IDB WORKING PAPER SERIES No. IDB-WP-564

Soap Operas for Female Micro Entrepreneur Training

Eduardo Nakasone
Maximo Torero

December 2014

Inter-American Development Bank
Multilateral Investment Fund (MIF)

Soap Operas for Female Micro Entrepreneur Training

Eduardo Nakasone
Maximo Torero

Inter-American Development Bank
2014

Cataloging-in-Publication data provided by the
Inter-American Development Bank
Felipe Herrera Library

Nakasone, Eduardo.
Soap operas for female micro entrepreneur / Eduardo Nakasone, Maximo Torero.
p. cm. — (IDB Working Paper Series ; 564)
Includes bibliographic references.
1. Entrepreneurship—Training of—Peru. 2. Women executives—Training of—Peru. I. Torero, Maximo. II. Inter-American Development Bank. Multilateral Investment Fund. III. Title. IV. Series.
IDB-WP-564

http://www.iadb.org

Eduardo Nakasone
International Food Policy Research Institute (IFPRI)
e.nakasone@cgiar.org

Maximo Torero
International Food Policy Research Institute (IFPRI)
m.torero@cgiar.org

Soap Operas for Female Micro Entrepreneur Training

Eduardo Nakasone and Maximo Torero*

December 9, 2014

Abstract

This paper analyzes the impact of the Strengthening Women Entrepreneurship in Peru (SWEP) program. SWEP trained female micro entrepreneurs in business management practices (such as accounting and marketing). The training, which was provided in 4- to 5-hour sessions, used soap operas and practical exercises specifically designed for the program. A field experiment was conducted among a group of micro entrepreneurs based in two Peruvian cities (Lima and Piura) to investigate whether SWEP had a positive impact on its beneficiaries. The results show that the program positively affected the adoption of business practices taught by the program. In particular, those who received the training were 4 to 6 percentage points more likely to assign themselves a fixed salary (rather than taking cash from their businesses based on personal needs) and 6 to 11 percentage points more likely to keep better records of potential business contacts. Some positive impacts were found on the adoption of bookkeeping practices (4 to 6 percentage points), although this result is not significant across all of the specifications. Although these changes in adoption rates were large compared with their baseline levels, they were rather small in absolute terms. Therefore, the study did not find any impact on average business performance, household expenditures, or women's empowerment in the household. Qualitative information suggests that micro entrepreneurs were satisfied with the training, but considered that many of the practices taught by the program were difficult to follow because of time constraints.

*Markets, Trade and Institutions Division, International Food Policy Research Institute (IFPRI). 2033 K St NW, Washington DC 20006. Nakasone: e.nakasone@cgiar.org; Torero: m.torero@cgiar.org. We are grateful to Miriam Vasquez and Giancarlo Cafferata from APRENDA; Ximena Querol, Director of the 10000 Women Program - Peru; and Sandra Darville, Claudia Gutierrez, and Carmen Mosquera from the IADB's Multilateral Investment Fund for their support in the evaluation of the SWEP program. USKAY EIRL facilitated the data collection. Yuri Soares, Miguel Almanzar, and anonymous reviewers from IADB provided insightful comments on a previous version of this paper. Mike Murphy provided research assistance with the data and the survey instrument. Maribel Elias and Luciana Delgado supervised all the logistical details of the field work. All errors remain our own.

1 Introduction

There is increasing interest in understanding the role of micro enterprises in developing countries. First, micro enterprises constitute the vast majority of firms in developing countries. For example, Li and Rama (2012) find that this group comprises 61percent of firms in Chile; 83 percent in Turkey; 95 percent in South Africa; 85 percent in India; and 84 percent in Pakistan. Although some micro enterprises will disappear, others may transition into larger companies and provide the foundation for a modern private sector. Second, micro enterprises are the most important source of employment: 72 percent of jobs are created in micro enterprises in developing countries.[1] Third, they have relatively low productivity levels, and this is especially concerning given their large share of employment. In this line, Angelelli, Moudry, and Llisterri (2006) calculate that, even though 77 percent of total employment is generated by micro and small enterprises in Latin America, they only contribute to 30-60 percent of the gross domestic product.

The economic literature provides a long list of potential culprits for the low levels of productivity of micro enterprises. A considerable proportion of previous studies has focused on the challenges that financial constraints can impose on small businesses: if firms are unable to borrow, they would be unable to finance optimal levels of capital (e.g., Evans and Jovanovic 1989; de Mel, McKenzie, and Woodruff 2008; Fafchamps et al. 2014; Banerjee et al. 2013). Restuccia and Rogerson (2008) argue that policies that create other distortions in input or output markets can lead to misallocation of resources and reductions in productivity. Low levels of human capital (at least, measured through formal schooling) have been another candidate to explain low levels of productivity in small firms in developing countries. De Soto (1989) argues that the

[1] There are some other calculations of the share of jobs in micro enterprises in developing countries (for example, see Ayyagari, Demirguc-Kunt, and Maksimovic 2011). However, most of these estimates are based on business surveys (which usually exclude firms in the informal sector) or are calculated for a single sector (e.g., manufacturing). We construct the average of the share of employment in micro enterprises for a sample of developing countries in the *World Development Report 2013*. (See the Report's Online Appendix). The World Bank's (2012) estimates are based on household surveys and, therefore, do include the informal sector. We estimate the average share of employment in micro enterprises (weighted by the size of their labor force) for 27 developing countries that have collected household surveys circa 2005 or 2010.

regulatory framework in developing countries creates an unnecessary burden for businesses in developing countries and promotes informality.

More recently, another potentially constraining factor in micro enterprise development has gained attention: managerial capital (Bruhn, Karlan, and Schoar 2010). For example, Bloom and Reenen (2010) argue that persistent differences in firm productivity can be explained by management practices. The authors surveyed around 6,000 firms in 17 countries and measure their practices in three broad areas: monitoring (e.g., production process tracking), targets (e.g., goal setting), and incentives (e.g., promotion workers with high performance). Their findings suggest that there is considerable variation in firms' practices within and between countries, and that better management is associated with stronger performance (in terms of size, productivity, and survival). The idea that managerial capital can spur firms' growth has powerful policy implications: even within their capital limitations or adverse regulatory environment, micro enterprises can improve if they put their inputs to a better use. Not surprisingly, this notion has promoted many business training programs in developing countries (Cho and Honorati 2013).

In this paper, we assess the impact of a large-scale program that provided short-term training to female micro entrepreneurs in Peru. The project, "Strengthening Women Entrepreneurship in Peru (SWEP)", was implemented by the Inter-American Development Bank's Multilateral Investment Fund, who partnered with APRENDA (a local institution that provides business training for micro entrepreneurs) and the Thunderbird School of Global Management (a United States-based business school). The general objective of the project was to improve the contribution of women-headed micro and small enterprises to family incomes and the economy of Peru, by providing assistance aimed at broadening access to business training.

Since its deployment in 2010, SWEP has provided business training to 100,000 women in Peru. Training sessions were free for participants, short (one afternoon), and provided only once. They relied on a combination of media, games, practical exercises, and take-home guides. The contents used a soap opera format and were designed by APRENDA and Thunderbird for the needs of female Peruvian micro entrepreneurs . The soap opera depicts the struggles of the owner of a cash-strapped grocery shop who decides to open a catering business when her husband gets fired from his job. The idea is that trainees could relate to the main character's problems.[2] Among other topics, the soap opera illustrates the advantages of timely cash flows, setting a fixed salary for the owners (rather than using the business's cash for personal needs), keeping record of important business clients and suppliers to develop a business network, and better work-life balance. The soap opera was complemented with instructors' in-class instructions for group activities and workbook exercises.

We designed a field experiment where we reproduced SWEP's recruiting process and conducted a baseline survey with around 2,500 female micro enterprise owners. Participants were randomly assigned to one of two groups: 60 percent were assigned to the treatment group and were immediately invited to participate in SWEP's training, while the remaining 40 percent were assigned to the control group and received the training a year later, at the end of the experiment. We conducted follow-up surveys 6 and 12 months after the baseline to investigate whether women who attended the SWEP sessions implemented the business management practices taught by the program and if these improved practices had any impact on their business outcomes (sales, expenditures, profits, size, and productivity). Given that the program should affect resources under women's control, we examine if changes in business outcomes translate to their household. Therefore, to some extent, we also analyze whether enhanced business profitability induces improvements in broad household welfare indicators and female bargaining power.

Recent papers have analyzed the impact of business training on micro entrepreneurial outcomes and have found mixed results.[3] For example, Calderón, Cunha, and Giorgi (2013) find that a 48 hour business training program in rural Mexico increased beneficiaries' profits, revenues, and clients. Karlan and Valdivia (2011) conduct a field experiment program among

[2]Previous evidence has shown that access to television and media can alter women's behavior. For example, Jensen and Oster (2009) show that access to cable TV increases female autonomy and lowers tolerance of domestic violence. La Ferrara, Chong, and Duryea (2012) show that areas where soap operas (*novelas*) are broadcasted experience reductions in fertility rates. The authors argue that the more "modern" attitude that female characters have in the novelas generates changes in women's role models and aspirations.

[3]For a more detailed review of the available evidence on the impact of training programs for small firms in developing countries, see McKenzie and Woodruff (2014) and Cho and Honorati (2013).

clients of a microfinance institution (MFI) in Peru to analyze the impact of training. The training was provided in 30-60 minute sessions after the clients' regular banking meetings over a period of one to two years. While they find that participants in the program improved their knowledge on business practices, they do not find changes in their business performance (i.e., revenue, profits, or employment). Field, Jayachandran, and Pande (2010) implement a similar field experiment with poor female bank clients in India, where a group of them participated in a two-day training on financial literacy, business skills, and aspirations. They find differential effects of the program for different castes: while Hindu women in upper castes increased their borrowing and business incomes, there was no effect among Muslims or Hindu women from historically disadvantaged castes. Bruhn, Karlan, and Schoar (2010) investigate the impact of a training program in Bosnia and Herzegovina and find increased financial basic knowledge among its beneficiaries. However, they also find that this increased knowledge did not translate into increases in the probability of business survival, business start-up, performance, or sales.[4]

This somewhat disappointing evidence led many to believe that training by itself would not suffice for businesses to thrive. One possibility is that, even with more entrepreneurial skills, financial constraints might still hinder firms' growth. de Mel, McKenzie, and Woodruff (2014) test this hypothesis by analyzing the impact of a program that provided a group of women with either training alone or a combination of training and a cash grant. Their results suggest that training by itself led to the adoption of enhanced business practices, but did not improve firms' profitability.[5] In contrast, the combination of human and financial capital led to increases in short-run improvements in business performance, but this effect dissipates a couple of years later. However, these effects were not found by Berge, Bjorvatn, and Tungodden (2011) and Giné and Mansuri (2011) in Tanzania and Pakistan, respectively. In their studies, male and

[4]The authors do find some improvements on sales and performance among entrepreneurs with ex-ante higher financial literacy. But, overall, they do not find any average effects of the program.

[5]Karlan, Knight, and Udry (2012) conduct a similar field experiment with tailors in urban Ghana, where they test the impact of a training program (business consultancy) and a cash grant among tailors. They find that the training led to a temporary change in business practices (which disappeared a year later) and that the cash grant led to business investments. However, neither intervention increased firms' profitability.

female MFI clients were offered a combination of training and financial capital. Training increased business knowledge among participants, but this only translated into higher sales for males' business sales and no such effect was found among women. In addition, both studies find that financial capital did not have any effect on firms' performance.

Another possibility is that the contents or methods of traditional business training are not suitable for micro entrepreneurs in developing countries. In this spirit, Valdivia (2011) tests whether general entrepreneurial training needs to be complemented with more personalized technical assistance to unleash firm growth. In his study, a group of female participants attended a traditional training program (36 three-hour sessions over 12 weeks), while a subgroup received additional technical assistance during three months. He finds that business performance effects were concentrated among those who received technical assistance (whose sales increased by at least 18 percent), suggesting a need to tailor the approach of traditional training programs. Drexler, Fischer, and Schoar (2014) take a different approach: rather than providing more comprehensive training, they compare the effect of traditional training vis-à-vis one based on simplified rule-of-thumb business rules. They find that the latter had a positive impact on business performance, while the former did not have any significant effect.

This paper presents three contributions to the existing literature. First, it provides new experimental evidence of the impact of training on female entrepreneurs' adoption of improved business practices and firm performance, an area in which there is no consensus in the available literature. Second, we provide evidence on alternative training methodologies to deliver business training. Along the lines of Drexler, Fischer, and Schoar (2014), we investigate the role of a simplified training program provided through SWEP. Training was provided for an afternoon at a relatively low cost.[6] Because the program was short, simple, and cheap, it was able to reach more than 100,000 female micro entrepreneurs. Third, to our knowledge, this is the first

[6]The implementation cost was about $27 per participant. This estimation considers the following four components of the original budget: (a) design of training model and curriculum, (b) development of training tools, (c) annual training of trainer and coordinators, and (d) implementation of training. The budget for these four components was $2.75 million. This excludes other indirect costs, such as promotion and outreach campaigns, creation of links with trade associations and markets, etc. The program trained 102,401 female micro entrepreneurs.

impact evaluation of a training program that extensively used classroom media. It included motivational videos and a soap opera about a struggling micro entrepreneur. The media contents were tailored to the needs of female micro entrepreneurs and highlighted many daily life aspects with which participants could identify.

Our results suggest that SWEP increased the adoption of business practices taught by the program. In particular, we find that participants were significantly more likely to assign themselves a fixed salary (rather than take cash from the business based on household needs). We also find larger shares of entrepreneurs that implement bookkeeping practices and who keep lists of potential contacts (i.e., clients, suppliers, etc.) for networking, although these results are not robust in all of our specifications. We also find that participants were quite satisfied with the training program: the average ranking was 7.8/10 points and, in general, reported that they were able to understand well the course contents. This leads us to believe that this media-based program was quite successful as a teaching strategy for micro entrepreneurs. However, we do not find that the implementation of these practices translates into enhanced firm performance. We find no significant treatment impact on sales, payroll, or other business expenditures.

The remainder of the paper is divided in the following sections. Section 2 describes the experimental design, training, and data collection process. Section 3 lays out the methodology and presents the results of our estimations. Section 4 discusses the results and examines alternative explanations for the results. The last section concludes and sketches some policy implications.

2 Experimental Design and Data Collection

In general, impact evaluation of most programs suffer from selection bias (i.e., those who decide to sign up may have different characteristics than those who do not) or reverse causality (i.e., those who enroll in the program might be the ones who anticipated larger gains from it) problems.[7] We try to overcome these shortcomings through an experimental design.[8]

Ideally, we would have conducted an experiment with participants of the SWEP program. However, this was unfeasible for several reasons. First, many participants in the training events were already clients or potential clients of local MFIs that referred them to SWEP. These institutions were reluctant to have their clients disclose their business and contact information in our surveys.[9] Second, the use of SWEP participants would have implied additional costs and delays for APRENDA. The program had to meet ambitious targets for the number of trained women. To maximize their recruiting effectiveness, the SWEP team would intensively recruit women in a certain area, train as many as they could, and move to the next area. A randomized controlled trial would have required the SWEP team to recruit participants in an area, only train those randomly allocated to the treatment group, wait for an appropriate time span for the impact evaluation, and go back to the same area to train the remaining women in the control group.

Because of these obstacles, we implemented an alternative experimental design that would replicate APRENDA's recruiting process on a smaller scale (see Figure 1). Between February and May 2011, we canvassed the most important markets, industrial clusters, and areas with dynamic commercial activity in two cities (Lima and Piura).[10] This enabled us to create a roster of 4,024 female micro enterprises, which constitute a sample of the target universe of SWEP and the population from which APRENDA could have recruited participants for its business training program. Legally, micro enterprises are defined as those with at most 10 employees

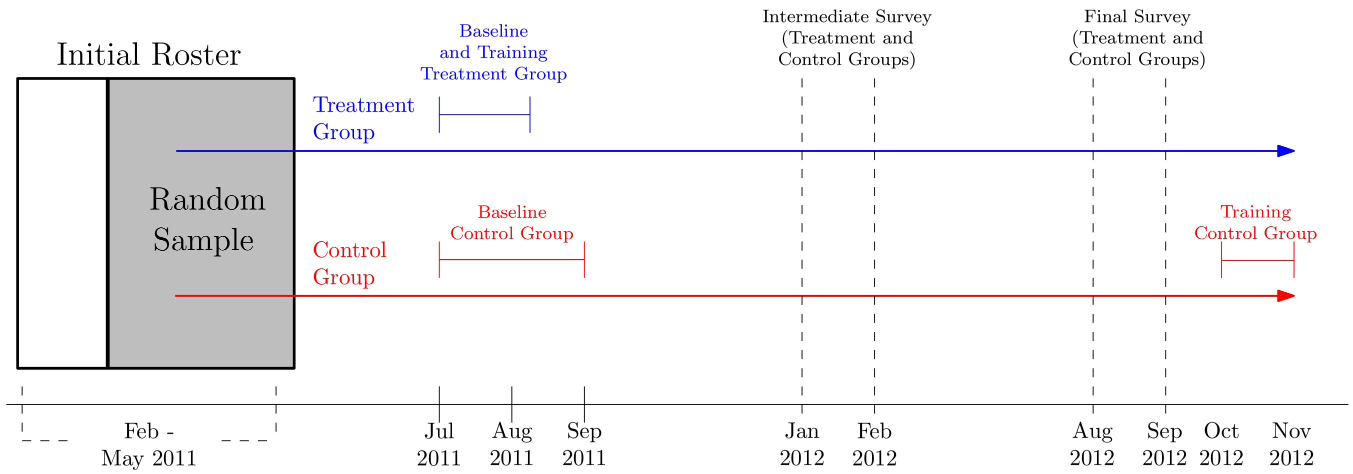
[7] For a discussion of methods to determine the causal impact of job training programs, see for example Heckman, Lalonde, and Smith (1999).

[8] See Duflo, Glennerster, and Kremer (2007) for a discussion of the advantages, disadvantages, and practical concerns of randomized controlled trials for evaluations.

[9] APRENDA is part of ACP, a local business holding that also owns an important share of an MFI (*Mi Banco*). Other MFIs referred their clients to SWEP as a complimentary business service, but feared that Mi Banco would contact their clients if they disclosed their personal information. Even when we guaranteed that we would not share any personal information with either APRENDA or Mi Banco, they were reluctant to participate in this experimental design.

[10] In Lima, our sample was concentrated in the following districts: Villa El Salvador (14 percent), Cercado de Lima (11 percent), San Juan de Lurigancho (7 percent), Comas (5 percent), Los Olivos (5 percent), Santa Anita (5 percent) and San Martin de Porres (5 percent). In Piura, the sample was highly concentrated in the district of Piura (64 percent).

Figure 1: Timeline of the Intervention



and with yearly sales of less than 150 tax units (about US$ 200,000). When business duties were shared by more than one household member, we asked that the primary decision maker be a woman. To mitigate attrition problems that have been prevalent in previous studies, we restricted our sample to businesses that had been operating for at least a year at the time of the roster collection. In addition, we explicitly asked micro entrepreneurs if they would be willing to participate in a free one-afternoon business training between July and September. Our roster was restricted to those who agreed to attend the business training.

We randomly sampled 2,600 micro entrepreneurs from our roster to participate in the evaluation: 1,500 were randomly assigned to the treatment group, while the remaining 1,100 were part of the control group. Each micro entrepreneur in the treatment group was invited to attend one out of 12 training sessions in our experimental design between July and August 2011. The ones in the control group did not receive the business training for the duration of the field experiment, but would be invited to participate at the end of it (October and November 2012). We collected some very basic information from women in the treatment and control groups during the collection of our roster. The first three columns of Table 1 show some of these characteristics: 61percent of them worked in retail (e.g., bodegas, clothing stores, etc.) 27 percent in services (e.g, beauty parlors, photocopy services, catering, etc.), and 12 percent in manufacturing (e.g., textile workshops, artisan jewelry and crafts, etc.). Their average age was around 42 years (65 percent of them were 30-50 years old), 30 percent had a bank account, and 70 percent owned a mobile phone.

Our treatment group was intentionally larger than the control group. We were concerned about entrepreneurs in the treatment group canceling or not showing up for the trainings. Previous recruiting experience of the APRENDA team suggested that many invitees would eventually decide not to attend the training.[11] Following APRENDA's approach, we included incentives for micro entrepreneurs to attend the training sessions. We hired buses to pick up women

[11]Imperfect compliance is usual in field experiments for training programs. Considerable no-show rates among micro entrepreneurs offered business training have also been documented by Bruhn, Karlan, and Schoar (2010), 61 percent; Drexler, Fischer, and Schoar (2014) 50~52 percent; Calderón, Cunha, and Giorgi (2013), 35 percent; Valdivia (2011), 49 percent; and Karlan and Valdivia (2011), 12~24 percent.

Table 1: Characteristics of Micro Entrepreneurs (Original Design vs. Effective Sample)

Characteristic	Design (Original Sample)			Effective Sample		
	Treatment	Control	Diff	Treatment	Control	Diff
City	0.79	0.79	0.00	0.74	0.80	-0.05***
(Lima=1)	(0.40)	(0.40)	(0.02)	(0.44)	(0.40)	(0.02)
Services	0.27	0.26	0.01	0.28	0.26	0.02
	(0.44)	(0.44)	(0.02)	(0.45)	(0.44)	(0.02)
Retail	0.61	0.62	-0.01	0.57	0.61	-0.04*
	(0.49)	(0.49)	(0.02)	(0.50)	(0.49)	(0.02)
Manufacturing	0.12	0.12	0.00	0.15	0.12	0.02
	(0.32)	(0.32)	(0.01)	(0.35)	(0.33)	(0.02)
Age	41.78	41.58	0.20	43.07	41.66	1.40**
	(11.18)	(11.58)	(0.45)	(10.63)	(11.52)	(0.55)
Has bank	0.30	0.30	0.01	0.31	0.30	0.01
account	(0.46)	(0.46)	(0.02)	(0.46)	(0.46)	(0.02)
Has mobile	0.70	0.69	0.01	0.72	0.69	0.03
phone	(0.46)	(0.46)	(0.02)	(0.45)	(0.46)	(0.02)
Has home	0.48	0.49	-0.02	0.51	0.50	0.01
land line	(0.50)	(0.50)	(0.02)	(0.50)	(0.50)	(0.02)
Has business	0.20	0.18	0.02	0.19	0.19	0.00
land line	(0.40)	(0.39)	(0.02)	(0.39)	(0.39)	(0.02)
Observations	1,500	1,100		703	1,035	

Note: The effective sample is comprised of women who attended the training in the treatment group and those who ones that completed the baseline questionnaire in the control group.

Standard deviations in parentheses. Significance levels of differences denoted by: *** p<0.01, ** p<0.05, * p<0.1.

from certain predetermined locations and provided them with lunch, a coffee break, and a symbolic present for their participation. Attendants would also enter a lottery where they could win appliances and other prizes. To maximize attendance, the training venues were carefully chosen considering the availability of public transportation, their central location, and closeness to the areas where the roster was collected.[12] Despite all our efforts , only 703 (47 percent) out of the 1,500 women invited to the trainings attended.

We collected a much more detailed baseline with the characteristics of the micro entrepreneurs in our field experiment. Our questionnaire included information on the women's micro enterprises (i.e., business practices, sales, costs, number of employees, payroll, access to credit, etc.) and households (i.e., composition, household expenditures, women's role in household decision making, etc.). The survey was administered with Android-based tablets to facilitate data collection. We took two approaches for the baseline collection. In one approach, women in the treatment group were interviewed in the training venues. Upon their arrival but prior to the beginning of the training the enumerators administered the survey. In the other approach, enumerators visited women in the control group at their houses or businesses to gather their information. Because enumerators were able to visit micro entrepreneurs multiple times and (re)schedule appointments as needed, 1,035 of the 1,100 women in our evaluation sample completed the baseline survey.

We were not able to collect data (baseline or follow-up) for invited micro entrepreneurs who did not attend the training or from the women in the control group who refused to be interviewed.[13] Therefore, our impact evaluation will be based on an effective sample of 703 micro entrepreneurs who did attend the training and 1,035 women in the control group. Because our sample and allocation deviated from our original random design, this creates some challenges to interpret our results. While the refusal rate (6 percent) in the control group was fairly small,

[12] There were ten training sessions in Lima in eight districts: Chorrillos (2), El Augustino (1), Independencia (2), Breña (2), Magdalena (1), Independencia (1), and Villa El Salvador (1). There were two training sessions in Piura.

[13] This prevents us from calculating standard Intention-to-Treat estimates as other papers in the literature (Karlan and Valdivia 2011; Valdivia 2011; Karlan, Knight, and Udry 2012; de Mel, McKenzie, and Woodruff 2014; Drexler, Fischer, and Schoar 2014).

the significantly larger no-show rates in the treatment group poses a problem: although women were randomly invited to the training, the actual decision to attend was not random. The last two columns of Table 1 shows that women in our final treatment group were 1.4 years older and more likely to work in retail. Table 2 presents more detailed characteristics of our effective sample at baseline. Women in both groups are similar in several dimensions (e.g., business value, education, access to formal credit, etc.). Importantly, they also have similar business outcomes in terms of sales, payroll, and other business expenditures. However, our effective treatment group has larger firms in terms of number of workers, has more access to informal credit, and is more likely to own their own stores. This seems to suggest that, if anything, women who attended the training were better off (and likely better able) than their counterparts in the control group. Therefore, the results of our paper constitute an upper bound of the actual effects of the training.

The APRENDA training sessions were typically 4-5 hours long. During these sessions, attendants watched a soap opera about Vicky, a struggling *bodega* owner.[14] The soap opera depicted wrongful management practices that are common among micro entrepreneurs.. Vicky's bodega was not able to keep up with the increasing household expenditures of her son and younger sister, and usually she covered these personal expenses with the bodega's day-to-day sales. Her situation turns critical when her husband is fired from his job and her sister unexpectedly gets pregnant. With the help of a friend, Vicky decides to start a catering business and get better organized. She sets a salary for herself (separating her personal expenses from the business's profits), sets a cash flow to decide any new investments, and collects a list of potential clients . These changes in practices eventually pay off and she is able to open a successful restaurant after a few months.

The soap opera was divided in six parts. After each part, the instructor reinforced the concepts the video illustrated with other examples and group games. Participants also received a workbook with exercises. After the soap opera and the instructor's explanations, participants

[14] The soap opera videos can be found in this website: http://www.programasalta.org/Programa/Material.

Table 2: Baseline Characteristics by Treatment Status

Characteristic	Control	Treatment	
A. Business Characteristics			
Age of business	98.8	101.7	
(months)	(91.8)	(91.4)	
Own store	0.48	0.52	*
	(0.50)	(0.50)	
Business value	26,674	25,780	
(self-reported, soles)	(34,008)	(33,796)	
B. Business Performance (soles)			
Yearly sales	43,887	42,003	
	(41,364)	(39,925)	
Yearly payroll	1,736	1,863	
	(4,024)	(4,068)	
Yearly expenditures	27,615	24,869	
(excl. payroll)	(36,266)	(32,354)	
C. Employment			
Total number of workers	0.81	1.10	***
	(1.17)	(1.36)	
Number of permanent	0.44	0.52	
salaried workers	(0.99)	(1.15)	
Number of permanent	0.28	0.42	***
non salaried workers	(0.63)	(0.72)	
Number of salaried	0.06	0.10	***
temporary workers [1]	(0.24)	(0.34)	
Number of temporary	0.03	0.06	***
non salaried workers [1]	(0.15)	(0.20)	
D. Credit			
Received formal loan	0.28	0.31	
in last 12 months	(0.45)	(0.46)	
Size of most important	6,193	5,691	
formal loan [2]	(7,476)	(9,403)	
Received informal loan	0.16	0.23	***
in last 12 months	(0.37)	(0.42)	
(cont'd)			
Size of most important	1,331	1,377	
informal loan [1]	(1,724)	(1,792)	
D. Household Variables			
Monthly per capita HH	358.5	363.7	
expenditure (soles)	(216.2)	(220.2)	
Value of household	4,132	4,927	
assets	(5,998)	(6,220)	***
D. Micro Entrepreneur's Characteristics			
Married	0.62	0.63	
	(0.49)	(0.48)	
Head of household	0.42	0.44	
	(0.49)	(0.50)	
Years of education	10.96	11.08	
	(3.36)	(3.28)	
E. Does woman decide..? [3]			
How to spend money	0.70	0.73	
	(0.46)	(0.44)	
Food purchases	0.83	0.85	
	(0.38)	(0.36)	
Furniture purchases	0.68	0.71	
	(0.47)	(0.46)	
Family outings	0.67	0.68	
	(0.47)	(0.47)	
Children's education	0.70	0.73	
	(0.46)	(0.44)	
Family discipline	0.66	0.71	**
	(0.47)	(0.45)	
What to do if any HH	0.82	0.82	
member is sick	(0.38)	(0.39)	
Observations	1,035	703	

[1] Number of temporary workers was adjusted by the number of months worked, i.e., number of temporary workers x (number of worked months) / 12.
[2] For the sample of those who received a loan in the last 12 months.
[3] Whether each of these decisions is primarily made by the micro entrepreneur.
Standard errors in parentheses. Significance levels denoted by: *** p<0.01, ** p<0.05, * p<0.1.

would start working on exercises. For example, the workbook would provide a hypothetical list of business sales, expenditures, and personal expenses. Participants would need to determine the cash flow, potential investments, and the salary they could assign themselves. The workbook also included take-home exercises where micro entrepreneurs were encouraged to work on their own cash flows and to create a list of potential clients to expand their business.

To measure the impact of the program, we collected two follow-up surveys. The first one was collected six months after the training (January and February 2012) and the second one was collected another six months later (August and September 2012). This survey was collected among those in the effective evaluation sample (703 in the treatment group and 1,035 in the control group). Enumerators visited micro entrepreneurs either at their home or business, depending on their availability. The questionnaires were similar to those in the baseline for comparability reasons.

3 Empirical Approach and Results

Our results focus on four main aspects on which the program might have affected micro entrepreneurs. First, we analyze whether training led to the adoption of the business practices that were taught to participants. The contents of the training prioritized three rules: (a) rather than taking cash based on the needs of their households, micro entrepreneurs should assign themselves a fixed salary every month; (b) they should register all sales and expenses in a cash flow to aid them in their investment and credit decisions; and (c) they should build a list of potential contacts (i.e. buyers, suppliers, etc.) to network. Second, we analyze whether there were any changes in business outcomes. In particular, we investigate if the training affected firms' sales, payroll, and other expenses. Third, we analyze if the program had any impact on micro entrepreneurs' household welfare as measured by participants' increased expenditures and the degree of the micro entrepreneurs' empowerment in the family.

We analyze the data with several specifications. Our basic specifications exploit the difference in outcomes Y_{it} for the post-intervention rounds of the data through a simple regression analysis:

$$Y_{i,t=T} = \beta Treat_i + \varepsilon_{it} \tag{1}$$

where $Treat_i$ is an indicator variable for the treatment group and $T = 1$ for the first follow-up (six months after the intervention) and $T = 2$ for the second follow-up (a year after the intervention). We also estimate the average effect over the impact evaluation period by pulling together both post-intervention rounds and estimating:

$$Y_{i,t=1,2} = \beta Treat_i + \theta D_{t=2} + \varepsilon_{it} \tag{2}$$

where $D_{t=2}$ is an indicator variable for the 12-month follow-up round of the data. Because our effective sample deviated from the original random design, we also estimate two additional versions of Equation 1 where we control for a set of characteristics of women in the control and treatment groups. In this spirit, we estimate ANCOVA (analysis of covariance) estimators, by including the baseline value of the dependent variable ($Y_{i,t=0}$) as a regressor.[15] We also estimate Equation 1 conditioning on a set X_i of micro entrepreneur's characteristics. This set X_i includes sector (i.e., services, manufacturing, or retail), city, age, head of household at baseline, whether the micro entrepreneur had received previous business training, and marital status.

$$Y_{it=T} = \beta Treat_i + \gamma Y_{i,t=0} + \varepsilon_{it} \tag{3}$$

$$Y_{it=T} = \beta Treat_i + \delta X_i + \varepsilon_{it} \tag{4}$$

Finally, we fully account for any (time-invariant) unobservable differences between the treatment and control groups, by estimating a fixed effects regression. Because this estimator only relies on within micro entrepreneur variability, it accounts the most for pre-intervention differ-

[15]McKenzie (2012) shows that the ANCOVA estimator accounts for pre-treatment differences in outcomes. In particular, $\hat{\beta}_{ANCOVA} = (\overline{Y^T_{Post}} - \overline{Y^C_{Post}}) - \hat{\gamma}(\overline{Y^T_{Pre}} - \overline{Y^C_{Pre}})$. He also shows that this procedure can increase the efficiency of the estimator.

ences of participants, but is also less efficient.

$$Y_{it} = \alpha_i + \beta_1 Treat_i D_{t=1} + \beta_2 Treat_i D_{t=2} + \theta_1 D_{t=1} + \theta_2 D_{t=2} + \varepsilon_{it} \tag{5}$$

Table 3 shows the results of our estimation framework for business practices. We find a strong significant impact of the program on the probability that the micro entrepreneur assigns herself a fixed salary. The coefficient is highly significant across all specifications and is large in magnitude: there is an increase of 5 percentage points, with a baseline value of 4 percent (more than doubling the share of women who implement this practice). In most specifications, we also find a positive impact of the program on the other two practices: bookkeeping (4-6 percentage points, from a baseline value of 45 percent) and networking list (3-10 percentage points, from a baseline value of 28 percent). However, the effect on networking is not significant in the fixed effects models, and the impact on bookkeeping is not significant in the ANCOVA and fixed effects models (arguably the two models with lower power to detect any effect).

Althoughwe find a general positive impact on the adoption of business practices, we do not find that these changes translate into any improvements in business performance. Table 4 shows the results for firms' sales, payroll, and business expenditures (excluding payroll). Our results suggest that firms did not experience any improvements and are, in general, consistent with previous results in the literature. As suggested by McKenzie and Woodruff (2014), most studies find some changes in business practices taught to training participants, but no effect on outcomes.

Finally, we estimate the impact of the program on household outcomes. First, we estimate whether there were any increases in household per capita expenditures (which is commonly used as a general measure of well-being). The results are shown in the top panel of Table 5, and suggest that the program did not have any significant impact on household expenditures. The coefficients are not statistically significant, and are rather small in magnitude compared with the means in the control group.

Next, we also test the impact of the program on women's empowerment. To capture the female micro entrepreneur's degree of empowerment in the household, we captured a set of questions

Table 3: Impact of Training on Business Practices

Specification	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Estimation	Post[1]	Post[1]	Avg Post[2]	ANCOVA[3]	Post[4]	ANCOVA[3]	Post[4]	F.E. [5]
Rounds[6]	1	2	1,2	1	1	2	2	0,1,2
Controls	No	No	No	No	Yes	No	Yes	No
N	1,480	1,371	2,851	1,480	1,480	1,371	1,371	4,589

A. Assigns Herself a Fixed Salary

Treat	0.040***	0.059***	0.049***	0.040***	0.039***	0.058***	0.057***	
	(0.013)	(0.014)	(0.011)	(0.013)	(0.013)	(0.014)	(0.014)	
Treat x $D_{t=1}$								0.036**
								(0.017)
Treat x $D_{t=2}$								0.056***
								(0.017)
$\bar{Y}_{ctrl,\ t=0}$								0.040
$\bar{Y}_{ctrl,\ t=1}$	0.051		0.051	0.051	0.051			
$\bar{Y}_{ctrl,\ t=2}$		0.044	0.044			0.044	0.044	

B. Bookkeeping

Treat	0.055**	0.046*	0.051**	0.038	0.044*	0.036	0.037	
	(0.026)	(0.027)	(0.021)	(0.025)	(0.026)	(0.026)	(0.026)	
Treat x $D_{t=1}$								0.000
								(0.030)
Treat x $D_{t=2}$								0.005
								(0.032)
$\bar{Y}_{ctrl,\ t=0}$								0.446
$\bar{Y}_{ctrl,\ t=1}$	0.413		0.413	0.413	0.413			
$\bar{Y}_{ctrl,\ t=2}$		0.446	0.446			0.446	0.446	

C. Networking

Treat	0.090***	0.111***	0.100***	0.063***	0.075***	0.090***	0.099***	
	(0.024)	(0.026)	(0.020)	(0.023)	(0.023)	(0.025)	(0.025)	
Treat x $D_{t=1}$								0.007
								(0.029)
Treat x $D_{t=2}$								0.028
								(0.031)
$\bar{Y}_{ctrl,\ t=0}$								0.279
$\bar{Y}_{ctrl,\ t=1}$	0.254		0.254	0.254	0.254			
$\bar{Y}_{ctrl,\ t=2}$		0.306	0.306			0.306	0.306	

1 Post estimators: $Y_{i,t} = \beta Treat_i + \varepsilon_{it}$, for each t=1 and t=2.
2 Average post estimator: $Y_{i,t=1,2} = \beta Treat_i + \theta D_{t=2} + \varepsilon_{it}$.
3 ANCOVA (analysis of covariance) estimators: $Y_{i,t} = \beta Treat_i + \gamma Y_{i,t=0} + \varepsilon_{it}$, for each t=1 and t=2.
4 Post estimators including baseline controls: $Y_{i,t} = \beta Treat_i + \delta X_i + \varepsilon_{it}$, for each t=1 and t=2. X_i includes sector, city, age, head of household at baseline, whether the micro entrepreneur had received previous business training, and marital status.
5 Fixed effects estimator: $Y_{it} = \beta_1 Treat_i D_{t=1} + \beta_2 Treat_i D_{t=2} + \theta_1 D_{t=1} + \theta_2 D_{t=2} + \alpha_i + \varepsilon_{it}$.
6 Rounds: baseline (0), Jan-Feb 2012 follow-up (1), and Aug-Sep 2012 follow-up (2).

Note: Standard errors in parentheses. Standard errors are clustered at the micro entrepreneur level in Columns 3 and 8. Significance levels denoted by: *** p<0.01, ** p<0.05, * p<0.1.

Table 4: Impact of Training on Business Outcomes

Specification	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Estimation	Post[1]	Post[1]	Avg Post[2]	ANCOVA[3]	Post[4]	ANCOVA[3]	Post[4]	F.E. [5]
Rounds[6]	1	2	1,2	1	1	2	2	0,1,2
Controls	No	No	No	No	Yes	No	Yes	No
N	1,468	1,366	2,834	1,453	1,468	1,361	1,366	4,536

A. Yearly Sales

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Treat	-1,673.3	-1,786.4	-1,728.0	-2,064.4	236.5	-525.9	-543.5	
	(3,660.0)	(3,174.7)	(2,751.4)	(2,234.5)	(3,629.1)	(3,028.6)	(3,159.3)	
Treat x $D_{t=1}$								-1,923.3
								(2,539.3)
Treat x $D_{t=2}$								570.3
								(3,042.8)
$\bar{Y}_{ctrl,\ t=0}$								43,886.7
$\bar{Y}_{ctrl,\ t=1}$	52,243.7		52,243.7	52,243.7	52,243.7			
$\bar{Y}_{ctrl,\ t=2}$		45,950.3	45,950.3			45,950.3	45,950.3	

B. Payroll

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Treat	394.2	-153.6	129.4	436.8	372.5	-175.0	-221.8	
	(397.0)	(235.4)	(269.0)	(360.3)	(395.5)	(220.7)	(233.6)	
Treat x $D_{t=1}$								439.3
								(407.9)
Treat x $D_{t=2}$								-41.9
								(304.7)
$\bar{Y}_{ctrl,\ t=0}$								1,736.2
$\bar{Y}_{ctrl,\ t=1}$	1,977.8		1,977.8	1,977.8	1,977.8			
$\bar{Y}_{ctrl,\ t=2}$		1,937.8	1,937.8			1,937.8	1,937.8	

C. Business Expenditures (excluding payroll)

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Treat	-1,532.5	-2,164.4	-1,837.0	-1,747.2	-144.2	-691.9	-1,272.4	
	(2,869.3)	(2,554.2)	(2,085.6)	(1,265.7)	(2,854.2)	(1,220.7)	(2,551.2)	
Treat x $D_{t=1}$								-1,787.4
								(1,487.4)
Treat x $D_{t=2}$								-797.4
								(1,431.7)
$\bar{Y}_{ctrl,\ t=0}$								27,615.1
$\bar{Y}_{ctrl,\ t=1}$	27,469.6		27,469.6	27,469.6	27,469.6			
$\bar{Y}_{ctrl,\ t=2}$		24,665.4	24,665.4			24,665.4	24,665.4	

[1] Post estimators: $Y_{i,t} = \beta Treat_i + \varepsilon_{it}$, for each t=1 and t=2.
[2] Average post estimator: $Y_{i,t=1,2} = \beta Treat_i + \theta D_{t=2} + \varepsilon_{it}$.
[3] ANCOVA (analysis of covariance) estimators: $Y_{i,t} = \beta Treat_i + \gamma Y_{i,t=0} + \varepsilon_{it}$, for each t=1 and t=2.
[4] Post estimators including baseline controls: $Y_{i,t} = \beta Treat_i + \delta X_i + \varepsilon_{it}$, for each t=1 and t=2. X_i includes sector, city, age, head of household at baseline, whether the micro entrepreneur had received previous business training, and marital status.
[5] Fixed effects estimator: $Y_{it} = \beta_1 Treat_i D_{t=1} + \beta_2 Treat_i D_{t=2} + \theta_1 D_{t=1} + \theta_2 D_{t=2} + \alpha_i + \varepsilon_{it}$.
[6] Rounds: baseline (0), Jan-Feb 2012 follow-up (1), and Aug-Sep 2012 follow-up (2).

Note: Standard errors in parentheses. Standard errors are clustered at the micro entrepreneur level in Columns 3 and 8. Significance levels denoted by: *** p<0.01, ** p<0.05, * p<0.1.

to determine whether she, her husband, or other family member decides:[16] (a) how to spend the household's income; (b) food purchases; (c) furniture purchases; (d) family outings; (e) children's education; (f) family discipline; and (g) what to do if a household member gets ill. Our analysis includes the estimation of seven empowerment measures. Independently testing such a large number of outcomes of the same family can lead to over-rejection: if we test a sufficiently large number of empowerment variables, we would considerably increase the probability that the treatment effect will be significant in at least one dimension (Duflo, Glennerster, and Kremer 2007; Schochet 2008).

We are really more interested in the effect of the training on women's empowerment, in general, rather than the strengthening of their bargaining power on specific household decisions. Therefore, we aggregate all seven decisions on a single index for this "family" of outcomes to avoid estimating too many outcomes. We adopt the framework proposed by Kling, Liebman, and Katz (2007). Consider G household decisions, each denoted by the subscript g. let σ_g be the standard deviation of the control group for each decision D_g. We estimate regressions $D_g = \theta_g \text{Treat}_i + \varepsilon_{ig}$ for each $g = 1, ..., G$. To account for the covariance of estimates θ_g across equations, Kling, Liebman, and Katz (2007) propose to estimate them as a system of Seemingly Unrelated Regressions (SUR), rather than estimating the parameters individually.[17] We then aggregate the effect on the family of decisions based on the treatment effects of each outcome

[16] Admittedly, measurement of women's empowerment is a complex issue and presents several challenges. Although we did not design a set of original questions for this purpose, we used several items from the Mexican Intrahousehold Violence Survey (*Encuesta sobre Violencia Intrafamiliar*): http://tinyurl.com/MexicoEnvif (see page 3).

[17] Evaluation of business training usually entails a large number of outcomes and is prone to family-wise error (FWE) problems. Karlan and Valdivia (2011) and Valdivia (2011) also analyze the impact of training programs on aggregate indexes of outcomes. Their approach is somewhat different: they normalize each outcome with the mean and standard deviation and the control group and aggregate the normalized variables: $\frac{1}{G}\sum_g \frac{(Y_{ig}-\mu_g)}{\sigma_g}$. They use this index as the dependent variable in a regression. This method is suitable when there are no missing variables in the analysis. In particular, (Kling and Liebman, 2004, p. 9) argue that "when an individual is missing data on an outcome however, the other non-missing outcomes implicitly are given more weight when the index is based on a simple average of non-missing standardized outcomes. The formulation described above based on the mean of estimated effects is a more direct summary of the estimates for each outcome." Some of our decisions are prone to missing data problems. For example, the question of whether women decide about their children's education is only answered by women with school-age sons or daughters. Therefore, we prefer to normalize the coefficients for different decisions rather than to create a single regressand based on normalized individual decisions.

and the control group standard deviations:[18]

$$\tau = \frac{1}{G}\left[\sum_{g=1}^{G}\frac{\theta_g}{\sigma_g}\right] \tag{6}$$

The sample variance of τ (used to test its significance level) is based on the full variance-covariance matrix of θ estimated through the SUR sytstem. Kling, Liebman, and Katz's (2007) estimation on a family of outcomes allows us to tell whether women experience any positive effect on the set of decisions as a whole. The bottom panel of Table 5 shows the results for the aggregate impact of the program an women's household decision-making. We find no statistically significant differences in the normalized aggregate empowerment index.

4 Discussion of Results

We discuss four threats to the validity of the above findings: (a) missing data in the treatment group caused by selective attendance to the training; (b) attrition in follow-up surveys; (c) statistical power limitations; and (d) low quality of the training and levels of practice adoption.

4.1 Selective Attendance to the Trainings

First, our identification strategy can be compromised by the high no-show rates among the invitees to the trainings. Because of the random assignment, had all the invitees attended the trainings, we would have expected very similar treatment and control groups. However despite all our efforts to increase attendance and our previous screening of each micro entrepreneur's interest only 703 (47 percent of invited women) attended the trainings. We could not interview the remaining 797 women in the treatment group who did not attend, and they

[18]Because we use different samples in our estimations, σ_g varies between specifications. In the cross-sectional post regressions with and without controls (i.e. Equations 1 and 4) we use the standard deviation of the control group in the period for which we estimate the regression (either $t = 1$ or $t = 2$). In the ANCOVA regressions, we standardize τ using the standard deviation in $t = 1$. Standardization in the ANCOVA is based on the baseline values of σ_g. It is not possible to estimate a difference-in-differences specification with this framework.

Table 5: Impact of Training on Household Outcomes

Specification	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Estimation	Post[1]	Post[1]	Avg Post[2]	ANCOVA[3]	Post[4]	ANCOVA[3]	Post[4]	F.E. [5]
Rounds[6]	1	2	1,2	1	1	2	2	0,1,2
Controls	No	No	No	No	Yes	No	Yes	No
N	1,498	1,418	2,916	1,454	1,498	1,376	1,418	4,600

A. Household per capita Expenditure (monthly, in Soles)

Treat	5.69	7.79	6.72	-1.94	6.77	4.56	9.01	
	(12.17)	(10.07)	(9.33)	(11.35)	(11.77)	(9.36)	(9.72)	
Treat x $D_{t=1}$								-12.88
								(14.67)
Treat x $D_{t=2}$								-8.34
								(12.16)
$\bar{Y}_{ctrl,\ t=0}$								362.68
$\bar{Y}_{ctrl,\ t=1}$	380.37		380.37	380.37	380.37			
$\bar{Y}_{ctrl,\ t=2}$		337.60	337.60			337.60	337.60	

B. Empowerment Index[7]

Treat (τ)	-0.031	-0.052	-0.042	-0.036	-0.017	-0.057	-0.051	N.A.
	(0.037)	(0.039)	(0.032)	(0.034)	(0.037)	(0.037)	(0.038)	-

[1] Post estimators: $Y_{i,t} = \beta Treat_i + \varepsilon_{it}$, for each t=1 and t=2.
[2] Average post estimator: $Y_{i,t=1,2} = \beta Treat_i + \theta D_{t=2} + \varepsilon_{it}$.
[3] ANCOVA (analysis of covariance) estimators: $Y_{i,t} = \beta Treat_i + \gamma Y_{i,t=0} + \varepsilon_{it}$, for each t=1 and t=2.
[4] Post estimators including baseline controls: $Y_{i,t} = \beta Treat_i + \delta X_i + \varepsilon_{it}$, for each t=1 and t=2. X_i includes sector, city, age, head of household at baseline, whether the micro entrepreneur had received previous business training, and marital status.
[5] Fixed effects estimator: $Y_{it} = \beta_1 Treat_i D_{t=1} + \beta_2 Treat_i D_{t=2} + \theta_1 D_{t=1} + \theta_2 D_{t=2} + \alpha_i + \varepsilon_{it}$.
[6] Rounds: baseline (0), Jan-Feb 2012 Follow up (1), and Aug-Sep 2012 Follow up (2).
[7] We calculate the impact on a set of household decision variables. In particular, we asked micro entrepreneurs whether they decide: (a) how to spend the household's income; (b) food purchases; (c) furniture purchases; (d) family outings; (e) children's education; (f) family discipline; and (g) what to do if a household member gets ill. We estimate a seemingly unrelated regressions system of decisions: $D_{ig} = \theta_g \text{Treat}_i + \varepsilon_i$ with decisions g=1,...,G. We estimate the overall impact on this family of outcomes through $\tau = \frac{1}{G}\Sigma_{g=1}^{G}\frac{\theta_g}{\sigma_g}$.

Note: Standard errors in parentheses. Standard errors are clustered at the micro entrepreneur level in Columns 3 and 8. Significance levels denoted by: *** p<0.01, ** p<0.05, * p<0.1.

were not part of the impact evaluation sample. This leaves us with an unbalanced composition of the treatment and control groups. However, as suggested by Table 2 relative to the control group it seems that our effective treatment group was comprised of women with larger businesses, somewhat more access to credit, and more household assets. This would suggest that the effective treatment group ended up having "better" business women and that, if anything, our estimates are an upper bound of the true impact.

Positive selection makes it improbable that those who did not show up (and for whom we do not observe outcomes) performed better than those in our sample. Nevertheless, even if that were the case, their performance would have needed to be exceedingly large compared with the ones who attended. We can do some back-of-the-envelope calculations to assess how much larger their outcomes would have needed to be to find an overall positive impact of the program. We illustrate this calculation for business sales. Take the first column of Panel A of Table 4. Average sales of the control group in $t = 1$ were S/. 52,244, and average sales in the treatment group were S/. 50,570 (i.e., S/. 52,244 + β_{Treat}). We can estimate what the average among those without data in the treatment group would have needed to be to generate, for example, an overall 10 percent impact (the minimum detectable effect for sales with the intervention, as discussed in section 4.3) of the program on sales is: $\left[\frac{(1.1X52{,}244X1500)-(703X50{,}570)}{757}\right]$. This would imply that those who did not attend would need to have had 26 percent larger sales than those who did to have a 10 percent effect in $t = 1$. Similarly, their sales should have been 29 percent higher than those who attended in $t = 2$. The magnitude of these estimates seem rather implausible.

4.2 Attrition in Follow-up Surveys

The results in Section 3 show that there are different numbers of observations in each round of the survey. There were 1,738 observations in the baseline that we tried to re-interview in the first and second follow-ups. However, only 80 percent of micro entrepreneurs participated in all three rounds of the survey. Table 6 shows the distribution of the sample attrition across rounds. Attrition is a usual problem in entrepreneurship studies because they take place in

Table 6: Attrition Rates across Rounds

Round[1]			Control		Treatment		All	
t=0	t=1	t=2	N	%	N	%	N	%
X	X	X	792	76.5	604	85.9	1,396	80.3
X			89	8.6	21	3.0	110	6.3
X	X		107	10.3	49	7.0	156	9.0
X		X	47	4.5	29	4.1	76	4.4
Total			1,035	100.0	703	100.0	1,738	100.0

[1] Baseline (t=0), first follow-up (t=1), and second follow-up (t=2).

urban settings and among relatively mobile individuals. For example, Karlan and Valdivia (2011), Valdivia (2011), and Berge, Bjorvatn, and Tungodden (2011) experienced attrition rates of 24, 28, and 18 percent, respectively. There were several reasons why we were not able to locate the micro entrepreneurs for follow-up interviews. Some of them were peddlers and street vendors (who could easily move their businesses) or went out of business.[19] We tried to locate them at home, but many of them had moved. A small share of them refused to cooperate with our survey. In general, sample attrition would not represent a problem for our analysis if it were uncorrelated with the treatment. However, Table 6 suggests that, in our case, our program was indeed associated with the attrition: the share of micro entrepreneurs who participated in the three rounds of our evaluation was 9 percentage points higher among those who received the training (i.

Because attrition is correlated with the treatment, we need to further investigate how it could affect our results. In particular, we would like to get an idea of what would have happened had those who did not participate in our follow-up surveys (and for which we do not have information after the intervention) remained in our estimation sample. For example, it is possible that those who dropped from the control group had smaller sales. This would imply an underestimation of the impact of the program, because the actual effect would have been larger

[19] Some of the micro entrepreneurs in our sample had their shops located in *Plaza Villa Sur*, a market that experienced a large fire while we were gathering the data for the first follow-up (Diario Correo 2012). Although we tried to interview them at home, many of them had moved away after the fire.

if we were able to incorporate the missing information for those who did not participate in the follow-up surveys.

Table 7 presents some baseline characteristics of the attrited micro entrepreneurs. The first four columns present some baseline variables within the treatment and control groups for those who attrited and those who remained in our estimation sample in the first follow-up. Columns 5-8 perform a similar analysis for the second follow-up survey (t=2). In general, micro entrepreneurs that attrited in the treatment group are similar to those who were re-interviewed. However, those who dropped from the sample in the control group appear to be better-off than those who remained: their businesses' values are larger, they have higher sales and payrolls, are more educated, and have larger per capita household expenditures. Although we cannot know with certainty what would have happened to those who attrited in the follow-up surveys, their baseline characteristics suggest that, if anything, attrition would have created an upward bias in our estimates. Because those missing in the control group were likely to have better businesses, we would be comparing the treatment group against a control group that had been artificially deflated.

Attrition may have biased our estimates of business performance (i.e., sales, payroll, and business expenses) and household outcomes (i.e., household expenditures and women's empowerment). In these cases, we do not find that the program had any positive impact, because these effects were likely to have been overestimated in Tables 3, 4, and 5; so the program had at most no positive impact.

We also follow Fitzgerald, Gottschalk, and Moffitt (1998) and adjust our estimates to account for selection based on micro entrepreneurs' observable characteristics. Basically, we explain attrition with a set of time-invariant and baseline characteristics and estimate each micro entrepreneur's probability of attrition. With these probabilities, we can construct estimation weights and, under certain assumptions, recover the true effect of the program. This procedure is discussed in more detail in Appendix B. Because micro entrepreneurs dropped from the sample (especially in the control group) were worse-off than those who did not, most of our attrition-adjusted estimates are smaller than the ones we find without the adjustment.

Table 7: Baseline Characteristics of Attrited and Non-Attrited Micro Entrepreneurs

Variable	First Follow-up (t=1)					Second Follow-up (t=2)				
	Control			Treatment		Control			Treatment	
	Sample[1]	Attrition		Sample[1]	Attrition	Sample[1]	Attrition		Sample [1]	Attrition
Age of business	98.3	102.4		103.2	83.0	97.8	103.4		103.3	87.6
(months)	(92.0)	(91.0)		(92.5)	(73.3)	(90.2)	(98.6)		(91.8)	(86.6)
Own Store	0.49	0.43		0.52	0.50	0.49	0.45		0.53	0.46
	(0.50)	(0.50)		(0.50)	(0.51)	(0.50)	(0.50)		(0.50)	(0.50)
Business value	25,700.1	33,132.8	**	25,548.9	29,005.6	25,757.8	30,621.7		25,734.3	26,197.3
(Self-reported, S/.)	(33,010.6)	(39,539.2)		(33,351.9)	(39,780.6)	(32,723.6)	(38,914.2)		(33,284.9)	(38,468.3)
Yearly sales (S/.)	42,816.1	51,117.0	**	41,785.0	44,900.2	43,050.2	47,523.5		41,806.7	43,842.7
	(40,203.0)	(48,047.3)		(40,294.4)	(34,910.9)	(41,115.8)	(42,353.3)		(40,033.8)	(39,164.3)
Yearly Payroll (S/.)	3,063.1	5,401.4	***	3,700.4	2,688.2	2,967.7	5,093.7	***	3,709.4	2,895.5
	(8,850.3)	(14,356.0)		(9,934.5)	(5,098.6)	(8,547.7)	(13,748.9)		(10,020.4)	(5,607.9)
Yearly expenditures	27,434.9	28,799.2		24,642.1	27,828.8	27,639.6	27,512.4		24,902.2	24,567.0
(excl payroll, S/.)	(36,393.5)	(35,535.0)		(32,259.2)	(33,775.0)	(36,921.8)	(33,418.1)		(32,320.6)	(32,907.3)
Total number of	0.81	0.80		1.11	1.03	0.79	0.86		1.10	1.10
workers[2]	(1.19)	(1.03)		(1.39)	(0.89)	(1.17)	(1.19)		(1.38)	(1.25)
Head of	0.43	0.37		0.45	0.34	0.42	0.42		0.43	0.50
household	(0.50)	(0.48)		(0.50)	(0.48)	(0.49)	(0.50)		(0.50)	(0.50)
Years of education	10.9	11.6	**	11.1	11.5	10.8	11.6	***	11.0	11.6
	(3.4)	(2.9)		(3.3)	(3.0)	(3.4)	(3.1)		(3.3)	(2.9)
HH Per capita monthly	352.9	395.7	**	360.0	414.9	351.2	390.2	**	358.4	413.1
expenditure (S/.)	(208.7)	(258.4)		(216.1)	(267.7)	(204.5)	(259.7)		(212.8)	(276.2)
N	899	136		653	50	839	196		633	70

[1] Those who we were able to interview and remain in the estimation samples.
[2] Number of total workers: permanent workers + adjusted temporary workers. Temporary workers were adjusted by the number of months worked, that is, the number of temporary workers x (number of worked months) / 12.

Note: Standard deviation in parentheses. Significance levels (for diferences between the sample and attrition groups) denoted by: *** p<0.01, ** p<0.05, * p<0.1.

Therefore, we are less worried about the impact of the program on business performance and household outcomes, where we do not find any positive effects. However, we are somewhat more concerned about the robustness of the positive and significant impacts we find on the adoption of business practices. If less skilled entrepreneurs dropped from the control group (as suggested by Table 7), they could have been less likely to have implemented any business practices and we would have overestimated the impact of the program. Alternatively, this could happen if those in the treatment group who dropped from the sample were those who would not have adopted the business practices taught during the training (i.e., the "actual" proportion of micro entrepreneurs in the treatment group that adopted a business practice would be smaller than the observed one).

Although both situations could happen simultaneously (i.e., the attrited controls could have had larger adoption rates and the attrited treated micro entrepreneurs could have had smaller adoption rates), we quantify both factors separately. First, we estimate how large the adoption rate in control group would need to be to wipe out any significant impact of the program. Assume a total sample of N (with N_c observations in the control group and $N - N_c$ observations in the treatment group). Suppose that n_c^A micro entrepreneurs attrited from the control group. Based on the available data for the $N_c - n_c^A$ micro entrepreneurs that we observe in the post-intervention survey, denote the estimated rate of adoption in the control group as p_c^o . The actual rate of adoption in the control group would be $\frac{n_c^A}{N_c} p_c^u + \left(\frac{N_c - n_A}{N_c}\right) p_c^o$, where p_c^u is the rate of adoption among the attriters. The actual treatment effect would then be: $\beta = p_T^o - \left[\frac{n_A}{N} p_c^u + \left(\frac{N - n_A}{N}\right) p_c^o\right]$,[20] where p_T^o is the observed rate of adoption in the treatment group. Of course we do not know p_c^u , but we can simulate different scenarios.

We want to determine how large the rate of adoption among attriters in the control group would need to be to have a zero-effect of program. To do so, we calculate the impact of the program through a regression, where we impute a rate of $p_c^u = (p_c^o + \delta)$ among attrited micro entrepreneurs in the control group for alternative positive values of δ (i.e. 1 percent, 2 percent,

[20]Note that the treatment effect for the estimations that consider only the micro entrepreneurs that were included in the follow-up survey would be $p_T - p_u^o$. This would overestimate the impact of the program as long as $p_c^u > p_c^o$.

...). To preserve the dichotomous nature of the variable, we impute either 0s or 1s to the attrited control group to simulate each hypothetical rate p_c^u. For each value of δ, we assign 1s to $\lceil n_c^A(p_c^o + \delta) \rceil$ (rounded up) observations, and assign 0s to the remaining $\lfloor n_c^A(1 - p_c^o - \delta) \rfloor$ observations. We are able to determine the value of $\tilde{\delta}$ for which the program no longer has any statistically significant effect, and calculate how much larger p_c^u would need to be with respect to p_c^o for this to happen (i.e., $\frac{p_c^o + \tilde{\delta}}{p_c^o}$).

This procedure can also be accommodated for the treatment group. In this case, instead of increasing the proportion of adopting micro entrepreneurs in the control group, we could simulate decreasing rates of adoption among the treatment group. Suppose that n_T^A micro entrepreneurs attrited from the treatment group and that their (unobserved) adoption rate is p_T^u, where $p_T^u = (p_T^o - \delta) < p_T^c$. We can calculate the impact of the program for alternative values of δ in: $\beta = \left(\frac{n_T^A}{N - N_c}\right)(p_T^o - \delta) + \left(\frac{N - N_c - n_T^A}{N - N_c}\right) p_T^o - p_c^o$. We impute 1s in $\lceil n_T^A(p_T^o - \delta) \rceil$ observations in the attrited treatment group, and impute 0s for the remaining $\lfloor n_T^A(1 - p_T^o + \delta) \rfloor$ observations. Again, this procedure will tell us the critical value of $\tilde{\delta}$ and how much smaller the adoption rate would need to be among the attrited group to eliminate the effect of the program.

We followed the same procedures for both post-intervention cross sectional data sets (t=1 and t=2). The estimates for assigning herself a fixed salary (A), keeping records of network contacts (B), and bookkeeping (C) are reported in Table 8. We find that it is very unlikely that attrition could eliminate the positive effect of the program on assignment of a fixed salary. For this to happen, the proportion of micro entrepreneurs that assigned themselves a fixed salary among attriters in the control group would need to be 3.1 to 5.3 times the proportion of those that we observe in the control group. Alternatively, even if no micro entrepreneurs in the attrited portion of the treatment group would have adopted this practice, the effect of the program would still be 3– 4 percent. The case is similar for keeping a record of potential network contacts: the proportion of micro entrepreneurs that adopt this practice among attriters in the control group would need to be twice what we observe in the data. We also find that even if no one among attriters in the treatment group adopts this practice, we would still get a 5 percent impact. All in all, we believe that it is very unlikely that those who attrited would be so substantially different that we would not find a positive impact of the program anymore.

However, our results for bookkeeping are not as robust. If those that attrited in the control group were 7 – 23 percent more likely to have a bookkeeping system (compared with those in the control group who remained in the sample), we would not be able to find a significant a positive effect any longer. Alternatively, this would also happen if the micro entrepreneurs in the attrited sample of the control group were 11% – 29 percent less likely to adopt this practice. Although this does not necessarily imply that the program did not have a positive impact on the adoption of bookkeeping practices, it could be the case that the positive effect that we previously found does not hold under some relatively plausible conditions.

4.3 Statistical Power

An alternative explanation is that the program was successful and had a positive impact on entrepreneurs, but our intervention migt be underpowered to detect any statistically significant change. To explore this possibility, we re-estimated our power calculations to determine the minimum detectable effect from our experimental design. We focus on micro entrepreneurs' adoption of practices and one outcome (yearly sales). Using the baseline information (i.e., 1,738 micro entrepreneurs: 703 in the treatment group and 1,035 in the control group), we simulate alternative hypothetical effects and determine the power to detect them at a 95 and 90 percent levels of significance. The results of these simulations are presented in Figure 2.

The power of our intervention reaches 0.8 (at a 90 percent level of confidence) for a 3 percentage-point increase in women's self-assignment of a fixed salary and a 5 percentage-point increase the share of micro entrepreneurs who keep records for business contacts. Because the impacts reported in panels A and B of Table 3 are above these minimum detectable effects, we find significant results. Our power calculations also suggest that considering a power of 0.8 and a 90% level of significance we would only be able to find a significant impact on bookkeeping if adoption would increase by at least 5.5 percentage points. The impact on this practice seems to be around this level between 3.7 and 5.5 percentage points (in panel B of Table 3) and the coefficients are statistically significant in some specifications. In other words, it might be that the impact on bookkeeping is positive, yet we would not be able to significantly detect it.

Table 8: Sensitivity of Business Practice Adoption to Attrition

A. Micro Entrepreneur Assigns Herself a Fixed Salary

t=1			t=2		
Adjustment (δ)[1]	Mean Imputed Group[2]	Effect [3]	Adjustment (δ)[1]	Mean Imputed Group[2]	Effect [3]
Impute missing observations in control					
0% (p_c^o)[4]	0.056	0.040***	0% (p_c^o)[4]	0.047	0.059***
		(0.012)			(0.012)
5%	0.102	0.033 **	5%	0.097	0.047**
		(0.013)			(0.013)
10%	0.153	0.024 *	10%	0.144	0.035*
		(0.013)			(0.013)
15%	0.203	0.015	15%	0.187	0.023
		(0.014)			(0.014)
20%	0.254	0.006	20%	0.195	0.010
		(0.014)			(0.015)
No Effect ($\tilde{\delta}$=12%)[5]	0.175	0.020	No Effect ($\tilde{\delta}$=14%)[5]	0.245	0.025
		(0.013)			(0.014)
$(p_c^o + \tilde{\delta})/\ p_c^o$ [6]	3.10		$(p_c^o + \tilde{\delta})/\ p_c^o$ [6]	5.25	
Impute missing observations in treatment					
0.0% (p_T^o)[4]	0.099	0.040 ***	0.0% (p_T^o)[4]	0.109	0.059***
		(0.012)			(0.012)
-2.5%	0.074	0.037 ***	-2.5%	0.082	0.055***
		(0.012)			(0.012)
-5.0%	0.049	0.035***	-5.0%	0.055	0.051***
		(0.012)			(0.012)
-7.5%	0.025	0.032 ***	-7.5%	0.036	0.048***
		(0.012)			(0.012)
-10.0%	0.000	0.029 **	-10.0%	0.009	0.044***
		(0.012)			(0.012)
No Effect [5]	NA	NA	No Effect [5]	NA	NA
$(p_T^o - \tilde{\delta})/\ p_T^o$ [6]	NA		$(p_T^o - \tilde{\delta})/\ p_T^o$ [6]	NA	

[1] Simulated rates of adoption of $p_c^o + \delta$ (or $p_T^o - \delta$) among attrited households in the control (or treament) group.

[2] We estimate a linear regression of the imputed variable on the treatment status and report the effect $\hat{\beta}$ and standard errors.

[3] While we impute a rate of adoption of $p_c^o + \delta$ (for the control group) or $p_T^o - \delta$ (for the treatment group), the rate in this column is not exactly the same because of rounding up when assigning 1s and 0s to the attrited observations. The rates reported here are: $\lceil n_c^A(p_c^o + \delta) \rceil / n_c^A$ and $\lceil n_c^T(p_T^o - \delta) \rceil / n_T^A$, respectively.

[5] When $\delta = 0$, the coefficient of the adjusted estimate is the same as the one calculated on the non-attrited sample.

[4] $\tilde{\delta}$ is the minimum value of δ for which the effect of the program is not significant.

[6] This ratio indicates how much larger (smaller) the adoption of practices should be among the attrited micro entrepreneurs in the control (treatment) group with respect to the observed mean in the control (treatment) group to yield a statistically insignificant effect.

B. Keeps Records of Network

t=1			t=2		
Adjustment (δ)[1]	Mean Imputed Group[2]	Effect[3]	Adjustment (δ)[1]	Mean Imputed Group[2]	Effect[3]
		Impute missing observations in control			
0% (p_c^o)[4]	0.258	0.089***	0% (p_c^o)[4]	0.310	0.111***
		(0.022)			(0.023)
10%	0.355	0.072***	10%	0.407	0.087***
		(0.022)			(0.023)
20%	0.457	0.053**	20%	0.508	0.062***
		(0.023)			(0.024)
30%	0.559	0.035	30%	0.609	0.037
		(0.023)			(0.024)
40%	0.656	0.018	40%	0.709	0.011
		(0.023)			(0.024)
No Effect ($\tilde{\delta}$=29%)[5]	0.548	0.037	No Effect ($\tilde{\delta}$=30%)[5]	0.609	0.037
		(0.023)			0.024
$(p_c^o + \tilde{\delta})/\ p_c^o$ [6]	2.12		$(p_c^o + \tilde{\delta})/\ p_c^o$ [6]	1.96	
		Impute missing observations in treatment			
0% (p_T^o)[4]	0.345	0.089***	0% (p_T^o)[4]	0.421	0.111***
		(0.022)			(0.023)
-10%	0.250	0.078***	-10%	0.325	0.095***
		(0.022)			(0.023)
-20%	0.155	0.066***	-20%	0.219	0.078***
		(0.022)			(0.023)
-30%	0.048	0.054**	-30%	0.123	0.063***
		(0.022)			(0.023)
-35%	0.000	0.048**	-40%	0.026	0.047**
		(0.022)			(0.023)
No Effect[5]	NA	NA	No Effect[5]	NA	NA
$(p_T^o - \tilde{\delta})/\ p_T^o$ [6]	NA		$(p_T^o - \tilde{\delta})/\ p_T^o$ [6]	NA	

[1] Simulated rates of adoption of $p_c^o + \delta$ (or $p_T^o - \delta$) among attrited households in the control (or treament) group.
[2] We estimate a linear regression of the imputed variable on the treatment status and report the effect $\hat{\beta}$ and standard errors.
[3] While we impute a rate of adoption of $p_c^o + \delta$ (for the control group) or $p_T^o - \delta$ (for the treatment group), the rate in this column is not exactly the same because of rounding up when assigning 1s and 0s to the attrited observations. The rates reported here are: $\lceil n_c^A(p_c^o + \delta)\rceil / n_c^A$ and $\lceil n_c^T(p_T^o - \delta)\rceil / n_T^A$, respectively.
[4] When $\delta = 0$, the coefficient of the adjusted estimate is the same as the one calculated on the non-attrited sample.
[5] $\tilde{\delta}$ is the minimum value of δ for which the effect of the program is not significant.
[6] This ratio indicates how much larger (smaller) the adoption of practices should be among the attrited micro entrepreneurs in the control (treatment) group with respect to the observed mean in the control (treatment) group to yield a statistically insignificant effect.

C. Bookkeeping

t=1			t=2		
Adjustment (δ)[1]	Mean Imputed Group[2]	Effect [3]	Adjustment (δ)[1]	Mean Imputed Group[2]	Effect [3]
Impute missing observations in control					
0% (p_c^o)[4]	0.418	0.054**	0% (p_c^o)[4]	0.447	0.047*
		(0.024)			(0.024)
5%	0.463	0.047*	5%	0.498	0.035
		(0.024)			(0.024)
10%	0.514	0.038	10%	0.549	0.022
		(0.024)			(0.024)
15%	0.565	0.029	15%	0.599	0.010
		(0.024)			(0.024)
20%	0.616	0.021	20%	0.650	-0.003
		(0.024)			(0.024)
No Effect ($\tilde{\delta}$=10%)[5]	0.514	0.038	No Effect ($\tilde{\delta}$=3%)[5]	0.479	0.039
		(0.024)			(0.024)
$(p_c^o + \tilde{\delta})/\ p_c^o$ [6]	1.23		$(p_c^o + \tilde{\delta})/\ p_c^o$ [6]	1.07	
Impute missing observations in treatment					
0% (p_T^o)[4]	0.469	0.054**	0% (p_T^o)[4]	0.500	0.047*
		(0.024)			(0.024)
-5%	0.420	0.049**	-5%	0.445	0.039
		(0.024)			(0.024)
-10%	0.370	0.043*	-10%	0.400	0.032
		(0.024)			(0.024)
-15%	0.321	0.037	-15%	0.345	0.023
		(0.024)			(0.024)
-20%	0.272	0.032	-20%	0.300	0.016
		(0.024)			(0.024)
No Effect ($\tilde{\delta}$=-13%)[5]	0.333	0.039	No Effect ($\tilde{\delta}$=-5%)[5]	0.445	0.039
		(0.024)			0.024
$(p_T^o - \tilde{\delta})/\ p_T^o$ [6]	0.71		$(p_T^o - \tilde{\delta})/\ p_T^o$ [6]	0.89	

[1] Simulated rates of adoption of $p_c^o + \delta$ (or $p_T^o - \delta$) among attrited households in the control (or treament) group.

[2] We estimate a linear regression of the imputed variable on the treatment status and report the effect $\hat{\beta}$ and standard errors.

[3] While we impute a rate of adoption of $p_c^o + \delta$ (for the control group) or $p_T^o - \delta$ (for the treatment group), the rate in this column is not exactly the same because of rounding up when assigning 1s and 0s to the attrited observations. The rates reported here are: $\lceil n_c^A(p_c^o + \delta)\rceil / n_c^A$ and $\lceil n_c^T(p_T^o - \delta)\rceil / n_T^A$, respectively.

[4] When $\delta = 0$, the coefficient of the adjusted estimate is the same as the one calculated on the non-attrited sample.

[5] $\tilde{\delta}$ is the minimum value of δ for which the effect of the program is not significant.

[6] This ratio indicates how much larger (smaller) the adoption of practices should be among the attrited micro entrepreneurs in the control (treatment) group with respect to the observed mean in the control (treatment) group to yield a statistically insignificant effect.

We reach a power of 0.8 (with a 90 percent confidence level) with an increase of about 10 percent in sales. Although this minimum detectable effect is somewhat larger, our results in Table 4 show that the impact was rather small and, in most cases, even negative. Our coefficients suggest that, if any, the effect of the program on sales was between -4 and 1 percent. Because of the magnitudes of our estimates, we do not believe that these results are driven by statistical power concerns.

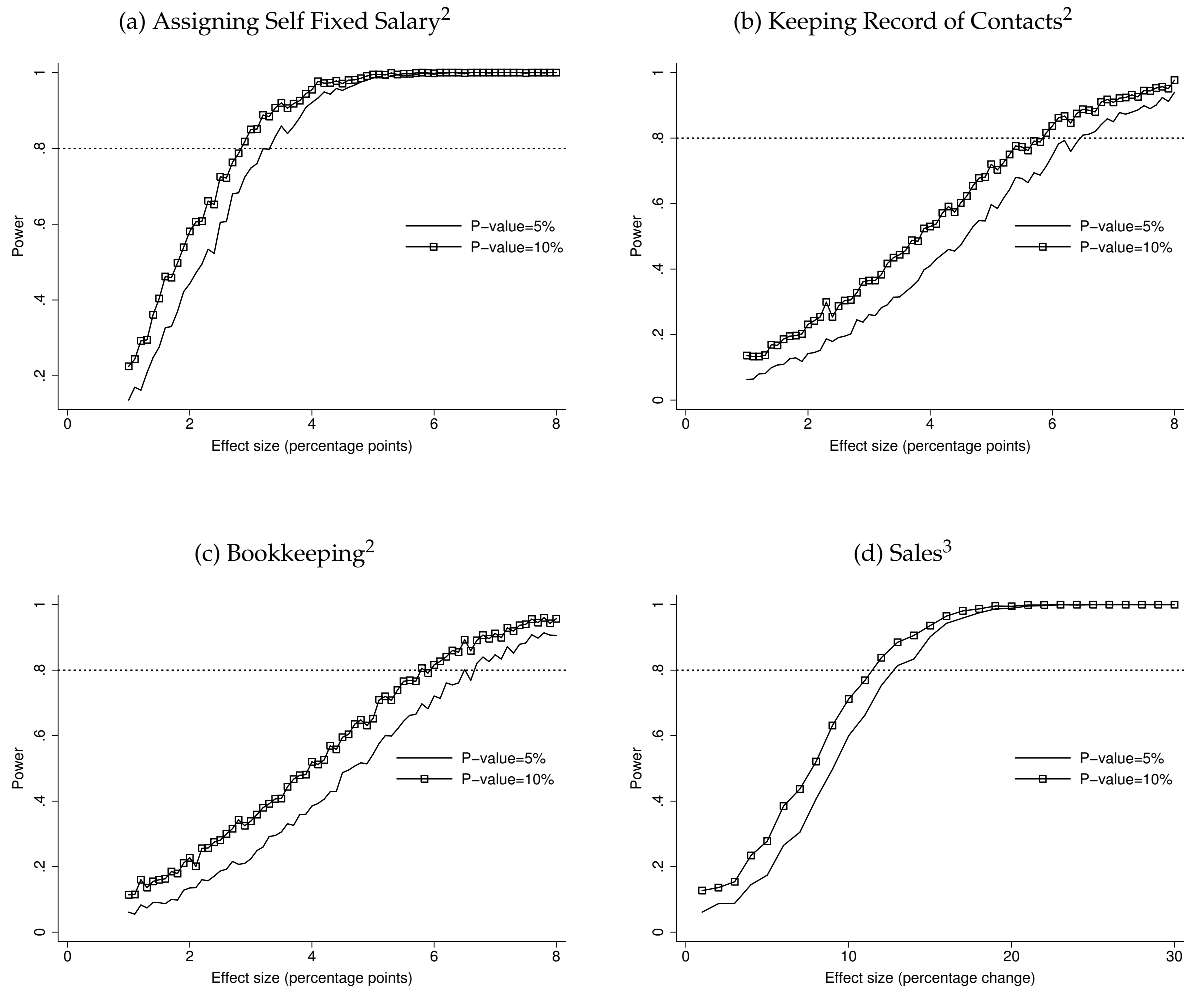
4.4 Quality of the Training and Practice Adoption

A possible explanation for our results is that the training was not useful for the participants. This could happen if the training did not explain how to implement their proposed business practices clearly enough or if it was unable to convey their usefulness. It could also be the case that even when participants did understand what they were taught the contents of the training were not helpful for their businesses or were difficult to implement.

It does not seem the case that micro entrepreneurs did not understand the training. In general, program participants seemed highly satisfied with the training they received. At endline, we asked the women in the treatment group (that had participated in the program about a year before) to rate their satisfaction with the training from 0 to 10 (with 0 being highly dissatisfied and 10 being highly satisfied). On average they assigned a score of 7.8 and 68 percent of them rated the training 8 or more. Albeit subjective, this score does not support the idea that participants did not value the training or that they did not find it useful.

If participants understood the training and considered there were any benefits from the contents they were taught, it might be that the contents of the programs were not practical for their businesses. This could explain why the the rate of practice adoption was relatively small: adoption rates of practices were relatively higher among those who received the training, but the increases were far from substantial in absolute terms. Table 3 shows that those who received the training were 4-6 percentage points more likely to assign themselves a fixed salary. This is an increase of 90–100 percent with respect to the baseline rate. Similarly, the share of micro entrepreneurs who kept a record of their business contacts was 6–11 percentage points larger

Figure 2: Power Calculations, Selected Variables[1]



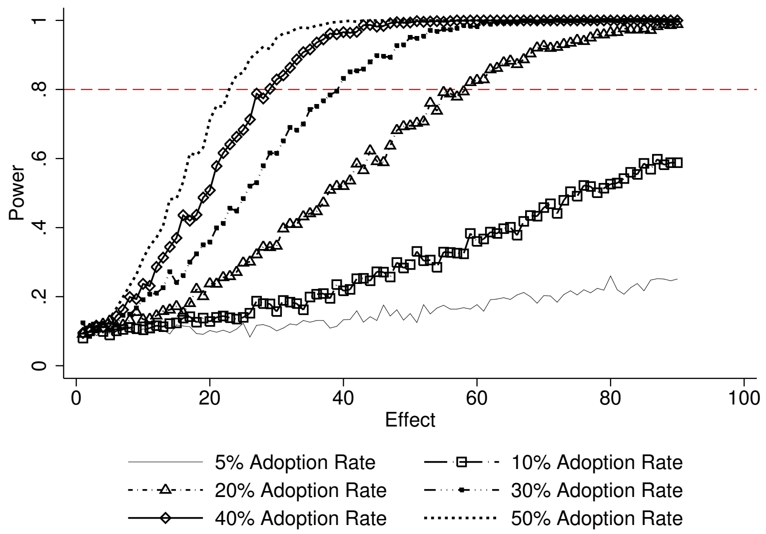
[1]The power calculations determine the proportion of times that we would reject the null hypothesis H_0 : Effect=0 (with p-values of 5 percent and 10 percent) for each effect size. Using the baseline means and standard deviations, we perform 1,000 simulations for each effect size. The simulations are based on our effective sample size at baseline: 1,738 micro entrepreneurs (703 in the treatment group and 1,035 in the control group).
[2]Effect size is expressed in percentage points for discrete variables.
[3]Effect size is expressed in percentage change for continuous variables.

(which represents a difference of 21–39 percent when compared with their initial situation) and the proportion of those who implemented bookkeeping was 4–6 percentage points higher (10–14 percent increase with respect to baseline) in the treatment group. Although all these changes were significant relative to the baseline levels, the changes are rather small in absolute terms. Hence, it is not completely surprising that these relatively mild changes would not necessarily translate into large improvements in the treatment group's average business performance, household outcomes, or female empowerment. In other words, if we think of these practices as the drivers of performance outcomes, then only women who marginally adopt these techniques would experience improvements, and these increases would be swamped on average.

We simulate how large the adoption rates of business practices would need to be to be able to detect a statistical impact on sales. Figure 3 estimates the statistical power of our intervention with hypothetical adoption rates and effects on sales. Although we do not estimate the power for any particular practice, we assume a general practice that would be adopted by 5, 10, 20, 30, 40, and 50 percent of the micro entrepreneurs in the treatment group. We suppose that the adoption of such practice would boost business sales of those who implement it by 1–90 percent of business sales (i.e., sales increases are generated by the adoption of the practice, so those who did not implement it would not experience such benefit). Assuming a significance level of 90 percent (i.e., a p-value of 10 percent), we find that any changes in sales values would be very difficult to detect with low levels of adoption. With a relatively high adoption rate of 50 percent, our intervention reaches a power of 0.8 when the effect of practices is about 20 percent of sales. With intermediate rates of adoption of 20–40 percent, practices would need to have an impact of 25–60 percent of income to reach that level of statistical power. When adoption is as low as 5–10 percent as in our intervention a power of 0.8 is not reached even if business sales double among adopters.

Among micro entrepreneurs in the treatment group who did not adopt each business practice in the endline, we did ask why they did not do so. This allows us to present some qualitative evidence and explore why the training did not yield larger adoption rates. The results of this analysis are presented in Figure 4. In all the cases, there were few women who did not perceive any benefits from the business practices taught by the program or that did not understand the

Figure 3: Adoption Rates and Power to Detect Increases in Sales Values[1]



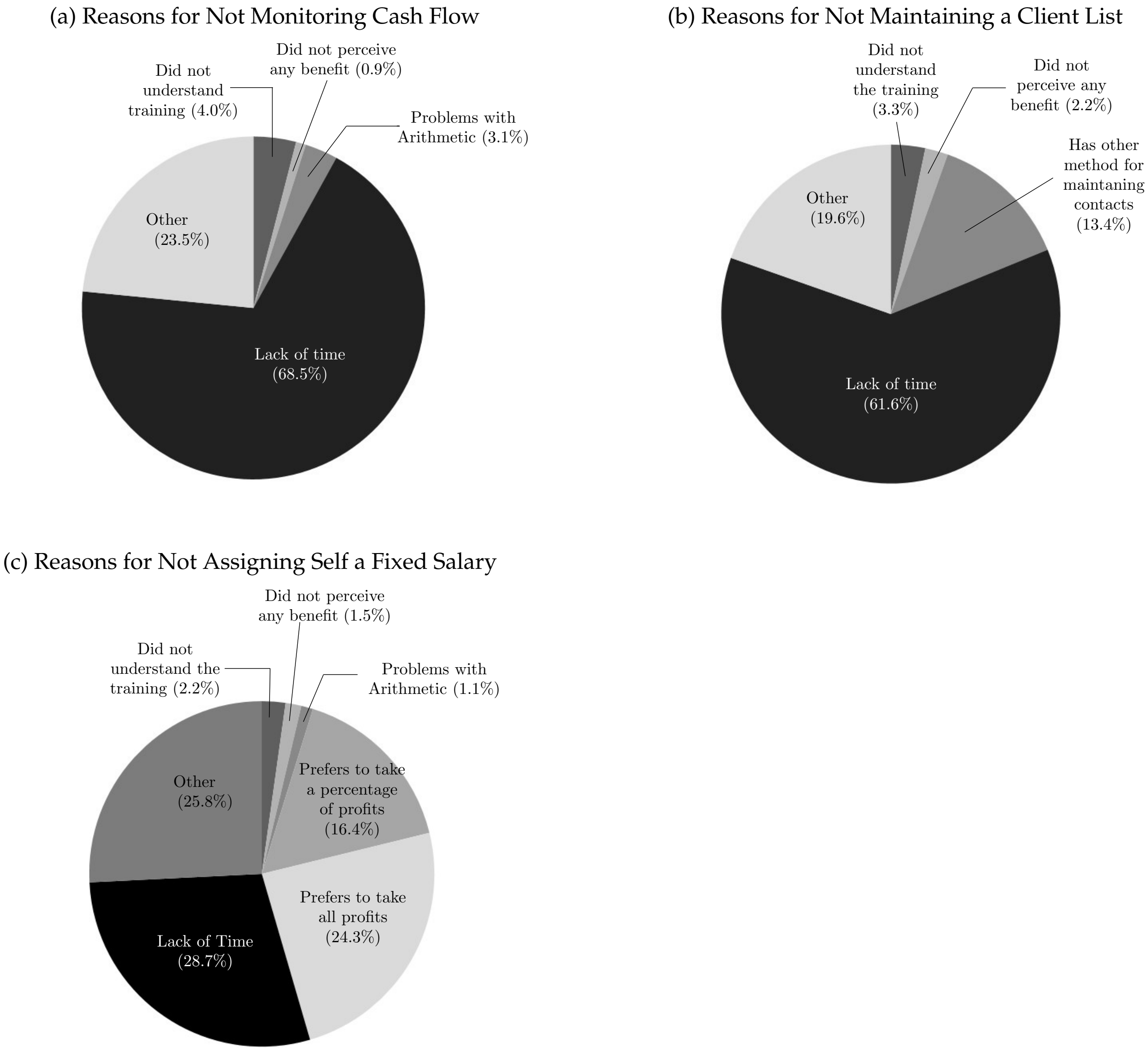
Note: Each simulation assumes the following: (a) X percent of micro entrepreneurs in the treatment group adopt a business practice; (b) the adoption of this practice leads to a 1%–50 percent increase in sales; (c) only micro entrepreneurs who adopt a practice experience sales increases. Potential values of X=5%, 10%, 20%,.. , 50%. We assume a p-value of 10% to determine the statistical power in these simulations.

training (between 3.7 and 5.3 percent). There are some specific reasons for not assigning themselves a fixed salary: entrepreneurs might prefer to take a percentage of profits (rather than a fixed sum) or they might just want to take all business profits. Also, there are some particular reasons to avoid keeping a client list (arguably they rely on other informal mechanisms, although they were not specified).[21]

However, in all cases, lack of time was an important reason not to adopt the business practices recommended by the program (between 29 and 69 percent, depending on the practice). This is an interesting finding, as it suggests that micro entrepreneurs value their time very highly and may be reluctant to pay the opportunity costs associated with taking on a particular practice. It may simply be the case that the training did not adequately convey the benefits associated with

[21] More than 20 percent of women mentioned other reasons not to adopt each practice. The "Other Reasons" category includes many distinct answers. Some of the reasons not to assign themselves a fixed salary were: particular and unexpected economic shocks, businesses that were not profitable enough, husbands taking away all business profits, neglect, etc. Examples of reasons not to keep a record of business contacts were: blurry vision, clients / suppliers coming anyway, single buyer / seller, etc. Other examples of reasons not to keep a record of sales and expenditures were: low levels of sales and expenditures that did not require bookkeeping, reluctance to know when business declines because it would demoralize them, entrepreneurs who forget to do so by the end of the day, etc.

Figure 4: Reasons for Not Adopting Business Practices



adopting the recommended practices, relative to the perceived cost. Equally possible, however, is that we do not have a sufficiently rich understanding of the time constraints that female micro entrepreneurs face. If this is indeed the case, future interventions could be tailored to focus on how one might adopt these practices in a busy environment without taking up too much of the individual's time.

5 Conclusion

We analyze the impact of a large-scale program (SWEP) that provided business training to female micro entrepreneurs in Peru. It is estimated that more than 100,000 women benefited from SWEP during its four years of operation. The program's innovative approach was to teach business practices based on a soap opera format, where women could identify with the struggles and daily problems of the main character. This soap opera was complemented with games and practical exercises conducted by an instructor. The cost of this media-based intervention was relatively low, providing a potentially cost-effective vehicle for entrepreneur trainings.

To assess the impact of the program, we conducted a field experiment in two cities in Peru (Lima and Piura). We collected a roster of female micro enterprises in three sectors (services, retail, and manufacturing) who would be willing to receive business training within a time frame of one year. From this roster, we randomly assigned 1,500 women to a treatment group and 1,100 women to a control group. Those in the treatment group were invited to participate in the training sessions immediately, while those in the control group would only be invited after our field experiment. Even when we anticipated cancellations and no-shows, unfortunately the rate of attendance among those invited to the training was only 47 percent (703 entrepreneurs).

We evaluate the adoption of three particular business practices emphasized by the program: assigning themselves a fixed salary (rather than taking money from the businesses based on their household needs), keeping a record of potential clients and contacts, and bookkeeping (registering business sales and expenditures to determine their cash flow). We find that the program increased the adoption rate of the first two practices by about 4–6 percentage points and 6–11 percentage points, respectively. The results were significant across multiple specifications. We also find that there were some increases in the rate of adoption of bookkeeping (between 4–5 percentage points). However, this result is not robust in all our regressions. While positive and (for the most part) significant, these relatively modest levels of adoption did not translate into improvements of average business performance (measured through sales, expenditures, and payrolls), household outcomes (per capita household expenditures), or women's empowerment (measured through a set of questions regarding women's decision-making within their households). We ran a series of simulations and robustness checks and verified that our results are not likely to be explained by: (a) the low level of attendance of those initially invited to the training sessions, (b) loss of statistical power induced by this low level of attendance, or (c) subsequent attrition in the sample.

All in all, we find that women who benefited from the program were highly satisfied with the trainings (the average rating was 7.8 of a maximum of 10 points). Women who did not adopt the business practices emphasized by the program were asked about their reasons. Only a few of them did not adopt the business practices because they did not feel they were important or did not understand the training. Instead, most of them indicated that they lacked the time to implement them in their businesses. This suggests the importance of understanding women's time constraints when designing business training curricula.

Appendix

A Multiple Hypothesis Testing for Individual Decisions

To test whether the program affected women's empowerment and ability to make more decisions in the household, we collected data on seven outcomes. Specifically, the survey included questions on whether women decide about the following: (a) how to spend the household's income; (b) food purchases; (c) furniture purchases; (d) family outings; (e) children's education; (f) family discipline; and (g) what to do if a household member gets ill.

The most apparent way to assess our hypothesis is to estimate seven individual regressions on the treatment. However, if we test individually a family of multiple outcomes, we face a problem: as the number of outcomes increases, there is a larger probability that at least one of them will be significant. In this case, if we consider a Type I error (α) of 0.1 and we test the effect of the program on seven outcomes, the probability that at least one of them is significant would be $1 - Pr[\text{no significant results}] = 1 - (1 - 0.1)^7 = 0.52$. Therefore, the probability that we observe a significant impact on at least one decision increases considerably just by chance.

If we are interested in the general impact of the program on a family of outcomes, we can estimate the mean standardized treatment effect (Kling, Liebman, and Katz 2007). To do so, we estimate a system of seemingly unrelated regressions and calculate the impact on each outcome (θ_g). We standardize θ_g by the standard deviation of each outcome in the control group and average across all G outcomes: $\tau = \frac{1}{G}\left[\sum_{g=1}^{G} \frac{\theta_g}{\sigma_g}\right]$.

However, we might still be interested in the impact on the particular variables. For example, a global zero-effect might be driven by positive impacts on certain variables and concurrent small (or even negative) effects on others. However, we need to adjust the p-values to reflect the multiple hypothesis testing problem. One way adjust p-values is the Bonferroni correction. Instead of using a cuttoff point of α, we can adjust it to α/G. However, this method has been criticized because it ignores dependencies among the data and is therefore much too conservative if the number of tests is large (Bland and Altman 1995). Because of this problem, several other methods to adjust for multiple hypothesis testing have been proposed.[22]

We choose a fairly simple ad-hoc adjustment proposed by Sankoh, Huque, and Dubey (1997). Denote the estimated p-value from the regression of outcome g as p, ρ_{gh} as the correlation between outcome g and another outcome h, and $r_{\cdot g} = \frac{1}{(G-1)} \sum_{g \neq h} \rho_{gh}$ is the average correlation of outcome g with the other $G-1$ outcomes. Then, the adjusted p-value would be $p_a = 1 - (1-p)^m$, where $m = G^{1-r_{\cdot g}}$. Intuitively, if $r_{\cdot g} = 1$ (because all variables contain the same information, we would be truly testing only one outcome), then p_a is just the p. However, if $r_{\cdot g} = 0$ (i.e. outcomes are independent of one another), then p_a would be consistent with a Bonferroni correction.

We apply this methodology in Table 9, where we report unadjusted p-values (in parentheses) and adjusted p-values (in brackets). All in all, the coefficients on all the variables are small and not statistically significant. Women's involvement in children's education is almost significant with an α of 0.1 (the unadjusted p-value is 0.11), but this effect disappears when we adjust the p-value to account for multiple hypothesis testing.

[22] Many of these methods (applied to educational interventions) are discussed by Schochet (2008).

Table 9: Effect of the Program on Individual Decisions

Coefficient	t=1			t=2		
	CS	ANCOVA	CS w/controls	CS	ANCOVA	CS w/controls
			A. Decides how to Spend Money			
Treatment	-0.018	-0.03	-0.011	-0.017	-0.027	-0.012
	(0.453)	(0.202)	(0.635)	(0.507)	(0.280)	(0.633)
	[0.844]	[0.501]	[0.955]	[0.875]	[0.618]	[0.947]
N	1548	1514	1548	1456	1421	1456
			B. Food Purchases			
Treatment	0.003	0.001	0.005	-0.003	-0.003	-0.004
	(0.843)	(0.950)	(0.762)	(0.860)	(0.881)	(0.843)
	[0.999]	[0.999]	[0.996]	[0.999]	[0.999]	[0.999]
N	1551	1536	1551	1470	1457	1470
			C. Furniture Purchases			
Treatment	-0.017	-0.017	-0.009	-0.019	-0.015	-0.017
	(0.481)	(0.455)	(0.712)	(0.445)	(0.539)	(0.498)
	[0.859]	[0.837]	[0.976]	[0.802]	[0.882]	[0.851]
N	1551	1503	1551	1445	1399	1445
			D. Family Outings			
Treatment	-0.006	0.001	0.001	-0.032	-0.024	-0.033
	(0.811	(0.972)	(0.975)	(0.197)	(0.340)	(0.180)
	[0.994]	[0.999]	[0.999]	[0.464]	[0.692]	[0.43]
N	1519	1425	1519	1412	1321	1412
			E. Children's Education			
Treatment	-0.038	-0.038	-0.025	-0.044	-0.052	-0.047
	(0.242)	(0.252)	(0.447)	(0.219)	(0.157)	(0.188)
	[0.566]	[0.584]	[0.833]	[0.534]	[0.41]	[0.475]
N	804	747	804	708	657	708
			F. Family Discipline			
Treatment	-0.040	-0.041	-0.034	-0.022	-0.034	-0.023
	(0.114)	(0.112)	(0.175)	(0.389)	(0.219)	(0.373)
	[0.289]	[0.284]	[0.417]	[0.744]	[0.496]	[0.725]
N	1453	1254	1453	1416	1205	1416
			G. What to do if Household Member Is Sick			
Treatment	0.012	0.007	0.013	-0.025	-0.026	-0.023
	(0.556)	(0.709)	(0.499)	(0.203)	(0.184)	(0.243)
	[0.931]	[0.983]	[0.898]	[0.524]	[0.485]	[0.598]
N	1,546	1,499	1,546	1,460	1,419	1,460

Note: Unadjusted p-values in parentheses. Adjusted p-values for multiple hypothesis testing, based on Sankoh, Huque, and Dubey (1997), in brackets.

B Adjustment for Attrition using Inverse Probability Weights

This section adjusts our previous estimates for attrition using inverse probability weights (IPWs). We follow the description of the problem and methodology described by Alderman et al. (2001).[23] Suppose we are interested in estimating the conditional density:

$$y_t = \beta_0 + \beta_1 x_t + \varepsilon_t \tag{7}$$

where y_t is the dependent variable (in our case, the adoption of business practices, business performance, or household outcomes), x_t is a scalar independent variable, and ε_t is a mean-zero random variable. However, y_t is only observed for a non-attrited subsample. In particular, we observe y_t when $A_t = 0$, where A_t is a variable that takes the value of one if an observation is missing and zero otherwise. Attrition is determined by a latent variable:

$$A_t^* = \delta_0 + \delta_t x_t + \delta_2 z_t + \eta_t \tag{8}$$

with $A_t = 1$ if $A_t^* \geq 0$ and $A_t = 0$ otherwise. Furthermore, suppose there is a vector z_t of auxiliary variables that is correlated with the probability of attrition, but is not included in x_t. Alderman et al. (2001) argue that z_t can include lagged values of the dependent variable as well as fixed characteristics of the respondents (e.g., age, location, etc.). Assume that both x_t and z_t are observed for the entire sample (either when $A_t = 0$ or $A_t = 1$).

If attrition happens at random, it does not pose any serious concern for our estimations. Although some of the sample might be lost, the estimates would still be unbiased and consistent. However, if attrition is not random, we need to make some further assumptions about its nature to correctly identify the parameters in the model. One possibility is to assume that there is selection on unobservables. This is similar to the estimation of the canonical Heckman (1979) selection model: z_t would be a set of instrumental variables that determine attrition but should

[23] For a more general discussion of the attrition problem and potential adjustments, see: Wooldridge (2002) and Wooldridge (2007).

be uncorrelated with ε_t. However, in general, it is difficult to identify appropriate variables for this estimation, and there are no variables in our data that would credibly meet any necessary exclusion restrictions.

Because we have a rich set of variables, we will instead suppose that our attrition exhibits selection on observables. This allows for the set of variables z_t to be correlated with both A_t and y_t. In particular, Fitzgerald, Gottschalk, and Moffitt (1998) state that, under selection on observables, this model is identified if ε_t and η_t are uncorrelated. This will hold if $Pr(A_t = 0|y_t, x_t, z_t) = Pr(A_t = 0|x_t, z_t)$: the probability of attrition is independent of the dependent variable y_t, once we account for x_t and z_t.

In practical terms, our approach is to estimate weights $w_u = Pr(A_t = 0|x_t, z_t)$ and normalize them by $w_r = Pr(A_t = 0|x_t)$, where x_t is the treatment indicator and z_t is a set of variables related to attrition. Our variables include: sector of the micro enterprise, sales at baseline, age of business, whether the micro entrepreneur had received previous business training, demographic composition of the household, whether the micro entrepreneur is the household head, her years of education, whether she is a migrant, her marital status, the per capita expenditure of her household, and a set of location (district) indicators. We calculate w_u and w_r as the predicted probabilities from probit models. The IPW is given by the inverse of the normalized weight: $W = \frac{w_r}{w_u}$. Then we can estimate 7 through weighted least squares with W as the regression weights. Intuitively, this estimation places higher weights on observations that share similar initial characteristics as those who attrit later on, and lower weights on those who are likely to remain in the sample.

We report estimates (including marginal effects) of the probit models of attrition in t=1 and t=2 in Table 10. The probability of remaining in the sample increases by 5–9 percent if a micro entrepreneur is in the treatment group. Participants are also somewhat more likely to remain in the sample when they are in the retail and manufacturing sectors (relative to the service sector), when there are more working-age members in the household, when they are the head of their household, and when they are less educated. We use the fitted values of this regression to estimate weights $\hat{W}_1$ and $\hat{W}_2$ for t=1 and t=2, respectively. These are used as weights in linear

Table 10: Probability of Remaining in the Sample (Probit Model)

Variable	t=1		t=2	
	(1)	(2)	(3)	(4)
	Probit Coeff	Marginal Effects[1]	Probit Coeff	Marginal Effects[1]
Treatment	0.3613***	0.0531***	0.4476***	0.0888***
	(0.094)	(0.013)	(0.085)	(0.016)
Retail [2]	0.0815	0.0126	0.1004	0.0211
	(0.104)	(0.016)	(0.094)	(0.020)
Manufacturing [2]	0.0702	0.0104	0.2306*	0.0434*
	(0.145)	(0.021)	(0.138)	(0.023)
Yearly sales (S/. 10,000s)	-0.0124	-0.0019	0.0028	0.0006
at baseline	(0.011)	(0.002)	(0.010)	(0.002)
Age of business (months, X100)	0.0163	0.0025	-0.0183	-0.0038
	(0.050)	(0.008)	(0.044)	(0.009)
Micro entrepreneur had received	0.3671**	0.0454***	0.2452	0.0452*
previous training	(0.180)	(0.017)	(0.151)	(0.024)
Number of members 0-14 y.o.	0.0660	0.0101	0.0078	0.0016
at baseline	(0.046)	(0.007)	(0.039)	(0.008)
Number of members 15-60 y.o.	0.0843**	0.0129**	0.0403	0.0084
at baseline	(0.034)	(0.005)	(0.030)	(0.006)
Number of members 61+ y.o.	0.0396	0.0061	0.1130	0.0235
at baseline	(0.074)	(0.011)	(0.072)	(0.015)
Household head	0.3423***	0.0510***	0.1282	0.0264
	(0.108)	(0.016)	(0.095)	(0.019)
Number of years of education	-0.0108	-0.0017	-0.0300**	-0.0062**
	(0.015)	(0.002)	(0.013)	(0.003)
Micro entrepreneur is a migrant	0.0742	0.0113	-0.0871	-0.0182
(not born in province)	(0.092)	(0.014)	(0.083)	(0.017)
Married	0.1396	0.0219	0.1453	0.0308
	(0.104)	(0.017)	(0.093)	(0.020)
HH per capita monthly	0.0058	0.0009	-0.0250	-0.0052
expenditure (S/. 100s) at baseline	(0.020)	(0.003)	(0.016)	(0.003)
Constant	0.9476***		1.3434***	
	(0.311)		(0.284)	
District dummies (29)	YES		YES	
Observations	1,738		1738	
Pseudo R-Squared	0.1045		0.1016	

Note: The dependent variable of the regression takes a value of 0 if the micro entrepreneur attrited from the sample, and 1 otherwise.

[1] Marginal effects of discrete variables are calculated as the discrete difference in $\Phi(D=1)$-$\Phi(D=0)$.

[2] Base category is services.

Standard errors in parentheses. Significance levels denoted by: *** p<0.01, ** p<0.05, * p<0.1.

cross sectional regressions: $Y_{it} = \beta Treat_i + \varepsilon_{it}$ for t=1 and t=2 and for each of our outcomes of interest (i.e., adoption of business practices, business performance, and household outcomes).

Estimates $\hat{W}_1$ and $\hat{W}_2$ are predicted values from a probit model, and have their own error structure. They could be thought of as a parallel to the instrumental variables or sample selection models- as "first stage estimates." Therefore, we need to adjust the standard errors of the linear cross sectional regressions (the "second stage") to incorporate the error structure of the first stage. To account for this, we bootstrap the estimation of both stages 1,000 times.

Our estimates adjusted for the IPWs are reported in Table 11. With two exceptions (assignment of own salary and payroll), all the impacts are smaller than the ones reported in Tables 3, 4, and 5.

Table 11: Cross Sectional Estimates (adjusted for IPW)

Dependent Variable	t=1	t=2
Own Salary	0.041***	0.061***
	(0.014)	(0.014)
Network	0.083***	0.104***
	(0.025)	(0.027)
Bookkeeping	0.050*	0.041
	(0.027)	(0.028)
Yearly Sales	-2,189.6	-2,390.6
	(3,907.9)	(3,149.4)
Payroll	441.6	-198.2
	(521.2)	(234.5)
Other Business Expenses	-2,002.2	-2,330.7
	(3,145.7)	(2,410.0)
Monthly HH per capita expenditure	5.23	3.76
	(13.0)	(10.5)
Women Household Empowerment[1]	-0.028	-0.053
	(0.038)	(0.041)
N	1,738	

Note: Bootstrapped standard errors (1,000 replications) to account for the error structure of the weights ($\hat{W}$) and the linear regression (second stage). Standard errors in parentheses. Significance levels denoted by: *** p<0.01, ** p<0.05, * p<0.1.

[1] We estimate a seemingly unrelated regressions system of decisions: $D_{ig} = \theta_g \text{Treat}_i + \varepsilon_i$ with decisions g=1,...,G, where each regression is weighted by $\hat{W}$. We estimate the overall impact on this family of outcomes through $\tau = \frac{1}{G}\Sigma_{g=1}^{G}\frac{\theta_g}{\sigma_g}$. σ_g is also estimated using W weights.

References

Alderman, H., Behrman, J.R., Kohler, H.P., Maluccio, J.A., and Watkins, S.C. (2001). "Attrition in Longitudinal Household Survey Data." *Demographic Research* **5**, 79–124.

Angelelli, P., Moudry, R., and Llisterri, J.J. (2006). "Institutional Capacities for Small Business Policy Development in Latin America and the Caribbean." Sustainable Development Department Technical Paper Series, Inter-American Development Bank, Washington, DC.

Ayyagari, M., Demirguc-Kunt, A., and Maksimovic, V. (2011). "Small vs. Young Firms across the World: Contribution to Employment, Job Creation, and Growth." Policy Research Working Paper 5631, World Bank, Washington, DC.

Banerjee, A., Duflo, E., Glennerster, R., and Kinnan, C. (2013). "The Miracle of Microfinance? Evidence from a Randomized Evaluation." Working Paper 18950, National Bureau of Economic Research, Cambridge, MA.

Berge, L.I.O., Bjorvatn, K., and Tungodden, B. (2011). "Human and Financial Capital for Microenterprise Development: Evidence from a Lab and Field Experiment." Discussion Paper 1/2011, NHH Dept. of Economics, Bergen, Norway.

Bland, J.M. and Altman, D.G. (1995). "Multiple Significance Tests: the Bonferroni Method." *BMJ: British Medical Journal* **310**(6973), 170.

Bloom, N. and Reenen, J.V. (2010). "Why do Management Practices Differ across Firms and Countries?" *Journal of Economic Perspectives* **24**(1), 203–224.

Bruhn, M., Karlan, D., and Schoar, A. (2010). "What Capital is missing in Developing Countries?" *American Economic Review* **100**(2), 629–33.

Calderón, G., Cunha, J.M., and Giorgi, G.D. (2013). "Business Literacy and Development: Evidence from a Randomized Controlled Trial in Rural Mexico." Working Paper 19740, National Bureau of Economic Research, Cambridge, MA.

Cho, Y. and Honorati, M. (2013). "Entrepreneurship Programs in Developing Countries: A Meta Regression Analysis." Policy Research Working Paper 6402, World Bank, Washington, DC.

de Mel, S., McKenzie, D., and Woodruff, C. (2008). "Returns to Capital in Microenterprises: Evidence from a Field Experiment." *The Quarterly Journal of Economics* **123**(4), 1329–1372.

de Mel, S., McKenzie, D., and Woodruff, C. (2014). "Business training and female enterprise start-up, growth, and dynamics: Experimental evidence from Sri Lanka." *Journal of Development Economics* **106**, 199–210.

De Soto, H. (1989). *The Other Path: The Economic Answer to Terrorism*. Harper and Row Publishers, Inc, New York, NY.

Diario Correo (2012). "Incendio consume cientos de puestos en mercado de Villa El Salvador." Online newspaper version, published February 17, 2012: http://tinyurl.com/pnj9a4g.

Drexler, A., Fischer, G., and Schoar, A. (2014). "Keeping it Simple: Financial Literacy and Rules of Thumb." *American Economic Journal: Applied Economics* **6**(2), 1–31.

Duflo, E., Glennerster, R., and Kremer, M. (2007). "Using Randomization in Development Economics Research: A Toolkit." In T.P. Schultz and J.A. Strauss, eds., "Handbook of Development Economics," Elsevier, volume 4, chapter 61, pp. 3895–3962.

Evans, D.S. and Jovanovic, B. (1989). "An Estimated Model of Entrepreneurial Choice under Liquidity Constraints." *Journal of Political Economy* **47**(4), 808–827.

Fafchamps, M., McKenzie, D., Quinn, S., and Woodruff, C. (2014). "Microenterprise growth and the flypaper effect: Evidence from a randomized experiment in Ghana." *Journal of Development Economics* **106**, 211–226.

Field, E., Jayachandran, S., and Pande, R. (2010). "Do Traditional Institutions Constrain Female Entrepreneurship? A Field Experiment on Business Training in India." *American Economic Review* **100**(2), 125–29.

Fitzgerald, J., Gottschalk, P., and Moffitt, R. (1998). "An Analysis of Sample Attrition in Panel Data." *The Journal of Human Resources* **33**(2), 251–299.

Giné, X. and Mansuri, G. (2011). "Money or Ideas? A Field Experiment on Constraints to Entrepreneurship in Rural Pakistan." Unpublished Manuscript, World Bank, Washington, DC.

Heckman, J. (1979). "Sample Selection Basis as a Specification Error." *Econometrica* **47**(1), 153–161.

Heckman, J.J., Lalonde, R.J., and Smith, J.A. (1999). "The Economics and Econometrics of Active Labor Market Programs." In O. Ashenfelter and D. Card, eds., "Handbook of Labor Economics," Elsevier, volume 3 of *Handbook of Labor Economics*, chapter 31, pp. 1865–2097.

Jensen, R. and Oster, E. (2009). "The Power of TV: Cable Television and Women's Status in India." *The Quarterly Journal of Economics* **124**(3), 1057–1094.

Karlan, D., Knight, R., and Udry, C. (2012). "Hoping to Win, Expected to Lose: Theory and Lessons on Microenterprise Development." Working Paper 312, Center for Global Development, Washington, DC.

Karlan, D. and Valdivia, M. (2011). "Teaching Entrepreneurship: Impact of Business Training on Microfinance Clients and Institutions." *The Review of Economics and Statistics* **93**(2), 510–527.

Kling, J.R. and Liebman, J.B. (2004). "Experimental Analysis of Neighborhood Effects on Youth." Working Paper 483, Princeton University, Industrial Relations Section (IRS), Princeton, NJ.

Kling, J.R., Liebman, J.B., and Katz, L.F. (2007). "Experimental analysis of neighborhood effects." *Econometrica* **75**(1), 83–119.

La Ferrara, E., Chong, A., and Duryea, S. (2012). "Soap Operas and Fertility: Evidence from Brazil." *American Economic Journal: Applied Economics* **4**(4), 1–31.

Li, Y. and Rama, M. (2012). "Firm Dynamics, Productivity Growth and Job Creation in Developing Countries: The Role of Micro- and Small Enterprises." Background Paper for the World Development Report 2013, World Bank, Washington, DC.

McKenzie, D. (2012). "Beyond baseline and follow-up: The case for more T in experiments." *Journal of Development Economics* **99**, 210–221.

McKenzie, D. and Woodruff, C. (2014). "What are We Learning from Business Training and Entrepreneurship Evaluations around the Developing World?" *World Bank Research Observer* **29**(1), 48 – 82.

Restuccia, D. and Rogerson, R. (2008). "Policy Distortions and Aggregate Productivity with Heterogeneous Establishments." *Review of Economic Dynamics* **11**(4), 707–720.

Sankoh, A., Huque, M., and Dubey, S. (1997). "Some Comments on Frequently used Multiple Endpoint Adjustment Methods in Clinical Trials." *Statistics in M* **16**, 2529–2542.

Schochet, P.Z. (2008). "Guidelines for Multiple Testing in Impact Evaluations of Educational Interventions." Unpublished manuscript, Mathematica Policy Research Inc., Princeton, NJ.

Valdivia, M. (2011). "Training or Technical Assistance? A Field Experiment to Learn What Works to Increase Managerial Capital for Female Microentrepreneurs." Unpublished manuscript, Lima, Peru.

Wooldridge, J.M. (2002). *Econometric Analysis of Cross-sectional and Panel Data*. MIT Press, Cambrdige, MA.

Wooldridge, J.M. (2007). "Inverse Probability Weighted Estimation for General Missing Data Problems." *Journal of Econometrics* **141**(2), 1281 – 1301.

World Bank (2012). *World Development Report 2013: Jobs*. World Bank, Washington, DC.