Furthering Financial Literacy:
Experimental evidence from a financial literacy program for microfinance clients in Bhopal, India

Anna Custers

MSc in Development Studies 2010

Dissertation submitted in partial fulfillment of the requirements of the degree

Furthering Financial Literacy:
Experimental Evidence from a Financial Literacy Training Program for Microfinance Clients in Bhopal, India

Abstract

This study presents evidence from a randomized field experiment that was conducted in Bhopal, India. It investigates the effect of a financial literacy training program on financial literacy levels of microfinance clients. This study finds a significant positive effect on financial literacy levels due to the training. In addition, it tests whether trainer gender influences the impact of the training. It does so by looking at learning outcomes and learning perceptions. For this respectively no and little significant causality is found. The sample for the experiment was drawn from a local microfinance organization that caters to women in the state of Madhya Pradesh.

Acknowledgements

Table of Contents

1. Introduction

Improving individual financial skills has increasingly been paid attention both in the developed and the developing world. In late 2006, the government of the Netherlands, in cooperation with the Dutch Central Bank, universities and several NGOs, introduced an on-line platform to strengthen consumer's financial self-sufficiency[1]. Around the world, similar projects have been launched. The Reserve Bank of India, for example started financial literacy initiatives in late 2007, consisting of a financial education website and free financial education and counseling to various target groups.[2] With 2010 being labeled the year of financial inclusion by the G20's Financial Inclusion Experts Group, the question how to develop financial literacy seems more pertinent than ever before.

The idea behind efforts to improve financial literacy is that it helps households to make better-informed choices in their demand for sound financial services, which in turn are believed to increase their well being. This idea is motivated by a compelling body of evidence that shows a strong association between financial literacy and a household's well being (Cole et al., 2009:3). In particular, financial literacy training can help to prevent problems such as cycle-lending and other forms of unbeneficial financial market participation. In addition, it could increase demand for welfare enhancing financial services like savings accounts and pensions.

This study analyzes the role of financial literacy in financial service demand by implementing a randomized field experiment of a financial literacy training program (FLTP) for microfinance clients in Bhopal, India. It compares individuals who did receive financial literacy training to individuals who did not receive training, but closely resemble the first group. The FLTP has been implemented by Samhita Microfinance, a microfinance organization that caters to women in the state of Madhya Pradesh.

A randomized evaluation is considered to give one of the most unbiased impact assessments of all research methods available.[3] It minimizes biases due to omitted variables, selection bias and reverse causality and is held to produce the optimum

---

[1] See http://www.wijzeringeldzaken.nl/english.aspx
[2] See http://www.rbi.org.in/scripts/PublicationDraftReports.aspx?ID=526
[3] For a discussion of this methodology, see Ravaillon (2009) and Rodrik (2008).

counterfactual inference (Duflo et al., 2006). The method strongly echoes Przeworski's view that counterfactual observations are required to infer causal relationships (Przeworski, 2004:536). Moreover, a randomized field experiment does not only measure impact, but departs from a theoretical framework and subsequently tests this theory in the field. It thus goes beyond a mere program evaluation (Duflo et al. 2006:22).

The aim of this study is to deliver a two-fold contribution to financial literacy theory. Firstly, it departs from the theoretical discussion surrounding financial literacy by extricating a specific link in the theoretical mechanism between financial literacy and household well-being: the link between financial literacy training and financial literacy levels. This link is subsequently tested in the field. To do so this study asks whether a FLTP affects financial literacy levels among microfinance clients.

Secondly, again departing from theory, this study tests how trainer gender may affect learning outcomes and perceptions. This is a relatively under-researched field, in particular in a context of developing countries. It does so by asking whether a differential design of a FLTP, based on trainer gender, differently affects financial literacy levels and learning perceptions among microfinance clients.

This study reports the following main findings: Firstly, it finds a significant positive effect on financial literacy levels as a consequence of the FLTP. Secondly, it finds respectively no and little significant causality of trainer gender on learning outcomes and learning perceptions. These results will be further outlined in the concluding sections of this study.

This study is structured as follows. The next section discusses the two theoretical departures for the randomized evaluation and introduces the three research questions following from this. Section 3 then discusses the empirical setting and research design. Section 4 focuses on measurements and section 5 outlines the sample selection and provides a test of randomization. In section 6 departures of perfect randomization are discussed, including the possible biases this introduces within results. Section 7 will further outline results, which will subsequently be discussed in section 8. Section 9 concludes.

2. Theoretical Framework and Research Questions

This section presents the theoretical discussion that informs the randomized field experiment. The first part discusses the theoretical debate surrounding financial literacy. The second part discusses the theoretical motivation behind the trainer gender experiment within the FLTP.

2.1. Theories on Financial Market Participation

Limited access to, or limited use of financial services generates persistent poverty traps. In that way financial exclusion can act as a break on economic development (Demirgüç-Kunt et al., 2007:22, Beck, Demirgüç-Kunt, 2008:384). Financial exclusion can be either voluntary chosen or involuntary suffered. In the case of involuntary exclusion, access is always restricted and supply fails to meet demand (Claessens, 2006:211). In the case of voluntary exclusion, using financial services is not restricted by availability, but by low levels of demand.[4] This problem is particularly manifest in developing countries. In many instances formal financial markets are emerging quickly in developing countries, yet demand remains limited, in particular among the poor (ATISG, 2010).

Broadly speaking two types of explanations have been offered to explain low demand levels. One view argues that these services are often still too costly for the poor in terms of transaction costs. Opening a savings account for example can be a significant time-investment, especially when traveling to the bank is required. In other words, formal financial participation can be hindered by external cost constraints. An alternative explanation is that demand for financial services is low due to limited awareness or limited financial literacy: if individuals lack knowledge of financial products they will simply have little demand for them (Cole et al., 2009:1, Kempson et al., 2000:50). In this view, formal participation is hindered by internal cognitive constraints.

These two explanations imply quite different policy options to increase demand for financial services. On the one hand, reducing external constraints such as transaction costs, for example via subsidies, could increase demand. On the other hand, reducing

---

[4] The classification of the reasons for exclusion used in this paper are based upon Claessens (2006) and Kempson et al. (2000).

internal cognitive constraints via a FLTP could increase demand. Cole et al. (2009:5) are the first to test these two competing theories in a randomized field experiment among unbanked households in a developing country, Indonesia. The authors find a large effect of small subsidy payments on the probability of opening a savings account, but no effect for a FLTP on opening a savings account for the population at large[5]. They do find a modest impact of the training program on opening a savings account among relatively uneducated and financially illiterate households, however. Nevertheless, their results suggests that the poor are more constraint by external factors, such as limited accessibility because saving products are too expensive, rather than by internal factors.

Given the attention policymakers pay to financial literacy training, these results are remarkable. They call into question the commonly assumed causal mechanism between financial literacy training and improved household well-being. This causal mechanism can be depicted as follows:

Financial Literacy Training $\rightarrow$ Financial Literacy Level $\Uparrow$ $\rightarrow$ Demand Welfare Enhancing Financial Services $\Uparrow$ $\rightarrow$ Household well-being $\Uparrow$
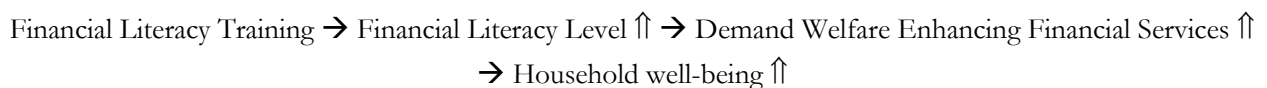
Source: Own Compilation

Cole et al. (2009) focus on this causal mechanism between financial literacy training and demand for welfare enhancing services (measured through the opening of a savings account), but do not measure the impact of training program on financial literacy levels directly. Cole et al. (2009) explore financial literacy levels before the intervention, but do not conduct a post-measurement of financial literacy levels. It is therefore unclear whether the authors find no effect of financial literacy training on demand for savings accounts because the training did not increase financial literacy levels, or because increased financial literacy levels did not culminate in a higher demand for savings accounts. In other words: there remains a possibility that financial literacy levels have gone up without this translating into a higher demand for savings accounts. This point will be elaborated on in the discussion.

---

[5] A savings account is regarded as a sound indicator for financial market participation, as this is generally the first service that people take up when entering the formal financial market (ATISG, 2010:11).

This study builds on the findings of Cole et al. (2009) by investigating the link between financial literacy training and financial literacy levels. It focuses on the determinants of low demand of financial services due to internal cognitive constraints. This is inherent to the focus on financial literacy training. To do so it asks research question 1:

1) Does a financial literacy training program affect financial literacy levels among microfinance clients in Bhopal, India?

In answering this question, this study makes use of a randomized field experiment, in which microfinance clients are randomly assigned to a FLTP or a control group. Section 3.2.1 describes this experiment in more detail.

2.2. Teacher Gender Theories

In addition the testing the effect of a FLTP on financial literacy levels, this study investigates whether teacher gender status has any effect on the training program. Samhita Microfinance has a preference for female trainers. They believe women make better trainers for an all female clientele. This assumption is common; male and female teachers are often presumed to have different teaching styles and capacities, which may have different influence on pupils (Sabbe, Aelterman, 2007:521).

Literature on gender effects in teaching provides at least two theories on how teacher gender may affect learning outcomes and perceptions. One view suggests that self-recognition plays an important role in learning. According to this view, women develop more positive attitudes towards traditionally non-female subjects such as mathematics, when a female teacher teaches them (Li, 1999). An alternative view suggests that the perception of authority over the subject matters for learning: perceived authority differs across subjects and gender. In particular, mathematics is often stereotyped as a male domain (Li, 1999).

These two theories suggest a differential design of financial literacy training for female clients. On the one hand, hiring female trainers could result in better learning outcomes and perceptions of the training because female trainer makes the female clients feel more comfortable in, for example asking questions. On the other hand, hiring male trainers

could give better results. In this case, the female clients may perceive men as more knowledgeable on the subject, or better teachers in terms of leadership.

The effect of gender in teaching is relatively well-researched. Sabbe and Aelterman (2007) provide a comprehensive overview of this research to date. One of their findings is that gender has little, if any effect on learning outcomes (Ibid., 526). They do find some evidence that teacher gender affects learning perceptions (Ibid., 527). The vast majority of research on gender in teaching has been conducted in developed countries, however (Warwick and Jatoi, 1994:377). Similar effects in a development context are relatively under-researched, in particularly with regard to financial literacy training. This might be explained by the fact that financial literacy is a new practice.

Given the above, it remains relevant to test the effects of trainer gender on learning outcomes and perceptions of Samhita's FLTP. This is done on the basis of research questions 2a and 2b:

2a)   Does a differential design of a financial literacy training program, based on trainer gender, differently affect financial literacy levels among microfinance clients in Bhopal, India?

2b)   Does a differential design of a financial literacy training program, based on trainer gender, differently affect perceptions of the training program in Bhopal, India?

In answering these questions, this study randomly allocates comparable male and female trainers to a comparable sub-group of clients that receive financial literacy training. Section 3.2.2 describes this trainer gender experiment in more detail. In addition, this study analyzes trainer evaluations in order to distill whether learning perceptions vary with trainer gender. The trainer evaluation will be discussed in section 4.

3. Empirical Setting and Experiment Design

This section describes the FLTP offered to microfinance clients in Bhopal. It subsequently describes the parameters of the field experiment and the empirical research strategy.

3.1 Empirical Setting: Samhita Microfinance and its Financial Literacy Training Program

To study whether financial literacy training could affect financial literacy levels I had the opportunity to work with Samhita Microfinance. Samhita is a non-profit, financially sustainable microfinance institution, which has been operating in Madhya Pradesh (MP), India, since September 2007. The state ranks among the eight poorest in India, as measured by the multidimensional poverty index, and is the sole Indian state that falls into the 'extremely alarming' category of the India State Hunger Index (UNDP, 2010, Menon et al., 2009). Samhita's microfinance model is based upon Bangladesh's Grameen Bank, with lending groups of five individuals[6]. Four to five of such lending groups typically form a center, which meet for weekly loan repayment. Samhita accepts only female clients.

Samhita's mission is to provide community development services that go beyond providing credit. Its financial literacy program is such a service. The goal of the training is to empower clients in their financial decision-making. To this purpose, Samhita developed a customized financial literacy training for its urban clients in Bhopal, consisting of four sessions: on budgeting, on savings and investments, on insurances and loans and a recap session. The training is tailored to poorly literate people and relies heavily on pictorial and interactive training materials, including a movie[7].

3.2 Field Experiment Design: Financial Literacy Intervention and Trainer Gender Experiment

3.2.1 Financial Literacy Intervention

The evaluation of Samhita's financial literacy training was conducted by means of a randomized phase-in experiment that took place during a field visit in July 2010. A randomized phase-in experiment is considered one of the fairest field experiment varieties within the method of randomized evaluation, as ultimately, all eligible beneficiaries will be subject to the intervention (Duflo et al., 2006:25). Selected clients for treatment received two weekly sessions of two hours for a period of two weeks. The

---

[6] For a more detailed description of this model, see Murdoch (1999).
[7] 48% of its targeted audience was illiterate. Source: baseline survey.

randomization was based upon a baseline survey that was conducted as part of Samhita's FLTP by a professional team between March and May 2010. Out of 9.940 clients, the base-line survey covered 9.195 cients. The members included in the baseline survey are a random subset of Samhita's Bhopal client population. No branches were excluded beforehand. However, some clients were left out due to repeated unavailability at the time of the survey.

In order to ensure internal validity of results, treatment and control groups have to be comparable prior to the intervention, in particular on those characteristics that are likely to highly correlate with the potential to change financial literacy levels. From the baseline survey, three indicators were chosen to form the basis of the randomization.[8] In order of importance:

1. Years of Schooling

Theory suggests that higher human capital, in particular higher cognitive ability, is strongly associated with financial literacy (Cole et al., 2009:12). In the absence of a direct measurement of cognitive ability, following standing academic practice, I use years of schooling as a proxy for human capital.[9]

2. Whether a client has a savings account in her name

A savings account is regarded as a sound indicator for financial market participation, as these are generally the first services people take up when entering the formal financial market (ATISG, 2010:11). Having a savings account is likely to affect initial financial literacy levels, but also the potential to increase financial literacy levels, due to familiarity with the subject matter.

3. Age

In their analysis of what predicts financial literacy, Cole et al. (2009:33) find a significant

---

[8] The indicators are not exhaustive in this respect, but they are held to be of main importance. Other plausible indicators for the randomization would be wealth, literacy levels, household size and other financial services uptake. See also Cole et al. (2009). These indicators were included in the checks of randomization once treatment and control group selection had taken place.

[9] This has long been accepted as a good proxy, but the academic debate in this field is far from settled (Cohen, Soto, 2007:52). For the purpose of this paper, however, years of schooling seems to be an adequate indicator – and in fact the only one available – for educational attainment.

effect for (squared) age on financial literacy scores. A possible explanation for this is that financial participation (and therefore financial literacy scores) varies with age. When testing this hypothesis, Cole et al. (2009:34) find that age is only a significant predictor for formal loans. Yet, even when women do not participate themselves in the financial market, others in their families might (varying with age, for example when women marry or their children grow up), which may in turn influence their financial literacy levels. I therefore decided to select age as one of the main three indicators for the randomization.

In addition, both groups would ideally have a similar variation in initial financial literacy levels prior to the intervention. The baseline survey does not contain a pre-measurement of financial literacy, however.[10] Yet, the three indicators chosen as the basis for the randomization can be considered accurate predictors of initial financial literacy levels. And are therefore considered to produce comparable groups with respect to initial financial literacy level variation as well.

3.2.2. Trainer Gender Experiment

In addition to the financial literacy intervention, I had the opportunity to study whether trainer gender could affect financial literacy levels or perceptions of the training. This by randomly allocating male and female trainers to comparable treatment groups. Working with Samhita enabled the identification of four comparable trainers, of both sexes. All trainers were similar in age, in education, and in working- and teaching-experience. The two more experienced trainers (male and female) were matched, as were the two younger trainers (male and female). Both couples were scheduled to train comparable treatment groups, making gender the most prominent factor of variation.

3.3 Empirical Strategy and Data Collection

The empirical strategy of this study is to compare average financial literacy levels of clients that have received financial literacy training, to those of who have not yet received the training. When assignment to treatment and control group is randomly determined and correctly implemented, the unbiased causal effect of the financial literacy

---

[10] This is often the case in educational interventions, when pre-measurements on educational attainment only take place after the randomization. See for example Banerjee et. al. (2005).

intervention can be obtained by estimating the following simple specification[11]:

$$FinLiteracy_i = \alpha + \beta * LitTraining + \varepsilon_i \qquad (1)$$

Where *FinLiteracy$_i$* is the average financial literacy level for client i after the training, and *LitTraining* is a dummy for assignment to the financial literacy training (i.e the treatment group).

In addition, this study compares average financial literacy levels of clients that have received financial literacy training by a female trainer, to those who have been trained by a male trainer. When assignment to a female and male trainer is randomly determined and correctly followed, the unbiased causal effect of the trainer gender experiment can be obtained by estimating the following simple specification[12]:

$$FinLiteracy_i = \alpha + \beta * FemaleTrainer + \varepsilon_i \qquad (2)$$

Where *FinLiteracy$_i$* is the average financial literacy level for client i after the training, and *FemaleTrainer* is a dummy for assignment to a female trainer.

To measure the effect of the financial literacy intervention on average financial literacy levels, original data collection for the purpose of this study took place during a field visit in July 2010. It concerns pre- and posttests, measuring financial literacy levels of treatment and control groups before and after the financial literacy intervention. To evaluate whether trainer gender could affect perceptions of the training, this study makes use of a trainer evaluation that was part of the posttests. The following section will describe these surveys in more detail.

4. Measuring Financial Literacy and Trainer Gender Perceptions

In their seminal paper on '*Financial Literacy and Planning*', Lusardi and Mitchell (2006) introduce three questions on financial literacy into the nationally representative U.S. Health and Retirement Study, in order to measure financial literacy levels of Americans

---

[11] This equation can be estimated with ordinary least squares. See also: Duflo et al., (2006:8).
[12] Ibid.

over the age of 50. In recent years, these questions have become standard measures for financial literacy levels. They have been extended to different target audiences and different countries. For example, the same questions have been adopted to measure financial literacy among American youth (Lusardi et al., 2009) and Dutch households (Van Rooij et al., 2009). Using the same approach, Cole et al. (2009) are the first to explore financial literacy levels in a developing countries' context; Indonesia and India.

4.1. Financial Literacy Survey

The pre- and posttest contained four questions on financial literacy. These include the three standardized questions as originally developed by Lusardi and Mitchell (2006), adapted to an Indian context by Cole et al. (2009). The questions refer to basic fundamental financial concepts: compound interest rate calculation, inflation and risk diversification (Lusardi and Mitchell, 2006:5). The following two questions have been literally adopted from Cole et al. (2009:7-8):

i) 'Suppose you borrow Rs. 100 from a moneylender at an interest rate of 2 percent per month, with no repayment for three months. After three months, do you owe less than Rs. 102, exactly Rs. 102, or more than Rs. 102?'

ii) 'Suppose you have Rs. 100 in a savings account earning 1% interest per year. Over the period of a year, the prices for goods and services rise 2%. With the money in your savings account, can you buy more than, less than, or the same amount of goods in one year as you could today?'

The third question asked by Cole et al. (2009:8) reads: 'Is it riskier to plant multiple crops or one crop?'.[13] Since the sample of this study concerns an urban, rather than rural population, I decided to adapt this question to an urban context. In consultation with Samhita, the question was adapted to the following:

iii) If you had Rs. 1000 would you put it in a bank that will pay you Rs. 1100 after a year or would you give it to an agent of a scheme who will give you Rs. 2000 after a year?

---

[13] Cole et al. have translated the question from Lusardi and Mitchell (2006) to the Indian context. The original question of Lusardi and Mitchell (2006:5) reads 'Do you think that the following statement is true or false? "Buying a single company stock usually provides a safer return than a stock in a mutual fund."'

Inevitably, this formulation has altered the nature of the question from risk diversification to risk aversion. Yet, the amounts have been chosen such to make the second option refer to a highly deceived financial investment. Choosing the second option, then, would imply lower financial literacy, as this option is unlikely to improve a household's financial well-being.

In addition to these three standardized questions, Cole et al. (2009) developed a fourth question, on simple interest calculation. This question was literally adopted in the survey (Cole et al., 2009:8):

   iv) 'Suppose you need to borrow Rs. 500. Two people offer you a loan. One loan requires you to pay back Rs. 600 in one month. The second loan requires you to pay back in one month Rs. 500 plus 15% interest. Which loan represents a better deal for you?'

4.2 Trainer Evaluation

In the last training session participants were asked to evaluate their trainer by means of a trainer evaluation. Besides covering questions on general perceptions of the training, such as training material and group dynamics, four questions were asked that specifically related to trainer gender perceptions. Responses to these four questions form the bases of the analysis to research question three, which explores how trainer gender may affect learning perceptions.

More specifically, the following two questions were asked in order to evaluate the importance of self-recognition and class comfort:

iv) How well did you think the trainer understood your specific needs in the program?

Very well   Well   Average   Not so well   Poorly

v) How comfortable did you feel with asking questions in class?

Very Comfortable   Comfortable   Somewhat Comfortable   Not very Comfortable   Not al all Comfortable

In order to evaluate perceptions related to subject authority, the following two questions were asked:

vi) Please rate the instructor's teaching and leadership (circle one)

Excellent   Good   Average   Fair   Poor

vii) Please rate the instructor's knowledge for today's training (circle one)

Excellent   Good   Average   Fair   Poor

5. Sample Selection and Test of Randomization

This section describes the experimental sample selection that formed the basis of the randomized field experiment. In addition it provides a test of randomization.

5.1 Experimental Sample Selection and Randomization

The randomized phase-in experiment was carried out within Samhita's logistical constraints. Branch selection for training is assumed largely exogenous to clients' (potential to improve) initial financial literacy levels: Out of five branches, two were selected for initiation of the program: Arera Colony and Nehru Nagar. Proximity to the head office informed this choice. Both branches are considered large and mature by Samhita's branch classification and are comparable in terms of operating lifetime. This precludes any bias due to length of a client's affiliation with Samhita. This is supported when looking at the average number of loans of clients in these two branches: comparing means, using a t-test assuming unequal variances, gives a p-value of 0.50353.[14] Also in terms of loan clients, the two branches are largely comparable: in June 2010, Arera Colony had 2468 loan clients, compared to 1926 in Nehru Nagar.

5.1.1. Treatment Group Selection
Within the two branches, clients were invited at the center-level to participate in the

---

[14] The p-value reported here refers to the experimental sample, as described in this section. Ideal sample would have consisted of all loan clients of these two branches.

FLTP. The availability of training locations, clients and weekly center meeting timings determined which centers were invited for the financial literacy training first. One of Samhita's employees was asked to pick centers for training, without having any prior knowledge about the randomized field experiment that would be implemented next to it.

This approach raises possible concerns of endogeneity: when treatment centers are systematically picked on the basis of (the potential to increase) financial literacy levels, results may be biased. In the case of this field experiment, program-timing bias seems not to be of major concern, however. The availability of a large pool of potential control groups allowed for picking control groups randomly that were very similar in all main baseline characteristics as described in section 3. Thus, even when treatment group selection was not perfectly random but rather determined by Samhita's logistical constraints, control group selection *was* random. This approach ensures the internal validity of results because it prevents any bias due to non-comparability of treatment and control groups, prior to the intervention. It may, however, affect external validity of the experiment, when the selected sample is no longer representative for the population of Samhita's 10.000 urban loan clients. This will be further discussed in section 8.

The final treatment group selection resulted in 17 client clusters for training, of which 9 clusters in Arera Colony and 8 clusters in Nehru Nagar. Each cluster consisted of 2 or more centers, based upon matching weekly center meetings. Over-sampling took place, as it was expected that not all clients would be instantly available. In Arera Colony this resulted in a total of 22 centers selected for treatment, compared to 23 centers in Nehru Nagar. This culminated in a sample of 501 clients for treatment. These clients were invited for the FLTP.

By chance, the treatment groups both within Arera Colony and Nehru Nagar could be matched in pairs which were very similar on the three main baseline indicators: education, savings account and age.[15] This allowed for random allocation of male and female trainers among comparable treatment groups.[16]

---

[15] Due to the odd number of treatment groups in Arera Colony, one group was left out. This resulted in a sample of 469 for the trainer gender analysis.

[16] For every pair, a number between one and two was randomly generated, using Microsoft Excel's unique random number generator. Using the third draw, a one resulted in a female trainer, a two in a male trainer.

5.1.2. Control Group Selection

Control groups were picked separately for Arera Colony and Nehru Nagar, in order to pursue optimum comparability of treatment and control groups. For each branch, a slightly different approach was taken with regard to control group selection. In Nehru Nagar the 152 centers that were left after treatment groups were selected, were colluded into similar 'potential training' groups as per weekly center meetings. This resulted in ideal control groups, as it is these compositions that Samhita would use when scheduling the financial literacy training. The exercise resulted in 64 potential control groups for Nehru Nagar. These 64 potential control groups were then compared on the three main baseline indicators to the treatment group. For each treatment group, two or three comparable control groups were identified, depending on the indicator variation. Actual control groups were subsequently selected by using Microsoft Excel's unique random number generator.[17] This generated four control groups in Nehru Nagar, comprising 8 centers and 98 clients.

Due to logistical constraints, centers could not be colluded as per weekly center meetings for control group selection in Arera Colony. Instead, centers were considered as the randomization level. This resulted in 166 potential control group centers.[18] Subsequently, following a similar approach for control group selection as in Nehru Nagar resulted in five control groups in Arera Colony, comprising 5 centers and 68 clients.

In sum, out of 9.940 potential beneficiaries, an experimental sample of 667 clients was randomly selected for treatment (501) and control groups (166).

5.2 Key Descriptive Statistics and Tests of Randomization

This section reports summary statistics for the experimental sample. In addition it provides a test of randomization. All statistics refer to the experimental sample as described in the previous subsection.

---

[17] For every cluster of comparable control groups, a number between 1 and 3 (2) was randomly generated. After the third draw, the group receiving number two was picked as control group.
[18] Out of 173 centers, 7 centers had to be dropped due to missing base-line data or size (smaller than 6).

The experimental sample in this study concerns female microfinance clients, in an urban setting. The baseline survey reveals that the average age of the women is 34 years (SD 8,0). The distribution of years of schooling is very skewed: 49,6% has had no education at all, whereas of those having attended school, 68% has completed class five or higher, and 56% has completed class eight or higher. Literacy levels reflect these figures: 52% of the women reported to be able to read and write.

Summary statistics on use of financial services prior to the training is presented in Table 1. Since the experimental sample comprises microfinance clients, everyone has at least one loan in her name. In addition, everybody holds a life insurance through Samhita, which covers the original loan amount over the course of a year.

Summary Statistics Experimental Sample	
Loan in her name	100,0%
Life insurance in her name	100,0%
Savings account in her name	37,2%
Debit card in household	14,7%
Insurance in household	28,5%
Knows about pensions	51,0%
Pension in household	9,3%

Table 1. Source: Own compilation from baseline survey. Figures are highly reliable, as whenever a respondent answered 'yes', she was asked to show documents of proof.

Table 2 provides a test of random assignment to treatment and control groups. In addition it provides a test of random assignment within the treatment group to female or male trainers. As can be derived from the table 2, the three main baseline indicators used as basis for the randomization do not vary systematically by treatment or trainer status. Neither do the other baseline indicators that may vary with the (potential to improve) financial literacy levels[19]. The only exception to this is the financial services index in the

---

[19] Note that average income per capita is statistically significantly different in both tests for random assignment, as well as the median income per capita in the test for random trainer gender allocation. While capita per income is reported for reference, it does not affect comparability. Samhita employs a strict income-policy when issuing loans, based on the Samhita Poverty Assessment Score Mechanism, which is more stringent that the globally followed $1.25 per capita PPP daily expenditure criteria. Yet, average per capita measurements came out relatively high, varying from 250 Rs. per month to Rs. 6750 per month. Most probably this was due to measurement errors at the time of reporting monthly income: many respondents may

Financial Literacy Intervention panel.[20] As this is not considered a main predictor for (the potential to change) financial literacy levels, the randomization is still by and large regarded successful.

---

have indicated monthly revenue, rather than profits. Income per capita should therefore not affect comparability.

[20] I constructed a financial services index by collating scores on the following five questions: Do you have a savings account in your name? Does anyone in your household have a Debit/ATM card? Does anyone in your household have an insurance policy? Do you know about pensions? Does anyone in your household have a pension? Per question, 'Yes' and 'No' were given numerical values of 1 and 2 respectively. The minimum total score is 5, the maximum 10.

Test of Random Assignment	Financial Literacy Intervention				Trainer Gender Status (Within Treatment Group)			
	Treatment	Control	Difference significant	p-value	Female Trainer	Male Trainer	Difference significant	p-value
Age								
(mean, t-test unequal variances)	34,1	33,6	No	0,501	34,3	34	No	0,67922
Years of schooling								
(median, Wilcoxon test)	0	4	No	0,4697	1	0	No	0,6512
(mean, t-test unequal variances)	3,77	4,05	No	0,484	3,66	3,57	No	0,82504
Savings account?								
(mean)	1,639	1,596			1,635	1,665		
Chi-square			No	0,328			No	0,495
Income / capita								
(median, Wilcoxon test)	1232,5	1200	No	0,1753	1300	1200	Yes, at 1%	0,009879***
(mean, t-test unequal variances)	1478,98	1299,77	Yes, at 5%	0,00469***	1589,65	1426,97	Yes, at 10%	0,05296*
Literacy levels								
(mean)	1,48	1,47			1,48	1,52		
Chi-square			No	0,788			No	0,404
Household size								
(mean, t-test unequal variances)	4,858	4,886	No	0,82911	4,897	4,864	No	0,80645
Financial services index								
(mean, t-test unequal variances)	8,627	8,446	Yes, at 10%	0,08014*	8,69528	8,65678	No	0,7173
Number of observations	501	166			233	236		

Table 2. Source: Own compilation from baseline survey. The p-value columns report p-values of the respective statistical tests for the hypothesis of equality of means / medians between the treatment (female trainer) and control (male trainer) group, as mentioned in the first column. All t-tests are performed assuming unequal variances. This is considered safer than the Student t-test assuming equal variances in case the true population variances really are unequal, and will still give robust results even when the variances of the true population are equal (Ruxton, 2006). For binary variables, where 'yes' corresponds to a value of 1 and 'no' to a value of 2, a Chi-square test is performed (given the large number of observations and slightly more conservative p-values compared to a Fisher Exact test). When distributions are somewhat skewed, a Wilcoxon median test is performed. In these cases t-tests are also reported for comparison purposes. *** indicates that the difference is statistically significant at the 1 percent level, ** at the 5 percent level, and * at the 10 percent level. Note that the total number of observations in the Trainer Status panel equals 469 instead of 501 individuals, as for the trainer gender analysis one treatment group in Arera Colony was dropped.

6. Departures from Perfect Randomization

The randomized field experiment suffered from both partial compliance and attrition. Paragraph 6.1 and 6.2 will describe how this has influenced internal validity of results. The analysis concerns the financial literacy intervention only. Paragraph 6.3 will discuss how these departures have affected comparability between trainer gender status groups.

6.1 Partial Compliance

The baseline survey reported 99,5% of the clients to be interested in participating in a FLTP. Of these, 222 out of 501 invitees eventually participated in the training program (44,3%). Compliance in the control group (i.e. completing the pretest) was 80,7%.[21] Partial compliance is a common issue in randomized field experiments, and even more so in educational interventions, as it is difficult to compel participation. This problem is usually tackled by producing an intention-to-treat (ITT) estimate from which a treatment on the treated effect (TOT) can be derived. Ideally, this study would have produced both these estimates. However, not all projected treatment and control groups have been fully covered by the pre- (and post-) tests. Only those who complied to the treatment have been subjected to the tests.[22] This prevents estimating an ITT estimate.

As an alternative, this study has cut the sample according to compliance to produce a TOT estimate. This does potentially reintroduce serious selection biases, in particular when individuals that do comply to the intervention have systematically different characteristics compared to those who do not comply. The results may then no longer be representative for all of Samhita's urban microfinance clients, let alone be generalizable to a higher level.

Bearing these reservations in mind there are reasons to believe that this experiment remains instructive. It gives very accurate estimates of the effect of the intervention on the compliers, as long as control and treatment groups are comparable. As Duflo et al.

---

[21] Imperfect compliance (i.e. control group clients receiving the intervention) was no issue in this experiment.

[22] In treatment groups the pretest was administered at the beginning of the first session. In control groups pretests were administered during surprise visits at home. Some control group clients could not be covered due to repeated unavailability at time of the survey.

(2006:51) argue, in some cases it is very valuable to understand the average effect of the intervention itself, rather than that of the ITT. This study is indeed interested in understanding the specific link between financial literacy levels and financial literacy training, rather than the broader policy question whether financial literacy training policies should be scaled up.

The question then is whether treatment and control groups are sufficiently similar to ensure internal validity of results. The comparability could be in jeopardy because the treatment groups were invited to participate in the FLTP and were therefore offered the choice to show up for the program or not, whereas the control group was not subject to this choice factor. The control group was given a surprise visit at home and subjected to the pretest. Compliance in the control group therefore depends less on individual choice.

The experiment does not reveal *why* some clients chose to not participate in the training. For example, a client may not have showed up because she perceived the training as too challenging, or quite the contrary: she may have felt it was of no need for her. Also, clients may just have been unavailable due to work or a visit out of town. If the latter were the case, the bias is likely to be small. In the other cases however, biased estimators in question (1) become problematic as those factors are most likely correlated with (the potential to improve) initial financial literacy levels. It is not obvious in which direction this choice factor could potentially bias results in this case, however.

The choice factor may be correlated with observable or unobservable characteristics that influence (the potential to improve) initial financial literacy levels. If the choice factor is correlated with observable characteristics, this should translate into differences in main baseline indicators between compliers and non-compliers in the treatment group. For example, if those who are higher educated or already participate in the financial market choose not to participate in the financial literacy training, one would expect to observe a significant difference in years of schooling or number of people having a savings account between these groups. This is not the case. When testing the hypotheses of equality of means (medians) of the main baseline indicators between compliers and non-compliers in the treatment group, no systemic differences appear. The same applies to comparing compliers and non-compliers in the control group. Results are reported in Appendix 1.

These results are reinforced when comparing treatment group compliers to control group compliers. Comparing both groups on equality of means (medians) of the main baseline indicators gives very robust results: all p-values are higher than 0,34.[23] For most indicators p-values even increased compared to the test of randomization, which suggests that variations of observed characteristics of compliers (both treatment and control) are more homogenously distributed than in the experimental sample. Moreover, analyzing average pretest scores for complier treatment and control groups also reveals no statistically significant difference between treatment and control group compliers. Pretest scores are evenly distributed, which suggests that outliers do not drive this result. This suggests that the main baseline indicators chosen as the basis for randomization are indeed accurate predictors of financial literacy levels. Results are reported in table 3.

---

[23] When testing for equality of means (medians) of income per capita, the null hypothesis could be rejected at the 5 (10) percent level.

	Financial Literacy Intervention			
	Compliers Treatment	Compliers Control	Difference significant	p-value
Age				
(mean, t-test unequal variances)	34,31	33,77	No	0,51709
Years of schooling				
(median, Wilcoxon test)	0	4	No	0,3551
(mean, t-test unequal variances)	3,71	4,18	No	0,34222
Savings account?				
(mean)	1,60811	1,6194		
Chi-square			No	0,832
Income / capita				
(median, Wilcoxon test)	1250	1170,83	Yes, at 10%	0,09998*
(mean, t-test unequal variances)	1449,78	1272,64	Yes, at 5%	0,02199**
Literacy levels				
(mean)	1,46364	1,46269		
Chi-square			No	0,986
Household size				
(mean, t-test unequal variances)	5,0045	4,9403	No	0,66894
Financial services index				
(mean, t-test unequal variances)	8,54505	8,47015	No	0,55334
Number of observations	222	134		

Table 3. Source: Own compilation from baseline survey. The p-value column reports p-values of the respective statistical tests for the hypothesis of equality of means / medians between the compliers treatment and compliers control group, as mentioned in the first column. For a motivation for the respective statistical tests used, see footnote table 2. *** indicates that the difference is statistically significant at the 1 percent level, ** at the 5 percent level, and * at the 10 percent level.

All the above suggests that despite the partial compliance and the presence of a choice factor for the treatment group, the comparability of complier treatment and control groups on the main baseline indicators and, importantly, on initial financial literacy levels has not significantly changed. These results only refer to observable characteristics, however. There may still be unobservable factors (such as motivation and interest in financial issues) that differ between complier treatment and control group. While this has not shown up in initial financial literacy levels, this may influence the extent to which participants are able to increase their financial literacy levels and thus remains a potential source of bias.

6.2 Attrition

Another source of potential bias is differential attrition between treatment and

comparison groups.[24] The testing procedure was designed to minimize attrition. Participants who attended two or more sessions but were not present at the posttest were visited at home to administer the posttest. Control group compliers were revisited when not at home the first time. Table 4 presents attrition levels and the difference in pretest scores between stayers and attriters.[25]

Attrition Financial Literacy Training		
	Treatment	Control
Financial Literacy Training		
Number of observations	43	32
Percent attrition (after pre-tests)	19%	24%
Pre-test score attritors	2,5349	2,6875
Pre-test score stayers	2,3547	2,3529
Difference in Average Score at Pretest Attriters-Stayers	0,1802	0,3346
t-test mean attriters vs stayers, p-value:	0,37524	0,05162

Table 4. Source: Own compilation. All t-tests are performed assuming unequal variances.

Attrition was 19% in the treatment group and 24% in the control group.[26] Differential attrition can affect internal validity when control and treatment groups are no longer comparable, that is, when different types of people drop out in respective groups (Angrist, 1996). In terms of financial literacy levels, this does seem to be the case: attriters have slightly higher average pretest scores than stayers in both treatment and control groups. This suggests that the less financially literate participants stayed in the experiment. Recalling that Cole et al. (2009) find a modest impact of the training among the relatively uneducated and financially illiterate households (and no impact for the general population at all), this attrition may bias results obtained from equation (1) upwards.

The question, then, is how large this attrition bias is. When testing the hypotheses of equality of means, average pretest scores are significantly different for attriters and stayers in the control group, yet not for attriters and stayers in the treatment group. This result is important with respect to the size of the bias. Within the treatment group, the

---

[24] Attriters are those who did fill a pretest (i.e. compliers), but no posttest.
[25] Similar to Banerjee et al. (2005).
[26] Attrition in the treatment group largely comprised participants that only came once (83,7%).

decision to drop out or not may be due to insufficient need for the program or to other unobservables that correlate with (the potential to increase) initial financial literacy levels. Yet, in as far as these unobservables culminate in different average pretest scores, these differences are statistically insignificant. This is also true for the three main baseline indicators.[27] This suggests a very modest bias due to non-random program drop out. This is supported by observations from the field, in which it appeared that many dropouts were caused by overlap in training timeslots with work or picking up children from school.

Within the control group attrition does seem to be a potential problem, given its magnitude and the significant difference in average pretest scores. However, recalling the testing procedure, this seems rather due to a random draw than to a systemic bias, as staff visits to administer the posttest were unannounced and repeated when not at home. In sum, although prudence in interpreting final results remains imperative, attrition seems unlikely to have had a large additional effect on the validity of results.[28] This result is bolstered when comparing stayers in treatment and control groups on all main baseline indicators: all p-values are in excess of 0,16.[29] Importantly, with a p-value of 0,79755 the difference in average pretest scores is statistically indistinguishable from zero. Results are reported in table 5.

---

[27] See also Appendix 2.
[28] While keeping in mind the potential bias due to the different nature of compliance as discussed in paragraph 6.1.
[29] An exception to this is a t-test for equality of means for income per capita (p-value = 0,05114).

	Financial Literacy Intervention			
	Stayers Treatment	Stayers Control	Difference significant	p-value
Age (mean, t-test unequal variances)	34,39	33,14	No	0,16856
Years of schooling (median, Wilcoxon test)	0	4	No	0,317
(mean, t-test unequal variances)	3,79	4,13	No	0,35414
Savings account? (mean)	1,58659	1,64706		
Chi-square			No	0,318
Income / capita (median, Wilcoxon test)	1200	1170,83	No	0,1622
(mean, t-test unequal variances)	1404,67	1244,91	Yes, at 10%	0,05114*
Literacy levels (mean)	1,46328	1,43137	No	0,67218
Chi-square				
Household size (mean, t-test unequal variances)	5,07263	4,96078	No	0,51592
Financial services index (mean, t-test unequal variances)	8,43575	8,5	No	0,6461
Average pretest score (mean, t-test unequal variances)	2,3547	2,3529	No	0,79755
Number of observations	179	102		

Table 5. Source: Own compilation from baseline survey. The p-value column reports p-values of the respective statistical tests for the hypothesis of equality of means / medians between the compliers treatment and compliers control group, as mentioned in the first column. For a motivation for the respective statistical tests used, see footnote table 2. *** indicates that the difference is statistically significant at the 1 percent level, ** at the 5 percent level, and * at the 10 percent level.

The remainder of this study will analyze effects of the financial literacy intervention and trainer gender experiment based on stayers only, reducing the sample size to 281 women (of which 179 in the treatment group and 102 in the control group). Table 6 summarizes the sampling procedure of this study.

Sampling Procedure		N	Percent
Baseline surveyed members		9195	100%
Baseline surveyed members selected branches		4153	45%
Of whom, Arera Colony		2359	25,7%
Of whom, Nehru Nagar		1794	19,5%
Experimental Sample		667	7,3%
Treatment Group Assignment		501	5,4%
Of whom, FEMALE Trainer		233	
Of whom, MALE Trainer		236	
Control Group Assignment		166	1,8%
Compliers Sample			
Compliers Treatment, according to Pretest		222	2,4%
Of whom, FEMALE Trainer		94	
Of whom, MALE Trainer		117	
Compliers Control, according to Pretest		134	1,5%
Stayers Sample, for final analysis			
Stayers Treatment, according to Pre- and Posttest		179	1,9%
Of whom, Treatment FEMALE		73	
Of whom, Treatment MALE		96	
Stayers Control, according to Pre- and Posttest		102	1,1%

Table 6. Source: own compilation. When randomly assigning treatment groups into FEMALE and MALE trainer status, one group in Arera Colony was left out, due to an uneven number of groups.

6.3. Effects of Departures from Perfect Randomization on Comparability by Trainer Gender Status

When focussing at stayers only, neither partial compliance, nor attrition affects internal validity of the estimators of the trainer gender effect, as both trainer gender status groups belong to stayers within the financial literacy treatment group. All participants are likely to share similar unobservable characteristics (the choice factor in particular) with respect to treatment participation.

Analyzing observable characteristics of stayers by trainer gender status gives improved results compared to the earlier test of randomization: all differences in means (medians) on baseline indicators are statistically indistinguishable from zero and report high p-values. Results are reported in table 7.

	Trainer Gender Status (Within Stayers Treatment Group)			
	Female Trainer	Male Trainer	Difference significant	p-value
Age (mean, t-test unequal variances)	34,07	34,73	No	0,58588
Years of schooling (median, Wilcoxon test)	0	0	No	0,794
(mean, t-test unequal variances)	3,78082	3,52083	No	0,71426
Savings account? (mean)	1,57534	1,59375		
Chi-square			No	0,81
Income / capita (median, Wilcoxon test)	1250	1200	No	0,2435
(mean, t-test unequal variances)	1454	1409,56	No	0,69683
Literacy levels (mean)	1,48611	1,436311	No	0,7704
Chi-square				
Household size (mean, t-test unequal variances)	5,09589	5,07292	No	0,91798
Financial services index (mean, t-test unequal variances)	8,45205	8,47917	No	0,88148
Average pretest score (mean, t-test unequal variances)	2,054796	2,55208	Yes, at 1%	0,00344
Number of observations	73	96		

Table 7. Source: Own compilation from baseline survey. The p-value column reports p-values of the respective statistical tests for the hypothesis of equality of means / medians between the compliers treatment and compliers control group, as mentioned in the first column. For a motivation for the respective statistical tests used, see footnote table 2. *** indicates that the difference is statistically significant at the 1 percent level, ** at the 5 percent level, and * at the 10 percent level.

Strikingly, when comparing average pretest scores by trainer gender status a statistically significant difference at the 1% level appears. This is remarkable as both groups share similar variations on the characteristics that are regarded as predictors for financial literacy levels (i.e. the main baseline indicators). There are three plausible explanations for this result. Differential financial literacy levels by trainer gender status could be due to either:

1. Observable characteristics that differ systematically between the two subgroups;

2. Random probability in draw from sampling distribution;

3. Unobservable characteristics that differ systematically between the two subgroups.

As already shown in table 7, observable characteristics seem not to explain this result, as the tests of equality of means (medians) report very high p-values on all indicators. The outcome could also simply be the result of a random draw from the underlying sampling distribution: the compliers treatment group. Although this is possible, it seems more likely that an unobservable factor is at work that is not captured in the main observable indicators, as the difference appears rather systemic.

Luckily, this pre-existing difference can be controlled for, by estimating a difference-in-difference estimator. This is possible because pre- and posttest data exists on financial literacy levels by treatment status. Assuming that in absence of the treatment both groups would have followed parallel trends in terms of financial literacy developments, the unbiased estimate of the treatment effect can be estimated by the following simple specification (See also Duflo et al., 2006:12):

$$\Delta FinLiteracy_i = \alpha + \beta * FemaleTrainer + \varepsilon_i \tag{3}$$

Where $\Delta FinLiteracy_i$ is the change in average financial literacy levels, and *FemaleTrainer* is a dummy for assignment to a female trainer. This specification will be estimated in the following section.

7. Results

This section reports experimental results for both the financial literacy intervention and the trainer gender experiment. When interpreting results, it is important to keep in mind that results concern the stayers sample only. For the possible biases that may result from this, see the previous section.

7.1 Financial Literacy Intervention

The FLTP as conducted by Samhita appears to have a very significant effect on participants' average financial literacy levels. The difference in average financial literacy scores on pre- and posttests for stayers in the treatment group is significant at the 1% level. Table 8 tabulates this result, in addition to presenting disaggregated scores per

question. Scores improved significantly on the three standardized questions based on Lusardi and Mitchell (2006), but not on question four.

Measured Financial Literacy, Stayers Treatment Group		Pretest	Posttest	Difference significant	p-value
1. Compound Interest	% Correct	72,6%	86,6%	Yes, at 1%	0,001***
2. If savings earns 1% and inflation is 2%, after one year is buying power greater, less or the same?	% Correct	73,7%	84,9%	Yes, at 5%	0,013**
3. Is a bank safer for savings than an agent?	% Correct	40,8%	77,7%	Yes, at 1%	0***
4. Borrowing Rs. 500, repaying Rs. 600 versus repaying 15 %	% Correct	51,4%	54,8%	No	0,525
Question 1 and 2 taken together	% Correct	49,2%	73,2%	Yes, at 1%	0,00004***
Question 1, 2 and 3 taken together	% Correct	24,6%	57,0%	Yes, at 1%	0***
All four questions taken together	% Correct	17,3%	33,0%	Yes, at 1%	0***
Mean share of correct answers question 1,2 and 3		62,4%	83,1%		
Mean share of correct answers all four questions		59,6%	76,0%		
All questions taken together	Avg. Score (out of 4)	2,385475	3,03911	Yes, at 1%	0***
Number of oberservations		179	179		

Table 8. Source: Own compilation from pre- and posttests. The p-value column reports p-values of a Chi-square tests, as it concerns binary variables: an answer can be either 'right' or 'wrong'. *** indicates that the difference is statistically significant at the 1 percent level, ** at the 5 percent level, and * at the 10 percent level.

These results are derived from looking at the treatment group only.[30] Although this gives a fair initial indication of the effect of the FLTP, it does not tell much about the magnitude of the effect. Also, it does not fully exclude the possibility that factors other than the FLTP may have affected financial literacy levels the time period in during which the training was took place. Causal inference therefore cannot be firmly established on

---

[30] For a similar methodology see Bertrand and Mullainathan (2004).

basis of the results of the treatment group alone. To obtain estimates of the *causal* effect, therefore, equation (1) is estimated. This includes control group data. These results are presented in table 9.

Table 9. Source: Own compilation from posttests.

The point estimate on LitTraining in equation (1) is large and statistically significant. This confirms the previous result and suggests that the FLTP has a large effect on participants' initial financial literacy level. A participant receiving the training program is likely to increase her average financial literacy level by an average score of 0,69 out of 4. This is equivalent to 17,25%.

7.2. Trainer Gender Effects

Testing the possibility that trainer gender affects learning outcomes results in no significant effect: when estimating equation (2) the point estimate on FemaleTrainer is negative but statistically insignificant. Remarkably, when estimating equation (3), controlling for the difference in pretests by trainer gender status, the point estimate on FemaleTrainer switches signs and turns positive. This suggests that pretest scores are an omitted variable in equation (2). This is not remarkable given the significantly lower average pretest scores for participants trained by a female. The result remains insignificant, however. This suggests that trainer gender status does not significantly affect learning outcomes in this experiment. Table 10 reports results.

QuickTime™ and a
decompressor
are needed to see this picture.

Table 10. Source: Own compilation from pre- and posttests.

Analyzing trainer evaluations by comparing cumulative ratings by trainer gender, a small systemic trend in learning perceptions can be detected. Comparing questions that are associated with self-recognition and class comfort reveals no substantial difference in learning perceptions by trainer gender. Yet, when comparing trainer evaluation questions associated with perceptions of subject authority, perceptions of male trainers slightly outstrip those of female trainers systematically. Figures 1-4 picture this trend.

QuickTime™ and a
decompressor
are needed to see this picture.

Figure 1. Source: Own Compilation. Note: percentages are cumulative.

QuickTime™ and a
decompressor
are needed to see this picture.

Figure 2. Source: Own Compilation. Note: percentages are cumulative.

QuickTime™ and a
decompressor
are needed to see this picture.

Figure 3. Source: Own Compilation. Note: percentages are cumulative.

QuickTime™ and a
decompressor
are needed to see this picture.

Figure 4. Source: Own Compilation. Note: percentages are cumulative.

8. Discussion

8.1. Financial Literacy Intervention

The results of the experiment show that a FLTP for urban microfinance clients positively affects average financial literacy levels. Combining this finding with earlier research on the correlation between financial literacy and a household's wellbeing, suggests that the increase in financial literacy levels as a result of the FLTP found in this study would culminate into increased demand for welfare enhancing financial services. This suggests a conclusion quite different from previous experiments with financial literacy training is possible. Cole et al. (2009) conclude that financial literacy training is hardly beneficial because it does not significantly increase demand for financial services. This study, however, has shown that a FLTP is beneficial because it can raise financial literacy levels.

Two possibilities follow from this: Either Cole et al.'s (2009) financial literacy training was unsuccessful in raising financial literacy levels. Or, it could have been successful, but, as they indeed argue, external constraints, rather than limited financial literacy, explain the low financial services uptake as found in their study. To be able to exclude either one of these possibilities, a follow-up study of the findings of my experiment needs to be

done, in order to see whether higher financial literacy levels translate into higher financial services uptake.[31] If this is the case, these findings in this study would contradict Cole et al. (2009). This, in turn, would give reason to continue and improve the practice of giving FLTPs as parts of development programs. If a follow-up study would not culminate into higher financial services uptake it would strengthen Cole et al.'s (2009) findings. This then, would point towards the need to invest in removing external constraints.

The positive outcome of the experiment in this study holds only if alternative interpretations of the results can be excluded. There is a possibility that the financial literacy measurement does not accurately capture financial literacy levels of the chosen population. Although this paper has used the best methods available, this does not mean they are foolproof. The questions used to measure financial literacy are heavily geared towards testing mathematical financial skills. While mathematical skills are important, there is more to financial literacy than calculation. The poor may be financially illiterate, but this does not prevent them from being clever about money. As Duflo aptly puts it, '*they are "incredibly smart" about day-to-day financial matters, "because the cost of errors is much bigger", but "so busy doing this effort, and optimizing on some margin, that they might entirely miss some huge elephant in the room," like the importance of buying fertilizer for their crops, or immunizing their children' (Quoted in Parker 2010)*. Put differently, knowing about the value of putting your money into a savings account, rather than keeping it under the mattress, is an expression of financial literacy that needs no calculations. Perhaps the key question is rather how to define financial literacy in terms of financial capability when considering the poor. This suggests that more work needs to be done on defining and measuring financial literacy in such a way that it captures forms of financial competences beyond mathematical-financial skills.

8.2 Trainer Gender Experiment

The experimental results show no significant effect of trainer gender on learning outcomes, suggesting it does not matter whether male or female trainers are employed. This result supports earlier findings by Sabbe and Aelterman (2007) who emphasize that research to date has found little, if any effect of gender on learning outcomes. However, this study does find some evidence that trainer gender affects learning perceptions. In

---

[31] Time-constraints prevented such follow-up, this would be the starting point for further research.

particular, it finds that subject authority perceptions are slightly higher for male trainers. This is also consistent with earlier findings of small evidence that trainer gender affects learning perceptions (Sabbe and Aelterman, 2007:527). It has to be kept in mind, however, that the trend based in this study is based on relatively small differences between subjective ratings of 'excellent' and 'good'. Surprisingly, no differential effects appear with regard to self-recognition and class comfort. Over 90% of participants by both male and female trainer report to feel (very) comfortable in class with asking questions, and over 97% reports that the trainer understood their specific needs (very) well. Rather, participants seemed to be very satisfied with the training overall, which may have eclipsed the evaluations of trainer gender effects in this case.

These results suggest that Samhita should continue using both female and male trainers. Samhita's preference for female trainers is not necessarily justified by reasons of improving learning impact alone. Again, the conclusions of this study are based on observations of two pairs of comparable teachers only, with highly satisfied training participants. Although this is not an exceptional research setting[32], more comparative study between male and female trainers is needed to better understand general patterns in trainer gender perceptions and outcomes in a development context.

Like in the case of the financial literacy intervention, the findings for trainer gender effects as presented above only hold provided that perceptions are measured accurately.[33] Although the measurement is designed on the basis of a coherent theoretical gender framework, it does not fully exclude the possibility of measuring perceptions that go beyond trainer gender, such as personal characteristics. When interpreting trainer gender perceptions, prudence is therefore essential. The possible bias has been limited, however, by hiring the most comparable set of trainers within the pool available.

---

[32] 'It is not unusual to see a study with samples of one or two teachers' (Sabbe and Aelterman, 2007:528).
[33] For learning outcomes the same reservations with regard to financial literacy measurements apply as discussed under subsection 8.1.

8.3. Validity of results: possible pitfalls

8.3.1. Internal Validity

The validity of results in a randomized field experiment strongly relies on comparability of treatment and control groups (Duflo et al.,2006:8). As extensively discussed in section 6, partial compliance and attrition have potentially reintroduced selection bias within the experimental sample. In particular the fact that randomization was based upon baseline characteristics, and not pretests, prevented eliminating the choice factor whether to participate in the financial literacy training or not. The choice factor that was present in treatment groups, but not in control groups, may have split the assigned treatment group along unobservable characteristics that could be correlated with the potential to increase financial literacy levels. This may have resulted in heterogeneous variation in unobservables between treatment and control groups. This variation is not heterogeneous between compliers and non-compliers, and stayers and attriters respectively, with respect to the observables 'years of schooling', 'previous financial market participation' and 'age'. This means there are good reasons to believe the direction of results holds. The magnitude, however, should be interpreted with care.

8.3.2. External Validity

There are three dimensions to external validity of a randomized evaluation: whether a program was implemented with special care, how specific the sample is, and how specific the program is to the context in which the randomized evaluation took place (Duflo et al., 2006:70). In this study, the last two dimensions are of particular concern.

The specificity of the stayers sample limits generalizability of results to the whole experimental sample as originally selected for the randomization. Results do not necessarily hold for those who did not comply to the treatment. The partial compliance encountered in this study, however is not exceptional. Rather, in any training program, some degree of partial compliance is likely to be present. In that respect, results may still be instructive for judging the effects of FLTPs on financial literacy levels.

Even if results hold for the experimental sample, the question is to what extent this sample is representative for Samhita's urban microfinance population. As discussed

under subsection 5.1, selection of the experimental sample is random: there are no reasons to believe that branch selection was endogenous to clients' financial literacy levels. Lastly, then, the question is to what extent Samhita's urban microfinance population is representative for urban microfinance clients across India. As microfinance institutions largely target similar groups of the poor, results could be relevant for other microfinance clients as well.[34] However, the specific district context of Madhya Pradesh has to be kept in mind when interpreting results. Results may be less instructive for poor non-microfinance populations in India as these populations are for example likely to be less financially experienced or, importantly, also includes males.

Closely related to specificity of the sample is specificity of the program. Possible program variations are infinite, yet this does not preclude deriving implications for future FLTPs. The large significant effect on participants' average financial literacy levels suggests that the content of the program is educative. In addition, the trainer gender experiment has revealed that trainer gender does not significantly affect learning outcomes. On the other hand, a different training frequency, for example one session per week, instead of two, could improve attendance. Although this has not been tested empirically, the relatively low compliance to the treatment (and the reasons mentioned) suggests that this could greatly improve reach of the training.

Further research would ideally evaluate the impact of FLTPs comparing various populations (both microfinance and non-microfinance), and various versions of the program. Only then claims of replicability of the training program effects can be made. The principal merit of this study lies in further testing two theoretical frameworks that underpin FLTPs and gender difference in teaching. Testing these theories in a different context provides further evidence on the benefits of financial literacy training programs. Furthermore, this study adds a new chapter to gender debate in research on teaching.

---

[34] In a similar way as Banerjee et al. (2009) suggest general inferences about the effects of microfinance based on a randomized evaluation of the impact of introducing microcredit in a new market.

9. Conclusion

This study has, to the best of my knowledge, been a first exploration of the specific link between financial literacy training and financial literacy levels, within the broader field of research into the causal mechanism between financial literacy and a household's well-being. It provides empirical evidence that a carefully designed financial literacy training program for microfinance clients in Bhopal increased average financial literacy scores of the participants. In addition, it finds that trainer gender does not significantly affect learning outcomes, but does affect learning perceptions: perceived subject authority is slightly higher for male trainers than for female trainers.

The data gathered for this study could be further explored with respect to heterogeneous treatment effects. For example, the data could be probed to see whether the increase in financial literacy varies by years of schooling, previous financial market participation and wealth. This goes beyond the purpose of this study, but could aid better understanding the dynamics of furthering financial literacy. In addition, it could be explored whether the effect of a financial literacy training program varies for different initial financial literacy levels (i.e. if those slightly less financially literate experience a larger impact of the training). Finally, the data could generate interesting comparative case study material to Cole et al. (2009), by studying financial literacy levels among different segments of the poor in India.

By studying the causal mechanism between financial literacy and a household's well-being this study has focused on microfinance clients in Bhopal. A next step is to repeat this study for other financial literacy training programs in developing countries and devising better ways of measuring financial capabilities of the poor. Such research, like this study, could be a stepping-stone towards improved financial literacy training programs that cater to the needs of the poor.

References

Angrist, J. (1996), *Conditioning on the Probability of Selection to Control Selection Bias*, NBER Technical Paper No. 181.

ATISG, (2010), *Innovative Financial Inclusion, Principles and Report on Innovative Financial Inclusion from the Access through Innovation Sub-Group of the G20 Financial Inclusion.* Retrieved From www.microfinancegateway.org/gm/.../Innovative_Financial_Inclusion.pdf. Accessed on 2nd August, 2010, 13:00PM.

Banerjee, A., Cole, S., Duflo, E., Linden, L. (2005), *Remedying Education: Evidence from Two Randomized Experiments in India*, NBER Working Paper No. 11904.

Banerjee, A., Duflo, E., Glennerster, R., Kinnan, C. (2009), *The Miracle of Microfinance? Evidence from a Randomized Evaluation*, Working Paper.

Demirgüç-Kunt, A., Beck, T., Honohan, P. (2008), 'Finance for All?' *World Bank Policy Research Report,* Washington, DC: The World Bank.

Chen, S., Ravaillon, M. (2008), *The Developing World Is Poorer Than We Thought, But No Less Successful in the Fight against Poverty*, World Bank Policy Research Working Paper 4703.

Claessens, S. (2006), 'Access to Financial Services: A Review of the Issues and Public Policy Objectives', *The World Bank Research Observer*, 21(2), pp. 207-240.

Cohen, D., Soto, M. (2007), 'Growth and fuman captial: good data, good results', *Journal of Economic Growth,* 12, pp.51-76.

Cole, S., Sampson, T. Zia, B. (2009), *Money or Knowledge? What drives demand for financial services in emerging markets?,* Harvard Business School Working paper No. 09-117.

Demirgüç-Kunt, A., Beck, T., Soledad Martinez, M. (2007), 'Reaching out: Access to and use of banking services across countries', *Journal of Financial Economics* 85, pp. 234-266.

Duflo, E., Glennerster, R., Kremer, M. (2006), *Using Randomization in Development Economics Research: A Toolkit*, NBER Technical Working Paper No. 333.

Lusardi, A., Mitchell, O. (2006), *Financial Literacy and Planning: Implications for Retirement Wellbeing*, Pension Research Council Working Paper No.1.

Lusardi, A., Mitchell, O., Curto, V. (2009), *Financial Literacy Among the Young: Evidence and Implications for Consumer Policy*, NBER Working Paper No. 15352.

Karlan, D., (2010) 'Helping the poor save more'. *Standford Social Innovation Review*, Winter 2010

Kempson, E., Whyley, C., Caskey, J., Collard, S. (2000), 'In or Out? Financial Exclusion: A Literature and Research Review', *Consumer Research Report 3*, London: Financial Services Authority.

Menon, P., Deolalikar, A., Bhaskar, A. (2009), *India State Hunger Index, Comparisons of Hunger Across States*, Washington, DC: IFPRI, Bonn: Welt Hunger Hilfe, Riverside: University of California.

Murdoch, j. (1999), 'The Microfinance Promise', Journal of Economic Literature, 37(4), pp. 1569-1614.

Li, Q. (1999), 'Teachers' Beliefs and Gender Differences in Mathematics: A Review', *Educational Research*, 41(1), pp. 63-76.

Parker, I. (2010), 'The Poverty Lab: Transforming development economics, one experiment at a time.' *The New Yorker*, May 17th.

Przeworski, A. (2004), Institutions Matter?, *Government and Opposition*, 39(4), pp.527-40.

Ravaillon, M. (2009), 'Should the Randomistas Rule?', *The Berkeley Electronic Press*, February.

Rodrik, D. (2008), *The New Development Economics: We Shall Experiment, But How Shall We Learn?*, HKS Working Paper No. RWP08-055.

Ruxton, G. (2006), 'The Unequal Variance T-test is an Underused Alterative to Student's T-test and the Mann-Whitney U Test', Behavioral Ecology, pp.688 -690.

Sabbe, E., Aelterman, A. (2007), 'Gender in Teaching: A Literature Review', *Teachers and Teaching*, 13(5), pp.521-538.

Van Rooij, M. Lusardi, A., Alessie, R. (2009), *Financial Literacy and Retirement Planning in the Netherlands*, DNB Working Paper No. 231.

Warwick, D., Jatoi, H. (1994), 'Teacher Gender and Student Achievement in Pakistan', *Comparative Education Review*, 38(3), pp.377-399.

Other sources

UNDP multidimensional poverty index
http://hdr.undp.org/en/reports/global/hdr2010/news/title,20523,en.html
Accessed on 30th July, 2010, 15:30PM

http://www.wijzeringeldzaken.nl/english.aspx
Accessed on 2nd August, 2010, 12:00PM

http://www.rbi.org.in/scripts/PublicationDraftReports.aspx?ID=526
Accessed on 2nd August, 2010, 12:00PM

Appendix 1

	Financial Literacy Intervention				Financial Literacy Intervention			
	Compliers Treatment	Non-Compliers Treatment	Difference significant	p-value	Compliers Control	Non-Compliers Control	Difference significant	p-value
Age								
(mean, t-test unequal variances)	34,31	33,93	No	0,59573	33,77	33	No	0,69761
Years of schooling								
(median, Wilcoxon test)	0	1	No	0,7347	4	3	No	0,5555
(mean, t-test unequal variances)	3,71	3,82	No	0,78433	4,18	3,4	No	0,38757
Savings account?								
(mean)	1,60811	1,66308			1,6194	1,5		
Chi-square			No	0,203			No	0,216
Income / capita								
(median, Wilcoxon test)	1250	1207,142857	No	0,8701	1170,83			
(mean, t-test unequal variances)	1449,78	1502,47	No	0,50415	1272,64	1413,37	No	0,23881
Literacy levels								
(mean)	1,46364	1,4964			1,46269	1,5		
Chi-square			No	0,467			No	0,705
Household size								
(mean, t-test unequal variances)	5,0045	4,74194	Yes, at 5%	0,0388**	4,9403	4,65625	No	0,36804
Financial services index								
(mean, t-test unequal variances)	8,54505	8,69176	No	0,15995	8,47015	8,32375	No	0,57255
Number of observations	222	279			134	32		

Appendix 1. Source: Own compilation from baseline survey. The p-value columns report p-values of the respective statistical tests for the hypothesis of equality of means / medians between the treatment and control group, as mentioned in the first column. For a motivation for the respective statistical tests used, see footnote table 2. *** indicates that the difference is statistically significant at the 1 percent level, ** at the 5 percent level, and * at the 10 percent level.

Appendix 2

	Financial Literacy Intervention				Financial Literacy Intervention			
	Stayers Treatment	Attriters Treatment	Difference significant	p-value	Stayers Control	Attriters Control	Difference significant	p-value
Age								
(mean, t-test unequal variances)	34,39	33,98	No	0,77468	33,14	35,78	No	0,11607
Years of schooling								
(median, Wilcoxon test)	0	2	No	0,7722	4	0	No	0,4992
(mean, t-test unequal variances)	3,79	3,37	No	0,55235	4,13	3,75	No	0,54641
Savings account?								
(mean)	1,58659	1,69767			1,64706	1,53125		
Chi-square			No	0,18			No	0,297
Income / capita								
(median, Wilcoxon test)	1200	1333,33	No	0,2679	1170,83	1183,33333	No	0,5945
(mean, t-test unequal variances)	1404,67	1637,58	No	0,14723	1244,91	1361,01146	No	0,43066
Literacy levels								
(mean)	1,46328	1,46512			1,43137	1,5625		
Chi-square			No	0,982			No	0,226
Household size								
(mean, t-test unequal variances)	5,07263	4,72093	No	0,14126	4,96078	4,875	No	0,74759
Financial services index								
(mean, t-test unequal variances)	8,43575	9	Yes, at 1%	0,00195***	8,5	8,375	No	0,62867
Number of observations	179	43			102	32		

Appendix 2. Source: Own compilation from baseline survey. The p-value columns report p-values of the respective statistical tests for the hypothesis of equality of means / medians between the treatment and control group, as mentioned in the first column. For a motivation for the respective statistical tests used, see footnote table 2. *** indicates that the difference is statistically significant at the 1 percent level, ** at the 5 percent level, and * at the 10 percent level.