This paper is a draft submission to the

WIDER Development Conference

# Human capital and growth

6-7 June 2016 Helsinki, Finland

# Investing in Boys and Girls: Schooling Decisions of Long-Run Microfinance Participants in Rural India*


**Jean-Marie Baland**
CRED, University of Namur

**Timothée Demont**†
Aix-Marseille University (Aix-Marseille School of Economics)

**Rohini Somanathan**
Delhi School of Economics


*DRAFT (this version: May 20, 2016)*


**Abstract**

This paper investigates the impact of the participation to self-managed microfinance groups, known as SHGs, on education outcomes in India. We analyze first-hand data collected from a panel of households in areas where new SHGs were formed in 2002. We observe these households up to 4 times over a seven-year period, which allows us to examine long-term effects of microfinance that are not observable in shorter panels. Combining propensity-score matching and panel regression techniques, we document important improvements over time in the enrollment rate of children beyond primary-school age, both at the village level (ITT estimates) and at the household level (ATT estimates), which are driven by a large reduction in drop-out rates beyond compulsory education. Notably, while we observe a significant gender gap at baseline, girls are as likely as boys to benefit from the program. Moreover, we find that SHGs are especially helpful at pushing up secondary enrollment rates in relatively poor and remote villages. Finally, we document a complementarity between children's school attendance and labor. Our analysis shows that human capital improvements take a long time to materialize, thus emphasizing the importance of long-term approaches to development interventions and their evaluation.



**Keywords:** Microfinance, Education, Child labor, Poverty.



*This research is part of the project (AdG-230290-SSD) supported by the European Research Council, as well as the project "Actors, Markets, and Institutions in Developing Countries: A micro-empirical approach" (AMID), a Marie Curie Initial Training Network (ITN) funded by the European Commission under its Seventh Framework Program - Contract Number 214705 PITN-GA-2008-214705. This work has benefited from the support of the Agence Nationale de la Recherche of the French government, through the program "Investissements d'avenir" (ANR-10-LABX-14-01).

†Corresponding author: timothee.demont@univ-amu.fr. We gratefully acknowledge helpful comments from Paul Reding, Xavier Giné, Catherine Guirkinger, Romain Houssa, and numerous participants to seminars at LSE, IIES, CRED, UCLouvain as well as to the 2014 Gender Workshop (Madrid), the 2014 AFSE meeting (Lyon), and the 2016 UNU-WIDER Conference on Human Capital and Growth. I also want to thank the NGO PRADAN (Delhi), and in particular Sanjay Prasad and Amit Kumar from the Data Research Unit, for outstanding assistance in data collection and cleaning.

**JEL Classification Numbers:** O15, G21, C33, R2

# 1 Introduction

Despite wide interest, the impact of financial services such as savings and credit on the economic and human development of poor users remains an open question to date (see Armendáriz and Morduch 2010, Kaboski and Townsend 2012, Karlan et al. 2013 or Duflo et al. 2015 for recent discussions). Summarizing the relatively large literature on the topic is however difficult because, one the one hand, there is a great diversity of microfinance institutions worldwide and effects are likely to vary by objectives and institutional type, and, on the other hand, some effects are hard to capture because of important time lags and heterogeneity. In particular, understanding whether and how microfinance can improve children's education in poor households is of crucial importance, given the widely-accepted links between human capital formation and long-term development outcomes such as poverty reduction and growth. Education beyond the elementary level is crucial to breaking intergenerational transmission of poverty. Unfortunately, access to secondary education has been lagging behind and remains far from universal in many developing countries. In India, secondary enrollment rates are estimated to be around 70% for the richest, urban households but as low as 30% for the poorest living in rural areas (World Bank, 2009). Moreover, a persistent concern in many developing countries has been the large gender bias in schooling performances, as emphasized by UN's fourth and fifth Sustainable Development Goals. In India, the problem is particularly acute and often attributed to the much larger financial returns that sons provide to parents as compared to daughters, due among others to the dowry system, the patrilocal family structure and the big differences in labor returns (e.g. Behrman, 1997; Rose, 2000; Kingdon, 2005, 2007; Zimmermann, 2012).

In this paper, we examine if long-run microfinance participants in one of the poorest Indian States changed their schooling decisions regarding boys and girls, and what were the mechanisms behind. Between 2002 and 2009, we collected detailed panel data about the living standards of members and nonmembers of informal credit and saving associations, known as Self-Help Groups (SHGs). While the non-randomized nature of the treatment does not allow straightforward causal inference, the richness and the long time span of the panel database permit the observation of a wide-ranging set of long-run evolutions and dynamic mechanisms at the individual, household and village levels. SHGs are small-scale village institutions that are engaged in a variety of collective activities, out of which savings and credit are the most important.[1] Bank-linked SHGs represent the dominant microfinance model in India, and the world largest. Yet, despite their very interesting characteristics (especially regarding sustainability and outreach), SHGs have not been largely studied so far.

There are good reasons to expect an impact of microfinance on schooling decisions among

[1] The model of bank-linked SHGs has been promoted by the National Bank for Agriculture and Rural Development since 1992 (the official guidelines are given by NABARD 1992 and RBI 1999). Section 2.2 presents the SHG program in more details.

member households. It can help affording education-related expenses by relaxing budget or liquidity constraints. It can modify households' preferences by empowering women and generating important peer effects. It can modify occupational choices and influence the opportunity cost of educating children. It can insure better households against income variation, thus decreasing the need to pull children out of school whenever shocks happen. Yet, the impact of microfinance on education outcomes is hard to measure, not only because of usual selection bias, but also because they can take a fair amount of time to materialize due to state dependence in school attendance, cohort effects and behavioral inertia. Solid empirical evidence on the matter is both meagre and mixed. Littlefield et al. (2003) review both quantitative and qualitative studies and conclude to a generally positive effect of microfinance on child schooling, though with important gender heterogeneity in some cases. However, none of the reviewed studies can reasonably be interpreted as causal estimates of the impact of microfinance. Recently, some randomized control trials have tried to tackle seriously the issue of selection bias. Though those studies were short-term and often lacked statistical power and external validity, they tend to indicate that microcredit has little impact on average education outcomes (see the summary of six recent RCTs by Banerjee et al. 2015). Augsburg et al. (2015) find even a reduction in school attendance for a subsample of marginal microcredit clients in Bosnia and Herzegovina with low education levels, due to a strong increase in the labor supply of young adults. By contrast, Angelucci et al. (2015) find evidence of positive community-level effects on school enrollment rates and negative effects on child labor in Mexico. Randomizing access to microbusiness loans at the individual level in Manila, Philippines, Karlan and Zinman (2010) find that (male) microentrepreneurs use loan proceeds to put their children at school instead of sending them to work (among a large set of other results). Randomizing the access to liquid savings accounts at the individual level in Nepal, Prina (2013) finds, one year after the intervention, a strong increase in household expenditures in education but no effect on enrollment rates. Non-randomized evidence is also mixed. Kaboski and Townsend (2012) use a large microcredit initiative by the Thai government as a natural experiment. They find that access to microcredit had a sizable effect on other consumption items but left education expenditures stable. By contrast, Maldonado and Gonzalez-Vega (2008) show that members of Bolivian MFIs are more likely to keep children in secondary school. Yet, they show that this probability decreases with agricultural land and argue that this reflects the fact that increasing opportunities to farm might increase the demand for child labor. Along the same line, Menon (2010) shows that credit for investment purposes does not increase the likelihood of going to school in Pakistan, if the children were already working part-time in a family business. Finally, in a study of two rural towns in Guatemala, Wydick (1999), while supporting positive effects from microloans on schooling, reports a statistically significant negative effect on schooling if the family enterprise produces a skill-demanding traditional cloth, such as a tailoring shop or a retail business.

This paper is one of the very few studies of the long-term effects of microfinance (other exceptions include Kaboski and Townsend 2012 and Berhane and Gardebroek 2011). Indeed, short-term effects tell little about long-term welfare. First, we know that development outcomes often take time to materialize, e.g. because behaviors need to adapt, investments need to mature before becoming profitable, or certain thresholds need to be crossed for effects to accelerate and eventually become sizable. Second, although transition paths are certainly relevant, it is the stable, equilibrium state that truly matters for evaluating the impact of microfinance. Some effects might be short-lived responses, ultimately reverting to the pre-intervention equilibrium, or to the contrary continue for a long time before they stabilize. Moreover, in many microfinance programs (and certainly in the informal groups we study), loan sizes gradually become larger and it is therefore natural to expect a lag between membership and the effects of credit. Third, microfinance certainly triggers general equilibrium effects, whereby credit, labor and/or good markets would need time to absorb the shocks. All of the above imply that aiming at the identification of long-run effects should rank high on the evaluation agenda, though it is extremely challenging (for instance, arguably no randomized experiment could ever maintain a pure control group long enough). Through to the collection of data on nonmembers both in treated and untreated villages, this is also one of the few papers studying explicitely the externalities of an intervention on non-participants (other examples include Angelucci and De Giorgi 2009; Flory 2012; Demont 2012).

Combining matching and panel-data techniques, we find a strong increase in children's school enrollment that appears in the last round of data, i.e. after six years of membership. Though there is some evidence that the process has probably started earlier, our finding implies that education outcomes and enrollment rates in particular can take a lot of time to materialize. Second, the detected effect is as strong for girls as for boys, despite the education gender gap observed at baseline. Third, the main channel explaining enrollment progress does not appear to be credit. By contrast, SHGs are especially helpful at pushing secondary enrollment rates in relatively poor and remote villages. We offer suggestive evidence that the increase in enrollment is at least partly driven by an increase in female awareness level.

The remaining of the paper is as follows. Section 2 presents the context of the study as well as the SHG program. Section 3 provides some background information about education in Jharkhand and about the potential role of SHGs therefore. Sections 4 and 5 presents the data and our empirical strategy. We then present results in section 6.1. Finally, we conclude.

# 2 The program and environment under study

## 2.1 Context

We study the introduction and the development of a large microfinance program in North India, between 2002 and 2009. The program was sponsored by a large development NGO called Professional Assistance for Development Action (PRADAN). The main stated objective of the organization is to promote and strengthen the livelihoods of socio-economically disadvantaged communities, such as indigenous people, women, scheduled castes, landless and the marginal and small cultivators. Central to this broad agenda is microfinance, which is considered as a means for the rural poor to make strategic investments in improving their livelihoods over time. Yet, unlike other microfinance models in which the NGO develops itself as the alternative credit provider, PRADAN organizes women in SHGs that become self-managed microfinance institutions. PRADAN started working with SHGs when the pilot program of NABARD was mainstreamed in 1996. By 2009, it was active in eight states of North India and had around 11,000 running SHGs.

This study focuses on the state of Jharkhand, which was carved out of Southern Bihar in 2000. It is among the poorest of the 27 Indian states, with more than half of its rural population living below the national poverty line (according to the latest official 2009 figures available). The UNDP's Multidimensional Poverty Headcount gives 75% of poor against a 54% national average. According to the 2011 census, its 66.4% average literacy rate is 8 points below the national average and entails an important gender bias, with 79.2% for males (82.1% nationally) and 52% for females (65.5% nationally). The state of Jharkhand is mostly rural (76% of its 33 million inhabitants) and its population consists of 26.2% tribals and 12.1% scheduled castes, which are known to be the most vulnerable groups of the Indian society. The present study focuses on villages only, which are very isolated on average. There, the main source of livelihood is subsistence agriculture and seasonal labor work. In its 2008 India State Hunger Index, the International Food Policy Research Institute (IFPRI) estimated that Jharkhand was suffering from the second highest level of hunger and malnutrition prevalence in India after only Madhya Pradesh - and higher than in Zimbabwe or Haiti (Menon et al., 2008). Historically, Jharkhand has been lacking lacking well performing local governmental and non-governmental organizations; today, it combines one the lowest SHGs to population ratio (less than 200 SHGs per 100,000 of population) with one of the highest percentage of poor population nationally (MicroSave, 2011).

Started in April 2002, PRADAN's intervention in 2009 was covering 12 out of the 24 administrative districts that constitute Jharkhand and had established microcredit groups in about two thousands villages in these areas or about 6% of all villages in the state (corresponding to over 60,000 women). An important aspect of PRADAN's strategy for expanding its activities has been to concentrate its programs in geographical clusters (targeting administrative blocks with high incidence of rural poverty), following a strategy of 'saturation' of the area. This strategy was chosen partly for administrative ease and economies of scale, but also to enable beneficiaries in different villages to interact and learn from their combined experience.

## 2.2 Self-Help Groups and PRADAN's intervention

The Reserve Bank of India gives the following definition:

> A Self-Help Group (SHG) is a registered or unregistered group of micro entrepreneurs having homogenous social and economic background, voluntarily coming together to save small amounts regularly, to mutually agree to contribute to a common fund

and to meet their emergency needs on mutual help basis. The group members use collective wisdom and peer pressure to ensure proper end-use of credit and timely repayment thereof. [2]

This section will describe in details how does this general definition translate into the environment under study. Establishing a group usually begins with a PRADAN representative holding a meeting at some public place in a village, such as the Panchayat office or the primary school, where the details of the program are described. Within geographical clusters around the local offices, PRADAN chooses to work with relatively disadvantaged communities and poor villages, where no other NGO has worked before. A study by CGAP found that PRADAN had indeed deeper-than-average outreach: almost all SHG members are tribal people or members of scheduled castes, 85% have no homestead land or only marginal nonagricultural landholdings and almost 90% live in thatched huts or are squatters (CGAP, 2007). After a few such meetings, a group of between 10 and 20 motivated women is formed[3]. One important rule suggested by PRADAN is that there may be only one member per household.[4] If a village is large, or interest in the program is widespread, multiple groups may be created. Moreover, new groups are sometimes created after a few years. After some initial training and capacity building from the NGO, the group chooses a name for itself, agrees on a weekly meeting time and determines other group rules, such as the minimum contributions per member at each meeting (usually 5 or 10 rupees - which amounts to about USD 0.5 - 1 per month), the interest rate charged on loans that are given to group members (usually 2% monthly), and fines for non-attendance or late payment.

After about half a year of smooth functioning, a savings account is opened at a commercial bank near the village to deposit group savings, and, usually after about one year, the groups showing mature financial behavior are enabled to take bank loans for a variety of income generating activities (the group is then said to be *linked*). At that point (i.e. after 1.5 to 2 years from the initial creation), groups are pretty much autonomous and the intervention of the NGO is only required to solve occasional problems. Rather, PRADAN then starts unrolling some livelihood programs, in which SHG members are trained to farming or self-employment activities of their

[2] http://www.rbi.org.in, website accessed in May 2011.

[3] Because Indian law requires that larger organizations be formally registered, PRADAN SHGs have never more than 20 members and are thus unregistered. Turning to the focus on women, it is interesting to note that the original NABARD SHG model has no built-in gender bias, although NABARD acknowledges a natural focus on women: "As regular meetings and savings are compulsory ingredients in the product design, it becomes more suitable for the women clients - as group formation and participatory meetings is a natural ally for the women to follow (NABARD website)". As a matter of fact, 90% of all SHGs which have borrowed from banks in India only have women members. PRADAN justifies its choice to focus exclusively on women both on equity and efficiency grounds, based on their long field experience: "[women] remain the most disadvantaged sector among the poor. Yet it is the women who prove to be most effective in fostering change in their families and communities (PRADAN website)".

[4] The operational definition of a household used by PRADAN is the set of people who live under the same roof ('choolah') and eat together. This is also the definition we use in our survey. Note that, in theory, people from the same households could join different groups, though this is very rare as it would impose serious burden on household chores during meeting times. Also, it should be emphasized that all 'rules' from PRADAN should rather be considered as indicative guidelines; ultimately, the groups decide on their own rules independently.

choice. Bank loans are always made to the group as a whole, without collateral and at market interest rates (fluctuating around 12% per annum).[5] The amount of the bank loan represents an increasing proportion of the group's savings (initially 1:1 or 2:1, it can go up to 10:1 after some years). Similarly, the repayment period is also gradually increased over time, from six months or one year initially up to three or even five years. Those dynamic features render particularly interesting and relevant the study of long-term impacts of SHGs, which is what this study aims at.

At a typical meeting, each member deposits the agreed minimum weekly savings or more, pays the interest on the loan she has taken (if any) and possibly pays back part of the principal. Interest earned on on internal loans remain within the group and becomes part of its corpus. Then, members who don't have a loan yet can require one to the group. Loans are individual but they have to be agreed on by the group and repayment is public. Moreover, there is a strong peer pressure ensuring due repayment, in order to preserve the group's resources. This being said, there is generally a lot of flexibility and understanding within the group when a member is not able to pay the weekly installment and asks, say, to pay double next time or when her cash flows become more favorable.[6] Finally, every year, the group agrees to distribute a portion of the realized profit to every member, through some kind of dividends that are proportional to savings. The remaining surplus is retained as a general reserve for adjustment against future losses or difficult contingencies.

As a conclusion, the bank-linked SHG model is a very decentralized, cheap and potentially sustainable way of providing access to small-scale savings and credit services in rural areas (not to mention other potential benefits from the group structure, such as peer support and other social services).[7] It is often thought as being well suited to progressively and safely bring the poor towards greater confidence, financial discipline, reduced vulnerability and smoother consumption profiles, access the formal financial sector, and eventually new livelihood activities able to improve

[5] Since 2000, SHG loans are considered as part of the priority sector lending of commercial banks, which encourages them to lend at relatively cheap rates. Indeed, the Reserve Bank of India holds that the interest rate charged to priority sectors should not exceed banks' prime lending rate (the rate at which the bank lends to its best customers). Even at those cheap interest rates, SHG lending is highly profitable for the banks, as transaction costs are low and the reported recovery rates are as high as 98% (Dave and Seibel, 2002). Moreover, commercial banks that lend to SHGs are eligible for subsidized refinance from NABARD. However, while this possibility was used extensively at start, "with increased bankers' confidence in SHG banking, banks started using their own resources and NABARD's refinancing dropped from 91 per cent in 1999 to 33 per cent in 2007 (Bali Swain, 2012)".

[6] A recent study by CGAP found that the average Portfolio at Risk > 90 days of PRADAN SHGs was over 20% (CGAP, 2007). They explain that, "although this level of loan delinquency would be disastrous for most microcredit providers, SHGs are surviving despite this. This has to do with the fact that a significant part of the SHG loans are used for crop cultivation and livestock rearing, neither of which offer a monthly cash flow. Yet, loan installments remain fixed at monthly [or even weekly] intervals, [...] sometimes out of a desire to keep a discipline of 'repaying something in each meeting'. Thus the high level of late repayments in SHGs does not always translate into defaults."

[7] CGAP (2007) estimated that the average cost of promoting and supporting SHGs in India is around 18 USD per group member (20 USD for PRADAN groups), and that the average return on assets after adjusting for loan loss provisions is around 9% (16% for PRADAN groups). Deducting the costs supported by the promoting NGO, SHGs break even on average, with an adjusted ROA of 0% (-1% for PRADAN groups). The study concludes that "The Indian SHG model can work sustainably in well-managed programs. Compared to other microfinance approaches, the SHG model seems to be producing more rapid outreach and lower cost."

their standard of living. This is at the core of PRADAN's strategy:

> Groups are useful to bring together poor people, build a synergy by pooling their energies and resources, promote values of mutual help and cooperation and also build economies of scale in providing inputs to the community. [...] The approach of PRADAN is twofold: making the SHGs creditworthy institutions, and developing a livelihoods focus in the groups so that the members utilize the credit available in setting up sustainable livelihoods. [...] PRADAN seeks to ultimately withdraw, as SHGs become viable and sustainable on their own as mutual support groups, successfully linked to financial service providers, and able to engage in a process of overcoming structural constraints and securing their rights to a sustainable livelihood (PRADAN, 2002).

Today, the bank-linked SHG model is considered as the largest and fastest-growing financial inclusion program in the developing world with, as of March 2012, almost 8 million bank-linked SHGs in India covering about 103 million families (NABARD, 2013)[8]. This represents a remarkable achievement, especially given the general acknowledgement that standard microfinance products remain more suited to urban and peri-urban areas than the to the rural world (e.g. Crépon et al., 2015; Angelucci et al., 2015). Yet, very little research has focused on the economic impacts of Indian SHGs.[9]

# 3 Education in Jharkhand and the potential role of SHGs

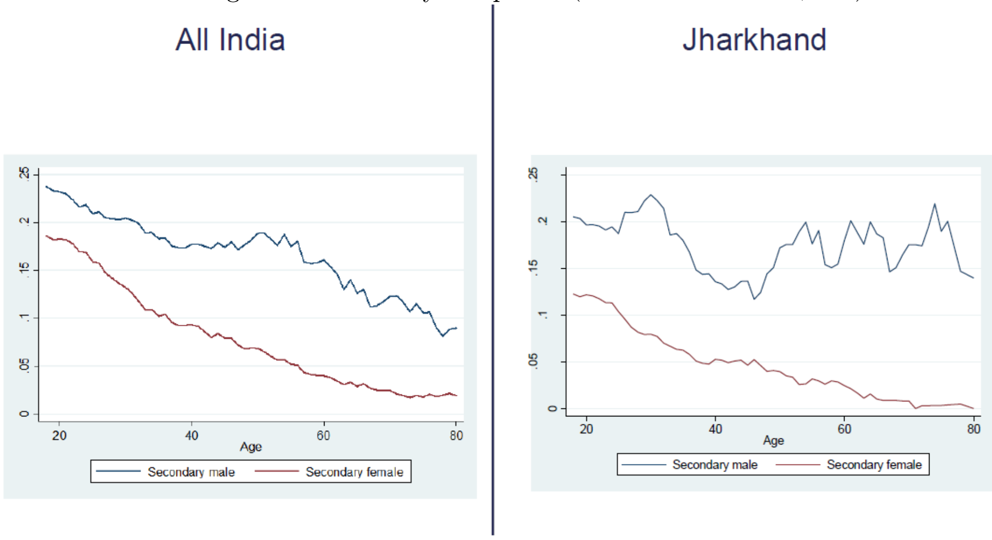
The education system in Jharkhand (as in most of India) follows a 5+3+2+2 schedule, the end of each step being sanctioned by an exam. Primary school (grades 1 to 5) starts at 6 years, followed by middle school (grades 6 to 8), lower secondary (grades 9 and 10) and higher secondary (grades 11 and 12). Average primary enrollment is 72.1% in Jharkhand, much lower than the national average at 83.3% (figures from DHS 2005-06). In addition, there is a very low transition rate into secondary school. As the graphs below make clear, Jharkhand experienced a slower increase of secondary enrollment and completion rates than the rest of the country, as well as a slower reduction of the gender gap. As a consequence, the state today has one of the lowest

[8]In Jharkhand, the number of bank-linked SHG is about 80,000. The success of the bank-linked SHG model has clearly been enabled by the dense network of commercial banks in rural India, which is largely due to the ambitious social banking program that created roughly 30,000 bank branches in unbanked rural locations between 1969 and 1990 (Burgess and Pande 2005 document extensively the program and its impact). Today, although it has actually decreased over the last decade, the number of rural branches of scheduled commercial bank is still at 20,773 (RBI 2010), and the number of branches per 1,000$km^2$ is one of the highest in developing countries (Burgess and Pande, 2005).

[9]One study in Andhra Pradesh compares newly-formed with mature groups and finds that longer-term exposure is associated with improvements in consumption and savings (Deininger and Liu, 2013). In Orissa, SHG-members are found to better coordinate in managing common pool resources (Casini and Vandewalle, 2011). In Rajasthan, Desai and Joshi (2013) assigned villages to SHG exposure and find that, after two years, treated women are more likely to save, work outside of agriculture, participate in household decisions and engage in civic activities. By comparing the impact on current borrowers vis-a-vis future self-selected borrowers in several states, Bali Swain and Varghese (2009) find that longer membership duration in SHGs positively impacts asset creation.

secondary enrollment rate of India, with a combined gross enrollment rate of about 18% and a persistent average gender gap of over 10 percentage points (World Bank, 2009).

Figure 1: Secondary completed (source: NSS 61st round, 2004)




The Indian Constitution directs the state to provide free and compulsory education for children until 14 years old (article 21A, added in 2002). Yet, this legal provision was translated into an executive act in 2009 only, with the Right of Children to Free an Compulsory Education Act (which came into effect on 1 April 2010). As a consequence, there was no major change in the legal environment during the seven years of our survey. However, various smaller and narrower schemes have always existed to provide different incentives to improve enrollment rates at the different education levels.[10] Given that they are always implemented at the district level, it is important to account for district fixed effects in the subsequent analysis.

Parents' decision to invest in the education of their children can be influenced both by altruistic and egoistic motives. Altruistic parents want the best for their child and choose education over work to secure the future of their offspring. On the other hand, egoistic parents might prefer to pay the upfront cost of schooling and reduce the current household income, in return for a higher income later on when the children are older and earn higher wages due to a better education. Note that this last motive is likely to entail important gender differences in the

[10]For instance, the Sarva Shiksha Abhiyan (Education for All) scheme was introduced in 2000-01 by the government of India to provide relevant elementary education for all children in the age group 6-14 (with a special focus on girls from 2003-04). It seeks to open new schools in those habitations without schooling facilities and to strengthen existing school infrastructure through provision of additional class rooms, teachers (and teacher training), toilets, drinking water, maintenance grant and school improvement grants. Another popular example is the Mid Day Meal scheme, which has been introduced as a centrally-sponsored program in 1995. It provides a free cooked meal for children enrolled in classes 1-5 of government schools (the scheme has been expanded in 2007 to include classes 6-8).

context under study, as girls usually leave the family household early to incorporate the in-laws' household after marriage. In some cases, however, in the absence of attractive schooling (e.g. low returns to education, poor infrastructure, social discrimination of backward classes, absenteeism of teachers, etc.), parents might prefer to send their children to work instead. Parents might also be deterred from sending their children to school due to the high costs associated with schooling. Direct costs can be a problem (e.g. school fees, expenditures for books, uniforms, transportation), in particular if they have to be paid in a lump sum. In addition, indirect costs are especially likely to be binding in poor households, as the opportunity cost of foregoing labor income or help from children is relatively more important in those households. For instance, older children (especially girls) are often required to look after their younger siblings and help with household chores, which might prevent them from attending school. For older children, their labor might be required in the household's business.

One way SHG membership can help educating children is by providing credit in order to pay for direct expenses. Because returns are realised many years after the period of investment, educating children is inherently an intertemporal problem, and parents' decision will depend upon their borrowing ability to borrow agianst future returns in order to finance investments in their children's early years (Rose, 2000). Nevertheless, direct costs are likely to be limited in our context, given that most schools in Jharkhand are government schools, which are supposed to be free up to the age of fourteen. Some evidence suggests, however, that this may not always be the case. Kingdon (2005) computes that school expenditures for a child, including tuition fees as well as expenditures on uniforms or stationary, are around 340 rupees in rural India per year, corresponding to more than 13 days of work for an agricultural laborer living in the neighboring state of Bihar. Transportation costs are also likely to be much larger at the secondary level, as, while almost all villages have a primary school, only 44.6% of Jharkhand villages have a lower secondary school within 5km of and 18.3% have a higher secondary school within 8km (Biswal, 2011). For instance, Muralidharan and Prakash (2013) show that a simple policy implemented by the state of Bihar, which provided girls who continued to secondary school with a bicycle, managed to increase girls' enrollment by 30%. In our sample, the median annual school expenses per child enrolled amounted to 480 INR in 2008, which represented between 9 and 10 days of work at the median wage rate or 3.3% of the median household income. When restricting the sample to households having at least one child enrolled in secondary school, the median school expenses per child enrolled becomes 1,747 INR (35 work days or 11.9% of median income), a very sizable amount. As a consequence, this direct cost channel can be important and will especially matter at the secondary (or post-compulsory) level.

The relaxation of credit constraints is not the only channel through which SHGs can affect education outcomes. If SHGs facilitate income generation in micro and small enterprises/farms, it can help to afford school expenses and potentially reduce the need to send out children to work in order to complement the household's income. Furthermore, if credit reduces the need for temporary migration in search of cash for adult members in the household, we might observe a decrease in child labor and potentially higher school enrollment and / or attendance. Perceived returns to education might also increase in the medium run if households switch to higher-return, larger-scale businesses or can afford to pay for (long-distance) migration of children. This said, the evidence regarding the potential of microfinance to generate new income remains mixed at best. Moreover, the overall effect need not be positive if the newly-created activities actually require more labor (i.e. if labor and capital are complements in the households' production function), which might take the form of child help if for instance the local labor market is failing.[11] Also, general equilibrium effects might cause changes in the wage structure, which could increase the opportunity cost of schooling as well as the expected returns to schooling.

Another way SHG membership could affect children's education is through providing insurance against income shocks. Indeed, Demont (2014) shows that SHGs help to smooth income in face of adverse rain shocks, and might therefore decrease the need for child labor as one traditional way of diversifying and smoothing family income. In the case of consumption smoothing, parents might decide to take their children out of school and put them into work to sustain a certain minimum consumption level if the parents' income drops or fails (supporting empirical evidence include Jacoby and Skoufias 1997, Jensen 2000, Beegle et al. 2006 and Hyder et al. 2012). The danger of this strategy lies in the irreversibility of child labor. Once children are taken out of school and integrated into the labor market, transition back to school is often very difficult (e.g. Guarcello et al., 2003; Cigno and Rosati, 2005; Duryea et al., 2007). Finally, SHG participation may also better enable women to foster the education of their children, by increasing their relative weight in household decisions. By having access to more resources and opportunities, participating women are better able to promote and invest in the education and the general well being of their family (Duflo, 2012). Peer effects and norms for better education conveyed by the NGO and the group can further reinforce the perceived priority of educating children.

Note that the resulting change in enrollment rates, if any, might entail important gender discrimination. For instance, given that males' returns are much higher on labor and marriage (e.g. because of the dowry system) markets, the additional resources might benefit to boys more than proportionally - thereby increasing the gender gap in education. This might revert when some target education level for boys is reached in the household, after which investment in girls might take over and the gender gap might decrease. On the other hand, given that the new resources come from the women, they might focus more strongly on girl's education.

# 4 Data

We performed four waves of household surveys. In August 2002, a first round was conducted to collect data on SHG participants and non participants in 24 villages where the program had

[11] Augsburg et al. (2015) give evidence of such an increase in the opportunity cost of schooling after accessing microcredit in Bosnia and Herzegovina.

just been introduced. It can therefore be considered as a baseline or pre-treatment survey, since participants were identified but did not receive any benefits yet. Indeed, SHGs need several months of smooth functioning and savings accumulation before being able to progressively start lending money. However, this first data collection had originally a narrower focus, namely to examine the targeting strategy of PRADAN's SHG program (see Dewan and Somanathan, 2007). This is why this first wave does not include control villages, which were introduced from the second round. We used a stratified random sampling strategy to select the surveyed villages in order to have a representativity of the entire state of Jharkhand. We formed four geographical clusters based on differences in demographic characteristics observed in census data (for instance, the religious and social composition of these areas is quite different): Northeastern (Santhal Parganas districts), Central (Hazaribagh and surrounding districts), Southwest (Ranchi-Lohardaga districts) and Southeast Jharkhand (Singhbhum districts). For each cluster, a random sample of 6 villages was chosen from the set of all villages with at least one SHG formed during the period April 1st - June 30th, 2002.[12] In each village, we selected randomly and surveyed 6 participants of microcredit groups and 18 non participants. Then, in January 2004, we added to the sample 12 extra member households from each village, in order to have a balanced number of members and non members. In addition, we randomly selected 12 new villages from the list of villages with no SHG and located in the same districts as the 24 previous villages. The objective was to have pure control, or untreated, villages that we could use as an alternative comparison group and to control for potential issues linked to village selection and externalities. In each of those control villages, 18 households were randomly selected and surveyed.

This constitutes the full sample that was followed from round 2 to round 4, which counts 1,080 households (40% members, 40% nonmembers, 20% controls) from 36 villages. This is the sample on which the econometric analyses will focus, though using data from the pre-treatment round to control for pre-existing differences between members and nonmembers through a matching strategy. Appendix A provides the full list of villages that were surveyed, as well as some basic descriptive statistics at the district and village levels. Though no difference is statistically significant, treated villages appear to be slightly worse off than control villages, which is consistent with the NGO's targeting of administrative blocks with high incidence of rural poverty. On average, treated villages have a higher proportion of landless households, are more heterogenous in terms of caste and tribes, are more remote and isolated.

The questionnaire aimed at measuring living standards along multiple dimensions. It recorded detailed information about household demographics, recurrent and durable expenditures, consumption, credit and savings, labor market participation and self-employment, migration, food vulnerability, landholdings and agriculture, the condition of the household dwelling, health, education, female empowerment and participation in village activities. All surveys were carried during the same period of the year, namely January-March, which corresponds to a rather slack, pre-harvest period at the end of the winter season.

[12] Groups created in July were excluded because membership tends to be unstable during the first month. Four villages were actually selected from a neighboring district of Bihar, see appendix A.

As a result of the two-stage sampling, all households were followed over five years at least, and some over seven years. Overall, 75% of the households were interviewed in all rounds. Table 1 gives some statistics about the dynamics of the sample. It shows that the retention rate is high, with an average attrition rate across rounds of 6%. A limited number of households could not be re-interviewed due to migration, identification problems, or refusal. In addition, two entire member villages had to be replaced for security reasons, due to extreme naxalite rebels' activism - one in round 2 (Sallaya) and one in round 4 (Kera). When removing those two villages, the average attrition rate falls to 4%, which is small enough not to be a major issue for the analysis.[13] As second issue that is relatively more important has to do with changes in membership status across rounds. Those occur mainly due to the opening of new groups or the closure of ill-functioning groups. Table 1 reports the relative numbers of exit from (row 3) and entry into (row 4) SHGs over time. Regarding new entry, the large majority of the cases concerns the creation of new groups. In fact, entering an existing group is relatively hard due to the size limit of the groups and the requirement that newcomers must contribute to the group an amount equal to the accumulated savings per member at that time. This deters prospective members from joining existing groups, and they often choose to motivate other people and start a new SHG instead. In addition, some SHGs were created in control villages after 2004. Cumulating the two types of change (i.e. non-member to member of SHG or drop-outs from SHG), the average rate of change of status across rounds is 13%, such that, in the last round, households who were members in round 2 are still much more likely to be part of an SHG than other households. In the econometric analysis, we make the conservative choice of defining members and non members according to their original membership status and perform various sensitivity analyses in order to account for imperfect treatment compliance (see next section).

Table 1: Sample dynamics, by survey round

| | round 2 | round 3 | round 4 |
|---|---|---|---|
| *With respect to the previous round:* | | | |
| % attrition (average)$^\dagger$ | 7.5 (3.8) | 4.6 (4.4) | 7.4 (4.4) |
| % attrition (SHG members)$^\dagger$ | 7.6 (4.6) | 1.9 (1.8) | 5.4 (1.8) |
| % change of treatment status (SHG members)$^\ddagger$ | 0 | 10.9 | 17.1 |
| % change of treatment status (non members)$^\ddagger$ | 0 | 17.0 (14.5) | 8.0 (6.0) |

$^\dagger$ Figures in parentheses exclude the two entire villages that had to be dropped for security reasons.
$^\ddagger$ Figures in parentheses indicate new groups.

[13] All our econometric results are robust to the exclusion of those two villages. Furthermore, we find that attrition is slightly lower for SHG members than for other households. To the extent that the households with worst outcomes tend to display higher attrition probability, this should lead to an attenuation bias, if anything. For instance, we find that the probability of attrition is positively correlated with round 2 enrollment rates in non-member households, s.t. those are the households with the lowest margin of enrollment improvement.

## 5 Empirical strategy

Our first approach is to estimate the effect of opening SHGs in a village, irrespectively of the eventual membership of households (intent-to-treat estimates, or ITT). We do this by simply averaging outcomes of all households living in treated villages and comparing their evolution to the situation in control villages, where the SHG program was not implemented (initially). In order to be representative of the village population from which they are drawn, observations are weighted in order to control for the different sampling probabilities of SHG and non-SHG households in treated and control villages. Using data from the three last survey rounds (2004, 2006 and 2009)[14], we adopt the following baseline specification:

$$Y_{ihvt} = \alpha + \beta T_v + \beta_3(T_v * R3_t) + \beta_4(T_v * R4_t) + C'_{it}\gamma + H'_{ht}\eta + V'_v\nu + \psi S_{vt} + \lambda_t + \delta_v + \epsilon_{ihvt} \tag{1}$$

where $Y$ is the outcome of interest (such as school enrollment) for child i living in household h and village v at year t, $T$ is a time-invariant dummy variable taking value 1 if village v is a treated village (i.e. where the SHG program was introduced from 2002) and 0 if it is a control village, $R3$ and $R4$ are round (time) dummies, $C$ and $H$ are vectors of control variables at the child and household levels respectively, and $V$ are pre-treatment village characteristics. The latter comprises the following variables that are all measured in 2001: village size (number of households), road access, distance to market, proportion of scheduled-caste and landless households, female working participation rate, male and female average literacy levels, presence of primary and middle schools, distance to secondary school (no village of the sample has got a secondary school). The vector $H$ is composed of the following household-level variables that are either pre-determined or very slow to change: land ownership, head's education and age, mother's education and age, household size , number of toddlers (0-5 years old), house size per equivalent adult, official scheduled caste or tribe and below-poverty-line statuses, religion. And the child's characteristics $C$ that we include are: his/her birth rank (and birth rank squared), age and sex. Finally, $S$ is a continuous variable controlling for the occurence of large village-wide income shocks during the year before each survey round, which could be a potentially important source of time-varying village heterogeneity[15], and $\lambda$ and $\delta$ are respectively time and district or

[14] As already explained, round 1 comprises only of a subset of households and variables. Moreover, the limited sample size would not allow us to get robust estimates for narrow age groups. In round 2, there is already a potential effect of some months of SHG membership, even though we believe this effect to be zero or very limited, because groups were still mostly in their savings phase and none were linked to a commercial bank yet (see section 2.2). If anything, by comparing round 4 to round 2, we may underestimate the total effect of the program, in which case our analysis should be interpreted as delivering a reliable lower bound.

[15] The vast majority of the people living in the survey area are small landholders, living out of a subsistence agriculture characterized by small marketable surpluses and little investments in infrastructure or inputs. In particular, the cultivation of rain-fed rice represents the main source of food and income for the households in the sample. As a consequence, the abundance of rain has a large positive impact on households' welfare and resources. We construct a continuous variable measuring the relative quality of the monsoon in the district during the year before loans were taken: $S = \frac{\overline{M}_d - Monsoon_{dt}}{\sigma_d}$, where $\overline{M}_d$ and $\sigma_d$ are respectively the historical average and standard deviation of the monsoon level in each district and are computed over a rolling window of the ten years

village fixed effects.[16] Descriptive statistics of all variables are given in appendix. In all specifications, we cluster standard errors at the household level (which corresponds to the treatment level), in order to allow for heteroskedasticity and correlation of errors within households (both between children and over time).[17]

The coefficients $\beta_3$ and $\beta_4$ are thus the main coefficients of interest, giving for each round the differential evolution of children living in treated villages as compared to children living in other villages from the same district (the base category being control villages in round 2). This ITT analysis is interesting because it estimates the true return of the NGO's intervention at the village level, taking into account the fact that a share of the population is not taking up the intervention. It has the advantage of bypassing entirely the participant selection issue, and of factoring in potential spillovers at the village level.

Yet, because only a fraction of treated villages' population participate in SHGs (between 4 and 75%; 29% on average), ITT estimates do not speak to the benefit of actual SHG participation. Therefore, our second approach aims at estimating the evolution of the impact of SHG membership over time for households who have decided to take part in the program (average treatment effect on the treated, or ATT). To do so, we drop from the sample the households in treated villages who (originally) decided not to participate in SHGs, and run the following regression:

$$Y_{ihvt} = \alpha + \beta SHG_h + \beta_3(SHG_h * R3_t) + \beta_4(SHG_h * R4_t) + C'_{it}\gamma + H'_{ht}\eta + V'_v\nu + \psi S_{vt} + \lambda_t + \delta_d + \epsilon_{ihvt} \quad (2)$$

where $SHG$ is a dummy indicating the original (time-invariant) membership status of household h (and the other variables are as defined above). As mentioned in the previous section, we have a small number of 'noncompliers' in the sample, i.e. households who joined or left SHGs after round 2. As a consequence, we define member and control households by their original (time invariant) status in round 2. This conservative choice avoids concerns about the endogeneity of membership changes and delivers a lower bound for the ATT.

Using this second specification, the coefficients $\beta_3$ and $\beta_4$ therefore give the differential evolution of children belonging to SHG household as compared to households in control villages

immediately preceding the survey year. Rainfall data come from the Global Precipitation Archive (Matsuura and Willmott, 2012). The nine districts present in the sample being spread across the different agro-climatic zones of the state, the variable captures important variation between villages and across years. As shown in Demont (2014), it is strongly correlated with income levels and access to traditional credit sources in the sample villages.

[16] District fixed effects appropriately account for our sample stratification strategy and for the fact that districts correspond to the basic unit of Indian administration (typically covering several dozen villages) that is in charge of the implementation of many policies (such as the coordination of development and welfare schemes, tax collection and maintenance of law and order). They also allow us to include our rich set of village-level characteristics to control for important school enrollment determinants explicitly. Nevertheless, to make sure that there is no remaining unobserved heterogeneity at the village level, we also report results using village fixed effects, in which case we drop the baseline difference $T$, as well the vector of village covariates $V$.

[17] Given the large number of households, cluster-robust inference will be accurate in case of serial correlation and costless in the absence of actual clustering (Wooldridge, 2010). Given that household and child levels are nested, it is useless to cluster for both, and clustering at the higher level is the conservative way of controlling for correlation of residuals within both levels (Bertrand et al., 2004).

from the same district (the base category being control households in round 2). They provide round-specific estimates of the ATT under the assumption that, after controlling for the initial difference between the member households and the control group, potential outcomes in the absence of treatment (or counterfactual levels) are independent of participation decisions.[18] This simple difference-in-difference framework might not be valid if selection is not constant over time, i.e. if the two populations do not follow parallel trends in the absence of treatment (given that we have only one pre-treatment round of data, we cannot directly test this), or is not linearly additive in its effect on outcomes. Moreover, simple regression estimators may be very sensitive to differences in the covariate distributions for treated and control units, because in that case they rely heavily on extrapolation to impute the missing potential outcomes from the estimated regression function (see Imbens, 2004). We therefore combine our diff-in-diff approach with propensity score matching, in order to directly balance treatment and comparison households in terms of the baseline characteristics that may have influenced selection into the program. In addition, matching estimators also impute the missing potential outcomes, but do so using only the outcomes of nearest neighbors of the opposite treatment group, in a non-parametric fashion. In our case, another advantage of matching is that it controls better for the fact that we have a heterogenous control group (by definition, the entire population of control villages is composed of both would-be non participants and would-be participants).

We regress equation (2) by weighted least squares, using an increasing function of a non-parametric estimate of the propensity score to weight the control observations. That is, we give more importance to control observations that are more similar to member households, according to their baseline characteristics. We start by estimating a logistic model for the probability of being a member in treated villages on a wide set of household characteristics and outcomes, using baseline (round 1) data - i.e. before any benefit from membership. Appendix B provides descriptive statistics about the variables used in the estimation of the scores and displays the estimation model. It shows that households who decide to participate in SHGs are more likely to come from a scheduled caste, to be female-headed, to be headed by a married couple, to have a head whose main occupation is a casual wage job, to have failed to benefit from the IAY program to construct / renovate their house[19], to own more land (though they are as likely to be landless), to have a smaller house, to have a per-capita consumption level that is below median (marginally significant) but not extremely low, to take more loans (though total credit is not different, implying a smaller average loan size), to have boys aged 12-14 in school, to have chil-

[18]Given self-selection of participants, it is more natural to consider that member households represent a particular subsample of the entire population and therefore focus on ATT. Estimating the average treatment effect on the entire population would require the additional assumption that potential outcomes under treatment are also identical across treatment groups, which is likely to be violated if e.g. participants self-select based on expectations about outcomes.

[19]The Indira Awaas Yojna scheme has been running since 1986 and provides financial assistance for construction or upgradation of dwelling units to below-poverty-line rural households, with a particular focus on scheduled tribes and castes (non scheduled castes or tribes can get access up to maximum 40% of the total IAY allocation). Note that the number of eligible BPL families is typically far greater than the funds available in any given year, so that there are important waiting lists and discretion in the final selection (done by local municipal councils).

dren aged 15-17 and to have few adults, as compared to other households from the same village. We then use the estimated coefficients to predict the scores $P(X)$ (or probability of joining) of control households in round 2. Our model gets a pseudo-R$^2$ of 17% and correctly classifies 78% of the households, which represents a fairly good predictive power. We then fix those scores over time and use them to construct the following weights for each observation:

$$w(SHG, X) = (1 - SHG)\frac{P(X)}{1 - P(X)} + SHG.$$

That is, member households receive a weight of one, while control households receive a weight that is proportional to their probability of being an SHG member. To interpret the effect of the weighting, note that $\beta$ in (2) will then pick up differences in the mean of the *latent* individual effects, such as would arise from initial selection into the program. As shown by Hirano et al. (2003), this weighting schedule leads to a fully efficient estimator under the conditional independence assumption discussed above. Note that weights are further normalized so that they add up to one, which helps ensuring convergence (Imbens, 2004; Busso et al., 2009). Appendix B provides different balancing diagnostics. For our data, the region of common support is large, as controls can be found over most of the score distribution of the member population. In such a context, the normalized weighting technique applies particularly well and has been shown to perform better than other matching estimators (Busso et al., 2009). Nevertheless, the proportions are naturally unbalanced, especially in the tails of the distribution. To focus on the common-support region and ensure that extreme observations do not exert undue influence on the estimates, we drop observations with scores below the minimum score of SHG members in round 1 (0.07) and above the maximum score of non-SHG members in round 1 (0.76).[20] The weighting technique makes the distribution of the estimated propensity scores look very similar between treated and control groups. Looking at individual covariates, matching observations on the estimated propensity score brings most of the means much closer together and appears to balance the control and treatment groups extremely well, removing all significant differences. In particular, there are no difference in baseline education outcomes left after matching.

In conclusion, the weighted-regression method represents an elegant way to obtain consistent estimates by balancing treatment and comparison households, while sticking to the efficient and transparent linear-regression framework presented above. It thus allows to show explicitly the evolution of the outcomes of interest round after round, while accounting for latent differences across treatment and comparison units that might affect the selection into the program as well as the potential outcomes. The combination of matching and regression leads to consistent estimates as long as either the regression model or the propensity score are specified correctly (Imbens, 2004). By addressing the correlation between the covariates and the outcome, the

[20] We check that our main results are robust to using instead the efficiency bounds recommended by Crump et al. (2009) - i.e. focusing on scores between 0.1 and 0.9 - though in our case that trimming strategy seems less conservative as it implies keeping virtually all the highest scores and dropping a lot of observations at the lower end of the score distribution.

linear regression framework helps eliminating any remaining bias and improving precision as compared to simple matching. Moreover, it allows the correction of standard errors and has better small-sample properties.

# 6 Results

## 6.1 Enrollment rates

In this section, we report the evolution of school outcomes for children in SHG households as compared to others.[21] For every children present in the household between 6 and 17 years old, our survey collected data on school enrollment, attendance, grade achieved, type and location of the school as well as school-related expenditures. We also have information about children who are temporary away from the household for study motives, e.g. on boarding schools or at relatives' place.

We start with the ITT analysis. We report the specifications with and without village fixed effects, and display the main estimates of interest. We focus on children beyond-primary-school age. This is because, as we explain above, this is were we expect most action to take place, because the baseline enrollment rates are much lower (41%, versus 75% for primary-school ages). We also add interection terms to investigate the girl-specific effects, because of the enrollment gender gap observed at baseline as well as the female orientation of the SHG program. We find a positive effect, detectable in the last round only, from having SHGs in the village on enrollment rates of children over 12 years, who are about 20 percentage points to be enrolled in SHG villages than in control villages. Interestingly, the effect is true for all children, and girls are as likely as boys to benefit from the program.

In table 3, we then turn to the estimation of the impact for direct participants, by comparing them to matched control households (thus dropping non-participants in SHG villages). We find once again strong positive effects of SHG membership, which now start already in round 3 - indicating more clearly a progressive improvement over time. By the last round, school-age children of SHG households are about 25 percentage points more likely to be enrolled. The effects tend to be stronger for girls, though interaction coefficients are not significant. Those are very sizable effects, representing more than a 50% improvement of baseline enrollment rates for children over 12 years old. Moreover, remember that this is likely to be a lower-bound estimate since we exclude round 1 data and we use the original membership status. Note that this effect is mostly driven by a fall in the dropout rate at secondary-school age; it does not come from the delayed effect of early entry into primary school, since children who start entering school from 2003-2004 are 11-12 years old in the last round. The fact the ATT coefficients are higher than the ITT ones indicate that SHGs benefit primarily to direct members, which was expected.

[21] Throughout the analysis of education outcomes, we focus on the direct children of the corresponding female member of the household.

However, we find evidence of positive spillovers on non-participants, and the ITT analysis reveals that the policy is beneficial at the community level as well.

Different reasons may explain the progressivity and the strength of the effects in round 4. First, educating children is fondamentally a forward-looking, long-term investment. For enrollment rates to increase, parents might need to progressively acquire higher expectations about the returns to education, in particular of girls. It is also likely that members first start to deal with the more pressing needs of present times, only to think about investing for the future in a second step. Indeed, the literature usually highlights the fact that, when deciding how children spend their time, parents usually have a strong preference for schooling, as long as the basic needs of the family are covered (Blume and Breyer, 2011). The last round, occurring three years after the previous one and about six years after the start of operation of the SHGs, appears to represent a sufficient time lapse. Moreover, the evolution of enrollment rates, and in particular the decrease in drop-out rates at the secondary-school level, is only the 'extreme' or extensive-margin effect of an increased investment in children's schooling. It is reasonable to think that it is determined by more intensive-margin outcomes occuring beforehand, such as parents making sure that enrolled children effectively attend classes, do homework, do not attend school with an empty stomach, etc. Those outcomes might explain survival into later years and therefore show up in enrollment rates with a delay. For instance, it could be that the member girls are observed to drop out less at the secondary-school level in the last round because of better school performance (and success rate in examinations) at the primary and midle school in the previous rounds. Another reason might stem from the fact that SHGs, by progressively accumulating the regular deposits of their members and becoming linked to a commercial bank after some years, require time to produce effects, if only in terms of access to credit. There might be more subtle reasons as well. For instance, participation to an SHG involves repeated interactions with other group members that may change one's attitudes or preferences with respect to schooling, and we do not expect those effects to be instantaneous. Finally, education decisions, particularly regarding girls, often involve some degree of intra-household bargaining between husbands' and wives' preferences about the allocation of scarce resources, as well as some reorganization of occupational choices within the household (below, we show that going to school and working are clear substitutes for children in the sample). Those processes take time to come into effect.

Some significant controls are interesting to point out (not displayed in the table). As expected, age is negatively associated with enrollment at all levels and especially at the secondary level, indicating the existence of an important drop-out issue. We observe a significant concave relationship between birth order and enrollment, indicating that children born second and third have the highest probability to be enrolled. Indeed, the first-born child is often required to contribute to households' resources, while the economic possibilities to send additional children to school might be exhausted (or the investment target might be reached) as the number of children gets large. This is partly confirmed by the fact the the number of toddlers (0-5 years) in the household affects negatively enrollment rates. We observe an important status-quo bias, as the education level of the head is strongly and monotonically associated with the enrollment of children (the higher the education of the head, the higher the probability that children are at school). The mother's education matters only for girls, for whom having an uneducated mother reduces significantly the enrollment probability. Finally, the per-capita house size is positively associated with enrollment rates, indicating that wealth is a significant constraint, though land ownership fails to be significant in most specifications.[22] We do find evidence that monsoon quality matters for the school enrollmenty of children in member househols, but rather in the opposite direction. Therefore, when controlling for monsoon quality, the average SHG effect on girls' enrollment rates increases even further. Replacing district fixed effects by time*district fixed effects delivers similar results (not shown here).

Table 2: Enrollment of children: ITT analysis

| | children aged 12-17 | | | | children aged 15-17 | | | |
|---|---|---|---|---|---|---|---|---|
| | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) |
| membervillage | -0.0851 | | -0.0431 | | -0.102 | | -0.0464 | |
| | (0.0619) | | (0.0797) | | (0.0831) | | (0.110) | |
| membervilXr3 | 0.0261 | 0.0156 | 0.00257 | -0.0198 | 0.103 | 0.109 | 0.0652 | 0.0633 |
| | (0.0748) | (0.0750) | (0.0961) | (0.0969) | (0.106) | (0.106) | (0.143) | (0.141) |
| membervilXr4 | 0.178** | 0.186** | 0.198* | 0.191* | 0.250** | 0.262** | 0.298** | 0.287* |
| | (0.0780) | (0.0797) | (0.101) | (0.103) | (0.115) | (0.115) | (0.147) | (0.149) |
| femaleXmembervil | | | -0.112 | -0.127 | | | -0.138 | -0.161 |
| | | | (0.112) | (0.113) | | | (0.158) | (0.162) |
| femaleXmvr3 | | | 0.0727 | 0.0972 | | | 0.0932 | 0.102 |
| | | | (0.130) | (0.131) | | | (0.215) | (0.210) |
| femaleXmvr4 | | | -0.0591 | -0.0200 | | | -0.132 | -0.0718 |
| | | | (0.150) | (0.148) | | | (0.201) | (0.200) |
| $N$ | 1704 | 1704 | 1704 | 1704 | 874 | 874 | 874 | 874 |
| $R^2$ | 0.222 | 0.245 | 0.229 | 0.252 | 0.181 | 0.213 | 0.190 | 0.223 |
| village controls | yes | no | yes | no | yes | no | yes | no |
| household controls | yes | yes | yes | yes | yes | yes | yes | yes |
| round FE | yes | yes | yes | yes | yes | yes | yes | yes |
| district FE | yes | no | yes | no | yes | no | yes | no |
| village FE | no | yes | no | yes | no | yes | no | yes |

Std errors clustered at the household level in parentheses (*p<0.10, **p<0.05, ***p<0.01).
All equations control for the sex, age and birth rank of children, as well as the monsoon quality in year t-1.
Observations weighted in order to account for the different sampling probabilities.

Being enrolled only represents the extensive margin of getting educated. Children then need to attend school regularly and to progress through grades smoothly. We now focus on enrolled children and analyze different intensive margin educational outcomes. First, we collected infor-

[22]While agriculture is by large the main economic activity in the region, there is no evidence that the opportunity cost of education is rising with land ownership, as for instance in Maldonado and Gonzalez-Vega (2008). In fact, the sign attached to the land ownership variables go quite in the opposite direction, indicating that landlessness tends to be associated with lower enriollment rates on average and that, by contrast, big land ownership tends to sustain higher enrollment rates.

Table 3: Enrollment of children: ATT analysis

| | children aged 12-17 | | | | children aged 15-17 | | | |
|---|---|---|---|---|---|---|---|---|
| | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) |
| membervillage | -0.154** | | -0.105 | | -0.0731 | | -0.188 | |
| | (0.0705) | | (0.0898) | | (0.103) | | (0.124) | |
| membervilXr3 | 0.269** | 0.239** | 0.201* | 0.162 | 0.176 | 0.106 | 0.352** | 0.232 |
| | (0.108) | (0.103) | (0.115) | (0.111) | (0.126) | (0.118) | (0.167) | (0.175) |
| membervilXr4 | 0.283*** | 0.276*** | 0.209* | 0.195* | 0.286** | 0.261** | 0.336* | 0.239 |
| | (0.0832) | (0.0861) | (0.113) | (0.112) | (0.138) | (0.122) | (0.182) | (0.166) |
| femaleXmembervil | | | -0.121 | -0.111 | | | 0.252 | 0.134 |
| | | | (0.122) | (0.111) | | | (0.192) | (0.189) |
| femaleXmvr3 | | | 0.158 | 0.175 | | | -0.365 | -0.244 |
| | | | (0.153) | (0.149) | | | (0.238) | (0.241) |
| femaleXmvr4 | | | 0.148 | 0.173 | | | -0.0964 | 0.0898 |
| | | | (0.172) | (0.166) | | | (0.261) | (0.249) |
| $N$ | 1066 | 1066 | 1066 | 1066 | 544 | 544 | 544 | 544 |
| $R^2$ | 0.248 | 0.282 | 0.254 | 0.286 | 0.204 | 0.249 | 0.220 | 0.260 |
| village controls | yes | no | yes | no | yes | no | yes | no |
| household controls | yes | yes | yes | yes | yes | yes | yes | yes |
| round FE | yes | yes | yes | yes | yes | yes | yes | yes |
| district FE | yes | no | yes | no | yes | no | yes | no |
| village FE | no | yes | no | yes | no | yes | no | yes |

Std errors clustered at the household level in parentheses (*p<0.10, **p<0.05, ***p<0.01).
All equations control for the sex, age and birth rank of children, as well as the monsoon quality in year t-1.
Control observations weighted by their propensity score.

mation about actual school attendance (days at school) during the week preceding the survey. We find that children in member villages are attending about 1 more school day on average, indicating that reported enrollment status indeed corresponds to school attendance.

In table 5, we follow cohorts of given children over time, who were enrolled in 2004, and check if we can detect a pure 'survival' effect at the middle and secondary school levels in rounds 3 and 4. Columns 1-4 study drop-out (i.e. being out of school) at middle and secondary school levels, by focusing on children who are all enrolled in primary school at baseline. We find that SHG participation has a large effect on reducing drop-out rates. In columns 5-8, children are all enrolled at baseline and reach secondary-school age (15-17) in round 4. As expected, effects are stronger in this case, as drop-out rates are the highest at the secondary-school level at baseline.

Table 4: School attendance: ITT analysis

| | children aged 12-17 | | | | children aged 15-17 | | | |
|---|---|---|---|---|---|---|---|---|
| | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) |
| membervillage | -0.283 | | -0.0620 | | -0.398 | | -0.198 | |
| | (0.312) | | (0.394) | | (0.388) | | (0.495) | |
| membervilXr3 | 0.0253 | -0.0597 | -0.217 | -0.342 | 0.289 | 0.331 | 0.126 | 0.126 |
| | (0.369) | (0.362) | (0.452) | (0.452) | (0.492) | (0.485) | (0.662) | (0.656) |
| membervilXr4 | 0.787** | 0.785** | 0.879* | 0.822 | 1.027** | 1.054** | 1.473** | 1.397** |
| | (0.388) | (0.388) | (0.511) | (0.509) | (0.521) | (0.516) | (0.644) | (0.655) |
| femaleXmembervil | | | -0.583 | -0.615 | | | -0.489 | -0.580 |
| | | | (0.538) | (0.545) | | | (0.735) | (0.732) |
| femaleXmvr3 | | | 0.627 | 0.704 | | | 0.381 | 0.435 |
| | | | (0.633) | (0.627) | | | (1.049) | (1.008) |
| femaleXmvr4 | | | -0.273 | -0.135 | | | -1.193 | -0.934 |
| | | | (0.772) | (0.766) | | | (0.967) | (0.956) |
| $N$ | 1704 | 1704 | 1704 | 1704 | 874 | 874 | 874 | 874 |
| $R^2$ | 0.214 | 0.238 | 0.220 | 0.244 | 0.186 | 0.219 | 0.198 | 0.230 |
| village controls | yes | no | yes | no | yes | no | yes | no |
| household controls | yes | yes | yes | yes | yes | yes | yes | yes |
| round FE | yes | yes | yes | yes | yes | yes | yes | yes |
| district FE | yes | no | yes | no | yes | no | yes | no |
| village FE | no | yes | no | yes | no | yes | no | yes |

Std errors clustered at the household level in parentheses (*p<0.10, **p<0.05, ***p<0.01).
All equations control for the sex, age and birth rank of children, as well as the monsoon quality in year t-1.
Observations weighted in order to account for the different sampling probabilities.

Table 5: School drop-out: ITT analysis

| | Children aged 7-12 and enrolled in 2004 | | | | Children aged 10-12 and enrolled in 2004 | | | |
|---|---|---|---|---|---|---|---|---|
| | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) |
| membervilXr3 | 0.0594* | | 0.0290 | | 0.0811 | | 0.0375 | |
| | (0.0359) | | (0.0434) | | (0.0586) | | (0.0661) | |
| membervilXr4 | -0.109* | -0.176*** | -0.145* | -0.208** | -0.220* | -0.294** | -0.258 | -0.337* |
| | (0.0556) | (0.0613) | (0.0786) | (0.0878) | (0.120) | (0.130) | (0.171) | (0.180) |
| femaleXmvr3 | | | 0.0666 | | | | 0.0841 | |
| | | | (0.0684) | | | | (0.0852) | |
| femaleXmvr4 | | | 0.0899 | 0.0809 | | | 0.0707 | 0.0832 |
| | | | (0.112) | (0.110) | | | (0.221) | (0.209) |
| $N$ | 1031 | 1031 | 1031 | 1031 | 447 | 447 | 447 | 447 |
| $R^2$ | 0.161 | 0.185 | 0.166 | 0.190 | 0.285 | 0.323 | 0.290 | 0.327 |
| village controls | yes | no | yes | no | yes | no | yes | no |
| household controls | yes | yes | yes | yes | yes | yes | yes | yes |
| round FE | yes | yes | yes | yes | yes | yes | yes | yes |
| district FE | yes | no | yes | no | yes | no | yes | no |
| village FE | no | yes | no | yes | no | yes | no | yes |

Std errors clustered at the household level in parentheses (*p<0.10, **p<0.05, ***p<0.01).
All equations control for the sex, age and birth rank of children, as well as the monsoon quality in year t-1.
Observations weighted in order to account for the different sampling probabilities.

In 6, we double check the enrollment results by looking at actual school expenditures, which we have collected for all school-age children. Using expenditures helps us to make sure that those children are not merely reported to be enrolled, but that they are indeed attending school and generating direct costs for their family. In columns 1-4, we regress the actual school expenditures on the same set of explanatory variables as previously (in the interest of space, we do not report the specification including female interaction variables, whose coefficients are never significant). Though expenditures data are a bit noisier and mix extensive and intensive margins effects, they tend to confirm the previous results, i.e. a significant increase for children of post-primary school age. Another strategy, having in mind the possibility of 'fake' enrollment, is to check whether significant education expenses have been incured. We construct a dummy variable that takes the value one if the school expenses are equal or larger than INR 100, which corresponds to the median of the conditional expenditure distribution for primary-school age children and to the modal price of a school uniform in the sample. That is, it can be considered as a minimum cost per child who is 'regularly' enrolled. Columns 5-8 show that the higher enrollment rates observed for SHG children are indeed linked with a higher incidence of significant school expenses. Finally, we observe that SHG participation has no intensive margin effect, as it does not change the amount spent by enrolled child (not shown).

Table 6: School expenditures: ITT analysis

| | Log of school expenditures | | | | Non-trivial school expenditures dummy | | | |
|---|---|---|---|---|---|---|---|---|
| | children aged 12-17 | | children aged 15-17 | | children aged 12-17 | | children aged 15-17 | |
| | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) |
| membervillage | -0.362 | | -0.559 | | -0.0371 | | -0.0939 | |
| | (0.378) | | (0.510) | | (0.0638) | | (0.0821) | |
| membervilXr3 | -0.0480 | -0.158 | 0.553 | 0.503 | -0.0324 | -0.0458 | 0.0806 | 0.0727 |
| | (0.450) | (0.453) | (0.644) | (0.655) | (0.0756) | (0.0754) | (0.106) | (0.108) |
| membervilXr4 | 0.801* | 0.788* | 1.371* | 1.378* | 0.105 | 0.102 | 0.197* | 0.200* |
| | (0.464) | (0.476) | (0.700) | (0.703) | (0.0802) | (0.0813) | (0.111) | (0.111) |
| $N$ | 1704 | 1704 | 874 | 874 | 1704 | 1704 | 874 | 874 |
| $R^2$ | 0.201 | 0.227 | 0.185 | 0.220 | 0.167 | 0.189 | 0.185 | 0.214 |
| village controls | yes | no | yes | no | yes | no | yes | no |
| household controls | yes | yes | yes | yes | yes | yes | yes | yes |
| round FE | yes | yes | yes | yes | yes | yes | yes | yes |
| district FE | yes | no | yes | no | yes | no | yes | no |
| village FE | no | yes | no | yes | no | yes | no | yes |

Std errors clustered at the household level in parentheses (*p<0.10, **p<0.05, ***p<0.01).
All equations control for the sex, age and birth rank of children, as well as the monsoon quality in year t-1.
Observations weighted in order to account for the different sampling probabilities.

## 6.2 Credit

One of the main role of SHGs is to bring access to credit to its members. Table 7 presents the different loan options available in a typical village of our sample. SHG borrowing appears clearly as a cheap option, especially as compared to moneylenders - which is the only other 'readily-available' alternative. SHG loans are also relatively smaller (1,270 INR, or about 20 USD, on average) and of shorter duration (7 months on average). Yet, they can be accessed much more frequently. From the number of loans, it is easy to verify that member households reduce dramatically their reliance on moneylenders, by substituting for SHG borrowing. This observation implies that (i) borrowing from SHGs is cheaper (and/or more convenient etc.) and (ii) SHGs meet most of the demand for credit of its members. Hence, SHGs do seem to improve the access to credit and can be expected to help enrolling members' children at school, especially at the secondary level that involves much larger costs.

In table 8, we show the average access and use of credit during the two years preceding the survey for member, nonmember and control households with school-age children. For the two first categories, we report round 1 data as well, in order to give an idea of the pre-SHG situation (though, as explained before, the sample is not entirely comparable between round 1 and the other rounds). It appears clearly that SHGs improve strongly the access to credit its members. From round 2, around 25% more households take credit as compared to round 1, implying that 9 member households out of 10 take at least one loan in two years. The average amounts borrowed from SHGs are relatively smaller, but increase over time as the volume of internal savings and bank credit increases. It is therefore not surprising that SHG members appear less credit constrained than other members when asked about the management of episodes of unexpected urgent cash needs. While very few of the other households take loans for education purposes, 14% and 7% of the SHG members did so in round 3 and 4 respectively. Figure 2 shows that SHG households having secondary-age children at school are more likely to borrow large amounts than other SHG households, which suggests that part of the credit taken by SHG members might be used to finance the secondary schooling of their children.

Table 7: Average conditions of different loan options available to villagers

| | SHG | Moneylender | Relative | Bank | Other |
|---|---|---|---|---|---|
| Interest rate (% monthly) | 2.5 | 7.8 | 2.3 | 3.3 | 0 |
| Amount (INR) | 1,270 | 3,013 | 3,850 | 21,567 | 3,171 |
| Duration (months) | 7.1 | 8.3 | 9.0 | 20.3 | 7.3 |
| Frequency SHG members (%) | 89.7 | 7.6 | 3.3 | 0.6 | 2.6 |
| Frequency other households (%) | 7.4 | 48.2 | 24.1 | 4.1 | 16.2 |
| number of loans | 3,241 | 1,008 | 518 | 86 | 340 |

Despite this dramatic increase in credit access, we do not find any specific effect of credit on enrollment rates for member households, which indicates that it is certainly not the main channel through which member households reach higher enrollment rates. In table **??**, we run

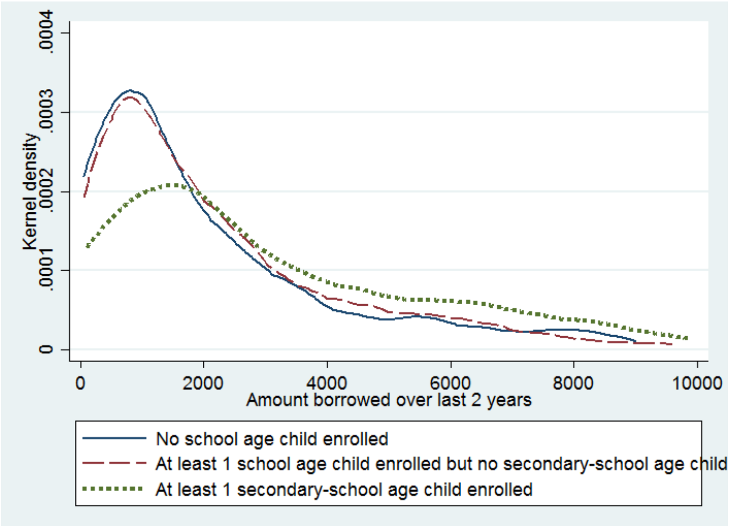
Table 8: Average credit access over the 2 previous years, by round and membership (hh. with school-age children)

| | SHG members | | | | Nonmembers | | | | Controls | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| | R1 | R2 | R3 | R4 | R1 | R2 | R3 | R4 | R2 | R3 | R4 |
| Has taken credit | 0.64 | 0.88 | 0.92 | 0.89 | 0.56 | 0.55 | 0.45 | 0.61 | 0.51 | 0.52 | 0.74 |
| Total credit if > 0 (INR) | 4,873 | 2,244 | 4,048 | 5,319 | 6,955 | 3,979 | 3,932 | 5,350 | 4,525 | 5,777 | 8,855 |
| Total credit (INR) | 3,209 | 1,970 | 3,705 | 4,769 | 4,026 | 2,247 | 1,792 | 3,392 | 2,297 | 2,988 | 6,474 |
| Number of loans | 1.0 | 1.9 | 2.1 | 1.7 | 0.9 | 0.7 | 0.5 | 0.7 | 0.6 | 0.6 | 0.9 |
| Credit constrained† | | 0.34 | 0.53 | 0.16 | | 0.55 | 0.75 | 0.28 | 0.49 | 0.78 | 0.30 |
| Has taken credit for educ.‡ | 0.02 | | 0.14 | 0.07 | 0.01 | | 0.02 | 0.03 | | 0.04 | 0.04 |

† Households who have experienced an unexpected urgent cash need during the year before the survey and did not borrow in order to meet it (but sold assets, exchanged labor or postponed the expense instead). Information not available in round 1.

‡ Education was not recorded as a separate purpose in round 2.

Figure 2: Credit in SHG households as a function of children schooling




our baseline enrollment regressions controlling for the total amount of credit taken in the two previous years (expressed in hundreds of roupies).[23] If credit was an important channel through which the SHG effect was coming, we should observe a positive coefficient for the credit variable and a reduction in the coefficient membervillageXr4. We find that the coefficient of the credit variable is positive but very small in all regressions (a credit of 10,000 INR being associated with a 4 percentage point increase in the probability of being enrolled). Interestingly, the effect is slightly larger for secondary-school age children. More importantly, controlling for credit does not modify the estimates of our main treatment coefficients. As explained in section 3, the major constraints on the education of children are unlikely to be financial, at least up to 14 years. In our data, not only households do not borrow much for education purposes, but they also very rarely mention education as the cause of an urgent need of cash (though this might

[23]Because credit explicitely earmarked for education purposes is rare in our data, the questionnaire in round 2 did not specify education as a separate purpose and credit is eminently fungible, we prefer to focus on total credit, i.e. all credit taken last year for any purpose.

be related to the predictability of education expenses).[24] Note that we find similar results using alternative measures such as having borrowed or not (dummy variable), number of loans taken, etc. Moreover, the interpretation remains difficult since loan amount is correlated with household income.

Table 9: Effect of credit on school enrollment: ITT analysis

| | children aged 12-17 | | children aged 15-17 | | children aged 12-17 | | children aged 15-17 | |
|---|---|---|---|---|---|---|---|---|
| | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) |
| membervillage | -0.0861 | | -0.0989 | | -0.111* | | -0.134 | |
| | (0.0608) | | (0.0812) | | (0.0641) | | (0.0845) | |
| membervilXr3 | 0.0214 | 0.0122 | 0.0902 | 0.0985 | 0.0248 | 0.0132 | 0.0868 | 0.100 |
| | (0.0747) | (0.0749) | (0.105) | (0.106) | (0.0753) | (0.0755) | (0.106) | (0.107) |
| membervilXr4 | 0.181** | 0.190** | 0.247** | 0.264** | 0.161** | 0.164** | 0.223* | 0.236** |
| | (0.0774) | (0.0792) | (0.112) | (0.112) | (0.0789) | (0.0805) | (0.115) | (0.115) |
| loanamount | 0.000374** | 0.000362** | 0.000540*** | 0.000538*** | 0.000364 | 0.000364 | 0.000547** | 0.000576** |
| | (0.000187) | (0.000175) | (0.000189) | (0.000184) | (0.000255) | (0.000234) | (0.000244) | (0.000234) |
| memberhh | | | | | 0.0467 | 0.0341 | 0.0593 | 0.0593 |
| | | | | | (0.0397) | (0.0412) | (0.0500) | (0.0501) |
| loanXmemberhh | | | | | 0.00000236 | 0.0000179 | 0.0000396 | -0.0000428 |
| | | | | | (0.000313) | (0.000300) | (0.000322) | (0.000314) |
| $N$ | 1704 | 1704 | 874 | 874 | 1663 | 1663 | 851 | 851 |
| $R^2$ | 0.226 | 0.248 | 0.189 | 0.221 | 0.229 | 0.254 | 0.191 | 0.229 |
| village controls | yes | no | yes | no | yes | no | yes | no |
| household controls | yes | yes | yes | yes | yes | yes | yes | yes |
| round FE | yes | yes | yes | yes | yes | yes | yes | yes |
| district FE | yes | no | yes | no | yes | no | yes | no |
| village FE | no | yes | no | yes | no | yes | no | yes |

Std errors clustered at the household level in parentheses (*p<0.10, **p<0.05, ***p<0.01).
All equations control for the age and birth rank of children, and the monsoon quality in year t-1.
Observations weighted in order to account for the different sampling probabilities.

As a final check of this financial channel, table 10 regresses measures of per-capita annual income and consumption on our usual treatment variables. Per-capita consumption is computed by adding up the annualized market value of all food and fuel consumption, as well as other expenses on clothing, sundry and other personal articles, services and rents, festivals and health. We focus on households with children of the relevant age (12-17). Our income aggregate incorporates the following sources of income that the households could earn over the last year: wage and casual labor income, self-employment income, monetary value of agricultural and livestock production (i.e. sales proceeds plus value of home-consumption, estimated at median market prices), net rental income from the supply of land, credit or other assets, net current transfers. Per-capita measures are computed using the equivalence scale proposed by Townsend (1994) and log transformed.[25] Surprisingly, there is no effect of SHG on the two measures of income / con-

[24]The main reported credit purposes are: medical (30%), farming/business (29%), consumption (25%) and social/family events (12%). Likewise, the major reported causes of urgent cash needs are: medical (54%), farming/business (12%), consumption (10%) and social/family events (9.5%). Education accounts only for 4% of credit events and 3% of urgent cash needs.

[25]Townsend computes male-adult equivalent consumption using the following age-sex weights (estimated from a dietary survey in rural Andhra Pradesh and Maharastra): for adult males, 1.0; for adult females, 0.9; for males and females aged 13-18, 0.94 and 0.83, respectively; for children aged 7-12, 0.67 regardless of gender; for children 4-6, 0.52; for toddlers 1-3, 0.32; and for infants 0.05.

sumption, neither at the village level (ITT) nor at the member level (ATT and ITT_member). Therefore, the increase in enrollment cannot be explained by an increase in income in treated villages. Moreover, the correlation between enrollment rates and consumption / income levels is never significant across various specifications. This therefore suggests that the direct cost of education is not a major issue up to 14 years (instead, the opportunity cost of education might represent a more serious constraint).

Table 10: Per-capita income and consumption

| | Per-capita consumption | | | Per-capita income | | |
|---|---|---|---|---|---|---|
| | ITT | ATT | ITT_member | ITT | ATT | ITT_member |
| | (1) | (2) | (3) | (4) | (5) | (6) |
| membervilXr3 | 0.0231 | 0.235** | 0.0246 | -0.256 | -0.112 | -0.261 |
| | (0.0874) | (0.0994) | (0.0922) | (0.158) | (0.151) | (0.186) |
| membervilXr4 | -0.122 | 0.0152 | -0.141 | -0.219 | 0.0674 | -0.228 |
| | (0.0925) | (0.116) | (0.0994) | (0.156) | (0.174) | (0.177) |
| memberhh | | | -0.102* | | | -0.125 |
| | | | (0.0615) | | | (0.125) |
| mbXr3 | | | 0.00341 | | | 0.0330 |
| | | | (0.0782) | | | (0.163) |
| mbXr4 | | | 0.0283 | | | 0.0452 |
| | | | (0.0749) | | | (0.154) |
| $N$ | 1707 | 1067 | 1666 | 1669 | 1040 | 1628 |
| $R^2$ | 0.291 | 0.276 | 0.301 | 0.259 | 0.318 | 0.265 |

Std errors clustered at the household level in parentheses (*p<0.10, **p<0.05, ***p<0.01).
All equations include village and round fixed effects, household controls, and monsoon quality in year t-1.
Observations weighted in order to account for the different sampling probabilities in col. 1,3, 4, and 6.
Observations weighted by their propensity score in col. 2 and 5.

## 6.3 Child labor

The household questionnaire also asked about the time allocation of children, which allows us to check more directly the evolution of child labor practices. We recorded detailed information on both income-earning activities (both in and out of the household) and domestic chores (including child care, fuel wood and water collection). As for education, we include children who migrated temporarily to work. Focusing on productive labor, 28% of school-age boys in the sample and 20% of school-age girls are reported as working outside of the household in round 2, respectively for 14.3 and 13.1 weekly hours on average (conditionally on working). In the sample, child labor is actually limited before 12 years and becomes fairly generalized afterwards, with participation rates reaching 34 and 25% for boys and girls respectively. We therefore focus on children aged 12-17 as above. Statistically, it is this type of productive work outside of the household that is strongly negatively associated with school enrollment, while helping with household chores is instead much easier to accomodate with school. Indeed, for girls, the correlation between enrollment and productive work is -14% (significant at 1% level), while it is actually significantly positive for household chores. Interestingly, such tradeoff seem to apply less for boys on average, for whom we find no significant correlation between school and productive labor.

Table 11: Productive labor of children aged 10-17: descriptive statistics

| | Boys | | | | | | Girls | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | Members | | | Other households | | | Members | | | Other households | | |
| | R2 | R3 | R4 | R2 | R3 | R4 | R2 | R3 | R4 | R2 | R3 | R4 |
| Any work (%) | 28.3 | 56.6 | 63.5 | 38.9 | 61.2 | 61.1 | 26.7 | 61.2 | 64.0 | 24.8 | 58.3 | 66.8 |
| Weekly hours: uncond. avg | 3.9 | 4.7 | 5.6 | 7.7 | 5.3 | 6.3 | 4.6 | 4.2 | 4.0 | 2.9 | 3.8 | 4.7 |
| Weekly hours: cond. avg | 13.8 | 8.2 | 8.7 | 19.7 | 8.6 | 10.2 | 17.2 | 6.8 | 6.2 | 11.7 | 6.6 | 7.1 |
| *Observations* | 254 | 242 | 252 | 244 | 242 | 285 | 240 | 245 | 275 | 234 | 211 | 220 |

We investigate the impact of SHG membership on three different measures of child labor: total number of hours worked, probability to work, and hours worked if working. We do not find any significant impact of SHG membership, which indicates that child labor and school enrolment are not necessarily substitutes in the context at hand. In fact, in table 13, we observe a positive correlation between enrollment and total hours worked and the probability to work, though enrolled children tend to work less hours. Only productive labor appears to be negatively correlated with enrollment. The complementarity between child enrollment and labor is also made clear from figure 3.

That is, unlike some studies that were carried in different contexts such as Augsburg et al. (2012) or Wydick (1999), we do not find any evidence that having access to microloans increases child labor.

We also investigated the relationship between child labor and land ownership. Interestingly, there is no relation between the total number of hours worked and land ownership. For productive labor, it appears that children in families with more land are more likely to work but for shorter times so that in total they do not work more. We also find no evidence of an inverted-U shape between child labor and land ownership as hypothesized by (Basu et al., 2010). Similarly, the

Figure 3: Hours of total labor and enrollment of children

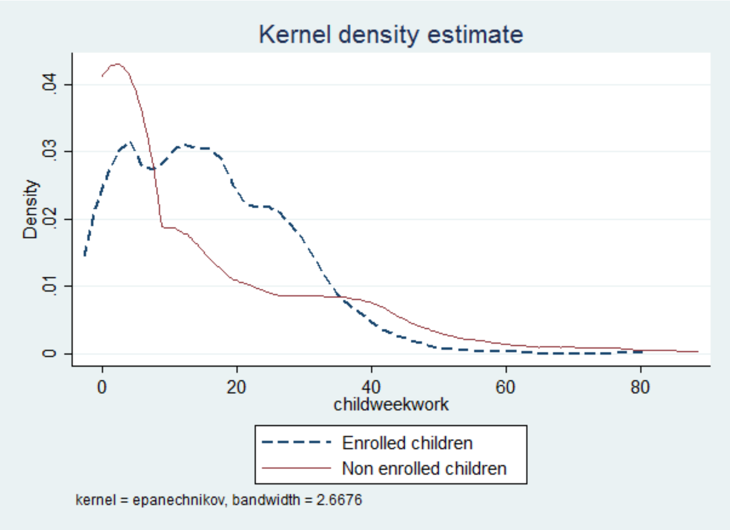



Table 12: Child labor

| | Total labor | | | Productive labor | | | Domestic labor | | |
|---|---|---|---|---|---|---|---|---|---|
| | hours | work dummy | hours if >0 | hours | work dummy | hours if >0 | hours | work dummy | hours if >0 |
| | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) | (9) |
| membervilXr3 | 0.927 | 0.0573 | -1.615 | 1.030 | 0.0719 | -2.293 | -0.104 | -0.0300 | -0.296 |
| | (2.963) | (0.0772) | (2.967) | (1.566) | (0.0806) | (3.145) | (2.096) | (0.0943) | (2.158) |
| membervilXr4 | -0.339 | 0.0789 | -4.020 | -0.672 | 0.0874 | -6.278* | 0.333 | 0.0167 | -0.323 |
| | (2.665) | (0.0735) | (2.948) | (1.669) | (0.0740) | (3.196) | (1.829) | (0.0904) | (1.969) |
| female | 5.723*** | 0.128*** | 4.315*** | -2.230*** | 0.00950 | -3.937*** | 7.953*** | 0.291*** | 5.579*** |
| | (0.949) | (0.0223) | (0.885) | (0.500) | (0.0291) | (0.713) | (0.747) | (0.0277) | (0.723) |
| $N$ | 1704 | 1704 | 1306 | 1704 | 1704 | 954 | 1704 | 1704 | 1097 |
| $R^2$ | 0.126 | 0.171 | 0.180 | 0.105 | 0.216 | 0.218 | 0.217 | 0.215 | 0.258 |

Std errors clustered at the household level in parentheses (*p<0.10, **p<0.05, ***p<0.01).
All equations include village and round fixed effects, household controls, the age and birth rank of children, and the monsoon quality in year t-1.
Observations weighted in order to account for the different sampling probabilities.

Table 13: Child labor with enrollment

| | Total labor | | | Productive labor | | | Domestic labor | | |
|---|---|---|---|---|---|---|---|---|---|
| | hours | work dummy | hours if >0 | hours | work dummy | hours if >0 | hours | work dummy | hours if >0 |
| | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) | (9) |
| enrolled | 3.479*** | 0.289*** | -3.856*** | -1.567*** | 0.0992*** | -4.586*** | 5.046*** | 0.288*** | 1.094 |
| | (0.917) | (0.0255) | (1.027) | (0.553) | (0.0295) | (0.969) | (0.685) | (0.0321) | (0.752) |
| membervilXr3 | 0.872 | 0.0528 | -1.502 | 1.055 | 0.0703 | -1.911 | -0.183 | -0.0345 | -0.304 |
| | (2.918) | (0.0728) | (2.950) | (1.566) | (0.0798) | (3.047) | (2.008) | (0.0900) | (2.152) |
| membervilXr4 | -0.986 | 0.0252 | -3.287 | -0.381 | 0.0690 | -5.049 | -0.605 | -0.0369 | -0.452 |
| | (2.637) | (0.0677) | (2.841) | (1.651) | (0.0736) | (3.077) | (1.778) | (0.0849) | (1.974) |
| female | 5.910*** | 0.144*** | 4.008*** | -2.314*** | 0.0148 | -4.255*** | 8.223*** | 0.306*** | 5.711*** |
| | (0.918) | (0.0225) | (0.903) | (0.509) | (0.0286) | (0.720) | (0.704) | (0.0278) | (0.702) |
| $N$ | 1704 | 1704 | 1306 | 1704 | 1704 | 954 | 1704 | 1704 | 1097 |
| $R^2$ | 0.136 | 0.257 | 0.193 | 0.110 | 0.224 | 0.249 | 0.258 | 0.282 | 0.260 |

Std errors clustered at the household level in parentheses (*p<0.10, **p<0.05, ***p<0.01).
All equations include village and round fixed effects, household controls, the age and birth rank of children, and the monsoon quality in year t-1.
Observations weighted in order to account for the different sampling probabilities.

effect of SHG membership does not differ according to the agricultural profile of the household, a result that has sometimes been highlighted in the literature (e.g Maldonado and Gonzalez-Vega, 2008).

## 6.4 Heterogeneity analysis and mechanisms

In what follows, we try to explore alternative channels.

First, we investigate the importance of the distance to secondary school, which arguably represents an important constraint on children's education. Indeed, while primary and middle schools are either in or close to every village in our sample, secondary schools can be relatively far and thus impose important travel costs and risks. For each village, we take the median travel time reported by children enrolled in secondary school, and then classify villages in two categories, below and above median distance (40 minutes). From table 14, we observe that the effect on enrollment is entirely driven by villages far from secondary school. It is also in those villages that child labor tends to increase, illustrating again the positive correlation between enrollment and child labor. In the interest of space, we only report total labor, but productive labor follows a similar pattern. At baseline, those villages tend to have a stronger agricultural occupational base (with households living predominantly out of agriculture representing 56% of the population as against 43% in other villages) and lower income levels (with an average per-capita income about 14% lower).

Table 14: Child enrollment and labor, by distance to secondary school: ITT analysis

| | Villages close to secondary school | | | | Villages far from secondary school | | | |
|---|---|---|---|---|---|---|---|---|
| | Enrollment | | Labor | | Enrollment | | Labor | |
| | 12-17 | 15-17 | hours | dummy | 12-17 | 15-17 | hours | dummy |
| | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) |
| membervilXr3 | -0.0258 | 0.0347 | 0.546 | 0.0811 | 0.0313 | -0.0760 | 3.827 | 0.0924 |
| | (0.105) | (0.143) | (3.416) | (0.101) | (0.136) | (0.212) | (5.263) | (0.0998) |
| membervilXr4 | 0.176 | 0.114 | -1.054 | -0.0231 | 0.399*** | 0.578*** | 5.900 | 0.313** |
| | (0.113) | (0.165) | (3.603) | (0.0953) | (0.133) | (0.195) | (4.105) | (0.131) |
| $N$ | 748 | 375 | 748 | 748 | 843 | 435 | 843 | 843 |
| $R^2$ | 0.279 | 0.227 | 0.147 | 0.183 | 0.237 | 0.235 | 0.152 | 0.189 |

Std errors clustered at the household level in parentheses (*p<0.10, **p<0.05, ***p<0.01).
All equations include village and round fixed effects, as well as household controls.
All equations control for the age and birth rank of children, and the monsoon quality in year t-1.
Observations weighted in order to account for the different sampling probabilities.

In table 15, we observe that SHGs tend to increase the emancipation of women over time, which points at a 'transformative' role of SHGs. The survey asked a series of questions to women (in the absence of their husband) aimed at assessing their relative awareness. Those consist of 11 questions assessing basic (financial) literacy, legal and health knowledge. The other measure of emancipation we use is the number of trips out of the village that women made over the month before the survey. In table 16, we observe that female emancipation is positively correlated with enrollment rates of children. Columns 1-4 re-run the main enrollment equations and control directly for our emancipation variables. We find a positive coefficient for emancipation, while the main treatment coefficients only slightly decrease. In columns 5-8, we condition the evolution of education outcomes on the level of baseline awareness. We observe that gains from SHG participation are larger in households in which women are relatively more aware.

Those observations indicate that the enrollment effect of the SHG treatment goes at least partly through an increased awareness of women. This is consistent with other studies that showed women's empowerment to be one of the most important consequences of female participation to savings and credit groups, because of the access to financial resources as well as the support from the group of peers (e.g. Ashraf et al., 2010; Bali Swain and Wallentin, 2012; Desai and Joshi, 2013). It is also in line with several papers showing that empowering women usually leads to improvement in household's and especially children's welfare (e.g. Ashraf et al., 2010; Duflo, 2012). In particular, our observations are close to recent evidence about gender quotas for elected positions on village councils in India. Indeed, Beaman et al. (2012) find that, several years after a woman held an electoral mandate, girls spend significantly less time on household activities and get more education, up to the point where the gender gap in education outcomes is completely erased (and even reversed). They argue that two main channels might be at work: first, women in leadership positions can undertake policies that make it easier for women to succeed, thus changing beliefs on what is possible for girls; and second, they can provide a role model of a successful woman. Given that they find no evidence of changes in young women's labor market opportunities and that they find strong effects on aspirations of parents and adolescents, they favor the second explanation. In the case of SHGs, Casini and Vandewalle (2011) show that member women, by undertaking collective actions, manage to somehow influence the political agenda of local authorities. Moreover, SHGs certainly serve to give more pride and confidence to their members, because of the access to financial resources as well as support from the group of peers.

Table 15: Emancipation indicators of females: ITT analysis

| | Awareness score | | No trip out of village last month | |
|---|---|---|---|---|
| | with 12-17 kids | with 15-17 kids | with 12-17 kids | with 15-17 kids |
| | (1) | (2) | (3) | (4) |
| membervilXr3 | 0.00896 | 0.214 | -0.249*** | -0.230** |
| | (0.396) | (0.478) | (0.0889) | (0.105) |
| membervilXr4 | 0.616 | 0.845** | -0.190** | -0.0777 |
| | (0.376) | (0.429) | (0.0967) | (0.111) |
| $N$ | 1679 | 865 | 1707 | 876 |
| $R^2$ | 0.371 | 0.413 | 0.189 | 0.242 |

Std errors clustered at the household level in parentheses (*p<0.10, **p<0.05, ***p<0.01).
All equations include village and round fixed effects, household controls and monsoon quality in t-1.
Observations weighted in order to account for the different sampling probabilities.

Table 16: Enrollment of children and mother's emancipation: ITT analysis

| | Controlling for current emancipation | | | | Comparing baseline emancipation levels | | | |
|---|---|---|---|---|---|---|---|---|
| | 12-17 | 15-17 | 12-17 | 15-17 | low_trip | high_trip | low_aware | high_aware |
| | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) |
| membervilXr3 | 0.0113 | 0.105 | 0.0123 | 0.0978 | -0.0863 | 0.0885 | -0.0743 | 0.0872 |
| | (0.0763) | (0.109) | (0.0756) | (0.107) | (0.146) | (0.0888) | (0.112) | (0.0965) |
| membervilXr4 | 0.175** | 0.258** | 0.183** | 0.259** | 0.0127 | 0.260*** | 0.0689 | 0.298*** |
| | (0.0808) | (0.117) | (0.0806) | (0.115) | (0.166) | (0.0984) | (0.118) | (0.108) |
| awareness_score | 0.0107* | 0.0170* | | | | | | |
| | (0.00627) | (0.0101) | | | | | | |
| no_trip | | | -0.0133 | -0.0474 | | | | |
| | | | (0.0287) | (0.0437) | | | | |
| $N$ | 1676 | 863 | 1704 | 874 | 504 | 1200 | 758 | 946 |
| $R^2$ | 0.246 | 0.215 | 0.245 | 0.214 | 0.396 | 0.228 | 0.308 | 0.280 |

Std errors clustered at the household level in parentheses (*p<0.10, **p<0.05, ***p<0.01).

All equations include village and round fixed effects, as well as household controls.

All equations control for the age and birth rank of children, and the monsoon quality in year t-1.

Observations weighted in order to account for the different sampling probabilities.

## 7 Conclusion

In this paper, we investigate the effects of participation to microfinance on education decisions in India. More specifically, we compare the evolution over time of school enrollment between participants and non participants to SHGs, the dominant form of microfinance in India, by using an original panel data set collected in Jharkhand (North India) between 2002 and 2009.

We find substantial evidence that SHG membership increases boys' and girls' school enrollment, through a lower dropout rate in non-compulsory education. The average treatment effect of SHG participation on children's post-compulsory education enrollment is estimated to be around 25 percentage points, a very impressive improvement. Moreover, we find that the SHG intervention was also successful in raising post-compulsory education enrollment rates at the village level.

There is some evidence that credit helps, though it is not main the main channel at work. Instead, SHG participation seems to induce a better organization and an empowerment of women within the household, and to allow villagers' coordination to send children to far-away secondary schools.

The effects are particularly important for the last round of the survey, which suggests that these effects may take a long time to materialize (in this case, 7 years after the start of the SHG program in the area). This may explain why most recent RCT evaluations have found very limited impact of microcredit. Moreover, we find that SHG participation is particularly important for female education, thus closing the important gender gap observed at baseline for primary and secondary levels.

In conclusion, informal microfinance groups as SHGs, by significantly improving girls' education levels, can have potentially wide-ranging development impacts in the long run. Further research is needed to understand better the channels and mechanisms at work.

## References

**Angelucci, Manuela and Giacomo De Giorgi**, "Indirect Effects of an Aid Program: How Do Cash Transfers Affect Ineligibles' Consumption?," *American Economic Review*, March 2009, *99* (1), 486–508.

_ , **Dean Karlan, and Jonathan Zinman**, "Microcredit Impacts: Evidence from a Randomized Microcredit Program Placement Experiment by Compartamos Banco," *American Economic Journal: Applied Economics*, January 2015, *7* (1), 151–82.

**Armendáriz, Beatriz and Jonathan Morduch**, *The Economics of Microfinance*, 2nd ed., Cambridge, Massachussets: The MIT Press, 2010.

**Ashraf, Nava, Dean Karlan, and Wesley Yin**, "Female Empowerment: Impact of a Commitment Savings Product in the Philippines," *World Development*, March 2010, *38* (3), 333–344.

**Augsburg, Britta, Ralph De Haas, Heike Harmgart, and Costas Meghir**, "Microfinance, Poverty and Education," IFS Working Papers W12/15, Institute for Fiscal Studies 2012.

**_ , _ , _ , and _** , "The Impacts of Microcredit: Evidence from Bosnia and Herzegovina," *American Economic Journal: Applied Economics*, January 2015, *7* (1), 183–203.

**Bali Swain, Ranjula**, *The Microfinance Impact*, London and New York: Routledge, 2012.

**_ and Adel Varghese**, "Does Self Help Group Participation Lead to Asset Creation?," *World Development*, 2009, *37* (10), 1674–1682.

**_ and Fan Yang Wallentin**, "Factors empowering women in Indian self-help group programs," *International Review of Applied Economics*, 2012, *26* (4), 425–444.

**Banerjee, Abhijit, Dean Karlan, and Jonathan Zinman**, "Six Randomized Evaluations of Microcredit: Introduction and Further Steps," *American Economic Journal: Applied Economics*, January 2015, *7* (1), 1–21.

**Basu, Kaushik, Sanghamitra Das, and Bhaskar Dutta**, "Child labor and household wealth: Theory and empirical evidence of an inverted-U," *Journal of Development Economics*, January 2010, *91* (1), 8–14.

**Beaman, Lori, Esther Duflo, Rohini Pande, and Petia Topalova**, "Female Leadership Raises Aspirations and Educational Attainment for Girls: A Policy Experiment in India," *Science*, 2012, *335* (6068), 582–586.

**Beegle, Kathleen, Rajeev H. Dehejia, and Roberta Gatti**, "Child labor and agricultural shocks," *Journal of Development Economics*, October 2006, *81* (1), 80–96.

**Behrman, Jere R.**, "Intrahousehold distribution and the family," in M. R. Rosenzweig and O. Stark, eds., *Handbook of Population and Family Economics*, Vol. 1 of *Handbook of Population and Family Economics*, Elsevier, 1997, chapter 4, pp. 125–187.

**Berhane, Guush and Cornelis Gardebroek**, "Does Microfinance Reduce Rural Poverty? Evidence Based on Household Panel Data from Northern Ethiopia," *American Journal of Agricultural Economics*, 2011, *93* (1), 43–55.

**Bertrand, Marianne, Esther Duflo, and Sendhil Mullainathan**, "How much should we trust in differences-in-diffrences estimates?," *The Quarterly Journal of Economics*, 2004, *119* (1), 249–275.

**Biswal, K.**, "Secondary Education in India: Development Policies, Programmes and Challenges," Research Monograph 63, Create April 2011.

**Blume, Jonas and Julika Breyer**, "Microfinance and Child Labour," Employment Working Paper 89, International Labour Organization August 2011.

**Burgess, Robin and Rohini Pande**, "Do rural banks matter? Evidence from the Indian social banking experiment," *American Economic Review*, 2005, *95* (3), 780–795.

**Busso, Matias, John DiNardo, and Justin McCrary**, "New Evidence on the Finite Sample Properties of Propensity Score Matching and Reweighting Estimators," IZA Discussion Papers 3998, Institute for the Study of Labor (IZA) February 2009.

**Casini, Paolo and Lore Vandewalle**, "Public Good Provision in Indian Rural Areas: the Returns to Collective Action by Microfinance Groups," Working Papers 1119, University of Namur, Department of Economics December 2011.

**CGAP**, "Sustainability of Self-Help Groups in India: two analyses," Occasional Paper 12, Consultative Group to Assist the Poor August 2007.

**Cigno, Alessandro and Furio Camillo Rosati**, *The Economics of Child Labour*, Oxford University Press, 2005.

**Crépon, Bruno, Florencia Devoto, Esther Duflo, and William Parienté**, "Estimating the Impact of Microcredit on Those Who Take It Up: Evidence from a Randomized Experiment in Morocco," *American Economic Journal: Applied Economics*, January 2015, *7* (1), 123–50.

**Crump, Richard K., V. Joseph Hotz, Guido W. Imbens, and Oscar A. Mitnik**, "Dealing with limited overlap in estimation of average treatment effects," *Biometrika*, 2009, *96* (1), 187–199.

**Dave, Harishkumar R. and Hans Dieter Seibel**, "Commercial Aspects of Self-Help Group Banking in India: A Study of Bank Transaction Costs," Working Papers 2002,7, University of Cologne, Development Research Center 2002.

**Deininger, Klaus and Yanyan Liu**, "Evaluating Program Impacts on Mature Self-help Groups in India," *The World Bank Economic Review*, 2013, *27* (2), 272–296.

**Demont, Timothée**, "Microfinance Spillovers: a Model and some Facts about Competition on Informal Credit Markets," CRED Working Paper, Namur University 2012.

_ , "Poverty, Access to Credit and Absorption of Income Shocks: Evidence from Self-Help Groups in India," Department of Economics Working Paper 2014/10, University of Namur 2014.

**Desai, Raj M. and Shareen Joshi**, "Collective Action and Community Development: Evidence from Self-Help Groups in Rural India," *The World Bank Economic Review*, 2013.

**Dewan, Isha and Rohini Somanathan**, "Poverty targeting in public programs: a comparison of some non parametric tests and application to indian microfinance," Working Paper, Delhi School of Economics 2007.

**Duflo, Esther**, "Women Empowerment and Economic Development," *Journal of Economic Literature*, September 2012, *50* (4), 1051–79.

_ **, Abhijit Banerjee, Rachel Glennerster, and Cynthia G. Kinnan**, "The Miracle of Microfinance? Evidence from a Randomized Evaluation," *American Economic Journal: Applied Economics*, January 2015, *7* (1), 22–53.

**Duryea, Suzanne, David Lam, and Deborah Levison**, "Effects of economic shocks on children's employment and schooling in Brazil," *Journal of Development Economics*, September 2007, *84* (1), 188–214.

**Flory, Jeffrey**, "Formal Savings Spillovers on Microenterprise Growth and Production Decisions Among Non-Savers in Villages: Evidence from a Field Experiment," Proceedings of the Annual Meeting, Agricultural and Applied Economics Association 2012.

**Guarcelllo, Lorenzo, Fabrizia Mealli, and Furio Camillo Rosati**, "Household vulnerability and child labor : the effect of shocks, credit rationing and insurance," Social Protection Discussion Papers 29136, The World Bank November 2003.

**Hirano, Keisuke, Guido Imbens, and Geert Ridder**, "Efficient estimation of average treatment effects using the estimated propensity score," *Econometrica*, 2003, *71* (4), 1161–1189.

**Hyder, Asma, Jere R. Behrman, and Hans-Peter Kohler**, "Negative Economic Shocks and Child Schooling: Evidence from Rural Malawi," PIER Working Paper Archive 12-039, Penn Institute for Economic Research, Department of Economics, University of Pennsylvania September 2012.

**Imbens, Guido W.**, "Nonparametric Estimation of Average Treatment Effects Under Exogeneity: A Review," *The Review of Economics and Statistics*, February 2004, *86* (1), 4–29.

**Jacoby, Hanan G and Emmanuel Skoufias**, "Risk, Financial Markets, and Human Capital in a Developing Country," *Review of Economic Studies*, July 1997, *64* (3), 311–35.

**Jensen, Robert**, "Agricultural Volatility and Investments in Children," *American Economic Review*, May 2000, *90* (2), 399–404.

**Kaboski, Joseph P. and Robert M. Townsend**, "The Impact of Credit on Village Economies," *American Economic Journal: Applied Economics*, April 2012, *4* (2), 98–133.

**Karlan, Dean, Aishwarya Lakshmi Ratan, and Jonathan Zinman**, "Savings by and for the poor: A research review and agenda," Discussion Paper 1027, Yale Economic Growth Center October 2013.

_ **and Jonathan Zinman**, "Expanding Microenterprise Credit Access: Using Randomized Supply Decisions to Estimate the Impacts in Manila," Working Paper, Innovation for Poverty Action 2010.

**Kingdon, Geeta Gandhi**, "Where Has All the Bias Gone? Detecting Gender Bias in the Intrahousehold Allocation of Educational Expenditure," *Economic Development and Cultural Change*, January 2005, *53* (2), 409–51.

_ , "The Progress of School Education in India," Working Paper 71, Global Poverty Research Group 2007.

**Littlefield, Elizabeth, Syed Hashemi, and Jonathan Morduch**, "Is Microfinance an Effective Strategy to Reach the Millennium Development Goals?," Focus Note 24, CGAP 2003.

**Maldonado, Jorge H. and Claudio Gonzalez-Vega**, "Impact of Microfinance on Schooling: Evidence from Poor Rural Households in Bolivia," *World Development*, November 2008, *36* (11), 2440–2455.

**Matsuura, Kenji and Cort J. Willmott**, "Terrestrial Precipitation: 1900-2010 Gridded Monthly Time Series (version 3.01)," Center for Climatic Research database, University of Delaware 2012.

**Menon, Nidhiya**, "Investment credit and child labour," *Applied Economics*, 2010, *42* (12), 1461–1479.

**Menon, P., A. Deolalikar, and A. Bhaskar**, "India state hunger index: Comparisons of hunger across states," Technical Report, International Food Policy Research Institute, Washington, D.C. 2008.

**MicroSave**, "Deposit Assessment in India," Technical Report, IFC 2011.

**Muralidharan, Karthik and Nishith Prakash**, "Cycling to School: Increasing Secondary School Enrollment for Girls in India," Working Paper 19305, National Bureau of Economic Research August 2013.

**NABARD**, "Guidelines for the Pilot Project for linking banks with Self Help Groups," Circular issued to all Commercial Banks, Mumbai 1992.

_ , "Status of microfinance in India 2012-2013," Technical Report, National Bank for Agriculture and Rural Development, Mumbai 2013.

**PRADAN**, "Handbook on promoting SHGs," Technical Report, New Delhi 2002.

**Prina, Silvia**, "Banking the poor via savings accounts: evidence from a field experiment," 2013. Mimeo.

**RBI**, "Swarnajayanti Gram Swarozgar Yojana," Circular issued to all Scheduled Commercial Banks, Mumbai 1999.

**Rose, Elaina**, "Gender Bias, Credit Constraints and Time Allocation in Rural India," *Economic Journal*, July 2000, *110* (465), 738–58.

**Townsend, Robert**, "Risk and Insurance in Village India," *Econometrica*, May 1994, *62* (3), 539–91.

**Wooldridge, Jeffrey M**, *Econometric Analysis of Cross Section and Panel Data*, 2nd ed., The MIT Press, 2010.

**World Bank**, "Secondary Education in India: Universalizing Opportunity," Document 48521, The World Bank 2009.

**Wydick, Bruce**, "The Effect of Microenterprise Lending on Child Schooling in Guatemala," *Economic Development and Cultural Change*, 1999, *47* (4), 853–869.

**Zimmermann, Laura**, "Reconsidering Gender Bias in Intra-Household Allocation in India," *Journal of Development Studies*, 2012, *48* (1), 151–163.

# A Sample districts and villages

Table 17: Sample villages and district

| Region | District | Village | Type |
|---|---|---|---|
| Northeast | Banka† | Fattapathar | Member |
| Northeast | Banka† | Kanibel | Member |
| Northeast | Banka† | Devhar | Control |
| Northeast | Banka† | Bagmunda | Member |
| Northeast | Dumka | Gwalshimla | Member |
| Northeast | Dumka | Sitasal | Member |
| Northeast | Dumka | Tetriya | Member |
| Northeast | Dumka | Barhet | Control |
| Northeast | Dumka | Ranga | Control |
| Central | Hazaribagh | Bigha | Member |
| Central | Hazaribagh | Debo | Member |
| Central | Hazaribagh | Ranik | Member |
| Central | Hazaribagh | Rupin | Control |
| Central | Koderma | Garhai | Member |
| Central | Koderma | Irgobad | Member |
| Central | Koderma | Saanth | Member |
| Central | Koderma | Lariyadih | Control |
| Southeast | E. Singhbhum | Haldipokhar | Member |
| Southeast | E. Singhbhum | Murasai | Member |
| Southeast | E. Singhbhum | Pukhuria | Member |
| Southeast | E. Singhbhum | Pathar Banga | Control |
| Southeast | W. Singhbhum | Baihatu | Member |
| Southeast | W. Singhbhum | Chandra Jarki‡ | Member |
| Southeast | W. Singhbhum | Kera | Member |
| Southeast | W. Singhbhum | Mermera | Member |
| Southeast | W. Singhbhum | Unchibita | Member |
| Southeast | W. Singhbhum | Jarki | Control |
| Southeast | W. Singhbhum | Nakti | Control |
| Southwest | Gumla | Jaldega | Member |
| Southwest | Gumla | Semra | Member |
| Southwest | Gumla | Umra | Member |
| Southwest | Gumla | Kurum | Control |
| Southwest | Khunti | Banabira | Member |
| Southwest | Khunti | Bhandara | Member |
| Southwest | Khunti | Udikel | Member |
| Southwest | Khunti | Irud | Control |
| Southwest | Khunti | Kamra | Control |

*Notes:* † Bihar. ‡ Chandra Jarki replaced Kera in round 4 due to insecurity reasons.

Table 18: District poverty (data from 2001 Census if not otherwise indicated)

| District | Population (thousands) | BPL households[1] | SC (%) | ST (%) | Female literacy (%) | Infant mortality (‰) | Households electrified (%)[2] |
|---|---|---|---|---|---|---|---|
| Banka | 1,608.8 | 215,784 | 12.4 | 4.7 | 28.7 | 56 | 4.7 |
| Dumka | 1,759.6 | 125,701 | 7.3 | 39.9 | 32.3 | 47 | 7.7 / 20.4 |
| Hazaribagh | 2,277.5 | 222,810 | 15.0 | 11.8 | 42.8 | 46 | 34.7 / 57.2 |
| Koderma | 499.4 | 51,282 | 14.4 | 0.8 | 33.6 | 46 | 21.7 / 31.2 |
| E. Singhbhum | 1,983.0 | 117,918 | 4.7 | 27.8 | 57.3 | 36 | 47.4 / 67.1 |
| W. Singhbhum | 2,082.8 | 152,560 | 4.9 | 53.4 | 34.4 | 54 | 16.5 / 22.5 |
| Gumla | 1,346.8 | 87,546 | 5.0 | 68.4 | 39.9 | 60 | 5.1 / 6.8 |
| Khunti | 2,785.1 | 207,187 | 5.2 | 41.8 | 51.7 | 45 | 29.9 / 48.1 |

*Notes:* [1] 2002-07, official BPL list from the Government of Jharkhand (Bihar for Banka).
[2] Figures on the right are from a household survey by the Ministry of Health and Family Welfare in 2002-04.

Figure 4: Districts and agro-climatic zones in Jharkhand

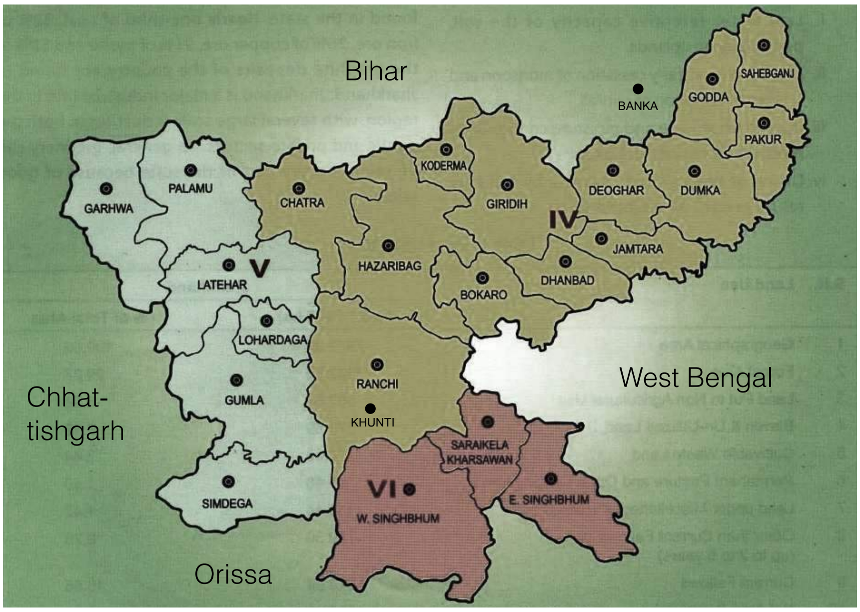


Table 19: Baseline characteristics of treated and control villages and balance check

| | Entire sample | | | Restricted sample | | |
|---|---|---|---|---|---|---|
| | control | treated | p-value treated = control | control | treated | p-value treated = control |
| Population (# households)[1] | 167.4 | 166.4 | 0.977 | 175.0 | 166.6 | 0.846 |
| SC population(%)[1] | 0.107 | 0.114 | 0.891 | 0.135 | 0.116 | 0.784 |
| ST population(%)[1] | 0.473 | 0.464 | 0.958 | 0.387 | 0.486 | 0.612 |
| Landless population (%)[1] | 0.246 | 0.300 | 0.577 | 0.364 | 0.303 | 0.589 |
| Illiterate population (%)[1] | 0.663 | 0.642 | 0.589 | 0.684 | 0.649 | 0.430 |
| Female illiterate population (%)[1] | 0.774 | 0.767 | 0.862 | 0.783 | 0.774 | 0.825 |
| Farming population (%)[1] | 0.352 | 0.366 | 0.892 | 0.235 | 0.353 | 0.232 |
| Working gender-parity index[1] | 0.472 | 0.512 | 0.785 | 0.352 | 0.493 | 0.387 |
| Unemployment (%)[1] | 0.408 | 0.353 | 0.591 | 0.495 | 0.365 | 0.272 |
| Female unemployment (%)[1] | 0.588 | 0.560 | 0.850 | 0.703 | 0.579 | 0.441 |
| Caste / tribe fractionalization[2,4] | 0.583 | 0.512 | 0.504 | 0.592 | 0.522 | 0.580 |
| Language fractionalization[2,4] | 0.347 | 0.358 | 0.888 | 0.347 | 0.352 | 0.957 |
| Religious fractionalization[2,4] | 0.402 | 0.298 | 0.246 | 0.379 | 0.299 | 0.446 |
| Hinduism is main village religion[3] | 0.637 | 0.596 | 0.761 | 0.685 | 0.612 | 0.645 |
| All-weather road reaches village[3] | 0.266 | 0.196 | 0.586 | 0.306 | 0.158 | 0.281 |
| Electricity available in village[3] | 0.403 | 0.439 | 0.840 | 0.500 | 0.413 | 0.683 |
| Irrigated land (%)[3] | 13.33 | 13.34 | 0.999 | 10.92 | 13.50 | 0.670 |
| Distance to nearest bank (km)[3] | 6.028 | 7.284 | 0.506 | 4.875 | 7.357 | 0.238 |
| Distance to nearest primary health center (km)[3] | 5.083 | 5.909 | 0.551 | 5.375 | 5.929 | 0.745 |
| Distance to nearest fair price shop (km)[3] | 2.611 | 4.509 | 0.272 | 2.583 | 4.724 | 0.314 |
| Distance to nearest market (km)[3] | 5.111 | 5.727 | 0.628 | 5.458 | 5.726 | 0.861 |
| Distance to nearest rail station (km)[3] | 23 | 20 | 0.780 | 14.50 | 20.90 | 0.553 |
| Presence of a bus stop in village[3] | 0.278 | 0.205 | 0.655 | 0.250 | 0.214 | 0.852 |
| Distance to nearest bus stop (km)[3] | 2.917 | 3.557 | 0.587 | 2.500 | 3.643 | 0.399 |
| Presence of a primary school in village[3] | 0.778 | 0.773 | 0.973 | 0.833 | 0.762 | 0.667 |
| Presence of a middle school in village[3] | 0.278 | 0.364 | 0.592 | 0.250 | 0.381 | 0.476 |
| Presence of a secondary school in village[3] | 0 | 0.0455 | 0.366 | 0 | 0.0476 | 0.452 |
| Distance to nearest secondary school (km)[3] | 8.333 | 7.182 | 0.559 | 8.917 | 7.262 | 0.501 |
| observations | 12 | 24 | | 9 | 22 | |

Sources of data: [1] Census of India 2001; [2] round 2 of our household survey; [3] our village survey. [4] Probability that two randomly-drawn individuals belong to different groups (commonly known as ethno-linguistic fractionalization index): $f = 1 - \sum_{i=1}^{n} s_i^2$, where $s_i$ refers to the sample share of the ith group. Std errors in parentheses.

# B Propensity score matching

Table 20: Summary statistics of the variables used in the estimation of p-scores (R1 data)

| variable | N | mean | p50 | sd | min | max |
|---|---|---|---|---|---|---|
| member household | 571 | 0.25 | 0.00 | 0.43 | 0.00 | 1.00 |
| scheduled caste household | 573 | 0.09 | 0.00 | 0.29 | 0.00 | 1.00 |
| tribal household | 573 | 0.41 | 0.00 | 0.49 | 0.00 | 1.00 |
| head male | 573 | 0.92 | 1.00 | 0.28 | 0.00 | 1.00 |
| head married | 573 | 0.85 | 1.00 | 0.36 | 0.00 | 1.00 |
| head no schooling | 573 | 0.52 | 1.00 | 0.50 | 0.00 | 1.00 |
| years of education of head | 573 | 2.69 | 0.00 | 3.49 | 0.00 | 15.00 |
| head unemployed | 573 | 0.11 | 0.00 | 0.32 | 0.00 | 1.00 |
| head self-employed | 573 | 0.25 | 0.00 | 0.43 | 0.00 | 1.00 |
| head casual wage | 573 | 0.24 | 0.00 | 0.43 | 0.00 | 1.00 |
| head salaried | 573 | 0.05 | 0.00 | 0.23 | 0.00 | 1.00 |
| IAY benefit | 573 | 0.18 | 0.00 | 0.39 | 0.00 | 1.00 |
| no land owned | 549 | 0.08 | 0.00 | 0.28 | 0.00 | 1.00 |
| land owned | 549 | 1.98 | 1.00 | 3.48 | 0.00 | 39.75 |
| age avg household | 573 | 25.51 | 23.86 | 9.53 | 10.12 | 70.00 |
| rooms in house | 573 | 3.25 | 3.00 | 2.23 | 1.00 | 12.00 |
| bicycles per productive adult | 559 | 0.24 | 0.20 | 0.27 | 0.00 | 2.00 |
| domestic assets | 573 | 0.28 | 0.22 | 1.05 | -1.07 | 9.80 |
| extreme poor consumption | 573 | 0.25 | 0.00 | 0.43 | 0.00 | 1.00 |
| poor consumption | 573 | 0.50 | 0.50 | 0.50 | 0.00 | 1.00 |
| food shortage during $\geq 1$ month last year | 573 | 0.56 | 1.00 | 0.50 | 0.00 | 1.00 |
| nb of loans during last 2 years | 573 | 0.90 | 1.00 | 1.09 | 0.00 | 7.00 |
| total credit during last 2 years | 573 | 2711 | 300.00 | 8124 | 0.00 | 80000 |
| nb of boys of prim. school age enrolled | 573 | 0.39 | 0.00 | 0.65 | 0.00 | 4.00 |
| nb of girls of prim. school age enrolled | 573 | 0.34 | 0.00 | 0.63 | 0.00 | 5.00 |
| nb of boys of mid. school age enrolled | 573 | 0.12 | 0.00 | 0.36 | 0.00 | 2.00 |
| nb of girls of mid. school age enrolled | 573 | 0.09 | 0.00 | 0.33 | 0.00 | 4.00 |
| nb of children 0-5 years | 573 | 1.06 | 1.00 | 1.17 | 0.00 | 8.00 |
| nb of boys of primary school age | 573 | 0.49 | 0.00 | 0.71 | 0.00 | 4.00 |
| nb of girls of primary school age | 573 | 0.48 | 0.00 | 0.72 | 0.00 | 6.00 |
| nb of boys of middle school age | 573 | 0.19 | 0.00 | 0.42 | 0.00 | 2.00 |
| nb of girls of middle school age | 573 | 0.14 | 0.00 | 0.39 | 0.00 | 4.00 |
| nb of children of secondary school age | 573 | 0.21 | 0.00 | 0.44 | 0.00 | 2.00 |
| nb of adults $\geq 18$ | 573 | 3.42 | 3.00 | 1.86 | 1.00 | 11.00 |
| adult participation in Lok Sabha elections | 573 | 56.92 | 66.67 | 40.53 | 0.00 | 100.00 |

Table 21: Probability of joining SHG at baseline: logit household-level regression

| | raw estimates | | marg. effects | |
|---|---|---|---|---|
| scheduled caste household | 1.042*** | (0.386) | 0.157*** | (0.0565) |
| tribal household | 0.348 | (0.274) | 0.0524 | (0.0413) |
| head male | -1.453*** | (0.549) | -0.219*** | (0.0830) |
| head married | 1.657*** | (0.427) | 0.250*** | (0.0656) |
| head no schooling | -0.537* | (0.309) | -0.0810* | (0.0460) |
| schooling of head | -0.206** | (0.0995) | -0.0310** | (0.0147) |
| headhighestgrade2 | 0.0114 | (0.00785) | 0.00172 | (0.00117) |
| head self-employed | 0.202 | (0.324) | 0.0305 | (0.0487) |
| head salaried occupation | 0.813 | (0.576) | 0.123 | (0.0863) |
| head casual wage occupation | 0.668** | (0.316) | 0.101** | (0.0467) |
| head unemployed | -0.427 | (0.606) | -0.0644 | (0.0912) |
| IAY benefit | -0.967*** | (0.360) | -0.146*** | (0.0525) |
| no land owned | 0.145 | (0.462) | 0.0219 | (0.0697) |
| landowned | 0.153** | (0.0711) | 0.0230** | (0.0104) |
| landowned2 | -0.00260 | (0.00200) | -0.000393 | (0.000299) |
| age average in household | -0.0339 | (0.0216) | -0.00512 | (0.00326) |
| nb of rooms in house | -0.455*** | (0.174) | -0.0687*** | (0.0259) |
| nb of rooms squared | 0.0371*** | (0.0144) | 0.00559*** | (0.00213) |
| nb of bicycles per productive adult (15-50) | 0.495 | (0.440) | 0.0746 | (0.0661) |
| domestic assets | 0.117 | (0.128) | 0.0176 | (0.0193) |
| extreme poor consumption (<p25) | -0.970*** | (0.342) | -0.146*** | (0.0508) |
| poor consumption (<p50) | 0.356 | (0.274) | 0.0538 | (0.0412) |
| food shortage during $\geq$ 1 month last year | 0.220 | (0.264) | 0.0333 | (0.0398) |
| nb of loans taken during last 2 years | 0.303** | (0.127) | 0.0458** | (0.0189) |
| volume of loans taken during last 2 years | -0.0000243 | (0.0000219) | -0.00000367 | (0.00000329) |
| nb of boys primary age enrolled | 0.287 | (0.359) | 0.0433 | (0.0541) |
| nb of girls primary age enrolled | 0.238 | (0.325) | 0.0359 | (0.0490) |
| nb of boys middle age enrolled | 2.148*** | (0.829) | 0.324*** | (0.124) |
| nb of girls middle age enrolled | -0.380 | (0.600) | -0.0574 | (0.0905) |
| nb of children 0-5 years | 0.115 | (0.148) | 0.0173 | (0.0222) |
| nb of boys primary age | -0.189 | (0.330) | -0.0285 | (0.0498) |
| nb of girls primary age | -0.0302 | (0.278) | -0.00456 | (0.0420) |
| nb of boys middle age | -2.103*** | (0.769) | -0.317*** | (0.115) |
| nb of girls middle age | 0.470 | (0.494) | 0.0709 | (0.0745) |
| nb of children secondary age | 0.667*** | (0.256) | 0.101*** | (0.0378) |
| nb of adults $\geq$ 18 | -0.192* | (0.109) | -0.0289* | (0.0162) |
| adult participation in Lok Sabha elections | 0.00146 | (0.00288) | 0.000221 | (0.000435) |
| Observations | 537 | | Pseudo $R^2$ | 0.185 |

For the dummy variables, marginal effects are computed for a discrete change from 0 to 1

Robust standard errors in parentheses (* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$)

Figure 5: Round-2 distribution of propensity scores, member vs. control households

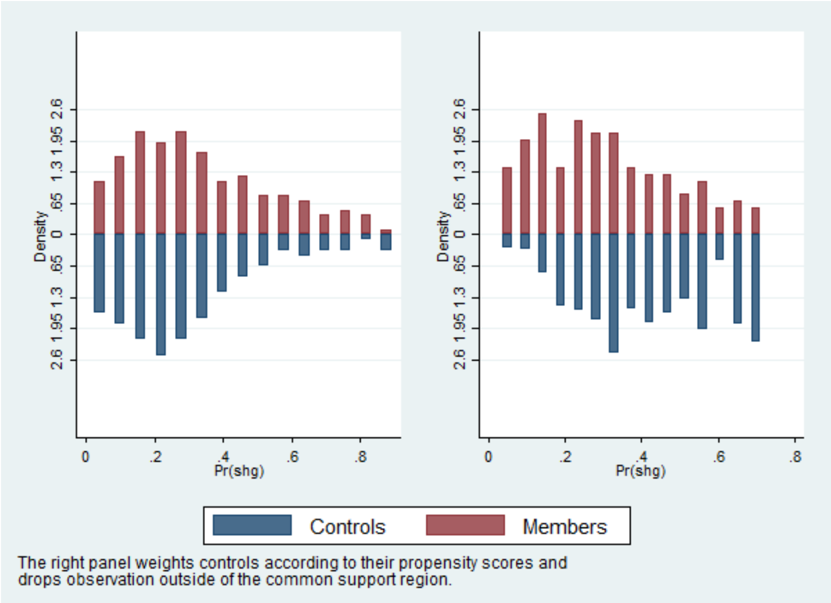



Figure 6: Round-2 distribution of propensity scores, member vs. all other households