
Daily and Monthly Wages		
Total Daily Labor Wages	4055	99.31 (142.01)
Total Monthly Labor Wages	4053	621.09 (2001.57)

Note: Descriptive statistics for household level outcomes of the main household survey sample. Column 1 gives the number of observations in the sample, Column 2 the mean and (standard deviation). Refer to the data appendix for variable definitions.