

Assessing the Impact of Training on Indian Self Help Groups with Nonexperimental Data

RANJULA BALI SWAIN* and ADEL VARGHESE**

**Department of Economics, Uppsala University, Box 513, Uppsala, Sweden, 75120.*

(e-mail: Ranjula.Bali@nek.uu.se)

***IFMR, Chennai, India & Department of Economics, Texas A & M University, College Station, TX, USA.*

(e-mail: avarghese@tamu.edu)

Abstract

We evaluate the impact of training provided by facilitators of Self Help Groups (SHGs). Indian SHGs are mainly NGO-formed microfinance groups but funded by commercial banks. We correct for membership selection bias with data on current as well as future SHG members. We then account for potential training endogeneity with propensity score matching. Regression and unadjusted matching results indicate that training does not aid in asset accumulation but can reverse the negative impact of credit on income. However, regression adjusted matching which controls for both participation and training selection bias reveals that training impacts assets but not income.

JEL Classification Numbers: G21, I32, O12.

Keywords: India, microfinance, training, impact studies, Self Help Groups.

I. Introduction

Recently, in India, Self Help Groups (SHGs) have emerged as a serious alternative to private microfinance institutions (MFIs). In a recent impact and sustainability study of the SHG Bank Linkage Program, NCAER (2008) claims that SHGs have significantly improved the access to financial services of the rural poor. Similarly, a RBI (2008) report on financial inclusion further emphasizes that SHGs are ‘the most potent initiative since Independence for delivering financial services to the poor in a sustainable manner.’¹ The NCAER study also finds that training improves members’ skills such as communication, marketing, and human development. However, the study does not indicate whether the training translated into better outcomes. In interviews, many members report on the inadequacy of the training received.

This paper aims to explore whether training services have had their intended impact. It tests this objective using a unique data set from five Indian states with SHGs. The data were not only collected on current members and non-members but also on newly enlisted SHG members who have not yet received loans. We examine whether training affects outcomes over and above membership (which measures loan access). We employ two different outcome measures, a long term (assets) and a short term (income).

Why are we interested in the effects of training? As Karlan and Valdivia (2009) note, one would like to know if microfinance institutions should teach skills. Some state that households already have the human capital and only need financial capital. Others claim that MFIs must also provide training, as households cannot effectively use the financial capital that they receive. Furthermore, since MFIs have already organized borrowers in order to

¹ RBI (2008), p.iii.

obtain loans, the cost of providing additional services is small.² A natural tension arises for MFIs on whether they should branch out to training or just lend.

An argument for MFIs to focus on lending is that membership by itself ‘trains’ participants in a number of ways. First, by working, saving, and repaying, members adopt a disciplinary ethic. Second, by actually working on projects, members ‘learn by doing’ without any need of training. Third, regular meetings provide a setting for members to discuss and learn from others about their work-related problems. Our data allows us to discern the effects of training from that of membership. We have data on new members (with no external credit) as well as mature members, thus controlling for member self-selection. We also have training data on the members, where not all mature and new members receiving training.

The paper begins by examining the impact of training on assets and income. We have deliberately chosen short term and long term variables to measure outcome.³ We then proceed to investigate whether the quantity of training matters, in terms of the amount of weeks of training of the borrowers. Third, we explore the type of training, specifically aimed at improving business skills. Finally, we use both matching and regression adjusted methods to adjust for training and participation bias.

The regression results (which correct for membership bias only) reveal that training has no effect on assets, but positively impacts income. We also observe that the amount of training has no impact on these outcome measures. We observe a similar impact when we account for training endogeneity only using matching methods. However, when corrected for both membership selection bias and training bias, we find that training positively impacts

² Here, we are not concerned with an actual extra product as in insurance or health (as in Smith (2002)). Here, the provision of human capital may help.

³ We will explore the impact on other outcome measures such as health and education in future work.

assets but has no impact on income. In other words, not correcting for member households who decide to train would provide a false impression of the impact of training.

In a broad sense, this paper falls under the umbrella of impact studies on microfinance. It contributes to the literature in both the methods and the topic studied. In terms of methods, it corrects for two different types of selection bias by combining two nonexperimental methods: the pipeline and matching methods. Coleman (1999) proposed the pipeline approach, where he compares current members to future members who have not yet received loans. We adapt Coleman's approach to the SHG framework. While Coleman surveyed borrowers in both treatment and control villages, we observe new and mature groups in SHGs in different villages but in the same district.⁴ We then use matching methods to control for training endogeneity. We propose the combination of these methods since for our data and setting, randomization is not feasible. In the absence of randomization, these methods provide an alternative for measuring impact.

Various authors have conducted a number of studies on the impact of SHGs with the NCAER study the most recent. Their report on 4,600 households from six states in India measures impact as the compound annual growth rate of the outcome measure from pre to post SHG participation. They find that income increases 6 %, assets increase 10 %, and participants are more empowered.⁵ Throughout this paper, we will draw on this study as it offers important insights into the functioning of SHGs and provides information on training

⁴ See Karlan and Goldberg (2006) for a review of the major studies and their methodologies. Another influential paper on microfinance impact is Pitt and Khandker (1999) which relies on Grameen's eligibility rule (although see the recent rebuttal by Morduch and Roodman (2009)).

⁵ Until the aforementioned NCAER study, impact studies on SHGs consisted mainly of the Puhazendhi and Badataya study (2002) commissioned by NABARD (India's rural development bank) with 115 members from three states. Their results find that SHG membership significantly increases the asset structure (30 %), savings, and annual net income.

and some other aspects of SHGs not covered in our data. Also, the NCAER study overlaps with our study in five of the six states (we do not have data on Assam).

Clearly, the NCAER analysis does not account for any changes in characteristics or broad economic changes through a control group. Furthermore, their study does not measure the effect of training per se except to note its inadequacy. Still, due to the scarcity of studies on SHGs, these types of studies have had much policy influence, widely quoted in a number of RBI and NABARD (India's agricultural development bank) documents. New insights in the methodology of impact studies can improve upon the previous studies of such an important credit institution.

Measuring the impact of training in general has spawned a large amount of literature, summarized in LaLonde (1995). The techniques have varied from GMM (Dearden *et al.* (2006)) to matching methods (for example, Heckman *et al.* (1997, hereafter HIT)). Training has actually provided the subject matter for a healthy debate on the effectiveness of matching estimators (see for example Smith and Todd (2005) and the rejoinder by Dehejia (2005)). The preponderance of the work has used data from developed countries and in particular, one training program from the US. When accounting properly for selection bias, the studies find a positive impact of training. In contrast to this paper on SHGs, these studies do not need to account for participation choice in a particular group before training choice.

In terms of measuring the training impact specifically on MFIs, Karlan and Valdivia (2009) provide the only rigorous study. Using the popular randomization method with data from Peru, they find that business training improved business practices and revenues. They also find that this increased knowledge led to greater repayments and client retention. They do not separately measure the effects of membership, so their results hold conditional on membership. In sum, for SHGs, microfinance and in general, for developing countries, not many studies have analyzed the contribution of training to outcome measures. Still, millions

of rupees are spent on training, where sixty percent of SHGs rely on outside support for training. An analysis of training can reveal whether this expenditure is worthwhile.

For those unfamiliar with SHGs, in the next section, we outline the basic information, design, and training. Section three discusses the methodology and explains potential biases. In the fourth section, we describe our data set with the results presented in Section Five. The last section concludes and draws policy lessons.

II. Self Help Groups and Training

Self Help Groups fall under the category of village banking which includes ten to twenty (primarily female) members. In the initial months the group members save and lend amongst themselves and thus build group discipline. Once the group demonstrates stability and financial discipline for six months, it receives loans of up to four times the amount it has saved.⁶ The bank then disburses the loan and the group decides how to manage the loan. As savings increase through the group's life, the group accesses a greater amount of loans.

Initiated in 1992, the SHG program faced slow progress up to 1999. Since then, it has mushroomed, growing to financing 687,000 SHGs in 2006-07 alone compared to 198,000 SHGs in 2001-02. The cumulative number of SHGs has grown to roughly three million by March 2007 reaching out to more than forty million families. According to NABARD (2006), 44,000 branches of 547 banks and 4,896 NGOs participate in the SHG bank linkage program. As with microfinance in India (or more generally with credit), the spread of SHGs has been concentrated in the Southern states.

As of March 2002, the cumulative number of linked SHGs in five states covered in this study follow a similar pattern. For these five states, their shares (in parentheses) of the

⁶ More recently, NABARD has allowed the banks to lend five to six times the savings amount to reflect the growing requirement of SHG members.

cumulative SHG links are the following: Andhra Pradesh (48.5), Tamil Nadu (12.5), Uttar Pradesh (6.6), Orissa (4.1), and Maharashtra (3.9). Well aware of the concentrated spread of SHGs, NABARD recently has identified thirteen poorer states in which they would like to expand their program. As the program is predominantly rural, NABARD has recently identified urban areas for experimentation. Both of these initiatives demonstrate NABARD's confidence in the program and call for a careful examination of the program's impact.

The SHG program links with the poor through Self Help Group Promoting Institutions (SHPIs), which primarily includes NGOs, but also banks, and government officials. The agencies survey the village, provide the details of the program, enlist borrowers, and sometimes organize the training. Three types of linkages have emerged as the most common. In Linkage Model 1, banks both form and finance SHGs. According to NABARD (2006), roughly twenty percent of SHGs fall under this linkage model. In the most popular linkage model 2 (roughly three-fourths of all SHGs), NGOs and others form the groups but banks directly finance them. In the third linkage model banks finance the SHGs through NGOs (but only 5 % of linkages fall under this model).⁷ For banks, SHG linkages allow them to expand lending which helps them to fulfill India's priority sector guidelines. For NGOs, the SHGs provide a vehicle to reach a larger number of the poor for a wider development agenda, such as women's development schemes and child development services.

The program features of small loan size, frequent meetings, frequent repayment installments, and savings dissuade the non-poor from joining SHGs. Thus, SHGs do not use explicit eligibility criteria but rather rely on indirect methods for attracting the poor. In many SHGs, SHPIs provide training and outreach to members in fields such as primary healthcare,

⁷ In our data, 70 % of SHGs follow Model 2 while 12 % and 18 %, respectively, follow the first and third models.

basic literacy, family planning, marketing, and occupational skills. The NCAER study finds that even after four years after formation, 61 % of SHGs are still dependent on these SHPIs.

Training and capacity formation can be broadly classified into two categories.⁸ General training to all SHG members covers group formation and an introduction to linkage methods, which includes basic literacy, bookkeeping, group formation and dynamics.⁹ The general training usually takes one day and each participant is awarded Rs. 250 (equivalent to a week of agricultural wages, a generous compensation). Since all participants receive this relatively homogeneous training, we do not include this aspect in our training measure.

The additional training module (the focus of our paper) relates to other types of training. These include skill formation training which aims at improving income-generating activities such as farming, craft or business. Skill formation programs include the REDP (Rural Entrepreneurial Development Program), designed for unemployed but educated rural youth.¹⁰ The training hours depend on the skill and training module but vary between ten and a half hours to seventeen and a half hours per day. Participants receive a stipend as well as free boarding and lodging. The REDP comprises of three distinct phases: pre-training (survey for identifying business activity), training (for six to eight weeks) and post training (with services rendered to trainees for at least two years for setting up units).

SHG members can demand the required skill training. However, their demand is not met in many cases because the viability of the training sessions require a critical number of potential trainees to make the 'demanded' training program cost effective. Moreover, SHPIs

⁸ Public information on the SHG training program is unavailable, The discussion below is our attempt to distill the information in a concise manner drawing from different sources. Much of this information was provided through visits with NABARD's regional office in Bhubaneswar, Orissa. Some of this information is also available through NABARD circulars.

⁹ More specifically it includes training on group formation and functioning; functions and qualification of office bearers; rules and regulations; planning, management and monitoring; financial service provisions, conditions and procedures; training of group leaders; and training of book keepers.

¹⁰ The MEDP (Microenterprise Development Program) began only after 2006, which is after our data had been collected. However, it will be discussed more at length in the conclusion.

need to find local trainers for that specific skill. Since some of the demand is internally driven, members participate out of interest and need. Actually, many members other than those that initially request the training participate in the sessions. Furthermore, as mentioned before, NABARD's stipend provides an added incentive to participate.

Other than the REDP, some SHPIs also provide additional education, health, and awareness creation training.. In contrast to the base training, these additional modules of training are more haphazard and not homogenous or focused. Thus, the training covered in this paper is 'as delivered' and not optimal in any sense. This notion of training contrasts with the Karlan-Valdivia study where meetings started with training, and with penalties such as fees for tardiness and threat of expulsion.¹¹

Different camps have touted the relative advantages of SHGs over private MFIs in training delivery. Defenders of NGO linked SHGs (linkage model 2) assert the following. First, that NGOs perform effective development activities within their own district and so are the best equipped to provide training services. They do not need an extra incentive mechanism to monitor and train SHGs (as suggested by the detractors). If NGOs choose to deny training services to a particular group, then the NGOs have identified that group as low quality. At times, members initiate the training which differs from many standard Grameen style models. Furthermore, to a certain extent, the group self-governs as the elected office bearers (such as the president and treasurer) maintain the financial records. These officers are usually more educated and share the records with the group.

In contrast to SHGs, several MFIs are donor-driven and face pressure to obtain high repayment rates which may stifle their training efforts. In particular, training may have payoffs later but add to costs now and may damage current outcome measures. Since the

¹¹ Even with these conditions, Karlan-Valdivia found many detractors who chose not to attend the training sessions.

government supports the SHG program with a development mission in mind, it may not face the same pressure. Overall, the SHG model reflects an institutional approach, while the private MFIs reflect a more market oriented outlook. We turn now to examine whether training impacts borrowers.

III. Estimation Strategy

In assessing impact, researchers seek to disentangle the causal effect from the potential selection bias. In particular, the decision to participate in SHGs and training depends on the same attributes that determine the impact variable (asset accumulation and income in this paper). We encounter a double selection problem: selection into participation and into training. We will first establish the correction for selection into the program and later discuss the treatment of training endogeneity. With nonexperimental data, we correct for the first selection bias using the pipeline method described below and account for the second using matching methods. In this section, we limit our remarks on impact assessment to those pertinent to this paper.¹²

In measuring the training impact of a well-established development program such as SHGs, the tools at our disposal are few. The increasingly popular method of randomization is complicated to implement for a national program such as the SHG program.¹³ Here we list some potential difficulties. Randomization entails disbursing new loans to some borrowers and denying loans to a new, additional set of borrowers. However, in cases such as SHGs the interest lies in finding how loans have affected the current borrowers who have borrowed for some time.

¹² Selection bias in impact studies on microfinance has been discussed at length in Goldberg (2005), Karlan and Goldberg (2006), and Coleman (1999). More generally, see the recent book by Angrist and Pischke (2009).

¹³ Duflo *et al.* (2007) explain the great benefits of randomization but they themselves admit, 'Randomized evaluations conducted in collaboration with governments are still relatively rare. They require cooperation at high political levels, and it is often difficult to generate the consensus required for successful implementation'.

A related comment is that randomization usually studies short term impact measures while here we are also interested in longer impact. Providing training to some borrowers and denying others will upset certain constituencies. Furthermore, one cannot hold a control group without training for long (as noted by Karlan and Goldberg). For large programs such as SHGs one would also need to synchronize the training randomization across different states. Instead of choosing not to perform evaluation of any kind, one can explore other alternatives with nonexperimental data.

A second strategy (as adopted by Pitt and Khandker) exploits an exclusion rule on credit access to estimate unbiased impact. However, SHGs follow no such exogenous rule. The third method, the Coleman pipeline approach, is most appropriate for our data. As mentioned in the introduction, at the data design stage Coleman interviewed both current and future members of a village bank. Presumably, both types of members have similar unobservables, thus correcting for member self selection. We can attribute any difference in impact to borrowing from the MFI. Adapting this pipeline method is not straightforward in that the SHG program is well established and not new as the program Coleman studied. Two differences with Coleman's study is that our treatment and control villages come from the same district and that the SHG program is the largest microfinance program in India in contrast to Coleman's study of a small, new village bank.

By design, SHG members have to wait to receive a loan from the bank (about six months) and we can exploit this design feature to identify the self-selected members who have not yet received a loan. Our method differs from Coleman in that he intentionally interviewed people at the data design stage while we exploit the wait some households have to face with later SHGs. Ideally, for perfect impact assessment, one would choose a control group from the same village (which would hold all external conditions constant) but then earlier signees of SHGs may have different reasons for joining than later signees.

To avoid the problems with villages, we raise the level of aggregation to another level, as in districts (where both mature and new SHGs reside). NABARD's choice to expand the SHG program occurs at the district phase without any specific announced policy targeting certain villages over others.¹⁴ Thus, we choose to aggregate at the district level, the basic administrative unit within a state where credit decisions are made.

We have data from districts where some members are currently active members of SHGs for at least one year but in the same district (but different villages), members from newly formed SHGs have been selected but not yet received financial services from the bank. Thus, to correct for selection into the program, the treatment group in our sample consists of mature SHG members, while new SHG members form our control group.¹⁵ We hypothesize that the mature and new SHG members have similar unobservables. We also have information on nonmembers from these districts so that we can condition on the selection to join the SHG.¹⁶

Program placement bias arises from non-random placement of programs, both for the SHG itself and the training programs. For example, the cost of arranging the training and availability of trainer will vary depending on the economic development and infrastructure of the region. For both SHG location and training, these differences across districts or regions due to non-random program placement may induce a bias in the impact results (i.e. members are not better (worse) off due to the program but simply because they live in a better (worse)

¹⁴ NABARD's or the bank's decision to link with a SHG might follow the NGO's choice. We do not have information whether NGOs favor certain villages over others within a certain district.

¹⁵ One caveat of this approach is that we need to assume the behavior of new SHG members has not changed while anticipating loans. In other words, while awaiting loans, SHG members do not begin asset accumulating knowing that they will receive SHG loans in the future. An advantage of SHGs is the following. Due to the slow incubation period of SHGs, members know for some time the nature of wait and will not change their behavior as radically as a one time boost in loans.

¹⁶ The dropout rate for SHGs is not severe in that the NCAER study estimated the dropout rate as 8.2 %, below the 20-30 % cited by Aghion and Morduch (2005) and Karlan as a severe problem. This dropout rate was calculated for SHGs with an average age of 5.4 years, nearly double our average age, so we conservatively estimate that the dropout rate in our data is below 5 %.

area). As described above, we hold these differences constant by drawing the treatment and control group from the same area, i.e. the same district.

We still need to account for nonmembers from these districts who may avail themselves of district specific policies, such as parallel government programs. We control for these differences with the use of district fixed effects. In that district-wide effects may spillover from mature members to new members and non-members, the estimates here would underestimate that impact. To account for the remaining village level variability, we employ village level characteristics.¹⁷

As mentioned in the earlier section, the SHPIs provide basic training to all SHGs. Then, the SHPIs organizes additional training for some of the SHGs. the training variable (T_{ijs}) indicates whether the household received such training. Thus, this variable captures whether training has impact beyond membership duration and self selection of the members.

Keeping in mind the outlined procedure, we estimate the following regression:

$$I_{ijs} = a + \alpha X_{ijs} + \beta V_{js} + \lambda D_s + \gamma M_{ijs} + \delta SGHMON_{ijs} + \phi T_{ijs} + \eta_{ijs} \quad (1)$$

Where I_{ijs} is the impact for household is measured in terms of asset accumulation or income generation, for household i in village j and district s , X_{ijs} are the household characteristics; V_{js} is a vector of village-level characteristics, and D_s is a vector of district dummies that control for any district level difference. Here, M_{ijs} is the membership dummy variable, which controls for the selection bias. It takes the value one for both mature and new SHGs. It takes the value of zero for those villagers that have chosen not to access the program. Here, $SGHMON_{ijs}$ is the number of months that SHG credit was available to mature members, which is exogenous

¹⁷ For this data set, we prefer this approach over village fixed effects. Here, with 218 villages and the available sample size, a regression with 218 dummies is simply infeasible. With aggregation at the district level, any differential impact of the program due to unobservables at the village level (e.g. village has a more dynamic leader or village has stronger political connections) cannot be taken into account.

to the households. The parameter of interest is ϕ which measures the impact of training. We now discuss how we address the selection bias of the trainees.

As described in the section on SHGs, for actual training delivery, it must pass a three step process. Only in the first step, the household takes part by requesting training. Actually, some households who do not demand training attend the training sessions. The other two steps of finding a trainer and hoping for a critical mass of trainees does not lie within the household's choice. The strength of endogeneity here contrasts sharply with the vast literature on US public training programs where trainees play more of a role in the choice of participation. Still, even here, training is endogenous and we need to take that into account.

One possibility is to use an instrument that would impact training directly and not impact outcomes. However, we cannot find such an instrument. For example, we have recall data on income in 2000 but that would certainly correlate with income in 2003. Furthermore, any village level shock would affect both the choice to train as well as the outcome measures. Without an instrument, we cannot test for endogeneity either, which would provide an early resolution of the strength of the endogeneity of the training variable.

An alternative is propensity score matching, which has received considerable attention lately.¹⁸ In contrast to regression methods, propensity score does not depend on linearity and has a weaker assumption on the error term. We first examine the viability of using propensity score matching for this data set. Heckman *et al.* (1997) have outlined three conditions for the least bias in this method. One, the survey questionnaire should be the same for participants and non-participants so that the outcome measures are measured the same for both. Two, both should come from the same local labor markets. Three, the availability of a rich set of

¹⁸ See the excellent survey by Caliendo and Kopeinig (2008) about the main issues on propensity score matching. Angrist and Hahn (2004) find that with propensity scores, there is a gain in precision in finite samples.

observables for both outcome and participation variables. Our data set satisfies all three conditions, the first and third as described in the data section and for the second, both treatment and control households reside in the same districts.

For the households in this data set, the estimators match the households who received training to those who did not. Except for the treatment, the households are very similar. Thus, any differential can be attributed to the impact. Households with low or high probabilities cannot be matched and generally, these are dropped. In matching terminology, we keep the households on the common support. The probability ($P(X)$) of being selected is first determined by a logit equation and then this probability (the propensity score) is used to match the households. Denote Y_1 as the outcome variable of interest for those with training ($T=1$), and denote Y_0 as the outcome variable of interest for those without training ($T=0$), then equation (2) denotes the mean impact of training:

$$\Delta = E[Y_1 | T = 1, P(X)] - E[Y_0 | T = 0, P(X)] \quad (2)$$

where the matched comparison group provides the data to calculate the second term, and the propensity score weights the whole expression for all households on common support.

For our purposes, we employ a version of matching which combines elements of regression. These regression adjusted matching estimators as in Barnow *et al.* (1980) allow for different covariates for the logit participation equation and the outcome equation. In our case these estimates are particularly important because of the need to account for the selection of participation into the program in which we use the pipeline method.¹⁹ The following procedure explains the steps for regression adjusted matching estimators. First, run a

¹⁹ We are aware that this specific type of selection is actually a sequential or dynamic selection process. In other words, the subsequent choice of training depends upon the choice of participation. Since as Caliendo and Kopeinig state: ‘practical experiences with sequential matching estimators are rather limited’ we estimate the static framework but using matching for the training selection problem.

regression for the outcome equation on the no training group $Y_0 = x\beta + \varepsilon$. We then and calculate the fitted values.²⁰ Second, subtract these values from the outcome variables for both the no training and training group (since these fitted values are free of the effect of training). Third, match the new variables, outcome variables minus the fitted values. The estimator is given by equation (3):

$$\Delta RAM = \sum_{j=1}^T w_j \left[\left(Y_{j1} - x_j \hat{\beta}_0 \right) - \sum_{i=1}^C W_{ij} \left(Y_{ij0} - x_i \hat{\beta}_0 \right) \right] \quad (3)$$

where RAM refers to regression adjusted matching estimators, T (C) refers to the total number of treated (not treated), and w refers to the particular weight used in matching.

We also have to choose which matching algorithms to use. Since the probability of two households being matched exactly is close to zero, distance measures are used to match households. Following Smith and Todd, we first choose the neighbor to neighbor (NN) algorithm. This algorithm is the most straightforward and matches partners according to their propensity score. We employ both one and ten person matching, where the latter uses more information to match the partners. The NN algorithm is only used for simple (or unadjusted) matching since its performance is not well known in regression adjusted matching.

For regression adjusted matching, we turn to the local linear regression (LLR) method (for bandwidths 1 and 4). The theorems in HIT which justify regression adjusted matching are based on LLR, a generalized version of kernel matching which allows faster convergence at the boundary points. The LLR method uses the weighted average of nearly all individuals in the control group to construct the counterfactual outcome. For regression adjusted matching, the analytical standard errors are tedious to compute. Bootstrapped standard errors for the

²⁰ HIT suggest a semi-parametric procedure which exploits a richer functional form. We attempted to fit this from our data with two candidates, age and SHGMON. We failed to reject the null hypothesis of linearity: P=0.664 and P=0.552 respectively for age and SHGMON.

LLR procedures are used since these are not subject to the general criticism of the use of bootstrap standard errors in matching models (see Abadie-Imbens (2007) and HIT).

IV. Data

The data used for the empirical analysis in this paper were collected by one of the authors and forms part of a larger study that investigates the SHG-bank linkage program.²¹ The household survey uses a quasi-experimental design, with pre-coded questionnaire to collect cross-sectional data for two representative districts each, from five states in India, for the year 2003.²² Within the states, the study avoided districts with over and under exposure of SHGs and only evaluated SHGs with good operational links. Furthermore, the years of SHG membership was restricted to three years or less. The reason was the following. First, the SHG movement as mentioned in the SHG description section only took off in 2000 and was not as well structured before 2000. Second, selecting SHGs of three years or less minimizes attrition and dropouts. The sampling strategy randomly chose the respondents from the SHG members at the district level. The non-members were chosen to reflect a comparable socio-economic group as the SHG respondents.

For this particular study, the collected data was further refined. Of the total respondents, 114 were from villages with no SHGs. Since these households were not provided the opportunity to self-select, these were dropped. Sixty old and new SHG respondents were from the same village and this would contaminate the sample since the earlier signees may be

²¹ The process involved discussion with statisticians, economists and practitioners at the stage of sampling design, preparing pre-coded questionnaires, translation and pilot testing with at least 20 households in each of the 5 states (100 households in total). The questionnaires were then revised, reprinted and the data collected by local surveyors that were trained and supervised by the supervisors. The standard checks were applied both on the field and during the data punching process.

²² These states (districts in parentheses) are Orissa (Koraput and Rayagada), Andhra Pradesh (Medak and Warangal), Tamil Nadu (Dharamapuri and Villupuram), Uttar Pradesh (Allahabad and Rae Bareli), and Maharashtra (Gadchiroli and Chandrapur).

of a different makeup than the later signees. Of the remaining sample, 593 respondents are from mature SHGs, 185 are from new SHGs, and 51 are non-members.

For SHGMON, or the number of months since a member has joined a SHG, we made the following adaptations. Since an SHG links with banks only six months after formation, we needed to account for those six months of no credit access. Almost all the new SHG respondents in our data had been members for less than six months and for these SHGMON=0. Only fourteen of these new respondents were members for more than six months, in which case SHGMON = date of formation - six months.²³ For the mature SHGs, their SHGMON = date of formation - six months. Some mature SHG respondents (forty six) did not report the date of their SHG formation. For these households, we used the number of the months since they received the first SHG loan for SHGMON.

The data were not collected specifically for a training study. An advantage is that respondents did not overemphasize training and answered training questions subsumed in their other questions. A disadvantage is that some specific answers to training are missing. We primarily have information on the total training weeks that a household has received. We set the training variable to 1 for all households who reported positive weeks of training. Since both mature members and new members received training, we can differentiate the impact of training from that of loan access.

The survey yields other measures of training. When comparing the means and variances of the training weeks for mature and new SHGs we find a significant difference: the amount of training weeks (1.52 versus 1.15) and variability in training is larger for mature

²³ These respondents have been a SHG member for less than a year and have been identified as a new member in the SHG list available at the district level.

SHGs (2.42 versus 1.87).²⁴ About half (48 per cent) of the mature SHGs received training while 39 per cent of the new SHGs reported the same.²⁵ These statistics are not surprising in that the longer length of membership of mature SHGs will provide them with more opportunities for training. Surprisingly, a sizeable percentage of new SHGs are receiving training indicating a new commitment by policymakers. We also account for a very small percentage of the non-members who also received training from other NGOs or programs.

Table 1 compares the characteristics of households who received training to those that did not (even including nonmembers). In general, those who received training were wealthier, had higher income, were older and lived in villages closer to paved roads and further from the market. These variables indicate that either more prosperous households receive training (who probably are not the target group of SHPIs) or training actually made households more prosperous. A t-test for the same variables before they received training and became members yield a similar difference: -4.34 for income and -2.26 for assets.²⁶ Still, we need to condition on the full set of covariates and control for member self-selection to study the full impact of training.

<Insert Table 1 here >

As suggested by Doss *et al.* (2007), we accumulate assets from six categories: land owned, livestock wealth, dwelling and ponds, productive assets, physical assets, and financial assets (includes savings and lending). Household income includes income from agriculture, poultry and livestock, wages, fisheries and forest resources, rent, remittances, and enterprise.

²⁴ A t-test with unequal variances revealed a t-ratio of 3.32 statistically significant at the 1 % level.

²⁵ NCAER(2008) also finds that nearly half of all the SHGs have had skill development training. About 35 per cent of the households received training only once in 2006 and another 15 per cent have received training multiple times.

²⁶ The data from the year 2000 is recall data and thus may have some measurement errors. Thus, we chose not to use it to create difference in difference estimators.

Household characteristics include age, gender, education dummies and number of earning members in the family. We also include dependency ratios in that we expect households with larger dependency ratios may have greater (lesser) incentive for asset accumulation (income generation). In order to control for initial wealth, we employ land owned three years ago.²⁷ For village characteristics, in addition to male wage, we include the following distance variables: paved road, market, primary health care center, and bus-stop. Table 2 presents the non-training related descriptive statistics of the independent and the dependent variables, respectively.

<Insert Table 2 here>

SHG members and non-members are of about the same age, dependency ratio, similar level of education and a higher amount of land on average. In terms of village level variables, members are closer to most amenities but not surprisingly non-members are relatively closer to banks. On other variables, we further find that members on average have a relatively higher income, own more land and dwelling but have slightly lower amount of assets. Combining these statistics with those of Table 1 indicate that SHG members who had training have higher income and possibly higher amount of assets. This selection bias will affect the results.

V. Results

This section presents and discusses the estimation results for the training impact of the SHG bank linkage program on asset accumulation and income. We first examine the results through regression methods since these serve as points of departure. Furthermore, these can be fully interpreted, along with the impacts of the covariates and interactions. We then compare these results to those obtained through matching methods. Table 3 provides the regression

²⁷ Since land forms the bulk of assets and land turnover is infrequent in India (see Pitt-Khandker (1998), for more discussion on this observation), this variable was the best choice for initial wealth.

results of Equation (1) for the impact of training on assets and income.²⁸ In Columns (1) and (2), we examine the impact of the binary training variable. Columns (3) and (4) focus on the amount of training received by SHG members and its impact on the gross assets and income.

<Insert Table 3 here>

Column (1) indicates that membership duration matters for asset accumulation and training does not provide any additional benefit.²⁹ Presumably, the in-built savings mechanism within the SHG framework provides sufficient incentives for members to build assets with training providing no additional advantage.³⁰ The results in column (2) indicate that surprisingly membership duration negatively affects income creation (though the point estimates are small). Thus, a trade-off emerges for SHG members to choose between short term income creation and long term asset creation.

Moreover, training can reverse the negative impact by positively increasing the income generation of the household. Roughly, up to three years of membership can be reversed with training (assuming constant returns to membership).³¹ The amount of training, however, does not have a direct impact on either income generation or asset creation as seen in columns (3) and (4). We also estimated a non-linear function of weeks of training with the assumption that initial weeks would have a larger effect but did not find any significance on these either. The results indicate that training may be more effective with a focused delivery, i.e. higher quality and different types of training. An analysis of quantity remains important

²⁸ The censoring of the income variable was roughly 7 %. We estimated Tobit regressions both with and without robust standard errors which yielded similar results.

²⁹ In related work, we examine the impact of SHG participation. We find that not including the membership dummy would actually erroneously indicate no impact of length of SHG membership leading one to conclude that length of time has no impact on assets.

³⁰ We have also estimated regressions with alternative asset specifications: gross assets minus SHG savings and also with net assets (gross assets – other borrowings). The results were substantially the same.

³¹ We also estimated an alternative specification testing whether training had greater impact on older SHG members. We failed to find any significance.

because the number of weeks translates into increased costs on the SHPIs as well as participants.

Of the household characteristics, we find that the dependency ratio positively impacts asset accumulation but negatively affects income generation. Households with a greater number of dependents have greater interest in creating assets than income. Education carries the expected signs in that households with greater education are more adept at asset creation (since no education is the dropped dummy) and less interested in current income generation. Initial wealth (as in the amount of land holdings) also clearly influences the current asset and income position of a household. Of the village characteristics, distance from paved road (as expected negatively) and distance from the bus stop are significant (though somewhat surprisingly positively) for assets but no influence on income.

Table 4 provides the impact of the level of infrastructure and business training with their respective interactions. Table 4, Columns (1) and (2), find that training has a much higher impact on assets when made available to SHGs in villages closest to paved roads. For effective asset accumulation, location of village matters and households need an avenue for marketable products. Without that outlet, households have no incentive to accumulate assets and training in these respects would not be of much use. For those with training, one kilometer less of paved road can drop assets by about 5000 rupees. With income generation, paved roads do not yield a similar impact, presumably because households can still consume their own products without relying as much on the market. Still, the result of no effect is surprising. In columns (3) and (4), we tested whether specific business related training

positively impacted outcome variables.³² We found no impact on business training per se but it becomes significant when impacted with months of membership. We find that for both assets and income, business training has similar effects. Initially, this type of training negatively impacts outcome but reverses itself after some months of membership. This may indicate that either a sufficient amount of credit is needed for the training to have effect (since loans increase with months of membership) or that the skills take some time to be absorbed before they have their desired impact. In general, the results of Table 4 indicate that a more nuanced approach to assets reveals that training impacts assets.

<Insert Table 4 here>

We now test the robustness of our results to the endogeneity of the training variable. As discussed in the methodology section, matching methods take into account the selection bias from training. A parsimonious logit equation determines the probability of participating in training.³³ Covariate candidates are variables that influence both the participation and the outcome variable and ones that are not affected by participation and its anticipation. The variables were chosen through a statistical significance and ‘hit or miss’ method and at the same time keeping the balancing in mind (see Caliendo and Kopeinig). We chose to include: age, age squared, gender, education dummies, shock in 2000, distance from bank, health care center, marketplace, and paved road, linkage model 2 and interaction of age and model 2. We will discuss the sensitivity of the results to the variables chosen later.

³² This variable is not precisely business training because households were asked ‘What other services, inputs, and facilities are available to members of the group?’ If households answered marketing information or skill training, this variable was set to 1.

³³ The issue of a simple versus a quasi-saturated logit model is a contentious one. As noted by many, though, the purpose of the logit equation here is not only to predict training participation (as in selection models) but also for covariate balancing.

We then matched the treated and comparison group based on the propensity score. The results in Table 5 confirm the regression results in terms of significance. We find no impact of training on assets but we find impact on income using both the local linear regression and neighbor to neighbor techniques. These estimates, however, hold conditional on participation and do not take into account member selection bias. The regression adjusted matching results in columns (3) and (4) take into account member selection bias. The impact effects reverse with the training impact on assets significant while the impact on income as not significant.

In Table 6, we compare the different estimates. In column (1), the unadjusted t-stat difference suggests that training impacts both assets and income strongly. With regression covariates, the impact of training on income falls by one half and the significance on assets disappears. The matching estimates suggest a greater impact of training but does not correct for participation bias. Finally, the regression adjusted estimates indicate a greater and significant impact of training on assets (on the order of 16 %) while the impact on income disappears. These last set of results contrast to previous studies who have found not much difference with unadjusted matching (for example, see HIT and Smith and Todd). Our results differ from previous work because of the different data set used (much of the previous studies use the same US training data) and the additional selection issue of participation.

We can now compare the impact of membership to that of training. We have run a large number of regression estimates in measuring the impact of membership on assets and consistently find point estimates similar to those found in Table 3. These results indicate that membership (evaluated at SHGMON means for mature members) provide a return of 15 % on assets. From the regression adjusted estimates above, training can double these returns. These estimates provide a partial resolution at least in this context to the question posed in the introduction of whether MFIs should only focus on lending. They should not. The regression estimates on income here and elsewhere suggest that membership has a zero or negative

impact on income. The regression adjusted estimates of income also indicate that training has no impact either.

Unlike previously researched US training programs, we do not have the luxury of comparing our estimates to experimental data. Theoretically, since regression adjusted estimates are the ones that correct for both types of selection bias, they would provide the most reliable estimates. Below we perform sensitivity analysis of various types to explore the robustness of the regression adjusted matching results. We also examine sensitivity to the inclusion of unobservables.

To begin with, we explore the sensitivity of the results in terms of the pscore specification and the matching algorithms. All of these results are available from the authors. In terms of the pscore specification, in general, when the logit equation is even more parsimonious than the one specified even excluding village level characteristics, the impact disappears. This result arises from the simplicity of the propensity score which does not correctly provide a proper match. When we add a large number of variables to the logit equation, we encounter balancing problems. Thus, our chosen equation balances the two but is robust to the addition and subtraction of a few variables. We have also used kernel algorithms for the matching and regression adjusted matching with different bandwidths yielding similar results. Finally, since the bootstrapped standard errors are not analytical, we ran the matching results a number of times for a check of the robustness of the bootstrapped standard errors.

Since for the conditional independence assumption, selection relies on observables, we also tested the sensitivity of our results to the inclusion of unobservables (Ichino *et al.* (2007)). Results available from the authors revealed not much difference using either a variable that mimics the distribution of the unobservables (if NGOs organized training) or

imposing a particular distribution on the unobservables. We also tested the impact of business training through matching methods (both unadjusted and adjusted). Similar to the regression results without interactions, we found no impact on assets or income.

In sum, as mentioned previously, regression is much more flexible than matching, allowing interactions and providing interpretation of covariates. Angrist and Pischke have noted that saturated regression models perform as well as simple matching models and indeed our regression estimates (though not a saturated model) resemble those of simple matching. However, regression adjusted matching results reveal that properly adjusting for both member and training selection bias will offer starkly different results on the impact of training. These results are robust to departures from our specification.

VI. Conclusion

In this paper, we evaluated the impact of training in Self Help Groups on two outcome measures, income and assets. Using regression adjusted matching methods, we find that training actually impacts assets and not income. These results are consonant with parallel work where we find that membership positively impacts asset creation and not income. We expected these impact results with membership but not necessarily with training. The impact of training on assets reveal that training strengthens members' skills in savings and asset accumulation. The lack of impact on income indicates that much more needs to be established for income generation. For example, marketable goods, infrastructure, and other factors play a part and that paradoxically, the effects on income generation may take more time than asset accumulation. Finally, the estimates on business training reveal that with longer membership, business training positively impacts both assets and income.

We now comment on future directions, both in terms of research and policy. In broad terms, the methodology here combines two nonexperimental methods. It provides a potential

alternative to randomization when for different reasons, experimental methods are unavailable. In the future, because of reasons mentioned in the introduction, randomization seems unlikely. Thus, we have to make use of nonexperimental data being as careful as the data allows with selection bias issues.

In terms of the survey, even though it has the best data to date on training on SHGs, more work needs to be done for data collection. One, our measure of quantity of training is provided in weeks, if one were to obtain a finer measure such as hours that may provide different results. Two, a better distinction of the types of training programs would help differentiate the ones that had most impact. Three, in future work we will examine the relationship between softer skills of training such as education and its impact on other outcome measures such as schooling. Though this type of training may incur costs now, it has future payoffs.

In terms of implementation, according to NCAER, more than eighty per cent of the SHGs face problems in developing the skills of their members. Major reasons cited are: lack of time, lack of interest, inadequate literacy among members and insufficient training facilities. This revelation is in line with our data, where the poorer borrowers who need the training more do not receive it. The SHGs in all the states suggested that the SHPIs allow more time in training and group discussions. They further require support from financial institutions in training on book keeping, reviewing and advice on SHG financial activities and health. Furthermore, the training program is not homogenized and varies by NGOs so it is difficult to grade the quality of the training program.

Two recent implementations offer future improvements for training programs. In a recent microenterprise study (Nussbaum *et al.* (2005)), trainers employed by SHGs were asked for feedback on how the training program could be improved. These same trainers were then asked to conduct training programs based on their insights. Recipients perceived these

programs as much higher quality. Another program initiated by the SHG program itself is Microenterprise Development Program (MEDP) which began in 2006 and thus does not impact this data set. This training program targets skill development for mature SHGs. Here, the initial demand for skill training comes from the SHGs and the SHPAs apply for grants to impart the relevant skill training. Another appealing aspect of this program is that the length of the training is limited to two weeks and can also be a minimum of three days. Future data collection on this program can evaluate its impact.

One of the limits of this study is that even if we have evaluated the benefits through the impact, we have not estimated the costs. Can another training mechanism deliver similar impacts at lower costs? A future study on SHGs can hopefully answer some of these question with a focus on more states, especially the newer ones in which NABARD forecasts SHGs to develop.

References

- Abadie, A. and Imbens, G. (2007). 'Bias Corrected Matching Estimators for Average Treatment Effects', Working Paper, Harvard University.
- Aghion, B. and Morduch, J. (2005). *The Economics of Microfinance*, MIT Press, Cambridge, Mass.
- Angrist, J. and Hahn, J. (2004). 'When to Control for Covariates ? Panel Asymptotics for Estimates of Treatment Effect', *Review of Economics and Statistics*, Vol. 86, pp. 58-72.
- Angrist, J. and Pischke, J. (2009). *Mostly Harmless Econometrics*, Princeton University Press, Princeton, NJ.
- Barnow, B., Cain, G. and A. Goldberger (1980). 'Issues in the Analysis of Selectivity Bias' in E. Stromsdorfer and G. Farkas (eds), *Evaluation Studies Review Annual*, Vol.5, Sage, San Francisco.
- Caliendo, M. and Kopeinig, S. (2008). 'Some Practical Guidance for the Implementation of Propensity Score Matching', *Journal of Economic Surveys*, Vol.22, pp. 31-72.
- Coleman, B. (1999). 'The Impact of Lending in Northeastern Thailand', *Journal of Development Economics*, Vol. 60, pp. 105-141.
- Dearden, L., Reed, H., and J. Van Reenen (2006). 'The Impact of Training on Productivity and Wages: Evidence from British Panel Data', *Oxford Bulletin of Economics and Statistics*, Vol. 68, pp. 397-419.
- Dehejia, R. (2005). 'Does Matching Overcome Lalonde's Critique of Nonexperimental Estimators? A Postscript', Working Paper, Columbia University.
- Doss, C., Grown, C., and C.D. Greene (2007). 'Gender and Asset Ownership', Working Paper, World Bank.

Duflo, E., Glennerster, R., and M. Kremer (2007). 'Using Randomization in Development Economics Research: A Toolkit' in T. Paul Schultz and J. Strauss (eds), *Handbook of Development Economics*, Vol. 4, Elsevier Science Limited, North Holland.

EDA Rural Systems. (2006). *Self Help Groups in India: A Study of Lights and Shades*, EDA Rural Systems, Gurgaon, India..

Goldberg, N. (2005). 'Measuring the Impact of Microfinance: Taking Stock of What We Know', Working Paper, Grameen Foundation USA.

Heckman, J., Ichimura, H., and Todd, P. (1997). 'Matching as an Econometric Evaluation Estimator: Evidence from Evaluating a Job Training Programme', *Review of Economic Studies* , Vol. 64, pp. 605-654.

Ichino, A., and Becker, S. (2002). 'Estimation of Average Treatment Effects Based on Propensity Score', *The Stata Journal* , Vol. 2, pp. 358-377.

Ichino A., Mealli F. and Nannicini T. (2007). 'From Temporary Help Jobs to Permanent Employment: What Can We Learn from Matching Estimators and their Sensitivity?', *Journal of Applied Econometrics* , Vol.23, pp. 305-327.

Karlan, D. (2001). 'Microfinance Impact Assessments: The Perils of using New Members as a Control Group', *Journal of Microfinance* , Vol.3, pp.76-85.

Karlan, D. and Goldberg, N. (2006). 'The Impact of Microfinance: A Review of Methodological Issues', Working Paper, Yale University.

Karlan, D. And Valdivia, M. (2009). 'Teaching Entrepreneurship: Impact of Business Training on Microfinance Clients and Institutions', Working Paper, Yale University.

LaLonde, R.J. (1986). 'The Promise of Public Sector-Sponsored Training Programs', *Journal of Economic Perspectives* , Vol.93, pp. 149-68.

Leuven, E. and Sianesi, B (2009). 'PSMATCH2: Stata module to perform full Mahalanobis and propensity score matching, common support graphing, and covariate imbalance testing', Statistical Software Components S432001, Boston College Department of Economics.

Morduch, J. and Roodman, D. (2009). 'The Impact of Microcredit on the Poor in Bangladesh: Revisiting the Evidence', Working Paper, Center for Global Development.

NABARD. (2006). 'Progress of SHG-Bank Linkage in India : 2005-06', Working Paper, NABARD.

NCAER. (2008). 'Impact and Sustainability of SHG Bank Linkage Programme', Working Paper, NCAER.

Nussbaum, M., Kumar, A., and A. Miehlbradt. (2005). 'Integrating Microenterprise into Markets: The Case of EDA's Leather Subsector Project in India', The Seep Network Case Study #2, Seep Network.

Pitt, M. and Khandker, S. (1998). 'The Impact of Group-Based Credit Programs on Poor Households in Bangladesh: Does the Gender of Participants Matter ?', *Journal of Political Economy* , Vol. 106, pp. 958-996.

Puhazendhi, V. and Badataya, K. (2002). *SHG-Bank Linkage Programme for Rural Poor – An Impact Assessment*, NABARD, Mumbai.

Reserve Bank of India (2008). *Rangarajan Committee on Financial Inclusion*, RBI, Mumbai.

Smith, S. (2002). 'Village Banking and Maternal and Child Health: Evidence from Ecuador and Honduras', *World Development*, Vol. 30, pp. 707-723.

Smith, J. and Todd, P. (2005). 'Does Matching Address Lalonde's Critique of Nonexperimental Estimators?', *Journal of Econometrics*, Vol. 125, pp. 305-353.

TABLES

TABLE 1

Training t-tests

<i>Variable Name</i>	<i>No Training (T=0)</i>	<i>Training (T=1)</i>	<i>T-test for equality of means</i>
	<i>Mean (S.D.)</i>	<i>Mean (S.D.)</i>	
N	367	475	---
Gross Assets (Rs.)	94535 (127418)	126710 (163750)	-3.11***
Income (Rs.)	13805 (14394)	19656 (18549)	-4.99***
Months in SHG	17 (16.05)	21.20 (15.33)	-3.95***
Age (yrs.)	33.76 (7.76)	35.87 (9.0)	-3.58***
Gender (Female=1)	0.95 (0.22)	0.95 (0.22)	0.21
Dep. Ratio	0.66 (0.21)	0.66 (0.22)	0.12
No Education	0.51 (0.50)	0.57 (0.50)	-1.75*
Primary Ed.	0.19 (0.39)	0.17 (0.37)	0.68
Secondary Ed.	0.20 (0.40)	0.14 (0.35)	2.30**
College Ed.	0.03 (0.16)	0.04 (0.19)	-0.86
Owned Land in 2000 (acres)	0.66 (1.24)	1.10 (1.61)	-4.32***
Distance from Paved Road (kms.)	3.22 (3.80)	2.85 (2.55)	1.67**
Distance from Bank (kms.)	7.26 (7.73)	7.45 (5.60)	-0.42
Distance from Market (kms.)	5.13 (4.16)	5.72 (3.81)	-2.14**
Distance from Healthcare (kms.)	3.49 (2.99)	3.62 (2.63)	-0.68
Distance from Bus Stop (kms.)	3.61 (3.72)	3.92 (3.31)	-1.26
Male Wage (Rs.)	46.32 (16.04)	46.25 (13.09)	0.07

*Notes:**** Significant at 1 % level. ** Significant at 5 % level. * Significant at 10 % level.

TABLE 2

Non- training related descriptive statistics

<i>Variable Name</i>	<i>Mature SHGs</i>	<i>New SHGs</i>	<i>Non-Members</i>
	<i>Mean (S.D)</i>	<i>Mean (S.D.)</i>	<i>Mean (S.D.)</i>
N	604	186	51
Gross Assets (Rs.)	109423 (145763)	104933 (136447)	111818 (170171)
Income (Rs.)	16841 (16458)	15460 (17942)	13905 (12269)
Months in SHG	26 (13)	0.31 (1.34)	0
Age (yrs.)	35.2 (8.70)	32.6 5(7.30)	35.60 (8.08)
Gender (Female=1)	0.96 (0.20)	0.92 (0.27)	0.96 (0.20)
Dep. Ratio	0.66 (0.22)	0.69 (0.19)	0.62 (0.23)
No Education	0.51 (0.50)	0.60 (0.50)	0.51 (0.50)
Primary Ed.	0.20 (0.40)	0.12 (0.33)	0.24 (0.43)
Secondary Ed.	0.17 (0.38)	0.18 (0.39)	0.12 (0.33)
College Ed.	0.03 (0.17)	0.04 (0.19)	0.02 (0.14)
Owned Land in 2000 (acres)	0.86 (1.43)	0.89 (1.50)	0.48 (1.12)
Distance Paved Road (kms.)	3.04 (3.43)	2.95 (2.99)	3.60 (3.04)
Distance from Bank (kms.)	7.90 (7.40)	6.30 (5.70)	4.96 (3.20)
Distance from Market (kms.)	5.70 (4.20)	4.34 (3.50)	5.50 (3.20)
Distance from Healthcare (kms.)	3.40 (2.64)	3.61 (3.21)	5.00 (3.30)
Distance from Bus Stop (kms.)	3.80 (3.70)	3.36 (3.15)	4.71 (2.80)
Male Wage (Rs.)	46.00 (12.41)	45.00 (20.00)	54.71 (16.40)

TABLE 3

Regression estimates of impact of training on asset creation and income ($x10^{-2}$)

	(1) <i>Gross Assets</i>	(2) <i>Income</i>	(3) <i>Gross Assets</i>	(4) <i>Income</i>
Member	-459.02 (2.32)**	19.38 (0.92)	-437.71 (2.28)**	25.48 (1.24)
SHGMON	6.34 (1.93)*	-0.74(1.68)*	6.37 (1.92)*	-0.72 (1.66)*
Training (Yes=1)	108.99 (1.18)	27.13 (1.83)*	----	----
Weeks of Training	----	----	14.87 (0.76)	3.01 (1.03)
Age	1.17 (0. 20)	1.28 (2.08)**	1.25 (0.21)	1.31 (2.13)**
Gender (Female=1)	101.17 (0.76)	-0.53 (0.02)	100.84 (0.76)	-0.88(0.03)
Dep. Ratio	402.15 (2.13)**	-109.8 (3.32)***	403.56 (2.15)**	-109.7(3.32)***
Primary Ed.	234.22 (1.92)*	-19.28 (1.10)	233.06 (1.90)*	-19.58 (1.11)
Secondary Ed.	292.87 (2.43)**	-33.15 (2.22)**	287.54 (2.36)***	-34.32(2.31)**
College Ed.	566.93 (2.09)**	-55.65 (1.70)*	567.37 (2.09)***	-55.50 (1.69)*
Land 3 years ago	423.55 (7.89)***	16.04 (2.74)***	426.00 (7.99)***	16. 73(2.86)***
Distance Paved Rd.	-74.96 (2.43)***	-0.27(0.08)	-77.53 (2.53)**	-1.04(0.30)
Distance Bank (kms.)	8.33 (0.72)	-0.92 (0.72)	7.93 (0.69)	-1.06(0.81)
Distance Market	-17.59 (1.57)	-0.002(0.00)	-18.22 (1.62)	-0.14(0.06)
Distance HealthCare	16.65 (0.68)	-1.83(0.66)	17.86 (0.72)	-1.54(0.55)
Distance Bus Stop	46.92 (1.53)	-1.03 (0.32)	47.91 (1.58)	-0.63(0.19)
Male Wage	-4.93 (1.07)	-0.02(0.03)	-4.85 (1.05)	-0.003(0.01)

Notes: *** Significant at the 1 % level. ** Significant at the 5 % level. * Significant at the 10 % level. All regressions include district fixed effects. Analysis based on 841 observations. Absolute t-ratios in parentheses computed with White heteroskedasticity-consistent standard errors clustered by village. See text for definitions of variables.

TABLE 4

Estimates of impact on asset creation and income with respect to infrastructure and provision of business training (x10⁻²)

	(1)	(2)	(3)	(4)
	<i>Gross Assets</i>	<i>Income</i>	<i>Gross Assets</i>	<i>Income</i>
Member	-460.64**	19.27	-373.65*	33.39
	(2.31)	(0.91)	(1.88)	(1.62)
SHGMON	6.92**	-0.70	3.84	-1.25
	(2.11)	(1.61)	(1.16)	(2.66)**
Training (Yes=1)	261.37**	37.40**	-----	-----
	(1.99)	(1.99)		
Distance Paved Rd. (kms.)	-60.41**	0.71	-----	-----
	(1.99)	(0.20)		
Distance Paved Rd.* Training	-50.79**	-3.42	-----	-----
	(2.09)	(1.00)		
Business Training (Yes=1)	-----	-----	-347.78*	-42.31*
			(1.73)	(1.73)
Business Training*SHGMON	-----	-----	14.69*	2.76**
			(1.72)	(3.08)

Notes: ** Significant at the 5 % level. * Significant at the 10 % level. All regressions include household characteristics and village level characteristics as in Table 3 and district dummies. Analysis based on 841 observations. Absolute t-ratios in parentheses computed with White heteroskedasticity-consistent standard errors clustered by village. See text for definitions of variables.

TABLE 5

Matching and regression adjusted matching estimates of training impact on assets and income ($\times 10^{-2}$)

<i>Matching Algorithm</i>	<i>(1) Gross Assets</i>	<i>(2) Income</i>	<i>(3) Gross Assets (Regression Adjusted)^a</i>	<i>(4) Income (Regression Adjusted)</i>
1 NN (S.E.)	176.50 (1.23)	42.34** (2.59)	-----	-----
10 NN (S.E.)	212.76* (1.92)	47.18** (3.75)	-----	-----
LLR (bw 1) (S.E.)	165.61 (1.49)	49.72** (3.54)	201.24** (1.99)	8.15 (0.60)
LLR (bw 4) (S.E.)	165.61 (1.52)	49.72** (3.83)	201.24** (2.12)	8.15 (0.64)

Notes: ** Significant at the 5 % level. * Significant at the 10 % level. NN = neighbor to neighbor, t-stats in parentheses. LLR= local linear regression, p-values in parentheses standard errors created by bootstrap replications of 200. ^aCovariates of regression same at Table 3, (1) and (2), omitting the training variable. See text for definitions of variables.

TABLE 6

Comparison of estimates of training impact on assets and income($\times 10^{-2}$)

<i>Variable</i>	<i>(1) Unadjusted (T-test)</i>	<i>(2) Regression</i>	<i>(3) Matching (LLR, bw 1)</i>	<i>(4) Regression Adjusted Matching (LLR, bw 1)</i>
Assets	321.75**(3.11)	108.99 (1.18)	165.61(1.49)	201.24** (1.99)
Income	58.51**(4.99)	27.13** (2.11)	49.72** (3.54)	8.15 (0.64)

Notes: ** Significant at the 5 % level. * Significant at the 10 % level. For (1) and (2), t-stats in parentheses. For columns (3) and (4), p-values with bootstrap standard errors of 200 replications. (1) is the simple t-test comparison. Column (2) is from Table 3, columns (1) and (2). (3) and (4) are from Table 5.