Can Self-Help Groups Really Be "Self-Help"?

Brian Greaney, Joseph P. Kaboski and Eva Van Leemput*

November 24, 2015

Abstract

We provide an experimental and theoretical evaluation of a cost-reducing innovation in the delivery of "self-help group" microfinance services, in which privatized agents earn payments through membership fees for providing services. Under the status quo, agents are paid by an outside donor and offer members free services. In our multi-country randomized control trial we evaluate the change in this incentive scheme on agent behavior and performance, and on overall village-level outcomes. We find that privatized agents start groups, attract members, mobilize savings, and intermediate loans at similar levels after a year but at much lower costs to the NGO. At the village level, we find higher levels of borrowing, business-related savings, and investment in business. Examining mechanisms, we find that self-help groups serve more business-oriented clientele when facilitated by agents who face strong financial incentives.

JEL: O1,O12,O16

Keyword: microfinance, self-help groups, privatized delivery

Over the past several decades microfinance services have expanded tremendously in developing countries. An increasingly common method of providing access to microfinance to the "poorest of the poor" are self-help groups (SHGs). In their most common form, SHGs essentially act as tiny savings and loan cooperatives. Currently these SHGs reach an estimated 100 million clients and this number has grown dramatically in recent years; active plans will nearly double this number by 2017.[1] Groups are not fully

*Greaney: Yale University, brian.greaney@yale.edu; Kaboski: University of Notre Dame and NBER, 717 Flanner Hall, Notre Dame, IN 46556, jkaboski@nd.edu; Van Leemput: Board of Governors of the Federal Reserve System, eva.vanleemput@frb.org. Research funded by the Bill & Melinda Gates Foundation grant to the University of Chicago Consortium on Financial Systems and Poverty. We are thankful for comments received from the editor and three anonymous referees, as well as presentations at Boston College, BREAD/Federal Reserve Bank of Minnesota Conference, Clemson University, the NBER Summer Institute, New York University, and UCLA. We have benefited from help from many people at Catholic Relief Services, especially Marc Bavois and Mike Ferguson, and the work of excellent research assistants: Luke Chicoine and Katie Firth in data collection, and Melanie Brintnall in data analysis. The views expressed are those of the individual authors and do not necessarily reflect official positions of the Federal Reserve Bank of St. Louis, the Federal Reserve System, or the Board of Governors.

[1] The National Bank for Agriculture and Rural Development (NABARD) program in India alone has grown from 146,000 clients in 1997 to 49 million in 2010.

"self-help", though. Although all funds are raised *internally*, the groups generally depend on *outside* assistance from administrative agents in their founding and continued administration. This motivates an important question: Can cost reduction or recovery in the delivery of *self*-help microfinance programs be effective in making them more fully self-help? The issue is common for many aid programs concerned with scalabilty, financial sustainability, and other types of aid, but the answer is not obvious. Recent research has shown that small costs to clients greatly reduce both take-up and program effectiveness.[2] We provide a theory and evidence that microfinance is different, even in very poor populations: A cost-recovery approach can actually be effective for SHGs.

The paper examines an innovation to the provision of NGO-sponsored microfinance services in three East African countries: Kenya, Tanzania, and Uganda. The status quo delivery mechanism was a typical "continuous subsidy" program, in which the NGO would train agents and then continually pay them a wage for starting up a fixed number of SHGs and providing financial services. In contrast, the innovation cut off payments to these agents after training, forcing them to become private entrepreneurs who start up any number of SHGs and earn their remuneration from their members. The hope was to not only lower costs to the NGO but also to expand access to services. Some programs already follow such an approach. A major World Bank/Indian government initiative with a goal of reaching 70 million new households is an important example.[3] Hence, the types of program and innovation we study are both of great interest. We examine the impact of this delivery innovation using a randomized control trial and a theoretical model in which control areas received the status quo program, while treatment areas received the private entrepreneur innovation.

The results are powerful and encouraging for the prospects of self-help groups indeed being "self-help", in the sense of being financially independent. Our randomization allows us to estimate the causal impacts on outcomes. The number of groups started by the treated agents, who charge fees, are slower to grow initially. However, after one year they reach the same number of clients and have more-profitable groups than the control agents, who provide their services for free. Moreover, the privatization treatment improves the outcomes of clients along many dimensions, leading to higher levels of: savings from business activities; credit, especially to business owners; employees; and business investment. It leads to households spending

[2] See Kremer and Miguel (2007) for an example with deworming pills or Cohen and Dupas (2010)'s analysis of insecticide-treated bed nets. These health-related programs have positive externalities that finance does not.

[3] The Rural Poverty Reduction Program in Andhra Pradesh, India, was a nearly $300 million project between 2000 and 2009, which has trained 140,000 "community professionals" (privatized providers), and reached 9 million women through 630,000 SHGs. In 2010, it was expanded into a nationwide program, the National Rural Livelihoods Mission, which is spending a combined $5.1 billion from the Indian government and $1 billion from the World Bank over seven years (World Bank, 2007, 2012).

a higher fraction of their time on their business, and correspondingly less in agricultural activities. These impacts are witnessed despite the fact that clients must pay for services under the private entrepreneur model. (Point-estimates of impacts on household income and consumption are insignificant.) Finally, the composition of the entrepreneurs' clientele is very different; the clients of the treated agents are more business-oriented and have a larger demand for financial services ex ante. Thus, the cost-sharing treatment appears to aim the program toward agents who are more likely to use the financial services for business purposes–the traditional goal of microfinance. Thus, cost-sharing could have important distributional consequences for NGOs with objectives beyond average impact and financial sustainability.[4]

To understand these results, we develop a theory in which cost recovery via membership fees can actually help solve an adverse selection problem that can plague credit cooperatives, especially when services are freely provided. Indeed, varying membership fees within a village can outperform a single membership fee. We provide suggestive evidence supportive of the role of membership fees: fees are strongly associated with greater levels of intermediation and agents appear to target their fees, varying them substantially both within and across villages. In contrast, we find no evidence that agent behavior along other dimensions drives these results. Because our experiment lacks variation in fees that is independent of other incentives, however, our results cannot definitively prove the proposed mechanism.

The specific variety of SHGs that we evaluate empirically are called SILCs (Saving and Internal Lending Committees). SILCs are promoted by Catholic Relief Services (CRS), a major non-governmental development organization, and are representative of other similar SHG programs sponsored by other agencies in the developing world.[5] In practice, SILCs are small groups of 10 to 25 members that typically meet on a regular basis to (1) collect savings, (2) lend to members with interest, (3) maintain an emergency "safety net" fund, and (4) share profits from lending activity. They do not receive external financial resources, only assistance from the outside agents who found and help administer the groups. In this sense, they effectively operate as small, independent, quasi-formal, self-financing credit cooperatives.[6]

Our empirical findings come from a large multi-country randomized controlled trial involving 276 agents who started a total of over 5,700 groups serving over 100,000 members across 11 districts in Kenya, Tanzania, and Uganda. All agents underwent a training phase. Upon completion, agents in the randomized

[4]We note two considerations, however. First, the entire population is quite poor. Second, we do not find that the program increases food vulnerability or negative responses to adverse shocks.

[5]These agencies include CARE, OxFam, Plan, World Vision and perhaps most importantly, NABARD, a large government agency in India.

[6]The "self-help" goals of these groups are not limited to self-intermediation. (Indeed, mature groups in some regions actually leverage their funds through outside loans.) They are also intended to help local communities by building social capital, empowering women, and fostering improved collective action. These aspects–at least in the data we study–are relatively minor compared with the financial activities.

treatment areas immediately became "Private Service Providers" (PSPs): entrepreneurs who need to start new groups and charge fees to group members in order to receive remuneration. In the control areas, field agents (FAs) received wages from CRS for establishing and administering a set number of groups but they were not allowed to charge their clients. This randomization was performed at a geographic level, so that treated agents did not compete with agents in the control group. Several sources of data verify that the randomization was indeed random.

In the post-treatment data PSP treatment increases group profitability by approximately 50 percent after one year. After three months, a PSP works with 4 fewer groups and 78 fewer clients, on average, than a traditional FA, but by one year the differences become statistically insignificant. The total amount of savings, number of loans, total credit disbursed, and profits all show similar patterns: They start out lower but increase over time and after one year the difference is statistically insignificant. PSPs earn only one-sixth of what FAs earn over the first year, but their earnings increase over time and agent attrition is low (less than 2 percent in either group). Overall, PSPs are substantially more cost effective, reducing the costs of providing services by over 40 percent after two years.

When considering the innovation's benefits to households, the results are even more encouraging. Despite the fact that households actually pay for services, the PSP approach is significantly more effective in delivering outcomes promoted by microfinance. (We estimate these impacts at the village level, where assignment is random, without reference to membership which is clearly non-random). On an intent-to-treat, per-household basis, the PSP leads to nearly \$24 (or 50 percent) more credit, \$19 (or 90 percent) more business investment, 0.13 (110 percent) more employees hired, and 3 more hours per week (33 percent) in business. The point estimate of \$7 (5 percent) more savings, is insignificant, but the \$13 estimates for savings from business profits and savings for business purposes are significant.

Related Literature

This paper contributes to several strands of literature. First, we contribute to a literature on cost recovery in development programs. Previous research on health-related cost sharing (i.e., Kremer and Miguel (2007) for deworming pills, Cohen and Dupas (2010) for insecticide-treated bed nets, and the simulations of Kremer et al. (2011) for clean water sources) highlighted the problem of small costs lowering the number of clients served.[7] Morduch (1999) conjectured that an emphasis on cost recovery in microfinance would limit its ability to reach the poorest households. We find that although costs alter the type of clients, the

[7]Not all prior empirical evidence has been negative, however, even for services with a public-good aspect. In Argentina, Galiani et al. (2005) found that privatization of water supplies reduced child mortality, especially in poor areas.

number of members is unchanged (and the cost-savings itself is important to expanding programs elsewhere). Low take-up of health interventions can lead to lower impacts, even on remaining clients, because the interventions have positive externalities. In contrast, we find higher aggregate impacts, and our theory suggests that financial services are qualitatively different than health interventions along this dimension.

Second, there are different theories of microfinance and a burgeoning empirical literature that has yielded mixed results regarding its impacts. Some theories follow the traditional narrative by modeling credit that enables entrepreneurship, investment, and growth (e.g., (Ahlin and Jiang, 2008), (Buera et al., 2012)), while others emphasize consumption smoothing or simply borrowing to increase current consumption (e.g., (Kaboski and Townsend, 2011), (Fulford, 2011)). Empirically, although there is evidence of very high returns to capital for some entrepreneurs,[8] with a few exceptions, the impacts of microfinance on consumption and entrepreneurial activity have generally been small (e.g., (Banerjee et al., 2015), (Crépon et al., 2011), (Kaboski and Townsend, 2012) (Karlan and Zinman, 2010)). Still, Kaboski and Townsend (2005) found that program details matter for impacts, as was the case for two recent studies with sizable impacts on businesses: Attanasio et al. (2011) and Field et al. (2009), who find impacts for joint liability loans and loans with repayment grace periods, respectively. We show that the delivery mode and incentives faced by institutions can greatly alter the entrepreneurial impact of microfinance.

Third, a large theoretical literature has examined credit markets under asymmetric information, including the design of cooperatives and lending groups (e.g., Banerjee et al. (1994), Ahlin and Townsend (2007), Wang (2013)). Seminally, Stiglitz and Weiss (1981) and De Meza and Webb (1987) analyze the impact of adverse selection on the provision of credit. The former show how it could lead to underprovision when entrepreneurs vary in the dispersion of returns, and the latter show that when entrepreneurs differ in their expected returns, adverse selection could lead to overprovision. Although our agents differ in their expected returns, as in De Meza and Webb (1987), our model has no room for overinvestment because we have a supply of funds in equilibrium, with the value of all projects exceeding their opportunity costs because agents with low returns also have low opportunity costs.[9] Our contribution is to show how two-part pricing can mitigate adverse selection in such a setting.

Finally, there is a recent literature on SHGs. Two recent randomized control trials of CARE's VSLA (Village Saving and Loan Associations) program found significant positive short-run impacts on food consumption in Malawi (Ksoll et al., 2012), and consumption, financial services, and assets in Burundi

[8]For example, de Mel et al. (2008) finds returns of 55 to 63 percent annually, substantially greater than market interest rates.

[9]The membership fees we propose are distinct from collateral because they are sunk; i.e., they are not contingent on borrowing or repayment.

(Bundervoet, 2012). Evaluations of OxFam's SHG program are ongoing but have found fewer impacts.

The remainder of the paper is organized as follows. Section 1 describes the program, experiment, data, and methods. Section 2 presents the results and evidence of selection. Section 3 develops a simple theory of a credit cooperative and the potential impact of membership fees, and suggestive tests of the role of membership fees in the data. Section 4 concludes.

1 Program and Methods

This section describes the operation of the SHG programs we study. We then document the details of the experiment, our data, and our regression equations.

1.1 SILC Program and PSP Innovation

Recall that the SHGs promoted by Catholic Relief Services are called SILCs (savings and internal lending committees). A typical SILC is a group of between 10 and 25 members who meet regularly to save, lend to members, and maintain a social fund for emergencies. SILCs allow those with limited access to financial services to save and borrow in small amounts, while earning interest on savings and borrowing flexibly. SHGs have gained wide support among development organizations because, in contrast to many traditional microfinance institutions, they emphasize savings as well as credit. Research has shown that many people in developing countries lack adequate savings capabilities, and some even value savings accounts that pay negative interest (e.g., Dupas and Robinson (2012)).

The advantage over more formal financial institutions is that SILCs are formed and meet locally, allowing members to avoid transportation and transaction costs that are prohibitive for those who save and borrow small amounts. For SILCs in Kenya, Tanzania, and Uganda, meetings are generally weekly, with a median weekly deposit of $1.25. A typical loan would be $20 for 12 weeks at a 12-week interest rate of 10 percent. The loan would be uncollateralized except for the personal savings in the fund.[10] Funds accumulate through savings, interest on repaid loans, and fines for late payments/other violations. These funds are held centrally. The funds follow cycles that generally last one year. All loans must be repaid at the end of each cycle, and the total fund is then temporarily dissolved with payouts to members made in proportion to their total savings contributed over the cycle. For SILC, the timing of payouts is typically arranged to coincide with school fees, Christmas, or some other time when cash is needed.

SILCs offer greater flexibility than rotating savings and credit associations (ROSCAs), and the fact

[10]Not all funds are lent out as loans; a portion is retained as a social fund available for emergency loans.

that funds can accumulate in a SILC allows for some of its members to be net savers, while others are net borrowers. The greater flexibility also makes their management nontrivial. They require strict record keeping to keep track of savings, loans, loan payments of various amounts, and payouts due. They also require judgment regarding who should receive loans, how much they should receive, and how to set interest rates. Risks of default are also potentially greater, since some members may borrow disproportionately, and this magnifies the importance of decisions on membership. In contrast to ROSCAs, SILCs do not arise spontaneously. Given their complexity, the role of trained field agents in founding, administering, and training the members themselves is critical. The services provided by field agents to these groups include initial training and follow-up supervision in the areas of leadership and elections; savings, credit, and social fund policies and procedures; development of a constitution and by-laws; record-keeping; meeting procedures; and conflict resolution.

CRS has traditionally catalyzed this process by training field agents (FAs) to start SILC groups. FA trainees are recruited from the more educated segment of existing SILC members. They receive initial training, begin forming groups within a month, and then receive refresher training three additional times; they are also monitored by a supervisor over the course of a year.[11] During the training phase, agents are required to form 10 groups. At the end of the training phase, the agents take an exam; if they pass they are certified. FAs receive a monthly payment during the training phase ($48 in Kenya, $31.50 in Tanzania, and $50 in Uganda), but this payment increases after completion of the training phase (to $54, $59.50, and $65, respectively). The required stock of groups also increases by 10 additional groups, which they meet with regularly. Both during and after the training phase, agents must report quarterly summary accounting data for each group (e.g., group name, number of members, total loans, total credit, profits, payouts, defaults) following a standardized MIS system. Beyond this data collection, there is little additional oversight from CRS after the training phase.

CRS introduced the PSP delivery innovation into this existing SILC promotion program; in the new program, fully trained FAs are certified as such and transition to PSPs, private entrepreneurs who earn payment for their services from the SILC groups themselves rather than from CRS. PSPs negotiate their own payment from the SILC members, with the most common form of payment being a fixed fee per member collected at each meeting.[12] After certification, payments from CRS to PSPs are phased out linearly over four months (75 percent of the training payment in the first month, 50 percent in the second

[11]Monitoring is done by checking over the constitutions and record books and occasionally sitting in on meetings of SILC groups of trainee FAs (generally at least once a month, rotating groups).

[12]For those groups that charge fees, the median quarterly fee per member is $0.50, which amounts to about 3 percent of the median member's quarterly deposits.

month, etc.). CRS' goal with this innovation is to lower the resources needed to subsidize SILCs, thereby improving both the long-term sustainability of the groups and CRS' ability to expand the program. The initial implementation of this delivery model was a large-scale Gates Foundation–funded program that involved training close to 750 agents to found roughly 14,000 SILCs and reach nearly 300,000 members. The FAs were recruited in three waves over three years, as different local partners (typically Catholic dioceses) in different regions of Kenya, Tanzania, and Uganda enter the expansion.

1.2 Experimental Design

The research focuses on the outcomes of a randomized set of FAs/PSPs from the first two of these waves. Agents in the first wave were recruited and began training in January 2009. This first wave was certified between December 2009 and January 2010. Agents in the second wave were recruited in either October 2009 (Kenya and Tanzania) or January 2010 (Uganda).[13] They were certified the following year, in October 2010 and January 2011, respectively. The second wave of agents represented an expansion of the program to new areas. After certification, those agents randomized as FAs earned monthly payments mentioned previously, which were chosen to compare well with anticipated PSP earnings after certification, and were required to start or assist 10 additional groups. (Unfortunately, PSP earnings fell short of these anticipations, as discussed in Section 4.1.)

The research includes data from multiple regions across Kenya, Tanzania, and Uganda. Within each region, a local partner supervised the implementation of the program in conjunction with CRS and our research team.[14] The randomization was stratified by country and assignment was done on a geographical basis, with all agents within a given geographical entity receiving the same assignment (FA or PSP). Treatment was assigned at the subdistrict level, with 50 subdistricts assigned to be served by the traditional FA program, and 108 subdistricts by the new PSP model. CRS had the goal of moving fully to the PSP delivery model in order to reduce costs, and all agents were recruited under the auspices of the PSP program. Both the partner organizations and their agents were notified of the particular randomized assignment just prior to certification. FAs did not remain FAs beyond the 12-month experimental phase, and out of concern for human subjects, the FAs were informed that they would transition to PSP assignment after

[13]The original plan was for all three countries to begin in October 2009, but the partners in Uganda experienced operational delays.

[14]The first-wave partners operated in Mombasa and Malindi (Kenya) and Mwanzaa and Shinyanga (Tanzania). Within Kenya, the second wave included expansion into Mombasa and Malindi, as well as new partners in Eldoret and Homa Hills. In Tanzania, the second wave expanded into three existing areas and added a partner in Mbulu. The Ugandan sample, all second wave, included partners in Gulu, Kasese, Kyenjojo, and Lira.

12 months.[15] The geographical levels were chosen to ensure that FAs would not compete against PSPs: sublocations in Kenya, wards in Tanzania, and subcounties in Uganda. (Within any area, PSPs could and did compete amongst themselves, however, including charging different fees.) The randomization was stratified by partner, with relatively more PSP regions.[16]

From among the expansion agents who were recruited, the initial sample included all agents who had not yet reached the certification step at the time of the initial randomization. The original year-1 sample included 51 agents in Kenya and Tanzania. In Kenya, the stratified randomization yielded a total of 9 PSPs and 9 FAs spread across two partners, while in Tanzania there were 20 PSPs and 13 FAs spread across two partners. The year-2 sample included 225 agents from Kenya, Tanzania, and Uganda. In Kenya there were 71 PSPs and 24 FAs spread across four partners, in Tanzania there were 44 PSPs and 19 FAs spread across three partners, and in Uganda there were 41 PSPs and 26 FAs spread across four partners.[17]

One downside of the experiment is that it lacks a "true" control, in the sense of a set of villages receiving no SILCs whatsoever. Unfortunately, from an evaluation design perspective, CRS declined to create pure control groups. For that reason, we can only make statements about impacts of the PSP program relative to the FA variety, but we have no experimental evidence on absolute impacts.

1.3 Data

Data were collected from four sources. The first, the MIS system, collects book-keeping accounting data at the level of SILC group. These group-level data (collected quarterly) include total membership, savings, credit, losses, interest rates, profitability, and payouts, as well as agent name and village. In order to pool the data across countries, we use exchange rates to put currency values into dollar equivalents. We analyze these data at the level of the SILC groups, but we also aggregate to (1) the level of the FA/PSP agents who operate them and (2) the level of village and analyze at these levels.

[15] In principle, this might reduce the likelihood of measuring differences between PSPs and FAs, since FAs may have behaved like PSPs (e.g., targeted clients, offering better services) in anticipation of this transition. However, while we do find significant differences in outcomes and selection, we do not find differences in targeting or services (see Section 2.3). Instead, we attribute the differences to fees, which were definitely not charged by FAs in anticipation.

[16] Relatively more of the geographical regions were assigned to PSPs for two reasons. First, the PSP program is less costly for the NGO. Second, the expectation was that the variance in outcomes would be higher under the PSP program. The second wave added relatively more agents into the evaluation sample, but similar numbers of FAs were chosen across each sample in an attempt to spread the costs of randomization. Because the randomization was done at a geographical level, the ratios of FAs to PSPs are not necessarily consistent across partners or countries.

[17] The randomization does not contain all of the recruited agents, particularly in the first year, for several reasons. First, the randomized evaluation was introduced somewhat late in the process (late December 2009). For the first wave, some partners had already certified their trained FAs as PSPs, and these were naturally excluded. Second, a small number were lost due to death or failure of the certification test. Finally, the initial randomized sample contained 268 agents, but unfortunately the agents from two of the partners in Tanzania had to be dropped from the sample after the partners ignored randomization assignments. These partners constituted 6 FAs and 8 PSPs in the first wave (from just one partner) and 29 PSPs and 6 FAs in the second wave.

The second source of data, an agent-level survey, supplements the MIS with agent-level characteristics (e.g., age, education, languages, work and family background, importance of FA income, and labor) as well as a smaller set of questions (e.g., on targeting of groups, time spent with groups, and negotiation of payments) collected every six months; additional group-level data were collected every six months covering membership characteristics, delivery of services, and the compensation scheme. Unfortunately, response rates on this survey were relatively low, so the sample is not as large and may suffer from biases in response rates.

The third and fourth sources of data are based on a set of 192 randomly chosen villages. The villages were selected as follows. A subset of 192 agents was chosen among the full-year sample of 225 second-wave agents in April 2010.[18] During this time, the agents were all in their training phase and had yet to be notified of their random assignment. For each of the 192 agents chosen, a village was randomly selected from among the villages in which they operated at least one SILC. In May 2010, a key informant survey was administered to that village chief. This survey (the third source of data) collected data on village infrastructure and proximity to important institutions (schools, markets, health clinics, banks, etc.), chief occupations, history of shocks to the village, and, most importantly, a village census of households.

Our fourth source of data, a household survey, obtained representative village samples to enable a comparison of village means, without reference to membership in order to identify causal "intention to treat" impacts. The data were stratified over likely initial SILC members and non-members using the village census.[19] In June, July, and August of 2010, the baseline survey was conducted among 1,920 households in eastern Kenya, Tanzania, and Uganda, respectively. (One village in Uganda was inaccessible and could not be surveyed.) A resurvey of the same households was conducted in Kenya and Tanzania (along with the Ugandan village not surveyed in the baseline) in June and July of 2011, approximately nine months after the agents had received certification. Uganda was resurveyed in October of 2011 , also nine months after agents had received certification. The household survey contained detailed data on household composition, education, occupation and businesses, use of financial services (especially SILC), expenditures, income, response to shocks, and time use, as well as some gross measures of assets, indicators

[18] These agents were stratified across country (83 of 96 in Kenya, 47 of 63 in Tanzania, and 62 of 67 in Uganda), but they were otherwise chosen randomly.

[19] Village censuses were matched with a list of known SILC members in order to select a sample for the fourth data source, the household survey. These members were members of agent groups during the training phase. From the list of SILC and non-SILC members, a sample of five households with matched SILC members and five households with no matched SILC members were chosen with weights assigned appropriately based on their proportions in the matched village census list. For households with matched SILC members, the respondent is the SILC member, while for the others it was generally the spouse of the head of household (appropriate since SILC members are disproportionately women). See Section A.10 in the online appendix for more detailed information on the construction of the weights.

of female empowerment and community participation, and questions about risk-aversion and discounting. Table A.1 in the online appendix presents some summary statistics in the baseline for households with and without SILC members. Although the population is quite poor, SILC members tend to be somewhat better off on a number of dimensions. Naturally, membership itself is endogenous, so this member vs. non-member comparison cannot separate the roles of selection and impact.

The data are high quality, but measurement error is always a concern with household-level survey data in a developing country. Our working definition of "household" relies on self-identification and is based on joint concepts: both eating from the same pot and living in the same home or compound. Among the data collected, expenditures, time use, and income are the most difficult to measure. Our measures of income probably suffer from the most measurement difficulties.[20] We focus on the respondent's income since it is presumably better measured, and respondents (many of whom are SILC members) are more likely to be affected directly by SILC.

1.4 Empirical Methods

We use simple regression methods tailored toward the different data sets. We first present our methods for estimating impact and then discuss our verification of the randomization.

1.4.1 Measuring Impact

Our estimation approaches differ slightly depending on the data source.

Agent and Group Impacts

For the agent-level data, we use the following regression equations:

$$Y_{idnt} = \alpha_{dt} + X_i\beta + \gamma wave_i + \delta PSP_n + \varepsilon_{itdn} \quad (1)$$

$$Y_{idnt} = \alpha_{dt} + X_i\beta + \gamma wave_i + \sum_{s=1}^{4} \delta_s PSP_{ns} + \varepsilon_{itdn} \quad (2)$$

Here, Y_{idnt} represents the outcome for agent i in district d, subdistrict n at time t. The outcomes we examine from the MIS data are total members, savings, number of loans, value of loans, profits, and agent pay. Here we control for several things by adding district-time fixed effects, α_{dt}; a dummy for the wave of agent i, $wave_i$; and the above agent i characteristics (gender, age, schooling dummies, number of dependents, and number of children), X_i. The variable PSP_n is a dummy that is positive for households

[20] See Section A.10 in the online appendix for measurement of these variables.

in treatment villages during the four quarters of treatment, while PSP_{ns} is specific to quarter s. Given eight quarters of data, we look for both an overall effect, δ (equation (1)), and duration-specific treatment effects, δ_s (equation (2)), for each of the four treatment quarters.

For the group-level data, the data are no longer aggregated across agents. We use the identical regression, however, except i now represents group i. For these regressions, the standard errors on estimates are clustered by subdistrict.

Household-Level Data

For the household-level data, we simply have two cross-sections. Rather than first differencing, which could exacerbate measurement error, we simply add the baseline outcome variable as a control and estimate impact using the following regression equation:[21]

$$Y_{jdn} = \alpha_d + X_j\beta + \rho Y_{jdn,t-1} + \delta PSP_n + \varepsilon_{jdn}. \tag{3}$$

The outcomes Y_{jdn} for household j, living in district d and subdistrict n, depend on a district-specific fixed effect, the characteristics of the household X_j (gender; age and age-squared; schooling dummies; and the number of adult men, women, and children in the household), and the baseline value for the outcome $Y_{jdn,t-1}$. Again, δ is the measure of the treatment effect. For the household data, we cluster standard errors by subdistrict, the level of treatment. We have 147 subdistricts in the household data.

Here, the impact of treatment is evaluated at the village level, without reference to SILC membership. The primary reason for this is that SILC membership itself is naturally endogenous. A secondary reason is that the overall impact of SILC could involve spillover impacts (either positive or negative) on non-members.

In the results section, we focus exclusively on the estimates of δ and δ_s.[22]

1.4.2 Baseline Randomization

The above methods rely on the exogeneity of the PSP treatment, PSP, which ought to follow from our randomization. We verify the randomization using several methods.

First, using the baseline data, we verify that the randomization was successful in terms of observables. We do this using three data sets: the village-level key informant data, the agent-level data (both MIS and

[21] Ignoring the panel aspect of the data and simply using the endline data expands the sample somewhat and produces very similar results.

[22] Section A.2 in the online appendix provides full regression results for a sample of each of the agent, group, and household regressions.

agent characteristics), and the household data.

For the agent-level data, we focus on a simple regression on the data used for explanatory variables:

$$X_{i,n} = \alpha + \gamma wave_i + \delta PSP_n + \varepsilon_{i,n},$$

where i again indexes the agent and n indexes the subdistrict in which the agent operates. We present the results for our independent variables used below. We control for the wave using $wave_i$. PSP_n is a dummy for whether subdistrict n received the PSP program, so that $\delta = 0$ is the null for the test of random assignment. We cluster the standard errors by subdistrict, the level of randomization.

Table 1 shows the baseline estimates for agents operating in the treatment (PSP) and control (FA) areas. We see no significant differences in gender, age, languages spoken, or number of children or dependents across the two samples. We do, however, see a significantly higher fraction of PSPs receiving secondary education and a correspondingly lower fraction of PSPs with primary school completion as the highest schooling attained. We believe this to be a purely random result rather than a problem with the implementation of the randomization.

Table 1: Agent-Level Randomization Results

	Age	Gender	Primary	Primary Complete	Secondary	Tertiary	Languages	Children	Financial Dependents
PSP	-0.33	-0.06	0.00	-0.14***	0.12*	0.02	0.09	-0.33	-0.40
s.e.	(1.2)	(0.07)	(0.01)	(0.05)†	(0.07)	(0.05)	(0.08)	(0.36)	(0.6)
FA Mean	36	0.69	0.01	0.46	0.43	0.10	1.9	4.6	6.4
Obs.	223	227	226	226	226	226	227	227	226

***, **, and * indicate statistical significance at the 1%, 5%, and 10% confidence levels, respectively. †††, ††, and † indicate statistical significance with a Bonferroni correction at the 1%, 5%, and 10% confidence levels, respectively. The results are estimated coefficients for a regression of the stated outcome on a PSP dummy and the following controls: age, age squared, gender, dummies for schooling (i.e., primary completed, secondary, and tertiary with a baseline of less than primary complete), number of languages spoken, number of children, number of financial dependents, cohort, and location fixed effects. The regressions are weighted by sampling weights. Standard errors are robust and clustered by subdistrict.

For the household-level data, we use only the first wave, and data are weighted appropriately (to account for the stratified sampling across likely members and non-members). Hence, a simple mean comparison suffices:

$$X_{j,n} = \alpha + \delta PSP_n + \varepsilon_{j,n}.$$

Here, j indexes household j, and the null of $\delta = 0$ is again the test for random assignment.

The top panel of Table 2 shows similar results for the household characteristics. Again, the assignment

Table 2: Household-Level Randomization Results - Baseline Demographics and Outcomes

	PSP			FA			PSP-FA
Demographics/ Controls	Mean	Std. Dev.	Obs.	Mean	Std. Dev.	Obs.	Mean △
Age	43	14	1362	42	13	536	0.31
Age Squared	2014	1284	1362	1986	1319	536	29
Gender	0.61	0.49	1363	0.58	0.49	534	0.03
# Adult Men	1.6	1.1	1380	1.5	1.1	539	0.06
# Adult Women	1.5	0.92	1380	1.6	0.95	539	-0.05
# Kids	2.6	2.0	1380	2.7	1.9	539	-0.07
No Schooling	0.22	0.41	1363	0.19	0.40	534	0.02
Some Primary	0.23	0.42	1363	0.21	0.41	534	0.02
Primary Completed	0.41	0.49	1363	0.49	0.50	534	-0.08*** †
Secondary	0.11	0.32	1363	0.09	0.29	534	0.02
Tertiary	0.03	0.16	1363	0.02	0.13	534	0.01
Outcomes (measured pre-treatment)	Mean	Std. Dev.	Obs.	Mean	Std. Dev.	Obs.	△ Coeff.
Total Savings	137	298	1380	137	293	539	-2.5
Savings for Business Owners	156	283	628	156	327	248	-8.6
Savings from Business Profits	30	169	1380	35	204	539	-5.1
Savings from Agric. Profits	30	215	1380	25	118	539	3.2
Savings from Salary/wage	17	104	1380	16	154	539	-4.2
Savings used for New Agric. Activity	37	223	1380	40	173	539	-6.6
Savings used for New Non-Agric. Activity	8.1	125	1380	5.4	31	539	3.9
Savings used for Existing Business	22	201	1380	16	187	539	2.1
Total Credit	48	214	1380	42	274	539	3.0
Credit for Business Owners	60	229	628	41	110	248	14
Credit from SILC	4.1	17	1380	3.7	18	539	0.48
Credit from Formal Lenders	35	210	1380	26	270	539	5.8
Credit from Informal Lenders	9.0	26	1380	12	52	539	-3.3
Credit used for Agric. Activity	13	131	1380	6.8	56	539	4.8
Credit used to Expand Business	13	126	1380	6.2	46	539	6.6
Credit used to Start New Business	1.5	19	1380	1.4	17	539	0.14
Start New Business	0.27	0.44	1380	0.25	0.43	539	0.03
Business Investment	39	147	1380	42	144	539	-2.8
Hours spent in Business	15	19	1380	15	19	539	0.53
Non-HH Employees	0.29	2.0	1380	0.42	4.4	539	-0.12
Hours spent as Employee	16	17	1380	15	18	539	0.60
Agric. Investment	55	454	1380	48	143	539	5.8
Hours spent in Agric.	27	15	1380	27	15	539	-0.47

***, **, and * indicate statistical significance at the 1%, 5%, and 10% confidence levels, respectively. †††, ††, and † indicate statistical significance with a Bonferroni correction at the 1%, 5%, and 10% confidence levels, respectively. The top panel presents the household randomization results for the demographic controls and the bottom panel presents the outcome variables. The last column of the bottom panel shows the estimated coefficients from a regression of the stated outcome on a PSP dummy and the following controls: age, age squared, gender, number of men, woman and children in the household, dummies for schooling (i.e., some primary, primary completed, secondary, and tertiary with a baseline of no schooling). All regressions utilize sampling weights. After weighting, the sample is representative at the village level, including all households within FA or PSP villages irrespective of SILC membership. Standard errors are robust and clustered by subdistrict.

of treatment appears to have been random with respect to the underlying characteristics of households, with the exception of education. Here, we see that the fraction of people whose highest attainment is primary school completion is significantly lower (0.08), and some of this is because the fraction with some secondary schooling is somewhat higher (0.02). Again, we believe this education result to be purely random.

We perform several exercises to ensure that our results are not driven by the higher schooling of either the agents or recipients in PSP areas. First, we include dummies for highest education attained in all regressions. Of course, if there are also significant differences in unobservables, this would not be sufficient. Second, the significant difference in education is concentrated in districts served by two partners: the Archdiocese of Mombasa in Kenya and Tahea in Mwanzaa, Tanzania. All of our significant results are robust to dropping these two areas, as Tables A.13 to A.17 in Section A.8.1 of our online appendix shows. Third, we examine the impact of dividing the sample by average village education rather than PSP/FA treatment. We discuss those results below.

Finally, we verify that our outcomes from equation (3) do not show "impacts" in the baseline household data–i.e., prior to treatment. The bottom panel of Table 2 verifies this for the 27 outcome variables we examine. Again, we see that the differences across the control and treatment are small and insignificant. The only exception is income, which is substantially higher in the PSP villages. This is only significant at the 10 percent level, and, again, we believe it to be purely random.[23]

1.4.3 Reasons for Impact

Using multiple methods, we explore three potential explanations for the impacts: (1) improved member selection by agents or households, (2) improved effort by agents, and (3) improved effort by members.

For the first explanation, our methodological approach is to include an interaction between a dummy for the PSP treatment with baseline variables, which are exogenous with respect to the randomized treatment, in regressions explaining endline membership. That is, we run

$$M_{jvn} = \alpha_v + X_j\beta + \eta_1 Z_j^{baseline} + \eta_2 PSP_n Z_j^{baseline} + \varepsilon_{jn}, \tag{4}$$

where M_{jvn} is endline membership (of household j in village v and subdistrict n) and $Z_j^{baseline}$ indicates various baseline household characteristics (income; business income; a dummy for whether the household

[23] Randomization results for 23 village characteristics from the village key informant survey data also support that the randomization was indeed random (Table A.3 of the online appendix). These data include the village population, the presence of various infrastructure, services, and facilities in the village–including financial institutions–and whether the village had experienced various natural disasters in the past five years. The differences between the control FA villages and treatment PSP villages are all small and insignificant. The lone exception is animal disease within the past five years, which occurred in 41 percent of PSP villages but only 21 percent of FA villages, statistically significant at the 1 percent level.

had positive savings; a dummy for whether the household had positive hours in business; and dummies for whether the household's estimated linear discount factor and hyperbolic discount factor are above the median).[24] Note that the fixed effects α_v are village-specific, so that the coefficient of interest η_2 will be identified from within-village variation in membership.[25] The η_2 coefficient estimates differential selection in PSP villages, i.e., the extent to which endline membership is more closely related to $Z_j^{baseline}$ in villages with PSPs.

For the second explanation, the agent questionnaire gives several measures of agents' behavior, including how households were targeted for new groups (based on demand, need, proximity, local connections, etc.), and three measures of "effort": the frequency of services provided to the group, the type of services provided, and the distance traveled to the group. We examine these as outcome variables in the agent-level regression equations, equations (1) and (2) above. The only difference is that these data are only available every six months, so our time-specific estimates in equation (2) are semi-annual rather than quarterly.

For the third explanation, we have data on the total hours per week spent working from the household time-use data. Although admittedly limited, hours working does give us some information on the respondents' overall levels of effort.

2 Results

We evaluate the impacts of the PSP program on PSP agents and groups themselves first and then on households. Finally, we examine potential explanations for the differential impact of PSPs.

2.1 Impact on Agents and Groups

Table 3 presents the agent-level results for various measures. These coefficients can be interpreted as treatment effects on the agents and the overall level of services intermediated by them.[26] As we are performing multiple testing, we also include Bonferroni corrections at a statistical level of $\frac{\alpha}{m}$ where α and m represent the significance level and the number of regressions in each table, respectively. The first row presents the overall impact δ from equation (1). With the exception of agent payments, which are quarterly

[24] The hyperbolic (δ^{hyp}) and linear (β^{lin}) discount rates are estimated by using indifference valuations (V) between time 0 and time t using the following formulas:

$$V_0 = \delta^{hyper} \beta^t V_t.$$

We obtained two estimates using two sets of questions: (1) tradeoffs between 0 and 1 month together with 12 and 13 months and (2) 0 and 3 months together with 12 and 15 months. We used the average of the two estimates.

[25] Alternative regressions without these fixed effects yield very similar results.

[26] With respect to the clients themselves, these treatment effects on the agents could encompass both the selection of different clients and causal impact on the clients. This estimation does not distinguish between the two.

flows, the dependent variables are accumulated stocks. On average across the four quarters, PSPs start 2.6 fewer groups, reach 62 fewer clients, and earn \$150 less in payments per quarter, all of which are significant at the 1 percent level. Based on these numbers, one might be skeptical that the PSP program will expand SILC services as well as the FA program will.

The remaining rows, which present the duration-specific estimates of δ_s from equation (2), offer stronger insight, however: PSPs start off more slowly than FAs, but they improve over time. This may be demand driven, as the PSP service is not free, but it may also be a supply side strategy of PSPs as they, for example, attempt to learn about their markets slowly. PSPs do significantly worse over the first three quarters in starting groups, reaching members, and intermediating loans, but these differences narrow over time and by the fourth quarter of treatment are not statistically distinguishable. Thus, by the end of the year, PSPs seem to be providing levels of services comparable to the FAs'. Payment for PSPs remains lower than for FAs, however, with the gap in cumulative payments widening over time. Indeed, if we calculate the cumulative payment to PSPs at the end of the year, the average is \$550 less than the \$708 average cumulative earnings of FAs.[27]

The smaller effects on the level of services in the early years are driven entirely by the fact that PSPs have fewer groups. Indeed, running comparable regressions at the group level yields no statistically negative impacts of PSPs, even early on (Table A.7 of the online appendix). Over time, we see relative improvements in membership, savings, loans, and profits over time, even at the group level. By the fourth quarter, a PSP's typical individual group has more savings, more credit, higher profits, perhaps an extra member, and statistically indistinguishable agent pay per group. In relative terms, these group-level differences are considerable–nearly 50 percent higher than the control means for savings and profit and 25 percent higher for loan value.

In sum, PSPs appear to have slower starts, but within four quarters they appear to be statistically indistinguishable from FAs in terms of the number of groups they start, clients they reach, savings they mobilize, and credit their groups provide. They earn substantially less, especially starting out, but their groups are ultimately more profitable.

Nevertheless, given the PSP's substantially lower cost to the NGO relative to FAs, after only one year the PSP's costs per member reached are substantially lower than the FA's costs. In the training year, both FAs and PSPs earn an average of \$518. In the year after certification, FA costs amounts to \$714, while PSP costs are only the first-quarter phase-in value: \$65. Thus, over two years, the cost of a PSP is just

[27]Table A.18 in Section A.8.2 of the online appendix shows the unweighted regression in Table 3. The point estimates and significance levels are highly robust.

Table 3: PSP Impacts on Agent-Level Outcomes

	Groups	Members	Savings	Loans	Loan Value	Profit	Earnings
All Quarters	-2.6***	-62***	-1090	-35*	-1050	-400	-150***
s.e.	(0.93) †††	(24) †	(800)	(18)	(730)	(310)	(5.0) †††
Quarter 1	-3.6***	-78***	-1150*	-46**	-1280**	-290	-160***
s.e.	(0.91)†††	(25)††	(650)	(18)†	(650)	(230)	(5.2)†††
Quarter 2	-2.4***	-63***	-1740*	-45**	-1720*	-870*	-150***
s.e.	(0.91)†	(23)†	(940)	(19)	(920)	(490)	(6.4)†††
Quarter 3	-3.1***	-70***	-1320	-37*	-1300*	-540	-140***
s.e.	(1.1)††	(27)†	(870)	(20)	(750)	(380)	(6.3)†††
Quarter 4	-1.6	-40	-170	-12	55	130	-150***
s.e.	(1.3)	(29)	(1090)	(22)	(910)	(370)	(4.9)†††
FA Mean	20	430	7610	230	7100	2140	180
Obs.	865	865	865	865	865	865	865

***, **, and * indicate statistical significance at the 1%, 5%, and 10% confidence levels, respectively. †††, ††, and † indicate statistical significance with a Bonferroni correction at the 1%, 5%, and 10% confidence levels, respectively. The results are estimated coefficients for a regression of the stated group-level outcome on a PSP (the randomized treatment) or PSP*Quarter dummy and the following controls: age, age squared, gender, number of languages spoken, number of children, number of financial dependents, dummies for schooling (i.e., primary completed, secondary, and tertiary with a baseline of less than primary complete), cohort, and location-date fixed effects. The regression is weighted by sampling weights. Standard errors are robust and clustered by subdistrict.

about half (i.e., $(518+65)/(518+714)$) of the cost of an FA. Since the cost of additional years is zero for PSPs, these numbers will almost certainly continue to fall over time. Averaged over the course of the first year, PSPs reach 62, or about 15 percent, fewer members (368 vs. 430), so the per-member cost of the PSP program is just over half (55 percent) of what the FA program costs. Since PSPs reach similar numbers of members by the end of the year, the relative cost of PSPs per-members reached will also almost certainly fall over time. If PSPs continue to grow relative to FAs, then their relative cost could fall even more rapidly.

A legitimate concern might be PSP retention, given their much lower earnings. So far, however, the drop out rates are very low across the board–1.6 percent (3 of 185) for PSPs and 1.1 percent (1 of 91) for FAs. In fact, this may indicate not that PSP pay is too low, but rather that FAs were paid more than the minimum amount needed to retain them.

2.2 Impact on Households

Although the PSP program appears to be cost effective in reaching households and providing services, another important question is whether it has similar effects on the households it reaches. We now turn to the household data to evaluate the relative impacts of PSP-run SILCs on households.[28] We first examine savings, credit, and productive decisions before examining the overall impact on income and expenditures.

Table 4 presents the impact estimates of δ in equation (3) for savings and borrowing behavior. With respect to savings (top panel), we see no significant impact on overall aggregate savings, notably including no effect on the savings of business owners, but the reported source and use of savings are both impacted. The PSP program leads to an additional \$13 of savings (per household in the village) coming from business profits (significant at the 5 percent level), but it has no impact on the amount of savings coming from agriculture or wage income. Similarly, an additional \$13 per household was saved by households that report using savings for existing businesses, and this estimate is significant at the 1 percent level. These estimates are substantial in percentage terms, amounting to increases of over 80 percent and 300 percent, respectively, relative to the FA villages. Thus, PSPs seem to have important effects on reported business-oriented savings.

[28] Here we have the village-level analog to the interpretation caveat in footnote 26. The causal impacts on villages encompass both causal changes on household-level behavior (e.g., reason for saving) and changes in the relative importance of behavior (e.g., levels of saving) across households.

Table 4: PSP Impact on Endline Household Savings and Credit

			Source			Purpose		
PANEL I: Savings								
	Total	Business Owners	Business Profit	Sell Agric. Product	Salary or Wage	New Agric. Activity	New Non-Agric. Activity	Existing Business
PSP	7.0	-14	13**	-9.0	7.8	-6.3	-2.4	13***
s.e.	(13)	(17)	(5.6)	(8.9)	(5.5)	(11)	(2.7)	(3.8)†††
FA Mean	142	173	15	44	11	43	4.6	4.3
Sample Mean	144	161	22	37	14	36	2.8	14
Median	64	88	0	0	0	0	0	0
PANEL II: Credit								
	Total	Business Owners	SILC	Formal	Informal	Agric. Activity	Expanding Business	Start New Business
PSP	24**	21**	5.1**	11	9.1***	4.8**	8.7***	2.1
s.e.	(11)	(9.5)	(2.1)	(10)	(3.1)††	(2.4)	(3.1)††	(1.4)
FA Mean	44	34	7.1	25	11	4.3	3.1	1.5
Sample Mean	56	50	11	28	16	7.1	9.3	2.9
Median	12	15	0	0	0	0	0	0
Obs.	1731	779	1731	1731	1731	1731	1731	1731

***, **, and * indicate statistical significance at the 1%, 5%, and 10% confidence levels, respectively. †††, ††, and † indicate statistical significance with a Bonferroni correction at the 1%, 5%, and 10% confidence levels, respectively. The results are estimated "intent to treat" coefficients for a regression of the stated outcome on a PSP dummy (the randomized treatment), the baseline outcome and the following controls: age, age squared, gender, number of men, woman and children in the household, dummies for schooling (i.e., some primary, primary completed, secondary, and tertiary with a baseline of no schooling). The regressions are weighted by sampling weights. After weighting, the sample is representative at the village level, including all households within FA or PSP villages irrespective of SILC membership. Standard errors are robust and clustered by subdistrict.

Turning to credit (the bottom panel), we see that the PSP program led to substantially higher levels of borrowing. The estimate of $24 is an increase of almost 60 percent relative to the FA mean, and it is significant at a 5 percent level. The comparably-sized estimate of $21 for reported business owners amounts to an even larger percentage increase, which is significant at the 5 percent level despite the much smaller sample size. The additional credit does not come exclusively from SILC, although per-household levels of credit from SILCs are $5 higher in PSP villages (significant at the 5 percent level). We see that borrowing from informal sources actually plays a larger role, with a coefficient of $9 that is strongly significant at the 1 percent level.[29] The reported purpose for borrowing is also affected by the PSP program with an additional $5 of credit for agricultural activities and $9 of credit for existing businesses (significant at the

[29] While some studies have found that microfinance substitutes for or crowds out informal borrowing (e.g., (Banerjee et al., 2015)), other programs have led to increases (e.g., (Kaboski and Townsend, 2011), (Kaboski and Townsend, 2012)). It is possible that funds are either leveraged, relent to others, or that repayment itself leads people to borrow from additional sources.

5 and 1 percent level, respectively). Both of these are increases of over 100 percent. In contrast, there is no impact on credit for new businesses.[30]

Table 5 delves more deeply into the impact of PSPs on the productive decisions of households. In general, PSPs lead to relatively more positive impacts on business efforts but, if anything, fewer positive impacts on agriculture. We find no significant impact of the PSP program on new business starts (although the power of the test is clearly weak).[31] We do, however, find significant impacts on the intensive margins of business. Business investment rises by $19 per household in response to the PSP treatment. Thus, business investment under the PSP treatment is roughly twice its level under the FA control. Likewise, time spent in business is higher by three hours per week, a difference of about 33 percent relative to FA control villages. The number of non-household members employed by the households in the sample is low overall (0.20 per household), with most households employing no outside workers. Still, the coefficient of 0.13 employees per household, significant at a 5 percent level, represents an increase of over 100 percent relative to the FA level. We do not see a significant corresponding increase in the hours spent as an employee, however. Respondents may be less likely than other household members to work as employees. The point estimate is positive, but insignificant and small relative to the mean. Finally, we look at agricultural decisions. Although credit for stated agricultural activities had been positively impacted by PSPs, the relative effect on agriculture investment is insignificant, and agricultural investment remains substantially larger than business investment. The lack of an impact on agriculture may be due to the fact that loan duration was typically shorter than crop cycles. PSPs lead to fewer hours spent in agriculture relative to FAs, however. The coefficient of –3 (hours per week per respondent) nearly offsets the positive impact on hours spent in business.[32]

We have also examined two simple summary measures of welfare: income and expenditures. Unfortunately, however, the evidence is weak, and our experiment lacks power along these dimensions. The point estimates are substantial but not significant for total income, business income, total expenditures, and total consumption (roughly 25 percent increases in total income and business income, and roughly 5 percent in consumption and expenditures). Nonetheless, the fact that these estimates are not negative

[30]Table A.20 in Section A.8.2 of the online appendix shows the equivalent unweighted regressions from Table 4. On the savings side, we find similar point estimates on savings coming from business profits and savings going to an existing business although the former is now only significant at then 10 percent level. With regards to credit, we lose significance on credit for business owners, credit from SILC and credit going to agricultural activities.

[31]The insignificant point estimate would indicate that the fraction of households starting new businesses in PSP areas was 6 percentage points higher. The rates of business ownership and business starts are high in the data. In the endline sample, 42 percent of households own a business and 24 percent reported starting a new business in the past 12 months.

[32]Table A.21 in the online appendix shows the unweighted regressions for Table 5. We lose significance on business investment, the number of employees hired and hours spent in agriculture.

Table 5: PSP Impact on Endline Household Productive Decisions

	Start New Business	Closed Business	Business Investment	Hours spent in Business	Employees (non-HH)	Hours spent as Employee	Agric. Investment	Hours spent in Agric.
PSP	0.05	-0.16**	19***	2.9*	0.13**	1.4	-3.5	-3.3**
s.e.	(0.06)	(0.07)†	(6.1)††	(1.5)	(0.06)	(1.5)	(9.7)	(1.5)
FA Mean	0.21	0.66	22	10	0.11	14	71	32
Sample Mean	0.24	0.53	34	12	0.20	15	66	29
Median	0	1	0	0	0	12	28	30
Obs.	1731	779	1731	1731	1731	1731	1731	1731

***, **, and * indicate statistical significance at the 1%, 5%, and 10% confidence levels, respectively. †††, ††, and † indicate statistical significance with a Bonferroni correction at the 1%, 5%, and 10% confidence levels, respectively. The results are estimated "intent to treat" coefficients for a regression of the stated outcome on a PSP dummy (the randomized treatment), the baseline outcome and the following controls: age, age squared, gender, number of men, woman and children in the household, dummies for schooling (i.e., some primary, primary completed, secondary, and tertiary with a baseline of no schooling). The regressions are weighted by sampling weights. After weighting, the sample is representative at the village level, including all households within FA or PSP villages irrespective of SILC membership. Standard errors are robust and clustered by subdistrict.

might be at least somewhat surprising, since clients must pay for their services out of pocket.[33]

One concern we had was that the PSP-led groups might have a greater impact on business-oriented behavior, but that this may come at the expense of other potential benefits of the program, such as consumption smoothing (through risk-sharing, credit, or dissaving in response to shocks) or financing the purchase of durables. To evaluate this, we looked at several measures, including (1) the probability of "ever going to sleep hungry," (2) the probability of experiencing various adverse shocks, and (3) the "number of weeks until finances returned to normal" following an adverse shock.[34] The point estimates on PSP were generally positive, but none were significant. That is, we certainly do not see evidence that PSPs underperform along this dimension. Instead, we view the two programs as comparable along this front.

2.3 Reasons for Differential Impact

In this section, we explore potential mechanisms that may drive the observed impacts. We find strong suggestive evidence that PSP programs cater to different people, and very weak evidence for the effect of incentives on other dimensions.

Table 6 presents the results of endline membership regressions on baseline household characteristics and their interaction with the PSP treatment following equation (4). The top panel shows the results for

[33]The results are shown in Table A.8 in the online appendix.

[34]The one exception was the probability of a business failure. If business investments are risky as our theory assumes, this is understandable, since the PSPs are making more business investments.

regressions with village fixed effects that highlight within-village selection, while the results in the bottom panel combine both within- and across-village selection. Focusing on the top panel, the first row presents the estimates of the direct impact of the household characteristic ($\hat{\eta}_1$). It shows that in general, characteristics such as income, business income, positive hours working in business, positive savings, and discount rates do not strongly predict membership. The second panel shows the differential selection within PSP treated villages ($\hat{\eta}_1$). Here higher baseline incomes, higher baseline business incomes, spending time working in business, having positive levels of savings, and having higher hyperbolic discount factors (i.e., suffering less from hyperbolic discounting) were all associated with a higher probability of SILC membership in PSP villages. These impacts are both statistically and economically significant. The mean impacts on the probability of membership range from 0.05 (business income) to 0.09 (income and hyperbolic discounting). Finally, our five measures are not independent, so the last column shows the results for the first principle component across all five measures. It is, again, strongly significant and quantitatively important.

Table 6: Endline Membership Selection on Baseline Characteristics

	Income	Business Income	Positive Savings	Positive Hrs. in Business	Linear Disc. Factor, β	Hyperbolic Disc. Factor, δ	1^{st} Princ. Comp.
Characteristic	-5e-06	-2e-06	-0.005	0.01	-0.12	-0.06	-0.01
s.e.	(4e-06)	(1e-04)	(0.08)	(0.06)	(0.08)	(0.04)	(0.02)
PSP*Characteristic	2e-04***	3e-04*	0.23**	0.12*	0.07	0.17***	0.10***
s.e.	(6e-06)††	(2e-04)	(0.09)†	(0.07)	(0.9)	(0.06)	(0.03)††
Village FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes
PSP	-0.08	-0.05	-0.30***	-0.11*	-0.10**	-0.08*	-0.03
s.e.	(0.05)	(0.05)	(0.09)††	(0.07)	(0.05)	(0.05)	0.05
Characteristic	-8e-05*	-6e-05	-0.17*	0.02	-0.16**	-0.02	-0.02
s.e.	(4e-05)	(8e-05)	(0.09)	(0.06)	(0.07)	(0.06)	0.02
PSP*Characteristic	1e-04**	3e-04*	0.30***	0.15**	0.11	0.09	0.09***
s.e.	(6e-05)	(2e-04)	(0.10)††	(0.07)	(0.08)	(0.06)	(0.03)††
Village FE	No	No	No	No	No	No	No
Obs.	1718	1718	1718	1718	1718	1718	1718

***, **, and * indicate statistical significance at the 1%, 5%, and 10% confidence levels, respectively. †††, ††, and † indicate statistical significance with a Bonferroni correction at the 1%, 5%, and 10% confidence levels, respectively. The results are estimated coefficients for a regression of household endline membership on on the stated baseline outcome, their interaction effects with a PSP dummy (the randomized treatment) and the following controls: age, age squared, gender, number of men, woman and children in the household, dummies for schooling (i.e., some primary, primary completed, secondary, and tertiary with a baseline of no schooling). The regressions are weighted by sampling weights and are reported both with and without village fixed effects. After weighting, the sample is representative at the village level, including all households within FA or PSP villages irrespective of SILC membership. Standard errors are robust and clustered by subdistrict.

Table 6 represents only a selection of the most salient results from our selection analysis. We conducted a wide variety of alternative specifications and included many different selection variables (i.e., $Z_j^{baseline}$) such as hours spent in agriculture or wage labor, presence and levels of credit, levels of savings, business investment, expenditures, and consumption. We also ran separate regressions, replacing membership with dummies for leavers (baseline members but endline non-members) or joiners (baseline non-members but endline members) separately and running regressions without fixed effects. Significance varied considerably across these different exercises, but both the significant results and the insignificant point estimates overwhelmingly support the story of the results presented in Table 6. The PSP groups appear more attractive to agents who are wealthier (i.e., who have higher consumption, higher savings and credit) and more business-oriented (i.e., who have fewer hours in wage labor and agriculture and more investment). Finally, we note that the fact that the results are equally strong when village fixed effects are included suggests that the selection we uncover is driven by patterns within the village, not differential membership rates across villages.

Nonetheless, we investigate alternative hypotheses. The first alternative hypothesis is that PSPs behave differently in either their targeting or their effort. The top panel of Table 7 shows no evidence that PSPs behave differently in targeting services to villages closer to their home (proximity), villages with which they have existing connections, or villages with greater perceived need (potentially driven by altruism) or demand (presumably driven by profit). In the bottom panel, we find no evidence of greater effort as measured by distance traveled to SILCs or number of services provided. Indeed, the only significant estimates in Table 7 show that PSPs are more likely to work only part-time, less likely to work more than part-time, and more likely to meet with their groups at least biweekly. It is nonetheless possible that unobserved effort varied across the two regimes.

The second alternative hypothesis is that the clients themselves put forth more effort. To evaluate this hypothesis, we can only focus on the total time spent working per week. While the composition of hours was impacted by the PSP program (recall Table 5), the overall total number of hours was not significantly affected.

In sum, we find no evidence that either the agents or members worked harder in response to the incentives of the PSP program, but we do find substantial evidence that the PSPs provide services to a wealthier, more business-oriented population.

Table 7: Impact of PSP Treatment on Agent Effort/ Behavior

Targeting							
	Proximity	Connections	Need	Demand	Other		
All Quarters	0.02	-0.03	-0.04	-0.03	-0.01**		
s.e.	(0.06)	(0.05)	(0.04)	(0.04)	(0.00)		
Quarters 1 & 2	0.01	-0.04	0.01	0.02	-0.01		
s.e.	(0.05)	(0.06)	(0.05)	(0.06)	(0.01)		
Quarters 3 & 4	0.03	-0.03	-0.10*	-0.08	-0.01		
s.e.	(0.09)	(0.07)	(0.05)	(0.07)	(0.01)		
FA Mean	0.32	0.54	0.84	0.30	0.01		
Obs.	3816	3816	3816	3816	3816		
Effort / Work Time							
	Average Distance	Average # Services	Work l.t. Half Time	Work Half Time	Work g.t. Half Time	Work Full Time	Biweekly Meetings
All Quarters	0.43	-0.05	0.25***	-0.11	-0.17*	0.02	0.08**
s.e.	(0.63)	(0.27)	(0.08)††	(0.09)	(0.09)	(0.04)	(0.04)
Quarters 1 & 2	-0.38	0.11	0.31***	-0.14	-0.24*	0.06	0.04
s.e.	(0.67)	(0.29)	(0.10)††	(0.13)	(0.12)	(0.05)	(0.06)
Quarters 3 & 4	2.3*	-0.42	0.13*	-0.04	-0.01	-0.08	0.13***
s.e.	(1.2)	(0.48)	(0.07)	(0.06)	(0.11)	(0.10)	(0.05)†
FA Mean	4.5	2.8	0.04	0.21	0.68	0.06	0.41
Obs.	146	146	135	135	135	135	3792

***, **, and * indicate statistical significance at the 1%, 5%, and 10% confidence levels, respectively. †††, ††, and † indicate statistical significance with a Bonferroni correction at the 1%, 5%, and 10% confidence levels, respectively. The results are estimated coefficients for a regression of the stated agent-level outcome on a PSP (the randomized treatment) or PSP*Half dummy and the following controls: age, age squared, gender, number of languages spoken, number of children, number of financial dependents, dummies for schooling (i.e., primary completed, secondary, and tertiary with a baseline of less than primary complete), cohort, and location-date fixed effects. The regressions are weighted by country-specific sampling weights. Standard errors are robust and clustered by subdistrict.

3 The Role of Membership Fees

Given the evidence that membership selection is based on business-oriented characteristics, we develop a model to provide a potential mechanism. More specifically, in Section 3.1, we develop a stylized model of a credit cooperative operating in the face of adverse selection where low-quality members can drive out high-quality members, lowering entrepreneurial activity (and aggregate output). Section 3.2 formulates the theoretical role of membership fees, showing that they can potentially solve this adverse selection and thereby increase entrepreneurial activity (and aggregate output).[35] We then empirically test the model predictions in Section 3.3 and we find that both agent intermediation outcomes and household outcomes

[35]We present these results but relegate the technical derivations and proofs to our online appendix (see Section A.9).

are strongly associated with fees, consistent with the theory.[36]

3.1 Model

We develop a one-period static equilibrium designed to capture a group cycle. In practice, groups generally continue beyond the cycle, but neither members nor groups need to commit to future participation.[37] Since the outcomes of interest involve entrepreneurial activity–rather than consumption smoothing or risk-sharing, for example–we model agents with entrepreneurial projects who, for simplicity and clarity, are risk neutral. In the data, we measured selection based on income, business activity, savings, and other measures of patience. In the model, we capture this using differences in average productivity and default. Although neither are directly observed in our data, higher productivity implies high average income and business income in the model, and the first principal component of our measures in the data is strongly negatively correlated with business failure and negative income shocks.[38]

3.1.1 Environment

Consider an economy with two stochastic project technologies that differ in their scale and productivity. The small-scale project requires one unit of capital, which it transforms, when successful, into $\underline{A}$ units of output. A large-scale project transforms $k > 1$ units of capital into $\overline{A}k$ units of output. The large-scale project is more productive in that $\overline{A} > \underline{A}$.

There is a unit measure of individuals, who are each endowed with one unit of capital for operating the technologies. All individuals have access to the small-scale project, but each person has access to the more productive large-scale project only with probability π. The individuals are divided into two types, $i \in \{L, H\}$ that differ in their inherent probability of success. A fraction of individuals, θ_L, have a lower probability of success in production, p_L, while the remaining $1 - \theta_L$ individuals succeed with probability $p_H > p_L$, where $p_L,\, p_H \in (0, 1)$. We assume that $\theta_L p_L < (1 - \theta_L)\, p_H$, so that the total number of potential successful type-H people exceeds the number of successful type-L people. When individuals fail, production yields zero output. In order to yield interesting results regarding adverse selection, we make the stronger assumption that $p_H\underline{A} > p_L\overline{A}$. That is, we assume the expected payoff of type-H individuals with the

[36]Thus the impact of our model credit cooperative is related to the charging of fees. A smoking gun empirical test of the model would involve randomizing membership fees. Lacking such exogenous variation - the model was developed ex post to understand the impacts.

[37]Bond and Rai (2009) show how the repeated game nature of lending groups can give incentives for future repayment in a model with strategic repayment. We rule out strategic repayment.

[38]The model has direct implications for default rates and returns on savings as well. Default rates are notoriously difficult to measure empirically, but we do find higher profitability among the PSP groups.

small-scale project exceeds that of type-L individuals with the large-scale project.

Since the large-scale project is more productive, but no individual is endowed with enough capital to operate it, there is potential demand for intermediation services. We model the timing and operation of a credit and savings cooperative as follows. Individuals decide whether to become members of a cooperative before finding out whether they have access to the large-scale technology. Members of a credit cooperative deposit their capital into the cooperative with a promised gross return of R_D. Individuals then find out whether they have access to the large-scale technology, and decide of whether to borrow at a gross interest rate of R_B, which equates demand for loans with available deposits. The cooperative is able to effectively distinguish between individuals borrowing for large-scale production and those borrowing for small-scale production. The member's type is unknown to the cooperative when it makes loan decisions, however, so that all borrowers pay the same borrowing rate. Unsuccessful members default on their loans and also forfeit their savings. Successful members repay their loans and then receive the return, R_D, on their savings, which effectively dissolves the fund.

Finally, consider that there is a minimum intermediation cost of C needed to remunerate the agent administrating the cooperative, but we assume that this cost is paid by an outside organization until Section 3.2.

3.1.2 Individual Decisions

Individuals simply maximize their expected income. A member of the cooperative receives in expectation

$$p_i \left(\overline{A}k - R_B k + R_D\right) \tag{5}$$

if she runs a large-scale project;

$$p_i \left(\underline{A} - R_B + R_D\right) \tag{6}$$

if she runs a small-scale project; and

$$R_D \tag{7}$$

if she simply saves. An individual not in the cooperative can neither invest in the large-scale activity nor save, so she simply earns

$$p_i \underline{A}$$

which we refer to as her outside option.

3.1.3 Equilibrium

An equilibrium must satisfy three conditions: (1) individuals' choices regarding joining the cooperative–whether and which project to undertake–must be optimal; (2) given R_B and R_D, the cooperative must earn zero profits; and (3) the market for funds must clear.

Given individuals' optimization and R_D, the demand for credit (per member) is a step function, involving willingness-to-pay thresholds that can be solved by equating the value of borrowing and investing (i.e., (5) or (6) for the large- or small-scale projects, respectively) with the value of simply saving (i.e., (7)) and solving for R_B.[39] Type-H individuals with the large-scale opportunity have the highest willingness to pay ($\overline{R}_{BH}$) while type-L individuals financing the small-scale project have the lowest willingness to pay ($\underline{R}_{BL}$). The ordering of the other thresholds depends on R_D and other parameter values. We focus on the interesting case, in which type-L individuals have a higher willingness to pay even though their expected payout is lower (i.e., $p_H\underline{A} > p_L\overline{A}$) because this case can lead to adverse selection. This arises because Type-Ls fail more often, and so they have less to gain, but limited liability (from less than full collateral) can give them a higher willingness to pay. Moreover, to simplify the analysis without losing the interesting features of the equilibrium, we focus on instances when market clearing is simplified because the total demand for loans from those with large-scale projects equals the supply of savings ($\pi k \to 1$).[40]

Defining f_L as the fraction of members of the cooperative who are type L, we can derive a break even condition for the cooperative showing that the borrowing rate must exceed the savings rate because of incomplete repayment:

$$\phi(f_L)R_B \quad = \quad R_D. \tag{8}$$

where $\phi(f_L) \equiv \frac{p_{avg}(f_L)}{\pi p_{avg}(f_L)+(1-\pi)} < 1$ is the effective repayment rate given the partial collateral of savings. Note that ϕ is increasing in f_L, since Type-Ls fail, and therefore, default more often.

We define $B(f_L;\tilde{p}_i)$ as the type-i individual's net benefit of joining the cooperative–i.e., the difference

[39] The borrowing thresholds for type-i individuals for the large- and small-scale project, respectively, are

$$\begin{aligned}\overline{R}_{Bi} &= \overline{A} - \left(\frac{1-p_i}{p_i k}\right) R_D \\ \underline{R}_{Bi} &= \underline{A} - \left(\frac{1-p_i}{p_i}\right) R_D\end{aligned}$$

[40] Formally, we make the following parameter assumption:

$$\pi k = 1 + \varepsilon.$$

where epsilon is an arbitrarily small, positive number, and we analyze the model under the conditions $\lim_{\varepsilon\to 0^+} \pi\overline{k} + \varepsilon$ and $\lim_{\varepsilon\to 0^-} \pi\overline{k} + \varepsilon$.

between the expected incomes of members and nonmembers in the cooperative. Given our simplifications, we can easily derive

$$\begin{aligned} B(f_L;\tilde{p}_i) &= \pi\tilde{p}_i\left[\left[\overline{A}-\overline{R}_{BL}(f_L)\right]k+\overline{R}_{DL}(f_L)\right]+(1-\pi)\,\overline{R}_{DL}(f_L)-\tilde{p}_i\underline{A} \\ &= \tilde{p}_i\left(\overline{A}-\underline{A}\right)+\left[(1-\pi+\pi\tilde{p}_i)\,\phi-\tilde{p}_i\right]\overline{R}_{BL}(f_L), \end{aligned}$$

where $\tilde{p}_i$ indicates the success probability of the particular individual (see Section A.9 in the online appendix). There are two forces at work in this equation. One force, clearly seen in the first term, is the fact that the cooperative allows the more productive large-scale projects to be financed, which is always an advantage and becomes larger as the individual's probability of success increases. The second term captures the compositional force, which depends on the average success rate in the cooperative compared to the individual's own success rate. The smaller the average success rate, the larger the wedge between borrowing and savings rates. For type-H individuals, this force is (weakly) negative, while for type-L individuals it is (weakly) positive.

Examination of $B(f_L;\tilde{p}_i)$ leads to the major results formalized in the following proposition.

Proposition 1 *Given the assumptions above,*
(i) Type-L individuals always join, $B(f_L;\tilde{p}_L)>0$.
(ii) Intermediate values of π *exist at which type-*H *individuals won't join a cooperative of all type-L members,* $B(1;\tilde{p}_H)<0$*, although they benefit more than type-L do from joining a cooperative of all type-H members,* $B(0;\tilde{p}_H)>B(0;\tilde{p}_L)$.

Proof of the proposition is straightforward and given in Section A.9 in the online appendix, but we offer some simple intuition here. Type-L individuals can only do better by joining, since both of the above-mentioned forces are positive for them. A poor composition lowers the benefits of joining because higher default rates lower the savings rate relative to the borrowing rate (i.e., lower ϕ). This wedge matters more for type-H, however, since borrowers only pay the borrowing rate and earn the savings rate when successful, and they succeed more often. Moreover, the type-H individuals have a higher outside option, so they benefit less from joining a cooperative with poor composition. Finally, when the composition is good, type-H individuals can have more to gain from financing large projects, since they succeed more often and earn a premium over the deposit rate.[41]

[41] If π is too high, type-H will always join. If it is too low, their benefits of joining will not exceed Type-L's. The requirement for intermediate values of π underscores the fact that the results rely on individual's having uncertainty over being a net borrower or net saver. If the timing were such that individuals knew whether they had a large- scale project before joining, then type-H individuals with the large-scale project would always join.

The left panel of Figure 1 shows these results graphically for an intermediate value of $\pi \in (\underline{\pi}, \overline{\pi})$. In such a case, although type-L individuals always join, type-H join only if type-L are less than some $\hat{f}_L$, defined by the root $B(\hat{f}_L; \tilde{p}_H) = 0$. If the proportion of type-L individuals in the population is high enough, $\theta > \hat{f}_L$, then type-H individuals never join and the equilibrium (denoted f_L^E) is $f_L^E = 1$. That is, all type-L join, but no type-H do.[42]

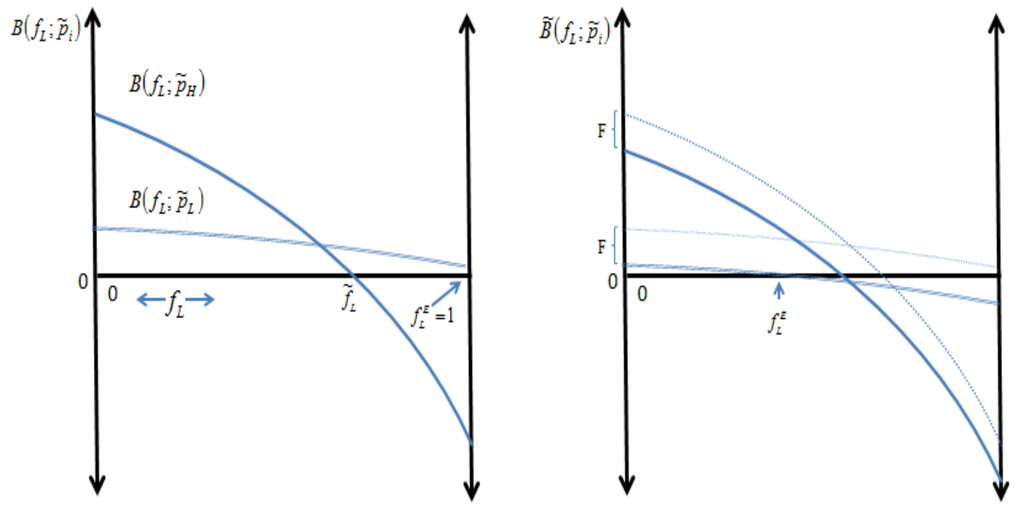

Figure 1: Benefits of Joining vs. Fraction Type-L

3.2 Theoretical Results on Membership Fees

Now consider the possibility of recouping the intermediation cost, C, by introducing a flat membership fee $F > 0$. We show how this could actually increase total output and the surplus of members.[43]

In the left panel of Figure 1, the benefit at lower levels of f_L is higher for type-H individuals. They therefore have a higher willingness to pay for membership in a cooperative with lower levels of type-H members. A membership fee has the potential of driving out type-L individuals, thereby inducing type-H individuals to join. Define $\tilde{B}(f_L; \tilde{p}_i) = B(f_L; \tilde{p}_i) - F$. The membership fee, F, can ensure that the intersection of $\tilde{B}(f_L; \tilde{p}_H) = \tilde{B}(f_L; \tilde{p}_L)$ is less than zero. If the relative benefits of type-H are high enough, this can actually increase average income even net of payments. In such a case, illustrated in the right panel

[42]If $\theta < \hat{f}_L$, multiple equilibria exist: two stable equilibria at either $f_L = 1$ or $f_L = \theta$, and an equilibrium, $f_L = \hat{f}_L$, that is unstable to perturbations around f_L. Of course, in all cases there are also additional trivial equilibria where no one joins.

[43]We do not include the cost, C, in the capital resource constraint of the model in order to maintain the simplicity of our stylized assumption that $\pi k - \varepsilon$ equals the amount of deposits available for loans. One could motivate these by an additional stylized assumption: introduce an initial endowment of $D > C$ output. We need to further assume that it cannot be used for investment, nor is it storable across the period of the model. Otherwise, the fund could demand this as collateral. We stress that entry costs differ from collateral in two important ways: (1) collateral is kept by the borrower in the case of repayment, and (2) for small π, entry costs are less than full collateral.

of Figure 1, the unique equilibrium value of f_L^E is at the point where type-L individuals are indifferent, $\tilde{B}(f_L; \tilde{p}_L) = 0$.[44]

We summarize this in the following proposition, and we give the details of parameter value requirements in Section A.9 of the online appendix.

Proposition 2 *There exist values of π, θ, and p_L satisfying Proposition 1 such that a membership fee, F, that induces some (or even all) type-L members not to join the cooperative will induce type-H members to join and increase the total income in the economy net of fees.*

Of course, if C is too large (that is, if it exceeds the potential benefits of type-H members, $(1 - \theta_L)\, B(0; \tilde{p}_H)$), requiring the cooperative to recoup costs through a flat membership fee would make the cooperative financially unsustainable.

Consider now the optimal policy, in the sense of maximizing total output. Total output is increasing in the number of individuals who finance the large-scale project, but the average output gain is larger for type-H individuals. The optimal single fee sets the intersection of $\tilde{B}(f_L; \tilde{p}_H) = \tilde{B}(f_L; \tilde{p}_L)$ in Figure 1 just below zero because it maximizes the number of type-L who enter, while also ensuring that all type-H enter. This fee leaves the members with no surplus, however.

Alternatively, we can solve for the equilibrium that maximizes total surplus. Under the assumption that $\theta_L < \frac{p_H}{p_L + p_H}$, type-$H$ members joining adds more to the surplus than type-L members joining. In this case, the fee that maximizes total surplus to members is the one that maximizes the surplus of type-H members. This is the lowest fee that keeps type-L members out; that is, it solves $\tilde{B}(0; \tilde{p}_L) = 0$. Call this F^*. The loss of type-L members does not lower the surplus because for any membership fee equilibrium in which type-H join, the surplus of type-L members is zero.

Finally, consider more flexible contracts that can achieve the first-best in the sense of maximizing total output by having everyone join the cooperative. A cooperative could effectively screen by offering two different contracts, $\{F, \phi\}$, which have the flavor of two-part tariffs. The cooperative can attract both types by offering a large F together with a large ϕ, which is attractive to type-H, and a small F together with a small ϕ, which is attractive to type-L.[45] Naturally, the output (net of fees) would be maximized

[44] Since total output is increasing in the number of agents who finance the large-scale project, the single fee that maximizes total output sets the $\tilde{B}(f_L; \tilde{p}_H) = \tilde{B}(f_L; \tilde{p}_L)$ just below zero. This maximizes the number of type-L who enter, while ensuring that all type-H enter. This leaves the members with no surplus, however. Since type-L members will never earn any surplus with the membership fee, the single fee that maximizes total surplus to members is the one that makes type-L members indifferent at $f_L = 0$.

[45] There are many such contracts that would accomplish this. There is also the possibility of adjusting R_B away from $\overline{R}_{BL}$, which we have focused on. In particular, any $R_B \in \left[\underline{R}_{BH}, \overline{R}_{BH}\right]$ for the first contract and $R_B \in \left[\underline{R}_{BL}, \overline{R}_{BL}\right]$ for the second would accomplish this. Since individuals are risk neutral, this only affects ex post inequality, not their ex ante valuation.

since all individuals would be in the cooperative.

A similar equilibrium, where total output is maximized and everyone joins the cooperative, could also be achieved by starting two different cooperatives with different membership fees. The contracts that maximize the member surplus in the cooperative attracting type-H members would charge $F = F^*$ (and have $\phi(0)$ as an equilibrium, break-even value), and the contract maximizing member surplus in the cooperative attracting type-L would have $F = 0$ (and have $\phi(1)$ as an equilibrium, break-even value).

Denote output (net of fees) under this equilibrium with two different cooperatives as Y_2^*, output (net of fees) under the single F^* fee equilibrium as Y_1^*, and output under no fees as Y_0^*; the following proposition summarizes how the benefits of the program vary with fee structure in these three examples.

Proposition 3 *For values of π and θ satisfying Proposition 2, the maximum output under two fees exceeds that under the single fee, F^*. Likewise, the maximum output under a single fee, F^*, exceeds that under no fee ($Y_2^* > Y_1^* > Y_0^*$).*

Propositions 2 and 3 motivate a simple test in Section 3.3.

Finally, we note that while the varying membership fees could potentially increase the total surplus of individuals, by including both types, the true social surplus would be net of the cost of financial intermediation, C. For a very high C that exceeds the benefit of serving the type-H population, $(1 - \theta_L)\, p_H(\overline{A} - \underline{A})$, social surplus is maximized with no cooperatives and no members. For a very low intermediation cost, C, that is less than the benefits of serving the type-L population, $\theta_L p_L(\overline{A} - \underline{A})$, social surplus is maximized with two cooperatives and everyone served. For intermediate values of C, social surplus is maximized with only one cooperative serving the type-H individuals.

One might thus interpret the model as illustrating a rationale for potentially excluding the poorest of the poor from microfinance: The benefits of their receiving microfinance do not exceed the costs, and their participation may actually drive out potential recipients who would benefit more.

3.3 Empirical Evidence on Membership Fees

The model suggests a strong role for membership fees in leading to the relative impacts of PSPs. In this section, we examine this empirically by focusing directly on these fees.

Our agent-level data contains information on PSP payments by the groups themselves. Although these payments are not randomized and so are potentially endogenous, our model suggests that variation could be driven by differences in intermediation costs (C) across villages which could reflect actual time and labor costs or the time and labor costs net of any altruistic motive (e.g., family, friends). Variation in either could plausibly be exogenous.

We document substantial fee variation across groups.[46] The quarterly per-group fee is \$5.80, on average; the standard deviation is 4.60; and the interquartile ratio is 3.1. Moreover, we find that 56 percent of fee variation occurs across villages, with just under half (44 percent) occurring within villages. Groups within the same village can be charged different fees, either by a single PSP or by different PSPs. PSP-specific fixed effects explain less than 40 percent of variation, however. Thus, the data indicate a high degree of price targeting within and across villages and by individual PSPs.

We perform two analyses to test the role of membership fees. First, we evaluate the extent to which financial intermediation in the agent data is associated with the charging of fees. Second, we examine whether the household outcomes are driven by the fee structure of villages.

3.3.1 Agent Outcomes

The model has predictions for economies of households, which we interpret as villages. To pursue the analysis further, we therefore aggregate the agent-level data by village and distinguish between villages where no fees are charged (413 villages), villages where a single uniform fee is charged (367 villages), and villages with variable fees–i.e., different groups are charged different fees (424 villages). We run regressions of the form

$$\begin{aligned} Y_{vdt} &= \alpha_{dt} + \gamma wave_d + \varpi_1 NoFeePSP_v \\ &\quad + \varpi_2 UniformPSP_v + \varpi_3 VariablePSP_v + \varepsilon_{itdn,} \end{aligned} \tag{9}$$

where Y_{vdt} are the same MIS outcomes (total members, total groups, savings, number of loans, value of loans, profits, and agent pay) aggregated by village, v. First, we run this regression using the per-group averages. The role of fees in the theory suggests that for the per-group averages, both ϖ_2 and ϖ_3 should be bigger than ϖ_1. That is, fees should enable higher levels of services, except perhaps for membership. Recall that these coefficients are all relative to the FA villages.

The results are shown in Table 8. They show how the typical group varies by the fees charged in the village. Except for loans, the positive significant estimates are all concentrated on the villages that charge fees, especially those with variable fees. Thus, fees seem to be closely related to the level of services that individual groups provide. Villages with uniform fees have significantly higher membership, and those

[46]To match the timing of the household data, we use only those groups that charged fees in the fourth quarter of the randomization, and we trim the lower and upper 5 percent of outliers.

with variable fees have significantly higher credit, membership, savings, and profits. The groups in PSP villages where no fees are charged are not statistically distinguishable from groups in FA villages. These PSP villages may be a combination of villages where PSPs offer free services out of social connections or altruism and villages in which PSPs anticipate introducing fees at a later date. Altruism would be an exogenous source of variation in the context of the model, while delayed fees might be potentially endogenous if correlated with willingness to pay, for example.

Table 8: Effect of "Village Type" on Per-Group Outcomes

	Members	Savings	Loans	Loan Value	Profit	Obs.
No Fee	-0.71	26	0.70	19	22	413
s.e.	(0.63)	(45)	(1.3)	(35)	(16)	
Uniform Fee	1.6**	26	0.26	44	14	367
s.e.	(0.69)	(33)	(1.1)	(31)	(8.3)	
Variable Fee	1.6**	110**	3.3	100**	20*	424
s.e.	(0.74)	(55)	(2.0)	(46)	(10)	
FA Mean	21	250	8.0	200	50	
Obs.	1760	1760	1760	1760	1760	

***, **, and * indicate statistical significance at the 1%, 5%, and 10% confidence levels, respectively. †††, ††, and † indicate statistical significance with a Bonferroni correction at the 1%, 5%, and 10% confidence levels, respectively. The results are estimated coefficients for a regression of the stated outcome (aggregated from the MIS group-level data to the village level and divided by the number of groups operating in the village) on a village type dummy, location-date fixed effects, and the cohort of the agent working in that village. The omitted village type is village served by an FA. The regressions are weighted by sampling weights. Standard errors are robust and clustered by subdistrict.

Second, we run regression (9) using the aggregate village totals. The theory suggests that the ability to vary fees should allow for more intermediation through a greater number of groups. Thus, ϖ_3 should exceed ϖ_2, even for membership. Again, these coefficients are all relative to the FA villages.

Table 9 presents the results and highlights the role of variable fees. Here, the positive estimates are almost exclusively in the villages where variable fees are charged. Indeed, uniform fees are associated with fewer groups, members, and services than in FA villages, but variable fees are associated with more groups, members, savings, loans, credit, and profits. This is consistent with the theory that variable fees can cater to larger populations than uniform fees, yielding higher levels of intermediation and larger total impacts (recall Proposition 3).

Table 9: Effect of "Village Type" on Total Outcomes

	Groups	Members	Savings	Loans	Loan Value	Profit	Obs.
No Fee	-0.43**	-9.1**	77	-1.3	18	64	413
s.e.	(0.19)	(4.2)	(140)	(4.2)	(110)	(55)	
Uniform Fee	-1.1***	-20***	-200**	-8.8***	-200**	-45**	367
s.e.	(0.14)†††	(3.4)†††	(88)	(3.3)††	(80)†	(19)	
Variable Fee	2.0***	46***	1020***	33***	960***	180***	424
s.e.	(0.23)†††	(5.2)†††	(210)†††	(7.4)†††	(180)†††	(35)†††	
FA Mean	2.4	51	610	21	530	120	
Obs.	1760	1760	1760	1760	1760	1760	

***, **, and * indicate statistical significance at the 1%, 5%, and 10% confidence levels, respectively. †††, ††, and † indicate statistical significance with a Bonferroni correction at the 1%, 5%, and 10% confidence levels, respectively. The results are estimated coefficients for a regression of the stated outcome (aggregated from the MIS group-level data to the village level) on a village type dummy, location-date fixed effects, and the cohort of the agent working in that village. The omitted village type is village served by an FA. The regressions are weighted by sampling weights. Standard errors are robust and clustered by subdistrict.

3.3.2 Household Outcomes

Next, we examine whether the fee structure of villages appears to be related to household outcomes. Recall that we have household data for only a small subset of villages. Among these, only two villages charge a single uniform fee, so we cannot divide the sample into three groups. Instead, we simply distinguish between PSP villages where fees are charged (952 villages) and PSP villages where no fees are charged (318 villages).

We first examine whether the charging of fees is related to underlying baseline characteristics by running regressions of the form

$$Y_{jdn0} = \alpha_d + X_j\beta + \delta PSP_{fee} + \varepsilon_{jdn}$$

where the baseline (time=0) outcome for household j (living in district d and subdistrict n), Y_{jdn0} depends on a district-specific fixed effect, and the characteristics of the household, X_j (gender; age and age-squared; schooling dummies; and the number of adult men, women, and children in the household). PSP_{fee} is a dummy for PSP villages where a fee is charged. The coefficient δ is relative to those PSP villages where no fees are charged. We add the district-specific fixed effects to control for any regional heterogeneity, since our fee analysis cannot fall back on randomization for exogeneity.

Using the above equation, we examine 27 different outcomes.[47] For the vast majority of these, we find that the PSP villages are statistically indistinguishable from the baseline, with the exception that villages

[47]The results are reported in Tables A.9, A.10, and A.11 in Section A.6 of the online appendix.

where fees are charged significantly have higher savings from business profit, have more credit from SILC and informal sources, and spend more hours in businesses than FA villages in the baseline. Our household results nevertheless control for any baseline differences, but we interpret our results as evidence that fee variation is reasonably exogenous.

We then examine how the endline results in Tables 4 and 5 are concentrated among the different fee villages. We run regressions that account for any baseline differences and take the form

$$Y_{jdn1} = \alpha_d + X_j\beta + Y_{jdn0} + \delta_1 PSP_{no\,fee} + \delta_2 PSP_{fee} + \varepsilon_{jdn}$$

The endline outcomes Y_{jdn1} (for household j living in district d and subdistrict n), depend on a district-specific fixed effect; the baseline outcome, Y_{jdn0}; and the characteristics of the household, X_j (gender; age and age-squared; schooling dummies; and the number of adult men, women, and children in the household). $PSP_{no\,fee}$ is a dummy for PSP villages where no fee is charged, and PSP_{fee} is a dummy for PSP villages where a fee is charged. The coefficients δ_1 and δ_2 are all relative to the FA villages.

The results are reported in Tables 10 and 11, which are exact analogs to Tables 4 and 5. We find that the overall positive results from the PSP are driven by the PSP villages where groups are charged a fee. For example, in Table 10, villages with fees use more savings for existing business, have more credit going to business owners, obtain more credit from SILCs, and use more credit for agricultural activities and expanding existing business. The results in Table 11 are a bit more mixed, however. We find a more statistically significant increase in business investment in PSP villages with fees, but the point estimate is actually smaller. We find fewer hours spent in agriculture. However, we find lower rates of business closures in villages without fees, and more hours spent as employees.[48]

To summarize, we find that the relative gains in intermediation and household outcomes that we measure for PSPs generally appear to be linked to the charging of fees, though not in every analysis. Although fees (rather than other aspects of PSP behavior) seem to be closely related to impacts, the high level of variability in fees suggest that it might be difficult to replicate the results of the privatization scheme by using a centrally mandated uniform fee, for example. A randomization on the fees themselves would add greater insight into these questions, however.

[48]Though not reported in the tables, we actually find a marginally significant positive impact in total household consumption in villages where fees are charged. We also note, however, that selection itself seems to occur in PSP villages regardless of whether fees are charged. Indeed, the differences between the fee and no-fee villages are not statistically significant. This is, admittedly, puzzling evidence for our fee interpretation.

Table 10: PSP Impact on Endline Household Savings and Credit by Village Type

			Source			Purpose		
PANEL I: Savings								
	Total	Business Owners	Business Profit	Sell Agric. Product	Salary or Wage	New Agric. Activity	New Non-Agric. Activity	Existing Business
No Fee PSP	29	-1.8	8.8	-11	14	-0.44	1.0	6.9
s.e.	(18)	(32)	(7.0)	(12)	(9.3)	(15)	(3.3)	(4.3)
Fee PSP	-10	-25	12*	-3.1	3.4	-9.7	-3.6	15***
s.e.	(14)	(17)	(6.3)	(9.1)	(4.4)	(11)	(2.9)	(5.2)††
PANEL II: Credit								
	Total	Business Owners	SILC	Formal	Informal	Agric. Activity	Expanding Business	Start New Business
No Fee PSP	16	0.72	3.8	2.2	9.8	-0.83	7.0	6.1
s.e.	(19)	(26)	(2.9)	(15)	(7.5)	(4.3)	(7.5)	(4.0)
Fee PSP	15	21**	6.0***	4.1	5.8	4.8*	6.2**	0.87
s.e.	(11)	(10)	(2.3)†	(9.1)	(3.1)	(2.9)	(3.1)	(1.4)
Obs.	1731	779	1731	1731	1731	1731	1731	1731

***, **, and * indicate statistical significance at the 1%, 5%, and 10% confidence levels, respectively. †††, ††, and † indicate statistical significance with a Bonferroni correction at the 1%, 5%, and 10% confidence levels, respectively. The results are estimated "intent to treat" coefficients for a regression of the stated outcome on two dummy variables: PSP households with no fee charged and PSP households with either a uniform or variable fee charged, the baseline outcome and the following controls: age, age squared, gender, number of men, woman and children in the household, dummies for schooling (i.e., some primary, primary completed, secondary, and tertiary with a baseline of no schooling). The regressions are weighted by sampling weights. After weighting, the sample is representative at the village level, including all households within FA or PSP villages irrespective of SILC membership. Standard errors are robust and clustered by subdistrict. All regressions include country and district fixed effects.

Table 11: PSP Impact on Endline Household Productive Decisions

	Start New Business	Closed Business	Business Investment	Hours spent in Business	Employees (non-HH)	Hours spent as Employee	Agric. Investment	Hours spent in Agric.
No Fee PSP	-0.03	-0.25***	26*	2.9	0.15	3.8**	-8.6	-3.4**
s.e.	(0.06)	(0.10)†	(14)	(1.9)	(0.11)	(1.6)	(11)	(1.5)
Fee PSP	0.06	-0.09	16**	1.8	0.09	1.3	1.2	-5.5***
s.e.	(0.07)	(0.07)	(6.4)	(1.7)	(0.07)	(1.2)	(10)	(1.7)†††
Obs.	1731	1731	1731	1731	1731	1731	1731	1731

***, **, and * indicate statistical significance at the 1%, 5%, and 10% confidence levels, respectively. †††, ††, and † indicate statistical significance with a Bonferroni correction at the 1%, 5%, and 10% confidence levels, respectively. The results are estimated "intent to treat" coefficients for a regression of the stated outcome on two dummy variables: PSP households with no fee charged and PSP households with either a uniform or variable fee charged, the baseline outcome and the following controls: age, age squared, gender, number of men, woman and children in the household, dummies for schooling (i.e., some primary, primary completed, secondary, and tertiary with a baseline of no schooling). The regressions are weighted by sampling weights. After weighting, the sample is representative at the village level, including all households within FA or PSP villages irrespective of SILC membership. Standard errors are robust and clustered by subdistrict. All regressions include country and district fixed effects.

4 Conclusion

We have presented evidence from a randomized trial of an innovation for privatized entrepreneurs who earn remuneration by charging fees for membership to their clients. The somewhat surprising empirical results indicate that through this cost-saving innovation, these microfinance services can indeed be "self-help" in the sense that, after initial training, the group administration can be financed through client-based fees. Relative to the continuously NGO-subsidized model, the private entrepreneurs expanded services more slowly but ultimately reached similar numbers of people and provide similar levels of services. The privately provided groups also had relatively stronger impacts in terms of the narrative microfinance dimensions of business entrepreneurship and investment.

As an example of a successful cost-recovery innovation–successful in terms of its intended goal of enabling NGO resources to stretch further, reaching greater numbers of people–the reasons for success are important for current and future microfinance programs. It does not appear that it was driven by the increased effort from improved incentives toward agents or members putting forth greater effort. Instead, it appears to be driven by privatized entrepreneurs catering to a more business-oriented population, who are willing to pay fees. We have developed a theory of the role of fees in mitigating adverse selection into credit cooperatives. Consistent with the theory, membership fees seem to be tightly linked to impact.

On the pro-side, this may target the services toward those who most benefit from them, and, indeed, better targeting may help improve the functioning of the groups. On the other hand, though, aid programs may still be interested in reaching those with a lower willingness to pay, who may be the truly poorest. Such tradeoffs may be more broadly important in moves toward sustainability or privatization. The distribution of benefits across the population and subpopulations is therefore an ongoing project of further investigation. A larger question is whether it is advisable to provide microfinance services to all populations. PSPs might improve welfare by limiting credit access to hyperbolic discounters, for example. Conversely, targeting populations that are more business-oriented may be riskier, potentially leading to future microfinance runs (Bond and Rai, 2009).

Another remaining question is whether privatization matters beyond the incentives to charge fees that it provides. If not, as the theory suggests, the favorable outcome and reduced costs could be attained by NGOs simply charging membership fees without privatizing. If the incentives do matter, then heterogeneous responses of PSPs to these incentives may provide insights. Unfortunately, our current data do not offer exogenous variation in payments or PSP behavior to further evaluate these issues. Experimental evidence on fees is an area for further research, as is examining the generality of these results for the provision of other financial services to the poor.

References

Ahlin, C. and N. Jiang (2008): "Can Micro-Credit Bring Development?" *Journal of Development Economics*, 86, 1–21.

Ahlin, C. and R. M. Townsend (2007): "Using Repayment Data to Test Across Models of Joint Liability Lending," *Economic Journal*, 117, F11–51.

Attanasio, O., B. Augsburg, R. De Haas, E. Fitzsimons, and H. Harmgart (2011): "Group Lending or Individual Lending? Evidence from a Randomised Field Experiment in Mongolia," Working Papers 136, European Bank for Reconstruction and Development.

Banerjee, A. V., T. J. Besley, and T. W. Guinnane (1994): "Thy Neighbor's Keeper: The Design of a Credit Cooperative with Theory and a Test," *Quarterly Journal of Economics*, 109, 491–515.

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

Bond, P. and A. S. Rai (2009): "Borrower runs," *Journal of Development Economics*, 88, 185–191.

Buera, F. J., J. P. Kaboski, and Y. Shin (2012): "The Macroeconomics of Microfinance." Tech. rep., National Bureau of Economic Research.

Bundervoet, T. (2012): "Small Wonders? A Randomized Controlled Trial of Village Savings and Loans Associations in Burundi." Manuscript, International Rescue Committee.

Cohen, J. and P. Dupas (2010): "Free Distribution or Cost-Sharing: Evidence from a Randomized Malaria Prevention Experiment," *Quarterly Journal of Economics*, 125, 1–45.

Crépon, B., F. Devoto, E. Duflo, and E. Parienté (2011): "Impact of Microcredit in Rural Areas of Morocco: Evidence from a Randomized Evaluation," Manuscript.

de Mel, S., D. McKenzie, and C. Woodruff (2008): "Returns to Capital in Microenterprises: Evidence from a Field Experiment," *Quarterly Journal of Economics*, 123, 1329–1372.

De Meza, D. and D. C. Webb (1987): "Too Much Investment: A Problem of Asymmetric Information," *Quarterly Journal of Economics*, 102, 281–292.

Dupas, P. and J. Robinson (2012): "Savings Constraints and Microenterprise Development: Evidence from a Field Experiment in Kenya." Tech. rep., UCLA.

Field, E., R. Pande, and J. Papp (2009): "Does Microfinance Repayment Flexibility Affect Entrepreneurial Behavior and Loan Default?" Manuscript, Harvard University.

Fulford, S. L. (2011): "Financial Access, Precaution, and Development: Theory and Evidence from India." Department of Economics Working Paper 741, Boston College.

Galiani, S., P. J. Gertler, and E. Schargrodksy (2005): "Water for Life: The Impact of the Privatization of Water Services on Child Mortality," *Journal of Political Economy*, 113, 83–120.

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.

——— (2011): "A Structural Evaluation of a Large-Scale Quasi-Experimental Microfinance Initiative," *Econometrica*, 79, 1357–1406.

——— (2012): "The Impact of Credit on Village Economies," *American Economic Journal: Applied Economics*, 4, 98–133.

Karlan, D. and J. Zinman (2010): "Expanding Microenterprise Credit Access: Using Randomized Supply Decisions to Estimate the Impacts in Manila," Manuscript, Yale University.

Kremer, M., J. Leino, E. Miguel, and A. Peterson Zwane (2011): "Spring Cleaning: Rural Water Impacts, Valuation, and Property Rights Institutions,," *The Quarterly Journal of Economics*, 126, 145–205.

Kremer, M. and E. Miguel (2007): "The Illusion of Sustainability," *The Quarterly Journal of Economics*, 122, 1007–1065.

Ksoll, C., H. B. Lilleor, J. H. Lonborg, and O. D. Rasmussen (2012): "The Impact of Community-Managed Microfinance in Rural Malawi. Evidence from a Cluster Randomized Control Trial." Manuscript, University of Southern Denmark.

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

Stiglitz, J. E. and A. M. Weiss (1981): "Credit Rationing in Markets with Imperfect Information," *American Economic Review*, 71, 393–410.

Wang, X. Y. W. (2013): "Risk, Incentives, and Contracting Relationships," Working paper.

World Bank, S. A. R. (2007): "SAR Regional Strategy Update," Tech. rep., World Bank.

——— (2012): "India's National Rural Livelihoods Mission an Overview," Tech. rep., World Bank.