Smoke and Mirrors: Evidence of Microfinance Impact from an Evaluation of SEWA Bank in India

Duvendack, Maren

University of East Anglia, Norwich, UK

August 2010

ISSN 1756-7904

INTERNATIONAL DEVELOPMENT
University of East Anglia

Working Paper Series

Smoke and Mirrors: Evidence of Microfinance Impact from an Evaluation of SEWA Bank in India

Maren Duvendack

2010

University of East Anglia

Working Paper 24

The School of International Development, University of East Anglia
Norwich, NR4 7TJ, United Kingdom

DEV Working Paper 24

Smoke and Mirrors: Evidence of Microfinance Impact from an Evaluation of SEWA Bank in India

Maren Duvendack

First published by the School of International Development in August 2010.

This publication may be cited as:
Duvendack, M., 2010, Smoke and Mirrors: Evidence of Microfinance Impact from an Evaluation of SEWA Bank in India, Working Paper 24, DEV Working Paper Series, The School of International Development, University of East Anglia, UK.

About the Author
Maren Duvendack is a Research Postgraduate in the School of International Development at the University of East Anglia, Norwich, UK.

Contact:
m.duvendack@uea.ac.uk
School of International Development
University of East Anglia
Norwich, NR4 7TJ
United Kingdom
Tel: +44(0)1603 592329
Fax: +44(0)1603 451999

ISSN 1756-7904

About the DEV Working Paper Series

The Working Paper Series profiles research and policy work conducted by the School of International Development and International Development UEA (see below). Launched in 2007, it provides an opportunity for staff, associated researchers and fellows to disseminate original research and policy advice on a wide range of subjects. All papers are peer reviewed within the School.

About the School of International Development

The School of International Development (DEV) applies economic, social and natural science disciplines to the study of international development, with special emphasis on social and environmental change and poverty alleviation. DEV has a strong commitment to an interdisciplinary research and teaching approach to Development Studies and the study of poverty.

International Development UEA (formerly Overseas Development Group)

Founded in 1967, International Development UEA is a charitable company wholly owned by the University of East Anglia, which handles the consultancy, research, and training undertaken by the faculty members in DEV and approximately 200 external consultants. Since its foundation it has provided training for professionals from more than 70 countries and completed over 1,000 consultancy and research assignments. International Development UEA provides DEV staff with opportunities to participate in on-going development work, practical and policy related engagement which add a unique and valuable element to the School's teaching programmes.

For further information on DEV and the International Development UEA, please contact:

School of International Development
University of East Anglia, Norwich NR4 7TJ, United Kingdom
Tel: +44 (0)1603 592329
Fax: +44 (0)1603 451999
Email: dev.general@uea.ac.uk
Webpage: www.uea.ac.uk/dev

Abstract

Microfinance has been on the development agenda for more than 30 years, heralded as the wondrous tool that reduces poverty and empowers women (Hulme and Mosley, 1996; Rutherford, 2001; Morduch and Haley, 2002; Khandker, 1998). Doubts, however, have recently been raised about the success of microfinance (Dichter and Harper, 2007; Banerjee et al, 2009; Roodman and Morduch, 2009; Karlan and Zinman, 2009; Bateman and Chang, 2009).

Given this context, this paper re-examines the microfinance impact evaluation of SEWA Bank conducted by the United States Agency for International Development (USAID) in India in 1998 and 2000. The USAID panel and a new cross-section data set are analysed using propensity score matching (PSM) and panel data techniques to address selection bias. Sensitivity analysis of the matching results is used to explore their reliability. Various sub-group comparisons between borrowers, savers and controls are also conducted to shed some light on the impact of savings versus credit.

The paper concludes that doubts remain about the quality of the impact estimates obtained through advanced econometric techniques. Direct observation and the outcome of sensitivity analysis of the PSM analysis suggest that the application of PSM and differences-in-differences (DID) to these observational data were probably unable to account for selection on unobservables[1].

Key words: Impact evaluation, evaluation methods, selection bias, microfinance, India

Forthcoming in T. Nair (ed.) *Development Promise of Indian Microfinance* (provisional title), New Delhi: Routledge.

[1] I would like to thank Richard Palmer-Jones for his time, patience and invaluable feedback throughout the analysis and write-up of this study. I would also like to thank Arjan Verschoor and Ben D'Exelle for critical reviews, and the University of East Anglia for financial support.
This paper is adapted from a chapter in the author's PhD thesis. The full version is available upon request. Correspondence: M.Duvendack@uea.ac.uk

Table of Contents

List of Tables

List of Figures

List of Boxes

1 Introduction

Microfinance interventions have become an important strategy in the fight against poverty for many developing countries. By the end of 2006, more than 10,000 microfinance institutions (MFIs) existed worldwide serving an estimated 100 million microfinance borrowers (Dieckmann, 2007). Despite the popularity of microfinance, however, there is little convincing evidence that microfinance programmes have positive impacts. A number of studies suggest positive social and economic impacts of microfinance (e.g. Hulme and Mosley, 1996; Rutherford, 2001; Morduch and Haley, 2002; Khandker, 1998) although recent studies (e.g. Dichter and Harper, 2007; Banerjee et al, 2009; Roodman and Morduch, 2009; Karlan and Zinman, 2009; Bateman and Chang, 2009) have raised doubts about the success of the microfinance phenomenon.

Rigorous microfinance impact studies are rare and most fail to control for selection bias, undermining estimates of impact. It is argued that microfinance clients commonly self-select into a programme or are selected by their peers or the microfinance loan officers, biasing access to loans against the poorer. It is further hypothesized that this selection or screening process is driven by unobservable characteristics such as entrepreneurial abilities, access to social networks, risk taking preferences and business skills (Coleman, 1999). Those characteristics are notoriously difficult to measure and are poorly dealt with or neglected by advanced econometric techniques. Many positive evaluations of microfinance may be misleading because of their failure to account for selection-bias, and selection bias may account for the exclusion of poorer and other marginal groups. Selection bias occurs when the characteristics of those who are treated, i.e. those who save with or borrow from SEWA Bank, are different to the population at large, or more specifically the control group with which they are compared in order to estimate impact (Manski, 1995). If selection bias cannot be controlled for then the impact assessment is biased. It is frequently hypothesized that participants in microfinance have characteristics which are difficult to observe in quantitative studies which account in part for their being or becoming better off through association with the MFI (Coleman, 1999).

This paper re-visits the evidence of the impact evaluation of SEWA Bank conducted by the United States Agency for International Development (USAID) in India in 1998 and 2000 to illustrate the broader challenges of measuring the impact of microfinance. In particular, the challenges of controlling for selection bias and the role of the unobservables in this context are discussed in depth. Existing panel data are subjected to propensity score matching (PSM) and panel data techniques which purport to eliminate selection bias in impact evaluations.[2]

The matching results are subjected to sensitivity analysis to assess their robustness; this is rather novel; sensitivity analysis of PSM was examined extensively by Rosenbaum (2002) and taken further by Ichino, Mealli and Nannicini (2006), and others, as discussed later in this paper. The few studies that have applied PSM to microfinance, e.g. Chemin (2008), Setboonsarng and Parpiev (2008), Arun, Imai and Sinha (2006) have not given sensitivity analysis much attention.

The analysis of the survey data is supported by direct observation; both the sensitivity analysis and the direct observation suggest that selection processes driven by unobservables strongly influence who becomes a participant in microfinance (and progression from saver to borrower) and cannot be fully controlled by econometric techniques. Further doubt is thrown on the impact claims by the sampling strategy of the control group for the original USAID study, which is not sufficiently described in the literature. Also, the panel design is problematic because it does not have a 'true' baseline which would allow a before and after comparison, with the control and treatments groups shown to be equivalent before joining SEWA Bank, since the treatment groups had already joined SEWA Bank well before the baseline period.

SEWA Bank members start as savers and the majority never progress to borrowing. The literature on the impact of savings on the well-being of the poor is scarce (see studies by Aportela, 1999; Ashraf, Karlan and Yin, 2006; and Devaney, 2006) since few MFIs offer savings products only. SEWA Bank is one of those MFIs that focuses on a savings approach and having more savers than borrowers (Chen and Snodgrass, 2001); thus a further selection process segregates SEWA Bank members into borrowers and lenders, in which there may well also be a role of unobservables. Thus this paper also contributes to the literature by conducting various sub-group comparisons between borrowers, savers and controls and with a potential to shed some light on the impact of savings versus credit. Finally, I draw conclusions as to what these findings imply for the reliability of the original impact estimates provided

[2] This is not the first attempt to re-examine the SEWA Bank study. Augsburg (2006) appears to be the first to re-investigate SEWA Bank's cross-section results. However, her study focuses on examining merely three household-level income-related outcome variables. She applies PSM and the differences-in-differences (DID) approach. Her results could not be fully replicated due to differences in the data re-construction; I contacted the author for her STATA do-files in order to comprehend her data analysis but she could not make them available. Her matching estimates were on average lower than the ones presented here in this study. I will not further refer to her study since this paper casts its net wider and re-analyses all of USAID's outcome variables on the household, enterprise and individual level.

by USAID. I find that while the results presented by USAID cannot be contradicted, there are doubts about the quality of the PSM impact estimates, and sensitivity analysis suggests significant over-estimation of impact due to the presence of unobservables. The USAID control group sampling procedure and the panel design leave questions of comparability with the treatment group unanswered. As a result, it is also not possible to reject the hypothesis that unobservable differences in characteristics account for some at least and possibly most of the observed impact. I draw conclusions as to appropriate impact evaluation procedures, and the need for qualification to even apparently robust estimates of impact.

This paper begins with by outlining the challenges of impact evaluations in the specific context of microfinance; this is followed by a brief description of SEWA Bank, its products, services and workings, drawing on my qualitative fieldwork during compilation of a repeat cross-section data set. Next, the research design and sampling procedures used by USAID and their data are described together with the estimation strategies; the analyses of USAID are replicated and PSM is applied to the data and sensitivity analysis conducted. Finally, the results are discussed in the light of my and other qualitative research on MFIs, before drawing conclusions.

2 Challenges of microfinance impact evaluations

Ultimately, microfinance impact studies are trying to find out how the lives of the poor would have turned out if microfinance had not been introduced. This is the problem of measuring the counterfactual which cannot be observed. Thus, every programme evaluation can only make an attempt at creating an estimate of such a counterfactual. Those estimates are then used to pinpoint the effect of the programme (Bryson, Dorsett and Purdon, 2002). The process of estimating counterfactuals commonly introduces biases which adversely affect impact evaluation results. Most impact assessors agree that future impact assessments should control for biases because ignoring them can greatly distort impact assessment results (Sebstad and Chen, 1996; Armendáriz de Aghion and Morduch, 2005; Pitt and Khandker, 1998; Coleman, 1999). Those are the challenges every impact evaluation has to grapple with and they are not unique to the context of microfinance impact evaluations.

Apart from the anecdotal evidence provided by qualitative studies (e.g. Todd, 1996), the majority of microfinance impact evaluations have applied a quasi-experimental design; well-known examples include Hulme and Mosley (1996) and Pitt and Khandker (1998), as well as Coleman (1999) who uses an innovative pipeline quasiexperiment to tackle selection bias. As in the experimental design, quasi-experiments identify a treatment and a control group. However, the treatment group already participates in a microfinance intervention, and the control group should be as identical as possible to the treatment group in terms of economic and social characteristics apart from the microfinance intervention (Hulme, 2000).There are a number of problems with using a quasi-experimental design, namely the identification of equivalent control groups and the challenge of overcoming selection bias.

In the case of microfinance, selection bias typically comes in the form of self-selection and non-random programme placement biases. According to Coleman (1999), self-selection bias refers to microfinance programme members that have self-selected into a programme. The decision to participate may have been influenced by certain unobservable characteristics such as entrepreneurial skills, organisational abilities and motivation which tend to increase the likelihood of individuals to self-select into a programme. Coleman (1999) further argues that prospective borrowers will not only have to make a decision on programme participation but they will also have to gain acceptance from incumbent borrowers, i.e. their peers, who have also self-selected into the programme. As a result, programme members will significantly differ from non-members in terms of motivation or wealth. In other words, programme participants usually self-select, meaning they choose to enter a microfinance programme in a non-random way or are selected by their peers. In addition, microfinance loan officers play a role in selecting borrowers. In other words, selection processes are driven by self-selection, selection by peers and by loan officers. Hence, impact studies need to address this problem because estimates obtained in the presence of selection bias will most likely be invalid. Moreover, programme placement can also be biased; MFIs assign new programme locations in non-random ways based, for example, on considerations for infrastructure or wealth (Hulme, 2000; Coleman, 1999). Some programmes, for example, are placed in areas which are easily accessible or the opposite could also be possible; programmes are placed in particular flood-prone areas (Pitt and Khandker, 1998).

Partly in response to critical reviews of evaluations using observational (qualitative and quantitative) data there has been a trend towards using experimental tools, i.e. conducting randomised control trials (RCTs) of many development interventions including microfinance (see Duflo and Kremer, 2003; Miguel and Kremer, 2004 for general RCTs and Banerjee et al, 2009; Karlan and Zinman, 2009 for microfinance RCTs). RCTs claim to resolve the issue of selection bias. However, they are vigorously debated and many microfinance interventions lack crucial characteristics necessary for valid RCTs. In brief, there are threats to internal and external validity caused by lack of double-blinding and/or the presence of Hawthorne and John Henry effects. For example, Hawthorne effects[3] refer to behavioural changes in the treatment group while John Henry effects[4] relate to behavioural changes in the control group. For example, individuals in the treatment group might positively change their behaviour for the duration of the study as they feel thankful for receiving treatment and as a response to being observed. The same behavioural changes might apply to members in the control group who might positively or in fact negatively alter their behaviour (Duflo, Glennerster and Kremer, 2007). Those weaknesses apply to microfinance RCTs as well and will adversely affect the reliability of the impact estimates obtained. Hence, critics of RCTs argue that there is a continuing role for observational methods (Deaton, 2009; Heckman and Urzua, 2009; Imbens, 2009 and Pritchett, 2009; see also Roodman and Morduch, 2009).

The analysis in this paper seeks to understand whether analysis of the original USAID observational data, supplemented by my own re-survey, can throw light in particular on the existence and effects of selection on unobservables. Before outlining the specifics of the research design, data, sampling procedure and estimation strategy, the next section introduces SEWA Bank and its microfinance programme.

3 The SEWA Bank context

SEWA Bank, a cooperative bank headquartered in the Indian city of Ahmedabad, is a sister organisation of the Self Employed Women's Association (SEWA). SEWA was established in 1972 as a trade union with the objective to organise self-employed women working in the informal sector. SEWA is not a mere trade union but a women's movement with its origins in Ghandian philosophy based mainly on principles of truth, non-violence and self-sufficiency (http://www.sewabank.com/aboutus-origin.htm). Various sister organisation grew out of the SEWA movement such as SEWA Bank which provides microfinance products, Vimo SEWA which provides insurance services, SEWA Academy, which is responsible for training and research, and a number of other organisations which

---

[3] In 1924 a series of experiments were conducted in the Hawthorne plant belonging to the Western Electric Company of Chicago. The aim of those experiments was to find out whether the productivity of workers could be improved with better lightening in the plant. Researchers found that workers increased their productivity irrespective of the lightening conditions which led to the conclusion that workers made an extra effort during their work precisely because of the knowledge of being observed (Duflo, Glennerster and Kremer, 2007; Levitt and List, 2009).

[4] John Henry effects refer to the "Ballad of John Henry" who was a rail worker and American folk hero. The ballad tells a tale of competition between rail workers and technical innovation which ultimately replaced rail workers. This tale can be related to the case of experimental design as discussed in Saretsky (1975).

offer a range of services to its female members (http://www.sewabank.com/aboutus-origin.htm).

Based on an initiative by Ela Bhatt, SEWA Bank was established in 1974 with the help of 4,000 SEWA members. The aim of SEWA Bank is to provide financial services such as savings and loan products to self-employed women. The bank has its base in urban Ahmedabad where it mainly operates individual savings and lending programmes but in the early 1990s it has also expanded into Gujarat's rural areas where it provides its services through self-help groups (SHGs) (http://www.sewabank.com/rural-activities.htm).

SEWA Bank emphasizes the provision of savings over credit (Chen and Snodgrass, 2001). That is illustrated by the following figures: as of fiscal year (FY) 2007, SEWA Bank had 163,187 clients out of which 143,806 were savers and the remainder of 20,011 were borrowers. SEWA Bank only targets women, not all of them are micro-entrepreneurs, i.e. women that sell goods and services on their own account, but many work as casual labourers or sub-contractors. Furthermore, SEWA Bank targets minorities, e.g. roughly 25% of its clients are Muslims and the remainder are from scheduled castes and tribes and other backward castes (this information is taken from SEWA Bank's internal management information system (MIS) which I had access to). To relate those figures to the overall population of Ahmedabad: according to the Census of India (2001), Ahmedabad has an overall population of close to 4.7 million out of which 82.1% are Hindus (out of which 13% are scheduled castes and 1% are scheduled tribes), 13% are Muslims and the remainder are Christians, Buddhists, Sikhs and Jains.

SEWA Bank offers a range of savings products such as current deposit, fixed term deposit and ordinary savings accounts as well as loan products. Loans can be secured which requires physical collateral such as jewellery or a savings account or unsecured which requires a guarantor as 'social' collateral. In addition, housing loans are offered as well as emergency loans which many clients use to pay for weddings, funerals or other consumption-related expenditure. As of FY 2007, approximately 49% of the loans disbursed were secured, 24% were unsecured and 27% were given as housing loans (this information is taken from SEWA Bank's internal MIS). Loan sizes vary from 5,000 Rupees to 50,000 Rupees. All loans are provided under an individual lending scheme, SEWA Bank does not operate any group lending schemes in its urban operations.

The loan application process works as follows: every potential SEWA Bank borrower is required to first open a savings account. SEWA Bank staff then monitors the savings behaviour of those potential borrowers, i.e. the size and regularity of their savings. Potential borrowers qualify for loans once they have regularly deposited money in the savings account for at least 6 months (Chen and Snodgrass, 2001). However, those rules are often relaxed. SEWA Bank loan officers, so-called Bank Saathi (as explained in Box 1), are living in the communities they service and are responsible for recommending future clients whose creditworthiness they then assess as well. If a Bank Saathi feels that a future client is bankable, then SEWA Bank commonly does not reject the loan application; this suggests that informal networks around those Bank Saathi drive the selection and loan approval processes as described in Box 1. In addition, if a future client can provide collateral, either physical or 'social', i.e. in the form of a guarantor, a loan is usually granted without the need of having to open a savings account.

Box 1: Social capital at SEWA Bank

The aim of the qualitative part of this study was to understand the selection processes from an ethnographic point of view. I wanted to understand how potential microfinance clients are recruited into SEWA Bank. To do this, I interviewed SEWA Bank staff and shadowed them in the field for several days in addition to interaction with them over design and fielding of a survey questionnaire.

The SEWA Bank staff I talked to explained that the Bank's 'recruitment' process works as follows: potential clients are generally recommended by people they know, i.e. family, friends or neighbours and are then referred to so-called Bank Saathi. The Bank Saathis are SEWA Bank's voice in the field; they are the first point of contact for the clients and function as advisors and mediators. In other words, they are the link between SEWA Bank and the clients. Saathi literally means companion. Bank Saathis are clients themselves and live in the same neighbourhoods as ordinary SEWA Bank clients. According to information from SEWA Bank, an individual can become a Bank Saathi when she has been saving and borrowing with SEWA Bank for several years, has displayed impeccable financial behaviour, is honest, trustworthy and good at managing relationships as well as has a certain social standing in the community. Bank Saathis are responsible for savings and loan collections but also for recommending and identifying future clients. Bank Saathis are paid on a commission basis, i.e. the more clients they 'recruit' the more they earn. Unfortunately, I could not obtain any further information on the nature of those payments and their size in proportion to the Bank Saathis income. Once a Bank Saathi has identified a potential client, a member of SEWA Bank staff visits the potential client to initiate the loan approval procedure. I shadowed SEWA Bank staff on some of those visits and was surprised by the informality of the loan approval process. It was more like an informal conversation with the potential client and was completed within ten minutes. These visits were often not documented and the loan was usually granted after such a visit – i.e. on the basis of information produced in this process and trust in the network that led to it and without further investigations. I formed the impression that SEWA Bank staff were usually inclined to follow the recommendations of the Bank Saathi, mainly because they (the Bank Saathi) are presumed to know the potential clients, their family, friends and neighbours as they are living in the same communities (and can thus observe those variables that are unobservable to formal data production techniques). As a result, there is a lot of room for the Bank Saathis to abuse their information advantage and power as suggested by Ito (2003). My enumerators saw evidence of such abuse and observed that some Bank Saathi demanded 10% to 15% of the loan amount granted as an additional commission from the client. It appears that this informal screening or selection process does indeed play a role in explaining microfinance participation. However, the econometric tools commonly applied in the context of impact evaluations do not seem to be able to control for those unobservables that seem to be driving the screening or selection process. Based on my observations I conclude that an ethnographic approach could possibly be more appropriate for providing further insights and this would be a recommendation for future research in this area.

SEWA Bank is unique in many ways. It is not only one of India's oldest and most established microfinance providers with a strong ideological base rooted in Ghandian traditions using struggle and development as a strategy to strengthen their members position in society but SEWA Bank also prefers to work in a cooperative structure and extend its financial services to individuals without the need for group formation. This cooperative structure allows SEWA Bank to focus on a savings approach. Typically, microfinance in India is offered by microfinance-NGOs (MF-NGOs) which are registered as non-profit organisations. The registration as a non-profit organisation limits their scope for providing financial services. The Reserve Bank of India (RBI), for example, prohibits all non-profit organisations from taking savings (Fisher and Sriram, 2002). Strictly speaking the MF-NGOs that do take savings are operating illegally. Thus, many organisations do refrain from mobilising savings because their organisational set-up simply does not allow it. Recently for-profit MFIs have emerged, so-called Non-Banking Financial Companies (NBFCs); examples include organisations like BASIX and SHARE. However, NBFCs, although regulated by the RBI, are also not allowed to take savings (Fisher and Sriram, 2002). Ghate (2007) as well as many practitioners (Karlan and Morduch, 2009) argue that this is a major drawback because the poor need savings more than credit as the next section will elaborate in more detail.

4 Impact of savings

As mentioned in section 3, SEWA Bank emphasizes savings over credit and had on average seven times more savers than borrowers as of FY 2007. SEWA Bank views credit merely as a complementary tool to savings, and hence I assess the impact of credit as well as savings to account for SEWA Bank's distinctive approach. The objective of this section is to briefly introduce the savings literature and to review some of the key studies that assessed the impact of savings.

Policy makers assumed for a long time that the poor are too poor to save and hence savings mobilisation was low on the agenda of many governments. This assumption has been questioned by Adams (1978) and von Pischke (1983) and further by Rutherford (2001) and Collins et al (2009). Rutherford (2001) claims that the poor have the capacity to save and traditionally used rotating savings and credit associations (ROSCAs) or other informal mechanisms to satisfy their savings needs. Indeed, savings are crucial for accumulating assets which in turn are used to finance future investments and consumption (von Pischke, 1983). Following Keynes (1936) and Browning and Lusardi (1996), Karlan and Morduch (2009) explain that individuals have various motives that encourage them to save such as "precautionary, life-cycle (to provide for anticipated needs), intertemporal substitution (to enjoy interest), improvement (to enjoy increasing expenditure), independence, enterprise, bequest, avarice, and downpayment" (p. 39). It is beyond the scope of this thesis to discuss the motivations of individuals to save in detail and hence I will refer the interested reader to a comprehensive review of the savings behaviour of individuals in developing countries which is provided by Rosenzweig (2001).

Studies evaluating the impact of savings are scarce, some notable exceptions include the studies by Aportela (1999), Ashraf, Karlan and Yin (2006) and Devaney (2006) for a review. Dupas and Robinson (2009) conducted the first and so far only randomised control trial (RCT) assessing the impact of savings products. Most MFIs focus on providing credit as well as savings and a range of other services which makes it rather challenging to disentangle the impact of savings from all the other products and services that clients use at the same time, e.g. Burgess and Pande (2005) showed that financial access can reduce poverty but they could not separate the impact of savings from the impact of credit.

Devaney (2006) reviewed eight impact studies - including the SEWA Bank study discussed in this paper - that focused on the impact of savings on the poor. The aim of most of those studies was to provide evidence that the poor have the capacity to save in the first place and to justify the need for savings products in addition to loan products. Moreover, the majority of those studies reviewed by Devaney (2006) investigated the impact of a particular savings product on the savings rate of the poor and found that savings had indeed a positive impact on the households' savings rate. The most recent study on savings impact by Dupas and Robinson (2009) based on a field experiment in Kenya testing for the existence of savings constraints concluded that access to savings has positive impacts on income and productive investments. Furthermore, Devaney (2006) claims that borrowers are more likely to save than non-borrowers. However, most of those studies mentioned here did not compare the impact of saving versus the impact of borrowing versus not saving or borrowing at all. Only the study by Rogg (2000) and the SEWA Bank study under discussion in this paper are exceptions in this regard. It can be concluded that the literature in particular on the impact of saving versus borrowing is still rather underdeveloped, but is generally positive.

In the Indian context, the RBI frequently turns a blind eye to the MFIs illegally mobilising savings because it recognised the importance of microfinance and savings in particular (Basu, 2006). A solution to the savings dilemma, i.e. the fact that MFIs are not officially allowed to mobilise savings but unofficially do so at times, is to register a MFI as a mutual benefit organisation, which allows it to be classified as a cooperative. SEWA Bank is one of the few microfinance providers that is registered as an urban cooperative bank which means that savings can legally be mobilised (Fisher and Sriram, 2002). This organisational set-up is suitable for SEWA Bank's activities since it allows its clients to save rather than access only credit. This paper returns to the issue of savings in section 7.6 when comparing the impact estimates of the various sub-groups, i.e. borrowers, savers and controls. The question is whether the data support the view that a savings approach - as advocated by Ghate (2007) and as implemented by SEWA Bank - is justified and desirable. After this brief introduction to SEWA Bank, the next section presents the research design and describes the data.

5 Research design and description of data

The study discussed here is one of three longitudinal USAID microfinance impact evaluations that were carried out between 1997 and 2000 on Mibanco in Peru, Zambuko Trust in Zimbabwe and SEWA Bank in India. All three studies share a similar research design and aim to examine the socio-economic impact of microfinance participation (Snodgrass and Sebstad, 2002). For the purpose of this study, the impact evaluation conducted on SEWA Bank is discussed in more detail in this section.

The original SEWA Bank (henceforth USAID) study assesses the impact of SEWA Bank's microfinance services on urban client households (Chen and Snodgrass, 2001). It examines hypotheses at the household, enterprise and individual level. The study hypothesized that microfinance participation at the household level leads to an increase in household income, more diversified income sources, housing improvements, an increase in household assets, better education of the household's children, an increase in food expenditure and improved mechanisms for coping with shocks. At the enterprise level, microfinance participation leads to an increase in informal sector income, an increase in revenues and fixed assets, employment generation as well as better transactional relationships. At the individual level, microfinance clients might gain more control over the household's resources and incomes, increase their self-esteem and self-confidence and increase personal savings and improve their ability to deal with the future (Chen and Snodgrass, 2001). Table 1 outlines the details of the hypotheses tested and the corresponding impact variables.

Table 1: Hypotheses and impact variables of USAID study

Household level	
Hypotheses	**Impact variable**
H1: increase of household income	• Total annual household income • Household income per capita
H2: more diversified income sources	• Inverse Simpson's index
H3a: housing improvements	• Expenditure on housing improvements and repairs in terms of material and labour
H3b: increase of household assets	• Expenditure on household assets, e.g. appliances, vehicles, jewellery
H4: better education of the household's children	• Net enrolment ratios
H5: increased food expenditure	• Per capita expenditure per day for food and beverages
H6: better coping with shocks	• Mechanisms used for dealing with shocks
Enterprise level	
E1: increase of informal sector income	• Microenterprise income of previous month from household head and respondent
E2: increase of revenues	• Gross sales revenue of previous month
E3: increase of fixed assets	• Value of all fixed assets used in microenterprise
E4: more employment generation	• Hours worked in previous week • Days worked in previous month
E5: better transactional relationships	• Types of suppliers and customers
Individual level	
I1: client gains more control of the households resources and income	• Who took decision to take last loan? • Who took decision how to spend loan amount? • Who took decision how to spend income?
I2: increase in self-esteem and self-confidence	• Respondent's feelings with regard to her contribution to household • Is this contribution respected by other household members?
I3: increase in personal savings	• Existence of personal savings
I4: better ability to deal with the future	• Respondent's feelings with regard to preparedness to deal with future • How does respondent prepare herself to deal with future?

Source: Chen and Snodgrass (2001, p. 58).

In order to test the hypotheses outlined in Table 1, researchers collected baseline data on 900 women from low-income households across ten wards in Ahmedabad. The sampling criterion required the selection of women who were above 18 years and economically active. An economically active person is defined as somebody who engages in informal economic activities in the home, on the street or on business premises and who is either self-employed, a dependent producer or a wage worker (on an irregular basis without written contracts and/or fixed wages) (Chen and Snodgrass, 2001). Out of the 900 sample women, 600 were SEWA Bank clients – consisting of borrowers and savers - and 300 non-clients. The sampling procedure was based on a three-step process (Chen and Snodgrass, 2001). First, a geographical area was selected; Ahmedabad is split into 43 wards and USAID limited its survey to 10 of those 43 wards due to budget constraints. The sample was drawn from the following ten wards in Ahmedabad: Behrampura, Jamalpur, Bapunagar, Rakhial, Asarwa, Khadia, Amraiwadi, Saraspur, Raikhad and Dudheshwar. The wards were selected based on the number of SEWA Bank clients residing in them. Almost half of all current clients live in those 10 wards (Chen and Snodgrass, 2001). Next, a random sample of borrowers and of savers was selected from a list provided by SEWA Bank which contained all its current borrowers and savers as of FY 1997, listed by ward. Savers should have made at least one deposit in a SEWA Bank savings account during FY 1997. Moreover, savers should not have taken out any loans in FY 1997. Replacements were made when the respondent could not be located, was not economically active, e.g. not self-employed anymore, or did not want to participate. Also, replacements were needed when respondents from the sample of current savers were not actively saving anymore or had taken out loans in FY 1998.

The rationale for sampling borrowers as well as savers is explained by SEWA Bank's emphasis of savings over credit - at the time of the USAID study there were ten savers for every borrower - hence USAID decided to gather a separate sample of savers. Chen and Snodgrass (1999) explain

> "that those clients who are savers only will benefit from having a secure place to deposit their savings. Since all borrowers have to save, it is hypothesized that there will be greater impact on those who borrow as well as save" (p. 16).

Finally a non-client sample was chosen. USAID carried out a pre-survey "in the neighbourhood [*Author's note: it is not clear whether neighbourhood and ward are used interchangeably or whether neighbourhood refers to something else*] of each of the 300 sample borrowers to identify 50 households in which there were economically active women over age 18 who were not SEWA members" (Chen and Snodgrass, 2001, p. 53). From those 15,000 households a random sample of 300 non-clients was drawn.

Rosenbaum (2002) argues that the sampling of an appropriate control group is crucial in observational studies and in view of this the robustness of the USAID sampling procedure of the control group is explored. Chen and Snodgrass (2001) argue that the neighbourhoods where most of SEWA Bank's clients reside are reasonably homogenous in terms of caste, occupation and class (p. 53) and hence the control group is relatively similar to the treatment group. However, if the households in the control group are so similar, then why are they not clients of SEWA Bank? This points towards a selection process that is driven by unobservable characteristics which account for why otherwise apparently eligible households did not belong to SEWA Bank, and hence the control group sampling of USAID does not convince. Chen and Snodgrass (2001) admit that SEWA Bank members

> "are not chosen at random but are in fact purposefully selected from a larger population, both by themselves and by SEWA Bank. A woman must first self-select by deciding to open a savings account and later to apply for a loan. Once she does so, SEWA Bank decides whether to provide her with the financial service in question" (p. 60).

The first round (henceforth round 1) of the USAID survey was conducted in January 1998 and a follow-up round (henceforth round 2) was then collected in January 2000. Between survey rounds, a rate of attrition of approximately 11 percent was observed, resulting in a final sample of 798 respondents (Chen and Snodgrass, 2001, p. 56). In addition to the two surveys, twelve case studies of SEWA Bank borrowers were conducted with the objective to provide a better understanding of the issues that SEWA Bank borrowers commonly have to deal with on a daily basis and how microfinance has helped them in the process (Chen and Snodgrass, 2001). Table 2 provides descriptive statistics for respondents in round 1 and 2 for illustrative purposes.

Table 2: Descriptive statistics of female research respondents

		Borrowers		Savers		Controls	
Data collection round		R1	R2	R1	R2	R1	R2
Sample size		264	264	260	260	262	262
Mean age (years)		37.81	40.28	34.55	36.88	35.36	37.51
Martial status in %	Married	89.77	88.64	87.31	85	80.92	79.39
	Never married	1.89	1.89	5.77	4.62	5.73	5.73
	Divorced	0.76	0.76	1.15	0.38	1.15	0.38
	Deserted	0.38	0	1.54	1.15	1.15	1.53
	Widowed	7.2	8.71	4.23	8.85	11.07	12.98
Religion in %	Hindu	72.35	72.73	76.54	76.92	77.1	77.48
	Muslim	27.27	26.52	23.46	23.08	22.52	22.14
	Other	0.38	0.76	0	0	0.38	0.38
Caste in %	Upper caste	15.15	14.39	16.15	15.77	22.9	23.66
	Backward caste	45.45	46.97	40.77	43.85	39.31	40.46
	Scheduled caste	29.92	31.82	35.38	35.77	29.77	32.06
	Scheduled tribe	9.09	6.44	7.31	4.62	8.02	3.82
	No response	0	0	0	0	0	0
Education	Never attended school in %	39.77	40.15	40	41.92	40.84	44.66
	Mean highest grade completed	3.9	3.9	4.3	4.24	4.2	4.01

Source: Author's own calculations.

The data collected from both survey rounds was subjected to Analysis of Variance (ANOVA) in order to examine cross-section differences and Analysis of Covariance (ANCOVA) to evaluate whether any personal characteristics possibly influenced any impact variables. Chen and Snodgrass (2001) argue that ANCOVA would reduce selection bias to a certain degree. In addition, gain score analysis was employed to estimate the degree of change over time between treatment and control groups and to assess whether such changes were significant. The findings of the USAID study provide evidence that microfinance leads to changes at the household level, i.e. higher household income in terms of total income and per capita income was observed. In addition, minor positive impacts could be observed on income diversification, food expenditure and the ability to cope with shocks. However, the evidence was rather mixed. Impact at the enterprise and the individual levels were negligible. Chen and Snodgrass (2001) admit that measuring impacts at the enterprise and individual level were rather challenging because SEWA Bank clients are not classical micro-entrepreneurs. Most clients do not have micro-enterprises but are dependent sub-contractors or labourers, thus do not require micro-enterprise capital. SEWA Bank provides loans for a range of purpose, e.g. business, housing improvements/repairs, repayment of other debts and consumption but without a particular focus on micro-enterprise development.

Fungibility of money is a central problem in the context of microfinance impact evaluations and notoriously difficult to control for (Hulme, 2000). Money is considered to be fungible within the household, i.e. once a loan has been taken out by the borrower, it is difficult to track in which way the loan has actually been used (Ledgerwood, 1999). Based on the findings of the USAID study, it appears that measuring impact separately at the enterprise and individual level does not lead to particularly satisfactory results, and this is likely at least partly to be because of fungibility.

As in the case of the USAID study, the majority of impact studies examine the impact of microfinance at multiple levels, i.e. at the household, enterprise and individual level; see Hulme and Mosley (1996), Sebstad et al (1995) and Gaile and Foster (1996) for a comprehensive overview of studies up to the mid 1990s. However, examining the impact at multiple levels requires sufficient funds and time. Moreover, solely looking at the individual, enterprise or community level has a number of disadvantages (see Table 3 for details). In particular the issue of fungibility (as mentioned earlier) has to be considered when assessing impact at the enterprise level. Hulme (2000) argues that "…for all studies except those that focus exclusively on 'the enterprise,' [then] a concern about fungibility may be irrelevant" (p. 85). He further argues that the most promising way to measure impact of microfinance appears to be at the household and institutional level. However, institutional level data are not available in this case. Hence, based on Hulme (2000) and after carefully examining Table 3, I conclude that re-examining the household level hypotheses of the USAID study appears to be the way forward due to issues of fungibility and difficulties of breaking down household level impacts to the individual level.

Table 3: Units or levels of evaluation and their advantages and disadvantages

Unit	Advantages	Disadvantages
Individual	• Easily defined and identified	• Most interventions have impacts beyond the individual • Difficulties of disaggregating group impacts on "relations"
Enterprise	• Availability of analytical tools (profitability, return on investments, etc.)	• Definition and identification is difficult in microenterprises • Much microfinance is used for other enterprises and/or consumption • Links between enterprise performance and livelihoods need careful validation
Household	• Relatively easy defined and identified • Permits an appreciation of livelihood impacts • Permits an appreciation of interlink-ages of different enterprises and consumption	• Sometimes exact membership difficult to gauge • The assumption that what is good for a household in aggregate is good for all of its members individually is often invalid
Community	• Permits major externalities of interventions to be captured	• Quantitative data is difficult to gather • Definition of its boundary is arbitrary
Institutional Impacts	• Availability of data • Availability of analytical tools (profitability, Subsidy Dependency Indices (SDIs), transaction costs)	• How valid are inferences about the outcomes produced by institutional activity?
Household Economic Portfolio (i.e. household, enterprise, individual and community)	• Comprehensive coverage of impacts • Appreciation of linkages between different units	• Complexity • High Costs • Demands sophisticated analytical skills • Time consuming

Source: Hulme (2000, p. 83).

As mentioned earlier, the aim of this study is to re-visit the evidence of microfinance impact evaluations; hence the USAID panel data set has been subjected to more advanced econometric techniques, i.e. PSM to account for selection bias (see claims by Dehejia and Wahba, 1999 and 2002). In addition, a new cross-section data set (henceforth Round 3) was produced, with the aim of exploring the potential of indicators of social capital to illuminate the role of the unobservables and to compare the USAID panel with Round 3 to get a clearer picture on short-term versus long-term impacts of microfinance. However, the Round 3 data have shortcomings. A further round of the panel could not be collected as neither the original panel sample SEWA Bank members not the controls could be identified. The sample size of 220 households is rather small because of budget and time constraints. Hence, the Round 3 data comparability to the USAID data is limited. Another point is that PSM requires rich and high quality data sets (Smith and Todd, 2005), but there are many missing data in Round 3, in particular in the social capital and the housing improvements section, which limits its explanatory power; hence Round 3 quantitative data are not further discussed in this paper.[5] To conclude, it appears that low-cost and small surveys such as the Round 3 survey do not necessarily add value and do not provide accurate impact estimates.

6 Estimation strategy

This study replicates the USAID analysis and subjects these data to PSM to control for selection bias in the hope of providing more reliable impact estimates. PSM matches participants to non-participants on the basis of observable characteristics and compares outcomes between the treatment sample and the sample of matches (Caliendo and Kopeinig, 2005 and 2008; Rosenbaum and Silber, 2001). The underlying assumption is that there is no selection bias due to unobservable characteristics, though, whether this assumption holds is questionable; this is examined using sensitivity analysis of the PSM results which can suggest the presence and likely size of the effect of selection on unobservables (Rosenbaum, 2002). We have seen earlier that in the case of microfinance unobservables are very likely to be present.

To begin with, the empirical model is outlined. As mentioned earlier, the USAID study collected data on three sub-samples: borrowers, savers and controls. The

[5] The quantitative results of Round 3 can be obtained from the author upon request. Qualitative information produced during the fieldwork are used.

objective is to assess the socio-economic impact of microfinance participation. Consider the following empirical specification; i stands for household in ward j:

$$(1) \qquad y_{ij} = C_{ij}\delta + X_{ij}\alpha + V_j\beta + \varepsilon_{ij}$$

<u>Where:</u>

y_{ij} = outcome on which impact is measured

C_{ij} = level of participation in microfinance, i.e. a membership dummy variable

δ = effect of the microfinance programme, main parameter of interest

X_{ij} = vector of household level characteristics

V_j = vector of ward level characteristics

α, β = unknown parameters

ε_{ij} = error term representing unmeasured household and ward characteristics that can influence outcomes

The characteristics of participants, i.e. borrowers and savers, are examined separately for round 1 and for round 2 using a logit model (Table 4). A treatment dummy denoting microfinance participation was created containing borrowers and savers to represent participants, i.e. C_{ij} as expressed in equation (1). This dummy is used as a dependent variable and assumes a value of 1 if an individual has self-selected into microfinance and a value of 0 if otherwise.

Table 4: Logit regression of probability of microfinance participation, without sampling weights[6]

Independent variables	Round 1	Round 2
Age household head	-0.040*** *0.000*	-0.002 *0.823*
Age respondent	0.019* *0.057*	0.010 *0.341*
Highest grade completed male	0.083*** *0.003*	0.001 *0.956*
Respondent married (yes=1)	0.789*** *0.004*	0.760*** *0.008*
Muslim (yes=1)	0.510** *0.013*	0.434** *0.030*
Upper Caste (yes=1)	-0.666*** *0.003*	-0.682*** *0.001*
Household size	0.011 *0.920*	-0.074 *0.422*
Nuclear household (yes=1)	-0.453** *0.042*	-0.293 *0.174*
Non-SEWA savings	-0.000 *0.190*	-0.000 *0.695*
Constant	-0.360 *0.713*	-0.413 *0.679*
Number of observations	768	785
Pseudo R-squared	0.059	0.032

Source: Author's own calculations.
Notes: p-values in italics. *statistically significant at 10%, **statistically significant at 5%, ***statistically significant at 1%. Please also note that the following control variables were included in the logit model: age, age squared, highest grade completed household head, highest grade completed respondent, sex household head (male=1), number of adult male in household, number of household members aged 0-14, subnuclear household (yes=1); all insignificant.

Table 4 presents the logit regression estimating the probability of microfinance participation. The logit model is required to predict the propensity scores so that the matching procedure can be implemented – section 7 outlines this in more depth. The

[6] Since SEWA Bank members are more likely to be savers than borrowers, i.e. as mentioned earlier there were ten savers for every borrower in rounds 1 and 2, there is a case for using appropriate weights. A separate set of logit regressions across rounds 1 and 2 were computed adjusted for sampling weights. However, the use of sampling weights led to very minor changes, i.e. slightly lower pseudo R-squared values and lower significance levels for few coefficients when sampling weights were applied. The results reported in this table do not consider sampling weights.

results presented in Table 4 show that the main variables associated with membership that are statistically significant in rounds 1 and 2 are being married, Muslim and upper caste. In addition, age of household head, age of respondent, highest grade completed male and nuclear household are significant in round 1. Few covariates are statistically significant and the values for the pseudo R-squared across both rounds are rather low which indicates that the model has limited explanatory power.[7]

7 Results

The findings with respect to selected hypotheses are presented in this section. The individual and enterprise level hypotheses of the USAID study led to mixed and rather insignificant results. As argued earlier, measuring impact at these levels is unsatisfactory, and the household level is the most promising way to obtain meaningful impact estimates. Hence, this section focuses on selected household level results only, i.e. on income, housing expenditure and children's education.[8]

Firstly, the selected household level results of the USAID study are replicated. Replication is an important step in validating results (Hamermesh, 2007). Hence, the USAID data[9] were subjected to ANOVA and ANCOVA. My replication closely reproduced the USAID study results and is thus not discussed further.

Next, PSM is employed on the USAID data to gauge whether more advanced econometric techniques than ANOVA and ANCOVA, which claim to account for selection bias (Chen and Snodgrass, 2001), would produce different results. The original household level results of the USAID study are compared with the results obtained when PSM was applied using 5-nearest neighbour matching and kernel matching with a bandwidth of 0.01 (Table 5). Participants, i.e. borrowers and savers together, versus controls are presented first. Further sub-group comparisons are presented later in this paper.

[7] I experimented with the logit model and tried various other control variables with the objective to enhance the explanatory power of the model but to no avail. The low explanatory power of the model has implications for the reliability of my PSM results; this is further investigated in section 7.3 where sensitivity analysis is introduced.

[8] The detailed and re-analysed results of all household level as well as individual and enterprise level hypotheses of the USAID study can be obtained from the author upon request.

[9] The data sets of all three USAID studies can be downloaded here: http://www.microlinks.org/ev_en.php?ID=4678_201&ID2=DO_TOPIC

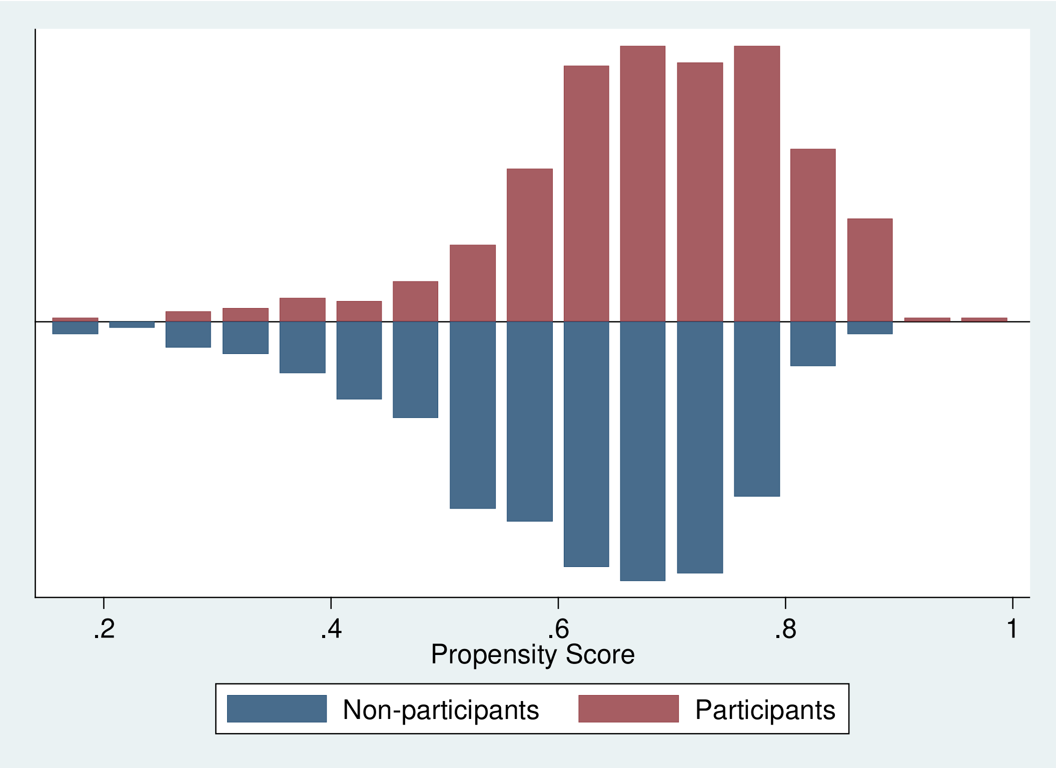
Before discussing the results presented in Table 5 and Table 8 in more detail, a few remarks with regard to the implementation of PSM are required. The basic idea of matching is to compare a participant with one or more non-participants who are similar in terms of a set of observed covariates X (Caliendo and Kopeinig, 2005 and 2008). This requires predicting propensity scores for each individual, i.e. participants as well as non-participants using a logit or a probit model. I used the logit model presented in Table 4 to predict those propensity scores. Then, before implementing the actual matching process, I examined whether the propensity scores I had obtained for participants and non-participants fulfil the common support assumption. Caliendo and Hujer (2005) express the common support assumption as follows:

$$(2) \qquad 0 < \Pr(D = 1 \mid X) < 1$$

This assumption indicates whether treatment and control groups provide equal support of X (Caliendo, 2006). This can be investigated graphically. Figure 1 presents the distribution of the propensity score for participants as well as non-participants; it shows that each participant with a certain propensity score has a corresponding non-participant. In other words, if the propensity scores for participants and non-participants overlap reasonably well, then the common support assumption is satisfied and it is recommended comparing those two groups. Next, the differences in the outcome variables for participants and their matched non-participants are calculated (Morgan and Harding, 2006) as presented in Table 5 and Table 8. Furthermore, I used t-tests[10] before and after matching for all results presented in Table 5 and Table 8 to examine the differences of the mean values for each covariate X across treatment and control groups. In addition, those t-tests calculated a 'bias' defined as the mean value of the treatment group and the (matched/unmatched) control group divided by the square root of the average sample variance in the treatment group and the (matched/unmatched) control group. The t-tests I employed indicate that the differences between treatment and control groups as well as the 'bias' were reduced considerably in most cases, hence the matching process was successful in generating a control group that was reasonably similar to the treatment group. Therefore, the use of PSM is justified in this case. As a consequence, I conclude that the balancing properties of the propensity scores were satisfied in all cases.

[10] The STATA command pstest was used.

Figure 1: Distribution of propensity scores for participants and non-participants



Source: Author's own calculations.

Table 5: PSM impact estimates of selected household level results – microfinance participants versus controls; without sampling weights[11]

	Round 1	**Round 2**
Total household income per annum in Rupees		
USAID	10,090***	15,302***
PSM - 5 nearest neighbour matching	8,944***	14,635***
PSM - kernel matching, bandwidth 0.01	8,638***	13,786***
Total household income per annum per capita in Rupees		
USAID	2,063***	2,685***
PSM - 5 nearest neighbour matching	2,019***	2,486***
PSM - kernel matching, bandwidth 0.01	1,913***	2,537***
Expenditure for housing improvements in Rupees		
USAID	3,748***	5,871
PSM - 5 nearest neighbour matching	3,701***	6,546***
PSM - kernel matching, bandwidth 0.01	3,484***	6,504***
School enrolment for girls aged 5 to 10		
USAID	-0.020	-0.005
PSM - 5 nearest neighbour matching	0.011	0.052
PSM - kernel matching, bandwidth 0.01	0.010	0.028
School enrolment for boys aged 5 to 10		
USAID	0.065	0.005
PSM - 5 nearest neighbour matching	-0.027	0.021
PSM - kernel matching, bandwidth 0.01	-0.007	-0.004
School enrolment for girls aged 11 to 17		
USAID	0.015	-0.015
PSM - 5 nearest neighbour matching	0.028	0.012
PSM - kernel matching, bandwidth 0.01	0.006	0.009
School enrolment for boys aged 11 to 17		
USAID	-0.075	-0.020***
PSM - 5 nearest neighbour matching	-0.025	-0.012
PSM - kernel matching, bandwidth 0.01	-0.045	-0.019

Source: Author's own calculations.
Notes: *statistically significant at 10%, **statistically significant at 5%, ***statistically significant at 1%. The results in this table refer to the differences in the mean values between matched samples; they were obtained using the STATA command psmatch2. I also ran the STATA command pscore with the

[11] With reference to footnote 6, the literature is unclear on how to accommodate sampling weights in the context of matching, hence the analysis was run with and without weights providing conclusive results. It is beyond the scope of this paper to explore this further but the interested reader is welcome to contact the author for more information.

objective to cross-check the psmatch2 results across the various matching algorithms. The results I obtained from the different STATA routines displayed minor differences in terms of the size of coefficient and the level of significance. Morgan and Winship (2007) argue that matching results can vary depending on the matching algorithm and PSM routine applied. Results are bootstrapped.

7.1 Income[12]

The results for income per annum[13], income per annum per capita[14] and expenditure for housing improvements are positive and statistically significant across rounds 1 and 2. Those results reflect the differences between participants and non-participants. For example, according to the USAID round 1 result; income per annum was higher by 10,090 Rupees for microfinance participants than for non-participants whereas the PSM results applying 5-nearest neighbour matching show that income per annum increased by 8,944 Rupees for participants. Similarly, the PSM results applying kernel matching with a bandwidth of 0.01 display an increase in income per annum by 8,638 Rupees for microfinance participants. It can be seen that the degree of impact depends on the econometric technique applied but even when the same technique is applied, i.e. PSM, the impact estimates still vary - though not substantively so - because of the different matching algorithms applied. For example, kernel matching estimates with a bandwidth of 0.01 for income per annum and for income per annum per capita are lower than the respective 5-nearest neighbour matching estimates for rounds 1 and 2. Also, when the bandwidth increases in the case of kernel matching, the impact estimate tends to increase as well. However,

[12] All income figures throughout this study have been deflated to January 1998 prices by using a deflator of 1.156 – as mentioned in the USAID study. This was the value of the Consumer Price Index (CPI-IW) for Ahmedabad in January 2000, expressed on a base of January 1998.

[13] Income per annum is confounded with household size and hence an unreliable measure of outcome. However, since USAID assesses the impact on income per annum and since I am comparing their results with mine, I will continue to report income per annum results throughout the paper. Nonetheless, the reader should treat those results with caution.

[14] An additional calculation not reported here was completed for total household income per annum per capita across rounds 1 and 2 as presented in Table 5; I made adjustments using an equivalence scale. Equivalence scales commonly allow the comparison of per capita income of households of various sizes and compositions on an equal basis. A range of equivalence scales exist and choosing one is a rather arbitrary process. The following equivalence scale adjusting for the various household sizes and compositions is used here: $(A + PK)^F$; where A = number of adults ≥ 18, K = number of children < 18, P = 0.7 which is the recommended percent value indicating how much each child contributes to the households consumption relative to the adults, and F = 0.65 – 0.75, a factor that accounts for economies of scales (Source: http://www.irp.wisc.edu/research/method/oakvos.htm). The application of this formula led to a minor increase in the size of the coefficient of per capita income per annum but the significance level remained the same. Since there is no clear recommendation as to which one of the equivalence scales to use and their application is debated, I decided to report total household income per annum per capita without making any adjustments.

these observations are not common to all outcome variables or rounds as Table 5 clearly demonstrates.

Overall, in the case of income per annum, income per capita and expenditure for housing improvements across round 1 and 2, the general trend and the statistical significance are similar across USAID and PSM results. In other words, the results obtained by applying PSM appear to support the original findings of USAID with the exception that the USAID results for expenditure for housing improvements in round 2 were not statistically significant while the PSM results are statistically significant.

7.2 School enrolment

The PSM results for school enrolment for girls aged 5 to 10 across rounds 1 and 2 display a positive trend, i.e. participants do better than non-participants, but none of the results are statistically significant. The respective USAID results are negative but also insignificant. The results for school enrolment for girls aged 11 to 17 across round 1 and 2 are equally meaningless, USAID argues that there is an increase in school enrolment in round 1 but the impact estimate is suddenly negative in round 2. The PSM results are all positive but insignificant.

The picture does not change dramatically when looking at school enrolment figures for boys aged 5 to 10 and 11 to 17. According to the PSM results across all three rounds, microfinance has negative impacts on the school enrolment of boys aged 5 to 10 with the exception of one value in round 2 obtained by applying 5-nearest neighbour matching which implies a negligible positive impact. The USAID results, on the other hand, indicate a positive impact but none of the results are statistically significant. Chen and Snodgrass (2001) argue that most boys in the age group 5 to 10 are in fact already enrolled in school irrespective of microfinance participation. All enrolment figures for boys aged 11 to 17 are negative across USAID and PSM results for round 1 and 2 with one figure being statistically significant. Hence, it can be concluded that microfinance participation does not seem to have any significant impact on children's education.

Overall, the most notable result is that there seems to be a positive impact on total income per annum and per capita as well as on expenditure for housing improvements. Before applying panel methods, the quality of the matching results needs to be assessed using sensitivity analysis.

7.3 Sensitivity analysis

The impact evaluation of SEWA Bank needs to answer the question whether the apparent effect of membership compared to the control group is due to the saving and borrowing enabled by membership of SEWA Bank or some unobserved characteristic of members compared to the control group, such as entrepreneurial abilities, access to social networks, etc. PSM allows control for observable characteristics included in the propensity score on which members and controls are matched, but it cannot control for unobservables (Caliendo and Kopeinig, 2005 and 2008; Rosenbaum and Silber, 2001). Rosenbaum (2002) developed the "conceptual advance" (ibid, p. 106) of Cornfield et al (1959) that the robustness of the estimate of the difference in outcome between treatment and control groups (the impact estimate) could be assessed by asking what magnitude of selection on unobservables (hidden bias) one would need in order to explain away the observed impact, thus: "[I]f the association [*Author's note: between treatment and outcome*] is strong, the hidden bias needed to explain it is large" (Rosenbaum, 2002, p. 106). In the context of death from lung cancer for smokers and non-smokers Cornfield et al (1959) suggested that if the ratio of the likelihood of death from lung cancer for smokers to the likelihood of death from lung cancer for non-smokers was high then a similar high ratio for the unobserved characteristic(s) would be required to make this unobserved characteristic the true cause of the higher prevalence of death from lung cancer by smokers.

Rosenbaum (2010) explains that "a sensitivity analysis in an observational study asks how the conclusions of the study might change if people who looked comparable were actually somewhat different…" (p. 367). In other words, the objective of sensitivity analysis is to explore whether the matching estimates are robust to selection on unobservables (Rosenbaum, 2002). Ichino, Mealli and Nannicini (2006) argue that "sensitivity analysis should always accompany the presentation of matching estimates" (p. 19).

Rosenbaum (2002) invites us to imagine a number Γ (gamma) (≥ 1) which captures the required degree of association, of an unobserved characteristic with the treatment, for it (the unobserved characteristic) to explain the observed impact. Γ is the ratio of the odds[15] that the treated have this unobserved characteristic to the odds that the controls have this characteristic.[16]

---

[15]Odds, which are widely used in assessing probabilistic outcomes, are derived from probabilities ($0 \le \pi_i \le 1$) by the following formula: $\pi_i/(1-\pi_i)$.

[16] Suppose two individuals j & k who are closely matched on observables so that $x_j = x_k$, but for whom p_j not equal to p_k - i.e. probability of being selected into SEWA Bank is not the same despite being equivalent on observables. The probability of being selected can be expressed as an odds ratio (the odds of probability of j/k (p_j/p_k) being selected $p_j/(1-p_j)$ or $p_k/(1-p_k)$). Then imagine there is a number Γ (gamma) such that $1/\Gamma \le \{p_j(1-p_k)\}/\{p_k(1-p_j)\} \le \Gamma$, then if $\Gamma = 1$ $p_j = p_k$ (i.e. there is no difference in the

This approach can be implemented using the rbounds procedure in STATA (Becker and Caliendo, 2007); this procedure uses the data to calculate the confidence intervals (for a given level of confidence – e.g. 95%) of the outcome variable for different values of Γ. A value of Γ that produces a confidence interval that encompasses zero is one that would make the estimated impact not statistically significant at the relevant level of confidence. If Γ is relatively small (say < 2) then one may assert that the likelihood of such an unobserved characteristic is relatively high, and therefore that the estimated impact is rather sensitive to the existence of unobservables (DiPrete and Gangl, 2004). If there is other evidence that there may be unobservables, such as my qualitative observations of the SEWA Bank selection processes, we cannot be confident that the estimated impact is not due to unobservables.

We can illustrate this approach by calculating Γ at which the estimated impact of SEWA Bank membership on household income per capita for round 1 is no longer statistically significant. Table 5 shows that the 5-nearest neighbour matching estimate for total household income per annum per capita in round 1 is 2,019 Rupees which is significant at 1%. This suggests that households participating in microfinance earn significantly more income per annum per capita than control households; however, this may not be due to membership *per se* but unobserved characteristics that account for membership (and or its impact). Sensitivity analysis explores the robustness of this impact estimate and demonstrates how it changes in the presence of selection on unobservables. The STATA procedure rbounds reports the estimates[17] and their 95% (or other) confidence intervals for matched pairs of SEWA Bank members and controls (see Table 6).

When Γ = 1 there is no selection on unobservables. If Γ increases to 1.2, then matched individuals differ in their odds of exposure to microfinance by a factor of 1.2 due to selection on unobservables. Table 6 shows that when Γ = 1.2 the statistical significance level ranges from < 0.0001 to < 0.0046. This implies that in this case selection on unobservables is not likely to explain the observed association between exposure to microfinance and higher income levels. However, when Γ = 1.3 or more, a relatively small difference in the odds of exposure implying that it is quite likely that such an unobserved confounding variable exists, the 95% confidence interval of

odds of being selected). Γ =2 means that individual j is twice as likely to be selected into SEWA Bank as individual k. This might be considered not unlikely based on my observations of the selection process operated by SEWA Saathi, and by my understanding of the requirements of households to be able to save, and for other to qualify for borrowing.

[17] In this case we use Hodges-Lehmann point estimates (see Rosenbaum, 2002). These are median shifts between treatment groups. Therefore, they are likely to be smaller than the mean shifts reported in Table 5 which provides the average treatment effects.

the point estimates encompasses zero. Consequently, we can argue that the observed impact of SEWA Bank membership on household income per capita is not significantly different from zero, and the association between microfinance exposure and higher income levels may well be due to unobservables.

Table 6: Sensitivity analysis for household income per annum per capita in Rupees for microfinance participants for round 1: magnitude of selection on unobservables, range of significance levels, Hodges-Lehmann point estimates and confidence intervals

	Significance levels		**Hodges-Lehmann point estimates**		**95% Confidence intervals**	
Gamma (Γ)	Minimum	Maximum	Minimum	Maximum	Minimum	Maximum
1	< 0.0001	< 0.0001	953	953	520	1,403
1.2	< 0.0001	< 0.0046	559	1,357	141	1,834
1.3	< 0.0001	< 0.0330	391	1,545	-27	2,033
1.4	< 0.0001	< 0.1292	241	1,717	-174	2,220
1.5	< 0.0001	< 0.3181	98	1,876	-311	2,394
1.6	< 0.0001	< 0.5561	-31	2,036	-436	2,562
1.7	< 0.0001	< 0.7637	-148	2,188	-556	2,732
1.8	< 0.0001	< 0.8967	-261	2,328	-669	2,891

Source: Author’s own calculations.
Notes: see footnote 17.

Sensitivity analysis was conducted on all the outcome variables presented in Table 5; testing the sensitivity of all impact estimates on the household, enterprise and individual level across rounds 1 and 2. The evidence provided by those tests are in agreement with the description above, namely that the impact estimates presented in Table 5 are sensitive to selection on unobservables and should be treated with caution.[18] In other words, the impact estimates presented here are all potentially overstated due to the presence of selection on unobservables. PSM appears to control for selection on observables only and fails to account for unobservables which are quite likely to exist. However, this is considered a worst case scenario and does not prove that there is no effect since it assumes both that the unobserved variable has

[18] The detailed results from those sensitivity tests are not presented here but the relevant STATA do-files can be made available upon request.

the specific effect on the odds ratio of treatment and that it has a strong effect on the outcome variable (DiPrete and Gangl, 2004, p. 291). Nevertheless, PSM and related tests allow the quantification of selection on unobservables which is helpful. These results lead me to concur with Ichino, Mealli and Nannicini (2006), namely that sensitivity testing should always complement the presentation of matching estimates. In this case caution in concluding that SEWA Bank membership has a causal effect on per capita income is warranted.

7.4 Panel analysis

The panel data analysis with or without PSM reveals nothing new and broadly confirms the results obtained from the cross-section analysis. As illustrated in Table 7, PSM using nearest neighbour matching on round 1 data caused some households which did not match on observable characteristics to be dropped, and only matched households were merged with round 2 data. Using the treatment and matched households a regression-adjusted differences-in-differences (DID) model was run on all outcome variables as set out by the following equation which is a fixed effects linear regression model; *i* stands for household in ward *j* at period *t*:

$$(3) \qquad y_{ijt} = \alpha_i + \delta_t + \beta C_{it} + \theta X_{it} + V_j + \varepsilon_{ijt}$$

<u>Where:</u>

y_{ijt} = outcome on which impact is measured at period *t*

C_{it} = level of participation in microfinance, i.e. a membership dummy variable, in period *t*

X_{it} = vector of household level characteristics in period *t*

V_j = vector of ward level characteristics

α_i = fixed effects unique to household *i*

δ_t = period effect common to all households in period *t*

β, θ = unknown parameters

ε_{ijt} = error term representing unmeasured household and ward characteristics at period *t*

Some evidence was found that there are positive and significant impacts at the household, enterprise and individual level as outlined by Table 7.

Table 7: PSM and DID results - impact of microfinance participation; without sampling weights[19]

Household level hypotheses	
Total household income per annum in Rupees	11,287***
Total household income per annum per capita in Rupees	2,181***
Inverse Simpson index	0.121
Expenditure for housing improvements in Rupees	5,628***
Expenditure on household assets in Rupees	734**
School enrolment for girls aged 5 to 10	-0.019
School enrolment for boys aged 5 to 10	0.019
School enrolment for girls aged 11 to 17	0.025
School enrolment for boys aged 11 to 17	-0.011
Food expenditure per day per capita in Rupees	0.93
Enterprise level hypotheses	
Informal sector income of whole household - per month in Rupees	3,091**
Informal sector income of respondent only - per month in Rupees	1,876**
Microenterprise revenues of all enterprises in household - per month in Rupees	3,050**
Microenterprise revenues of microenterprises for which respondent is primarily responsible - per month in Rupees	1,559**
Current value of fixed assets of all microenterprises in household in Rupees	211
Current value of fixed assets of microenterprises for which respondent is primarily responsible in Rupees	482
Hours worked in previous week in all microenterprises in household	13.78***
Days worked in previous month in all microenterprises in household	10.17***
Main types of suppliers - inferior[20] suppliers? Yes=1, No=0	0.060*
Main types of customers - inferior[21] customers? Yes=1, No=0	0.083**
Individual level hypotheses	
Respect by other household members? Yes=1, No=0	0.015
Prepared to deal with future? Yes=1, No=0	0.019

Source: Author's own calculations.
Notes: *statistically significant at 10%, **statistically significant at 5%, ***statistically significant at 1%. The STATA procedure xtreg was applied to implement DID.

[19] As discussed in footnote 6 and 11.
[20] Individuals/households and retailers are inferior sources of supply as defined by Chen and Snodgrass (2001).
[21] Individual consumers are considered to be inferior customers as defined by Chen and Snodgrass (2001).

These findings are broadly in agreement with the results that USAID presented. However, caution is required when interpreting the panel data findings. One would have expected some differences between the cross-section and panel data results since panel analysis should account for the unobservables but, this appears not to have been the case in this study. However, one could argue that the USAID dataset is not a real panel with a 'true' baseline because round 1 respondents were already microfinance clients when the baseline dataset was collected. The same clients and control households were then re-surveyed two years later. Strictly speaking, the baseline should have collected data on households that were not participating in microfinance at the time of the baseline data collection but became microfinance clients between survey rounds. This would have allowed a before and after comparison which would have been better suited to the analysis because it would have been possible to compare the treatment and control households in terms of all observables including outcomes; this would have allowed one to assess whether these two samples were broadly similar in these terms, although it would still not be possible to control for unobservables which affect response to the treatment.

7.5 Summary

Most of the PSM results confirm the findings of the USAID study if one ignores unobservables. What does this outcome mean for the issue of selection bias and the utility of PSM? Chen and Snodgrass (2001) argue that ANCOVA using a suitable control group accounts for selection bias to a certain degree. Based on the PSM results presented in Table 5, the first impression is that this assessment is indeed accurate. However, doubts remain as there are strong qualitative and theoretical reasons to think that unobservables have not been fully controlled for. This notion is confirmed by the sensitivity analysis which shows that the matching estimates are quite sensitive to selection on unobservables.

Also, the quality of the matches is doubtful; PSM requires rich and large datasets in order to function properly (Heckman, Ichimura and Todd, 1997; Heckman, Ichimura, Smith and Todd, 1998; Smith and Todd, 2005). Moreover, the panel does not resolve the issue because it is not a 'true' panel, and, even if it were, might not control for the effects of unobservables. Microfinance clients might have been better off than non-clients even before participating in microfinance, i.e. in terms of access to social networks, wealth, skills or motivations and this might have led them to self-select or to be selected into microfinance either by their peers or the staff of the microfinance organisation, and to be able to benefit more from membership that otherwise observationally similar households.

7.6 Sub-group comparisons

Having reached these preliminary conclusions with regard to membership of SEWA Bank, whether as saver or (saver and) borrower, sub-group comparisons were conducted to understand the impact of savings compared to saving and borrowing. As argued in section 3 and 4, Ghate (2007), as well as many microfinance practitioners, believe that the poor need savings more than credit. The following comparisons were investigated: borrowers versus controls, savers versus controls, borrowers versus savers, one-time borrowers versus savers, repeat borrowers versus savers, one-time borrowers versus controls and repeat borrowers versus controls. Again, only the key findings are presented. The results of the comparisons of the various borrower groups with savers are similar to the various borrower group comparisons with controls in terms of absolute numbers and level of significance, hence only the latter comparisons are discussed.[22]

[22] The detailed and re-analysed results of all household level as well as individual and enterprise level hypotheses across all sub-group comparisons of the USAID study can be obtained from the author upon request.

Table 8: Selected household level PSM results – sub-group comparisons; without sampling weights[23]

	Total household income per annum		**Total household income per annum per capita**		**Expenditure for housing improvements**	
	5-nearest neighbour	Kernel matching, bandwidth 0.01	5-nearest neighbour	Kernel matching, bandwidth 0.01	5-nearest neighbour	Kernel matching, bandwidth 0.01
Borrower versus control						
Round 1	12,323***	12,323***	2,364***	2,347***	5,046***	5,069***
Round 2	17,915***	18,256***	3,222***	3,378***	8,137**	8,160**
Borrower versus saver						
Round 1	9,152***	9,020***	1,567**	1,405*	3,700**	3,349**
Round 2	11,014***	10,141***	1,634**	1,634**	4,547	4,115
Saver versus control						
Round 1	7,236**	6,472**	1,545**	1,431**	2,212**	1,858*
Round 2	10,162***	10,085***	1,899***	1,909***	5,044***	4,508*
One-time borrower versus control						
Round 1	11,196**	12,212***	1,998**	2,186**	5,125***	5,107***
Round 2	30,099***	27,700***	5,500***	5,669***	18,619*	18,468*
Repeat borrower versus control						
Round 1	17,556***	15,738***	3,410***	3,203***	6,059***	6,027***
Round 2#	2,319		1,112		-825	

Source: Author's own calculations.

Notes: *statistically significant at 10%, **statistically significant at 5%, ***statistically significant at 1%. All figures in Indian Rupees. The results in this table refer to the differences in the mean values between matched samples; they were obtained using the STATA command psmatch2. I also ran the STATA command pscore with the objective to cross-check the psmatch2 results across the various matching algorithms. The results I obtained from the different STATA routines displayed minor differences in terms of the size of coefficient and the level of significance. Morgan and Winship (2007) argue that matching results can vary depending on the matching algorithm and PSM routine applied.
No values for kernel matching in round 2, the sample was too small with propensity scores outside the common support region, no adequate matches were found. Results are bootstrapped.

The discussion of sub-group comparisons focuses on selected household level hypotheses, namely income per annum, income per annum per capita[24] and

[23] As discussed in footnotes 6, 11 and 19, the literature is unclear with regard to accommodating sampling weights in the context of matching. Hence, as before, the analysis across all sub-group comparisons across rounds 1 and 2 was conducted with and without sampling weights and the results obtained in both cases were conclusive. The author can be contacted for more information on this topic.

expenditure for housing improvements. The outcome variables with regard to children's education were dropped because their results were mostly insignificant across all rounds and across all sub-group comparisons, hence confirming the earlier findings of the cross-section and panel data analysis. Again, 5-nearest neighbour matching as well as kernel matching with a bandwidth of 0.01 were the matching algorithms of choice.

The PSM results in Table 8 indicate that borrowers do significantly better than controls across all three outcome variables. In detail, in the borrower versus control comparison the results of the outcome variables income per annum and income per annum per capita are consistent across rounds 1 and 2 in terms of size of impact and level of significance but with slightly higher absolute impact figures in round 2. This suggests that impact strengthens over time, i.e. the longer a client is participating in microfinance the more likely he or she is to reap benefits, but these additional advantages are slight once a saver has become a borrower. Similar trends can be observed in the borrower versus saver comparison where the coefficients are smaller than in the borrower versus control comparison.

Similarities can also be observed in the savers versus control comparison where the outcome variables income per annum and income per annum per capita are consistent across rounds 1 and 2 in terms of size of impact and level of significance but their absolute impact figures are slightly lower than the ones reported in the borrower versus control comparison. To conclude, savers do better than controls but are slightly worse off than borrowers, which is to be expected. As mentioned earlier, the SEWA Bank model is built around mobilising savings. As of FY 1997, SEWA Bank had on average ten times more savers than borrowers. Hence, it appears that focusing on a savings approach is indeed a desirable strategy since savers have significantly higher impact estimates than control group members.

However, no clear picture seems to emerge when comparing one-time borrowers versus controls with repeat borrowers versus controls; in part this is because sample sizes are small. Comparing round 1 figures only, it appears that repeat borrowers do significantly better than one-time borrowers who do better than controls, and both

[24] Total household income per annum per capita was re-calculated using the formula for the equivalence scale described in footnote 14. The equivalence scale adjusted results in Table 8 are conclusive with the ones presented in Table 5, and hence following the earlier procedure the results that are not adjusted are reported.

do better than savers. When using round 2 figures the reverse appears to be true, with repeat borrowers worse off than one-time borrowers; and savers.[25]

The evidence provided by those sensitivity analyses on selected outcome variables is presented in Table 8 (i.e. only significant matching estimates were tested). These results concur with those presented earlier in this paper, namely that the results are very sensitive to selection on unobservables;[26] and the results of the sub-group comparisons overstate the impact of microfinance participation.

8 Conclusion

This study contributes to the impact evaluation literature by providing new insights from re-analysing the existing USAID panel data with PSM and DID; it contributes to the microfinance literature by throwing doubt on the claims of impact of a well known microfinance project. The basic PSM results presented in this paper approximate those obtained by USAID, i.e. borrowers do better than savers who in turn do better than controls. Presented in this way, these findings broadly support the existing belief that savings by themselves are desirable and that savings tools are complementary to a credit approach. The evidence is inconclusive whether repeat borrowers do better than one-time borrowers, and it appears that this is not always the case. The estimates obtained from the repeat borrower versus control comparison are unreliable due to small sample sizes which did not allow implementation of an adequate matching procedure.

However, sensitivity analysis of the PSM estimates shows that the matching estimates do not appear to be particularly reliable as they indicate high sensitivity to selection on unobservables. This supports qualitative evidence from this study and the literature (Ito, 2003; Fernando, 1997) of the presence of strong selection on unobservables. Sensitivity analysis suggests that the more likely true impact estimates would be significantly lower, and possibly not significantly different from

[25] A word of caution, round 1 defines repeat borrowers as borrowers who have taken out more than two loans. In round 2, repeat borrowers refer to borrowers who have repaid their earlier loan and have taken out a new loan between survey rounds. There are only 56 repeat borrowers between the two survey rounds. The definition of repeat borrowers differs across rounds which would explain the inconsistency of the results. The sample size of 56 is simply too small to provide any meaningful PSM results.

[26] The detailed results from those sensitivity tests are not presented here but the relevant STATA do-files can be made available upon request.

zero, for all outcome variables across all sub-groups and data collection rounds if selection on unobservables could be controlled for.

Further questions are raised about the ability of these methods to control for unobservables with these data because the USAID panel data set is not a 'true' panel. It does not allow a before and after comparison; what is being compared is the change in the outcome variable between a group that was already a member of SEWA Bank in round 1 and a control group surveyed at the same time, with both groups at a later date. While compared to a proper before and after comparison this may underestimate the total impact assuming the two groups are indeed comparable, it at the same time reduces the possibility of controlling for unobservables because any differences between the participants and controls in the absence (before) SEWA Bank cannot be empirically observed in these data. It cannot be shown that the treatment group before treatment was indistinguishable in terms of outcome variables, or, of course, unobservables, from the control group because there are no data from before treatment. Further doubts are raised by the way in which treatment and controls were sampled. This failed to explicitly rule out bias because the method of sampling of controls is not reported sufficiently. Indeed, the description of the procedure lays open a strong possibility that control households may have less ability to benefit from SEWA Bank than participating households because they had that opportunity but either chose not to participate or were selected out by self, peers or SEWA Bank staff.

The collection of the new cross-section data has been of limited help since their results were rather inconclusive and yielded little explanatory power mainly due to the shortcomings of the data pointed out earlier in this paper. However, the effort to collect a new wave of data was highly instructive, including giving insights into current selection processes, which were likely to have been operative to some degree in the past.

Based on the findings in this paper, it can be argued that a selection or screening process could be at work which is driven by the unobservables, e.g. entrepreneurial drive, business skills, possibly social capital, which together affect microfinance participation and cross section and (not true) panel differences between the treatment and control groups. The qualitative results presented here indicate a strong presence of social capital influencing (as outlined in Box 1) participation but the quantitative results cannot confirm this view due to a lack of adequate data. This leads to the question of how well social capital can be measured in the first place; this is a topic that the World Bank extensively dealt with from the mid 1990s onwards and measurement tools such as the Social Capital Assessment Tool (SOCAT) were developed (Grootaert and Bastelaer, 2002). Those tools, however, are mostly too general to yield any useful data, as found in this study. This is perhaps not very surprising; the concept of social capital is rather fuzzy (Harriss and de Renzio, 1997; Molyneux, 2002), and hence difficult to measure.

The debate on the appropriateness of the evaluation methods currently used to account for selection bias is far from over; but it is clear there is no miracle cure. The discussion in this paper demonstrates that the evaluation techniques currently available have drawbacks in one way or another. PSM is not the wondrous tool as advocated by many and the impact estimates presented in this paper should be taken with appropriate qualifications. There is qualitative evidence that there are strong unobservable effects, and that the unobservables have not been accounted for by any of the econometric techniques employed. Thus, controlling for biases due to unobservable characteristics remains a major challenge and a clear-cut solution to this issue has not yet been found. One point is clear, however, it is recommended to complement strictly quantitative approaches with qualitative ones.

Finally, not only do these data and methods not provide support for the idea that microfinance is highly beneficial to the poor, rather than perhaps benefitting a slightly better off group, but it leaves open whether microfinance is of any real benefit at all, since much of the apparent difference between microfinance participants and controls is likely due to differences in their characteristics rather than the intervention per se, not withstanding “inspiring stories” (Armendáriz de Aghion and Morduch, 2005, p. 199). This raises the question of under what circumstances, and for whom microfinance has been, and could be of real rather than imagined benefit to the poor.

Bibliography

Adams, D. W., 1978. Mobilizing Household Savings through Rural Financial Markets. *Economic Development and Cultural Change,* 26 (3), p.547-560.

Aportela, F., 1999. Effects of Financial Access on Savings by Low-Income People. *Available at: http://www.lacea.org/meeting2000/FernandoAportela.pdf.*

Armendáriz de Aghion, B. & Morduch, J., 2005. *The Economics of Microfinance.* Cambridge: MIT Press.

Arun, T., Imai, K. & Sinha, F., 2006. Does the Microfinance Reduce Poverty in India? Propensity Score Matching based on a National-Level Household Data. Economics Discussion Paper, The University of Manchester, September.

Ashraf, N., Karlan, D. S. & Yin, W., 2006. Female Empowerment: Impact of a Commitment Savings Product in the Philippines. *Available at: http://www.econ.yale.edu/growth_pdf/cdp949.pdf.*

Augsburg, B., 2006. Econometric Evaluation of the SEWA Bank in India: Applying Matching Techniques based on the Propensity Score. Working Paper MGSoG/2006/WP003, Maastricht University, October.

Banerjee, A., Duflo, E., Glennerster, R. & Kinnan, C., 2009. The Miracle of Microfinance? Evidence from a Randomized Evaluation. *Available at: http://econ-www.mit.edu/files/4162.*

Basu, P., 2006. *Improving Access to Finance for India's Rural Poor.* Washington D.C.: The World Bank.

Bateman, M. & Chang, H.-J., 2009. The Microfinance Illusion. *Available at: http://www.econ.cam.ac.uk/faculty/chang/pubs/Microfinance.pdf.*

Becker, S. O. & Caliendo, M., 2007. Sensitivity Analysis for Average Treatment Effects. *The STATA Journal,* 7 (1), p.71-83.

Browning, M. & Lusardi, A., 1996. Household Saving: Micro Theories and Micro Facts. *Journal of Economic Literature,* 34 (4), p.1797-1855.

Bryson, A., Dorsett, R. & Purdon, S., 2002. The Use of Propensity Score Matching in the Evaluation of Active Labour Market Policies. Policy Studies Institute and National Centre for Social Research, Working Paper No. 4, Department for Work and Pensions.

Burgess, R. & Pande, R., 2005. Do Rural Banks Matter? Evidence from the Indian Social Banking Experiment. *The American Economic Review,* 95 (3), p.780-795.

Caliendo, M., 2006. *Microeconometric Evaluation of Labour Market Policies.* Berlin: Springer.

Caliendo, M. & Hujer, R., 2005. The Microeconometric Estimation of Treatment Effects - An Overview. Forschungsinstitut zur Zukunft der Arbeit (IZA) Discussion Paper No. 1653, July.

Caliendo, M. & Kopeinig, S., 2005. Some Practical Guidance for the Implementation of Propensity Score Matching. Forschungsinstitut zur Zukunft der Arbeit (IZA) Discussion Paper No. 1588, May.

Caliendo, M. & Kopeinig, S., 2008. Some Practical Guidance for the Implementation of Propensity Score Matching. *Journal of Economic Surveys,* 22 (1), p.31-72.

Census of India, 2001. *District Census Handbook, Part XII A&B.* Ahmedabad: Government of India.

Chemin, M., 2008. The Benefits and Costs of Microfinance: Evidence from Bangladesh. *Journal of Development Studies,* 44 (4), p.463-484.

Chen, M. A. & Snodgrass, D., 1999. An Assessment of the Impact of SEWA Bank in India: Baseline Findings. Report submitted to USAID Assessing the Impact of Microenterprise Services (AIMS), August.

Chen, M. A. & Snodgrass, D., 2001. Managing Resources, Activities, and Risk in Urban India: The Impact of SEWA Bank. Report submitted to USAID Assessing the Impact of Microenterprise Services (AIMS), September.

Coleman, B. E., 1999. The Impact of Group Lending in Northeast Thailand. *Journal of Development Economics,* 60 (1), p.105-141.

Collins, D., Morduch, J., Rutherford, S. & Ruthven, O., 2009. *Portfolios of the Poor: How the World's Poor Live on $2 a Day.* Princeton: Princeton University Press.

Cornfield, J., Haenszel, W., Hammond, E. & Lilienfeld, A., 1959. Smoking and Lung Cancer: Recent Evidence and a Discussion of Some Questions. *Journal of the National Cancer Institute,* 22, p.173-203.

Deaton, A., 2009. Instruments of Development: Randomization in the Tropics, and the Search for the Elusive Keys to Economic Development. *Available at: http://www.princeton.edu/~deaton/downloads/Instruments_of_Development.pdf.*

Dehejia, R. H. & Wahba, S., 1999. Causal Effects in Nonexperimental Studies: Reevaluating the Evaluation of Training Programs. *Journal of the American Statistical Association,* 94 (448), p.1053-1062.

Dehejia, R. & Wahba, S., 2002. Propensity Score-Matching Methods for Nonexperimental Causal Studies. *The Review of Economic Studies,* 84 (1), p.151-161.

Devaney, P. L., 2006. Microsavings Programs: Assessing Demand and Impact, A Critical Review of the Literature. Assessing the Impact of Innovation Grants in Financial Services, IRIS Center, June.

Dichter, T. & Harper, M. eds., 2007. *What's Wrong with Microfinance?* Warwickshire: Practical Action Publishing.

Dieckmann, R., 2007. Microfinance: An Emerging Investment Opportunity: Uniting Social Investment and Financial Returns. Report completed for Deutsche Bank Research, December.

DiPrete, T. A. & Gangl, M., 2004. Assessing Bias in the Estimation of Causal Effects: Rosenbaum Bounds on Matching Estimators and Instrumental Variables Estimation with Imperfect Instruments. *Sociological Methodology,* 34 (1), p.271-310.

Duflo, E., Glennerster, R. & Kremer, M., 2007. Using Randomization in Development Economics Research: A Toolkit. Centre for Economic Policy Research, Discussion Paper No. 6059, January.

Duflo, E. & Kremer, M., 2003. Use of Randomization in the Evaluation of Development Effectiveness. Unpublished mimeo.

Dupas, P. & Robinson, J., 2009. Savings Constraints and Microenterprise Development: Evidence from a Field Experiment in Kenya. NBER Working Paper No. w14693.

Fernando, J. L., 1997. Nongovernmental Organizations, Micro-Credit, and Empowerment of Women. *The ANNALS of the American Academy of Political and Social Science,* 554 (1), p.150-177.

Fisher, T. & Sriram, M. S., 2002. *Beyond Micro-Credit: Putting Development Back into Micro-Finance.* New Delhi: Vistaar Publications.

Gaile, G. L. & Foster, J., 1996. Review of Methodological Approaches to the Study of the Impact of Microenterprise Credit Programs. Report submitted to USAID Assessing the Impact of Microenterprise Services (AIMS), June.

Ghate, P., 2007. Consumer Protection in Indian Microfinance: Lessons from Andhra Pradesh and the Microfinance Bill. *Economic and Political Weekly,* 42 (13), p.1176-1184.

Grootaert, C. & Bastelaer, T. v. eds., 2002. *Understanding and Measuring Social Capital: A Multidisciplinary Tool for Practitioners.* Washington D.C.: The World Bank.

Hamermesh, D. S., 2007. Viewpoint: Replication in Economics. *Canadian Journal of Economics,* 40 (3), p.715-733.

Harriss, J. & de Renzio, P., 1997. "Missing Link" or Analytically Missing? The Concept of Social Capital. *Journal of International Development,* 9 (7), p.919-937.

Heckman, J. J., Ichimura, H., Smith, J. & Todd, P., 1998. Characterizing Selection Bias Using Experimental Data. *Econometrica,* 66 (5), p.1017-1098.

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

Heckman, J. J. & Urzua, S., 2009. Comparing IV with Structural Models: What Simple IV Can and Cannot Identify. NBER Working Paper No. 14706.

Hulme, D., 2000. Impact Assessment Methodologies for Microfinance: Theory, Experience and Better Practice. *World Development,* 28 (1), p.79-98.

Hulme, D. & Mosley, P., 1996. *Finance against Poverty.* London: Routledge.

Ichino, A., Mealli, F. & Nannicini, T., 2006. From Temporary Help Jobs to Permanent Employment: What Can We Learn from Matching Estimators and their Sensitivity? Forschungsinstitut zur Zukunft der Arbeit (IZA) Discussion Paper No. 2149, May.

Imbens, G., 2009. Better LATE Than Nothing: Some Comments on Deaton (2009) and Heckman and Urzua (2009). NBER Working Paper No. 14896.

Ito, S., 2003. Microfinance and Social Capital: Does Social Capital Help Create Good Practice? *Development in Practice,* 13 (4), p.322-332.

Karlan, D. S. & Morduch, J., 2009. Access to Finance. *Available at: http://karlan.yale.edu/p/HDE_June_11_2009_Access_to_Finance.pdf.*

Karlan, D. S. & Zinman, J., 2009. Expanding Microenterprise Credit Access: Using Randomized Supply Decisions to Estimate the Impacts in Manila. *Available at: http://karlan.yale.edu/p/expandingaccess_manila_jul09.pdf.*

Keynes, J. M., 1936. *The General Theory of Employment, Interest and Money.* London: Macmillan.

Khandker, S. R., 1998. *Fighting Poverty with Microcredit: Experience in Bangladesh.* New York: Oxford University Press.

Ledgerwood, J., 1999. *Microfinance Handbook: An Institutional and Financial Perspective.* Washington D.C.: The World Bank.

Levitt, S. D. & List, J. A., 2009. Was there Really a Hawthorne Effect at the Hawthorne Plant? An Analysis of the Original Illumination Experiments. NBER Working Paper No. 15016.

Manski, C. F., 1995. *Identification Problems in the Social Sciences.* Cambridge: Harvard University Press.

Miguel, E. & Kremer, M., 2004. Worms: Identifying Impacts on Education and Health in the Presence of Treatment Externalities. *Econometrica,* 72 (1), p.159-217.

Molyneux, M., 2002. Gender and the Silences of Social Capital: Lessons from Latin America. *Development and Change,* 33 (2), p.167-188.

Morduch, J. & Haley, B., 2002. Analysis of the Effects of Microfinance on Poverty Reduction. NYU Wagner Working Paper No. 1014, June.

Morgan, S. L. & Harding, D. J., 2006. Matching Estimators of Causal Effects. Propects and Pitfalls in Theory and Practice. *Sociological Methods & Research,* 35 (1), p.3-60.

Morgan, S. L. & Winship, C., 2007. *Counterfactuals and Causal Inference. Methods and Principles for Social Research.* Cambridge: Cambridge University Press.

Pitt, M. M. & Khandker, S. R., 1998. The Impact of Group-Based Credit Programs on Poor Households in Bangladesh: Does the Gender of Participants Matter? *Journal of Political Economy,* 106 (5), p.958-996.

Pritchett, L., 2009. The Policy Irrelevance of the Economics of Education: Is "Normative as Positive" Just Useless, or Worse? In Cohen, J. & Easterly, W., eds. *What Works in Development? Thinking Big and Thinking Small.* Washington D.C.: Brookings Institution Press.

Rogg, C. S., 2000. The Impact of Access to Credit on the Saving Behavior of Microentrepreneurs: Evidence from 3 Latin American Countries. *Available at: http://idbdocs.iadb.org/wsdocs/getdocument.aspx?docnum=1481486.*

Roodman, D. & Morduch, J., 2009. The Impact of Microcredit on the Poor in Bangladesh: Revisiting the Evidence. Center for Global Development, Working Paper No. 174, June.

Rosenbaum, P. R., 2002. *Observational Studies.* New York: Springer.

Rosenbaum, P. R., 2010. *Design of Observational Studies.* New York: Springer.

Rosenbaum, P. R. & Silber, J. H., 2001. Matching and Thick Description in an Observational Study of Mortality After Surgery. *Biostatistics,* 2 (2), p.217-232.

Rosenzweig, M. R., 2001. Savings Behaviour in Low-Income Countries. *Oxford Review of Economic Policy,* 17 (1), p.40-54.

Rutherford, S., 2001. *The Poor and Their Money.* New Delhi: Oxford University Press.

Saretsky, G., 1975. The John Henry Effect: Potential Confounder of Experimental vs Control Group Approaches to the Evaluation of Educational Innovations. *The American Educational Research Association's Annual Meeting.* Washington, D.C., 2 April 1975.

Sebstad, J. & Chen, G., 1996. Overview of Studies on the Impact of Microenterprise Credit. Report submitted to USAID Assessing the Impact of Microenterprise Services (AIMS), June.

Sebstad, J., Neill, C., Barnes, C. & Chen, G., 1995. Assessing the Impacts of Microenterprise Interventions: A Framework for Analysis. Center for Development Information and Evaluation, Working Paper No. 7, USAID, March.

Setboonsarng, S. & Parpiev, Z., 2008. Microfinance and the Millennium Development Goals in Pakistan: Impact Assessment Using Propensity Score Matching. Asian Development Bank Institute (ADBI) Discussion Paper No. 104, March.

Smith, J. A. & Todd, P., 2005. Does Matching Overcome LaLonde's Critique of Nonexperimental Estimators? *Journal of Econometrics,* 125, p.305-353.

Snodgrass, D. & Sebstad, J., 2002. Clients in Context: The Impacts of Microfinance in Three Countries: Synthesis Report. Report submitted to USAID Assessing the Impact of Microenterprise Services (AIMS), January.

Todd, H., 1996. *Women at the Center: Grameen Bank Borrowers After One Decade.* Boulder: Westview Press.

von Pischke, J. D., 1983. Towards an Operational Approach to Savings for Rural Developers. In Von Pischke, J. D., Adams, D. W. & Donald, G., eds. *Rural Financial Markets in Developing Countries.* Baltimore: Johns Hopkins University Press.